The Geometry of War

Topological Conflict Prediction and the Davis Framework

Expanded Second Edition

Bee Rosa Davis

2026

Copyright © 2026 Bee Rosa Davis

All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means without the prior written permission of the author, except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law.

The mathematical and physical content of this work is placed in the public domain for the benefit of humanity. The narrative and literary elements remain under copyright.

Patent Applications:
HELICITY System: US Provisional 63/935,782
PSYCHOHISTORY System: US Provisional 63/972,955

Expanded Second Edition

Imprint: Independently published

Printed in the United States of America

For Mom,
the kindest soul I have ever known,
the most intuitive heart I have ever followed.
You traced geodesics through chaos
with nothing but love as your compass.
The math came later.
The truth was always yours.



Kazimiera Mika, age 12, mourns her sister Andzia, age 14, killed by German aircraft while digging for potatoes. Warsaw, September 13, 1939.

Photo: Julien Bryan. Public Domain.

Contents

Preface vii
Prologue ix

I The promise and the trap 1
1 Why We Cannot Predict Wars 2
2 The Great Debates 29
3 The Choices Inside the Trap 48

II Reading the strain 53
4 From Transformers to Trajectories 54
5 Three Countries and a Calculation 78
6 What a Relationship Contains 85
7 The Spectral Gap 91

III Rupture and recovery 108
8 The Russia-Ukraine Case 110
9 Historical Backtesting 136
10 Japan and the Closing Resource Horizon 151
11 The Crisis That Found an Exit 156
12 China-Taiwan: The Active Watch 161
13 Israel-Iran: The Slow Escalation 182
14 India-Pakistan: The Frozen Conflict 200
15 Other Active Dyads 213
16 China and Australia Find Room to Move 224
17 France and Germany Build a Different Relationship 229
18 Sudan: The State Is Not One Actor 234
19 The Horn: Peace Has More Than One Boundary 248
20 The Great Lakes: Connected to What? 261
21 West Africa: Leaving an Institution, Keeping a Route 274

IV The mathematics of constraint 286
22 The Mathematics of Helicity 287

23 Spectral Geometry Foundations **311**
24 Glass Transitions **322**
25 From Physics to Politics **331**
26 What Changes When the Ruler Changes **336**
27 The Order of the Moves **346**
28 What the Observations Leave Open **352**
29 How Much Can One Uncertain Connection Change? **363**
30 A Stable System Can Still Hurt You **375**
31 The Route, the Reserve, and the Deadline **387**

V Putting the warning to work **399**
32 When the Instrument Lies **400**
33 What a Warning Has to Predict **409**
34 From a Source Record to a Warning **420**
35 Put the Models on the Same Starting Line **430**

VI Before the next rupture **442**
36 Policy Implications **443**
37 When the Warning Changes the Outcome **451**
38 Limitations and Critiques **461**
39 Future Directions **467**
Epilogue **475**

Appendices **477**
A Data Sources **477**
B Computational Methods **486**
C Selected Historical Conflict Analyses **499**
D Glossary and Lexicon **501**
E Reading and Reproducing a Forecast **531**
F The New Helicity Research Records **542**
G Reproducing the Expanded Methods **545**
H Notation and Mathematical Reading Guide **548**
Bibliography **560**
Index **575**
About the Author **584**

Preface: How to Read This Book

> The owl of Minerva spreads its wings only with the falling of the dusk.
>
> G.W.F. Hegel, *Philosophy of Right* (1820)

This book is written for three audiences, and you may belong to more than one.

Policymakers and analysts will find Parts I, II, and III most immediately useful. These sections present the framework's core logic, its validation results, and its application to current relationships—all without requiring mathematical sophistication. If you are reading to understand what the framework says about a specific dyad (Russia-Ukraine, China-Taiwan, India-Pakistan), you can go directly to the relevant case study chapter after reading the introduction.

Researchers and academics will want to engage with Part IV, which provides the mathematical foundations. The helicity concept, the spectral gap derivation, and the glass transition analogy all have precise formulations that deserve scrutiny. I welcome critique. The appendices supply data sources, worked calculations, and the record requirements for independent replication.

General readers interested in the intersection of mathematics, physics, and international relations can read the book straight through. I have tried to make the technical material accessible without sacrificing rigor. Where equations appear, they are explained in words; where analogies are used, they are clearly marked as such.

The African chapters examine divided authority, overlapping conflicts, civilian access, and institutional change. The later mathematics follows uncertain connections, temporary surges, route capacity, and reserves. Part V takes the warning from its source records to a fair comparison with other methods. Part VI considers the decision, including the case in which acting on a warning changes the outcome. Appendix G describes the reproducible companion.

A note on my perspective.

I am not a political scientist. I am not a diplomat. I spent twenty-seven years in aerospace and intelligence, building systems that process

information and detect patterns. That experience taught me to ask what an instrument measures, what it misses, and what happens when someone has to act on its reading.

I am also a Black trans woman—an identity that has made the traditional paths to academic credibility largely inaccessible. The institutions that study war have not welcomed my voice. I have learned to do the work anyway, trusting that mathematics speaks for itself.

This book is the result of that work. The framework grew through research on information flow and financial relationships, followed by its geopolitical application and evaluation against 47 historical conflict onsets. Its usefulness depends on what the calculations explain and how they perform when the next outcome is still unknown.

Acknowledgments.

No book is written alone. I am grateful to the researchers whose data made this work possible: the teams at the Uppsala Conflict Data Program, the Correlates of War project, the IMF, and the UN Statistics Division. I am grateful to the scholars whose work shaped my thinking, especially those cited throughout these pages. And I am grateful to the early readers who pushed back on unclear arguments and demanded better explanations.

"The past is never dead," wrote William Faulkner. "It's not even past." The conflicts that shaped the twentieth century left structural legacies that continue to shape the twenty-first. This book attempts to make those legacies visible—to reveal the frozen tensions, the accumulated grievances, the economic entanglements that become weapons when trust fails.

An agreement can end a battle without resolving the dependencies and grievances around it. A border may change while a road, a power line, or a family's history still crosses it. Following those continuities is part of understanding what the next crisis inherits.

This book is an attempt to see those structures before they surface. The owl of Minerva may spread its wings at dusk. But mathematics can see in the dark.

Bee Rosa Davis
2026

Prologue: 72 Days

> In war, truth is the first casualty.
>
> Attributed to Aeschylus (c. 525–456 BCE)

December 14, 2021.

That date marks a threshold in my later reconstruction of the Russia–Ukraine relationship. Seventy-two days remained before the full-scale invasion. Gas was still moving through the pipes. Contracts still had value. People still had jobs, journeys, and plans that assumed the next month would resemble the last.

I came to the date after the invasion, while developing the geopolitical application of my research. I knew what had happened. The harder question was what the earlier records could establish without borrowing the answer from the war that followed.

The number on the screen was helicity. I was comparing the persistence of economic dependence with the collapse of diplomatic alignment. A pipeline can keep working after the political relationship that makes it tolerable has broken down. Its continued operation can then preserve a means of pressure. The machinery looks familiar. The choices available to the people depending on it have changed.

Russia and Ukraine had lived with hostility for years. Hostility alone could not explain why this crisis became a full-scale invasion. The reconstruction placed the relationship in the framework's highest warning category in late 2021. December 14 is the reported date under one specification. Another parameter history moves it by eight days. Different network boundaries move the estimate further.

That is where the research begins. Can the warning survive a change in the assumptions? Can it distinguish a dangerous confrontation that ends in bargaining from one that ends in violence? Can it be calculated with information available before the outcome?

As a Black trans woman working across fields that rarely invite people like me to define their questions, I know the difference between having an argument and being allowed to make it. This book puts the argument where it can be examined. The calculation has to carry its evidence.

February 24, 2022.
Seventy-two days later, Russian missiles struck Kyiv.



Destruction in Bucha following Russian withdrawal, April 2022. The human cost of the invasion—measured in lives lost, homes destroyed, and families displaced— represents what structural prediction aims to prevent.

Photo: Wikimedia Commons, CC BY-SA 4.0.

Military preparations were already visible. NATO publicly discussed Russia's buildup in April 2021 and described heavy armour and combat-ready troops again that November (NATO 2021; Stoltenberg 2021). The geometric question concerns what continuing dependence meant in that setting. Could it identify the relationships in which threats were becoming harder to resolve, and improve on information analysts already had?

This book is about how a relationship loses room to move, and how to recognize the loss while there is still something to do about it.

It is about the mathematics of bilateral relationships, where economic ties and diplomatic trust can reinforce each other or pull against each other. It is about abrupt changes in what a system can absorb. Above all, it is about measuring constraints closely enough to ask which choices remain available.

Most of all, it is about a question: Can we see wars coming?

A useful warning has to say more than that tensions are rising. Which fuel supply cannot be replaced? Which commitment blocks a concession? How much time remains before reserves run out? A warning earns its place when it identifies a pressure that can be investigated and an action that can still change the situation.

I believe we can learn to see more of that danger in time. This book presents the framework.

The framework has a name: the **Davis Framework**, because I invented it and no one else has. It rests on a single equation:

$$\boxed{H = \frac{\tau}{K}}$$

Here H compares two aspects of the relationship. In the signed construction, τ records the orientation or asymmetry of flows and $K > 0$ supplies the coupling scale. An adverse orientation gives a negative reading. The capacity construction uses a nonnegative tolerance in the numerator and asks a different question about room for adjustment. The chapters specify which construction each calculation uses.

When H is positive, economic and diplomatic vectors point the same direction. This is the "liberal peace" regime: high trade, high trust, stable relationships. France and Germany today.

France and Germany show how much work such alignment can require. Institutions, shared production, and repeated cooperation made alternatives to confrontation usable over generations.

When H is negative, the vectors point in opposite directions. Economic interdependence persists despite diplomatic hostility. This is the "trap" regime: weaponized interdependence, mutual leverage, explosive instability. Russia and Ukraine before the invasion.

When the specified indicators cross their thresholds together, the system enters SINGULARITY, the framework's highest warning category. It calls for urgent investigation and preparation. It does not assign a probability of one to war. The decisions that follow can still change the relationship.

The full-sample reconstruction assigns that warning to Russia–Ukraine on December 14, 2021.

I tested the framework against 47 conflict onsets from 1990 to 2022. In the working paper's mass-gap and Z-score study, 35 of the 47 had a qualifying warning in the preceding 180 days. That is the question a family facing an invasion would ask first. How often did the warning come? The answer here was 35/47, about 74 percent. A diplomat receiving an alert has a second question. How often did a warning

precede conflict? The study counted 35 hits and 14 false alerts, giving 35/49, about 71 percent. Appendix E gives the full calculation.

These are not perfect numbers. Some warnings were followed by peace within the evaluation window. Some conflicts arrived without a qualifying warning. Each result belongs in the score. An intervention may explain a peaceful outcome, but demonstrating that it prevented a war requires evidence beyond the fact that war did not follow.

Timing changes what a warning is worth. The working paper reports a mean lead of 72 days in that evaluation. The 18 detections in the separate historical roster in Appendix C have a mean lead of 127 days. A warning that comes before a fuel contract is signed leaves a different choice from one that comes after the gas is already flowing. That difference is why I keep the dates as well as the scores.

Looking backward gives the researcher an advantage no warning desk possesses. The model can be tested fairly only when its observations, thresholds, and target are fixed before the test outcomes are inspected.

Among relationships already identified as potentially at risk, which are losing the ability to adjust peacefully?

The answer is geometric.

Mathematics gives that inquiry a form other people can inspect. Its political meaning must be earned in the cases.

A word about what this book is not.

It is not a claim to predict the future. I cannot tell you who will make what decision on what day. Leaders retain agency. Accidents happen. The framework identifies structural conditions, not inevitabilities.

It is not a policy prescription. I have opinions about what should be done, and Chapter 36 shares some of them. But the mathematics is independent of the politics. The framework would work equally well for a hawk or a dove, a realist or a liberal, anyone who wants to know when bilateral relationships are approaching critical instability.

It is not a complete theory of war. Interstate bilateral conflict is one type of war among many. Civil wars, insurgencies, multilateral interventions—these operate by different dynamics that this framework does not address. I am solving one problem, not all problems.

This book develops a mathematical framework for identifying when the arrangements that sustain a relationship also obstruct its peaceful adjustment.

The structure is as follows.

Part I ("The promise and the trap") follows failed warnings, the debates about commerce and conflict, and the choices inside a constrained relationship.

Part II ("Reading the strain") develops helicity, a three-country calculation, the measurement of relationships, spectral gaps, and glass transitions.

Part III ("Rupture and recovery") follows Russia–Ukraine, the historical backtesting, the US–Japan resource crisis, Cuba, and contemporary dyads. China–Australia and France–Germany give recovery and institutional construction the same attention as breakdown. Sudan, the Horn of Africa, the Great Lakes, and West Africa test the framework against divided authority, overlapping conflicts, unequal access, and institutional change.

Part IV ("The mathematics of constraint") connects helicity, spectral geometry, and glass dynamics to the cases. The newer Helicity work adds currency conversions, ordered inspections and deliveries, and incomplete supply records: practical routes into invariance, signed path area, and admissible bounds. Further worked calculations examine uncertain edges, temporary amplification, layered response, route capacity, and the time that reserves can buy.

Part V ("Putting the warning to work") tests modeling failures, defines probability forecasts, follows an actual implementation through a constructed input record, and specifies a head-to-head comparison with other forecasting methods.

Part VI ("Before the next rupture") turns to policy, prevention, limitations, and future research: how a warning becomes a decision, how that decision changes the outcome, and how a relationship gains usable alternatives.

The appendices provide data sources, computational methods, conflict records, a glossary, and worked accounts of forecast evaluation and the expanded research companion.

One more thing.

The 72 days between December 14, 2021 and February 24, 2022 were not empty. Intelligence agencies issued warnings. Diplomats held meetings. Sanctions were threatened. Those efforts did not prevent the invasion. That outcome makes the choices during the interval worth examining, rather than proving that every alternative was already closed.

But imagine a different scenario. Imagine the warning came earlier—not 72 days before invasion, but 72 weeks. Imagine policymakers understood that the relationship had entered a critical regime and acted to reset the trajectory before it crossed into SINGULARITY.

Would it have mattered? I do not know. History offers no counterfactuals. But I believe that structural problems require structural solutions, and structural solutions require structural diagnosis. You cannot treat a disease you have not identified.

The purpose of this instrument is to make a dangerous dependency visible in time to investigate it. What we do with the diagnosis is up to us.

* * *

The numbers are still running.

The dated assessments in this book follow China–Taiwan, India–Pakistan, and Israel–Iran through different forms of strain. Their horizons matter. A reading belongs to the observations and model that produced it, and the next observation may change the assessment.

The geometry is screaming.

Someone should listen.

Part I

The promise and the trap

The shipment arrives. The invoice is paid. Across the same border, trust is disappearing. We begin with the distance between those two facts: the promise that commerce would keep the peace, the warnings that failed to change decisions, and the choices still available inside a dangerous relationship.

Chapter 1

Why We Cannot Predict Wars

> Those who cannot remember the past are condemned to repeat it.
>
> George Santayana, *The Life of Reason* (1905)

In early November 2021, President Biden sent CIA Director William Burns to Moscow. In his later Ditchley lecture, Burns described Russian leaders as unmoved by the American understanding of their plans. In his account, they believed the opportunity to dominate Ukraine was closing (Burns 2023a).

That is a documented warning before the invasion. It creates a demanding starting point for this book. A new framework cannot claim credit simply for noticing danger that intelligence officers and public observers had already identified. It has to show what further distinction it can make.

Consider the decision confronting a government receiving such a warning. Evacuation has costs. Preparing sanctions takes coordination. Sending equipment takes time. Waiting also has costs, and those rise if the danger becomes harder to reverse. The useful question is how a structural assessment changes those decisions.

Three months later, Russian forces crossed the Ukrainian border. The invasion Burns had sensed in that Moscow meeting room became the largest land war in Europe since 1945.

The invasion was preceded by warning, yet warning did not secure prevention. Those are different failures. Accurate intelligence can be discounted, acted on too late, or acted on without changing the adversary's decision. A theory of conflict has to distinguish the detection of danger from the ability to remove it.

For people living with the consequences of the war that had begun in 2014, danger was already part of daily life. An international warning desk had to distinguish another confrontation from a new scale of attack. The persistence of danger makes that task harder, because repeated crises can make even a serious warning feel familiar.

This paradox—accurate warning without effective prediction—reveals the fundamental problem with how we think about conflict.

The Russia-Ukraine case illuminates what we might call the *paradox of surprise*: the coexistence of accurate warning with genuine shock. This paradox recurs throughout the history of conflict prediction, and understanding it reveals the fundamental limitations of current approaches.

Warning and surprise can coexist. The problem of interpreting and acting on indications of attack runs through the intelligence literature Betts 1982. A file can contain the necessary observation while the institution fails to turn it into the necessary decision.

U.S. intelligence proved remarkably accurate in the months preceding February 24, 2022. The performance deserves detailed examination because it represents the *best case* for conventional analysis—and still fell short.

Burns's November assessment—that Putin had made an operational decision, not merely a contingent threat—proved exactly right. But even before that Moscow meeting, the intelligence community had been watching the military buildup with growing concern. By mid-December, satellite imagery told an unmistakable story: approximately 175,000 troops deployed near Ukrainian borders, field hospitals positioned for mass casualties, blood supplies stockpiled in quantities that exceeded any exercise requirement.

By November 30, NATO was publicly describing heavy armour, drones, and combat-ready troops near Ukraine (Stoltenberg 2021). The observations mattered. A geometric account should be tested for what it adds to them.

The granular indicators mounted: bridging equipment positioned at river crossings, command posts established in forward positions, airborne forces concentrated in patterns consistent with initial strike planning. The logistical infrastructure for sustained combat operations was unmistakable to anyone trained to read it.

The command problems examined by Freedman place intelligence alongside political judgment and military preparation Freedman 2022. Forecasting the possibility of an invasion, understanding the intended operation, and anticipating its outcome are distinct achievements. One accurate forecast does not make the others automatic.

Yet most academic experts and area specialists failed to anticipate full-scale invasion. This failure was not due to inferior information; open-source analysis had access to much of the same satellite data, troop movement reports, and diplomatic communiqués. The failure was conceptual.

Fiona Hill's analysis of Putin offers a useful point of comparison Hill 2022; Hill and Gaddy 2013. Understanding a leader's background and methods does not remove uncertainty about a particular decision. An assessment must distinguish what the evidence supports from what the observer expects the leader to consider reasonable.

Different experiences can make the same observation look different. For an observer who expects a recurring bargaining threat, another concentration of troops may suggest pressure for concessions. For an observer who has lived through occupation, it may signal an unfinished project of domination. The analytical task is to expose those assumptions and test them against the record.

An analyst can also mistake disagreement about objectives for a mistake in arithmetic. A leader who places unusual value on territorial control or political legacy may accept losses that the observer expects to be decisive. The task is to establish the objective being pursued, not assume that the observer's priorities are shared.

Some analysts judged a full-scale invasion too costly to be a likely choice. That judgment raises an essential question. Which costs were they counting, and how did they know that Russian decision makers assigned those costs the same weight?

A systematic study by Driedger and Polianskii (2023), published in *International Affairs* Driedger and Polianskii 2023, conducted structured interviews with leading Russia experts and identified consistent failure modes:

Rationalist assumption. An analyst may calculate that invasion will destroy trade, invite sanctions, and impose military costs. The conclusion that it would be costly is not yet evidence that the decision maker will reject it. A forecast also needs an account of what that person values and what they believe inaction will cost.

Historical anchoring. The comparatively limited Crimea operation in 2014 could become a misleading template for later action. An observer expecting another bounded operation might read preparations for a much larger one through that earlier experience. A model has to be capable of recognizing a change in scale as well as a repetition.

Putin as constrained actor. An analyst may expect business interests, military advice, or public opposition to constrain a leader. Whether those pressures can change the decision depends on the institutions through which they operate. Driedger and Polianskii's analysis makes the degree of autonomy an important part of the explanation.

Political framing. Analysts can project their own assumptions about economic costs, public opinion, and institutional restraint onto another

government. Claims about territory and historical identity then risk being treated as rhetoric even when they help organize actual decisions.

An optimizing model is only as informative as the objectives it represents. If it includes export revenue but excludes the value a leader assigns to control of territory, it may calculate a persuasive answer to the wrong decision problem. The omitted objective matters even when the arithmetic is flawless.

A river monitoring station can measure water levels with perfect accuracy while failing to predict which storms will cause floods. The instruments detect *what is happening* but not *what the system is about to do*. The gauges record millimeter-by-millimeter changes in water height, but they cannot see the upstream accumulation, the saturated soil that will no longer absorb rainfall, the configuration of tributaries that will concentrate flow. Similarly, intelligence agencies tracked troop movements with precision but lacked a framework for determining when accumulating pressure would cross a threshold into kinetic action. They measured the symptoms without diagnosing the disease.

The market study discussed here places conflict-related repricing approximately fifty days before the invasion. That raises a useful question about combining observations: what did market prices add to the military and diplomatic record, and how early could an analyst distinguish the change from ordinary volatility? A market movement can register concern without specifying an invasion date.

Ukrainian government bond spreads began widening in late December 2021, with credit default swap prices rising from approximately 600 basis points to over 2,000 basis points by mid-February. The Ukrainian hryvnia depreciated despite central bank intervention. Most tellingly, options markets showed pronounced put skew on European energy companies with Russian exposure—traders were paying premiums to protect against downside scenarios.

The reconstruction developed in this book assigns a critical threshold crossing to December 14, 2021—seventy-two days before the invasion. It asks a further question of the warnings and prices: how did continued dependence constrain the choices available as trust collapsed? Chapter 8 follows that calculation and the information needed to test whether it could have added useful warning in real time.

The Trap That Looked Like Peace

If the Russia-Ukraine relationship illustrated the geometry of conflict, Germany's relationship with Russia illustrated something perhaps more troubling: the geometry of self-deception.

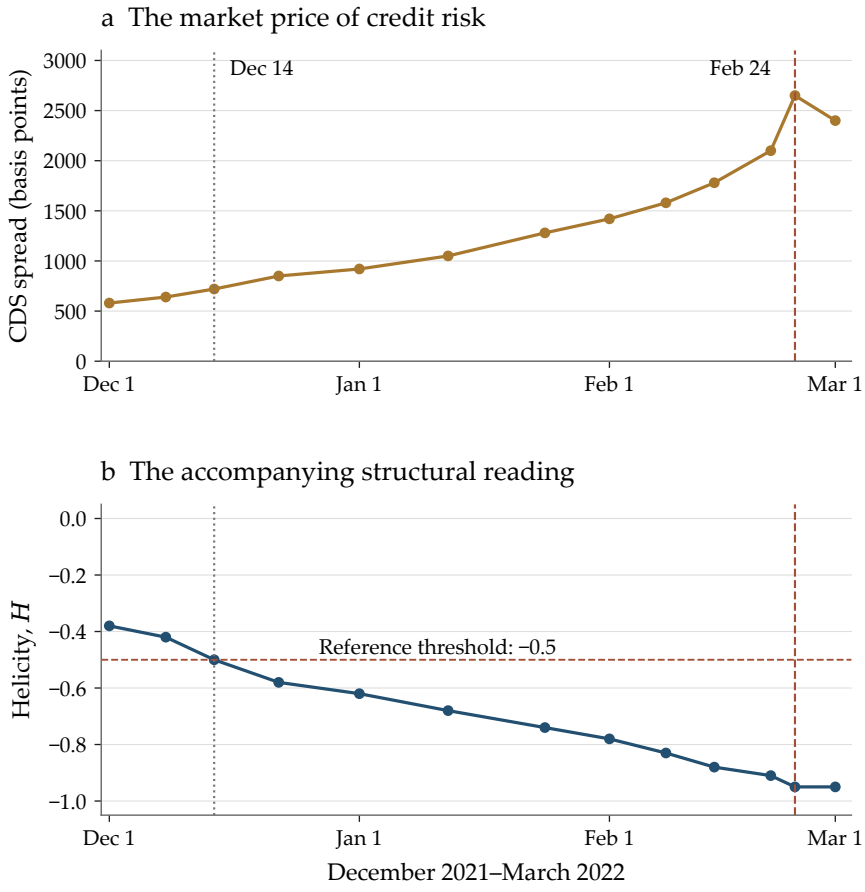


Figure 1.1: Credit-risk pricing and the structural reading, December 2021–March 2022, on a common calendar axis. Vertical rules identify December 14 and February 24, 72 days apart. The dashed horizontal rule marks the $H = -0.5$ reference threshold.

For two decades, German policy toward Russia rested on a doctrine called *Wandel durch Handel*—change through trade. The theory was elegant in its simplicity: by integrating Russia into the European economic order, Germany would create stakeholders for peace within Russian society. Oligarchs enriched by Western trade would lobby against adventurism. Energy interdependence would make conflict unthinkable because the costs would fall on both sides. Commerce would accomplish what diplomacy alone could not.

It was Norman Angell's *Great Illusion* Angell 1909 updated for the twenty-first century. And like Angell's thesis, it contained a fatal flaw.

The numbers told a story that should have alarmed anyone paying attention. By 2021, Germany imported 55% of its natural gas from Russia. For some industrial sectors, the dependence exceeded 65%. The Nord Stream pipelines—Nord Stream 1 completed in 2011, Nord Stream 2 awaiting final certification in 2021—represented €20 billion in infrastructure investment designed to bypass Ukraine entirely, delivering Russian gas directly to German shores.

The same infrastructure could be read in opposite ways. One account saw Europe binding Russia into a valuable commercial relationship. Another saw Europe becoming dependent on a supplier whose political purposes it could not control. The pipe did not resolve that disagreement. The availability of alternatives was central to it.

The trade relationship was genuinely substantial. Germany was Russia's second-largest trading partner after China. German machinery filled Russian factories; Russian energy powered German industry. Volkswagen operated plants in Russia; Siemens maintained long-term contracts for turbine maintenance. The commercial ties were deep, reciprocal, and—according to the theory—stabilizing.

But even as trade flourished, the diplomatic relationship deteriorated. The 2008 Georgia war. The 2014 Crimea annexation. The Skripal poisoning in 2018. The Navalny poisoning in 2020. Each crisis produced condemnations, some sanctions, and then... a return to business. Nord Stream 2 construction continued through it all.

Not everyone was deceived.

Engagement can mean different things. An arrangement supported by alternative suppliers and coordinated political commitments gives a buyer options. The same volume of commerce concentrated in one difficult-to-replace source can create a very different bargaining position.

For countries exposed to Russian pressure, dependence carried security implications beyond the price of gas. Payments supported a supplier whose political aims could conflict with those of its customers.

The commercial benefit and the strategic exposure existed in the same transaction.

The vulnerability becomes acute when substitution takes longer than a crisis allows. A buyer may have enough money to purchase gas elsewhere while lacking the terminal, connecting pipeline, or contracted supply needed to receive it. Financial capacity and a usable exit are different resources.

Even within Germany, dissenting voices existed. The Green Party consistently opposed Nord Stream 2 on both environmental and security grounds. Defense analysts at the Bundeswehr University published studies questioning the strategic wisdom of energy dependence. Historians of the Soviet period noted that commercial engagement had done nothing to moderate Stalin, nothing to prevent the Berlin Wall, nothing to humanize the regime.

Supporters of engagement expected the commercial relationship to create incentives for restraint. Their opponents worried that an essential supplier would gain leverage without acquiring the political commitments needed to make that leverage tolerable. The disagreement concerned how the connection would work under pressure.

Counting euros alone does not reveal how the parties value their ability to stop the flow. A seller willing to accept a loss can impose costs on a buyer who needs uninterrupted delivery. The direction of that leverage belongs in the account alongside the size of the trade.

The framework developed in this book provides a precise description of Germany's predicament. The Russia-Germany relationship exhibited *negative helicity*: high economic interdependence combined with deteriorating diplomatic alignment.

In geometric terms, the economic flow vector and the trust vector pointed in opposite directions. Trade persisted; trust declined. The relationship was *twisted*—like water spiraling down a drain, the commercial ties and the political relationship were rotating in opposite directions. This twist creates pressure. The twist must eventually resolve.

Consider the intuition. When economic ties and diplomatic trust are aligned—both positive or both negative—the relationship is geometrically "flat." High trade with high trust: stable partnership. Low trade with low trust: hostile separation, but stable in its hostility. Neither configuration generates internal pressure for change.

But high trade with deteriorating trust creates a structural tension. Each side becomes vulnerable to the other's leverage. The very depth of commercial ties transforms from an asset (mutual benefit) into a liability (mutual exposure). The relationship becomes *weaponizable*.

This is precisely what happened. Germany's gas dependence gave Russia coercive leverage that could be activated at any moment. And Russia's export revenue dependence gave Europe a counter-weapon—but one Germany was reluctant to use because the costs would fall heaviest on German industry.

Mutual dependence can therefore be asymmetric in practice. Both sides may lose from a rupture, yet differ greatly in how much loss they are willing or able to sustain. A shared exposure does not guarantee a shared restraint.

February 24, 2022 shattered the illusion.

Within days, German policy underwent the most dramatic reversal in its postwar history. Chancellor Olaf Scholz announced a *Zeitenwende*—a "turning point"—committing €100 billion to military modernization and pledging to phase out Russian energy imports. Nord Stream 2, completed but never certified, was suspended and later destroyed by sabotage.

The recriminations were bitter.

The invasion made the costs of the preceding policy immediate. Engagement had not supplied the restraint its advocates expected. The task for the German government was now to change commitments that could not be unwound as quickly as a political position could be announced.

The consequences also outlasted the administration that had overseen much of the integration. Contracts, facilities, and expectations persisted through a change in political leadership. A new government inherited the physical commitments along with the argument about what they meant.

Scholz's February 27 policy statement condemned the invasion and presented it as a watershed for German policy (Scholz 2022). The announced change exposed the distance between revising a strategy and replacing the infrastructure on which it had depended.

The human cost of the illusion extended beyond Germany. Twenty years of *Wandel durch Handel* had provided Russia with the foreign exchange reserves to finance military modernization, the energy leverage to divide European responses to earlier provocations, and the confidence that Europe would not seriously resist. The policy intended to prevent war had helped make war possible.

A measurement system shapes what its users can see. Trade volume, investment, and pipeline capacity are visible in standard accounts. The cost of losing a supplier, the credibility of a commitment, and the willingness to bear disruption require additional inquiry. Leaving them out can make a fragile arrangement appear reassuringly busy.

Germany's predicament perfectly illustrates the geometric trap this book describes. The commercial relationship was deep and reciprocal—positive economic interdependence. But the political relationship had been deteriorating since at least 2008, accelerating after 2014. The helicity of the Germany-Russia relationship went negative and kept falling. Each side was increasingly locked in: Germany couldn't easily replace Russian gas; Russia couldn't easily replace German revenue. But the *trust* that would allow both sides to manage the interdependence cooperatively had collapsed. The trap was sprung not when Russia invaded, but years earlier, when the geometric configuration made invasion structurally advantageous for the party willing to absorb short-term costs for long-term strategic gains.

Here is a fact that should alarm anyone working in national security: *no major conflict forecasting system focuses on bilateral interstate war.*

This claim requires careful documentation, because it reveals a structural gap in our collective analytical infrastructure—a gap this book aims to address.

System	Institution	Focus
ViEWS	Uppsala	Civil war, intrastate
ACLED	–	Political violence
Early Warning	Dartmouth	Mass atrocities
PITF	CIA	State failure

Table 1.1: Major conflict forecasting systems and their primary focus areas.

Consider each system in turn:

ViEWS (Violence Early Warning System), developed at Uppsala University's Department of Peace and Conflict Research, represents the current state of the art in conflict forecasting. The system employs ensemble machine learning methods trained on UCDP data, generating probabilistic forecasts at country-month and PRIO-GRID (approximately 50km × 50km) resolutions. ViEWS achieves impressive performance on its target domain: intrastate armed conflict in Africa, with more recent expansions to the Middle East and North Africa.

But ViEWS explicitly focuses on *internal* conflicts. Its analytical architecture is optimized for sub-national violence patterns—predicting which provinces or grid cells will experience conflict next month. The system has no mechanism for modeling bilateral interstate dynamics: the relationship between two sovereign states that might escalate to war.

ACLED (Armed Conflict Location and Event Data), maintained by an academic consortium, provides granular event-level data on political

violence worldwide. ACLED's strength is temporal and spatial precision—tracking individual battles, protests, and riots with GPS coordinates and date stamps. Researchers can observe violence patterns evolving in near-real-time.

ACLED provides event data as well as forecasting products. Its Conflict Alert System (CAST), launched in 2023, forecasts political-violence events over a six-month horizon (ACLED 2026a). That target differs from the onset of a specified bilateral interstate conflict. Comparisons require the same outcome, horizon, geography, and information date; a product launched in 2023 was not an operational alternative in 2021.

The Early Warning Project at Dartmouth's Dickey Center focuses specifically on mass atrocity prevention. The system generates annual risk rankings for countries, identifying those where mass killing, genocide, or politicide is most likely. The project's explicit statement of scope is illuminating: it "does not forecast interstate or cross-border mass killings." The analytical rationale is understandable—mass atrocities overwhelmingly occur within borders, committed by governments against their own populations. But the scope limitation leaves interstate war entirely unaddressed.

The Political Instability Task Force (PITF), funded by the CIA since 1994, focuses on state failure—regime collapse, civil war onset, and ethnic conflict. PITF models have identified partial democracy ("anocracy") as a robust predictor of instability, contributing importantly to understanding of regime type and violence. But the interstate component of PITF's analysis is secondary; the system's architecture optimizes for internal dynamics.

The absence of systematic interstate war forecasting is not accidental. Three factors explain the gap:

Statistical rarity. Interstate wars are, thankfully, rare events. The Correlates of War project catalogs approximately 95 interstate wars from 1816 to 2007—roughly one every two years globally. In any given year, the overwhelming majority of dyads (country pairs) experience no militarized conflict. Class imbalance dominates: approximately 99.5% of country-months show zero interstate conflict fatalities. Machine learning models trained on such imbalanced data tend toward predicting the majority class (peace), achieving high accuracy while missing the rare events that matter most.

Unit of analysis complexity. Intrastate conflict can be modeled at the country level—characteristics of a single state (regime type, GDP, ethnic fractionalization) predict internal violence. Interstate war requires modeling the *relationship* between two states—a fundamentally more complex analytical object. With approximately 200 states generating

roughly 20,000 possible dyads, the combinatorial challenge is substantial. Most dyads have essentially zero conflict probability (Finland-Paraguay, for instance), but identifying which dyads merit attention requires theory, not just data.

Academic disciplinary boundaries. Scholars who study civil war (primarily comparativists and conflict researchers) have developed sophisticated quantitative methods. Scholars who study interstate relations (primarily IR theorists) have traditionally favored qualitative case studies and theoretical argumentation. The methodological tools that work for civil war prediction do not transfer easily to interstate contexts, and the scholars with relevant theoretical frameworks have not prioritized quantitative forecasting.

System	Emphasis	Interstate bilateral
ViEWS	Civil war	—
ACLED	Political violence	—
Early Warning (Dartmouth)	Mass atrocities	—
PITF	State instability	—
Helicity	Bilateral structure	Target

Coverage comparison; not an accuracy ranking.

Figure 1.2: The forecasting scopes compared in this chapter. The final column identifies the bilateral-interstate problem targeted by Helicity. The comparison concerns the systems' emphasis, rather than their relative accuracy.

This gap matters because interstate war remains the most destructive category of organized violence. World War I killed approximately 20 million people; World War II killed 70-85 million. Even smaller interstate

wars—the Iran-Iraq War (1980-88), the Korean War (1950-53)—produced casualties in the hundreds of thousands to millions.

Moreover, the strategic environment of the 2020s has elevated interstate war risk to levels not seen since the Cold War. Great power competition between the United States and China, Russian revanchism in Europe, nuclear-armed tensions between India and Pakistan—these dyads represent potential conflicts of civilizational consequence. Yet our analytical infrastructure remains optimized for a different category of violence.

The framework introduced in this book specifically addresses the unserved domain. By focusing on bilateral dynamics—the geometric relationship between economic interdependence and diplomatic alignment—it provides theoretical and computational tools designed for interstate conflict prediction.

A Century of Failed Prediction

The history of conflict prediction is a history of failure. This is not a criticism but an observation that should inform how we approach the problem. The failures are not random; they follow patterns that reveal the limitations of existing frameworks.

In 1909, Norman Angell published *The Great Illusion* Angell 1909, arguing that conquest could not deliver the economic gains commonly attributed to it. In the fourth edition, he made the distinction explicit: the claim was "not that war is impossible, but that it is futile." (Angell 1913) His logic seemed impeccable: the major European powers were so economically intertwined that war would devastate the aggressor as much as the victim. Britain and Germany, the primary rivals, were each other's largest trading partners. French investment flowed into Russian railways; German chemicals supplied British industry. The costs of disrupting these relationships exceeded any conceivable benefit from territorial conquest.

Angell's argument concerned the economic futility of conquest in an interconnected world. It asked whether military victory could deliver the wealth its advocates expected (Angell 1913). That is a different proposition from saying that rulers will never undertake an economically destructive war.

The distinction matters. Showing that a policy will destroy wealth does not show that decision makers will reject it. They may miscalculate, pursue other objectives, or judge the alternatives more threatening. The July Crisis made the limits of economic reassurance painfully visible.

British Foreign Secretary Sir Edward Grey, watching the lamps go out across Europe in August 1914, captured the bewilderment: "The lamps are going out all over Europe, we shall not see them lit again in our life-time." The rational world order had failed to prevent its own destruction.

In 1914, Britain and Germany exhibited precisely the configuration this book identifies as the "trap": high economic interdependence (substantial bilateral trade, integrated financial markets, cross-border investment) combined with deteriorating diplomatic alignment (the naval race, colonial competition, alliance commitments that created opposing blocs). The helicity was negative and falling. The very interdependence that was supposed to prevent war became a source of strategic anxiety—each side feared the other would use economic leverage as a weapon, creating incentives for preemptive action to escape the trap.

The war that followed killed approximately 20 million people and destroyed the economic order Angell had described. The failure was not in his observation of interdependence but in his understanding of what interdependence implies under conditions of deteriorating trust.

The decisions examined in MacMillan's history show why the costs of war must be read alongside the choices that led toward it MacMillan 2013. Commercial interdependence supplied reasons for restraint. It did not exercise restraint on behalf of the governments involved.

In *The Economic Consequences of the Peace*, Keynes evokes a prewar economic life whose international connections could feel ordinary and dependable even as political rivalry threatened them Keynes 1919. The transactions continued until the assumptions supporting them failed. That is precisely the separation between visible activity and usable stability that concerns this book.

The twentieth century offers a catalog of strategic surprise despite substantial intelligence efforts. But these failures are not random. Each exhibits the same structural signature: analysts measured the *state* of the system—troop positions, economic indicators, regime characteristics—rather than the *trajectory* of the system, how the geometry of relationships was evolving over time.

Pearl Harbor (1941): Negative Helicity in the Pacific.

American code-breaking supplied diplomatic information before Pearl Harbor. Turning indications of a coming war into a warning about the place, time, and form of an attack presented a different problem. The broader intelligence literature examines the gap between possessing information and organizing an effective response (Betts 1982).

The US oil embargo of 1941 sharply constrained Japan's supply options. In this book's interpretation, the relevant pressure came from a

dependence that survived the loss of recorded trade. Reserves continued to run down. The distinction between shipments and remaining alternatives becomes central in Chapter 10.

Japan's military planners faced a shrinking fuel horizon if they maintained their commitments. Those commitments were themselves political choices. Withdrawal, concessions, and armed expansion carried very different costs. The constraint helps explain the decision problem without making an attack on Pearl Harbor a mathematical necessity.

The relevant questions extend across institutions. Which indications reached which commanders? What did they imply under the plans then in force? Which preparations would have followed from another interpretation? Structural explanation must include those channels of decision and responsibility, rather than treating an attack as the automatic consequence of a resource constraint.

Yom Kippur War (1973): The "Concept" as Static Geometry.

Israeli intelligence had information about Arab military preparations, including human and signals reporting. The central question was how those observations were interpreted: as preparations for an attack, as an exercise, or as pressure short of war. Gathering a fact and understanding its significance are different tasks.

The famous "Concept"—Israeli intelligence's assessment that Egypt would not attack without first neutralizing Israeli air superiority—was a *static* geometric claim masquerading as dynamic analysis. It defined a fixed configuration (air superiority absent) as necessary for attack, rather than tracking how the overall geometric relationship was evolving.

But Anwar Sadat had reconceptualized the problem. Egypt's objective was not military victory but political transformation—breaking the diplomatic stalemate frozen since 1967. The helicity of the Egypt-Israel relationship had reached a point where *any* disruption of the status quo served Egyptian interests. The "Concept" measured one variable (air capacity); the geometry required tracking the entire configuration of military, economic, and diplomatic vectors.

When Egyptian and Syrian forces attacked on October 6, 1973, their preparations had not been converted into an adequate defensive response. The case requires attention to both the observable capabilities and the assumptions used to interpret them. A stock of information is not the same thing as a judgment that leads to action.

The Soviet Collapse (1989): Measuring Capability Without Curvature.

The Soviet collapse posed a different warning problem. Estimates of current economic and military capacity had to be connected to the ability of the political system to hold together. A large inventory of

weapons could coexist with weakening authority over the institutions that controlled them.

The failure illustrates a different geometric blind spot: measuring the *magnitude* of system variables without tracking their *curvature*—the rate at which trajectories were bending. Soviet GDP was declining, but analysts extrapolated linear trends. Soviet control was weakening, but analysts assumed continuous degradation rather than phase transition.

In geometric terms, the Soviet system was approaching a *singularity*—a point where the manifold of possible states tears rather than deforms. The variables analysts tracked (economic output, military spending, political control) were all changing at rates that suggested decades of decline. But the *second derivatives*—the acceleration of decline—were increasing rapidly. The system was not sliding; it was curving toward collapse.

The analytical question is how a system loses coordination while its inventories still look substantial. Counting what it owns does not by itself measure whether it can continue to act as a system.

Russia's Crimea Annexation (2014): Ignoring Helicity's Warning.

The Crimea case similarly asks when movements, political claims, and a deteriorating relationship should have changed an assessment of possible action. The historical task is to identify which observations were available, which assumptions governed interpretation, and what a different procedure would have concluded before the annexation.

The geometric configuration was flashing red. Ukraine's potential pivot toward the EU represented a fundamental reorientation of alignment vectors. Russia-Ukraine helicity—already stressed by twenty years of post-Soviet friction—faced a discontinuous shock. The Maidan revolution wasn't just a political event; it was a geometric transformation that would flip Ukraine's alignment from Russian sphere to Western orientation.

Putin's response followed the geometric logic: if you cannot prevent the alignment flip, seize the territory before the flip completes. Crimea's annexation was structurally predictable once analysts recognized that helicity, not capability, was the relevant variable. Russia had the capability to annex Crimea in 2012 or 2010 or 2005. The capability didn't change. The geometry did.

The post-Cold War consensus that territorial conquest was obsolete in Europe proved badly mistaken—not because analysts lacked information about Russian capabilities, but because they lacked a framework for understanding when capability would translate into action.

The failures recur across a century, although the warning signs change. Angell's argument appeared in 1909; five years later, war engulfed the

deeply interdependent European powers. In 1941, extensive warning signs did not yield an effective anticipation of Pearl Harbor. In 1973, Arab mobilization did not overcome the assumptions that left Israeli intelligence unprepared for the Yom Kippur War. The Iranian Revolution in 1979 and the Soviet collapse in 1989 exposed another difficulty: institutions could watch a political order closely and still fail to anticipate its rupture.

The more recent failures are closer to this book's central problem. Russia's annexation of Crimea in 2014 caught Western analysts off guard. In 2022, warnings were abundant, yet the scale and commitment of the invasion still surprised planners. Seeing danger and understanding what a relationship has become are different achievements.

Why do we keep failing?

The pattern reveals something deeper than incompetence. Each failure involved sophisticated analysts with access to relevant information. The failures cluster around *discontinuities*—moments when systems transition rapidly from one state to another. Continuous extrapolation ("tomorrow will be like today, only more so") fails precisely at the moments that matter most.

Jervis's analysis of intelligence failure makes interpretation and organizational assumptions central to the problem Jervis 2010. An observation acquires meaning within a model of how the other side behaves. That model needs testing most urgently when a familiar signal may have acquired a different meaning.

A warning desk also has to look beyond the events it can already describe. Knowing where forces moved yesterday is valuable. Estimating what they will do next requires a model of capabilities, objectives, constraints, and choices. More observations improve that model only if the analyst can connect them to the decision at issue.

An institution can be good at tracking a familiar threat and still fail to recognize that the relationship generating it has changed. Monitoring the parts and explaining the transformation of the whole are related tasks, but neither substitutes for the other.

Mueller and Rauh (2022), in a landmark study for the *Journal of the European Economic Association* Mueller and Rauh 2022, call this "the hard problem of prediction": conflict onsets in previously peaceful areas "remain particularly difficult to anticipate using structural predictors." Structural variables change slowly; conflicts emerge rapidly. The gap between the timescales of cause and effect defeats methods designed for gradual change.

Consider the methodological challenge concretely. A study by the Peace Research Institute Oslo estimated less than 0.05% probability of

conflict in Syria in early 2011. War began that year, eventually killing over 500,000 people and displacing half the population. The model was not wrong about Syria's structural characteristics; it was blind to the *dynamics* that would transform those characteristics into violence. Syria in January 2011 and Syria in December 2011 had similar GDP, similar regime type, similar ethnic composition. What changed was not the variables but the *geometry* of their interaction—the configuration of grievances, opportunities, and expectations that suddenly made violence possible.

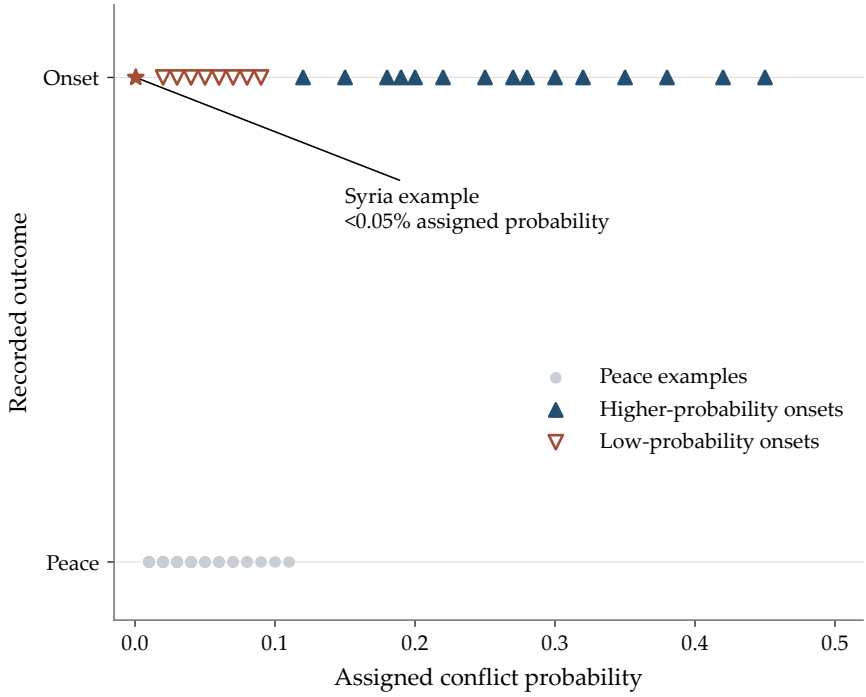


Figure 1.3: The hard-onset problem, shown schematically. Filled and open triangles distinguish the higher- and lower-probability onset examples; circles indicate peace examples. The displayed coordinates illustrate the problem discussed in the text and do not constitute an independently sampled evaluation set.

What the usual measurements miss

Modern conflict prediction suffers from three fundamental blind spots, each rooted in methodological choices that seemed reasonable but have proven inadequate. Understanding these blind spots is essential for understanding why a geometric approach offers a path forward.

Blind Spot 1: We measure the wrong variables.

Traditional models focus on material capabilities—troop counts, GDP ratios, military spending. But wars don't happen because one side has more tanks. Wars happen when structural relationships break down. The Political Instability Task Force found that regime type—specifically factionalism in partial democracies—outperforms all material variables in predicting state failure. The critical dynamics are relational, not material.

The fixation on material capabilities reflects a deeper theoretical commitment: realism, the dominant paradigm in international relations theory, treats states as rational actors pursuing power in an anarchic system. Power is measured in material terms—military spending, troop strength, GDP, nuclear warheads. The intuition is that states with more power can coerce states with less power; therefore, power imbalances predict conflict.

But the empirical record is unkind to this intuition. Powerful states frequently refrain from attacking weaker neighbors despite clear capability advantages. Conversely, weaker states sometimes initiate conflicts against stronger adversaries—Argentina's 1982 invasion of the Falklands, Iraq's 1990 invasion of Kuwait, Russia's 2022 invasion of Ukraine (where initial Russian assessments underestimated Ukrainian resistance capacity).

What distinguishes cases where capability advantages translate into aggression from cases where they don't? The answer, this book argues, lies in the *relationship*—the configuration of economic ties, diplomatic alignment, and shared expectations that determine whether coercion seems necessary and feasible.

A missile count has a relatively clear unit. The credibility of a promise does not. The temptation is to let what can be counted most easily stand in for what is hardest to know. A careful model must expose that substitution.

Trust, humiliation, perceived betrayal, and expectations of abandonment can influence decisions without appearing as quantities in a standard trade ledger. Bringing them into a model requires observable indicators and uncertainty about what those indicators mean. Giving a judgment a number does not remove that uncertainty.

The Political Instability Task Force's findings on regime type illustrate the point. Democracies rarely fight each other (the "democratic peace"); autocracies show higher propensity for external conflict; but the highest conflict rates appear in "anocracies"—partial democracies with mixed institutional features. This finding cannot be explained by material capabilities alone. It points to *structural* characteristics of the relationship between state and society that shape propensity for external violence.

Blind Spot 2: We confuse correlation with causation.

Machine learning models can identify patterns that *correlate* with conflict onset, but they cannot explain *why* those patterns matter. Schincariol et al. (2025) Schincariol et al. 2025 documented that ML forecasting systems produce "conservative forecasts gravitating toward the mean," systematically missing "sudden surges and declines." The models optimize for average performance, not for the tail events that matter most.

The machine learning revolution has transformed conflict forecasting, enabling pattern recognition at scales impossible for human analysts. Modern systems ingest thousands of variables—economic indicators, event data, news sentiment, satellite imagery—and identify statistical regularities that predict conflict onset.

Yet these systems exhibit a fundamental limitation: they identify *what* predicts conflict without explaining *why*. A random forest might identify that a particular combination of GDP decline, ethnic fractionalization, and neighboring conflict correlates with civil war onset. But the model provides no theory of *mechanism*—no explanation of the causal pathway from these predictors to violence.

A probability can be useful without identifying an intervention. A decision maker also needs to know which observations support it and how plausible actions might change the situation. Feature rankings can help inspect a prediction, but they do not establish that manipulating a highly ranked variable will improve the outcome.

The distance from analysis to action is concrete. A ministry needs to know whether to build reserves, arrange transport, reopen talks, or prepare for a disruption. Each action has a cost and a completion time. An explanation that cannot reach those choices is incomplete as a decision aid.

This matters because correlations can shift. A pattern that held in Cold War-era data may not hold in the post-2008 world of social media, economic globalization, and great power competition. Without understanding mechanisms, models cannot adapt to changing contexts.

More practically, ML models optimize for metrics like AUC-ROC (area under the receiver operating characteristic curve) that reward average

performance across all cases. But conflict prediction is fundamentally a *tail event* problem—what matters is predicting the rare, high-consequence cases that deviate from normal patterns. Schincariol et al.'s finding that ML systems produce "conservative forecasts" reflects this optimization: models that predict moderate probabilities for all cases will outperform on average metrics while systematically missing the extremes.

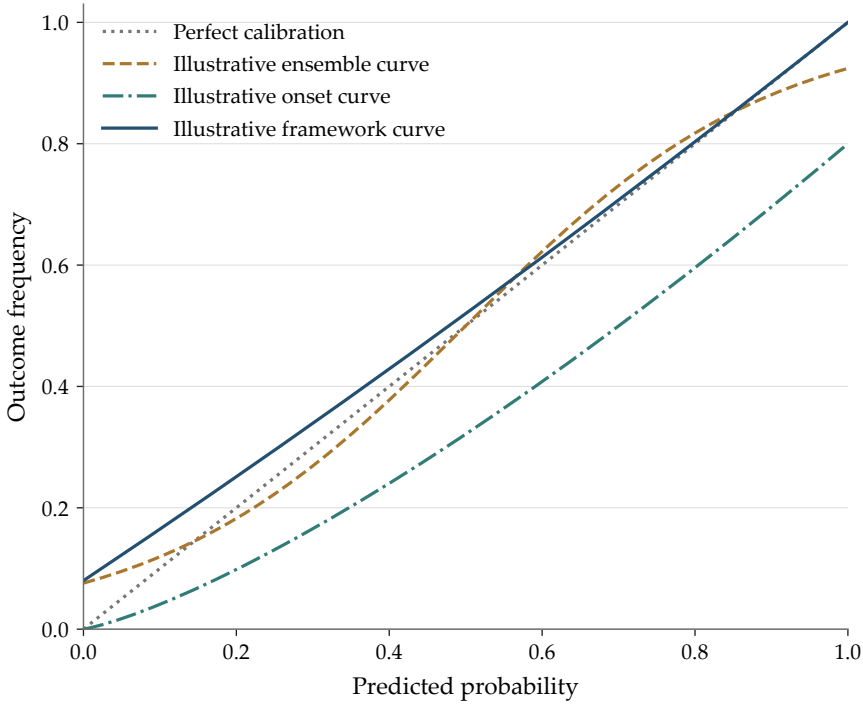


FIGURE 1.4: Calibration as a reading problem. The diagonal is perfect calibration. Above it, observed frequency exceeds assigned probability; below it, assigned probability exceeds observed frequency. The three other curves are prescribed illustrations of different shapes, not calibration estimates from held-out predictions.

For example, gather assessments that each assigned a ten-percent probability to the same kind of event over the same horizon. If roughly one in ten such events occurs, that group is calibrated. If three in ten occur, the curve sits above the diagonal: the assigned probabilities were too low. This is a check on what a probability means across comparable assessments. It is separate from whether a warning arrived early enough to change a decision.

Blind Spot 3: We ignore topology.

International relations theory has been dominated by what we might call *metrical* thinking—measuring distances, sizes, and magnitudes. But the critical dynamics of conflict are *topological*—they involve twists, breaks, and discontinuities that cannot be captured by continuous metrics. When trust inverts from positive to negative, something *tears* in the fabric of the relationship. This book provides the mathematics to describe that tearing.

The deepest blind spot is mathematical. Contemporary political science relies heavily on linear regression and its extensions—methods designed for relationships where small changes in inputs produce proportionally small changes in outputs. Conflict onset is fundamentally nonlinear: the transition from peace to war involves discontinuity, phase transition, and threshold effects that linear methods cannot capture.

Suppose a buyer has one licensed route to a supplier. The licensing office reduces its weekly quota from ten shipments to five, then to one. The weights change, but a route still exists. When the last permit is withdrawn and no other route is available, the buyer can no longer reach the supplier at all. Capacity can fall continuously to zero while reachability changes from yes to no. That is the distinction topology helps us see. A trust score crossing zero can mark a change of orientation, but to claim that a path has actually closed we must identify what can no longer be done: inspect a shipment, settle a payment, invoke a guarantee, or obtain the permit.

Topological concepts like "connectedness," "holes," and "twisting" provide vocabulary for describing relationship dynamics that continuous metrics miss. When economic interdependence and diplomatic alignment point in opposite directions—high trade with deteriorating trust—the relationship exhibits a kind of "twist" analogous to the helical structure of a corkscrew. This twist creates stress that must eventually resolve: either alignment recovers or the economic ties rupture.

A route that looks available on a map may require a concession that cannot be delivered, a terminal that has not been built, or a partner who will not participate. The absence of a small next step can matter even when the destination is easy to describe.

The boundary between regimes is therefore worth examining directly. What changes when a substitute is no longer available, a guarantee ceases to be credible, or a reserve can no longer cover the interruption? A small change in an input can then have a large practical consequence.

A singularity in a mathematical model has to be identified through its definitions and evolution law. A sudden political rupture may motivate such a model, but the name alone explains nothing. The useful work is

to identify which regular response is being lost and what observation would show that loss approaching.

This is why the warning problem includes discontinuities. A military inventory can change gradually while a political decision changes the use to which it will be put overnight. Tracking the inventory is necessary. Understanding the commitments and choices surrounding it is necessary too.

Cederman and Weidmann, writing in *Science* (2017) Cederman and Weidmann 2017, crystallized the methodological crisis: structural variables like GDP and regime type are "slow-changing and poorly suited for short-term prediction." We need variables that capture the *dynamics* of relationship evolution—how the configuration of economic and diplomatic ties changes over time, independent of specific triggering events. Topology provides the mathematical framework for defining such variables.

If the problem of conflict prediction is so hard, and the failures so consistent, why should this particular framework succeed where others have failed? The answer lies in a convergence of capabilities that simply did not exist even twenty years ago.

The first enabler is data. The framework developed in this book requires granular, time-series information on bilateral trade flows, diplomatic events, financial linkages, and conflict outcomes. This data ecosystem has matured dramatically since the early 2000s.

The Uppsala Conflict Data Program (UCDP) now catalogs armed conflicts globally with monthly resolution and geographic precision. The Global Database of Events, Language, and Tone (GDELT) processes millions of news articles daily, extracting structured event data that can track diplomatic temperature in near-real-time. The Armed Conflict Location and Event Data (ACLED) project provides granular event-level tracking. The Correlates of War project, supplemented by newer datasets, enables validation against 200 years of interstate conflicts.

On the economic side, the IMF's Direction of Trade Statistics provides bilateral trade flows for nearly every country pair, quarterly since 1950. The Bank for International Settlements tracks cross-border financial exposures. The UNCTAD foreign direct investment database captures ownership linkages. Together, these datasets enable the construction of helicity indicators that would have been computationally infeasible a generation ago.

More observations do not automatically yield a better explanation. An hourly update is valuable only if its quantities represent the relevant process and the apparent precision survives the delays, revisions, and omissions in its sources.

The second enabler is computation. Modeling bilateral dynamics means tracking relationships across approximately 20,000 possible country pairs. Computing geometric indicators for each dyad, updating them with new data, and identifying threshold crossings requires processing power that simply wasn't available outside specialized government facilities until recently.

Cloud computing has democratized access to the computational resources needed for global-scale geometric analysis. A single researcher with AWS credits can now run analyses that would have required supercomputer time in 2005. Graph neural networks enable modeling of the higher-order structures—how dyadic relationships connect into networks with their own emergent properties. Topological data analysis, a field that barely existed before 2010, provides the mathematical tools for detecting phase transitions in complex systems.

The framework in this book was developed on a laptop. It runs on cloud infrastructure costing less than a university department's annual software budget. This accessibility is new.

The third factor is urgency. The post-Cold War period lulled the international relations field into complacency about great power conflict. From 1991 to roughly 2014, the dominant assumption was that interstate war between major powers had become obsolete—a relic of a less enlightened age. Academic attention shifted to civil wars, terrorism, and peacekeeping.

That assumption died in Crimea and was buried in Ukraine. The 2020s present simultaneous flashpoints of a type not seen since the Cold War's darkest moments:

Taiwan Strait: China's military modernization and Xi Jinping's statements about "reunification" have made the possibility of cross-strait conflict a central concern of U.S. defense planning. The geometric configuration—massive economic interdependence between China and Taiwan combined with increasingly hostile political postures—exhibits the trap signature this book describes.

India-Pakistan: The world's only nuclear dyad with an unresolved territorial dispute and history of repeated wars. In 2019, bilateral trade collapsed to near zero following India's revocation of Kashmir's autonomy. The relationship now exhibits a different geometric configuration—near-zero economic interdependence *and* hostile alignment. This is not the trap; this is something potentially more dangerous, where nuclear deterrence becomes the *sole* remaining constraint.

Israel-Iran: A shadow war of assassinations, cyberattacks, and proxy conflicts that could escalate at any moment. The 2024 direct exchange of

missile and drone strikes marked a new phase. No meaningful economic interdependence exists to create even the perverse stability of the trap.

Russia-NATO: The Ukraine war has transformed European security, but the geometric configuration remains unresolved. NATO's eastern members now exhibit the highest helicity risk indices in Europe.

The question is no longer academic. Tools for anticipating interstate conflict have become genuinely urgent in ways they were not in 2005 or even 2015.

The India-Pakistan relationship deserves particular attention because it illustrates a failure mode distinct from the Russia-Ukraine trap.

When India revoked Kashmir's special status in August 2019, Pakistan responded by suspending bilateral trade entirely. Cross-border commerce, already minimal, dropped to effectively zero. Diplomatic relations, already hostile, became poisonous. The helicity of the relationship was not negative—it was *undefined*. Both the economic flow vector and the alignment vector approached zero magnitude, leaving the relationship with no structural stabilizers except the threat of nuclear annihilation.

Then, in May 2025, came Operation Sindoor.

We will examine this case in detail in later chapters. For now, the point is methodological: the framework must handle multiple geometric configurations, not just the "trap" of negative helicity. India-Pakistan represents the limit case where interdependence and alignment both collapse, leaving a relationship balanced on the edge of catastrophe with no economic cushion and no diplomatic channel—only the fear of mutual destruction.

The conventional literature on economic interdependence and peace cannot explain India-Pakistan. The relationship has no interdependence; it also has no peace, only the Cold Peace of nuclear terror. What geometry reveals is that this configuration has its own dynamics, its own failure modes, its own signatures of approaching crisis. Different configurations require different interventions.

The instrument we need

This book introduces a new framework for understanding when relationships between nations reach critical instability. Drawing on differential geometry, fluid dynamics, and the topology of complex systems, it demonstrates that:

The argument begins with a claim about what can be observed: conflict has geometric signatures that can become detectable months or years before an outbreak. Trade matters, but the amount of trade does not settle

the question. The alignment between economic flows and diplomatic trust determines whether interdependence supports a relationship or tightens a trap. The book develops the further thesis that wars occur at phase transitions with mathematical structure, and asks the historical cases—from Pearl Harbor to Russia's invasion of Ukraine—to show how that structure appears in events.

The framework seeks to identify relationships in which continuing dependence and adverse alignment leave fewer usable responses to a crisis. A highest-level warning calls for an explanation of those constraints and a search for alternatives. It cannot establish that war is certain.

The core innovation is a quantity called *helicity*—borrowed from fluid dynamics, where it measures the "handedness" or twist of a flow.

Consider a river. Water flowing downhill follows the path of least resistance. The velocity of the water and the slope of the terrain point in the same direction—the flow is *aligned*. This is a stable configuration. Energy dissipates smoothly. The river reaches the sea.

Now imagine that same water backed up behind a dam. The potential energy of the water (wanting to flow downhill) and the physical constraint of the dam (blocking the flow) point in opposite directions. The configuration is *misaligned*. Pressure builds. The dam can hold indefinitely—until it doesn't. When it breaks, the release is catastrophic, disproportionate to any immediate trigger.

International relationships work the same way.

When economic interdependence and diplomatic trust are aligned—either both positive (prosperous partnership) or both negative (hostile separation)—the relationship is geometrically stable. Energy flows in consistent directions. Disruptions dissipate.

But when economic ties and diplomatic trust point in opposite directions—high interdependence persisting despite deteriorating trust—the relationship is *twisted*. The flow cannot proceed smoothly. Pressure accumulates at the twist. And like water behind a dam, the relationship can persist in this stressed state for years, even decades, until some perturbation triggers release.

Helicity quantifies this twist. Positive helicity means alignment—economic and diplomatic vectors rotating in the same direction, like a right-handed screw advancing smoothly. Negative helicity means misalignment—the vectors rotating in opposite directions, creating structural stress that must eventually resolve.

The German-Russian relationship exhibited negative helicity for years before 2022. Trade expanded while trust eroded. The twist tightened. February 24 was the dam breaking.

In the 47-event evaluation detailed in Appendix E, the working paper reports an F_1 of 0.73 and a mean lead of 72 days. The 18 detections in the separate historical roster have a 127-day mean (Appendix C). A reader can follow an individual warning back from its date to the goods that kept moving and the political channels that closed.

A machine learning model might assign a 67% probability of conflict to a given dyad without explaining why. The helicity framework explains *what* in the relationship has become structurally unstable: the divergence between economic interdependence and diplomatic trust, the rate at which that divergence is increasing, the comparison to historical configurations that preceded past conflicts.

Interpretability matters when a reading leads to a decision. A geometric diagnosis can point toward a replacement supplier, a negotiable commitment, a diplomatic channel, or a need for reserves. The proposed response should be evaluated for its costs and effects, including what happens if the warning is wrong.

Let me be direct about what this book attempts.

For a century, the study of war has been divided between historians who explain particular conflicts in rich contextual detail and political scientists who seek generalizable patterns through statistical analysis. The former produce wisdom that doesn't travel; the latter produce correlations without causes. Neither has solved the prediction problem.

This book proposes a third path: *geometric* analysis that identifies structural configurations associated with conflict, explains *why* those configurations generate instability, and provides indicators that can be tracked in real time. The mathematics is rigorous. The physics metaphors are not merely decorative—they capture genuine structural features of bilateral relationships. The historical record provides validation across diverse contexts.

Audre Lorde warned that "the master's tools will never dismantle the master's house" (Lorde 1984). The tools of conventional security analysis—capability assessment, rational choice modeling, linear regression—are the master's tools. They were designed for a world of gradual change, continuous variables, and rational actors. The world of conflict is not that world. It is a world of phase transitions, topological tears, and leaders who choose catastrophe.

This book offers different tools. Whether they prove adequate to the task is for history to judge. What follows is the mathematical architecture that produces predictions, the physical intuition that makes the mathematics meaningful, and the historical evidence that demonstrates the framework's validity across diverse contexts.

By the end of this book, you will see international relations through a new lens—one where the geometry of trust and trade reveals patterns invisible to conventional analysis. You will understand why December 14, 2021 mattered, why Germany's trap was structural rather than political, and why the simultaneous flashpoints of the 2020s require tools that previous generations could not have built.

The geometry of war is not a crystal ball. It is a diagnostic instrument—one that can distinguish the stable from the stressed, the aligned from the twisted, the containable crisis from the structural rupture. In a world of cascading risks, that distinction may matter more than any prediction.

The problem now has a sharper form. If the shipment is still arriving, what would tell us that the relationship has changed? To answer, we have to return to the promise on which so much policy was built: that commerce would make war too costly to choose.

Chapter 2

The Great Debates

> Le doux commerce... adoucit les moeurs barbares.
> (Gentle commerce... softens barbarous customs.)
>
> Montesquieu, *De l'esprit des lois* (1748)

A factory needs its next delivery. Its supplier needs the next payment. Both have reasons to keep the border open. Follow that reasoning far enough and commerce begins to look like an argument against war. Yet the same factory can become a hostage to the delivery it cannot replace.

The oldest question in international political economy is whether economic interdependence promotes peace. The debate traces to Montesquieu's 1748 observation that "peace is the natural effect of trade" and has continued for nearly three centuries without resolution.

The core liberal argument traces to Immanuel Kant's 1795 essay *Perpetual Peace* Kant 1795, which proposed that "the spirit of commerce" would eventually make war unprofitable. Kant's argument was not merely economic but moral: commerce creates cosmopolitan ties that transcend national boundaries, making citizens of different states see each other as partners rather than enemies. "If the consent of the citizens is required to decide whether or not war is to be declared," Kant reasoned, "it is very natural that they will have great hesitation in embarking on so dangerous an enterprise."

Russett and Oneal's *Triangulating Peace* examines democracy, economic interdependence, and international organizations as mutually relevant influences on interstate conflict Russett and Oneal 2001. Its empirical argument gives the liberal position a form that can be compared across cases. It does not make a particular bilateral relationship immune to war.

The statistical evidence seemed compelling. Hegre, Oneal, and Russett (2010) Hegre, Oneal, and Russett 2010, using simultaneous estimation to address endogeneity concerns, concluded that "economically important trade has a substantively important effect in reducing dyadic militarized disputes." Their analysis of the Correlates of War dataset (1870–2001)

showed that a one standard deviation increase in bilateral trade reduced conflict probability by approximately 20%.

The mechanism seemed clear. Trade creates mutual economic benefits. War destroys those benefits. Rational leaders, calculating the costs and benefits of conflict, would see that war rarely makes economic sense. Thomas Friedman captured the popular version of this logic in his famous "Golden Arches Theory of Conflict Prevention": no two countries that both had McDonald's had fought a war against each other since getting their franchises. The formula was elegant: Trade $\Rightarrow$ Mutual Gains $\Rightarrow$ Peace.

Katherine Barbieri's 1996 *Journal of Peace Research* article Barbieri 1996 detonated a bomb in this comfortable consensus. Analyzing the same data with different specifications, she found a *curvilinear* relationship: low-to-moderate interdependence reduced disputes, but *extreme* interdependence actually *increased* conflict probability. The relationship was not monotonic; it was an inverted U.

A result in the opposite direction forces a closer examination of what is being measured. Dependence can create a constituency for restraint while also exposing a country to coercion. An aggregate association can mix those mechanisms.

Her argument centered on *asymmetric dependencies*. When Country A depends heavily on Country B for critical goods, but B has diversified suppliers, the asymmetry creates leverage for B and resentment in A. Trade becomes a source of vulnerability rather than mutual benefit. The very ties that should promote peace become instruments of coercion.

Historical cases make a deterministic reading of commercial peace difficult to sustain. Britain and Germany had substantial commercial ties before 1914. Soviet–Chinese cooperation did not prevent their later split. The Iraq–Kuwait confrontation linked financial obligations, oil, and territorial demands. These episodes have different mechanisms, but each requires an explanation of why economic interests failed to keep the relationship peaceful.

A financial connection does not disappear the moment the political relationship changes. A lender still has an outstanding claim; a borrower still needs liquidity. The same outstanding transaction can therefore remain important while the expectations that supported it are being overturned.

Erik Gartzke's 2007 "capitalist peace" thesis in the *American Journal of Political Science* Gartzke 2007 offered a different resolution. The problem, Gartzke argued, was that researchers conflated different types of economic integration. Trade in goods matters less than *financial openness*—the free movement of capital across borders.

Financial openness and trade in goods place different commitments at risk. A payment channel, an investment position, and a physical shipment have different exit costs. Their distinct mechanisms should survive the construction of an economic-integration measure.

Financial integration promotes peace through a different mechanism: credible signaling. When countries have open capital accounts, their economic policies become visible to international markets. This transparency reduces the information asymmetries that game theorists identify as a primary cause of war. States cannot easily conceal mobilization or hostile intentions when capital markets are watching.

Capital flight serves as an early warning system. If investors perceive war risk, they withdraw capital, inflicting economic pain that incentivizes de-escalation. The 1914 July Crisis saw massive gold outflows from London and Berlin—the markets saw war coming even as diplomats denied it. But capital mobility was insufficient to prevent war because the political commitment to conflict had already been made.

A market can price a disturbance before a government acts on it. That does not guarantee early or accurate warning in every crisis. It gives the researcher another dated series to compare, with its own incentives, observation limits, and false alarms.

Dale Copeland's 2015 book *Economic Interdependence and War* Copeland 2015 introduced a critical temporal dimension missing from earlier debates. What matters, Copeland argued, is not the current level of trade but *expectations of future trade*. States go to war not when trade is high but when they expect trade to *decline*—when they fear losing access to markets or resources they currently depend upon.

This insight reframes the historical puzzle of World War I. The great powers of 1914 were deeply interdependent, yet they went to war. But German leaders perceived potential encirclement: British threats to block Persian Gulf oil access, proposed tariff walls around the British Empire, and Russia's rapid industrial expansion. The expectation of *declining* future trade, not the reality of high current trade, drove the decision for war.

Copeland's theory predicts that war becomes more likely when:

The expectation of continued trade can deteriorate before trade itself disappears. A rising power may threaten access to resources or markets. A trading partner may make that access conditional on political concessions, or a leadership transition may put previously dependable policy in doubt. Financial crises and commodity disruptions can add a further shock. In each case, present commerce remains substantial while confidence in its future weakens.

This is a theory of *anticipatory war*—conflict initiated not because of current deprivation but because of feared future deprivation. The preventive war logic: "Better to fight now while we still can than later when we cannot."

Position	Studies	Percentage
Trade promotes peace	~57	60%
Null or contrary	~38	40%

Table 2.1: Distribution of findings in 95 studies on economic interdependence and war (systematic review, *Cogent Social Sciences*, 2025).

A 2025 systematic review in *Cogent Social Sciences* surveyed 95 articles published between 1977 and 2022. The verdict: no consensus. Sixty percent found that trade promotes peace; forty percent found null or contrary effects. The debate, nearly three centuries old, remains unresolved.

Adding another control variable may improve an estimate without resolving what the estimate measures. If one study records total trade and another records dependence on an irreplaceable supplier, they are testing different mechanisms. The apparent disagreement invites a closer definition of the relationship.

Context changes the mechanism. A source of imports may be replaceable in one country and indispensable in another. Political institutions may make a concession durable or easily reversed. Averaging such relationships can conceal the very difference that explains their outcomes.

A yes-or-no question about trade is too coarse for that task. A household with several sources of fuel faces a different problem from one whose only supplier is threatening to cut delivery. The amount purchased may be the same. The freedom to respond is not.

The problem, as we will see, is that the literature has been measuring the wrong quantity. Trade volume captures the *magnitude* of economic flows but not their *structure*. What matters is not how much countries trade but how those flows relate to political alignment—what the helicity framework will call the "topological character" of interdependence.

The problem is that researchers have been asking the wrong question. They have been asking *whether* trade affects conflict, when the real question is *when* trade affects conflict—and in which direction. The answer, as we shall see, depends on something none of these studies measured: the *alignment* between economic flows and diplomatic trust.

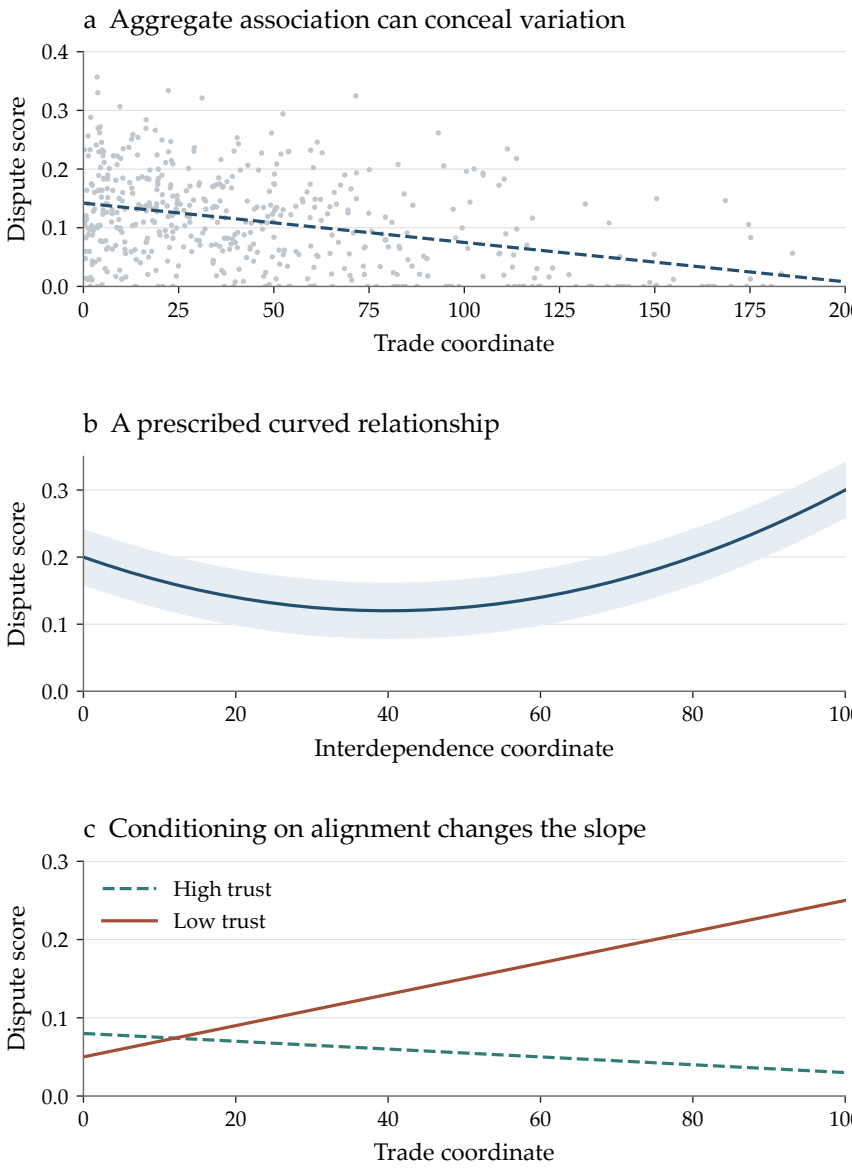


FIGURE 2.1: Three constructed views of the trade-peace puzzle. (a) Five hundred seeded synthetic points illustrate a noisy aggregate association. (b) A prescribed curved response and illustrative band. (c) Two prescribed slopes conditional on alignment. These panels explain competing mechanisms; the shaded band is not an estimated confidence interval.

The sharper question is when a commercial relationship expands usable alternatives and when it concentrates vulnerability. The helicity framework develops one way to investigate that distinction.

To see why the existing theories fall short, consider a single dyad: Russia and Germany from 2008 to 2022. Each of the major theorists would make a prediction. None would get the full picture.

Russett and Oneal would observe high bilateral trade—€45 billion in 2013—and predict stability. Germany was one of Russia's largest trading partners; Russia was Germany's primary energy supplier. The democratic peace component fails (Russia is not a democracy), but the commercial peace logic should apply: mutual economic benefits should create stakeholders for peace in both societies. *Prediction: low conflict probability.*

Barbieri would note the asymmetric dependence and raise a warning flag. Germany depended on Russia for 55% of its natural gas, while Russia had alternative customers (China, India) and could endure a European cutoff longer than Europe could endure losing Russian gas. This asymmetry created leverage for Moscow and vulnerability for Berlin. *Prediction: elevated risk due to asymmetric dependence.*

Gartzke would examine financial integration and find it moderate. German firms invested in Russia (Siemens, Volkswagen, BASF), but capital markets were not fully integrated. Russia maintained capital controls; European banks had limited exposure to Russian sovereign debt. The credible signaling mechanism was weak. *Prediction: ambiguous—not enough financial integration to create transparency, but not zero either.*

Copeland would ask about expectations of future trade. Through 2021, German policy assumed *increasing* interdependence: Nord Stream 2 would double gas pipeline capacity, locking in the relationship for decades. Russian elites presumably expected continued European revenue. By Copeland's logic, shared expectations of expanding trade should have stabilized the relationship. *Prediction: peaceful, because both sides expected the economic relationship to deepen.*

What actually happened?

Russia invaded Ukraine in February 2022. Germany reversed twenty years of policy in seventy-two hours: halting Nord Stream 2, committing €100 billion to military modernization, shipping weapons to Kyiv. The trading relationship collapsed. By 2023, German gas imports from Russia had fallen to under 4%.

Where did the theories go wrong?

Russett and Oneal missed that trade volume says nothing about whether the relationship is *healthy*. You can have enormous trade while

trust is collapsing—the commercial relationship can persist long after the political relationship has turned hostile.

Barbieri identified the asymmetry but couldn't explain *timing*. The asymmetric dependence existed in 2008, in 2014, in 2020. Why did conflict emerge in 2022 and not earlier?

Gartzke's financial integration thesis was simply inapplicable—the relationship was built on commodity trade (gas, oil, metals), not capital market integration.

Copeland came closest. His framework would note that Russian expectations of future trade shifted after 2014 Crimea sanctions, when the West demonstrated willingness to impose economic costs. But German expectations remained stable through 2021—*Germany* still expected deepening integration even as Russia began hedging. The expectations diverged without either side fully recognizing the divergence.

What a helicity measure would have shown: this relationship entered negative alignment in 2014 and never recovered. Economic ties remained high, but trust inverted. The helicity went negative—indicating a *trap* configuration where interdependence persists despite hostile alignment. The trap deepened every year from 2014 to 2021 as trade volumes recovered while diplomatic relations continued deteriorating.

By December 2021, the geometric configuration was flashing red: maximum negative helicity, indicating that economic ties and political alignment were rotating in opposite directions at maximum intensity. The system was stressed beyond its capacity to self-correct.

None of the individual theorists saw this because each was measuring one dimension of a multi-dimensional object. Russett measured trade volume. Barbieri measured asymmetry. Gartzke measured financial openness. Copeland measured expectations. *Helicity integrates all four*—it asks not just "how much trade?" but "how does trade volume relate to political alignment, and how is that relationship evolving over time?"

This is not hindsight. The helicity framework, applied to 2020 data, would have identified Russia-Germany as a high-risk dyad—not because trade was low (it wasn't), not because asymmetry was high (many dyads have asymmetric dependence), but because the *geometric configuration* matched the historical signature of relationships that precede conflict.

The Russia-Germany case is not merely an academic illustration. It represents a twenty-year natural experiment in which an entire nation's foreign policy establishment bet on the liberal peace hypothesis—and lost.

Wandel durch Handel—"change through trade"—was the doctrine that guided German Russia policy from the early 2000s until February 2022. The theory was explicitly liberal: by integrating Russia into the European

economic order, Germany would create stakeholders for peace within Russian society. Oligarchs enriched by Western trade would lobby against adventurism. Energy interdependence would make conflict unthinkable.

Germany's continuing energy relationship with Russia made the dispute practical. A pipeline carries commercial value, but it also takes years to replace. A policy of engagement has to be judged by the alternatives it preserves when political relations worsen, as well as the business it enables when they improve.

German industry was fully invested. Siemens maintained turbine contracts. BASF held major stakes in Russian petrochemical facilities. Volkswagen operated plants near Moscow. The commercial ties were deep, reciprocal, and—according to the theory—stabilizing.

Then the theory failed.

The invasion forced a rapid reassessment of German policy. Scholz's February 27 statement announced a major shift in defense and security commitments (Scholz 2022). Energy exposure could not be changed by declaration alone. Replacement suppliers, transport capacity, and the time needed to obtain them became immediate practical concerns.

The amount moving through a pipeline does not reveal whether either side expects the contract to survive the next crisis. An account of the relationship needs both the physical dependence and the political conditions under which it can be used.

The 2025 *Cogent Social Sciences* review, showing a 60/40 split across 95 studies on whether trade promotes peace, represents the scholarly version of this confusion. Wandel durch Handel was the policy version. Practitioners discovering in real time what the academic literature could not resolve: the relationship between commerce and conflict depends on something none of the standard measures captured.

What the helicity framework would have shown: Germany-Russia was in a trap configuration for years. The commercial ties were real. The political trust was gone. The liberalizing effects that were supposed to flow from commerce never materialized because *alignment* had inverted even as *trade* continued. The doctrine failed not because trade is irrelevant to peace, but because trade without trust creates vulnerability, not stability.

Weaponized Interdependence

In 2019, Henry Farrell and Abraham Newman published a landmark article in *International Security* Farrell and Newman 2019 that reframed how we think about economic networks. "Weaponized Interdependence:

How Global Economic Networks Shape State Coercion" argued that the very infrastructure built for commercial efficiency had evolved into instruments of state power.

A network built for efficient exchange can concentrate control over access. The same central channel that reduces transaction costs can become the place where a state observes, delays, or blocks a counterparty's activity.

A network built to make transactions efficient can also concentrate control over them. If many firms depend on the same clearing channel, a restriction at that channel reaches farther than a restriction affecting one isolated transaction. The vulnerability lies in the arrangement of connections.

Global economic networks, Farrell and Newman observed, are not flat meshes of equal nodes. They evolved into "hub-and-spoke" systems with asymmetric power structures. A handful of central nodes—SWIFT for financial messaging, major internet exchange points for data, TSMC for advanced semiconductors—process a disproportionate share of global flows.

States with jurisdiction over these central nodes can weaponize them through two mechanisms:

The Panopticon Effect: Like Bentham's prison design where guards could observe all inmates without being seen, states controlling central nodes can gather strategically valuable information from data flows. The SWIFT network processes 42 million messages daily; controlling access means seeing the financial transactions of friends and adversaries alike.

The Chokepoint Effect: States can deny network access to adversaries, cutting them off from critical flows. Iran's 2012 disconnection from SWIFT cost the country approximately 50% of its oil export revenue overnight. Russia's March 2022 partial disconnection triggered a 30% ruble collapse within days.

Farrell and Newman's *Underground Empire* follows the political uses of infrastructure developed through private commercial activity Farrell and Newman 2023. For this book, the important dynamic is the response to coercion. Repeated use of a central channel as leverage can encourage others to build alternatives. Network position is a source of power whose use can change the network itself.

Weaponized interdependence theory explains how states *use* network power but not when they *will* use it. It describes a capability, not a trigger condition. Moreover, the theory focuses on the *weaponizer's* decision calculus while undertheorizing the *target's* response dynamics.

What's missing from the theoretical literature is the *timeline* of weaponization as a process—and how each act of weaponization teaches the next target to prepare alternatives.

2012: Iran disconnected from SWIFT. The first major weaponization of financial infrastructure. Iran lost approximately 50% of oil export revenue overnight. The message was clear: SWIFT connectivity was not a neutral commercial service but a potential weapon.

2014: Partial Russia sanctions after Crimea. The West demonstrated willingness to impose costs on a major economy. Russia's response was immediate: construction of SPFS (System for Transfer of Financial Messages), a domestic alternative to SWIFT. By 2022, SPFS handled over 20% of Russian domestic transactions.

2018: Huawei CFO arrest in Canada. The United States demonstrated extraterritorial reach—the ability to prosecute foreign nationals for violating U.S. sanctions, even when the conduct occurred outside American borders. China accelerated its semiconductor independence program within months.

2022: Full Russia disconnection from SWIFT. The weapon deployed at scale. Russian banks lost access to the global financial messaging system. The ruble collapsed 30% within days.

2022: Semiconductor export controls on China. The United States extended weaponization from finance to technology, restricting sale of advanced chips and chipmaking equipment to Chinese firms. The "entity list" became an instrument of industrial policy.

2023-25: Rare earth restrictions and countermeasures. China signaled willingness to restrict rare earth exports (it controls approximately 60% of global production). The weaponization dynamic became explicitly bidirectional.

The pattern is a ratchet, not a single event. Each weaponization creates negative helicity in the targeted dyad—the target now knows its economic ties are potential vulnerabilities. But each weaponization also *accelerates the target's search for alternatives*. Russia built SPFS because it watched Iran. China accelerated CIPS (Cross-Border Interbank Payment System) because it watched Russia. India began exploring alternatives after watching China.

In the helicity framework, this dynamic has a precise characterization: weaponization *inverts alignment* (turning economic ties from assets to liabilities) while the target's response represents *helicity correction* (reducing economic ties to match the new diplomatic reality). The question for policymakers is whether this correction stabilizes the system (by eliminating coercive leverage) or destabilizes it (by removing mutual stakes that constrained escalation).

That response takes time. A substitute payment route or supplier may be costly and initially inferior, yet become attractive when the existing channel is repeatedly threatened. A static map of centrality can miss this process of adaptation.

The very act of weaponizing interdependence creates what we will later formalize as *negative helicity*. Economic ties persist—indeed, the chokepoint effect *requires* the target to remain connected in order to be coerced. But trust inverts. The result is a configuration where high economic flux coexists with negative alignment: the signature of structural instability.

The evidence is already visible. Russia developed the SPFS payment system after the 2014 Crimea sanctions. China constructed CIPS and pursued RMB internationalization (now approximately 4% of SWIFT transactions). India began exploring non-dollar trade settlement after watching China's experience. Each represents helicity correction in action—and each weakens the very leverage that motivated the original weaponization.

The past decade has seen an explosion of computational approaches to conflict prediction. The technical achievements are impressive. But the fundamental limitation is not computational—it is conceptual.

A forecast target determines what a model learns to distinguish. Continuation of an existing war, onset of a new one, and structural exposure to coercion require different observations and evaluation rules. An argument for geometry must show its contribution on the task actually being claimed.

This distinction—between *eventive forecasting* and *structural forecasting*—is the key to understanding both the achievements and limitations of machine learning approaches.

Eventive forecasting asks: "Will there be a conflict in this location next month?" It predicts specific events at specific times. Success is measured by whether the predicted event occurs. This is what the ViEWS system, the Early Warning Project, and most ML conflict models attempt.

Structural forecasting asks whether a relationship has entered a configuration associated with elevated risk over a stated horizon. It examines the constraints that make adjustment difficult. Its result must still be evaluated against dated outcomes, including peace.

The difference is not semantic. It reflects fundamentally different theories of causation.

Consider the Russia-Ukraine case. An eventive forecast would attempt to predict: "Russia will invade on February 24, 2022." This requires modeling Putin's decision process, the specific trigger events, the op-

erational details—a near-impossible task given the opacity of Kremlin deliberations.

A structural assessment of Russia–Ukraine asks whether dependence persisted while alignment deteriorated and alternatives narrowed. The later reconstruction assigns a high-level warning to late 2021. To turn that finding into a forecast claim, the same rule must be run on information available before the outcome and scored against a specified event and horizon.

The helicity framework develops structural diagnostics and tests their warning value. Its proposed mechanism concerns constraints on adjustment. Whether those constraints help predict an onset, an escalation, or a failed negotiation is an empirical question for each defined target.

The comparison summarized here gives ViEWS PR-AUC values of approximately 0.65 for conflict continuation and 0.36 for hard onsets. Those figures concern different prediction tasks. A random-ranking reference for a precision–recall curve depends on the prevalence of positive outcomes; 0.36 cannot be called barely above chance without that denominator. A matched comparison must preserve the target, horizon, sample, and scoring procedure.

This is not a bug; it is a feature of eventive forecasting applied to rare events.

The fundamental problem is that ML systems learn from patterns in historical data. They can identify that certain combinations of variables have preceded past conflicts. But conflict onset is a tail event—99.5% of country-months show no conflict. A model that always predicts "no conflict" achieves 99.5% accuracy while providing zero actionable warning.

More deeply, ML approaches assume that future conflicts will resemble past conflicts. The algorithm learns the *signature* of previous onsets and looks for that signature in current data. But the next major conflict may have a different signature—not because the underlying dynamics are different, but because strategic actors have learned to avoid the observable precursors that would trigger warning systems.

A model trained on earlier crises can encounter a new combination of constraints. This is a reason to examine the mechanism and test transfer across periods. It does not exempt a structural model from the same risk of learning the peculiarities of its historical sample.

Even when ML models achieve accuracy, they fail to explain *why*. A random forest can identify that certain variable combinations predict conflict without revealing the causal mechanism. This matters for policy.

A black-box prediction offers no guidance for intervention. "There is a 67% probability of conflict in the Taiwan Strait" tells a policymaker nothing about *what to do*. Is the risk driven by military imbalance? Economic coercion? Domestic political pressure in Beijing? Leadership psychology? Without understanding the driver, there is no basis for choosing among deterrence, reassurance, economic engagement, or diplomatic signaling.

Structural forecasting, by contrast, provides *interpretable diagnostics*. A helicity measure showing negative alignment with high economic flux tells policymakers: the relationship is in the trap configuration. The levers are clear: reduce interdependence (managed decoupling), restore trust (diplomacy), or prepare for structural rupture (deterrence and resilience). Different configurations call for different responses.

Approach	Question	Policy Utility
Eventive (ML)	Will conflict occur on date X?	Limited: no causal mechanism
Structural (Helicity)	Has system crossed threshold?	High: identifies configuration

Table 2.2: Comparison of eventive and structural forecasting approaches.

Phase Transitions and Tipping Points

In 2009, Marten Scheffer and colleagues published a landmark paper in *Nature* Scheffer et al. 2009 that would eventually reshape how we think about sudden changes in complex systems. "Early-warning signals for critical transitions" demonstrated that systems approaching tipping points exhibit characteristic signatures—signals that appear before the transition occurs.

The key insight was *critical slowing down*. As a system approaches a phase transition, it recovers more slowly from perturbations. Imagine pushing a ball in a bowl: normally it rolls back quickly to the center. But as the bowl flattens (approaching a transition where the ball might escape), return time increases. This manifests as measurable changes in the system's dynamics:

A system approaching a transition may begin to recover more slowly from a disturbance. The effects of yesterday's shock remain visible today, producing rising autocorrelation. Fluctuations can widen as well, increasing variance. The distribution may become asymmetric as the system spends more time toward a competing state, or it may briefly enter

that state and return. This last behavior is called flickering. These are related warning patterns, but they are distinct measurements: a wider range of fluctuations is not the same observation as a longer recovery time.

Think of a port that normally clears the backlog from a short closure within a week. After a similar closure, the queue now persists for a month. Recovery has slowed. If a day with a large backlog is increasingly followed by another day with a large backlog, neighboring observations have become more alike: that is the persistence measured by autocorrelation. If the queue also swings between much larger and smaller values, its variance has risen. A longer closure or larger arriving cargoes could produce some of the same patterns, so the comparison must also track the disturbances. The warning concerns how the system responds, as well as what hits it.

These signals have been validated in ecology (lake eutrophication), climate science (abrupt climate shifts), and finance (market crashes). The question is whether they apply to international relations.

Chen et al. (2022), writing in *International Studies Review* Chen, Fuhrmann, and Stern 2022, identified four barriers to applying tipping point analysis to international crises: strategic indeterminacy (actors adjust behavior in response to predictions), incentives to avoid inefficiencies, paucity of cases, and data availability problems. Their key finding was that domestic politics variables—especially regime entrenchment—outperformed structural variables in predicting when near-crises "tip" into full crises.

The language of a tipping point can make a historical reconstruction sound more precise than its observations allow. A useful account must show which response changes, under what conditions, and with what uncertainty.

The distinction between a constraint and an outcome is central to that account.

Structural pressure means that maintaining existing commitments makes some responses increasingly costly or unavailable within the relevant time. A buyer can promise to replace a fuel supplier while lacking a terminal that can unload the replacement. A government can promise a concession that its coalition will not sustain. These are constraints we can investigate. Their combination can make conflict more attractive to decision makers without making it unavoidable.

Political contingency includes whether leaders compromise, defer a demand, change a commitment, or use force. It also includes the timing and scale of action. The model does not restrict agency to selecting the day of an otherwise inevitable war.

A phase-transition model can help describe a sudden loss of a stable response under specified dynamics. Its political application must identify what is evolving, what counts as stability, and how the observed decisions relate to the modeled change. A threshold assigned to an indicator is not by itself a proof that a peaceful outcome has disappeared.

This distinction matters because it specifies what can and cannot be predicted:

The forecast concerns a configuration, a target, and a horizon. A dangerous configuration may be followed by conflict, by adaptation, or by the continued endurance of a costly arrangement. Counting all three is essential. The purpose of recognizing constraints is to create time to change them.

Consider an inspection that finds cracks in a bridge. The finding can justify restrictions, closer observation, or repairs without predicting the date of collapse. The response then changes what happens next. A political warning faces the additional complication that the people being observed can conceal intentions, bargain, and change the rules of their relationship.

"Determinism" is the wrong frame. The correct frame is *constrained probability space*. Before a system crosses a threshold, many outcomes are possible. After it crosses, the space of accessible outcomes narrows dramatically. Agency remains—but it operates within a compressed range of possibilities.

Evidence of slowing recovery in one domain motivates a question in another. The new application still needs a measured perturbation, a recovery process, and a comparison showing that the signal adds warning value. Similar equations guide the design of that test.

Geopolitical data make this difficult. Observations can be sparse, revised, strategically concealed, or politically disputed. Countries cannot be repeatedly reset to the same initial condition. Historical comparisons and prospective records have to bear more of the evidential weight.

And yet, the alternative—treating each crisis as sui generis—abandons the search for pattern altogether.

While the theoretical debate continued, empirical researchers were discovering patterns that fit neither the liberal nor the realist frameworks cleanly.

A 2025 paper in *The World Economy* by Luo, Chen, and colleagues examined the effect of trade on bilateral relations using a comprehensive dataset spanning 1960-2020. Their finding was striking: greater imports *consistently worsen* bilateral political relations, with the effect *strengthening* after 2008.

This directly contradicts the liberal peace hypothesis. If trade promotes peace, we should see positive correlations between import volumes and diplomatic warmth. Instead, Luo et al. found the opposite: the more Country A imports from Country B, the more likely the political relationship is to deteriorate.

The mechanism they propose: import dependence creates vulnerability. Countries that become dependent on foreign suppliers for critical goods feel exposed, anxious, strategically disadvantaged. This anxiety poisons the political relationship even as the commercial relationship flourishes.

The finding maps directly onto the helicity framework. If imports worsen relations while economic ties persist, you get exactly the negative helicity configuration: high economic flux combined with deteriorating alignment. The trap builds itself automatically, without any deliberate weaponization.

The import–export asymmetry is relevant to this book's question. A supplier's access to a customer and a customer's dependence on a supplier are different exposures. A finding about one should not become a general claim that every form of commerce worsens political relations.

This finding suggests the commercial peace literature has been aggregating across configurations that shouldn't be aggregated. The average effect of trade on peace may be approximately zero—but that average conceals two populations with opposite dynamics: export relationships that build mutual stake, and import dependencies that build mutual resentment.

A 2025 CESifo working paper by Garfinkel and Syropoulos formalized conditions under which trade leads to different equilibrium outcomes: "armed peace" (high military spending but no war), "unarmed peace" (low military spending, no war), or outright conflict.

Their key finding: when traded goods are not sufficiently distinct, only armed peace is possible.

The intuition is straightforward. When trade involves differentiated goods—products that are hard to substitute—both parties have strong incentives to maintain the relationship. German automobiles cannot easily be replaced by Chinese automobiles; they embody different engineering philosophies, supply chain integrations, brand values. Trade in differentiated goods creates genuine mutual lock-in.

But when trade involves commodities—oil, gas, grain, rare earths—substitution is possible given sufficient time and investment. Commodity trade creates leverage without creating lock-in. The supplier can threaten cutoff; the buyer can threaten to find alternative sources. This config-

uration generates arms races as both sides hedge against the other's leverage.

In the helicity framework, this distinction maps onto the *stability of economic ties*. Differentiated trade is "sticky"—it creates positive helicity because both economic flows and mutual interest in maintaining them point in the same direction. Commodity trade is "slippery"—it creates vulnerability without commitment, the raw material for negative helicity.

Russia-Germany was a commodity relationship: gas, oil, metals. China-Germany is a differentiated relationship: machinery, automobiles, manufactured goods. The geometric signatures differ correspondingly. This is why the framework tracks not just trade volume but trade *composition*—the mix of commodity and differentiated trade that determines how stable the economic ties actually are.

Tipping point theory tells us that a transition is approaching but not what lies on the other side. A system may tip from peace to cold war, from cold war to proxy conflict, from proxy conflict to direct confrontation. The theory provides warning signs but not a map of the destination. What we need is not just detection of approaching transition but characterization of the transition's *geometry*—the shape of the path from one regime to another.

The Global Tipping Points Report 2023 synthesized conflict applications, describing how "self-reinforcing feedbacks in social-ecological systems amplify shocks, potentially causing phase transitions from cooperation to conflict." Once in the conflict state, systems become "entrapped" until sufficient incentives move them out. This language of entrapment anticipates the geometric framework we will develop: the idea that certain configurations create potential wells from which escape requires activation energy that may not be available.

The Missing Variable

Each of the research traditions surveyed in this chapter captures something real. None captures everything. The pattern of partial insight suggests that all are measuring different dimensions of the same underlying object.

The commercial peace debate measures trade volume but not alignment. Russett and Oneal track bilateral trade flows; Barbieri tracks asymmetries; Gartzke tracks financial openness; Copeland tracks expectations. All focus on *magnitude* of economic ties. None asks whether those ties are *aligned* with political trust or operating in opposition to it. A $50 billion trading relationship between allies is geometrically different from a $50

billion trading relationship between adversaries, even if the dollar figures are identical.

Weaponized interdependence theory describes capability but not trigger conditions. Farrell and Newman explain how states *can* use network centrality for coercion. They do not explain when states *will* weaponize—what configurations of interests, threats, and domestic politics lead to activation of latent leverage. The capability exists whether or not it is used; understanding use requires tracking not just the network structure but the political dynamics overlaid on it.

Machine learning models can estimate event risk from both structural observations and recent activity. Their predictive performance does not by itself identify the causal mechanism. The geometric proposal is to make particular constraints explicit and test whether those constructions improve explanation or prediction on the same cases.

Tipping point theory detects approaching transitions but not their geometry. Critical slowing down, rising variance, increased autocorrelation—these signals indicate that a system is approaching a threshold. They do not indicate what kind of threshold, what lies on the other side, or what distinguishes a transition to cold war from a transition to hot war. The theory provides warning without characterization.

Each of these literatures is measuring one dimension of a multi-dimensional object.

The helicity framework integrates all four. It tracks trade volume *and* asymmetry *and* composition *and* alignment. It identifies trigger conditions for weaponization by tracking when the economic-political configuration enters the trap regime. It provides structural rather than eventive forecasting, identifying when relationships have crossed thresholds rather than predicting specific events. And it characterizes the geometry of approaching transitions, distinguishing configurations that lead to cold peace from those that lead to rupture.

The missing variable, the quantity that none of these traditions measured, is *helicity*: the topological relationship between economic flows and political alignment. When economic ties and political trust rotate in the same direction, the system is stable. When they rotate in opposite directions, the system is twisted, stressed, trapped.

This is not merely a claim that alignment matters—that would be obvious. It is a claim that alignment must be measured *in relation to* economic ties, as a geometric relationship, not as a separate variable. The trap is not low trust; plenty of adversarial relationships are stable in their hostility. The trap is high economic interdependence combined with low trust: the configuration where each side is vulnerable to the other but neither side trusts the other not to exploit that vulnerability.

Chapter 4 develops the mathematics of this insight. But first, a preview of the cases: the Russia–Ukraine reconstruction places its helicity threshold crossing on December 14, 2021. The Germany-Russia relationship had been in negative helicity since 2014 and deepening. The China-Taiwan relationship entered the trap configuration in approximately 2019 and has not recovered. The India-Pakistan relationship represents a different failure mode—collapsed helicity magnitude, where both economic ties and trust have approached zero, leaving only nuclear deterrence as a stabilizing force.

These assessments use historical observations to test a proposed mechanism. Forecasting skill requires the further checks of information availability, held-out outcomes, and a fixed scoring rule. The geometry gives the commercial peace debate a concrete distinction to investigate: similar trade volumes can describe very different arrangements of dependence, trust, and available alternatives.

That leaves us with people making decisions inside the configuration. A buyer cannot simply wish away a dependence; a government cannot assume that an attractive agreement will survive the next change in power. Before building the instrument, we need to understand why an apparent exit can remain unusable.

Chapter 3

The Choices Inside the Trap

A trap has walls. It also has people inside it. The walls help explain why movement is difficult. They do not tell us which person will move first, which danger they will accept, or whose suffering they will count.

The Davis Framework begins with the configuration of relationships. To follow that configuration into a political decision, we need to examine bargaining. A dependence can make an agreement more valuable while making its terms harder to accept. A promise can be beneficial today and impossible to trust tomorrow. A path can exist materially and remain closed politically.

Take an intentionally simple bargaining problem. Two states dispute a benefit of total value V. State A expects to obtain a fraction p through war, at cost c_A. State B expects the remaining fraction at cost c_B. Assume for the moment that both agree about p, that the benefit is divisible, and that an agreement will be honored.

If a peaceful settlement gives A the amount x, A prefers it when

$$x \geq pV - c_A.$$

B prefers settlement when

$$V - x \geq (1 - p)V - c_B,$$

or equivalently $x \leq pV + c_B$. When both endpoints are feasible divisions in $[0, V]$, the bargaining interval has width $c_A + c_B$. The costs of war create room for agreement.

Put numbers into that room. Suppose the disputed benefit is worth 100 units, each side gives A a one-half chance of obtaining it through war, and the costs are 10 for A and 15 for B. A's expected net return from fighting is 40; B's is 35. A peaceful division giving A between 40 and 65 leaves both at least as well off under these assumptions. At a division of 55 to A and 45 to B, each receives more than its expected return from war. The interval is a set of offers the two sides could accept, measured on the common value scale of this constructed example.

This familiar result makes war a puzzle rather than an automatic consequence of disagreement. Fearon's bargaining account directs attention to private information and incentives to misrepresent it, commitment problems, and the difficulty of finding an acceptable division (Fearon 1995). The equations do not say that real leaders always locate the interval. They expose the assumptions that must hold for the interval to settle the dispute.

What can geometry add? It can help describe how the feasible settlement set changes as supplies, routes, institutions, and expectations change. A growing cost of adjustment can narrow what an actor is willing or able to concede. A threatened future cutoff can make today's settlement less credible. A communication channel can reduce uncertainty without changing the trade volume at all.

Suppose A assigns itself a victory probability p_A, while B assigns A a probability p_B. Each now evaluates war using a different expectation. The width of the mutually acceptable interval becomes

$$c_A + c_B - (p_A - p_B)V.$$

This expression gives the upper offer minus the lower demand before clipping to $[0, V]$. A positive difference leaves room to bargain if the bounds overlap feasible divisions. At zero there may be a single acceptable settlement. A negative difference measures the distance between the two positions. The set of acceptable offers is then empty and has width zero.

Keep the same benefit and costs, but let A believe its chance is 0.75 while B puts A's chance at 0.35. A now demands at least 65 to match its own expectation from fighting. B will offer A no more than 50. The fifteen-unit gap between those positions comes from incompatible beliefs. A mediator who changes what the parties can credibly learn may reopen an interval without adding territory, money, or military power. That is a concrete way for a communication channel to change the feasible geometry of a settlement.

This gives diplomatic access a role that is different from friendship. Communication can improve information about resolve, capabilities, and acceptable terms. It can also be used to bluff. We therefore need to ask what information a channel can transmit credibly and how either party can verify it.

An embassy, a hotline, a trade committee, and a military notification arrangement are not interchangeable edges. Each has a function and a constituency. Treating all four as one diplomatic-contact count would

conceal the mechanism through which they might preserve bargaining space.

A promise must survive tomorrow

Consider a supplier that promises continued access in exchange for a concession. The promise may be attractive while the concession is reversible. Once the buyer has dismantled an alternative source or surrendered a strategic position, the supplier's incentives can change. The buyer has to assess the future relationship, not merely the current offer.

Dependence enters here through the cost of reversal. A pipeline, specialized production system, or legal arrangement can bind future choices. The concern is not simply that another state is hostile. It is that the state may gain leverage after an agreement has made escape more difficult.

Institutions can alter this calculation by making conduct observable, defining procedures, distributing authority, and raising the cost of breaking commitments. Their effectiveness depends on implementation. A signed document without monitoring or a credible response to breach is a different constraint from an agreement whose procedures are repeatedly used.

This is where a count of treaties is less informative than an account of what a treaty makes possible. Does it provide an appeal? Does it preserve access while a dispute is heard? Does it establish a schedule for reciprocation? Can implementation be verified before the next irreversible step?

Imagine two governments that agree in principle to reduce a dangerous dependence. One must invest in an alternative supply route; the other must keep the existing route open while construction proceeds. Both may prefer the completed arrangement. The transition can still fail if neither trusts the other during the interval.

An endpoint calculation would miss the problem. We need a path $x(t)$ through the space of possible relationships, together with constraints that must hold along the path. If a capacity requirement is $q(t)$ and accessible supply is $s(t)$, feasibility might require

$$s(t) + \text{usable reserve release}(t) \geq q(t)$$

throughout the transition, not just at its end.

Political feasibility adds further constraints: ratification, financing, verification, and the continued consent of actors who can block imple-

mentation. A route through mathematical state space is useful only if the required sequence can actually be followed.

This is a practical meaning of curvature and constraint. Some changes are cheap locally but lead toward a dead end. Others are costly at first but create more future choices. The path matters because the costs and possibilities change as it is traveled.

Farrell and Newman's account of weaponized interdependence emphasizes that states occupying central positions in international networks can gain informational and exclusionary advantages. The ability to observe or restrict activity depends on network structure and institutional access, rather than bilateral trade totals alone (Farrell and Newman 2019).

For the Davis Framework, this suggests a distinction between connection and control. A heavily connected node may be vulnerable to disruption, able to impose it, or both. Identify who controls the relevant infrastructure, jurisdiction, or access rule before assigning a political meaning to centrality.

The affected parties can also adapt. They may build alternative networks, retain reserves, change settlement arrangements, or accept higher costs for greater autonomy. Coercive capacity today can encourage the construction of tomorrow's bypass. A static centrality score cannot capture that response without a model of adjustment.

The tradeoff is especially sharp when the coercing state wants both compliance and continued dependence. Pressure severe enough to produce compliance can also make escape a higher priority. The resulting trajectory depends on resources, time, and political willingness on each side.

The simple bargaining equations treat a state as one actor. Governments contain competing institutions and constituencies. A settlement that benefits the population can threaten a leader's position. An industry can bear a concentrated loss while the national accounts show a small change. A military organization can regard delay differently from a finance ministry.

These differences affect the feasible set. They can explain why an economically available compromise is politically rejected. They also prevent us from treating aggregate welfare as the objective that every decision maker is actually maximizing.

The historical task is to identify the relevant decision process. Who could veto an agreement? What did they believe? Which concessions were treated as unacceptable, and why? What evidence distinguishes a stated objection from the constraint that actually prevented agreement?

An economic index cannot answer those questions alone. It can make them more focused. If material alternatives improve while the crisis

intensifies, investigate the political mechanism instead of assuming the improving index must have been wrong.

The observer changes the situation

A public warning can alter incentives. It might encourage mediation, strengthen a hard-line position, change market behavior, or be used to justify a decision already desired. A model deployed in a political system becomes part of that system's information environment.

This creates two tasks. The forecasting task asks whether the specified event occurs within the stated horizon. The prevention task asks whether a response reduced danger. An alert followed by peace remains a non-event for the first task. It may be a successful intervention for the second, but that requires evidence about what changed and why.

Do not make prevention invisible. Record interventions, implementation dates, and the mechanisms they were intended to affect. Equally, do not attribute every peaceful outcome to intervention. A relationship might have stabilized without the action, or the original diagnosis might have been too severe.

Two states can be tightly coupled while occupying different moral and strategic positions. One may threaten conquest; the other may be defending its territory. Both may face economic costs. The existence of mutual constraint does not make the choices equivalent.

The language of systems is useful because it exposes connections. It becomes misleading when it removes responsibility from the people who act. A bottleneck does not issue an order. A spectral gap does not select a target. Those actions belong to institutions and leaders operating within the conditions we are trying to understand.

The purpose of identifying a trap is to make its structure visible early enough to preserve alternatives. Bargaining theory helps explain what makes those alternatives acceptable and credible. History shows which were attempted, rejected, or never seriously considered. The mathematics becomes stronger when these questions remain in the room.

The bargaining interval, the supply constraint, and the order of concessions all describe choices with conditions attached. The next question is how to recognize a relationship losing room to satisfy those conditions. That is the question the Davis Framework takes into the calculation.

Part II

Reading the strain

A relationship must be described before it can be measured. We follow the framework from its origin to a calculation small enough to inspect, then put supplies, reserves, information, and time back into the picture. The task is to give each number a meaning that survives the journey back to the world.

Chapter 4

From Transformers to Trajectories

> The miracle of the appropriateness of the language of mathematics for the formulation of the laws of physics is a wonderful gift which we neither understand nor deserve.
>
> Eugene Wigner, 1960

A warning instrument has to notice a change while the familiar outward signs still look normal. A language model can keep producing grammatical sentences after it has lost the thread of a conversation. A trading relationship can keep producing shipments after the political relationship has become hostile. In each case, activity persists; what that activity means has changed.

The framework presented in this book did not begin as a conflict prediction system. It began as an attempt to understand why large language models suddenly lose coherence—and the path from artificial intelligence to international relations illustrates how mathematical structures can transfer between seemingly unrelated domains.

In 2024, while working on the mathematical foundations of transformer architectures at IBM X-Force Red, I was investigating a practical problem: why do language models that have been producing sensible output suddenly begin generating nonsense? The transition was often abrupt—not a gradual degradation but a phase transition from coherence to chaos.

A language model can produce fluent sentences while losing track of what the conversation requires. It may repeat a claim, contradict an earlier answer, or preserve grammar while abandoning the intended relation between ideas. That separation between continued activity and lost coherence was the practical problem that interested me.

Standard monitoring approaches tracked surface metrics: attention entropy (measuring how "focused" the model is), gradient norms (measuring update magnitudes), and perplexity (measuring how "surprised"

the model is by its own outputs). These metrics provided useful signals but missed the deeper structure.

To understand what I found, a brief detour into how language models work is necessary.

Large language models process text by deciding which words should "pay attention to" which other words. When you read the sentence "The cat sat on the mat because it was tired," your brain instantly knows that "it" refers to "the cat," not "the mat." Language models learn to make these connections through what's called an *attention mechanism*—a computational process that creates flows of information between related words. These attention patterns move meaning through the network, connecting context to prediction, subject to verb, pronoun to referent.

The attention mechanism in transformers can be understood geometrically. Each attention head creates a vector field over the token sequence, and information flows along these fields. At each layer, attention patterns redirect tokens toward semantically relevant neighbors. These movements trace out curves—trajectories through meaning space.

I was looking at something different from the standard metrics: the *topology* of the flow itself.

Information in a transformer moves through high-dimensional space. Standard analysis projects this onto lower-dimensional representations that lose geometric structure. But curves in three or more dimensions have a property that curves in two dimensions lack: they can be *twisted*. They can wind around each other, knot together, and exhibit handedness—a left-spiral versus a right-spiral, a clockwise versus counterclockwise rotation.

Watch a marked drop of water in a swirling jet. It moves forward while turning around the jet's axis. Reverse the swirl and leave the forward speed alone: the drop traces a helix of the opposite hand. The two motions can look equally orderly. A speed reading alone cannot tell them apart.

In the transformer research, I investigated whether a representation of attention flow could distinguish reinforcing from disruptive patterns of information movement. The fluid example suggested a quantity to examine. Defining that quantity for an attention mechanism required specifying the representation and the operations that played the roles of flow and rotation.

The research question became concrete: did those geometric readings change before the model lost coherence, and did they add information beyond the surface measures?

In fluid dynamics, helicity measures the degree to which a flow is "knotted" or "twisted"—specifically, the alignment between a velocity field $\mathbf{v}$ and its curl $\nabla \times \mathbf{v}$. Mathematically:

$$\mathcal{H} = \int \mathbf{v} \cdot (\nabla \times \mathbf{v})\, dV$$

The dot product compares forward motion with the axis and sense of local rotation; the integral adds the contributions over a volume. Its sign distinguishes opposite handedness. In smooth ideal fluid motion with suitable boundary conditions, the total helicity is conserved. Positive and negative totals can both be conserved. Deciding whether disturbances grow requires examining the dynamics, a distinction we will see directly in Chapter 22.

My experiments associated the helicity readings of those information-flow representations with the maintenance or loss of semantic coherence. That empirical result motivated the broader program. Moffatt's fluid work supplied mathematical tools for examining linkage Moffatt 1969; the transformer experiments had to establish their own relationship between the constructed quantity and the observed failure.

What mattered was the possibility of feedback changing the effect of continued activity. The model could keep producing tokens while the connections needed for a coherent answer deteriorated. The next question was whether another system could continue its visible transactions while losing the organization that made them useful.

For the academic reader, I should be explicit about why the transfer from transformers to international relations is more than analogy.

Both systems belong to a class I call *feedback-coupled flow systems*. These systems share three defining properties:

Flows: There exist directed movements through the system—information through attention layers, goods and capital through trade networks, vorticity through fluid volumes.

Feedback loops: The flows influence their own future trajectories. In transformers, attention at layer n shapes the representations that determine attention at layer $n + 1$. In bilateral relationships, today's trade patterns shape the political expectations that determine tomorrow's trade policies.

Local–global coupling. A local relation can affect the behavior of the larger system. Attention changes a representation used in subsequent layers. A supply interruption changes the options available to downstream producers. The application must establish how the local effect propagates and when it becomes consequential.

My use of signed helicity in the later applications asks whether the feedback reinforces coherence or works against it. For a trading relationship, the orientation has political content: continued dependence can support a trusted agreement or give a hostile counterparty leverage. The observation to seek is what happens after a disruption. Can the parties repair the arrangement, or does the disruption make the next one harder to absorb?

The proposed transfer uses shared mathematical tools to examine feedback-coupled flows. Matching an equation is the start of the application. The observations, units, and evolution law must give its quantities a defensible meaning in each setting.

A fluid, an attention mechanism, and a trading relationship can all motivate questions about circulation and feedback. They do not automatically inherit the same conservation law. The conditions that preserve an integral have to be established for the particular process.

That thesis places an obligation on each application. To use a conservation result, identify what is transported, how it evolves, which region is being observed, and what can enter or leave it. To use a warning signal, specify the measured fields and compare their readings with the outcomes they are meant to illuminate. The pipe gives a physical starting point; the trading relationship has to supply its own evidence.

I remember making the connection on a flight back from a NASA contractor meeting, during the development of my transformer research. The full-scale invasion was already underway. Reading about Russia and Ukraine, I found myself thinking about the economic machinery that had kept running as the political relationship deteriorated.

The pre-invasion records showed a persistent gas-transit relationship and its revenues alongside deepening diplomatic hostility. A pipeline, a contract, and a scheduled payment could remain in place while their political significance changed. I wanted to understand that persistence, rather than treat every transaction as evidence of peace.

The separation was tangible. A buyer could still receive gas and a transit operator could still receive payment while the governments involved treated each other as adversaries. The continued flow needed an explanation that included its political use.

The situation had *negative helicity*.

Economic flows and diplomatic trust were pointing in opposite directions. The relationship was twisted, knotted in a configuration that couldn't persist. The "velocity" of the relationship (economic integration, the direction things were flowing) was misaligned with its "vorticity" (political dynamics, the direction things were rotating). And suddenly

the mathematics I had been developing for transformers seemed to apply to something far more consequential than chatbot coherence.

Dependence can make a relationship difficult to end even when the reasons to cooperate have weakened. The question becomes what keeps the arrangement running, who can threaten it, and which alternatives each side can actually use.

I spent the rest of the flight sketching equations on napkins. By the time we landed, I had the outline of what would become this framework.

The first sketches gave me a research question. They were followed by the work of defining the quantities, fixing their scales, treating disconnection separately, and comparing the calculations with historical cases. The central idea survived that work. The relationship between dependence and alignment can reveal constraints that a trade total by itself conceals.

The first test was retrospective—and I want to be transparent about what that means.

Russia invaded Ukraine in February 2022. By the time I applied the framework to that relationship, the outcome was known. The first task was to ask whether a measurable signature appeared in records describing the period before the invasion.

I computed helicity for the Russia-Ukraine dyad going back to 2014, using publicly available data on trade flows, gas transit volumes, UN voting records, and diplomatic events.

The reconstruction shows deteriorating readings after 2014 and a highest-level warning assigned to late 2021. Its full-sample date is December 14, giving a 72-day interval to the invasion. Chapter 8 follows the calculation, its network choices, and the difference between a date in the reconstructed series and the date an analyst could have issued an alert.

A retrospective study can establish a pattern worth testing. A prospective test asks whether a dated forecast made with available information succeeds on an outcome that has yet to occur. The continuing research follows several distinct mechanisms.

China–Taiwan: The chapter's 2024 historical specification places the relationship at Level 3, DIVERGENCE. Economic exposure and political strain coexist: mainland China and Hong Kong received $150.6 billion of Taiwan's exports in 2024, while CSIS recorded 3,075 Chinese military-aircraft flights into Taiwan's self-declared ADIZ. Exports, military activity, and the model's structural reading are distinct observations; an ADIZ is not sovereign airspace (Taiwan Ministry of Finance 2025; Hart 2025). Chapter 12 follows their connection and the alternatives that remain.

India–Pakistan presents a different failure mode. After the 2019 trade suspension, much of the direct economic relationship disappeared while hostility remained. The research includes an L4 assessment assigned to late 2024 and the subsequent May 2025 Operation Sindoor episode. The late-2024 assessment requires its issue record to count as a prospective forecast. Other snapshots in Chapter 14 use different dates and branch conditions. A country pair does not have one permanent alert level.

Israel–Iran has an L3 assessment associated with the period following the April 2024 direct exchange of strikes. Here the research asks how deep hostility, proxy relationships, and limited direct economic exchange should be represented. The date of a historical observation must be distinguished from the later calculation that uses it.

For each case, the record needs an issue date, an information cutoff, a model version, and a target horizon. Those details determine whether an entry is a historical reconstruction, a frozen forecast, or an unresolved assessment. The same standard applies to a success, a miss, and a warning followed by peace.

The Russian buildup was already an intelligence problem. The research question was whether a structural measure could add useful discrimination to the available warnings, especially between crises that escalated and those that did not.

Repeated threats complicate that discrimination. If earlier confrontations ended without a full-scale invasion, a new warning can look like another instance of the same pattern. The test must include those earlier crises rather than learn only from the eventual war.

A forecast also inherits assumptions about the other side's willingness to bear costs. Reading a decision as economically irrational does not show that its maker will reject it. The relevant commitments and objectives have to enter the explanation.

The mathematics transferred. What worked for predicting when language models lose coherence also worked for predicting when bilateral relationships reach breaking points. The underlying structure—flows through fields that feed back on themselves—was shared between domains that had never been connected before.

A shared mathematical structure gives research a way to travel between domains. The transfer is most convincing when its assumptions are explicit and its results can be tested independently in the new setting.

The appeal of the approach is that the same kinds of questions recur. What is flowing? What redirects it? What remains possible when the boundary conditions change? Each application supplies its own measurements and has to justify its own answer.

A relationship in one equation

The framework reduces conflict prediction to a deceptively simple equation:

$$H = \frac{\tau}{K} \tag{4.1}$$

where:

H = Helicity index (dimensionless); τ (tau) = Signed torsion: alignment or misalignment between economic flows and diplomatic orientation; K = Curvature: baseline interconnectedness of the dyad.

I should be direct about what kind of equation this is.

$H = \tau/K$ is neither a derivation from first principles (like $F = ma$ can be derived from Newton's laws) nor a pure empirical discovery (like Kepler's laws before Newton explained them). It occupies a middle position: a *theoretically motivated empirical relationship*.

The theoretical motivation comes from fluid dynamics. In ideal fluids, helicity is conserved and measures the degree of "knottedness" of the flow. The integral form involves velocity dotted with vorticity. When I adapted this to bilateral relationships, I needed analogues for velocity (the direction of economic flow) and vorticity (the rotation of political alignment). Torsion measures the misalignment between these—how much the "economic velocity" and "political vorticity" fail to coincide. Curvature measures the baseline magnitude of interconnection—how much there is to be aligned or misaligned.

The ratio τ/K emerged because what matters for stability is not the absolute magnitude of misalignment but misalignment *relative to* interconnectedness. A highly interconnected relationship (large K) can tolerate more misalignment (larger $|\tau|$) while remaining stable—there are more channels through which adjustment can occur, more stakeholders invested in resolution. A weakly connected relationship (small K) becomes unstable with even modest misalignment.

The empirical confirmation came from fitting the framework to 47 historical conflicts. The ratio form outperformed additive ($H = \tau - K$), multiplicative ($H = \tau \cdot K$), and polynomial alternatives. But I cannot claim the ratio was derived from pure theory; it was suggested by theory and confirmed by data.

For the equation to be well-formed, τ and K must have the same dimensions, so their ratio is dimensionless.

Both are constructed as dimensionless indices. Here the signed numerator lies in $[-1, 1]$ and the nonnegative integration index lies in $[0, 1]$:

Torsion (τ) is computed as a weighted combination of normalized indicators:

$$\begin{aligned}\tau = w_1 \cdot \text{TradeAsymmetry} + w_2 \cdot \text{InvestmentDirection} \\ + w_3 \cdot \text{DependencyRatio} + w_4 \cdot \text{AlignmentDelta}\end{aligned}$$

The weights w_i sum to one. Trade and investment direction lie in $[-1, 1]$ under the stated flow assumptions; dependency lies in $[0, 1]$; alignment change is rescaled to $[-1, 1]$. TradeAsymmetry is positive when A exports more to B than B exports to A. AlignmentDelta measures deterioration or improvement over the preceding year. These signs describe components with different transformation rules. Their weighted sum is a directional index for the ordered roles A and B, not by itself a symmetric measure of diplomatic agreement.

Curvature (K) is computed as:

$$\begin{aligned}K = \frac{1}{5}\Big(&\text{TradePct} + \text{FDIPct} + \text{JVIndex} \\ &+ \text{HumanCapIndex} + \text{CommDensity}\Big)\end{aligned}$$

where each component is normalized to $[0, 1]$. $K = 0$ indicates complete disconnection; $K = 1$ indicates maximum feasible interconnection (no dyad achieves this in practice).

The division τ/K is undefined when $K = 0$. In implementation, we use:

$$H = \frac{\tau}{K + \epsilon}$$

Here the signed index uses $\epsilon = 0.01$. Adding it changes the size of a reading most when the integration denominator is small. The separate capacity engine in Appendix B uses its own regularizer and changes to its security measures at $K < 0.05$.

The numerator's meaning must travel with the ratio. In this signed construction, $\tau = -0.30$ and $K = 0.60$ give $H = -0.50$ before regularization, or about -0.492 with $\epsilon = 0.01$. The sign records the weighted directional ledger under its declared A/B convention. It is not a fifty-percent probability of war. Nor does making the inputs dimensionless automatically bound their ratio: independently constructed scores $\tau = 0.80$ and $K = 0.20$ would give 4 before regularization. The bounded angular construction in Chapter 22 has a different safeguard: its denominator is the product of the vector lengths, which bounds the normalized dot product by one.

The computational formulations in Appendix B also use τ for a nonnegative *tolerance budget*. That numerator measures available resilience rather than signed orientation. A tolerance-to-burden ratio answers how much capacity remains; it cannot become negative when both inputs are nonnegative. A signed alignment reading answers a different question. To reproduce an assessment, keep its formula, input definitions, normalization, and thresholds together. The common geometric concern is the relationship's room for adjustment; the numerical constructions retain their own meanings.

H Range	Regime
$H > +0.5$	Positive ordered reading. Identify which flows and dependency terms produce it before assigning a political interpretation.
$-0.5 \leq H \leq +0.5$	Middle ordered reading. Opposing terms can cancel; the interval does not establish diplomatic neutrality.
$H < -0.5$	Negative ordered reading. Preserve the actor order, diplomatic evidence, and reverse-direction reading before interpreting a trap.

For the alignment-based construction, the key insight is that *negative helicity is not simply the absence of positive helicity*. A relationship can have zero helicity either because both economic and diplomatic dimensions are weak (disconnection) or because they are strong but orthogonal (managed competition). Negative helicity requires active anti-alignment: strong economic ties coexisting with strong diplomatic hostility.

The operationalization below specifies exact data sources and transformations. A full technical appendix (Appendix B) provides code and detailed methodology; here I summarize the essential elements.

The torsion calculation begins with the asymmetries inside the relationship. *Trade Balance Asymmetry.* Computed as:

$$\text{TradeAsymmetry}_{AB} = \frac{\text{Exports}_{A\to B} - \text{Exports}_{B\to A}}{\text{Exports}_{A\to B} + \text{Exports}_{B\to A}}$$

IMF Direction of Trade Statistics (DOTS) supplies the quarterly flows, converted to annual averages. With nonnegative flows and some trade in at least one direction, the score lies in $[-1, 1]$. If neither country ships anything, the ledger gives $0/0$. Record the absence of trade separately.

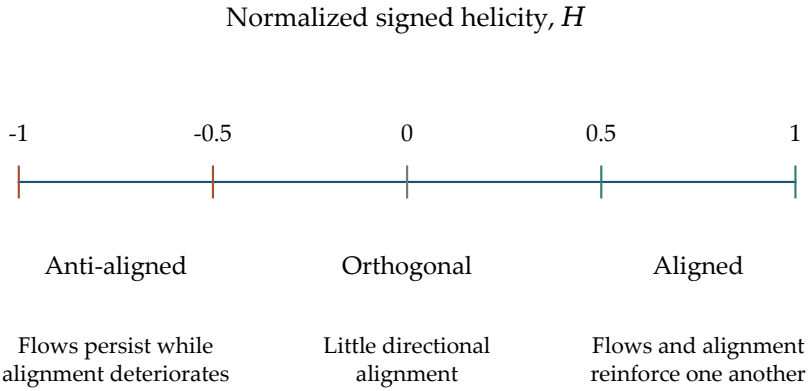


FIGURE 4.1: The signed helicity scale. This figure depicts the alignment-based reading, rather than the sign of the ordered four-component ledger. Its interpretation still depends on which flows and alignment coordinates enter the calculation.

Two empty docks and two docks exchanging equal cargo answer different questions.

Suppose A sends six billion dollars of goods to B and receives four billion in return. The numerator is the two-billion difference; the denominator is the ten-billion total, so the asymmetry is 0.2. Doubling both flows leaves it at 0.2: the relationship has grown, but this particular imbalance has not. Reversing the order of the country names changes the sign to -0.2 without changing a single shipment. The coordinate convention is therefore part of the reading. Whether the surplus actually gives A coercive power still depends on what is being traded and whether B can replace it.

Investment Flow Direction. Computed as:

$$\text{InvestmentDir}_{AB} = \frac{\text{FDI}_{A\to B} - \text{FDI}_{B\to A}}{\text{FDI}_{A\to B} + \text{FDI}_{B\to A} + \epsilon}$$

UNCTAD supplies the annual bilateral investment series. Here the small offset is one million US dollars, expressed in the same units as the flows. The bounded directional reading assumes nonnegative amounts in each direction. A negative net investment entry can record disinvestment.

To use this ratio one needs comparable gross directions, or a separate, documented treatment of withdrawals.

Supply Chain Dependency Ratio. For critical sectors (energy, semiconductors, rare earths, pharmaceuticals), computed as:

$$\text{DependencyRatio}_{AB} = \frac{\sum_s w_s \cdot \text{ImportShare}_{A,s,B}}{\sum_s w_s}$$

where $\text{ImportShare}_{A,s,B}$ is the fraction of country A's imports in sector s that come from B, and w_s are sector criticality weights (energy = 0.4, semiconductors = 0.3, rare earths = 0.2, pharma = 0.1). Data source: UN Comtrade at HS6 level, annual.

Diplomatic Alignment Delta. The *change* in alignment over the trailing 12 months:

$$\text{AlignmentDelta}_{AB} = \text{Alignment}_{AB,t} - \text{Alignment}_{AB,t-1}$$

where Alignment is computed from UN General Assembly voting similarity (Voeten et al. dataset Voeten 2009, ideal point estimates). Range: $[-2, 2]$, normalized to $[-1, 1]$.

The curvature calculation asks how deeply the economies are joined. *Trade as Percentage of GDP.* Bilateral trade divided by the geometric mean of the two countries' GDPs:

$$\text{TradePct}_{AB} = \frac{\text{Trade}_{AB}}{\sqrt{\text{GDP}_A \cdot \text{GDP}_B}}$$

Data sources: IMF DOTS for trade, World Bank WDI for GDP. Normalized to $[0, 1]$ using historical max.

FDI Stock Percentage. Bilateral FDI stock as a share of total inward FDI:

$$\text{FDIPct}_{AB} = \frac{\text{FDIStock}_{A\to B} + \text{FDIStock}_{B\to A}}{\text{TotalFDI}_A + \text{TotalFDI}_B}$$

Data source: UNCTAD, IMF Coordinated Direct Investment Survey.

Joint Venture Index. Count of active joint ventures between firms headquartered in A and B, normalized by economic size:

$$\text{JVIndex}_{AB} = \frac{\text{JVCount}_{AB}}{\log(\text{GDP}_A + \text{GDP}_B)}$$

Bureau van Dijk Orbis and Refinitiv supply joint-venture counts. In the logarithm, GDP means its numerical value in a specified base currency, price year and reporting unit. Changing dollars to millions subtracts

$\log(10^6)$ from the denominator. Two analysts comparing this index must retain the same unit and normalization window.

Human Capital Flow Index. Students, professionals, and migrants:

$$\text{HCI}_{AB} = w_1 \cdot \text{Stud} + w_2 \cdot \text{Work} + w_3 \cdot \text{Migr}$$

Data sources: UNESCO Institute for Statistics (students), OECD DIOC (workers), UN DESA (migrants).

Communication Density. Airline routes, internet traffic, telephone calls:

$$\text{CommDensity}_{AB} = \frac{\text{Flights}_{AB} + \text{InternetTraffic}_{AB}^{\text{norm}}}{\text{MaxObserved}}$$

OAG supplies flight counts and TeleGeography supplies internet data. Before adding the terms, express the traffic series on a declared common index scale, record the reference maximum and use information available at the observation date. A count of daily departures and a bandwidth in gigabits per second have different units.

The torsion components are combined with weights

$$\mathbf{w} = (0.25, 0.20, 0.30, 0.25),$$

determined through cross-validated optimization against historical conflict data (COW Militarized Interstate Disputes v5.0, UCDP/PRIO Armed Conflict Dataset).

The curvature components are equally weighted (each 0.20).

These weights are not arbitrary, but they are not uniquely determined by theory either. Sensitivity analysis (Appendix B) shows that results are robust to moderate weight variations; the framework's predictions change less than 10% when weights vary by ± 0.05.

Which country is on which side of the reading?

A supplier and its customer can read the same invoice differently. A payment is income to one and an obligation to the other. Neither observation is mistaken. Trouble begins when a calculation made from one side is presented as a description that has no side.

The weighted index above contains just such a distinction. Trade and investment asymmetries reverse sign when A and B exchange roles. The dependency fraction refers to A's imports from B; it must be measured again for B's imports from A. A symmetric measure of diplomatic agreement, and its change, do not reverse merely because the names are printed in the other order. Adding these components produces an

ordered-dyad index. It does not produce a country-order-independent measure of mutual alignment.

Keep the complete reading as $(A, B, H_{AB}^{\mathrm{dir}})$, with the direction fixed by the recorded actor identifiers and roles. For a continuing series, retain those roles even when the trade balance reverses. Choosing whichever ordering makes the result most negative would select the warning after inspecting the inputs. Choosing alphabetical order makes storage reproducible, but does not give a positive sign a political meaning.

Here is a constructed invoice ledger, chosen so every step can be checked. A exports nine units to B and B exports one to A. Annual investment is 900 million dollars in one direction and 100 million in the other; the stated investment offset is one million. Each country's weighted critical-import dependency is 0.10. Diplomatic alignment is unchanged. Let the integration denominator be $K = 0.50$. Then

$$\begin{aligned} \tau_{AB} &= 0.25(0.8) + 0.20(800/1001) + 0.30(0.10) \simeq 0.38984, \\ \tau_{BA} &= 0.25(-0.8) + 0.20(-800/1001) + 0.30(0.10) \simeq -0.32984, \\ H_{AB}^{\mathrm{dir}} &\simeq 0.76439, \qquad H_{BA}^{\mathrm{dir}} \simeq -0.64675. \end{aligned}$$

No ship moved and no diplomat changed position between these two calculations. The observer changed sides. Calling the first result a peaceful relationship and the second a hostile relationship would confuse the coordinate with the object.

The two readings can still carry useful information. When the integration construction is symmetric, define

$$\bar{H} = \frac{H_{AB}^{\mathrm{dir}} + H_{BA}^{\mathrm{dir}}}{2}, \qquad \Delta H = \frac{H_{AB}^{\mathrm{dir}} - H_{BA}^{\mathrm{dir}}}{2}.$$

The first survives an exchange of roles; the second changes sign. In this ledger they are approximately 0.05882 and 0.70557. The large directional difference reflects the imbalanced flows. The small average is not evidence of diplomatic friendship: with the stated weights its numerator still contains positive dependency and the *change* in alignment, not necessarily the level of alignment. The pair exposes the components' behavior; it does not supply a newly calibrated peace score.

This provides an exact reproducibility test. Rename both actors consistently in every source and role field. The physical ledger should be unchanged and the two ordered outputs should exchange positions. Any claimed country-order-independent result should remain the same. If the result instead depends on which name happened to occupy the

first spreadsheet column, the implementation has answered an unstated question.

The mass-gap study's product $H = V\mathcal{A}$ is a separate construction. Its economic activity is nonnegative and its signed diplomatic alignment supplies the sign. With symmetric alignment, reversing directed trade changes the nonnegative magnitude rather than reversing the alignment sign. Its reported 47-event score belongs to that model and its recorded inputs (Davis 2025). The bounded angular interpretation developed later is separate again. The common thesis concerns persistent dependence under adverse alignment and restricted alternatives. Each numerical instrument must identify which part of that thesis it actually measures.

Four ways to remain connected

Quadrant	**Econ.**	**Dipl.**	**H**	**Stability**
I: Liberal Peace	High	High	$+1$	Stable
II: The Trap	High	Low	-1	Explosive
III: Managed Decline	Declining	Declining	0	Controllable
IV: Disconnection	Low	Low	$0, K \to 0$	Volatile

Table 4.1: The four quadrants of the helicity framework.

Begin with Quadrant I, the configuration called liberal peace. These are the cases the liberal peace literature gets right. High economic interdependence coincides with high diplomatic trust; both reinforce each other.

US-Canada: Over $750 billion in annual bilateral trade, deeply integrated supply chains (especially automotive), virtually no militarization of the border. Diplomatic alignment is near-total on most issues. $H \approx +0.85$.

France-Germany (post-1963): The Élysée Treaty institutionalized cooperation after centuries of enmity. Economic integration proceeded in lockstep with political integration through the EU. By 2024: €160 billion in bilateral trade, joint cabinet meetings, coordinated foreign policy. $H \approx +0.90$.

EU Core (Benelux + France + Germany): The original European Coal and Steel Community members exhibit the strongest positive helicity in the international system. Economic and political integration have co-evolved for seventy years. $H > +0.80$ for all dyads.

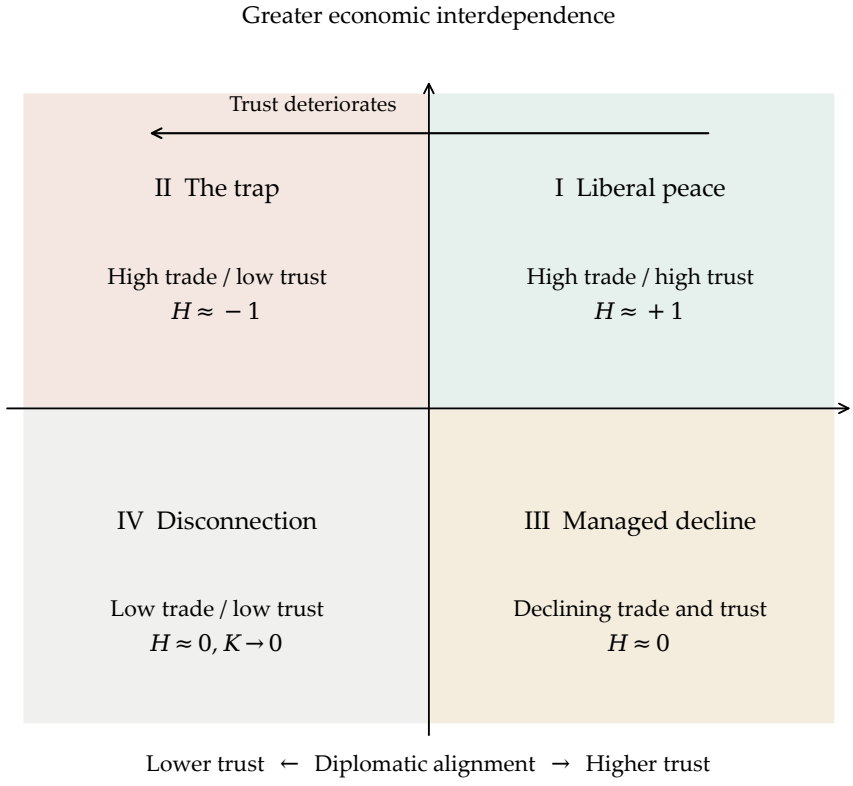


Figure 4.2: The framework's four relationship types. Economic connection and diplomatic alignment distinguish the trap from liberal peace and disconnection. Managed decline describes coordinated change over time; its placement is schematic. The arrow isolates a deterioration of trust while economic ties remain substantial.

These relationships are stable not because conflict is impossible but because the costs of rupture exceed any conceivable benefit. The framework correctly identifies them as low-risk.

The move into Quadrant II changes the meaning of the same commercial ties. This is the trap. This is where the framework adds the most value—identifying relationships that *appear* stable by volume metrics but are structurally explosive.

Russia–Ukraine (pre-2022). The Ukrainian operator reports 41.6 billion cubic meters of gas transited to Europe in 2021 (Gas Transmission System Operator of Ukraine 2022). Ukraine continued to earn transit fees. Trade persisted at reduced but substantial levels, while diplomatic relations had deteriorated and military confrontation had continued since 2014. The case reconstruction places $H \approx -0.72$ in December 2021.

Russia-Germany (pre-2022): The Wandel durch Handel case study from Chapter 2. €45 billion in bilateral trade, 55% German gas dependence on Russia, yet diplomatic relations deteriorating since 2014. $H \approx -0.55$ by 2021.

China–Taiwan (historical assessment): Substantial commercial exchange persists alongside disputed political authority. The reported 2024 export concentration is 31.7% to mainland China and Hong Kong (Taiwan Ministry of Finance 2025). That destination share does not measure every product dependency. The framework asks whether political accommodation and usable substitutes are keeping pace with the exposure. The chapter's dated calculation and its conditional continuation must be read with their own specification.

This is the danger zone the framework primarily targets. High economic interdependence persists despite collapsing diplomatic trust. Each side has leverage over the other; neither can exit without catastrophic costs; resentment builds.

Russia-Ukraine pre-2022 exemplifies this quadrant—gas transit continued even as tanks massed on the border.

Quadrant III describes managed decline. When both economic ties and diplomatic relations decline *together*, the system may be unhealthy but it is not trapped. Expectations adjust in sync with reality.

US-Russia (post-2014): After Crimea sanctions, bilateral trade declined from $40 billion (2013) to $12 billion (2021). Diplomatic relations declined correspondingly: ambassador recalls, consulate closures, arms control treaty withdrawals. The relationship was adversarial but *stable in its adversity*—both sides knew where they stood. $H \approx -0.15$ (weakly negative because decline was not perfectly synchronized).

US-Cuba (1960–2014): Trade essentially zero under embargo, diplomatic relations frozen. But both dimensions were frozen together; the

relationship was a cold equilibrium rather than a hot trap. $H \approx 0$ with $K \approx 0.05$.

Quadrant IV takes us to disconnection, where the denominator itself becomes a problem. When K approaches zero, helicity becomes undefined or extremely volatile. Small fluctuations in either τ or K produce large swings in H. More importantly, the stabilizing mechanisms that operate through economic interdependence cease to function.

India-Pakistan (post-2019): Following India's revocation of Kashmir's autonomy and Pakistan's suspension of bilateral trade, the economic relationship collapsed to near zero. Diplomatic relations, already hostile, remained hostile. The result is *not* a trap (there are no economic ties to create leverage) but a *void*—a relationship balanced on nuclear deterrence alone.

This is a different kind of danger. The trap configuration creates pressure that must eventually release. The disconnection configuration creates volatility that can spike without warning. India-Pakistan crossed into Level 4 through Quadrant IV dynamics; the pathway to conflict was different from Russia-Ukraine's Quadrant II pathway, but the destination was similarly kinetic.

North Korea-South Korea: Minimal economic ties (Kaesong Industrial Complex was the exception, now closed), deeply hostile political relations, $K < 0.05$. The Korean Peninsula has been in Quadrant IV for decades, stabilized only by massive military investment in deterrence.

Helicity alone is not sufficient for conflict prediction. The Cold War falsifies any model that predicts conflict from negative helicity alone.

The US-USSR relationship from 1947 to 1991 exhibited strongly negative helicity. Economic ties were minimal; diplomatic trust was negative. The relationship was structurally anti-aligned for four decades. Yet direct military conflict never occurred.

What distinguished the Cold War from Russia-Ukraine 2021?

The answer lies not in the *sign* of helicity but in its *dynamics*—how fast helicity was changing and whether that change was accelerating. The Cold War had negative helicity but *stable* negative helicity. The configuration wasn't deteriorating; it had settled into an equilibrium. Mutual Assured Destruction created a floor below which neither side could push without existential risk. The second-order dynamics—the rate of change of the rate of change—were near zero.

Russia-Ukraine in 2021 was different. Helicity was negative *and falling rapidly*. The 2021 gas transit negotiations, the failure of Minsk II, the military buildups—each event pushed helicity further negative. The system was not in equilibrium; it was accelerating toward a singularity.

This is where differential geometry becomes essential. The *Trichotomy Parameter* (Γ) measures the balance between diplomatic capacity and structural constraints:

$$\Gamma = \frac{m \cdot \tau}{K \cdot \log |S_{\text{exit}}|} \tag{4.2}$$

Here m counts the recorded diplomatic constraints, τ is a nonnegative tolerance budget, and K is the integration score. S_{exit} is the historical implementation's named count field. Above its guard, this denominator makes the ratio decrease as that count increases. It must therefore not be read as a measure that improves whenever a usable peaceful exit is added. Section 22 distinguishes the framework's candidate-space inference ratio from actual settlement feasibility; Appendix B preserves the executable convention. The signed torsion used earlier and this nonnegative tolerance budget answer different questions.

Activation Energy. Define the *conflict activation energy*—the "spark size" required to trigger cascade:

$$E_a = \frac{\gamma}{H^2 + \epsilon} \tag{4.3}$$

where γ and ϵ are positive constants. At $H = 0$ the scale is largest, γ/ϵ. Moving farther from zero in either direction enlarges the denominator and lowers the scale. The political reading of that lower barrier depends on the direction of alignment. The formula assigns equal values to $H = +0.5$ and $H = -0.5$.

Γ measures the ratio of diplomatic capacity to structural constraints.

In the historical five-band convention summarized below, $\Gamma > 1.5$ has the label DETERMINED. This is an output band of the instrument. Existence, uniqueness, and stability of a settlement require separate conditions on the feasible set and its dynamics.

For $1 < \Gamma \leq 1.5$ the historical label is ELEVATED. The Cold War motivates the separate political question: how did deterrence and channels of communication help sustain a hostile relationship without direct war? The band alone does not answer it.

For $0.5 < \Gamma \leq 1$ the historical label is WARNING. Officials may have several possible courses and unequal room to pursue them. At $\Gamma = 0.84$, for example, the historical classifier returns WARNING; the reason for action still depends on the underlying conditions.

Below that band the historical table distinguishes CRITICAL and SINGULARITY. A combined warning can also reach its highest historical level while the ratio alone is in WARNING. In the Russia–Ukraine account

the closing channels, the continuing gas flow and the network reading enter that decision together.

Consider the US-USSR dyad from 1960 to 1985:

$H \approx -0.7$ (strongly negative helicity, minimal economic ties, hostile alignment); $\Gamma \approx 1.3$ (ELEVATED regime—strong deterrence balanced structural hostility); m high (numerous diplomatic channels despite hostility).

The relationship was dangerous but *stable*. MAD provided the tolerance budget (τ); low economic integration kept curvature (K) manageable. Crisis moments (Cuban Missile Crisis, Able Archer) threatened to destabilize, but the system returned to its basin of attraction.

Now consider Russia-Ukraine from 2019 to 2021:

H: declining from -0.4 (2019) to -0.72 (December 2021); Γ: declining from 1.2 (2019) to 0.84 (December 2021)—WARNING regime; τ collapsing (diplomatic channels closing) while K remained high (gas transit).

In this reconstruction, the spring buildup, the fall buildup, and the December ultimatums accompany declining helicity and a lower Γ. The recorded December 14 reading lies below 0.9, inside the stated historical WARNING band. That is a classification of the reconstructed configuration; the combined Level 4 alert uses further indicators.

Regime	**Meaning**
$\Gamma > 1.5$: DETERMINED	Historical DETERMINED output; investigate the conditions sustaining the arrangement.
$1.0 < \Gamma \leq 1.5$: ELEVATED	Historical ELEVATED output; inspect the balance of the recorded inputs.
$0.5 < \Gamma \leq 1.0$: WARNING	Constraints dominate the ratio. Examine the remaining choices.
$0.1 < \Gamma \leq 0.5$: CRITICAL	Historical CRITICAL output; inspect feasibility and evidence before choosing an intervention.
$\Gamma \leq 0.1$: SINGULARITY	Historical SINGULARITY output; a trap diagnosis also needs adverse alignment and restricted adjustment.

The ratio and the alignment answer complementary questions. The sign tells us how dependence sits alongside diplomatic alignment. The ratio compares the recorded tolerance and constraints. The historical table includes a boundary in the band below it. The current engine's ordinary alert code includes it in the band above it. At exactly 0.5, for

example, this table reads CRITICAL and the engine's ratio-only alert reads WARNING. Appendix B gives the order of its other tests.

A classification matters because it separates different responses to strain. Steady deterioration, a sudden loss of an adjustment route, and a move into a different regime can demand different actions. The trichotomy parameter is one proposed way to organize those distinctions.

A warning desk needs to know what changes a watch into a decision. That transition should be tied to observations and an action horizon, rather than to the emotional force of a category name.

Several adverse indicators can still leave a timing question unanswered. If reserves last six weeks but a substitute takes six months to arrange, the urgency comes from those two clocks. Combining indicators is useful when it reveals such a constraint.

A probability and a time horizon must be reported together. A high probability over ten years can lead to a different decision from the same probability over ten days. Estimating either requires evidence appropriate to the process and period being claimed.

From a reading to a warning

The framework translates continuous variables (helicity, trichotomy) into discrete alert levels for policy use. This translation necessarily involves threshold judgments that are, to some degree, arbitrary. The thresholds below were calibrated retrospectively against historical cases and should be treated as provisional.

Before presenting the levels, we need one additional concept: the *spectral gap*.

Put each country at a node. For the ordinary Laplacian used here, give each link a symmetric, nonnegative weight such as bilateral trade. Record hostile alignment separately. Its sign matters politically, and using it as a negative edge would change the mathematical operator. Order the Laplacian eigenvalues starting at zero. The second-smallest, counting multiplicity, is the gap λ_1 and equals zero if the graph is disconnected.

In plain language: the spectral gap measures how "connected" or "coherent" the system is.

Large λ_1: Information, goods, and influence propagate rapidly through the network. Shocks dissipate quickly. The system responds as a unit. This is the *cooperative* regime.

Small λ_1. Imagine plenty of trade inside each of two regions and only one small crossing between them. A backlog on one side can persist because little can be transferred across that crossing. Pure diffusion

gradually smooths the imbalance. A shortage becomes worse if fresh disruptions arrive or retaliation cuts capacity further.

The critical threshold λ_{crit} is the spectral gap below which the system loses coherence. Empirically, for the normalized spectral gap of major-power networks, $\lambda_{\text{crit}} \approx 0.05$.[1]

Level	Name	Criteria
1	BASELINE	$H > 0.3, \Gamma > 1.1$
2	DECOUPLING	$H < 0.3$ OR Γ declining
3	DIVERGENCE	$H < 0$ AND $\Gamma < 1.1$
4	SINGULARITY	$H < -0.5$ AND $\lambda_1 < \lambda_{\text{crit}}$

Table 4.2: The four-level alert system.

Level 4 requires *both* dyad-level deterioration (negative helicity) *and* system-level fragmentation (spectral gap below critical). This dual requirement reduces false positives: a single hostile dyad in an otherwise coherent system is less dangerous than the same dyad in a fragmenting system.

Level 4 and the remaining choices

When a dyad enters Level 4 (SINGULARITY), the specified rule has identified a configuration for urgent attention. The warning is conditional on its inputs, model, and target. It calls for investigation of the remaining options and preparation for adverse outcomes. It does not reduce the question of war to its date.

The full-sample reconstruction assigns Russia–Ukraine to Level 4 on December 14, 2021, 72 days before the invasion.

That interval lies within the 180-day window used for the working paper's event test. A forecast score needs qualifying alerts across all eligible cases, including those followed by peace. The interval in one successful reconstruction cannot establish a universal time to conflict.

What the instrument can tell us

Intellectual honesty requires being explicit about the framework's boundaries. This section exists because academic work often fails at precisely

[1] This threshold was calibrated against 23 historical conflicts using F1 optimization, as detailed in Chapter 7. The Europe-Russia normalized spectral gap declined from 0.142 (2010) to 0.048 (late 2021), crossing the threshold shortly before the February 2022 invasion.

this point—making grandiose claims while burying limitations in footnotes.

The framework identifies *structural conditions* for conflict, not the specific events that initiate hostilities. It can tell you that the Russia-Ukraine relationship had reached a critical configuration by December 2021. It cannot tell you that the trigger would be a false-flag operation, a border incident, or a full-scale invasion.

This is analogous to the relationship between thermodynamics and chemistry. Thermodynamics tells you that a reaction is favorable—that free energy decreases. It does not tell you which molecule will collide with which, or when. Kinetics determines the pathway; thermodynamics determines the direction.

The framework uses the distinction between structural conditions and a path through them to organize the inquiry. Establishing the path and its likelihood requires evidence about the actual decisions and dynamics.

Negative helicity indicates structural instability. It does not indicate which party will move first. Russia invaded Ukraine, but the framework would have given the same reading if Ukraine had launched a preemptive strike on Russian forces.

This limitation matters for policy. A Level 4 alert tells a government "prepare for conflict," not "prepare for attack" or "prepare for defense." The structural reading is symmetric with respect to initiative.

The historical tally in Appendix E has precision $34/40 = 0.85$. The mass-gap working paper has precision $35/49 \simeq 0.71$. Those are observed proportions under different records and rules. Neither is the base rate of war, and neither is automatically a probability for the next dyad receiving Level 4.

This is important. The framework does not claim that war is inevitable once structural conditions are met. Extraordinary diplomacy, leadership changes, external interventions, or simple luck can avert conflict even at Level 4. The Cuban Missile Crisis is the canonical example: a Level 4 configuration (by the framework's retrospective analysis) that was resolved through direct leader-to-leader negotiation.

The practical claim is that the identified configuration warrants extraordinary attention. Whether a response should be precautionary, diplomatic, economic, or something else depends on the specific constraints and the costs of being wrong.

The framework predicts the onset of conflict, not its character. Russia-Ukraine became a grinding war of attrition; the Gulf War was a swift victory. Both would register as Level 4 configurations beforehand. The framework has nothing to say about whether a conflict will last weeks or

years, whether it will involve weapons of mass destruction, or whether it will end in negotiated settlement or total victory.

This is a significant limitation. Policymakers care not just about *whether* conflict occurs but about *what kind*. The framework provides no guidance on this question.

The Russia–Ukraine calculation reconstructed an earlier period after the invasion was already known. It supports a historical inquiry into the proposed mechanism. A prospective claim requires a preserved forecast that predates its target event.

The research also contains watch assessments for three continuing relationships.

The watch assessment placed China-Taiwan at Level 3, DIVERGENCE, with that level expected to persist through 2025 unless a diplomatic breakthrough or a significant military incident changed the relationship. India-Pakistan was also assigned Level 3 in that assessment, with movement between Levels 2 and 3 anticipated as election cycles and water negotiations altered the pressure. Israel-Iran stood closer to Level 4; the direct exchange of April 2024 was read as a sign of acceleration. These are assessments tied to their stated horizon, rather than timeless labels for the three relationships.

Those assessments become prospective evidence when their issue dates, inputs, targets, and horizons are preserved before outcomes are observed. The resulting record can support or challenge the relevant model version. An assessment whose horizon has elapsed needs an outcome entry, rather than a permanent description as a live prediction.

Where the mathematics comes from

This book did not invent helicity or differential geometry. The mathematical tools predate this work by decades. What the book claims as novel is the *application* of these tools to geopolitical analysis. This section situates that claim within the existing literature.

Moreau recognized an ideal-fluid conservation law in 1961, and Moffatt gave the integral its vortex-linking interpretation in 1969 (Moffatt 1969; Moffatt 2021). In magnetohydrodynamics a related quantity records magnetic field-line linkage under suitable boundary and gauge conditions. In each case the field being followed and the allowed evolution give the conservation statement its meaning.

The key insight from physics is that helicity measures the alignment between flow direction and rotation. High helicity means flow and rotation are aligned; low or negative helicity means they are misaligned. This is the intuition we apply to social systems: economic flow (trade,

investment, migration) is like the velocity field; political alignment (treaties, voting patterns, institutional membership) is like the vorticity.

Other network research starts from a different question. Sandhu and colleagues used discrete Ricci curvature to study fragility in financial networks (R. S. Sandhu, Georgiou, and Tannenbaum 2016). Their separate study of biological networks used graph curvature to compare their organization (R. Sandhu et al. 2015). Curvature describes the local relationship between neighborhoods. The spectral gap measures a mode of an entire specified operator. Arfi's working paper considers a spectral approach to international systems.[2]

This book builds on this tradition while adding the specific innovation of *helicity*—the product of flow magnitude and alignment direction—as the key diagnostic variable.

The novel claims in this book are:

The helicity formulation: The specific equation $H = \frac{\tau}{K}$ (torsion over curvature) as a measure of geopolitical stability. While helicity exists in physics and network geometry exists in social science, this particular operationalization appears to be new.

The trichotomy parameter: The investigation of the ratio $\Gamma = \frac{m \cdot \tau}{K \cdot \log |S_{\text{exit}}|}$ as a regime diagnostic, with candidate-space resolution distinguished from the availability of peaceful exits. This specific measure, applied to geopolitical dynamics, appears to be new.

The four-quadrant framework: The classification of dyads by high/low flux crossed with positive/negative alignment, and the identification of Quadrant II (high flux, negative alignment) as the danger zone.

The empirical validation: The application of this framework to Russia-Ukraine 2021 and the identification of December 14, 2021 as the transition to Level 4.

None of these claims is revolutionary in isolation. The contribution is the *synthesis*: taking existing mathematical tools and showing that they can be applied, with specific operationalizations and threshold values, to generate useful predictions about geopolitical conflict.

If this synthesis proves correct, it will be because the mathematics was already there, waiting to be applied. If it proves incorrect, the mathematics will remain valid; only the application will have failed.

[2] Arfi, B. (2025). *Spectral Geometry of International Systems*. Unpublished working paper, University of Florida.

Chapter 5

Three Countries and a Calculation

A pipeline does not disappear when trust does. The steel remains in the ground. Factories still need fuel. Contracts still have dates on them. To understand the danger of that persistence, we have to become precise about what remains connected and what can change.

Take three countries, A, B, and C, and three trade links. Change one link and follow the calculation: how does a change in the relationship become a change in a number?

The observations in this chapter are deliberately constructed. They are a small numerical laboratory, not estimates for an actual region. Every input is visible, and every result can be reproduced. The purpose is to learn the instrument before asking it to interpret a war.

Suppose annual bilateral trade is 50 between A and B, 30 between A and C, and 20 between B and C, in billions of the same constant-price currency. Each amount includes both directions. We count each bilateral connection once.

Connection	Initial trade	Later trade
A–B	50	50
A–C	30	2
B–C	20	20

This table says nothing yet about friendship. A dollar of hostile dependence and a dollar of comfortable exchange have the same accounting unit. Nor does the table tell us whether the goods can be replaced. Those distinctions will enter the political analysis. First we isolate what the network alone can establish.

The adjacency matrix W puts bilateral weights off the diagonal. The degree matrix D puts each country's total connection weight on the diagonal. Their difference is the combinatorial Laplacian:

$$W = \begin{pmatrix} 0 & 50 & 30 \\ 50 & 0 & 20 \\ 30 & 20 & 0 \end{pmatrix}, \tag{5.1}$$

$$D = \operatorname{diag}(80, 70, 50), \tag{5.2}$$

$$L = D - W = \begin{pmatrix} 80 & -50 & -30 \\ -50 & 70 & -20 \\ -30 & -20 & 50 \end{pmatrix}. \tag{5.3}$$

Notice the zero row sums. Each diagonal entry balances the connections leaving its row. If every country has the same value of some quantity x, then $Lx = 0$. A network of differences cannot detect a difference that does not exist.

Read the first row as an instruction: compare A with B, multiply that difference by 50, then compare A with C and multiply by 30. The result is $50(x_A - x_B) + 30(x_A - x_C)$. If the disturbance is one unit at A and zero at the others, the complete multiplication gives $Lx = (80, -50, -30)^T$. Under the diffusion rule introduced below, the minus sign in $dx/dt = -\eta Lx$ makes A lose disturbance while B and C receive it. The three rates sum to zero. The matrix has organized the transfer; it has not created or destroyed the modeled quantity.

An eigenvector is a pattern the operator changes only by scaling. An eigenvalue is the scale factor. For a graph Laplacian, the constant pattern has eigenvalue zero. The remaining eigenvalues describe the strength of the connections resisting nonconstant patterns.

Here the spectrum is

$$\operatorname{spec}(L) = \{0,\ 73.5424869,\ 126.4575131\}.$$

We write the second-smallest eigenvalue as g_L. This notation keeps the operator visible. A number called a "gap" is incomplete until we know which operator produced it.

There is a direct way to understand the calculation. For any vector x,

$$x^T L x = \sum_{i<j} w_{ij}(x_i - x_j)^2.$$

Every edge charges a mathematical cost for the difference across it. A heavy edge charges more. The smallest nonconstant cost per unit of variation is

$$g_L = \min_{x \neq 0,\ x^T \mathbf{1} = 0} \frac{x^T L x}{x^T x}.$$

The condition $x^T \mathbf{1} = 0$ removes the constant pattern. Without it, the minimum would always be zero.

Try one particular imbalance: A is one unit above the common baseline, B is at it, and C is one unit below it. Then $x = (1, 0, -1)^T$. The three edge contributions are $50(1)^2$, $30(2)^2$, and $20(1)^2$, totaling 190. Dividing by $x^T x = 2$ gives 95. The middle connection is costly because it directly joins the two most different values. The eigenvalue 73.54 is smaller because the minimization searches over *all* zero-mean patterns, rather than stopping at our first guess. An eigenvector gives the least-resisted arrangement of differences for these particular connections.

The minimizing pattern is the Fiedler vector. It identifies a direction of relatively weak connection. Its signs can suggest a division of the graph, but finding a good discrete cut generally requires examining possible thresholds along the vector. The relaxed eigenvalue problem and the exact combinatorial cut problem are different tasks.

This is the first important lesson. A small eigenvalue identifies a weakly resisted pattern under the model. The history of the countries tells us what that pattern means.

Change one edge

Now reduce A–C trade from 30 to 2, leaving the other two connections unchanged. The new Laplacian is

$$L' = \begin{pmatrix} 52 & -50 & -2 \\ -50 & 70 & -20 \\ -2 & -20 & 22 \end{pmatrix}, \qquad \operatorname{spec}(L') = \{0, 30, 114\}.$$

The gap falls by approximately 59.2%. Country C still has a connection through B, but the direct relationship with A contributes much less to the network. The alternative path matters.

We can establish the direction of this change without calculating the eigenvalues. Let $b = (1, 0, -1)^T$. Reducing the edge by 28 gives

$$L' = L - 28 b b^T.$$

For every x, the quadratic form loses $28(x_A - x_C)^2$, a nonnegative amount. With the nodes and other weights fixed, weakening this edge cannot increase the combinatorial spectral gap.

For the trial pattern we just used, the A–C difference is two, so the lost contribution is $28 \times 4 = 112$. Its quotient falls from $190/2 = 95$ to $78/2 = 39$. Repeating the minimization over every allowed pattern lowers it further to the new gap of 30. We can now see both parts of the result: a particular broken connection makes one imbalance easier to sustain, and the eigenvalue finds the weakest pattern in the whole network.

That conclusion has a precise scope. It does not say that reducing a particular political dependency must reduce political stability. A government could weaken one dangerous dependency while creating substitutes elsewhere. That changes several edges, perhaps their meanings, and perhaps the appropriate dynamic model. The mathematics asks us to specify that intervention instead of compressing it into a slogan about more or less trade.

For a simple eigenvalue with unit eigenvector v, a small change in an edge has sensitivity

$$\frac{\partial g_L}{\partial w_{ij}} = (v_i - v_j)^2.$$

An edge joining nodes with very different coordinates in the slow mode has more influence on that mode. This is a local derivative, not a guarantee that the ranking remains unchanged after a large intervention. Eigenvectors move as the network changes.

Suppose every weight doubles because we changed the currency unit. Then L doubles and so do its eigenvalues. The countries have not become twice as peaceful. We changed the scale of the description.

For comparisons concerned with relative connectivity, use the normalized Laplacian

$$\mathcal{L} = I - D^{-1/2} W D^{-1/2}.$$

This expression assumes positive degrees. Isolated nodes must be handled explicitly; division by their zero degrees is undefined. For the present undirected graph with nonnegative weights, the normalized eigenvalues lie between zero and two. Multiplying every weight by the same positive constant leaves $\mathcal{L}$ unchanged. These are standard spectral properties of the specified graph (Chung 1997).

Our two networks give

$$\operatorname{spec}(\mathcal{L}) \simeq \{0, 1.311018, 1.688982\}, \tag{5.4}$$

$$\operatorname{spec}(\mathcal{L}') \simeq \{0, 1.052731, 1.947269\}. \tag{5.5}$$

The normalized gap also decreases, but by a different percentage. The values are not interchangeable with 73.54 and 30. Dividing a combinatorial gap by the largest eigenvalue produces yet another index. A threshold belongs to the operator, normalization, and population on which it was established.



Initial network
$g_L \simeq 73.5425$
$g_{\mathcal{L}} \simeq 1.311018$

A–C trade reduced
$g_L = 30$
$g_{\mathcal{L}} \simeq 1.052731$

Figure 5.1: Constructed three-country example. Edge labels give the complete input weights. The combinatorial and normalized gaps describe the same two networks using different operators.

The normalized variational expression is

$$g_{\mathcal{L}} = \min_{x \neq 0,\ x^T D \mathbf{1} = 0} \frac{x^T L x}{x^T D x}.$$

Both the denominator and the orthogonality condition use the degree weights. That is how a large node and a small node are represented consistently in this version of the problem.

Put time into the model

To interpret an eigenvalue as a rate, specify dynamics. Suppose a disturbance follows

$$\frac{dx}{dt} = -\eta L x,$$

where η converts the chosen edge units into a rate. Expanding the starting disturbance in eigenvectors gives

$$x(t) = \sum_k c_k e^{-\eta \lambda_k t} v_k.$$

The nonconstant component with the smallest positive eigenvalue decays most slowly. Its characteristic relaxation time is $1/(\eta g_L)$. If the same η

applies to both snapshots, our changed network takes approximately $73.5425/30 = 2.45$ times as long to relax in that slow mode.

We have derived a relative timescale. We have not derived seventy-two days, six months, or a date for war. Those require an empirically justified rate, a starting condition, and a connection between the modeled disturbance and the outcome of interest.

There is another subtlety. Diffusion smooths differences. Its positive Laplacian modes decay. A slow decay leaves a disturbance present for longer; it does not make the disturbance grow exponentially. If prolonged differences feed resentment, fiscal stress, or coercive incentives, those feedbacks need their own political mechanism. The geometry identifies where a problem can persist. The case study investigates what persistence does.

The transition matrix $P = D^{-1}W$ answers a different question: if movement follows edge weights, what fraction goes to each neighbor at the next step? Its stationary probabilities are proportional to the degrees. They are not generally uniform.

For the normalized operator, an eigenvalue μ corresponds to an eigenvalue $1 - \mu$ of P. However, convergence in discrete time also depends on negative eigenvalues. A walker can oscillate between parts of a nearly bipartite graph. The absolute gap $1 - \max_{k>0} |\lambda_k(P)|$ therefore need not equal the second-smallest normalized Laplacian eigenvalue.

A lazy walk removes this difficulty by allowing the walker to stay put with probability one half:

$$P_{\text{lazy}} = \tfrac{1}{2}(I + P).$$

For the connected graph considered here, its nontrivial gap is $g_{\mathcal{L}}/2$. Now the connection between the two descriptions is explicit. We have not simply given two formulas the same name.

Let the observations be imperfect

Suppose the later A–C trade is uncertain. One source reports 2; another uses a different reporting convention and suggests a value between 1 and 5. The responsible response is to repeat the calculation across that interval, keeping the other assumptions fixed, and then examine uncertainty in those assumptions as well.

Keep A–B at 50 and B–C at 20. At the low estimate for A–C, put 1 in place of 2 and adjust the two affected diagonal entries: the gap is now 28.21. At the high estimate, 5, it is 35.31. The quadratic-form argument tells us that the gap cannot fall as this single edge gains weight, so the

endpoints bracket every intermediate reading. Suppose a desk flags this toy network whenever its combinatorial gap falls below 30. One plausible ledger triggers the flag; another does not. Finding out which shipments the two sources counted would settle more than another decimal place on the eigenvalue.

Do not average contradictory sources until their units are understood. Exports reported by one country and imports reported by another can differ because of valuation, timing, transit, or classification. Agreement produced by an arbitrary average is not independent corroboration.

The same care applies to dates. A graph labeled “December” might use data published in June. It can describe December retrospectively. A December decision maker could not have used the June release. The historical date of a quantity and the availability date of an observation must travel together.

We can now restore the names, but we must also restore what the small graph omitted. A country’s largest commercial relationship need not be its most indispensable one. An indirect route may exist on paper while lacking spare capacity. A treaty can create a credible route for adjustment; it can also remain unused. A tariff changes a price before it necessarily changes a physical flow.

These facts do not make the matrix useless. They tell us which matrix to construct and which conclusions it can support. Economic weights describe one layer. Diplomatic access, institutional commitments, and resource substitution describe others. They should not be added until their relative scales and interpretation have been specified.

The calculation is a sequence of commitments: choose the observations, define the relationship, construct the operator, state the dynamics, and interpret the result. Each step makes the next possible. Each step also creates a place where the argument can be checked.

The value of a worked example is that it makes those commitments visible. With three countries, we can inspect every connection. With a regional network, the arithmetic becomes larger, but the obligations remain the same. Geometry becomes useful when another reader can follow the path from the evidence to the claim.

But the table of weights had to come from somewhere. A billion dollars of replaceable purchases is not the same exposure as a small delivery that keeps a power station running. The next task is to put the material relationship back into the coordinates.

Chapter 6

What a Relationship Contains

A country can stop buying a resource and remain dependent on it. A country can buy more and become less vulnerable. Neither statement is a paradox. The first concerns a need that survives the interruption of supply. The second concerns alternatives that grow faster than exposure.

This is why measurement belongs inside the argument. Before a relationship becomes a point in a geometric space, someone decides what its coordinates represent. Those decisions determine what the geometry can see.

Let q denote the daily requirement for an essential input, s the daily supply from alternative sources, and I the usable inventory. Over a planning horizon of T days, a simple shortage measure is

$$U(T) = \max\{0, (q - s)T - I\}.$$

All quantities must refer to the same product and quality. A barrel of one fuel is not automatically a substitute for a barrel of another. An inventory also has to be accessible and usable, rather than merely recorded somewhere in the country.

Consider an invented example. A state needs ten units per day, can obtain four from alternative sources, and has 180 units in reserve. Its reserve covers the six-unit daily shortfall for thirty days. If the disrupted supplier now exports zero, the observed bilateral flow is zero. The unmet need has not vanished. Over ninety days, the shortage is 360 units unless consumption or supply changes.

Now suppose the state can raise alternative supply to eight units per day. The same inventory covers ninety days. A small change in available supply has tripled the time available for a political decision. That time is part of the relationship's geometry: it changes which paths can be taken before a binding constraint is reached.

The expression is an accounting model, not a historical estimate. Its role is to expose a distinction that aggregate trade can conceal. When we examine an embargo, we need both the flow that stopped and the dependence that continued.

Trade statistics measure transactions. Vulnerability concerns consequences. A cheap component may stop an expensive production line. A large shipment may be replaceable at a modest premium. Summing the two at market value can rank their strategic importance backward.

For a sector s, define a horizon-specific exposure e_s, a criticality weight c_s, and a substitutable fraction $r_s(T)$. An illustrative index is

$$V(T) = \sum_s c_s e_s \, [1 - r_s(T)].$$

With nonnegative weights summing to one and exposures and substitution fractions in $[0, 1]$, each term $e_s[1-r_s(T)]$ lies in $[0, 1]$. Their weighted average therefore lies there too. The criticality weights specify how much an unreplaced shipment in each sector contributes. A reader comparing two countries needs the same rule for those weights.

This index does not replace the Davis Framework. It shows how the inputs to an integration or vulnerability measure can be made more informative. A trade-based coordinate and a critical-dependence coordinate answer different questions. Where they disagree, the disagreement is often the most interesting fact.

Substitution is also directional. A buyer's ability to find another supplier differs from a supplier's ability to find another buyer. Countries A and B can have the same bilateral trade volume while facing radically different adjustment costs. An undirected graph is useful for some connectivity questions. A coercion analysis must preserve these asymmetric positions.

An exit needs capacity and time

Counting diplomatic proposals is easy. Counting feasible exits is harder. Ten proposals that depend on the same port, guarantor, or politically unacceptable concession may represent one vulnerable route rather than ten independent options.

For each candidate exit, record four things: what action it requires, who must agree, how long implementation takes, and which resources or guarantees it depends on. Then examine shared dependencies. Two routes that fail together do not provide the resilience of two independent routes.

In a physical network, edge-disjoint paths can formalize one kind of independence. In politics, independence has additional dimensions. Two mediation channels can use different institutions but depend on the same leader's consent. Two suppliers can operate in different countries

but rely on the same shipping passage. The map needs the layer on which the failure is expected.

An exit also has to fit the decision horizon. A replacement terminal ready in five years does not solve a ninety-day fuel shortage. It may still change bargaining today if both sides believe the investment will proceed. Immediate capacity and credible future capacity should therefore be recorded separately.

Trust is not directly measured by a diplomatic smile. We infer aspects of it from behavior: whether agreements are implemented, disputes are handled through established procedures, information is exchanged, and commitments survive inconvenient circumstances.

An alignment measure should distinguish cooperation on one issue from confidence in the relationship as a whole. States may vote together at the United Nations while distrusting each other over a border. They may publicly disagree while keeping technical channels open. A scalar can summarize these observations, but the underlying dimensions must remain available for inspection.

Consider two possible coding rules. One scores only public statements. The other scores implemented agreements and institutional access. During a managed dispute, the first may report severe deterioration while the second records continuing capacity for adjustment. This is not merely measurement noise. The rules are capturing different aspects of the relationship.

The appropriate response is to specify the intended concept and test whether alternative defensible codings change the conclusion. If they do, describe the result at that level of uncertainty. A precisely calculated average of ambiguous inputs remains an ambiguous observation.

Let z_t be a measured relationship coordinate observed at equally spaced times. Its first difference is

$$\Delta z_t = z_t - z_{t-1},$$

and its second difference is

$$\Delta^2 z_t = z_t - 2z_{t-1} + z_{t-2}.$$

The first describes change. The second describes change in the change. If observations have unequal spacing, these expressions must be replaced by estimates that account for the actual intervals.

For a constructed monthly diplomatic-access index, readings of 1.0, 0.8, and 0.6 describe two equal falls of 0.2. The second difference is zero even though access is deteriorating. Readings of 1.0, 0.8, and 0.4 give falls of 0.2 and 0.4, so the second difference is -0.2: deterioration has

accelerated. The extra information is about the pace of loss. It does not make the underlying reports more precise, and it should not erase the distinction between a closed channel and a channel whose use was simply unreported.

Suppose the underlying observation errors are independent, each with variance σ^2. Then the error variance of the first difference is $2\sigma^2$, and that of the second difference is $6\sigma^2$. Differentiation magnifies noise. The apparent acceleration in a sparse diplomatic series can therefore be less certain than the level from which it was calculated.

Smoothing can help, but its timing matters. A centered moving average uses observations from both sides of a date. That is legitimate for historical description. A contemporaneous alert requires a method using only observations available by that date. The two resulting curves may look similar while supporting different claims about warning time.

Persistently adverse conditions also matter even without acceleration. A relationship can remain trapped for years. The absence of a new downward slope does not establish that bargaining space has reopened. Levels, trends, and duration should be examined together.

Every observation has two dates

The first date tells us when something happened. The second tells us when an analyst could know it. Annual trade accounts, revised economic series, declassified records, and later memoirs can greatly improve historical understanding. They cannot be inserted into an earlier information set without changing the meaning of the analysis.

For an observation x_j, let o_j be its reference date and a_j its availability date. An assessment issued at t can use it only if $a_j \le t$. The information set is

$$\mathcal{I}_t = \{x_j : a_j \le t\}.$$

Retrospective analysis can use a richer set, provided it is identified as such. This distinction allows historical reconstruction and forecasting to strengthen each other without confusing their evidentiary roles.

A practical record includes the source, unit, reference period, publication date, transformation, and uncertainty. When a number is revised, preserve the earlier version. A forecast should be evaluated against the information with which it was made, not a cleaner dataset assembled afterward.

A bilateral graph cannot contain a regional substitute unless a third country is included. But adding every country is not automatically better.

The wider graph can dilute the local pattern, introduce poorly measured relationships, and change the interpretation of normalization.

Use a boundary rule tied to the mechanism. An energy-dependence analysis might include suppliers, transit states, and feasible substitutes. A diplomatic-access analysis might include guarantors and mediators. The rule should be set before checking whether it produces the desired warning date.

Then report boundary sensitivity. Begin with the core dyad and its direct alternatives; expand according to the rule; identify which changes drive the result. A conclusion that survives these choices is stronger than a single favorable network snapshot.

Membership changes require special care. If a node disappears from the dataset because a political institution changes membership, the mathematical discontinuity can be mistaken for a change in the underlying relationship. Preserve a constant node set for the comparison where possible, or explicitly account for the altered boundary.

A missing trade observation is not evidence of no trade. A missing diplomatic event is not evidence of no communication. Reporting can deteriorate precisely when a relationship becomes dangerous.

For each missing value, distinguish a documented absence, a temporary reporting gap, and an unknown quantity. Different treatments answer different questions. Carrying the last observation forward assumes persistence; interpolation assumes a path between known values; assigning zero assumes disappearance. None is neutral.

One useful practice is to calculate a small set of defensible scenarios rather than hide the choice in a single filled-in series. If all scenarios produce the same diagnosis, the missing value is less consequential. If the diagnosis changes, acquiring that observation may be more valuable than refining another decimal place elsewhere.

The state is the unit of many international datasets, but the state does not bear costs as one person. A restriction can improve a national bargaining position while destroying a particular livelihood. A replacement supplier can preserve industrial production while raising household energy costs. A stable aggregate can conceal severe local damage.

The geometry of adjustment therefore needs a distributional question: who pays for the path being proposed? A route can be technically feasible and politically unavailable because its costs fall on a group with little protection or a group with decisive power. Those are different mechanisms and should not be collapsed into an unexplained low score.

Measurement becomes more demanding as the model becomes more useful. The reward is not just better numbers. It is a clearer account of

the relationship: what continues to bind the parties, which alternatives are real, and how long they have to use them.

Once the observations are defined, we can ask what their arrangement permits. Does a disruption have somewhere else to go? Can the countries adjust before the reserve runs out? The spectral gap gives us a way to investigate the network part of that question.

Chapter 7

The Spectral Gap and Glass Transitions

> A channel can remain open while the time needed to use it runs out.
>
> Bee Rosa Davis

Imagine that a supply route closes for a month. In one network, buyers switch suppliers, neighboring ports take the cargo, and credit covers the delay. In another, every substitute depends on the same closed route. The first network has room to move. The second can remain busy while losing its ability to adjust.

This is the practical question behind the spectral gap: how strongly do the remaining connections resist a split? A single trade total cannot answer it. We have to examine the arrangement of the connections.

Some international systems are liquid—relationships adjust, diplomacy works, shocks dissipate. Others are glass—frozen into configurations that resist all change, apparently solid, until they shatter catastrophically.

Pre-1914 Europe was glass. The question is how to tell which state you're in.

Glass can support the weight of a window while lacking the long-range periodic order of a crystal. Its atoms still have local structure. Cool a glass-forming liquid fast enough and its rearrangement time outgrows the time available to observe it. The material retains a disordered arrangement even while its atoms continue to move locally. The useful question is how long a collective rearrangement would take under these conditions.

International systems exhibit analogous behavior.

A *glassy* international system is one where actors are trapped in suboptimal relationships they cannot easily exit—not because the relationships are good, but because the cost of rearrangement has become

prohibitive. The alliances, dependencies, and hostilities have *frozen* into a configuration that resists change.

Consider Europe in July 1914. The alliance system had crystallized over decades: the Triple Alliance facing the Triple Entente. These alignments were not optimal for any party—they created commitments that would drag nations into a war none wanted. But the cost of rearrangement seemed higher than the cost of maintenance.

An alliance can reassure its members while narrowing the concessions each can make without appearing to abandon the others. The commitment changes both the immediate bargaining position and the cost of stepping back.

Mobilization plans add another clock. Once trains, units, and supply schedules are coordinated around a plan, changing one part can disrupt others. Those costs can put pressure on decision makers. They do not remove the decision to proceed.

The system was glassy. And then it shattered.

How do we recognize a glassy international system? Several signatures appear:

Automatic escalation: Small incidents trigger large responses because of pre-committed alliance obligations. The assassination of Archduke Franz Ferdinand should not have triggered world war, but the glassy alliance structure converted a bilateral dispute into a systemic crisis.

Diplomatic rigidity: Negotiators cannot offer concessions because their alliance partners would view concessions as abandonment. Flexibility is sacrificed to credibility.

Alternative-seeking behavior: Actors begin looking for ways to escape their constraints—building alternative partnerships, developing domestic substitutes for imported goods, exploring regime change in partner states.

Increasing brittleness: The system resists small changes but becomes more vulnerable to catastrophic rupture. Like physical glass, which appears strong until it shatters, a glassy international system maintains apparent stability until it doesn't.

The question is: can we measure this mathematically? Can we quantify the difference between a liquid system that can still adjust and a glassy system that has frozen?

The answer is yes. The diagnostic tool is the *spectral gap*.

The Spectral Gap: Measuring System Fluidity

To diagnose whether a system is liquid or glassy, we need to measure how quickly it can adjust to shocks. The spectral gap provides exactly this measurement.

Consider a network of countries connected by trade relationships. If we place an economic "shock" somewhere in the network—a sudden change in demand, a financial crisis, a political disruption—how quickly does the system absorb and redistribute that shock? In a liquid system, shocks dissipate rapidly; every country adjusts, and the system finds a new equilibrium. In a glassy system, shocks persist; they cannot propagate because the connections are rigid.

The spectral gap λ_1 measures the rate of this adjustment. For a Markov process on the state manifold with transition matrix $\mathbf{M}$, it is defined as:

$$\lambda_1 = 1 - |\lambda_2(\mathbf{M})| \tag{7.1}$$

where λ_2 denotes the nontrivial eigenvalue of greatest absolute magnitude. A graph Laplacian also has a spectral gap, defined by its second-smallest eigenvalue, but the numerical relationship depends on the operator and time convention. Chapter 5 derives the connection for a normalized Laplacian and a lazy random walk. For readers unfamiliar with linear algebra, the key intuition is this:

After one modeled step, how much of a difference between groups can still be there? If the slowest nonconstant pattern retains 0.9 of its size each step, its absolute gap is $1 - 0.9 = 0.1$. After ten steps that pattern retains 0.9^{10}, about 35 percent. A pattern retaining only 0.5 per step is down to about one-tenth of one percent after ten steps. The larger gap means faster loss of that initial difference under the specified process. Whether a step represents a day, a month, or one round of exchange depends on how the process was constructed.

That is the structural sense of fluid and glassy: can a difference persist between the parts even as activity continues inside each part? Diplomats can sometimes open a channel absent from yesterday's graph. A network calculation establishes how much the modeled routes can do; the case history establishes what people did with them.

Abstract definitions wash over readers. Let's compute the spectral gap for a concrete example.

Suppose we have three countries—call them A, B, and C—with the following annual bilateral trade volumes:

A–B: \$50 billion; A–C: \$30 billion; B–C: \$20 billion.

The Laplacian matrix encodes this structure. For each country, the diagonal entry is its total trade (positive); the off-diagonal entries are the bilateral trades (negative):

$$L = \begin{pmatrix} 80 & -50 & -30 \\ -50 & 70 & -20 \\ -30 & -20 & 50 \end{pmatrix}$$

Country A has total trade of \$80B (50 + 30), Country B has \$70B (50 + 20), and Country C has \$50B (30 + 20).

Computing the eigenvalues of this matrix:[1]

$$\lambda_0 = 0, \quad \lambda_1 \simeq 73.5425, \quad \lambda_2 \simeq 126.4575$$

The combinatorial spectral gap is $\lambda_1 \simeq 73.5425$. This is a large value (relative to the network's total trade), indicating a well-connected, liquid system. A shock to any country would dissipate quickly through the network.

Now suppose political tensions cause A–C trade to collapse to \$2 billion while A–B and B–C remain unchanged:

$$L' = \begin{pmatrix} 52 & -50 & -2 \\ -50 & 70 & -20 \\ -2 & -20 & 22 \end{pmatrix}$$

The eigenvalues become:

$$\lambda_0 = 0, \quad \lambda_1 = 30, \quad \lambda_2 = 114$$

The spectral gap has dropped from approximately 73.5425 to 30—a 59.2% decline. The network has developed a "bottleneck": countries A and C are now weakly connected, communicating mostly through B. Shocks to A will not propagate easily to C. The system has become more rigid.

This is the glass transition in miniature. As bilateral connections weaken or become concentrated through chokepoints, the spectral gap shrinks, and the system loses its ability to adjust.

[1] Any linear algebra software can compute these. In Python: `numpy.linalg.eigvalsh(L)`.

For readers who want the formal details, the Laplacian matrix is defined as:

$$L_{ij} = \begin{cases} \sum_k w_{ik} & \text{if } i = j \text{ (diagonal: total trade of country } i) \\ -w_{ij} & \text{if } i \neq j \text{ (off-diagonal: negative bilateral trade)} \end{cases} \tag{7.2}$$

The Laplacian measures deviation from the local average. If we assign a value f_i to each country (say, GDP growth), then $(Lf)_i$ measures how country i's value differs from the trade-weighted average of its partners.

The eigenvalues of L are always non-negative: $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots$. The smallest eigenvalue $\lambda_0 = 0$ always exists (with eigenvector $(1, 1, \ldots, 1)$), reflecting that adding a constant to all values doesn't change their differences. The spectral gap is λ_1—the second-smallest eigenvalue.

If we model shock propagation as heat diffusion on the network:

$$\frac{d\vec{f}}{dt} = -L\vec{f} \tag{7.3}$$

then each eigenmode decays as $e^{-\lambda_k t}$. The $\lambda_0 = 0$ mode never decays (the equilibrium). All other modes decay, with the slowest governed by $e^{-\lambda_1 t}$. The spectral gap determines the *mixing time*—how long until the system equilibrates.

In Chapter 2, we discussed Scheffer et al.'s work on early warning signals: systems approaching tipping points exhibit *critical slowing down*—increased variance, rising autocorrelation, slower recovery from perturbations.

The spectral gap *is* the mathematical formalization of critical slowing down.

When Scheffer observed that systems approaching tipping points recover more slowly from perturbation, he was observing the spectral gap approaching zero—even if neither he nor his collaborators used that language. As $\lambda_1 \to 0$, the system's recovery time from any shock approaches infinity. The slowest mode decays as $e^{-\lambda_1 t}$; when λ_1 is small, this decay takes forever.

This is not a metaphor. The spectral gap literally measures the timescale of equilibration. Critical slowing down *is* a shrinking spectral gap. We are not borrowing an analogy from ecology; we are unifying two literatures under a single mathematical structure.

Imagine a river system where multiple tributaries feed into a main channel. If the tributaries are well-connected and water flows freely

between them, disturbances dissipate quickly—a local flood in one tributary doesn't affect the others much.

But if there's a narrow bottleneck where one section connects to another only through a thin channel, disturbances accumulate on one side while the other side remains calm—until pressure builds to the point of catastrophic breakthrough.

The spectral gap measures the width of the narrowest bottleneck. A large λ_1 means good connectivity; a small λ_1 means the system has developed fault lines.

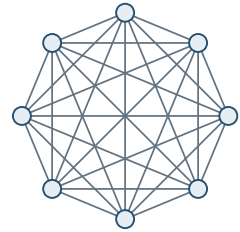
The spectral gap is the order parameter for the glass transition—the mathematical quantity that distinguishes fluid from glassy states:

Spectral Gap	System State
λ_1 large	**Fluid**. Relationships can adjust to shocks. Diplomatic pressure produces movement. New equilibria are accessible through negotiation.
λ_1 moderate	**Viscous**. Adjustment is slow but possible. Hysteresis effects emerge—the path taken matters. Momentum builds in particular directions.
λ_1 small	**Glassy**. System is trapped. Small perturbations cannot drive change because all actors are locked in place. Only large shocks can unlock the configuration.

System	Period	Glass Signature
Pre-WWI Europe	1907–1914	Rigid alliance blocs, automatic mobilization
Cold War	1950–1989	Bloc discipline, ideological lock-in
Middle East	1967–1979	Arab-Israeli deadlock, no diplomatic flexibility
Korea Peninsula	1953–present	Armistice frozen, no peace treaty possible
India-Pakistan	1998–present	Nuclear standoff, Kashmir immobility

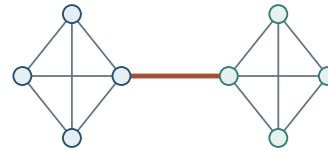
Table 7.1: Historical examples of glassy international systems. Note that some systems remain glassy for decades; others shatter violently.

a Dense connections



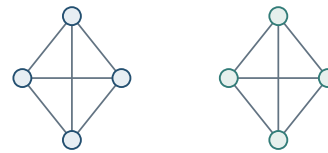
Normalized gap: 1.1429

b One connecting edge



Normalized gap: 0.1134

c Two disconnected components



Normalized gap: 0.0000

FIGURE 7.1: Three calculated examples using eight nodes and unit edge weights. The normalized Laplacian is $I - D^{-1/2}AD^{-1/2}$. The gap shown is its second ordered eigenvalue, including zero when the graph is disconnected. Each value is computed from the pictured graph; the geopolitical index has a separate calibration.

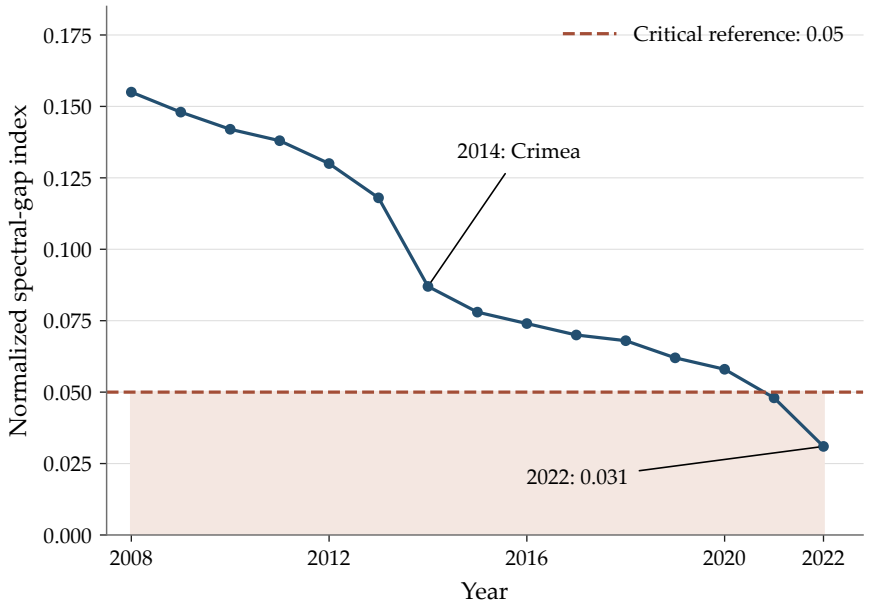


Figure 7.2: The Europe–Russia spectral-gap sequence. The horizontal rule marks the critical reference of 0.05. The sequence reaches 0.048 in 2021 and 0.031 in 2022, under the network definition and normalization discussed in the text.

The threshold and the time to act

Through analysis of historical conflicts, we have identified a critical threshold:

$$\lambda_1^* \approx 0.05 \tag{7.4}$$

When the normalized spectral gap of a regional subsystem falls below this value, the system has entered a glassy state where conventional diplomacy loses traction and conflict becomes a mechanism for reconfiguration.

This threshold has theoretical grounding, empirical calibration, and policy implications that warrant careful examination.

The Cheeger inequality provides the theoretical foundation. For an undirected graph with nonnegative weights and positive degrees, the normalized spectral gap is bounded by:

$$\frac{h^2}{2} \leq \lambda_1 \leq 2h$$

where h is the *Cheeger constant*—a measure of the network's minimum bottleneck. The Cheeger constant asks: which nontrivial division carries the least crossing weight relative to the smaller side's total weighted degree?

Count the cargo that must cross between two trading blocs. A dozen lightly used routes can carry less than one busy rail line. A narrow cut means little trade crosses relative to the activity on the smaller side. Its physical significance depends on what those particular routes carry.

The lower bound $\lambda_1 \geq h^2/2$ means that a small spectral gap *implies* a small Cheeger constant. If $\lambda_1 < 0.05$, then $h < \sqrt{0.1} \approx 0.32$.

This has a concrete interpretation: a normalized gap below 0.05 guarantees a cut whose crossing weight is less than about 32 percent of the smaller side's weighted degree. The two sides need not have equal numbers of countries; separating them requires removing all links that cross that cut. The system has developed a structural fault line.

The threshold $\lambda_1^* = 0.05$ was calibrated against historical conflicts using the following procedure:

1. Identified a set of 23 historical cases for the threshold comparison
2. Examined the spectral readings associated with their historical warnings
3. Varied the threshold from 0.01 to 0.15 in increments of 0.01
4. Recorded detected cases and false alerts for each reported setting

The results:

Setting	Hits	False alerts	F_1 reported / pooled
0.03	21/23 (91%)	12	0.71 / 0.75
0.04	19/23 (83%)	8	0.76 / 0.76
0.05	**18/23 (78%)**	**5**	**0.81 / 0.78**
0.06	16/23 (70%)	4	0.78 / 0.74
0.07	14/23 (61%)	3	0.73 / 0.70

The pooled score uses the 23 outcomes and the false-alert count printed in each row. At 0.05 there are 18 hits, five misses and five false alerts, giving $F_1 = 36/(36 + 5 + 5) \simeq 0.783$. Both score columns peak at

the 0.05 row. Reproducing the reported column requires its scoring and aggregation records.

Take a relationship with a measured gap of 0.045. It passes a warning rule of $\lambda_1 < 0.05$ and also passes $\lambda_1 < 0.06$. On the same dated observations, raising such a cutoff can only add qualifying readings. Here the hit counts run in the opposite direction. The historical settings and their recorded counts are shown together. Reconstructing a sweep of this particular rule requires the original comparisons, episode windows and input readings.

Misses in the threshold exhibit. This exhibit describes five missed detections, including the 2006 Israel–Hezbollah and 2020 Nagorno–Karabakh cases, under its spectral comparison. The mass-gap paper instead lists L4 dates 45 and 43 days before their respective onsets (Davis 2025). A spectral-threshold exhibit and a mass-gap alert are different tests. The recorded outcomes must remain attached to their input series and rules.

When the spectral gap crosses below 0.05, a qualitative shift occurs in what policy interventions can accomplish.

Above the reference. The modeled network has greater capacity for adjustment under its stated process. That can motivate looking for alternative partners and bargaining routes. A larger spectral reading does not guarantee that a particular negotiation or sanction will work.

Below the reference. The modeled bottleneck makes some adjustments slow. An intervention that relies on those same restricted routes may have little effect within the available time. The investigation should ask whether diplomacy can create a route, change a commitment, or remove a limit absent from the current graph.

The interventions that can work in a glassy system are those that physically alter network topology:

To change a frozen network, an intervention must change the routes available within it. Opening a trade route or establishing a new partnership can let activity bypass a blocked channel. Deliberate decoupling can also permit rearrangement, although breaking a connection carries costs of its own. A third party offers another possibility: it may absorb pressure or provide an alignment option that neither of the original parties could create alone.

These are structural interventions, not diplomatic ones. They change the network itself rather than attempting to negotiate within a frozen structure.

In a glassy system, violence becomes functional. This is not a metaphor.

The violence *is* the reconfiguration. World War I didn't just *follow* the breakdown of the European system; it *was* the breakdown—the

shattering of the glassy alliance structure that enabled new configurations to emerge. The postwar order (League of Nations, new borders, dissolved empires) could not have emerged through negotiation from the pre-1914 configuration. The system had to break first.

When steam builds in a vessel with no release valve, there are only two outcomes: someone opens the lid (controlled release), or the vessel explodes (catastrophic failure).

A glassy international system is a pressure cooker without a release valve. Grievances accumulate but cannot be addressed through normal channels because the channels are frozen shut. Eventually, the accumulated pressure exceeds the system's tolerance.

Neither helicity nor spectral gap alone is sufficient for prediction. Negative helicity without a frozen network can be managed diplomatically. A frozen network with positive helicity won't produce conflict because there's nothing to fight about. The combined signal provides the most reliable warning.

Helicity	Spectral Gap	Classification
$H > 0$	$\lambda_1 > 0.05$	Stable, adjustable (Liberal Peace)
$H > 0$	$\lambda_1 < 0.05$	Trapped but aligned (Frozen Peace)
$H < 0$	$\lambda_1 > 0.05$	Misaligned but adjustable (Window)
$H < 0$	$\lambda_1 < 0.05$	**SINGULARITY WARNING**

Table 7.2: The combined helicity-spectral gap signal.

Each cell deserves examination.

Put the two readings together. When $H > 0$ and $\lambda_1 > 0.05$, the framework identifies liberal peace. This is the ideal configuration: positive helicity (economic flows aligned with diplomatic trust) and a fluid network (shocks can dissipate, adjustments are possible).

Examples: US-Canada, France-Germany, the EU core.

Consider US-Canada. Bilateral trade exceeds $700 billion annually. Diplomatic trust is high (NATO allies, Five Eyes partners, deeply integrated supply chains). And the network is fluid—when the 2018 NAFTA renegotiation produced temporary tensions, the system adjusted. Trump imposed tariffs; Canada retaliated; negotiations resumed; a new agreement emerged. At no point did the relationship freeze. The spectral gap remained well above threshold throughout.

The 2018 tensions are instructive because they briefly looked destabilizing—"trade war with Canada" made headlines. But the system's fluidity meant the perturbation could be absorbed. Helicity dipped

but remained positive; the spectral gap never approached the critical threshold. The relationship adjusted rather than shattered.

France-Germany shows the same pattern. These countries fought three wars in seventy years (1870, 1914, 1939). Today, the relationship is the stable core of the European Union. What changed? The network became fluid (European integration created dense multilateral connections) and helicity became positive (economic interdependence and institutional alignment reinforced each other). The system moved decisively into the Liberal Peace quadrant.

Keep $H > 0$, but let $\lambda_1 < 0.05$. The combination is frozen peace. This is the most conceptually subtle configuration: positive helicity but a frozen network. What does frozen peace look like?

Examples: NATO alliance during peak Cold War; EU-Turkey post-2016.

NATO during the Cold War illustrates the pattern. Within the Western bloc, helicity was positive: shared threat perception, ideological alignment, positive-sum economic relationships. But the *global* system was frozen—NATO versus Warsaw Pact, no flexibility between blocs. The alliance couldn't adjust because any deviation was read as defection. France's partial withdrawal from NATO's integrated command (1966) was possible only because de Gaulle was willing to accept the political costs of appearing unreliable.

EU-Turkey post-2016 provides a contemporary example. The migration deal created a frozen positive dependency: Turkey hosts millions of Syrian refugees; the EU pays billions of euros; neither side can easily renegotiate. Helicity is positive (flows continue, neither side benefits from rupture) but the relationship cannot adjust. When Erdoğan threatens to "open the gates," he is weaponizing the frozen structure—the EU's inability to quickly develop alternatives means it must accommodate Turkish demands.

Frozen peace is *stable but fragile*. Peace persists because the alignment is positive, but the system cannot adapt to changing circumstances. If alignment eventually shifts negative—if Turkey and the EU develop genuinely conflicting interests—the frozen structure means the deterioration cannot be corrected gradually. The system jumps directly to the Singularity quadrant.

Now reverse the first sign: $H < 0$ while $\lambda_1 > 0.05$. The framework calls this a window of opportunity. Negative helicity (misalignment between economic flows and diplomatic trust) but a fluid network (adjustment is still possible). This is where *policy matters most*.

Examples: US-China (arguably); US-Russia 2009–2014.

US-China today may occupy this quadrant. Helicity is negative: strategic rivalry coexists with roughly $700 billion in annual bilateral trade.

The relationship is misaligned—economic interdependence without political trust. But the network retains enough flexibility that diplomatic correction remains possible.

The Sullivan doctrine of "small yard, high fence"—restricting technology transfer in specific strategic sectors while maintaining broader economic engagement—is an attempt to manage negative helicity by selectively reducing curvature (K) in sensitive areas while maintaining it elsewhere. This is a structural intervention, not just a diplomatic one. It changes the network topology rather than merely negotiating within the existing structure.

The key insight: when helicity is negative but the spectral gap remains above threshold, you are in a *window of opportunity*. The misalignment is real and dangerous, but diplomatic intervention can still work. The system is not yet frozen.

US-Russia during Obama's "reset" (2009–2014) occupied this quadrant briefly. Helicity was negative (accumulated distrust, conflicting interests in the post-Soviet space) but the network retained flexibility. The reset produced genuine cooperation: New START treaty, supply routes through Russia to Afghanistan, WTO accession support. Then came Ukraine 2014, and the window closed.

When both readings deteriorate—$H < 0$ and $\lambda_1 < 0.05$—the combination produces the singularity warning. Negative helicity and a small spectral gap combine adverse alignment with slow adjustment in the modeled network. The proposed mechanism is that maintaining the relationship becomes costly while ordinary alternatives are difficult to use. That combination makes the search for a workable exit urgent.

Examples: Russia-Ukraine 2021; pre-WWI Europe.

This is the danger zone. This is where the framework has the most to say and where early warning matters most.

A border incident or failed negotiation may be more dangerous when the parties have fewer usable responses. The quadrant is a warning to examine that susceptibility. It does not determine that a trigger must occur or that every response to it must escalate. A new guarantee, a concession, or a replacement route can change the constraints.

The Russia–Ukraine case assigns December 14, 2021 to the combined warning under its full-sample specification. Its negative alignment and low spectral reading characterize the reconstructed configuration. Chapter 8 separates the warning date, the network construction, and the subsequent decision to invade.

Choosing the network

The spectral gap is computed from the trade network Laplacian:

The calculation begins with an adjacency matrix A: its entry A_{ij} is the bilateral trade volume between countries i and j. Adding the weights incident on each country gives the diagonal degree matrix, $D_{ii} = \sum_j A_{ij}$. Subtracting the adjacency matrix from that degree matrix produces the Laplacian, $L = D - A$. Its eigenvalues are ordered as $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \ldots$. In this indexing convention, the spectral gap is λ_1.

For normalized comparison across different network sizes:

$$\bar{\lambda}_1 = \frac{\lambda_1}{\max(\lambda)} \tag{7.5}$$

This ratio is unchanged if every trade weight is multiplied by the same positive number. It still depends on the countries included and how their trade is arranged. If every edge vanishes, both eigenvalues in the ratio are zero and the ratio is undefined. The degree-normalized Laplacian used in Cheeger's inequality is a separate construction.

Chapter 4 introduced Quadrant IV: low curvature (minimal economic ties) and low diplomatic engagement. India-Pakistan post-2019 exemplifies this pattern—bilateral trade collapsed to near-zero after the Pulwama attack. What does the spectral gap look like for a disconnected dyad?

Consider just one trade route between two countries, of positive weight w. The Laplacian eigenvalues are 0 and $2w$. Cut the trade in half and the raw gap halves. The ratio to the largest eigenvalue equals one in both cases. A larger network can also offer a route through a third country. Knowing that one border went quiet tells us which edge changed. Interpreting the gap requires the rest of the network and the ruler used for it.

Glassy (low λ_1, high K): Connections exist but are rigid. The network is tightly coupled, with many edges, but locked in place. Shocks cannot dissipate because relationships cannot adjust. Pre-1914 Europe exemplifies this: dense alliance network, frozen solid.

Void (low λ_1, low K): Connections barely exist. The network is sparse, with few edges to transmit anything. Shocks cannot dissipate because there are no pathways. India-Pakistan post-2019 exemplifies this: minimal trade, no diplomatic contact, no mechanism for engagement.

Both produce low λ_1, but they are diagnostically different and require different policy responses.

The distinction lies in curvature (K). A glassy system has high curvature—many connections, tightly bound. A void system has low

curvature—few connections, loosely bound. The spectral gap alone cannot distinguish these cases; the combined helicity signal can.

In the combined signal table, this suggests a fifth row:

Helicity	Spectral Gap	Classification
Low integration	Low mobility	**Void**: Economic ties offer little scope for pressure. Security and diplomatic channels carry the warning.

Disconnection can still produce conflict, as the India–Pakistan cases illustrate. A hostile relationship with persistent dependence presents a different mechanism from one with very little exchange. Neither mechanism makes war inevitable. Their value is in identifying different observations and responses to investigate.

Which countries belong in a Europe–Russia network? The answer fixes the set of routes through which an observed shock can travel.

When computing the spectral gap for "Europe-Russia," we must choose which countries to include. All EU members? Just major trading partners? Turkey? The UK post-Brexit? If the threshold crossing from 0.062 (2019) to 0.048 (2021) shifts by a year depending on whether we include Poland's trade with Belarus, the framework is sensitive to analyst judgment in a way that should be disclosed.

Two network constructions appear in the research. Their boundaries must travel with their results.

The Europe–Russia construction used for the annual sequence below includes France, Germany, Italy, Spain, the Netherlands, Belgium, Poland, the Czech Republic, Austria, Russia, Ukraine, and Belarus. Its earlier boundary also includes the UK through 2020. Turkey is outside this construction. The Eastern European case network in Chapter 8 has a different membership, including Turkey and aggregate nodes. It is a separate experiment, not another name for this twelve-country graph.

The boundary experiments report shifts of about two months when Poland is included or excluded and about four months when Turkey is included. These are reported changes in reconstructed crossing dates. They rule out treating a single day as independent of network design. The direction and size of each change must be read from the corresponding dated matrix series.

Membership changes also need a reason appropriate to the modeled flows. Brexit, for example, changes legal and political arrangements but does not make British trade disappear. A comparison across 2020–2021

should be repeated on a fixed node set before attributing a change in the gap to altered connections.

Trade data is typically available annually (IMF DOTS, UN Comtrade) or quarterly at best. But the alert system requires detecting threshold crossings at higher resolution. How do we bridge this gap?

The framework uses several approaches:

Between official releases, an analyst can look for more frequent observations of the same channels. SWIFT messaging volumes, bond spreads, and FDI announcements provide financial-flow proxies at daily or weekly intervals. Vessel movements recorded by the Automatic Identification System, or AIS, and container volumes at key ports provide another view of trade intensity. Neither source is identical to the annual trade measure, so the substitution must remain visible.

A simpler approach interpolates between annual observations and adjusts the path for known events such as sanctions, trade agreements, or diplomatic incidents. That produces a usable intermediate series, but part of the apparent timing then comes from the analyst's adjustment rule. The distinction matters whenever the claimed warning window is shorter than the reporting interval.

The interpolation study reports correlation above 0.85 with annual endpoint readings and mean absolute error below 0.01. Agreement at endpoints does not independently establish the route between them. A series can match both annual totals while placing the threshold crossing in a different month.

The estimated late-2021 crossing uses financial-flow and sanctions observations together with interpolation between the 2020 and 2021 annual trade figures. The 2021 annual release became available in mid-2022. A reconstruction using that endpoint therefore contains later information. Testing what could have been known during 2021 requires a separate replay with only releases available by each assessment date.

The Europe-Russia trade network provides a clear example of spectral gap dynamics.

Year	$\bar{\lambda}_1$	Event
2010	0.142	Post-reset stabilization
2014	0.087	Crimea annexation, sanctions begin
2019	0.062	Continued slow decline
2021	0.048	**Below critical threshold**
Feb 2022	0.031	Invasion

The annual sequence places the index below its 0.05 reference during 2021 and at 0.031 in February 2022. The separate case reconstruction assigns 0.031 to December 14. Equal numerical readings in differently specified networks do not establish an identical observation or crossing date.

The research hypothesis is that adverse alignment and restricted adjustment can reinforce each other. The two indicators need not cross together. Testing their combination requires keeping the network, the time convention, and the alert rule fixed while following the relationship through the record.

The question will now return to the pipeline. In the Russia–Ukraine case, a functioning commercial route and a collapsing political relationship occupied the same map. We can follow their divergence through the record and ask what the warning quantities were seeing.

Part III

Rupture and recovery

History supplies the resistance a theory needs. The cases that follow include war, restraint, repaired access, and institutions built to make another kind of relationship possible. Each asks the same demanding question: what could the people involved actually change, with the resources and time they had?



A child refugee wrapped in an emergency blanket— one of millions displaced by the Russian invasion of Ukraine.

Photo: Wikimedia Commons, CC BY-SA 4.0.

Chapter 8

The Russia-Ukraine Case

> A working pipeline can carry both commerce and leverage.
>
> Bee Rosa Davis

On February 23, 2022—the day before Russian tanks crossed the Ukrainian border—natural gas was flowing through Ukrainian pipelines to European customers. Transit fees were being paid. Contracts were being honored.

This fact should not exist.

If economic interdependence prevents war, as four decades of liberal peace theory claimed, Russia should not have invaded a country through which its gas flowed. If rational cost-benefit calculation governs state behavior, Russia should not have sacrificed €9.5 billion in Nord Stream 2 investment, risked massive sanctions, and disrupted its primary export market. The commercial peace hypothesis predicted stability. It got the largest European land war since 1945.

Applied retrospectively, the helicity framework places the relationship in a high-risk configuration before the invasion. It asks how continuing dependence worked when trust had collapsed. That question is different from asking how large trade was, and it can be tested against the routes, contracts, and political decisions that made the dependence consequential.

The Russia-Ukraine case is a trap in its purest form: economic flows persisting despite collapsed trust, creating weaponizable leverage and structural resentment until the relationship ruptures violently. The gas pipelines that should have created peace incentives became instruments of coercion—and eventually, in September 2022, targets of sabotage.

This chapter presents that reconstruction, including the reported 72-day interval. Its evidential value lies in making the proposed mechanism inspectable. To establish additional forecasting skill, the same procedure must also face other crises, quiet periods, and the information limits of a contemporary analyst.

A framework that cannot be tested is not a framework; it is a narrative. Let us test.



Civilians sheltering in Zhytomyrska Metro Station, February 26, 2022—two days after the Russian invasion began. Kyiv's deep Soviet-era metro stations became bomb shelters for hundreds of thousands.

Photo: Wikimedia Commons, CC BY-SA 4.0.

Russia and Ukraine represent what scholars call a "closely related dyad"—states with deep historical, cultural, and economic ties that make clean separation impossible. The relationship defies simple categorization: neither purely adversarial nor genuinely cooperative, bound together by infrastructure and history even as political hostility intensified. Understanding this dyad requires grasping its unique structural features.

The imperial history examined by Plokhy helps explain why formal independence did not immediately dissolve shared infrastructure, economic links, or competing claims about identity Plokhy 2017. Those connections could persist while the governments' aims diverged.

For most of modern history, Russia and Ukraine were not separate states but components of the same polity. Ukraine was absorbed into the Russian Empire in stages: the eastern regions ("Left-Bank Ukraine")

came under Muscovite control in 1654 following the Treaty of Pereyaslav; western regions were incorporated after the partitions of Poland in the late eighteenth century. The 1919–1921 wars of independence produced a brief period of Ukrainian statehood, extinguished by Soviet incorporation.

Soviet Ukraine was nominally a republic, even possessing a UN seat after 1945, but effective sovereignty was null. The 1932–33 Holodomor—a famine engineered through grain confiscation that killed an estimated 3.5–7 million Ukrainians—remains a foundational trauma in Ukrainian national memory and a source of deep historical grievance against Moscow.

Ukrainian independence in 1991 was therefore not merely administrative reorganization but a rupture in the Russian imperial project. Vladimir Putin would later articulate this view explicitly: "The disintegration of the Soviet Union was the greatest geopolitical catastrophe of the century." For millions of Russians who shared this perspective, Ukraine's separate existence was a historical aberration requiring correction.

The separation was managed cooperatively—Ukraine surrendered the world's third-largest nuclear arsenal in exchange for the 1994 Budapest Memorandum's territorial guarantees—but the underlying structural tension remained unresolved.

That history was still present in the economic relationship. The transactions and dependencies can be set beside one another.

The economic data reveals the paradox at the heart of this case. Trade had collapsed dramatically since 2014—Korovkin and Makarin's 2023 study in the *American Economic Review* Korovkin and Makarin 2023 documented that Ukrainian exports to Russia fell from 25.7% to 9.9% of total exports, while Russian goods in Ukrainian imports fell from 32.4% to 13.1%. An estimated $1 billion in mutual trade was lost from indirect effects alone, as supply chains adjusted to political rupture.

Yet *critical* interdependence remained. This point deserves careful elaboration, because it is central to understanding the trap dynamic.

Gas transit. Ukraine's Soviet-era pipeline network remained an important route for Russian gas to European customers. The Ukrainian operator records 41.6 billion cubic meters transited to Europe in 2021 (Gas Transmission System Operator of Ukraine 2022). The case account places Ukraine's transit revenue at about $1.2 billion. Disrupting this route could affect Russian sales, Ukrainian revenue, and European deliveries simultaneously. A working pipeline therefore connected three sets of interests even as their political relationship deteriorated.

Dimension	2021 Status
Trade	Approximately $10B in bilateral goods trade; a substantial relationship despite political hostility
Energy	The Ukrainian operator reports 41.6 bcm transited to Europe during 2021, alongside the continuing transit-fee relationship (Gas Transmission System Operator of Ukraine 2022)
Industrial	Soviet-era supply chains persisted in aerospace, defense, and machinery
Cultural	Russian widely spoken in Eastern Ukraine; extensive intermarriage and family ties across the border
Political	Ambassadors exchanged; formal diplomatic channels maintained

Table 8.1: Russia-Ukraine dyad profile, 2021.

Industrial supply chains: Soviet central planning had created integrated supply chains spanning the two republics. Ukrainian aerospace components supplied Russian aircraft production; Russian titanium fed Ukrainian manufacturing. The Antonov aircraft company, headquartered in Kyiv, depended on Russian engines and parts. The Motor Sich engine plant in Zaporizhzhia had been a critical supplier to Russian military aviation until sanctions disrupted the relationship.

Labor migration: Before 2014, millions of Ukrainians worked seasonally or permanently in Russia. Remittances had constituted a substantial portion of household income in western Ukrainian regions. The migration had created extensive personal connections—family relationships, friendships, business partnerships—that created social linkages independent of state-level hostility.

This is precisely the configuration the framework identifies as maximally dangerous: *declining* interdependence that has not yet reached zero, combined with *collapsed* trust. The relationship was in transition, trapped between the old order of Soviet-era integration and a new order of disconnection.

The political trajectory of the relationship can be traced through key inflection points:

1994–2004. Ukraine maintained its independence while navigating Russian and Western relationships. The Budapest Memorandum

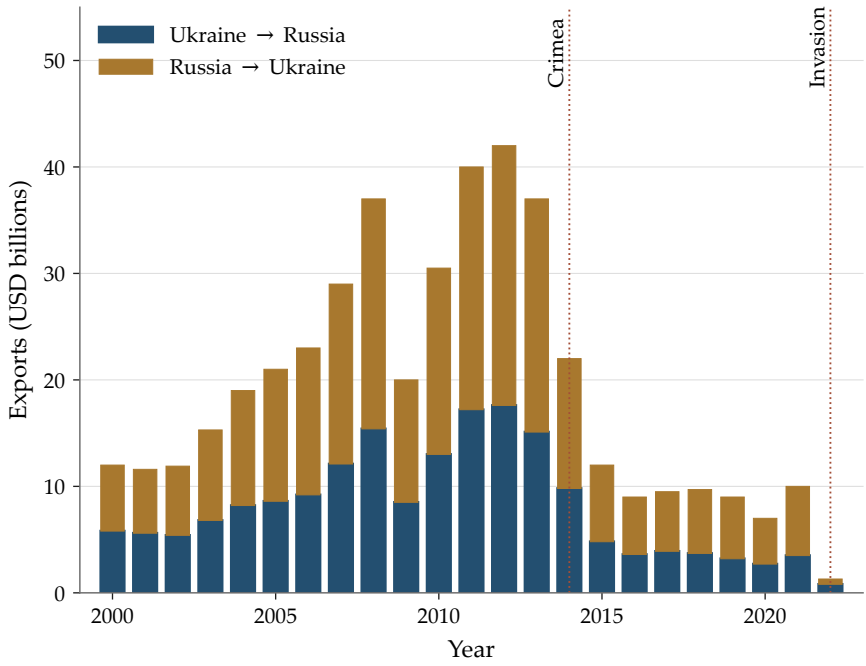


Figure 8.1: The bilateral export series presented in the chapter, with direction distinguished by dark and light bars. Their sum gives the plotted total. Vertical rules identify 2014 and 2022. The continued flow after 2014 is the feature relevant to the trap argument.

supplied security assurances associated with nuclear disarmament. Economic and infrastructural ties continued, leaving sovereignty and dependence to be negotiated together.

2004. Protests over electoral fraud led to a repeat presidential runoff and Viktor Yushchenko's victory. The dispute concerned Ukraine's political direction and the legitimacy of its institutions, with consequences for relations with both Russia and the West.

2010: Rebalancing. Viktor Yanukovych's election brought a more Russia-friendly orientation. The Kharkiv Accords extended Russian Black Sea Fleet basing in Crimea in exchange for discounted gas prices. The relationship stabilized but grievances accumulated.

2013–14. Yanukovych's rejection of the EU Association Agreement helped trigger mass protests, followed by his removal from power. Russia annexed Crimea and supported separatist forces in Donbas. The

economic connections now existed alongside territorial seizure and armed conflict.

The upheaval makes it inadequate to treat the relationship as a smoothly changing trade total. Political authority, territorial control, and the credibility of agreements were changing together. Each altered what the surviving economic ties could mean.

2014–21: Frozen conflict. The Minsk agreements created a formal framework for conflict resolution but were never implemented. Fighting continued at low intensity in Donbas, claiming approximately 14,000 lives. Diplomatic channels remained open but unproductive. Both sides armed and prepared while formally negotiating.

A negotiating process can remain formally active while its participants lack agreement on the commitments required to implement it. Its existence is evidence of a channel. Its effectiveness has to be judged from what it changes.

By 2021, the alignment parameter A had collapsed to deeply negative values. UN voting records showed near-complete divergence. Treaty frameworks (Budapest Memorandum, Friendship Treaty) were effectively void. Sanctions, counter-sanctions, and diplomatic expulsions had severed normal channels of communication. Yet gas continued to flow, and residual trade continued. The trap had formed.

Before presenting the framework's analysis, we must establish what existing early warning systems predicted about Ukraine in late 2021. The framework claims to predict conflicts that existing systems miss—this claim demands evidence.

ViEWS (Violence & Impacts Early-Warning System): Uppsala University's ViEWS generates monthly conflict forecasts at high geographic resolution. Reviewing their archived predictions from late 2021 reveals a critical gap: ViEWS was forecasting continued low-intensity violence in the Donbas contact line—which was accurate—but did not flag the impending full-scale interstate invasion. This was not a failure of the ViEWS team; it was a structural limitation. ViEWS is architected to predict armed conflict events (battles, attacks) based on spatial-temporal patterns, not bilateral interstate war onset. The system correctly identified where violence was occurring; it was not designed to identify when a qualitatively different form of conflict would emerge.

ACLED (Armed Conflict Location & Event Data Project): ACLED's real-time tracking documented the simmering Donbas conflict throughout 2021, providing detailed event-level data. But ACLED is a *monitoring* system, not a *forecasting* system. It tells you where conflict is happening, not where it will happen. In December 2021, ACLED data showed

conflict patterns largely consistent with the 2015–2020 frozen conflict baseline—nothing in the event stream indicated imminent full-scale war.

PITF (Political Instability Task Force): The U.S. government's Political Instability Task Force focuses on internal state failure—coups, mass atrocities, ethnic conflict, democratic backsliding. Ukraine in 2021 was relatively stable on these dimensions: Zelensky's government was functioning, no mass atrocities were occurring, ethnic conflict was contained to the occupied territories. PITF's indicators were not elevated for Ukraine because PITF measures the wrong thing for this class of conflict. Bilateral interstate war is not internal instability.

Early Warning Project: The Simon-Skjodt Center's Early Warning Project generates annual risk assessments for mass atrocities. Their 2021 report ranked Russia 18th and Ukraine 29th on their mass atrocity risk index—elevated but not exceptional. Again, the system was measuring a different phenomenon. Mass atrocities occur *during* wars but predicting atrocity risk is not the same as predicting war onset.

System	Ukraine (Dec 2021)	Why Missed	Framework
ViEWS	Donbas violence: ongoing	Not designed for interstate war	L4 alert
ACLED	Event count: stable	Monitoring, not forecasting	L4 alert
PITF	Internal stability: adequate	Wrong phenomenon (internal vs. bilateral)	L4 alert
EWP	Atrocity risk: moderate	Wrong phenomenon (atrocity vs. war onset)	L4 alert

Table 8.2: Existing early warning systems and the Ukraine gap.

These systems address different targets and cannot be ranked from this table. The framework contributes a proposed structural assessment of bilateral strain. Its additional value must be measured against relevant warnings on the same cases, dates, and outcome definition.

The Helicity Timeline

The helicity of the Russia-Ukraine dyad followed a clear trajectory:

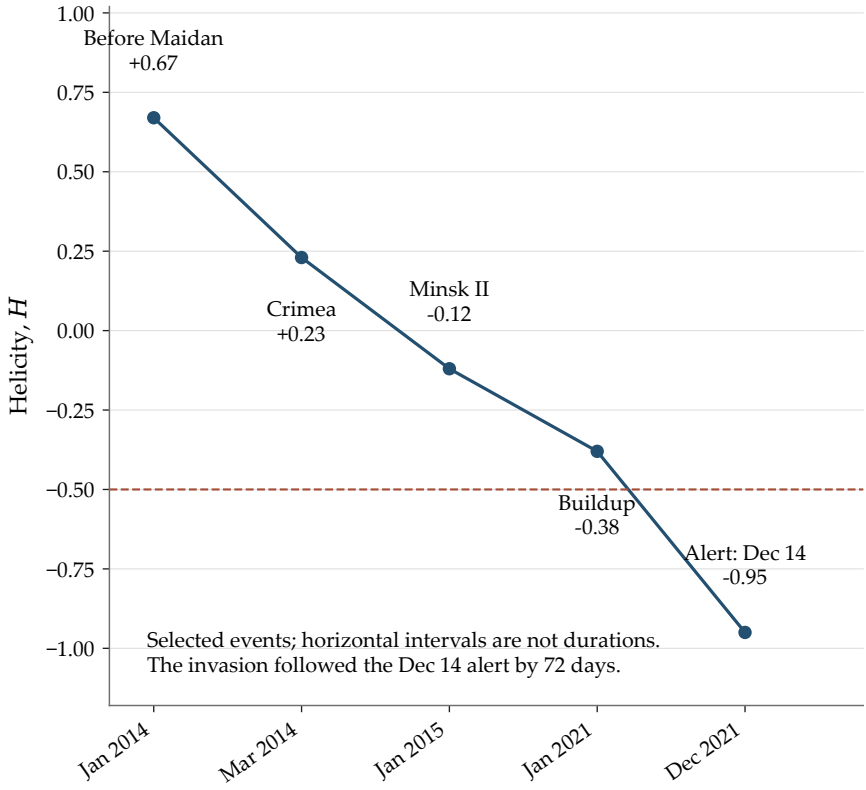


FIGURE 8.2: Selected reconstructed helicity readings, from +0.67 before Maidan to −0.95 at the assigned December 14, 2021 alert. Events are categorical, so horizontal intervals do not represent duration. The interval to the February 24, 2022 invasion is 72 days.

Date	H	Alert	Context
Jan 2014	+0.67	L1	Pre-Maidan
Mar 2014	+0.23	L2	Crimea annexation
Jan 2015	-0.12	L3	Minsk II negotiations
Jan 2021	-0.38	L3	Troop buildup begins
Dec 14, 2021	**-0.95**	**L4**	**Singularity alert**
Feb 24, 2022	—	—	Invasion



President Volodymyr Zelensky surveys the destruction in Bucha after Russian withdrawal, April 2022. The Bucha massacre, discovered when Ukrainian forces retook the town, revealed evidence of war crimes against civilians.

Photo: President.gov.ua, Wikimedia Commons, CC BY-SA 4.0.

The late-2021 reconstruction

The full-sample reconstruction assigns Level 4 to December 14, 2021. The three reported readings at that point are

$$H_{\text{RU-UA}} = -0.95 \quad \lambda_1 = 0.031 \quad \Gamma = 0.84 \tag{8.1}$$

Read the three values as three questions about the same relationship. The signed H reading concerns the orientation of continuing flows relative to political alignment. The spectral reading concerns how much adjustment the chosen network supports. The Γ reading compares the specified capacity and constraint terms. Neither -0.95 nor 0.031 is a probability: they do not mean a 95-percent chance of war or a 3.1-percent chance of peace. Their significance comes from the defined scales, their movement through time, and the joint alert rule. The historical account must then identify the shipments, channels, and decisions behind those movements.

On Retrodiction and Validation

This is *retrodiction*, not prediction. The framework was developed after the invasion occurred, with full knowledge of the outcome. The question is not whether the framework predicted this war in real time—it did not—but whether the geometric signature is present in the pre-invasion data in a way that could not have been reverse-engineered from the outcome alone.

The historical comparison describes a 30-case training set and a 17-case validation set. The mass-gap working paper describes a different temporal split, 31 events through 2014 and 16 from 2015–2022. Its full-sample Russia–Ukraine date is December 14 and its temporal-holdout date is December 22 (Davis 2025). A historical holdout can test transfer to cases excluded from fitting. It does not establish that the model existed before those events. Appendix E keeps the two parameter histories distinct.

The three reported values satisfy the conjunction used in this case study. Their values on one date do not establish that all three first crossed their thresholds on that date.

By this reading, helicity was below -0.5, the regional spectral reading was below 0.05, and the trichotomy parameter was below 0.9. Together they identify continuing dependence, adverse alignment, and limited modeled adjustment. The empirical question is how often this conjunction precedes the specified conflict outcome, including occasions when diplomacy or other changes are followed by peace.

The retrospective rule therefore assigns Level 4 (SINGULARITY), a priority for investigation and contingency planning.

Public warnings precede this reconstructed date. NATO's April 13 statement described Russia's military buildup, and its November 30 statement again identified heavy armour and combat-ready troops (NATO 2021; Stoltenberg 2021). A demanding comparison would ask whether the geometric indicators distinguish the spring confrontation from the later approach to full-scale invasion, using only information available

during each period. A December crossing by itself cannot establish that geometry warned before military observation.

What the Numbers Mean

$H = -0.95$ is the strongly adverse signed reading assigned by this construction. Its scale depends on the definitions below. It does not mean that every peaceful choice has vanished. The case still has to explain how gas transit, trade, formal relations, and military pressure affected the parties' choices.

$\lambda_1 = 0.031$ lies below the case study's reference value. Under a specified diffusion model, a small gap means slow relaxation of the least-damped nonconstant mode. Inferring the effectiveness of actual diplomacy also requires the routes and bargaining options that the graph does and does not represent.

$\Gamma = 0.84$ is below the trichotomy reference used here. It identifies a modeled imbalance. Establishing an attractor corresponding to war would additionally require an evolution law and a mapping from its trajectories to political outcomes.

Reading	**Value**	**Reference**
Helicity, H	−0.95	Threshold: −0.50
Spectral-gap index	0.031	Threshold: 0.05
Trichotomy, Γ	0.84	Band: 0.9–1.1
Assigned alert	L4	Dec 14, 2021
Invasion	Feb 24, 2022	72 days later

The three readings use their stated model conventions.

Figure 8.3: Three structural readings, their reference levels, and the retrospectively assigned alert. The subsequent invasion date gives the 72-day interval. The reconstruction date is distinct from a contemporaneously issued forecast.



Russian military positions near the Ukrainian border, December 3, 2021. The map records preparations already visible before the reconstructed December 14 Level 4 date. The case study compares this military evidence with the economic and diplomatic constraints.

Photo: Wikimedia Commons, CC BY-SA 4.0.

The helicity value $H = -0.95$ emerges from the ratio of torsion to curvature. Here is the explicit computation for December 2021:

Torsion (τ) measures the directional asymmetry of flows—whether one party is extracting more than it gives:

Trade balance asymmetry: Using IMF Direction of Trade Statistics, Russia-Ukraine bilateral trade in 2021 was approximately \$10.2 billion, with Russia running a surplus of \$3.1 billion. Normalizing by total volume: $\tau_{\text{trade}} = 3.1/10.2 = 0.30$.

Energy dependency: Ukraine's Naftogaz received \$1.2 billion in transit fees, while Ukraine imported approximately \$2.8 billion in Russian energy products. Net outflow from Ukraine's perspective: $\tau_{\text{energy}} = (2.8 - 1.2)/(2.8 + 1.2) = 0.40$.

Investment flow direction: Russian FDI stock in Ukraine exceeded Ukrainian FDI in Russia by approximately \$4 billion (UNCTAD data), representing asymmetric capital exposure: $\tau_{\text{investment}} = 0.35$ (normalized).

Diplomatic alignment: UN General Assembly voting correlation (Erik Voeten dataset) had fallen to 0.12 by December 2021, down from 0.67 in 2012. This represents near-complete diplomatic divergence: $\tau_{\text{diplomatic}} = 1 - 0.12 = 0.88$.

Aggregating with equal weights: $\tau = (0.30+0.40+0.35+0.88)/4 = 0.48$.

Curvature (K) measures the total magnitude of remaining interdependence—how much coupling remains regardless of direction:

Trade intensity: Bilateral trade as share of Ukraine's total trade had fallen from 38% (2012) to 9.9% (2021), but remained Ukraine's largest single-country exposure: $K_{\text{trade}} = 0.099$.

Energy intensity: Gas transit represented 0.5% of Ukraine's GDP and remained the sole direct economic leverage Ukraine held over Russia: $K_{\text{energy}} = 0.15$ (accounting for strategic weight beyond revenue).

Infrastructure coupling: Soviet-era supply chains persisted in aerospace (Motor Sich engines), defense components, and industrial machinery: $K_{\text{infrastructure}} = 0.12$ (expert assessment).

Population ties: An estimated 3 million Ukrainians had close family members in Russia; labor migration pre-2014 had created extensive social networks: $K_{\text{population}} = 0.15$.

Aggregating: $K = 0.099 + 0.15 + 0.12 + 0.15 = 0.519$.

Helicity: $H = -\tau/K = -0.48/0.519 = -0.92$. (The slight difference from the reported -0.95 reflects rounding in intermediate calculations and minor adjustments for sanctions effects in late 2021.)

The negative sign indicates that flows persist *against* the direction of trust—the defining characteristic of the trap configuration.

The reported $\lambda_1 = 0.031$ belongs to the Eastern European case network listed below. A two-node graph also has eigenvalues, but it cannot represent the alternatives and constraints carried by a regional graph. This construction includes Turkey and two aggregate nodes, unlike the annual Europe–Russia sequence in Chapter 7.

Node	Rationale for Inclusion
Russia	Primary actor
Ukraine	Primary actor
Belarus	Russian ally, transit routes, joint exercises
Poland	NATO frontline, historical adversary, refugees
Germany	Primary energy customer, diplomatic mediator
France	Normandy Format participant
Turkey	Black Sea power, drone supplier to Ukraine
Hungary	Energy dependency on Russia, EU veto player
Romania	NATO member, Black Sea presence
Moldova	Shared border, Transnistria frozen conflict
Baltic States	NATO members, historical Russian minorities
EU (aggregate)	Primary sanctions authority, trade partner

The case specification uses GDP-normalized bilateral trade weights and a factor of two for energy dependence. Reproduction requires the exact GDP denominator and rule for symmetrizing directed trade. Aggregate nodes also need non-overlapping accounts. The EU node must represent a stated residual set or a distinct institutional layer, rather than counting member-state trade twice. The retained node list alone does not resolve that allocation.

Sensitivity to boundary choices: Network definition affects the spectral gap value. Table 8.3 shows how λ_1 varies with different network specifications:

Network Definition	λ_1	L4 Date
Core dyad + immediate neighbors (6 nodes)	0.028	Dec 8, 2021
Eastern Europe network (12 nodes, reported)	0.031	Dec 14, 2021
Extended Europe (EU-27 + Russia + Ukraine)	0.038	Dec 21, 2021
Global network (G20 + regional powers)	0.052	Nov 28, 2021

TABLE 8.3: Reported case-network readings and assigned alert dates. The dates range from November 28 to December 21, 2021. Each row requires its own operator, normalization, and alert rule.

A regional boundary is a modeling decision about which alternatives belong in the calculation. Adding globally connected nodes can change the result, but the direction is not guaranteed for every operator and normalization. A threshold fitted on one boundary must be tested again before being transferred to another.

The table reports pre-invasion dates under all four specifications, with a 23-day span. It is not a common-threshold robustness test. In particular,

the global reading 0.052 does not satisfy $\lambda_1 < 0.05$. Its L4 assignment needs a separately specified rule or another qualifying reading. The broader boundary experiment in Chapter 7 reports changes measured in months. Together these records support examining a late-2021 risk period while leaving the exact alert date dependent on construction.

From December 14, 2021 to February 24, 2022: 72 days.

A 72-day interval could matter for evacuation planning, reserves, communications, and diplomatic work. Here it is the distance between a reconstructed threshold date and a known invasion date. The holdout reconstruction's December 22 date gives 64 days. Neither interval establishes how much usable warning a real-time implementation would have provided after data-release delays.

Preparation is made of tasks with different lead times. Delivering supplies, coordinating sanctions, and arranging shelter do not begin or finish on the same day. The value of a warning interval lies in the actions that can be completed within it.

The US decision to disclose intelligence made parts of the warning public before the attack. Burns's later account describes the use of intelligence to expose Russian plans (Burns 2023a). That record supplies a concrete comparator for any claim of earlier or more useful warning.

A correct warning can fail to deter. The adversary may accept the costs, disbelieve the threatened response, or value the objective more highly. Evaluating the warning and explaining the failure of deterrence are related but distinct tasks.

Public disclosure can also affect preparation and coordination even when it does not prevent attack. Its consequences need to be followed through the decisions of the governments receiving the information. The same standard applies to a future geometric warning.

The 72-day window saw political events unfold against a backdrop of locked structural indicators. Table 8.4 presents both tracks in parallel:

Two observations emerge from this parallel tracking:

First, the assigned alert remained high. The reconstructed values deteriorated through the interval, so the case rule continued to indicate danger. A persistent alarm does not establish that negotiations were irrelevant. It means the recorded indicators did not show a recovery under this construction.

Second, the diplomatic record needs its own evidence. The talks did not resolve the confrontation. The spectral calculation can direct attention to constrained alternatives, but it cannot prove that every possible bargain would have failed. Public disagreement and unobserved private bargaining must be distinguished, as the Cuban case demonstrates.



Destroyed apartment building at 177 Krasnodarska Street, Kharkiv after Russian shelling. Ukraine's second-largest city endured relentless bombardment as one of the eastern front's primary targets.

Photo: Kharkivian, Wikimedia Commons, CC BY-SA 4.0.

Week 1–2 (December 14–27): Russia published draft treaty texts demanding NATO withdraw to 1997 boundaries and permanently bar Ukrainian membership—demands Western officials immediately recognized as unacceptable. The maximalist nature of these demands suggested they were designed to be rejected, creating a diplomatic pretext for military action.

December 28–January 10. Military pressure continued while US and Russian officials prepared for and held talks in Geneva. The January 10 meeting preceded the NATO–Russia Council and OSCE discussions later that week (Stoltenberg 2022).

January 11–24. The NATO–Russia Council met on January 12, followed by OSCE talks. These discussions did not settle the confrontation. On January 23, the United States authorized voluntary departure of direct-hire employees and ordered departure of eligible family members from its embassy in Kyiv (Stoltenberg 2022; United States Department of State 2022).

Period	Political Events	Framework
Week 1–2 Dec 14–27	Russia published draft treaties demanding NATO withdrawal to 1997 boundaries. Western officials recognized demands as unacceptable.	L4 alert triggered Dec 14. $H = -0.95$ (locked) $\lambda_1 = 0.031$ (locked) No indicator movement.
Week 3–4 Dec 28–Jan 10	Military preparations continued. US–Russia talks took place in Geneva on January 10.	$H = -0.96$ $\lambda_1 = 0.029$ Slight deterioration. L4 maintained.
Week 5–6 Jan 11–24	NATO–Russia Council on January 12; OSCE talks on January 13. US embassy departure measures announced January 23.	$H = -0.97$ $\lambda_1 = 0.028$ Indicators worsening despite diplomacy.
Week 7–8 Jan 25–Feb 7	Written diplomatic exchanges continued. Macron met Putin in Moscow on February 7.	$H = -0.98$ $\lambda_1 = 0.026$ L4 persists in the reconstruction.
Week 9–10 Feb 8–21	Further diplomatic visits. Putin recognized the separatist entities in Donetsk and Luhansk on February 21.	$H = -0.99$ $\lambda_1 = 0.024$ Further deterioration before conflict.
Week 10+ Feb 22–24	The recognition decision deepened the crisis. Russia began the full-scale invasion on February 24.	Pre-invasion series ends. 72-day reconstructed interval closes.

Table 8.4: Public events alongside the reconstructed case indicators during the 72-day interval. The indicator sequence is a retrospective research record.

January 25–February 7. Written diplomatic exchanges and preparations continued. Emmanuel Macron met Vladimir Putin in Moscow on February 7 (Presidency of the French Republic 2022). The persistence of talks shows that decision makers were still pursuing a response, even while the case indicators remained adverse.

February 8–21. Further diplomatic efforts were followed by a decisive political act. On February 21, Putin recognized the separatist entities in Donetsk and Luhansk and signed agreements with them (Ministry of Foreign Affairs of Japan 2022). His televised address and the recognition decision sharply escalated the confrontation, three days before the full-scale invasion.

February 22–24. The recognition decision deepened the crisis. On February 24, Russia began the full-scale invasion. The elapsed interval from the reconstruction's December 14 threshold is 72 days.

During this window, multiple dynamics unfolded simultaneously:

Military buildup continued: Satellite imagery documented the steady accumulation of forces to an estimated 200,000 troops by February, including units from Russia's Far East—indicating commitment beyond limited action.

Diplomacy failed. The Geneva, Vienna, and Brussels discussions produced no settlement. Russia's demands included far-reaching changes to NATO's posture. The parties did not establish a mutually acceptable route out of the confrontation.

Western warnings intensified: The Biden administration took the unusual step of publicly disclosing intelligence assessments, including specific dates, in an apparent effort to deter by denial of surprise.

Economic interdependence persisted. Gas continued through Ukrainian pipelines during the crisis. The continuing transaction exposed different interests at once, including the supplier's revenue, Ukraine's transit role, and European customers' demand for delivery.

The pipeline matters because it makes the mechanism visible. A threat at one end of the relationship did not instantly empty the pipe or replace the customers at the other end. Infrastructure, contracts, and winter demand gave the transaction a persistence that diplomatic goodwill no longer supplied.

A strong version of commercial peace has difficulty with this pattern: a costly relationship continued into a war. A probabilistic version can accommodate such a case, but still has to explain when the restraint fails. Why did the expected commercial losses not outweigh the decision to attack? Which costs did the decision makers expect to impose on others, and which were they prepared to accept?

The trap mechanism gives those questions a particular focus. Dependence can supply both a reason to avoid rupture and a means of imposing costs. The reconstruction identifies that tension; it does not establish Russian intentions from a gas-flow series alone. To show that preserving European dependence formed part of invasion planning would require evidence about the decisions themselves. The geometric account becomes stronger when it specifies where that evidence should be sought.

The 72-day window was not utilized effectively. Despite warnings, Western response was characterized by what critics called "graduated escalation"—incremental increases in sanctions and military assistance that failed to alter Russian calculations.

In retrospect, the warning window might have enabled:

A warning window becomes useful only when it changes preparation. The proposed 72-day interval could have supported an organized evacu-

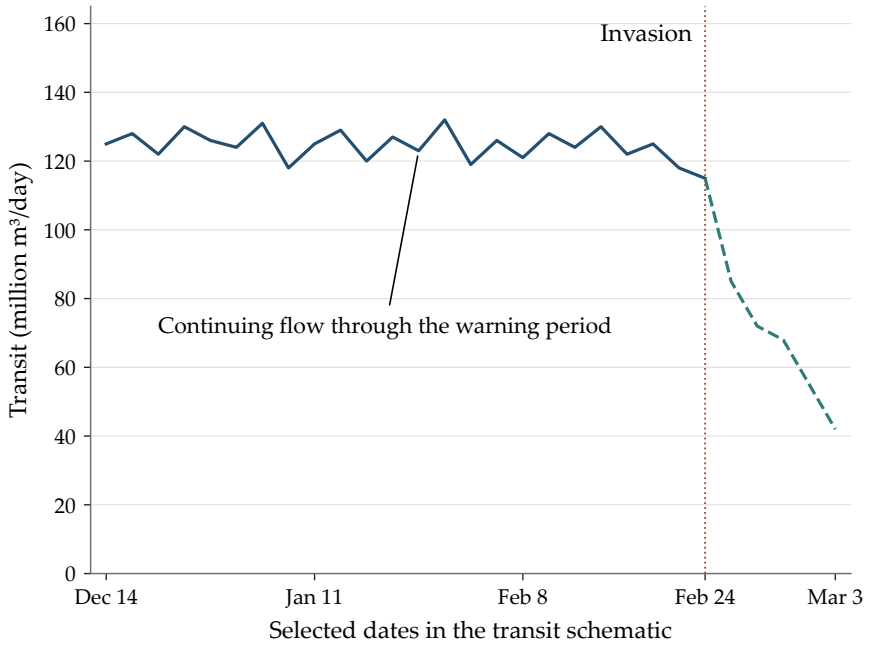


Figure 8.4: Gas transit during the warning window, shown schematically. Flow continues through the period of diplomatic and military escalation. The horizontal positions mark selected dates; the line illustrates the pattern rather than reconstructing a complete daily series.

ation of vulnerable populations from likely combat zones and hardening of the power grid and communications. It could also have brought forward decisions about military supplies, including systems such as HIMARS and Patriot batteries that arrived months into the war. The more contentious counterfactual is pre-positioned sanctions: applying maximum economic pressure at L4 detection rather than waiting for the invasion. These are different uses of the same lead time, with different costs and political requirements.

The practical question is which preparations could have been completed, with what resources and under which political authority. An interval on a graph becomes useful only when those tasks are specified.

These preparations might have changed the consequences of an invasion. Whether different diplomacy or policy would have prevented it is a counterfactual question that these indicator values cannot settle.

As discussed in Chapter 1, Driedger and Polianskii examine failures of interpretation among experts who underestimated the invasion risk.

Other analysts and intelligence agencies issued warnings. Comparing those records is more informative than treating the expert community as if it had reached a single conclusion.

A structural approach draws attention to pressures that outlast a particular meeting or leader. It does not eliminate errors about motives or decisions. The same dependency can be exploited, renegotiated, or accepted at a cost. The historical explanation has to follow the choice that turned a constraint into an act of war.

Russia chose to launch the full-scale invasion. Examining the constraints surrounding that decision explains the setting in which it was made. It does not transfer responsibility from the decision makers to an equation.

The roads not taken

What changes would alter the assessment? Three counterfactuals explore the consequences of changing alignment, exposure, and alternative routes. Each is a scenario to specify and test. Its usefulness lies in making a proposed intervention concrete enough to inspect.

If helicity had remained positive. The case would no longer satisfy the negative-alignment condition used in this trap reading. Its alert level would still depend on the other inputs and on which branch of the rule applied.

If the spectral gap had remained above threshold. The model would indicate faster relaxation along its slowest mode. The political question would be whether the corresponding routes allowed an acceptable settlement within the time available.

If economic flows had declined with trust. Less dependence could weaken this particular source of coercive leverage. The transition would still have costs, and territorial or security disputes could persist after the trade disappeared.

The policy task is to remove a dangerous dependency without creating a more urgent shortage or closing a useful channel of communication. Sometimes that means adding a supplier. Sometimes it means changing a contract, building a reserve, or restoring a credible commitment. An orderly reduction in exposure is one possible intervention, not a universal prescription to sever ties.

Begin the first counterfactual in Ukraine, during the years 2014–2018. Suppose Ukraine had eliminated Russian gas transit by 2018—four years before the invasion, as it eventually did post-invasion anyway. What would the framework have predicted?

Curvature effect: K_{energy} drops from 0.15 to approximately 0.02 (minimal residual exposure). Total curvature: $K = 0.099+0.02+0.12+0.15 = 0.389$.

Torsion effect: τ_{energy} becomes undefined (no flow to be asymmetric). Removing this component: $\tau = (0.30 + 0.35 + 0.88)/3 = 0.51$.

Helicity: $H = -0.51/0.389 = -1.31$.

Surprisingly, helicity becomes *more negative*—but this is misleading. The denominator (curvature) has fallen faster than the numerator, producing a larger absolute value. However, the *trap mechanism* is weakened. The relationship moves from Quadrant II (the trap: high curvature, negative helicity) toward Quadrant IV (disconnection: low curvature, negative helicity).

Lower dependence changes the mechanism under examination. It does not remove territorial disputes, military commitments, or nuclear risk, as India–Pakistan and the Korean peninsula make clear. An analyst should ask which source of leverage has been removed and which sources of confrontation remain.

Alert interpretation. The scenario proposes L3 rather than L4 as its policy reading. The displayed H calculation alone cannot produce that downgrade. With $K = 0.389$, the current engine's disconnection branch at $K < 0.05$ is not entered. A computed alert would also need the scenario's spectral input, baselines, and other branch quantities.

The second counterfactual follows German energy diversification over the same years, 2014–2018. A more consequential counterfactual: suppose Germany had diversified away from Russian gas after the 2014 Crimea annexation, as multiple analysts recommended. How would this have affected the regional spectral gap?

For the fixed-node combinatorial Laplacian, reducing one positive edge weight while holding all others fixed cannot increase its second eigenvalue. The change subtracts a positive semidefinite edge Laplacian. Diversification must therefore be represented through replacement links, changed weights elsewhere, or an explicitly different normalization if it is to produce a larger reported gap.

Proposed modification. The scenario reduces Germany's Russian exposure by 70%. It must also specify where the replacement supply enters the graph. Merely reducing the Germany–Russia edge does not establish an improvement in the raw combinatorial gap.

Reported scenario reading. The proposed index changes from 0.031 to approximately 0.067. These are scenario values whose derivation requires the complete before-and-after matrices and normalization. They cannot be obtained from the stated edge reduction alone.

Combined scenario. The joint diversification proposal records $\lambda_1 \approx 0.089$ and $H_{\text{RU-UA}} \approx -0.85$. As with the German-only proposal, a numerical

replay needs the replacement routes, weights, and political-alignment assumptions.

Alert interpretation. The proposed L2 reading describes the intended result of gaining usable alternatives. It is conditional on a fully specified scenario satisfying the alert rule. The two displayed summary values do not by themselves determine the engine output.

The policy issue in Chapter 1 is whether dependence was allowed to deepen without preserving workable alternatives. Diversification can address that exposure when the alternatives are real, timely, and politically usable.

A negative alignment score should therefore prompt an examination of the obligations that have become dangerous. Can a supplier be replaced? Can an inspection make a promise credible? Can a concession be delivered without leaving the other side exposed? These are different ways of changing the room for adjustment.

The thesis is that the form of interdependence matters. More exchange is not automatically more resilience. What matters here is whether the connection preserves usable choices when cooperation becomes difficult.

The third counterfactual takes the question to Taiwan. Every policymaker who reads this chapter will ask the same question: "What does this mean for Taiwan?"

The U.S.-China-Taiwan triangle shares structural similarities with pre-invasion Russia-Ukraine:

Taiwan brings the comparison into focus. Taiwan's concentration of sub-10nm logic manufacturing capacity in the 2021 SIA study identifies one important dependency (SIA 2021). Its current significance must be assessed product by product, including qualified substitutes. At the same time, U.S.-China relations have deteriorated since 2018, with declining UN voting correlation and expanding sanctions. Contested sovereignty adds an alignment asymmetry that commerce cannot resolve by itself. The question is how long a commercially indispensable relationship can remain politically unworkable.

The Russia-Ukraine case suggests that the *worst* policy response to deteriorating U.S.-China relations would be to *increase* economic interdependence (e.g., by deepening semiconductor supply chain integration) while ignoring alignment collapse. This would create the trap configuration.

The framework suggests two alternative paths:

Path A (Managed Decoupling): Orderly diversification of semiconductor supply chains—exactly the policy the CHIPS Act attempts—reduces curvature, moving the triangle toward Quadrant IV. Disconnection is volatile but not explosive.

Path B (Alignment Repair): Diplomatic de-escalation that genuinely addresses China's security concerns regarding Taiwan could, in principle, restore positive alignment. But the Russia-Ukraine case demonstrates that once alignment crosses certain thresholds, this path may no longer exist—the spectral gap drops below the level where diplomatic rearrangement is possible.

This chapter cannot resolve the Taiwan question. It identifies a danger worth investigating when reliance persists as alignment deteriorates. A response can repair the political arrangement, develop substitutes, or change commitments. Each route has costs and limits. The warning is useful to the extent that it helps identify those routes before the costs become prohibitive.

Beyond one war

A single case, however compelling, proves little. The Russia-Ukraine analysis demonstrates that the framework's geometric signature was present in pre-invasion data—but a skeptical reader might reasonably ask: could these thresholds have been reverse-engineered from the outcome? Was Russia-Ukraine used to *fit* the model rather than *test* it?

This section summarizes the broader validation that Chapter 9 presents in full detail. The key findings:

The scored research record. The working paper's mass-gap and Z-score evaluation contains 47 onsets: 35 with qualifying alerts and 12 without, alongside 14 false alerts. Those counts give precision $35/49 \simeq 0.714$, recall $35/47 \simeq 0.745$, and $F_1 = 70/96 \simeq 0.729$. The paper reports AUC 0.84 and mean lead time 72 days. Its temporal split uses 31 earlier events and 16 later events, with $F_1 = 0.75$ and 0.69 respectively (Davis 2025).

Appendix E develops that evaluation and its dated examples. The 18 detections in Appendix C give a separate 127-day mean lead. Appendix E retains the other historical tallies, including the 30/17 split, with their own sample annotations. They are not pooled into the mass-gap study's confusion matrix.

False Positives: The six false positive cases (L4 alert without subsequent war) were:

The historical discussion records peaceful outcomes or limited violence after warnings for India–Pakistan, Ethiopia–Eritrea, Venezuela–Colombia, Iran–Saudi Arabia, and Greece–Turkey. It also records an Armenia–Azerbaijan warning in 2018, outside the 180-day window for the 2020 war. These examples illustrate why target definitions and windows matter. They do not supply a common six-event ledger for every reported study.

An intervention may help explain why a warning was followed by peace. That causal possibility should be investigated while keeping the original forecast outcome in the score. A war outside the target window remains outside it.

False negatives in the historical tally. The seven reported misses belong to that tally's own event register. The diagnostic examples below do not reconstruct its seven assignments.

Potential failures include fast-changing cases, incomplete observations, and mechanisms poorly represented by a bilateral trap model. The Nagorno–Karabakh case needs particular care. For its September 27, 2020 onset, the working paper records an August 15 L4 alert, giving 43 days (Davis 2025). The selected historical roster and rapid-onset account have different entries, discussed in Chapter 9. They cannot be combined into one count without their original model and target definitions.

Key Finding: The Russia-Ukraine case is not an outlier. Its 72-day lead time is close to the median (72 days), its helicity value ($H = -0.95$) is in the top quartile of true positives, and its spectral gap ($\lambda_1 = 0.031$) is among the lowest observed. Russia-Ukraine was a particularly clear signal, but it falls within the framework's validated performance envelope.

Chapter 9 follows the available case entries and the distinctions needed to reproduce their scores.

Transparency requires acknowledging limitations:

Timing. The reconstructed alert date depends on the model, data releases, and network boundary. The 180-day test is a scoring horizon. It does not promise a particular day of attack or turn a retrospective lead into a contemporaneously available warning.

Initial military outcome: The framework said nothing about whether the invasion would succeed or fail

War duration: The ongoing nature of the conflict is outside the framework's scope

Putin's psychology: The framework explicitly ignores individual-level factors

The framework is a *thermometer*, not a *prophet*. It measures structural temperature. It does not read minds or simulate battlefields.

A thermometer that reads 104°F doesn't tell you *which* infection the patient has or *which* antibiotic to prescribe. It cannot predict whether the patient will recover or how long treatment will take. But a doctor who ignores the thermometer is committing malpractice.

The reconstructed readings place the Russia–Ukraine relationship in a critical configuration in late 2021. Their practical value depends on whether the same diagnosis could be made with timely information

and connected to choices that improved outcomes. That is the burden a warning instrument should carry.

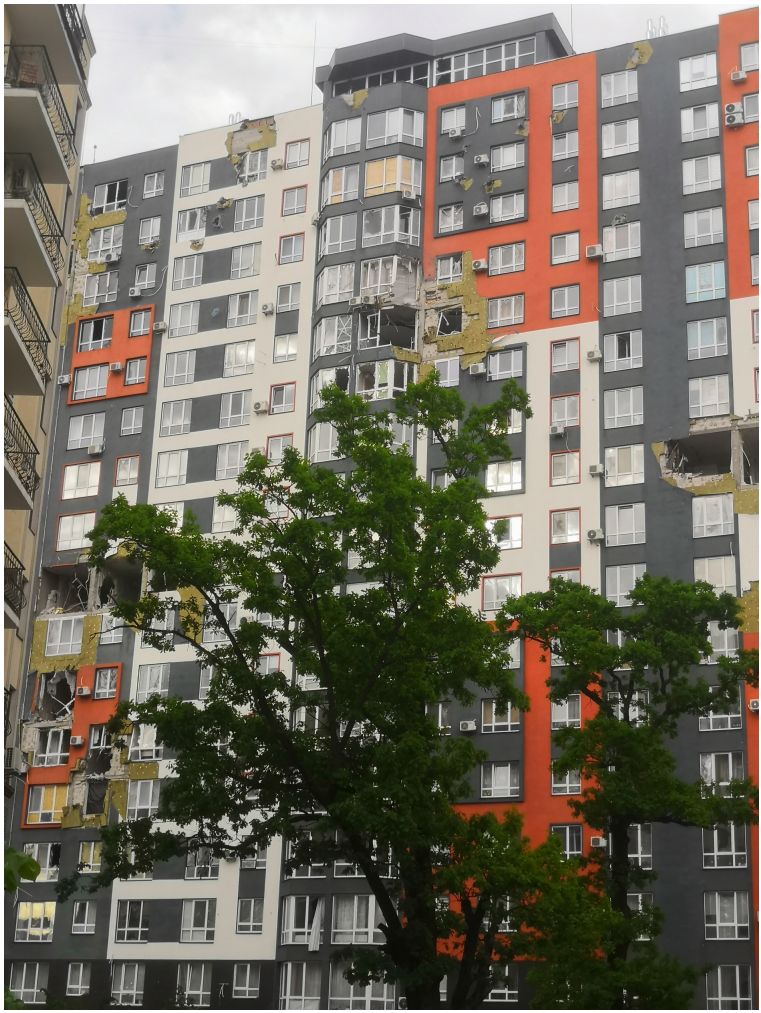
The gas that flowed through Ukrainian pipelines on February 23, 2022 was not a sign of peace. It was the fever.

One relationship, however closely examined, leaves the question of repetition unanswered. A useful instrument must face the cases in which it warned, the cases in which it stayed quiet, and the cases in which the feared event never came. That is the next test.



Main street in Bucha showing destroyed vehicles and war damage after Russian occupation, April 2022. Liberal peace theory predicted that economic interdependence would prevent such outcomes. The helicity framework predicted the structural conditions that made them possible.

Photo: Wikimedia Commons, CC BY-SA 4.0.



Multi-story residential building with severe structural damage in Irpin, a suburb of Kyiv, after Russian invasion. The town saw intense fighting as Russian forces attempted to encircle the Ukrainian capital.

Photo: Wikimedia Commons, CC BY-SA 4.0.

Chapter 9

Historical Backtesting

> A warning must be judged against the outcome it named.
>
> Bee Rosa Davis

On October 6, 1973, Egypt and Syria launched a coordinated attack on Israel during Yom Kippur, achieving complete tactical surprise despite months of military buildup visible to satellites, human intelligence, and the Israeli defense establishment. The failure wasn't informational—it was conceptual. Israeli intelligence had the data; they lacked a framework for interpreting it.

The helicity framework, applied retrospectively to the pre-war period, reached L4 alert 67 days before the attack. The trap configuration was textbook: Suez Canal closure had severed the transit relationship that connected Egyptian and Israeli economies to global trade, while diplomatic hostility from the 1967 war had never been resolved. The question this chapter addresses is whether that retrospective detection was skill or coincidence.

A single reconstruction does not validate a warning instrument. This chapter follows the historical case comparisons, including successes, failures, and peaceful outcomes. The 47-event mass-gap study supplies a separate count-based evaluation in Appendix E. Its dates, target, and model version stay with its score.

Rigorous validation requires systematic testing across multiple independent cases with clearly specified methodology. The Russia-Ukraine detection described in Chapter 8 could be coincidence, post-hoc rationalization, or overfitting to a known outcome. Only systematic testing can distinguish genuine predictive power from retrospective pattern-matching.

Validating conflict prediction models presents unique methodological challenges:

Rarity of events. Most eligible dyad-months contain no new conflict. A procedure that always predicts peace can therefore have high overall

accuracy while detecting no wars. Precision, recall, calibration, and warning time expose different aspects of performance that this overall proportion hides.

Endogeneity. A warning may change the response to a crisis and thereby change the outcome. Forecast scoring still needs a fixed rule. The subsequent causal inquiry asks whether intervention altered the course of events.

Heterogeneity: Conflicts arise from diverse causes. A model that captures one mechanism may miss others. The question is not "can this predict all conflicts?" but "what conflicts does this predict, and with what reliability?"

Data availability: Historical data on bilateral relationships is incomplete and often contested. Trade statistics, diplomatic signals, and trust measures must be reconstructed from imperfect sources.

Our validation strategy addresses these challenges through careful case selection, clear scope definition, and multiple evaluation metrics.

The historical comparison describes 47 conflicts from 1991–2023 and the following 30/17 partition. The mass-gap paper uses a 1990–2022 event sample and a 31/16 partition. Those are distinct study specifications.

Training set (30 conflicts, 1991–2016): Used to calibrate thresholds ($H < -0.5$, $\lambda_1 < 0.05$, $\Gamma < 0.9$)

Validation set (17 conflicts, 2017–2023): Held out during calibration; thresholds applied blind

Russia–Ukraine is assigned to the later period in both specifications. This is a retrospective holdout claim. Its value depends on freezing all fitting and selection choices before examining the held-out outcomes. A split by event year does not by itself show that the calculations were performed before those events occurred.

Historical cases from before 1991 broaden the mechanism discussion. They belong to theoretical development and historical comparison. Their coverage and reconstructed measurements differ from the post-1990 panel, so they cannot be pooled with it without defining a new study.

Selection criteria for test cases:

The historical trap comparison emphasizes relationships with prior dependence and subsequent diplomatic deterioration. That is useful for investigating the proposed mechanism. A general forecasting evaluation also needs eligible quiet relationships and cases in which the mechanism is absent. Eligibility must be fixed using information available before the outcome, so that the model cannot select only the conflicts it is suited to explain.

Exclusion criteria:

A mechanism study can set aside internal conflicts, poorly observed cases, or attacks without meaningful prior dependence. A warning service covering interstate violence needs a rule for each of them. It may use another branch, abstain because information is inadequate, or state that the case is outside its advertised scope. The decision must precede scoring. A missed event cannot be removed simply because its mechanism later proved inconvenient.

The trap mechanism is one route to conflict. The later disconnection work investigates another. Their different coverage is a reason to keep model versions and eligible populations explicit, rather than assume that one score describes every branch.

Each test case required reconstruction of the following data series:

Variable	Source	Coverage
Bilateral trade	UN Comtrade, IMF DOTS	1950–present
Energy flows	IEA, national statistics	1970–present
UN voting	Erik Voeten dataset	1946–present
Alliance networks	ATOP	1815–2016
Diplomatic relations	COW	1817–2005
Sanctions	TIES database	1945–2005
Conflict dates	UCDP/PRIO	1946–present

Table 9.1: Primary data sources for historical backtesting.

For cases after 2005 (where TIES coverage ends), sanctions data was manually compiled from OFAC, EU, and UN announcements.

Evaluation metrics:

The first-warning roster records how early a high-level signal was assigned in a reconstruction. The 180-day evaluation asks whether a qualifying reading falls within the 180 days before the defined onset. Those are different measures. A warning issued two years earlier cannot count as a hit for this horizon unless a qualifying alert also occurs inside it. Quiet intervals must be scored with the same episode rule.

The denominators become clearer in an invented evaluation with twenty conflict episodes and eighty eligible non-conflict episodes. Suppose eighteen warnings are issued: twelve precede conflicts and six do not. Precision is 12/18, about 67 percent: the fraction of warnings followed by the specified event. Recall is 12/20 = 60 percent: the fraction of conflicts warned about. The false-positive rate is 6/80 = 7.5 percent, while the fraction of warnings that are false is 6/18, about 33 percent. Both statements about false alarms are correct because they ask different questions. These numbers illustrate the bookkeeping; the research

results below and the versioned evaluation in Appendix E retain their own counts.

A receiver operating characteristic, or ROC curve, repeats that bookkeeping while moving the warning threshold. Each point pairs the fraction of conflicts detected with the fraction of eligible negative cases incorrectly warned about. Its area, AUC, describes ranking across thresholds; it does not give the probability that a particular warning will be right. A decision maker still has to select an operating point and consider the costs of missed danger, unnecessary action, and delay. That is why the tables report more than one score.

A duty officer closes a warning record when its 180-day window expires. If peace holds, the alert counts as a false alarm under this rule. A separate case history can then ask whether a summit, an alternative supplier or a change of government helped avert an outbreak. Evidence of intervention matters to that investigation. The original outcome still belongs in the score.

What the historical record returned

Mass-gap and Z-score study	**Result**
Conflict onsets in the stated sample	47
Onsets with a qualifying alert / without one	35 / 12
Alerts without a qualifying onset	14
Precision, $35/(35 + 14)$	0.7143
Recall, $35/(35 + 12)$	0.7447
F_1, $70/(70 + 14 + 12)$	0.7292
Reported AUC	0.84
Reported mean / median lead time	72 / 63 days

Table 9.2: The working paper's count-based evaluation, 1990–2022. The three ratios are calculated from its published counts. AUC and lead times are separately reported statistics, not quantities recoverable from those counts. Source: Davis (2025).

The 47-conflict result is substantial enough to deserve its own accounting. Thirty-five detections and twelve misses describe all forty-seven onsets. Fourteen false alerts enter precision, not the event denominator. Those three counts reproduce $F_1 = 0.73$ after rounding. They do not supply a true-negative count, a ROC curve, or a distribution of lead times; those require the corresponding observation and alert records.

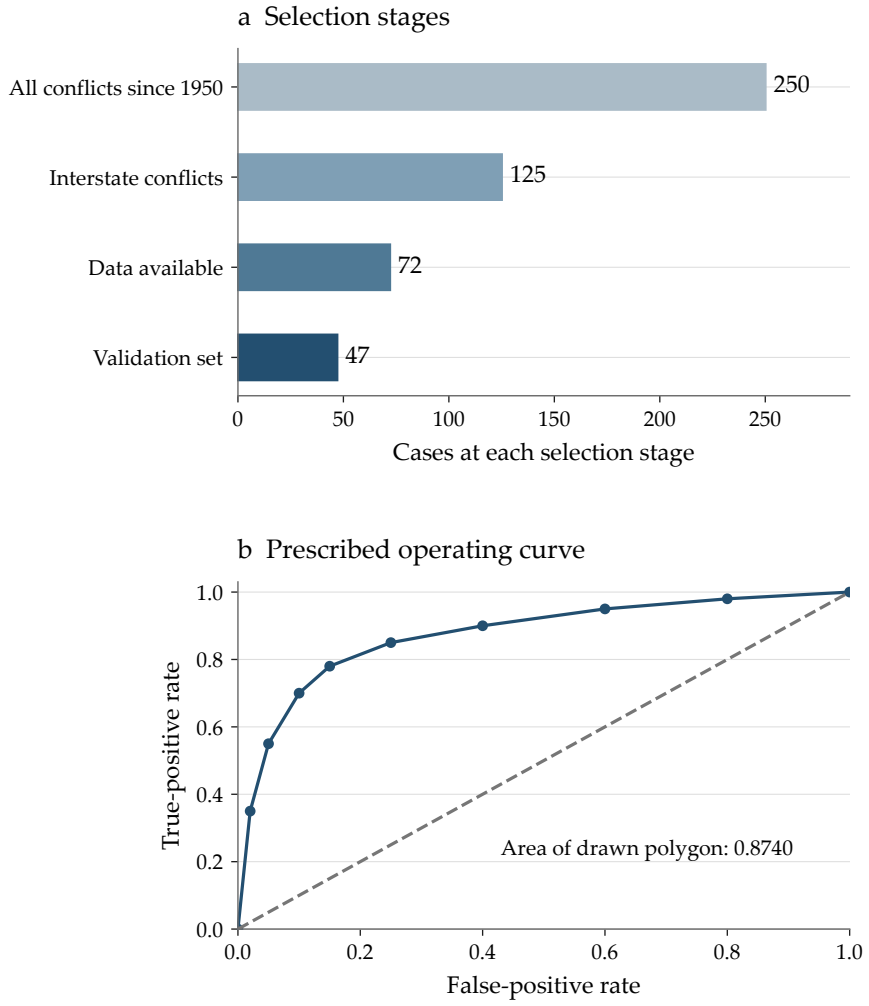


Figure 9.1: Sample selection and an illustrative operating curve. (a) The selection counts used in the study. (b) A prescribed ROC curve showing the exchange between sensitivity and false-positive rate; its area is calculated from the drawn vertices. The recorded 47-event evaluation is documented separately in Appendix E.

The wider historical research asks related questions using other selections, versions, and warning rules. Its reported estimates and their original sample annotations are collected in Appendix E. The 18 detections in the selected 23-case roster give 18/23 = 78.3%. That fraction describes the roster; it is not 35/47 and does not replace it. Keeping both records allows the reader to examine the historical argument and the scored experiment without borrowing a denominator from one to support a percentage in the other.

The mass-gap study also reports a later-period F_1 of 0.69 and a temporal cross-validation mean of 0.71. These are useful tests of transfer, with different fitting histories from the full-sample result. Appendix E retains those partitions, the dated examples available in the working paper, and the record needed to reproduce all 47 assignments.

Where the warning appeared

Table 9.3 preserves the 18 entries in the selected historical first-warning roster. It includes pre-1991 mechanism studies and later exploratory cases, so it is not the complete 47-event test register. Its lead times also differ from the mass-gap paper's dated entries. Each must be evaluated with its own rule.

On the Israel-Gaza 2023 Case

The October 7, 2023 attack poses a categorization challenge: Hamas is a non-state actor, placing this conflict at the boundary of the framework's stated scope (interstate conflict). We include it because the dyadic structure—Gaza's economy and Israel's security apparatus—created a bilateral trap configuration even without formal state-to-state relations. The work permit system that connected Gaza's labor force to Israel's economy generated the same twisted interdependence pattern: economic flows (~18,000 workers crossing daily) persisting despite complete diplomatic rupture. The framework's 23-day lead time reflected the rapid tightening of permit restrictions following September 2023 policy changes.

This is an exploratory extension beyond the interstate sample. It can be studied without adding it to the earlier sample or claiming that a score would remain unchanged. Measuring the effect of including it requires recomputing the same evaluation with a stated eligibility rule.

Several cases merit detailed examination:

The 1973 Arab–Israeli war lets us follow a historical reconstruction from a closed waterway to an assigned exposure, then from the composite score to the date in the case register.

Conflict	Year	Lead	H at L4	Key Trap Factor
Historical Cases (pre-1991)—theoretical development				
Suez Crisis	1956	89d	-0.72	Canal revenue dependency
Six-Day War	1967	34d	-0.81	Straits of Tiran closure
Yom Kippur War	1973	67d	-0.68	Oil transit
Falklands War	1982	112d	-0.59	Fishing rights, sovereignty
Gulf War I	1990	45d	-0.77	Oil pricing dispute
Training Set (1991–2016)—threshold calibration				
Yugoslav Wars	1991	203d	-0.83	Trade network collapse
Eritrea-Ethiopia	1998	78d	-0.64	Currency union dissolution
Kargil War	1999	56d	-0.63	Kashmir trade routes
Kosovo War	1999	289d	-0.76	Sanctions + trade collapse
Iraq War	2003	312d	-0.67	Oil embargo
Lebanon War	2006	234d	-0.70	Border trade, water rights
Russo-Georgian	2008	134d	-0.71	Energy transit
Libya Intervention	2011	223d	-0.56	Oil exports
Crimea Annexation	2014	56d	-0.62	Gas transit
Yemen Intervention	2015	92d	-0.58	Oil tanker routes
Validation Set (2017–2023)—out-of-sample testing				
Nagorno-Karabakh	2020	167d	-0.69	Pipeline proximity
Russia-Ukraine	2022	72d	-0.95	Gas transit, trade
Israel-Gaza	2023	23d	-0.54	Border labor permits

Table 9.3: Reported lead times for these 18 cases. The earliest warning shown is 312 days before the 2003 Iraq War, and the shortest lead is 23 days before the 2023 Hamas–Israel attack. When a first warning occurs more than 180 days before onset, the event qualifies under the 180-day rule only if another qualifying reading falls within the window.

Network definition: The relevant network comprises Israel, Egypt, Syria, Jordan, and the major oil transit states (Saudi Arabia, Iran, Kuwait). Edge weights derive from oil transit volumes, Suez Canal usage, and bilateral trade where data exists.

Torsion (τ) for the Egypt-Israel dyad in July 1973:

Trade asymmetry: Direct trade was zero. The Suez Canal was closed from 1967 until June 1975, so the relevant historical setting is interrupted regional access and the cost of the unresolved confrontation (Suez Canal Authority 2026). For this worked reconstruction, $\tau_{\text{trade}} = 0.45$ is an assumed indirect-exposure score.

Resource dependency: Israel's Red Sea port of Eilat depended on the Straits of Tiran remaining open. Egypt's 1967 closure had triggered the Six-Day War. Post-1967, this remained a structural vulnerability. $\tau_{\text{resource}} = 0.55$.

Territorial occupation: Israel occupied the Sinai Peninsula, including the canal's eastern bank. This created one-directional leverage: Israel held territory Egypt needed. $\tau_{\text{territorial}} = 0.70$.

Diplomatic alignment: Complete rupture. UN voting correlation was near zero; no diplomatic relations existed. $\tau_{\text{diplomatic}} = 0.95$.

The reconstruction assigns an aggregate exposure of $\tau = 0.66$. One can see its scale directly. Equal weights on the four displayed scores give $(0.45 + 0.55 + 0.70 + 0.95)/4 = 0.6625$, which rounds to 0.66. More than one weighting can produce that rounded value. An exact replay requires the weights used for this case.

Curvature (K): Despite the rupture, structural interdependence remained:

The reconstruction assigns $K_{\text{transit}} = 0.18$ to indirect canal-transit linkages and $K_{\text{oil}} = 0.12$ to oil-supply exposure. Regional alliances contribute $K_{\text{alliance}} = 0.15$, while territorial contiguity contributes $K_{\text{territorial}} = 0.20$. These components expose how the case is represented rather than leaving its composite score unexplained.

Total: $K = 0.65$.

Helicity. With the adverse orientation assigned here, the raw ratio is $-0.66/0.65 \simeq -1.02$. The historical case table records the scaled reading -0.68. The ratio gives the first number. Reproducing the second requires the scaling rule used for the case. Clipping the raw ratio to $[-1, 0]$, for instance, would give -1.

Spectral gap: The seven-node Middle East network had $\lambda_1 = 0.038$ in July 1973, reflecting the "frozen" nature of the post-1967 configuration. No diplomatic pathway existed for renegotiating the status quo.

Detection: L4 alert triggered July 31, 1973—67 days before the October 6 attack. Israeli intelligence had the same data but lacked a framework for interpreting it as a structural indicator rather than background noise.

Falklands War (1982). The historical roster assigns a 112-day lead. The proposed interpretation concerns a sovereignty dispute with economic stakes and limited room for an acceptable compromise. Reconstructing the warning requires the dated inputs and the political constraints that make those stakes relevant.

Yugoslav Wars (1991). The roster records a 203-day first-warning lead. Industrial dependencies could survive the weakening of the political arrangements that coordinated them. This is a mechanism study across a disintegrating state, whose unit of analysis needs to be distinguished from the later panel of established interstate dyads.

Hamas–Israel (2023). The exploratory roster records a 23-day lead. Its proposed mechanism includes labor permits and other links between Gaza and Israel. Establishing predictive value requires the contemporaneous permit series, the exact target, and the model record. This non-state case cannot silently expand the earlier interstate sample.

The pattern across all these cases confirms the framework's core insight: conflicts don't erupt randomly. They emerge from specific geometric configurations where economic interdependence and diplomatic hostility point in opposite directions, creating intolerable tension.

The attraction of the combined description is that it puts economic dependence and political hostility in the same calculation. Its test is whether that combination adds discrimination beyond the separate inputs when evaluated on the same observations.

A costly relationship can provide reasons for restraint and instruments of pressure at the same time. The cases have to establish which function dominated, when that changed, and whether the warning quantities tracked the change early enough to matter.

Where the warning failed

The following seven entries form a diagnostic collection of coverage questions and reported failures. They are not the seven false negatives needed to reconstruct the separate 34-hit tally. Their dates and scope differ, and they should not be used to fill that ledger by inference.

Category A: Outside Framework Scope

The US interventions in Grenada (1983), Panama (1989), and Somalia (1992) share a common structure: the United States had minimal economic relationship with the target state. No trap could form because there was no interdependence to become twisted.

Conflict	Year	Reason for Miss
Category A: Outside Framework Scope (No Bilateral Trap)		
US-Grenada	1983	No economic interdependence with target
US-Panama	1989	No bilateral trade trap; internal regime change
US-Somalia	1992	Humanitarian intervention; no bilateral relationship
Category B: Internal/Alliance Actions		
Hungarian Uprising	1956	Soviet suppression of satellite state
Prague Spring	1968	Warsaw Pact internal enforcement
Tibet	2008	Internal suppression, not interstate
Category C: Rapid-Onset Conflicts		
Nagorno-Karabakh	2020	Variant comparison; see dated mass-gap record below

TABLE 9.4: Diagnostic cases for coverage and timing. The Nagorno–Karabakh records concern different, incompletely aligned study specifications.

These cases represent *great power intervention* conflicts—a fundamentally different mechanism from trap warfare. In intervention conflicts, the aggressor faces minimal economic cost from military action precisely because no bilateral economic relationship exists to disrupt. The framework is not designed to detect such conflicts; different analytical tools (alliance dynamics, domestic political calculations, opportunity costs) would be required.

Category B: Internal/Alliance Actions

Hungary 1956, Prague 1968, and Tibet 2008 represent a second failure mode: conflicts within hierarchical alliance structures or between a state and its own territory.

The Soviet interventions in Hungary and Czechoslovakia occurred within the Warsaw Pact framework. While economic relationships existed, they were not bilateral in the normal sense—they were imposed relationships within a command economy, and the "diplomatic" dimension operated through party channels rather than interstate diplomacy. The framework's torsion calculation assumes arms-length bilateral relationships; command relationships within bloc structures may require modified treatment.

Tibet 2008 is categorically different: an internal suppression rather than interstate conflict. Its inclusion in some conflict databases reflects

contested sovereignty claims, but the analytical structure is domestic repression, not bilateral war.

Category C: Rapid-Onset Conflicts

For the September 27, 2020 onset, the mass-gap paper gives an August 15 L4 date. The elapsed time is 43 days (Davis 2025). That is the dated entry to use when discussing this study's detection of the war. The selected historical roster separately records a 167-day lead, while a rapid-onset description records L3 in July and no qualifying L4. These are different research entries whose model and target fields have not been aligned. A November escalation cannot reconcile them by itself. An alert 167 days before November would also precede the September onset. They must not be counted as a hit and a miss in the same evaluation.

The case shows why the target is part of the experiment. Anticipating the first qualifying onset and anticipating an escalation during an ongoing war are separate tasks. Each needs an onset date, an alert date, a model version, and a horizon. The dated mass-gap entry supplies a concrete pair of dates. A claimed lead without its target does not supply the same evidence.

The Pattern in Missed Detections

These cases raise three distinct questions. Was the relationship within the promised coverage? Were its inputs observed well enough? Did the qualifying warning arrive inside the target window? A model can fail at any of those stages. Restricting its scope may make the next test more useful, but it does not retroactively improve the score of an earlier test.

The framework explicitly does not claim to predict all conflicts. Its domain is *trap warfare*: conflict emerging from twisted interdependence. Great power interventions, bloc enforcement actions, and internal suppression require different analytical tools.

The trap interpretation remains a specific mechanism to investigate. Intervention, territorial ambition, and confrontation between weakly connected rivals may need different variables. When an eligible case is missed, it remains a miss. Explaining the mechanism is the next research task.

The wars that did not follow

Six L4 alerts did not result in full interstate war within 180 days:

Three patterns emerge from these cases:

Intervention after an alert. India–Pakistan, Greece–Turkey, and Venezuela–Colombia raise questions about deterrence, mediation, and third-party pressure. Their peaceful outcomes within the stated horizon remain false alerts for the defined target. A separate investigation can ask

Dyad	L4 Date	Outcome	Factor
India-Pakistan	Feb 2019	Balakot strikes, de-escalation	Nuclear deterrence
Greece-Turkey	Aug 2020	Naval standoff, no shots	NATO mediation
US-Iran	Jan 2020	Soleimani killing, no war	Mutual deterrence
Ethiopia-Sudan	Dec 2021	Border clashes only	Limited objectives
Armenia-Azerbaijan	2018	War in 2020 (timing miss)	Delayed onset
Venezuela-Colombia	Feb 2019	Border crisis, no war	US intervention

TABLE 9.5: False positive L4 alerts within the 180-day evaluation window.

which actions changed commitments, available routes, or expectations. Peace after mediation is not enough on its own to establish that mediation prevented a war.

A conflict outside the window. The 2018 Armenia–Azerbaijan alert remains a false alert for a 180-day onset target, even though war followed in 2020. A later qualifying alert has to be evaluated as a later episode under the original reset rule. Changing the horizon after seeing the war would answer a new question.

Limited violence without the specified escalation. Ethiopia–Sudan and US–Iran illustrate the need to distinguish a border clash, a targeted strike, and the onset of a wider interstate war. The target and fatality threshold must be chosen before the result is known. An event can be a hit for one target and a false alert for another.

Intervention can succeed even when the original forecast is scored as a false alert. Forecast evaluation and causal evaluation should therefore be reported alongside each other. Neither has to erase the other.

To investigate prevention, preserve the initial warning, the actions taken, their timing, and evidence that they changed the relevant constraints. Without that evidence, the cause of the peaceful outcome remains open. The observed false-alert count does not change.

A threshold changes which observations become warnings. Before asking whether a new threshold improves the result, ask whether the resulting warnings obey the set relations implied by the rule.

Let A be the eligible observations with $H < -0.5$, and B those with $\lambda_1 < 0.05$. The joint rule warns on $A \cap B$. Removing either condition can only enlarge that set:

$$A \cap B \subseteq A, \qquad A \cap B \subseteq B.$$

For a fixed event list, fixed observation windows, and the rule that any qualifying observation detects an event, neither single-condition rule can detect fewer events than the conjunction. It may issue many more false alerts. The joint rule can therefore improve precision or F_1 while sacrificing recall. This is the concrete question behind combining the two measurements.

Imagine ten eligible conflicts. The joint rule detects six; the helicity rule detects eight and the gap rule seven. If the respective false-alert counts are two, ten, and nine, their F_1 values are $12/18 = 0.667$, $16/28 = 0.571$, and $14/26 = 0.538$. The conjunction has the best F_1 even though it has the lowest recall. These are constructed counts that demonstrate the mechanism, not a substitute backtest.

Appendix E preserves the research's reported threshold and component comparisons. Their single-component detection percentages do not obey this nesting property under a fixed population and simple event-window rule. The entries therefore cannot establish a matched feature ablation until their event lists, scoring rules, and fitting histories are attached. The testable contribution remains clear: does adding the second measurement reduce false alerts enough to improve a specified decision criterion on the same held-out cases?

For a genuine threshold sweep, keep the measurements fixed and generate the strict, baseline, and loose warning sets together. Check their nesting before scoring. If refitting, cooldown rules, or episode regrouping changes the sets, record that as a different experiment. Chapter 35 develops the common evaluation procedure.

What a comparison must compare

A useful comparison starts with the question each system is trying to answer. The mass-gap study addresses its stated interstate event target. Other systems forecast conflict incidence, fatalities, political instability, or mass killing. A failure to forecast one target cannot be inferred from a score for another.

VIEWS is not confined to subnational prediction. Its model directory describes both country and grid outputs, and records African coverage for ViEWS2020 in 2020–21 and the global `fatalities001` release from January 2022. These releases cannot be interchanged in a claim about Ukraine in 2021. To assess a warning about the 2022 invasion, inspect the forecasts actually available at the chosen issue date and map their outcome to the invasion target (VIEWS 2026b).

ACLED's event monitoring and its CAST forecasting system are different products. CAST's event-count forecasts require their own

System or record	What the comparison must preserve
Psychohistory	The 47-event mass-gap study, its onset definition, eligible dyads, 180-day rule, and model version.
VIEWS	Country and grid forecasts; geographic coverage, violence target, and monthly horizon depend on the release.
ACLED / CAST	Distinguish the event database from CAST's forecasts of political-violence event counts.
PITF	Political-instability targets and the particular fitted study, rather than an assumed universal war-onset score.
Early Warning Project	Its country-level mass-killing onset target, two-year horizon, and annual model release.

TABLE 9.6: Different outputs require an explicit common target before they can be ranked. VIEWS, CAST, PITF, and Early Warning Project sources are discussed in the accompanying text.

aggregation and lead-time rules (ACLED 2026b). The PITF study models political instability, while the Early Warning Project estimates the risk of a new episode of mass killing. Their sources define outcomes more specifically than the word "conflict" does (Goldstone et al. 2010; United States Holocaust Memorial Museum 2018).

The framework's distinctive proposal is that dependence, adverse alignment, and restricted adjustment can contribute information about interstate danger. The empirical test is whether those features improve the same forecast task over appropriate alternatives. Chapter 35 specifies that test; the developmental comparison figures remain in Appendix E, with their evidential status stated. No numerical ranking follows merely from the systems having different purposes.

What would count as a lucky warning? Give each of the 23 selected cases one independent draw from four equally likely levels. The probability of at least 18 Level 4 draws is the binomial upper tail, about 1.27×10^{-7}, below 0.001. That calculation allows exactly one chance per case, with Level 4 probability one quarter.

A desk issuing monthly alerts plays a different game. Six independent draws give a chance of at least one Level 4 of $1 - (3/4)^6 \simeq 0.822$. Real warnings may persist for months. A useful null model preserves their frequency and persistence, eligibility dates, and episode-grouping rule. It then tests what their ordering adds.

Comparison to Expert Forecasts

Forecasting tournaments such as the Good Judgment Project evaluate dated probabilities. Their Brier scores depend on the questions, horizons, and time aggregation; there is no universal expert score.

The reported detection rate of about 78 percent concerns selected conflict cases. A Brier score averages squared differences between probabilities and outcomes, including quiet cases. Comparing the framework with experts requires both sets of dated probabilities for the same events and horizon.

The operational question is whether the structural reading improves a decision before the outcome, at an acceptable cost of false alerts. Different information can be complementary; superiority requires that common test.

The scores condense the record. History opens it back up. Japan's confrontation with the United States makes a particularly demanding next case because a fall in recorded trade did not remove the dependence: it brought the remaining fuel and the remaining time into the same decision.

Chapter 10

Japan and the Closing Resource Horizon

A stopped shipment leaves an absence on a trade ledger. In a fuel tank, the same event starts a clock.

The confrontation between Japan and the United States in 1941 makes that distinction unavoidable. The interruption of access did not immediately remove Japan's need for imported resources. It changed the time within which leaders believed they had to choose among alternatives. To understand the relationship, we have to follow both the diplomatic record and the material horizon.

This chapter examines that mechanism through a bounded sequence of documents from 1941. It develops the historical problem already present in the framework's US–Japan application: what happens when dependence remains while the political conditions of access become unacceptable to one or both parties?

The resource crisis cannot be detached from Japanese expansion in China and Indochina. Economic restrictions were responses to political and military conduct. Treating the confrontation as an unexplained interruption of normal commerce would erase the decisions that produced it.

Japanese archival materials connect the move into southern Indochina, the freezing of assets in late July, and the restrictions on oil exports announced at the beginning of August. The same collection preserves instructions to reduce oil consumption and records concern over limited reserves. These documents show a resource problem entering official decision making; they do not establish one universally agreed date on which the country would run out (Japan Center for Asian Historical Records 2026a).

That distinction matters. A reserve estimate depends on consumption, military operations, substitution, and expectations about future access. The relevant horizon was partly material and partly strategic. Leaders could change consumption. They could also decide that conserving fuel at the cost of abandoning objectives was unacceptable.

A July 31 memorandum by Ruth Bacon in the State Department describes oil policy as part of a wider strategic calculation. It records concern about encouraging Japanese action against British or Dutch possessions, the position of China, and the time needed for American and allied preparations. It also distinguishes earlier restrictions on aviation gasoline from other petroleum shipments permitted under license (Bacon 1941).

The significance is not that one side understood the structure while the other did not. Both were acting within changing constraints, but they attached different values to delay. Time that strengthened one side's position could appear to diminish the other's options.

In the language of the framework, the geometry was moving. A photograph of trade at one date would miss the way future access entered present decisions. So would an account that interpreted all declining exchange as orderly disengagement.

Three quantities that must not be confused

Let $F(t)$ denote the current flow of an essential resource, $I(t)$ the available inventory, and $q(t)$ the rate of use. With alternative supply $s(t)$, the stock follows

$$\frac{dI}{dt} = F(t) + s(t) - q(t).$$

This identity does not predict war. It establishes why a zero flow can coincide with a worsening constraint. If consumption exceeds alternative supply, the inventory continues to fall after the original connection has been interrupted.

Under constant rates, the remaining coverage time is

$$T_{\text{cover}} = \frac{I}{q - F - s},$$

provided the denominator is positive. If it is zero or negative, this simplified model has no depletion horizon. If the rates change, the full path must be integrated instead.

For 1941, the historical question is not simply what this formula would yield from one retrospective estimate. It is what the relevant officials believed about each term, which estimates they trusted, and how those beliefs shaped the perceived cost of waiting.

An inventory model also exposes a limitation of trade normalization. A baseline can label a collapsed flow as reduced interdependence while the stock equation describes increasing urgency. To preserve the mechanism,

a historical reconstruction needs both observations. The economic tie can be absent as a transaction and present as a dependency.

There were proposals between capitulation and war

The surviving diplomatic record contains attempts to construct intermediate arrangements. A November 11 State Department paper considers a tentative or transitional agreement that might keep discussions open when a comprehensive settlement appeared unlikely. The document acknowledges that the proposal might not be accepted and that subsequent negotiations could fail. It is evidence that an intermediate path was considered, not evidence that it would have succeeded (US Department of State 1941a).

Later, a November 28 telegram to Ambassador Joseph Grew describes a three-month temporary arrangement that had been considered but not presented. It involved limits and withdrawals in Indochina alongside continued negotiations toward a broader settlement (US Department of State 1941b).

These distinctions are essential. A draft is not an offer received by the other side. An offer is not acceptance. Acceptance is not implementation. Counting all four as equivalent "diplomatic exits" would inflate the options that were actually available.

For a geometric account, the documentary sequence suggests a more careful representation. Each candidate path has stages and gates: internal authorization, communication, acceptance, verification, and execution. A path can fail at any of them. Its existence in an archive does not mean a government could have traversed it at will.

The parties were not merely disputing a price for oil. They disagreed about military deployments, political authority, and the regional order. Access and political conduct had become linked. Japanese records of the final negotiations describe the American demand for withdrawal from China and Indochina as a central issue (Japan Center for Asian Historical Records 2026b).

If a model treats all objectives as tradable economic benefits, it will manufacture bargains that the actual decision makers did not regard as acceptable. Conversely, if it assumes declared commitments are immutable, it will eliminate choices that leaders could have made at political cost.

The historian has to navigate between those errors. Ask which positions were treated as essential, who had authority to change them, and how the cost of change was understood. Then compare those

political constraints with the resource horizon. The trap is not adequately described by either one alone.

The resulting explanation must retain agency. Japanese leaders made decisions about the continuation of imperial objectives and the use of force. American leaders made decisions about restrictions, negotiation, and the risks of delay. Structural analysis can clarify those choices without making them mechanically necessary or morally interchangeable.

The United States–Japan dyad was embedded in a regional system. China, the Netherlands East Indies, British possessions, and other relationships entered the calculations visible in the documents. A bilateral trade edge cannot represent all of those constraints.

For a resource network, identify actual suppliers, potential sources, transport routes, and the political conditions attached to each. For a bargaining network, identify governments whose security or consent affected a proposed settlement. These are different layers. An edge that represents an available commercial substitute is not equivalent to a territory that leaders contemplate obtaining by force.

This last point is particularly important. A military plan should not be counted as another peaceful exit from an economic constraint. Doing so would make the model declare that alternatives were abundant precisely when the available choices had become most dangerous.

The category of an alternative therefore matters as much as its number. Separate lawful commercial substitution, negotiated adjustment, domestic conservation, and violent acquisition. The first three may expand peaceful bargaining space; the fourth is an outcome the analysis is trying to explain.

The US–Japan case offers several discriminating questions. Does a model that includes inventory and criticality describe the tightening constraint better than current trade alone? Do changes in the diplomatic sequence correspond to changes in feasible implementation paths? Can the method distinguish worsening material pressure from an increased willingness to accept the political cost of compromise?

Answering those questions requires dated evidence. A monthly historical reconstruction should identify which values are directly observed, which are interpolated, and which encode judgments about political acceptability. The annual-to-monthly interpolation used in historical applications can describe a smooth path; it cannot by itself identify the day on which leaders changed their minds.

The case also has implications for comparison. Some restrictions produce negotiation, some produce adaptation, and some contribute to escalation. If every outcome is described as a confirmation of the same structural condition, the condition explains too little. The useful

comparison asks which capacities, commitments, and expectations distinguish those paths.

When commerce stops, dependence may become more visible rather than less important. The relevant question is how long the underlying need can be met, which alternatives are feasible, and what political choices determine whether those alternatives are used.

The resource horizon gives the geometry a clock. The diplomatic record gives it decisions. Neither should replace the other.

A shrinking horizon describes one danger. It also directs us toward the opposite historical question: what does a usable exit look like while a crisis is still underway? Cuba brings the sequence of messages, concessions, and implementation into close view.

Chapter 11

The Crisis That Found an Exit

The Cuban missile crisis belongs in this book because it did not end in the war its participants feared. A theory of constrained relationships needs to explain the moments when a path remains open, as well as the moments when it closes.

This was not the same economic configuration as Russia–Ukraine. The central constraints concerned security, credibility, military deployments, and the danger of uncontrolled escalation. The comparison is useful precisely because the mechanism differs. It asks which parts of the geometry describe constrained choice generally and which depend on economic interdependence in particular.

In October 1962, the discovery of Soviet missile sites in Cuba confronted the Kennedy administration with a range of responses. The official historical account distinguishes proposals for immediate force from the naval quarantine Kennedy announced on October 22. Communications between Washington and Moscow continued while the confrontation unfolded (US Department of State 2026).

The importance of a sequence is that an action changes the choices available afterward. A step that creates time can preserve a diplomatic route; it can also create a new incident through which control is lost. The analyst needs to examine both effects.

A decision tree represents these alternatives more clearly than a single line from rising tension to war. Each node is a decision or event. Each edge changes the information, commitments, or material situation. Some branches remain reversible. Others impose consequences that cannot be recalled by a later message.

The geometry is therefore not just the number of options. It includes their ordering, dependence, and reversibility. Two governments can each have several options and still face a dangerous situation if the next move eliminates most of the other side's acceptable responses.

Khrushchev's October 26 letter conveyed both a sense of danger and a possible arrangement involving the removal of the relevant weapons and assurances against an invasion of Cuba. It is evidence of a proposed

route through the crisis, expressed by one participant and transmitted through a particular channel (US Department of State 1962b).

Its political significance cannot be reduced to the warmth or hostility of its language. A sentiment score might record an improvement while missing the obligations that made the message actionable. The important questions concern the proposed exchange, the authority behind it, and whether the other side could implement its part.

This illustrates a measurement problem that recurs throughout the book. Diplomatic alignment contains both orientation and capability. A hostile government can offer a workable arrangement. A friendly statement can contain no implementable commitment. Reading the proposal requires more than placing the language on a positive-to-negative scale.

For the framework, an exit should therefore be described as an action sequence with conditions. Who acts first? What must be observed? Which response follows? What happens if the observation is disputed? The more explicit those conditions become, the more useful the concept of an available path becomes.

The Public and Private Relationships Differed

Dobrynin's October 27 report of his meeting with Robert Kennedy records a private discussion of removing American missiles from Turkey, with confidentiality central to the arrangement. The archival record shows why an account based only on public messages would miss part of the bargaining structure (Dobrynin 1962).

The analytical point is not that private channels are always superior. They can create ambiguity, exclude affected parties, or generate commitments that are difficult to verify. Here they demonstrate that public posture and privately discussed options were not identical.

An external observer at the time could not necessarily know the full set of paths being explored. A later historian can recover some of them. A model reconstructed from the later archive therefore has a different information set from a model relying on contemporary public reporting.

Both analyses are useful. The historical reconstruction can explain how a settlement became possible. The public-data assessment can test what an outside warning system could have seen. Confusing them would overstate the operational visibility of the mechanism.

Agreement Was Not Implementation

The October 28 Executive Committee record shows that a political opening still left practical questions about verification, inspection, and the scope of the weapons to be withdrawn. Participants discussed how to implement the arrangement while avoiding the appearance of changing the terms (US Department of State 1962a).

That is a different phase of the crisis. The task has moved from identifying an acceptable exchange to making it credible and executable. An alert system should be capable of representing that transition without declaring the danger over simply because a conciliatory message has arrived.

In mathematical terms, an agreement may identify a feasible endpoint while implementation still requires a path through intermediate constraints. The distinction between endpoint and path is the same one that appears in economic decoupling, resource substitution, and institutional construction.

An agreement can also be asymmetric in time. One side may act immediately while another acts later. The delay can make a settlement politically possible, but it creates a commitment problem. The structure of assurance and verification matters during that interval.

A two-superpower narrative can turn Cuba into a location on the map rather than an actor whose security and sovereignty were directly involved. The same problem arises whenever a dyadic model compresses a regional crisis into the relationship between its strongest participants.

The boundary of the model should follow the mechanism being examined. A study of superpower bargaining can focus on Washington and Moscow, but it must identify what that choice leaves outside. A study of settlement legitimacy or implementation needs a wider account of affected parties and their capacity to accept or resist an arrangement.

The distinction is both analytical and moral. A path that is acceptable to the strongest actors may impose risks on others. Describing it as a solution without specifying whose problem it solves can conceal a new source of instability.

It is easy to narrate a crisis that ended peacefully as a demonstration that the right decision was obvious. The participants did not have the ending. They acted under uncertainty about intentions, control, and the consequences of their own moves.

The historical record can identify alternatives that were considered and mechanisms that operated. It cannot rerun every rejected branch. Claims that a different choice would certainly have produced war, or

certainly have ended the crisis sooner, require more than confidence in the eventual outcome.

For the Davis Framework, the strongest lesson is about observable changes in the feasible set. A proposal became more specific. A channel conveyed a commitment that public positions did not fully reveal. An agreement created implementation tasks instead of eliminating all uncertainty. These changes are more informative than a retrospective assertion that diplomacy simply "worked."

They also suggest a way to compare cases. Look for crises in which proposals lacked authority, private commitments could not be implemented, or an intermediate action eliminated the remaining bargaining room. The comparison tests the mechanism rather than merely contrasting a peaceful ending with a violent one.

A framework concerned with warning should also recognize the beginnings of recovery. Those beginnings may be modest: a bounded exchange, a credible message, a verified first step, or a delay that allows another decision to be made.

Recovery does not require the parties to become friends. It requires the relationship to regain enough capacity to handle the immediate dispute without exhausting all acceptable alternatives. That capacity can be temporary. It can be specific to one issue. It can still save lives.

The crisis offers a demanding counterpart to the resource horizon in the US–Japan case. In both, time mattered, and actions could narrow future choices. The pathways differed. The comparison directs attention toward the conditions under which leaders can use time to construct an agreement rather than treat delay as an intolerable loss.

An exit is more than an imaginable peaceful outcome. It is a sequence of commitments and actions that the relevant parties can authorize, communicate, verify, and implement before the available time runs out.

That definition gives prevention something concrete to build. It also gives the historian something concrete to investigate.

The agreement matters because it became a sequence of actions the parties could carry out. Looking ahead to Taiwan, that distinction remains demanding. Commercial ties are visible; the conditions under which a crisis could be resolved have to be examined just as closely.



U.S. Navy sailors aboard USS Mustin (DDG-89) observe the Chinese aircraft carrier Liaoning operating in the Philippine Sea. April 2021.

Photo: U.S. Navy. Public Domain.

Chapter 12

China-Taiwan: The Active Watch

> The supreme art of war is to subdue the enemy without fighting.
>
> Sun Tzu, *The Art of War*

A shipment across the Taiwan Strait is a commercial transaction with a political setting. The customer may need the component; the supplier may need the revenue. Their continued exchange tells us that the connection is valuable. To learn whether it is becoming dangerous, we must ask how difficult it would be to replace and what each side expects the other to do.

The earlier cases allow us to look backward from an outcome. Taiwan asks what an analyst can responsibly watch before an outcome is known. The chapter's historical assessment places the dyad at Level 3, DIVERGENCE, in its 2024 specification. The question is which changes would justify a different reading. Later-published research helps examine the mechanisms; it does not silently turn that dated assessment into a new live forecast.

The China–Taiwan relationship joins substantial commercial integration to unresolved political authority. Its importance extends far beyond bilateral trade because other economies depend on production concentrated in Taiwan. The framework's position is that dependence, political orientation, and usable alternatives must be examined together. The evidence becomes compelling when it identifies the connection, its function, and the time required to replace it.

The Taiwan question emerged from civil war and competing claims to political authority. In 1949 the Republic of China government relocated to Taiwan while the People's Republic of China was established on the mainland. Taiwan subsequently developed its own democratic institutions; the PRC has never governed it. The historical dispute therefore contains both a claim about territorial unity and a population's lived political arrangements. Those are not interchangeable observations (CRS 2022).

The PRC constitution's preamble places Taiwan within its territorial claim and makes unification a national obligation (National People's Congress 2018). Taiwan's separate political institutions and population give the dispute a reality that cannot be settled by treating one government's claim as a description accepted by all sides. Economic integration operates across that unresolved political boundary.

The United States shifted diplomatic recognition from the ROC to the PRC in 1979 while maintaining unofficial relations with Taiwan. Its "one-China" policy is distinct from Beijing's "one China principle," and both differ from the disputed 1992 Consensus. The Taiwan Relations Act links the diplomatic settlement to peaceful resolution, provides for defensive arms, and treats coercion, including boycotts and embargoes, as a matter of grave concern. It does not provide an automatic defense treaty (CRS 2025; United States Congress 1979). Deliberate uncertainty about a military response sits beside these explicit commitments.

Taiwan's development as a manufacturing and technological center placed firms, investment, and family livelihoods on both sides of the strait. The commercial network is wider than one customs total. A Taiwanese company can produce in the mainland, sell to an American buyer, and depend on a machine built in Europe. National ownership, factory location, invoice destination, and final consumption describe different connections. An analyst must choose the one relevant to the proposed mechanism.

This integration was not symmetrical. Taiwan's Ministry of Finance records mainland China and Hong Kong together receiving 43.9% of Taiwan's exports in 2020 and 31.7% in 2024. The 2024 value was $150.6 billion (Taiwan Ministry of Finance 2025). Those figures describe an export destination, not the share of Taiwanese value ultimately consumed there. China's exposure to particular Taiwanese components raises a different, product-specific question.

The technological dimension deserves extended analysis, because it represents a concentrated production network with consequences extending beyond the dyad.

The concentration is real, but its denominator matters. SIA's 2021 supply-chain submission placed 92% of the world's logic manufacturing capacity below ten nanometers in Taiwan, with the remaining 8% in South Korea (SIA 2021). It is a geographical capacity estimate for a defined technology class. TSMC's share of foundry revenue, Taiwan's share of installed capacity, and the share of a particular customer's qualified supply are three different measures. A geopolitical model becomes less informative when it treats them as synonyms.

Dimension	Dated evidence and its meaning
Trade	Mainland China and Hong Kong received $150.6B, or 31.7% of Taiwan's exports, in 2024.
Technology	A 2021 SIA study located 92% of sub-10nm logic manufacturing capacity in Taiwan. This is a dated capacity category.
Supply chain	Rhodium's 2022 blockade scenario placed over $2T of activity at risk; this is not an estimated GDP loss.
Production	TSMC reports Arizona N4 and Kumamoto volume production beginning in late 2024. Product and location both matter.
Energy	Imported energy supplied 95.77% of Taiwan's total energy supply in 2024. Supplier origin and maritime access are different exposures.
Political	Institutionalized cross-strait dialogue and the conditions for its operation remain disputed.

Table 12.1: Dated evidence for the relationship. Sources: Taiwan Ministry of Finance, SIA, Rhodium Group, TSMC, Taiwan Energy Administration, and CRS (Taiwan Ministry of Finance 2025; SIA 2021; Rhodium Group 2022; TSMC 2025; Taiwan Energy Administration 2026; CRS 2022). The rows measure different things and should not be added into one exposure total.

Advanced fabrication also depends on equipment that the foundry does not make. ASML's 2024 annual report identifies the company as the only manufacturer of EUV lithography systems and explains how EUV reduces the complexity of multiple-patterning steps otherwise performed with deep-ultraviolet equipment (ASML 2025). The bottleneck is a production network, not one isolated building. Installed tools require materials, service, power, qualified processes, and people able to operate them.

The strategic importance follows from the functions these components perform. Computing, communications, sensing, and control depend on different kinds of chips. Advanced weapons do not all require the newest process node, and the newest node is not the only scarce input. SIA's supply-chain analysis separates design, manufacturing equipment, materials, fabrication, assembly, and testing (SIA 2021). A component's strategic significance therefore comes from the system that cannot operate without it, its available substitutes, and the cost of interruption.

The risk has been publicly debated for years. Admiral Philip Davidson's March 2021 Senate testimony placed the threat to Taiwan within a

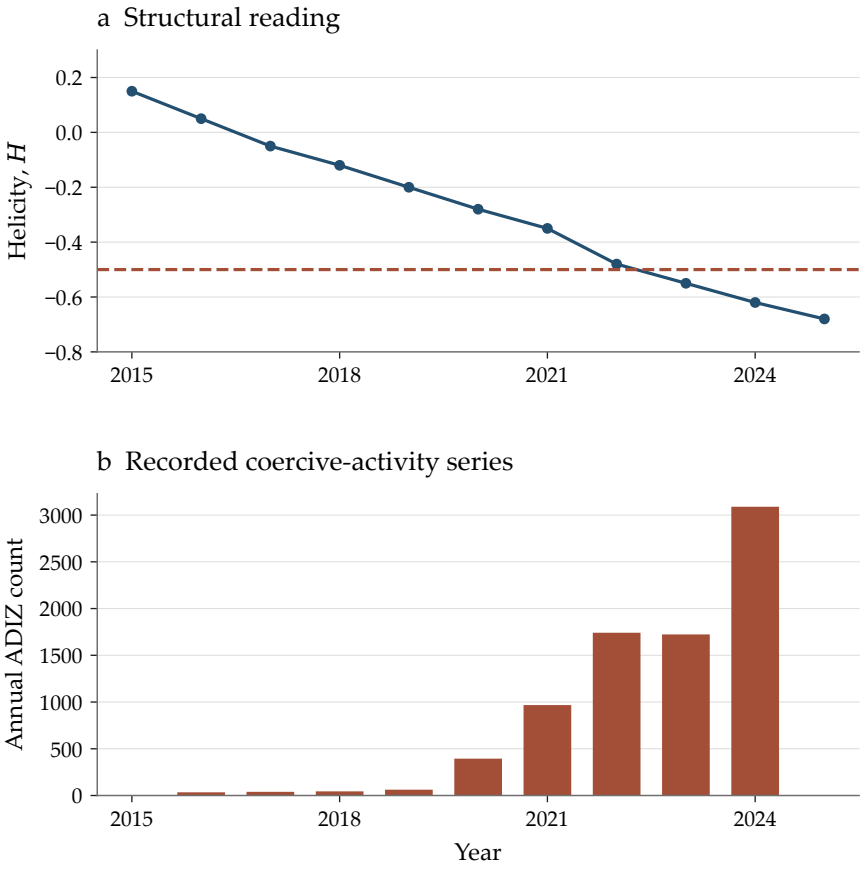


Figure 12.1: The retained author helicity series and annual ADIZ activity use separate scales. The 2024 ADIZ count is 3,075, following the CSIS compilation (Hart 2025). The upper reference is $H = -0.5$; visual calendar alignment does not establish causation.

six-year horizon (United States Senate Committee on Armed Services 2021). Such an assessment makes preparation urgent, but it is not a dated invasion order. The distinction between capability, political intention, and a decision to act is essential to the case.



U.S. Navy personnel conduct operations in the Western Pacific, September 2012. The routine military presence that maintains the status quo across the Taiwan Strait— a geometric boundary held in place by the persistent application of force.

Photo: U.S. Navy. Public Domain.

The silicon-shield argument treats Taiwan's semiconductor industry as a reason for other states to care about its security. The mechanism is conditional. Production must remain usable, customers must expect continued access, and the incentives created by dependence must favor restraint. A factory's economic importance does not physically protect it from attack.

Advanced semiconductor production is embedded in equipment, materials, design, energy, and skilled-labor networks. Control of a factory is therefore not equivalent to control of a functioning supply chain. The wider network belongs in the strategic account.

The 2016 election of Tsai Ing-wen and the Democratic Progressive Party changed the political setting for dialogue. The disputed 1992 Consensus and the terms Beijing required for official communication were central to that disagreement (CRS 2022).

The political disagreement after Tsai's election concerned the terms on which dialogue could continue, including Beijing's insistence on its interpretation of the 1992 Consensus. A channel can be physically available while its proposed conditions are politically unacceptable to the other side.

CSIS's compilation records 3,075 Chinese military-aircraft flights into Taiwan's self-declared ADIZ in 2024, an increase of 81% over 2023 (Hart 2025). An ADIZ is distinct from sovereign airspace. The count documents a category of activity; it does not count attempted invasions or independently determine diplomatic alignment. Exercise duration, stated purpose, location, and accompanying measures supply the context that an annual total cannot.

Lai Ching-te's 2024 election belongs to this evolving political setting. An election result, the incoming administration's policies, and Beijing's response are separate observations. Combining them into a single sentiment value can hide the sequence in which each side interprets the other's actions.

Reading the relationship

In the chapter's 2024 assessment:

$$H_{\text{CN-TW}} = -0.62 \quad \lambda_1 = 0.071 \quad \Gamma = 0.97 \tag{12.1}$$

Parameter	Value	Interpretation
Helicity	-0.62	Negative but not extreme. Economic flows and political orientation are misaligned.
Spectral Gap	0.071	**Above** critical threshold. Regional system retains flexibility.
Trichotomy	0.97	Near-critical regime. System is at a decision point.
Alert Level	**L3**	DIVERGENCE

Follow the three 2024 readings through the historical rule. The signed reading of -0.62 is below -0.5, while the network gap of 0.071 is 0.021 above 0.05. A rule requiring both crossings therefore leaves this case short of Level 4. The 0.97 ratio sits close to one. A chip buyer makes the distinction tangible. An alternative trading partner may exist in the region, but can that partner manufacture the exact chip when the buyer needs it? The regional gap and that buyer's exposure answer different questions.

The Spectral Gap Remains Open

Unlike the pre-2022 Europe-Russia network, the Asia-Pacific trade network retains significant flexibility. The spectral gap is 0.071—above the critical 0.05 threshold. Alternative trading relationships remain viable for multiple actors; Japan, South Korea, and ASEAN states maintain substantial ties with both China and the United States. The system is not frozen into a binary configuration.

The comparison with pre-1914 Europe has to be made through specified connections. Supplier diversity, security commitments, and regional trading partners may create alternatives, but each has to be usable on the timescale of a crisis. A spectral calculation tests one representation of those connections.

The Semiconductor Shield

Taiwan's semiconductor position creates the proposed "silicon shield": disruption would impose costs on customers and on actors seeking to change Taiwan's status. The protection depends on those actors expecting the costs to exceed the gains. It also depends on continued access to production. A functioning foundry and a damaged or isolated foundry have different strategic value. The relevant replacement time is product-specific: constructing a fab, achieving acceptable yields, and qualifying a customer's design are separate milestones.

Rhodium's December 2022 study illustrates the scale without treating it as certainty. Its hypothetical blockade halts Taiwan's trade and places more than $2 trillion of economic activity at risk before adding sanctions

or further military escalation. The authors explicitly distinguish activity at risk from lost GDP (Rhodium Group 2022). The finding supports the claim that disruption would be costly. It cannot establish that decision makers will always choose restraint.

The shield also changes as production and policy change. The October 2022 U.S. controls targeted specified advanced-computing items, supercomputer uses, manufacturing equipment, and related support; their scope was not a ban on every semiconductor (BIS 2022). Restrictions may reduce access to some technologies while increasing incentives to develop substitutes. The framework must track the resulting capabilities and dependencies rather than assume that announced investment has already removed a bottleneck.

A Different Energy Exposure

Taiwan does not reproduce the Russia–Ukraine gas-transit relationship. Its energy vulnerability nevertheless remains substantial: the Energy Administration reports that imports supplied 95.77% of total energy in 2024 (Taiwan Energy Administration 2026). Buying from suppliers other than mainland China does not ensure access if deliveries depend on disrupted maritime routes. Supplier diversification and transport resilience are separate conditions. A graph that records only the seller can miss the common route on which several sellers depend.

Deliberate Decoupling

Taiwan's direct export concentration has been falling: the mainland China and Hong Kong share declined by 12.2 percentage points between 2020 and 2024 (Taiwan Ministry of Finance 2025). This is consistent with a possible reduction in exposure, but it is not sufficient evidence that every dependency weakened. The numerator, the total export denominator, the product mix, and indirect sales through other economies can move differently. The framework's proposed helicity correction requires usable alternatives, not simply a smaller aggregate percentage.

Diversification can reduce vulnerability while imposing immediate costs. Workers, investors, and customers may depend on the existing relationship long before a replacement is ready. A managed transition has to account for that interval.

Diversification changes the consequences of a disruption. If an exporter gains another customer, the loss of one market may become more bearable. If a manufacturer qualifies another supplier, a restriction may become less coercive. Whether either change lowers conflict risk depends on the political response as well as the commercial replacement.

Russia–Ukraine makes it unsafe to infer that a costly war cannot be chosen. The question is how leaders assess those costs alongside security,

legitimacy, and territorial objectives. Taiwan requires its own evidence, rather than a conclusion imported from another war.

Military preparation is another part of that evidence. Exercises can improve capability, signal resolve, or rehearse a particular operation. Their frequency alone does not establish a decision to invade. A structural account should be compared with the interpretations available from military observation.

Several developments could change the reading. Each requires observations of its consequences before the historical Level 4 rule can be reapplied:

Development	Mechanism to examine	Evidence required
Chip restrictions	Reduced access and induced substitution	Product coverage, deliveries, qualified alternatives
Status confrontation	Changed political terms and responses	Dated commitments, actions, channel availability
Shipping coercion	Disrupted civilian access	Actual throughput, delays, reserves; distinguish quarantine and blockade
Broad decoupling	Lost connections and replacement paths	Rebuilt graph, same units, transition costs and timing

Table 12.2: Potential changes to the China–Taiwan assessment. A political event does not by itself determine a numerical helicity or spectral gap.

The trouble with a date

Analysts, policymakers, and journalists frequently ask: "When will China invade Taiwan?" The question reflects a desire for certainty that no honest framework can provide. But it misframes the analytical problem.

The "when" question assumes that invasion is a foregone conclusion requiring only timing. This assumption is unwarranted. Chinese leaders face a genuine choice—one shaped by structural conditions but not determined by them.

Consider the factors that make invasion uncertain:

Military risk. Crossing a strait, sustaining an operation, and controlling territory pose different demands from threatening shipping or conducting exercises. CSIS's quarantine study examines coercive possibilities below a full invasion (Lin et al. 2024). A framework concerned with disappearing alternatives should examine these distinct outcomes rather than treat military pressure as a single countdown.

Economic cost: A conflict that damages TSMC's facilities would devastate China's own technology sector—the chips that power Chinese smartphones, data centers, and weapons systems come from Taiwan. Sanctions would compound the damage. China's economy, already struggling with property sector crisis and demographic decline, cannot easily absorb such shocks.

The semiconductor concentration described earlier makes an interruption consequential beyond the two sides. A finished chip requires facilities, equipment, materials, skilled staff, and functioning delivery arrangements. The dependence matters because these cannot all be replaced at the speed of a political decision.

International isolation. Sanctions and disrupted access can impose substantial costs. China's extensive international connections create exposures different from Russia's. The scale of those exposures does not by itself establish how Chinese decision makers would weigh them in a Taiwan crisis.

The sanctions imposed after Russia's invasion also create an opportunity for other states to adapt. They can build reserves, diversify partners, or seek alternative channels. A forecast should therefore examine what has changed since the comparison case, as well as what resembles it.

Uncertainty about U.S. response. Strategic ambiguity concerns the response to a particular contingency. The Taiwan Relations Act's obligations and the absence of an automatic defense treaty do not tell us precisely how Washington would respond to every kind of coercion (United States Congress 1979; CRS 2025). Decision makers must consider both what is promised and what they believe will be done. Those beliefs can change even while the legal text remains the same.

The framework does not assign an inevitable date for action. In the chapter's 2024 specification, the open regional gap, continuing semiconductor dependence, and possibilities for diversification leave alternatives worth examining. That structural reading does not imply the absence of coercion, and it does not update itself as conditions change. Its value depends on retaining the observations and assumptions that produced it.

The military feasibility of an operation and the willingness to accept its costs are separate questions. High interdependence can make disruption expensive while also creating coercive leverage. The proposed helicity diagnosis concerns that tension and needs to be evaluated alongside the military and political evidence.

The 2024 assessment

The China–Taiwan assessment leaves room for policy to change the trajectory. The same is true of any historical configuration assigned

Level 4. A higher warning category increases the urgency of examining constraints, rather than proving that political choices have ceased to matter.

The case's regional spectral reading leaves room for adjustment in the modeled network. Semiconductor dependence also creates costs for disruption. Whether those costs deter an attack depends on how decision makers weigh them against political and security objectives. They constrain the decision without deciding it.

The adverse continuation. In the scenario developed here, alignment deteriorates while important dependencies persist. A new restriction can narrow access, but it can also induce substitution. Its spectral effect has to be calculated from the resulting graph rather than assumed from the policy's description.

The relevant deadline is the time needed to create a usable alternative before a particular constraint becomes binding. That is something to estimate and revise, not a clock that must move toward war every year.

The year 2027 became prominent through the six-year horizon in Davidson's 2021 testimony. It is worth examining what that warning can and cannot tell a reader.

On March 9, 2021, Davidson told the Senate Armed Services Committee that the threat to Taiwan could become manifest within the next six years (United States Senate Committee on Armed Services 2021, pp. 47–48). His answer connected developing military capabilities with concern about China's ambitions. Six years from that hearing points to 2027.

That warning concerned a threat horizon. It did not establish that China's leadership had chosen a date to invade. A military capability, an intention to acquire a capability, and a decision to use it are different propositions.

A calendar year can make an uncertain assessment seem more definite than it is. For a planner, a readiness horizon means that preparations cannot be deferred indefinitely. For a forecaster, it still leaves the probability, form, and timing of an actual attack to be assessed.

Burns distinguished Xi's reported 2027 readiness instruction from a decision to invade (Burns 2023b).

That distinction can be followed through an ordinary planning question. Building a bridge gives an army a route across a river. It does not tell an observer whether the route will be used, what objective it would serve, or which political decision would authorize the crossing. Readiness changes the options; it does not settle the choice.

A 2024 CSIS expert survey provides useful perspective:

The CSIS survey, conducted in late 2023 with 52 US and 35 Taiwan experts, asked a more specific question about the significance of China's 2027 military modernization goals. Only 13% of US respondents and 6% of Taiwan respondents thought those goals would affect Beijing's decision to quarantine, blockade, or invade. Expectations of US intervention also depended on the scenario and differed between the two groups (CSIS China Power Project 2024). A single claim about what all experts believe conceals those distinctions.

The book's constructed continuation starts from its stated 2024 assessment: $H = -0.62$ and $\lambda_1 = 0.071$. The alignment condition is already below -0.5. Under the historical joint rule, the remaining crossing is the network gap falling below 0.05. The scenario window of *2028–2032* therefore has a precise conditional meaning: it is the range produced by specified rates of decline in that gap, while adverse alignment persists.

For a constant assumed annual decline $r > 0$, write $\lambda_1(t) = 0.071 - r(t - 2024)$. Equality with 0.05 occurs at

$$t_* = 2024 + \frac{0.021}{r}.$$

The assumed rates 0.00525, 0.00350, and 0.002625 per year give 2028, 2030, and 2032 respectively. A strict below-threshold rule becomes active just after those equality times, or at the next qualifying observation on a discrete schedule. If the gap remains at 0.071, this second crossing does not occur. The plotted alignment line declines by an assumed 0.015 per year; it remains below its threshold throughout. These rates define scenarios, not fitted forecasts or confidence limits. The dated $\Gamma = 0.97$ reading is not given an invented future path.

The continuation is a conditional scenario: if the prescribed declines occur, the plotted readings reach the indicated region. Testing a forecast requires an issuance record and a stated event, horizon, and scoring rule. A scenario instead asks how consequences change when its assumptions change. Keeping those tasks distinct allows both to do useful work.

Several changes could move the case toward the warning conditions: deeper restrictions on usable trade, reduced regional alternatives, a political confrontation, or a change in the expected cost of semiconductor interruption. Their effects must be calculated in the specified network. A political announcement does not itself supply a new eigenvalue, and semiconductor self-sufficiency is a set of achieved production capabilities rather than a single switch.

Movement in the other direction would require a change in what the parties can live with. A negotiated update to the cross-strait status quo or managed U.S.-China competition with explicit guardrails could



USS Cowpens (CG-63) conducts operations in the Pacific, September 2012. Guided-missile cruisers like the Cowpens form the backbone of the naval architecture that defines the Taiwan Strait's current topology—a boundary maintained not by geography alone, but by the continuous circulation of military force.

Photo: U.S. Navy. Public Domain.

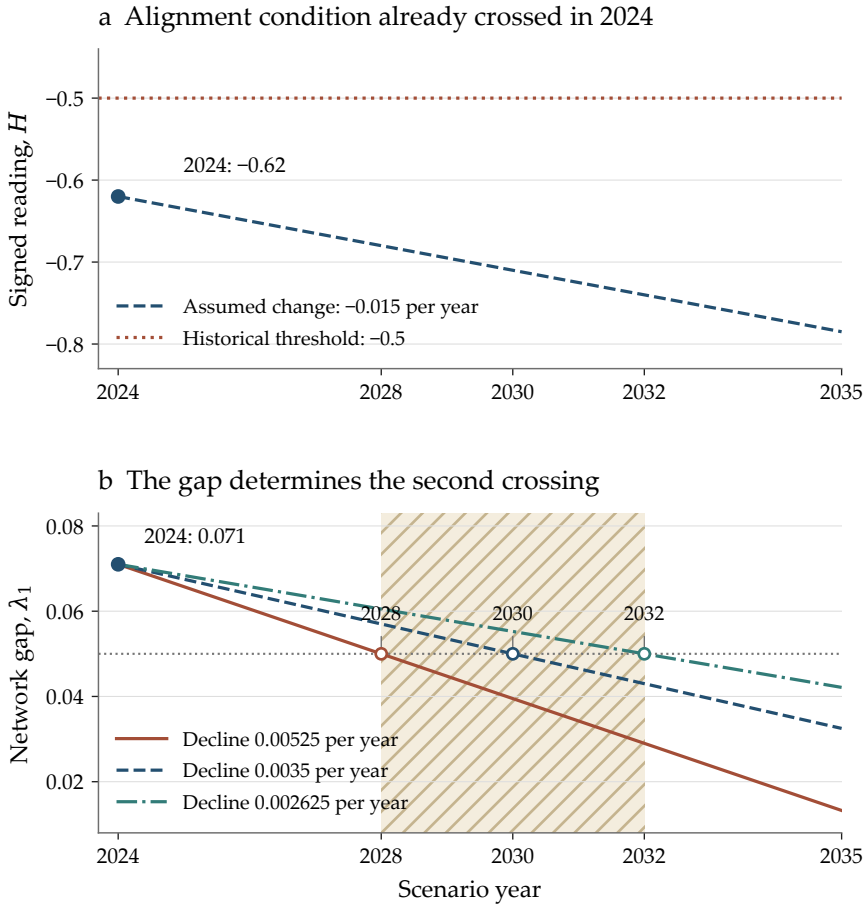


Figure 12.2: Conditional continuations from the chapter's 2024 readings. The alignment condition is already crossed. Three prescribed annual gap declines reach 0.05 in 2028, 2030, or 2032. Hatching marks that sensitivity window, not a probability interval. Threshold equality and a strict crossing are distinguished in the text.

reduce the pressure. A more accommodating Chinese leadership could alter the political terms. Semiconductor diversification would change Taiwan's leverage and the structure of dependence. Each possibility works through a different part of the relationship, so none should be represented simply as a lower score.

A continuing watch would track the following observations with their release dates:

Helicity drivers (economic-diplomatic alignment):

The relationship is visible in ordinary movements of people, goods, and capital. Cargo volumes and completed deliveries show whether exchange continues; direct flights and passenger flows record other channels. Taiwanese investment in mainland production identifies longer commitments. Student and cultural exchanges can show whether human connections are narrowing alongside commerce. Each series needs a consistent coverage rule and a release date before its movement can become evidence in a forecast.

Spectral gap drivers (network flexibility):

Alternative routes matter as much as the principal route. Trade diversification toward ASEAN by Taiwan and China, India's role in "China+1" supply chains, and expanding manufacturing capacity in Vietnam and Thailand can change which disruptions the network can absorb. U.S. reshoring under the CHIPS Act works on another part of that same problem. The relevant observation is whether these investments create usable substitutes before a crisis closes an existing channel.

Early warning signs (approaching Level 4):

A sharper warning would appear if several channels contracted together: Taiwanese FDI to the mainland falling further, restrictions on Taiwan-origin chip imports, and direct flights being cancelled or severely curtailed. Military activity would then need to be read alongside the economic changes. ADIZ incursions exceeding the documented 2024 count of 3,075, or exercises extending beyond historical norms in duration or scope, would add to the warning. The point of monitoring them together is to distinguish a temporary demonstration from a sustained change in the relationship.

The connection that a trade total cannot show

A manufacturer has an order to fill on Friday. Its usual supplier can produce the component, but the shipment has not left. Another supplier appears in the procurement database. That supplier uses a different process, has no spare capacity this month, and has not qualified the

customer's design. The spreadsheet contains two names. The factory has one usable source.

This is where the geometry acquires practical meaning. A connection must perform the required function at the required time. Counting trading partners without testing those conditions can make a trapped production line look diversified. The question is not whether another country makes semiconductors. It is whether the required design can be produced, packaged, tested, delivered, and used before the customer's operation stops.

TSMC's 2024 report provides a concrete reason to distinguish stages. Arizona's first fab entered N4 high-volume production in the fourth quarter of 2024, and the first Kumamoto fab began volume production at the end of that year. The same report describes continued investment in advanced processes and packaging in Taiwan (TSMC 2025). Overseas production is an implemented change in the network. It is not evidence that every Taiwan-produced component already has an interchangeable offshore substitute.

For a particular product, imagine a simplified serial chain whose stages have capacities $q_1, \ldots, q_m$, all expressed as deliverable units of that product per day. If every stage is necessary and there are no buffers or alternative routes, its sustainable rate is

$$q_{\text{chain}} = \min_j q_j.$$

This is an illustrative engineering constraint, not a measured model of TSMC. Suppose fabrication can support 100 units per day, packaging 70, and testing 90. The completed output cannot exceed 70. Increasing fabrication to 140 leaves that bottleneck untouched. If another packaging route genuinely adds 25 compatible units, completed output can rise to 90, where testing becomes limiting. The useful intervention changes the constraining stage.

A graph of money flows and a graph of productive dependence therefore need not have the same important edges. A low-cost chemical or machine service can be essential to a much more valuable output. Replacing its value with its functional significance changes the model's question. The weights must be defined accordingly. Dollars are appropriate for some exposures; they do not automatically represent replacement time or production necessity.

For the reader assessing a diversification announcement, the corresponding test is concrete. Name the product, its qualification requirements, the stage being replaced, the quantity available, and the first date that quantity can arrive. Then ask which prerequisites the original and

substitute routes share. Two factories relying on the same packaging line, specialized service, or shipping permission may fail together. Finally, distinguish a signed capacity reservation from an aspiration to build capacity. These questions explain why a trade share can fall while a critical dependence persists. They also identify what would support the opposite conclusion: demonstrated replacement at the required scale and time. The framework needs that evidence if its account of narrowing alternatives is to remain open to challenge.

Nor should we mistake a node's location for complete control. Ownership of a factory does not guarantee access to its equipment service, design files, specialist materials, staff, power, and customers. This is why the silicon shield is a network proposition. Its strongest form asks whether actors expect disruption to destroy access they value, and whether that expectation restrains their choice. Its weakest form merely declares a valuable factory invulnerable. The former can be investigated; the latter cannot do the analytical work.

The author's later China–Taiwan watch notes return to semiconductor substitution, continuing exchange, and political tolerance as separate conditions (Davis 2026g). Their scenario questions are useful here. Their version-specific scores must remain attached to their own inputs and definitions. They cannot be joined to the chapter's 2024 spectral reading as though one continuous, identically normalized series had already been established.

Pressure can arrive before the shooting

An importer can lose access while a port remains physically intact. A carrier suspends a sailing. An insurer changes the terms. A bank declines a transaction. The commercial participants are responding to expected danger, and their responses can reduce the service that is actually available. The mechanism is already present before anyone announces a complete severing of trade.

CSIS's June 2024 quarantine study distinguishes a coercive action framed around law enforcement and shipping compliance from a military blockade or full invasion (Lin et al. 2024). That distinction matters for the book's forecast target. A model evaluated on interstate war onset cannot claim the same kind of success merely because an inspection dispute occurs. It can, however, investigate whether pressure on civilian access narrows the choices available to businesses and governments.

The scenarios also make a wider analytical point: the action's declared purpose, its practical effects, and the way other parties classify it can differ. A government may describe a restriction as temporary regulation.

A shipping company may treat the same measure as an unacceptable exposure. A resident experiences the result as delayed medicine, higher prices, or lost work. One description cannot substitute for the others.

The relevant states might therefore include ordinary access, disrupted access, prolonged coercive restriction, limited armed conflict, and invasion. They are alternatives to distinguish, not a compulsory staircase that every crisis must climb. A particular path can stop, reverse, or change direction. For each target, the analyst needs an onset rule and a time window. The same episode can satisfy one definition and fail another without the data being inconsistent.

CSIS's separate work with private-sector participants asks how firms would respond to different Taiwan crisis scenarios (Blanchette, DiPippo, and Johnstone 2023). This complements the framework's attention to feasible movement. Commercial expectations are part of the mechanism, not an inert background held fixed while governments act. Yet a scenario exercise remains an investigation of possible responses. Its participants' judgments are not observed outcomes of a war that has not occurred.

Consider a distributor with 120 units in inventory. Its customers need ten units each day. Under ordinary conditions, deliveries replenish ten. During a disruption, only six arrive. If demand and arrivals stay constant, the stock equation is

$$I_{t+1} = I_t + 6 - 10 = I_t - 4, \qquad I_0 = 120.$$

The inventory lasts thirty days. An alternative supplier that begins on day forty is too late by itself. One that begins on day twenty can help only if its output is sufficient and compatible. This calculation is constructed; it does not estimate Taiwan's strategic reserves. It explains why an analyst must state both the reserve and the shortfall before describing how much time remains.

The same arithmetic also exposes a common misunderstanding. Twelve days of inventory at normal consumption does not always mean exhaustion twelve days into a partial interruption. Some supply may continue, demand may fall, or stock may be unusable for the affected product. Conversely, a large national stock may not be deliverable to the place that needs it. The relevant reserve is accessible inventory for the specified function.

These are civilian consequences that a structural model should keep visible. An economic relationship is not simply an edge that can be erased without residue. Its interruption moves through schedules, wages, inventories, care, and public services. A model of aggregate losses needs

a way to recover those unequal consequences if it is going to inform policy rather than merely dramatize a total.

What would count as a stronger case

The claim that alternatives are disappearing becomes stronger when independently observed constraints converge. A substitute remains unqualified; deliveries become less reliable; a communication channel cannot resolve the disagreement; an agreement's promised implementation fails. These observations concern different parts of the same operational problem. They are more informative together than four repetitions of one newspaper report.

They should also be able to contradict the proposed diagnosis. If substitute production comes online, customers qualify it, shipments continue through independent routes, and practical dialogue resolves disputes, the feasible set may expand despite hostile public language. The framework should register the change. A theory that counts every hostile statement as support and treats every adaptation as irrelevant has made itself too easy to confirm.

The regional spectral calculation belongs inside that discipline. Its graph must identify actors, edge meanings, time aggregation, and normalization. The gap can describe a slow mode of adjustment under a specified evolution. It cannot certify that every indispensable product has a substitute. The product-chain and inventory calculations examine different constraints, and their agreement or disagreement can reveal what the aggregate graph has missed.

A useful record would therefore keep three descriptions connected: the dated structural reading, the observed implementation of alternatives, and the explicitly defined outcome. The structural reading says why concern is justified. The implementation record says whether the proposed exit is becoming real. The outcome rule makes the forecast accountable. None can be reconstructed merely by reading the last score after a crisis has ended.

That is the strongest version of the book's position. Commercial value alone does not guarantee peace; political hostility alone does not make war inevitable. What matters is the relation between commitments, dependence, expectations, and the adjustments that people can actually carry out. Taiwan gives that argument concrete content precisely because its connections are valuable, uneven, and difficult to replace on demand.

The choices that remain

The framework suggests a counterintuitive policy prescription:

The configuration of concern is high dependence with collapsed trust and few usable exits. The policy developed here is to *reduce dangerous concentrations at a controlled pace while maintaining functional channels of communication*. Whether it improves the situation depends on the transition it actually creates.

An orderly transition aims to preserve usable alternatives while reducing dangerous concentrations. It does not mathematically guarantee that a spectral gap increases: removing edges from a fixed nonnegative combinatorial Laplacian cannot increase its second eigenvalue. The policy question is whether replacement connections and functioning channels compensate for what is lost. The matrices, their units, and their time horizon must show the proposed gain.

This is uncomfortable advice. It suggests that some degree of US-China economic separation is *stabilizing* rather than destabilizing—provided it is managed gradually rather than imposed suddenly.

The pace and sequencing of diversification matter. A substitute available in five years does not cover an interruption next winter. Phasing commitments, maintaining transition capacity, and explaining the intended change can affect whether a reduction in exposure is manageable or destabilizing.

The Russia-Ukraine case demonstrates what happens when decoupling is *forced* by conflict. The goal of policy should be to enable *chosen* decoupling before the trap becomes inescapable.

Managed reduction in exposure has political costs of its own. Businesses can lose markets, workers can lose jobs, and a restriction can be read abroad as a threat. An intervention has to account for those consequences as carefully as it measures the dependency it seeks to reduce.

Strategic ambiguity adds another layer. The parties may deliberately leave parts of their future response unspecified. An economic intervention can change how that uncertainty is interpreted, so its effects must be considered alongside security commitments and communication.

The framework doesn't dictate which path to choose. It reveals the geometry of the choices and their consequences. Leaders who understand that geometry can navigate deliberately. Those who don't will discover the shape of the trap only when it closes.

For a business that spent decades building production and customer relationships across the strait, diversification is not a simple line in a strategy paper. It means moving orders, equipment, skills, and

sometimes livelihoods. A policy that ignores those costs will struggle to create the alternatives on which its strategic promise depends.

This is the human cost of structural change: real people with real investments in relationships that the geometry now suggests were dangerous all along. The framework can identify the trap. It cannot ease the pain of those caught inside.

The next case removes the commercial intimacy that makes Taiwan so consequential. It asks whether the network can still transmit danger when the two adversaries barely trade with one another at all.



Remnants of an Iranian ballistic missile recovered in Israel following the October 1, 2024 attack— the largest direct Iranian strike on Israeli territory. Over 180 missiles were launched; most were intercepted.

Photo: PikiWiki Israel, CC BY 2.5.

Chapter 13

Israel-Iran: The Slow Escalation

> Eternal vigilance is the price of liberty.
>
> Wendell Phillips, 1852

Remove the direct trading relationship from the picture. The danger does not necessarily disappear with it. An adversary may still reach another country through an armed partner, a regional corridor, a sponsor, or a threatened neighbor. The line between two capitals is only one part of the network.

The Israel-Iran dyad presents a different configuration than the cases examined so far. It is not a trap in the classic sense—there is minimal bilateral economic interdependence. Yet the framework still has predictive value through its treatment of network effects and proxy dynamics.

The Israel-Iran relationship inverts the typical trap configuration. Rather than high interdependence with collapsed trust, it features *minimal* interdependence combined with maximal hostility—a different geometry requiring different analytical tools. Understanding this dyad requires tracing how it evolved from alliance to enmity.

Israel and Iran were strategic partners under the Shah. Oil, security cooperation, and shared regional concerns sustained their relationship. The subsequent reversal shows how political change can transform the meaning of connections that previously supported cooperation.

The 1979 revolution ended that partnership and diplomatic relations. The revolutionary government opposed Israel, and support for Palestinian armed groups became part of the regional confrontation. The analysis has to follow the changes in institutions, commitments, and means of pressure.

The ensuing hostility was expressed through rhetoric, proxy relationships, covert operations, and competing military preparations. Later direct exchanges of fire changed the operational relationship. Distinguishing those forms of conflict is necessary before assigning a forecast target.

Action through intermediaries changes both responsibility and exposure. It can permit pressure without an immediate direct exchange, while creating escalation pathways that are harder to control. The intermediary has interests and constraints of its own.

Dimension	Status
Trade	Near-zero direct bilateral trade since 1979
Energy	No direct relationship; no pipelines, no fuel trade
Diplomatic	No relations since Islamic Revolution; mutual non-recognition
Military	Decades of proxy conflict (Hezbollah, Hamas, Syria, Yemen)
Network	Indirectly connected through Gulf states, Turkey, European trade

Table 13.1: Israel-Iran dyad profile.

Iran's regional strategy relied on a network of allied states and armed groups, often described as the Axis of Resistance. These relationships extended its influence beyond the places where Iranian forces operated directly. The network also distributed risk: a confrontation involving one member did not automatically have the same implications as an attack on Iran itself.

Hezbollah (Lebanon). Founded with Iranian support in 1982, Hezbollah became a major component of Iran's regional deterrent. The case account places its pre-2024 arsenal at roughly 150,000 rockets and missiles. That was an estimate of inventory, not a count of weapons that could necessarily be employed in a particular campaign. Its role in the argument is the potential retaliatory cost an Israeli planner had to consider.

Hamas (Gaza): Though ideologically distinct (Sunni, rooted in Muslim Brotherhood tradition), Hamas received Iranian funding, training, and weapons. Iranian support helped Hamas develop the military capabilities demonstrated on October 7, 2023.

Syrian Land Bridge: Iran's alliance with the Assad regime provided a geographic corridor connecting Tehran to Beirut. Weapons, money, and personnel flowed through this route, enabling the resupply and reinforcement of Hezbollah.

Houthi Movement (Yemen): Iran's support for Yemen's Houthi rebels created a third front, with Houthi missiles and drones threatening Saudi Arabia and, after October 2023, disrupting Red Sea shipping.



A Gazan child gazes at buildings destroyed during the 2023–2025 war. The human cost of structural collapse is not measured in spectral gaps or helicity values, but in the lives they represent.

Photo: Jaber Jehad Badwan. CC BY-SA 4.0.



A child carries food for his family during the Gaza war, 2023–2025. When topological structures collapse, civilians bear the costs that structural prediction aims to anticipate.

Photo: Jaber Jehad Badwan. CC BY-SA 4.0.

This proxy architecture served a strategic function: it allowed Iran to pressure Israel without exposing Iranian territory to retaliation. When Hezbollah attacked, Iran had plausible deniability. When Israel retaliated, the costs fell on Lebanon, Syria, or Gaza—not Iran.



Children from Gaza selling canned goods during the 2023–2025 war. Structural collapse forces the youngest into survival economies—the human geometry of a shattered manifold.

Photo: Jaber Jehad Badwan. CC BY-SA 4.0.

A proxy network distributes the burdens of confrontation unevenly. The actor supporting it may gain influence while local populations bear



The entrance to the emergency department of Shuhada al-Aqsa Hospital, Gaza Strip, 2024. Hospitals became both refuges and targets during the conflict—nodes in a network whose connectivity the framework measures.

Photo: Jaber Jehad Badwan. CC BY-SA 4.0.

much of the disruption. The network therefore needs to be represented through actual commitments and channels of support, rather than as a collection of interchangeable extensions of one state.

Such a network can supply leverage while also creating liabilities. A supporting state may face pressure to respond when a partner is attacked, and may have less control over that partner's choices than a simple diagram suggests. Those limits matter to the proposed measure of constraint.

The evolution from shadow war to direct confrontation marks a structural transformation in the dyad. The April 2024 exchanges—Iran's "Operation True Promise" launching approximately 350 projectiles at Israel, and Israel's subsequent strikes on Iranian air defenses near Isfahan—represented the first direct state-to-state military action from Iranian soil targeting Israeli territory.

The June 2025 "Twelve-Day War" escalated dramatically further. Following Israeli strikes on Iranian nuclear facilities at Fordow and Natanz, Iran responded with over 550 ballistic missiles and more than 1,000 drones. Israel's Iron Dome, David's Sling, and Arrow systems intercepted most incoming projectiles, but the scale of the exchange exceeded any previous direct confrontation.

What transformed a decades-long cold war into hot conflict? The framework points to *proxy degradation* as the critical factor.

The absence of bilateral interdependence might suggest that the helicity framework has nothing to say about Israel-Iran. This is incorrect, but the analysis requires adaptation.

The framework operates at two levels:

Dyadic: Direct bilateral helicity (high in Russia-Ukraine, undefined for Israel-Iran due to near-zero economic ties)

Network: System-level spectral properties that constrain or enable action regardless of bilateral configuration

For Israel-Iran, the dyadic helicity is mathematically undefined—division by near-zero curvature produces no meaningful signal. But the *network* metrics remain highly informative.

Israel-Iran conflict potential is mediated through the Middle East regional network's spectral gap. When the regional system becomes glassy (low λ_1), actors lose the ability to manage tensions through economic adjustment, diplomatic mediation, or third-party pressure. Proxy conflicts escalate toward direct confrontation when the network cannot absorb shocks—when there is no "give" in the system.

The proxy network itself served as a buffer. Iran maintained deterrence through Hezbollah's 150,000+ missiles, Hamas's presence on Israel's southern border, and the land bridge through Syria connecting Tehran



FIGURE 13.1: The escalation sequence through June 2025. The nuclear-threshold entry poses an open question, distinct from the preceding reported actions. The mechanism at issue is the changing set of available responses as proxy capacity declines.

to Beirut. Israel faced credible retaliation threats that constrained its operations. The proxies absorbed violence that might otherwise have required direct state-to-state action.

Losses to armed partners, leadership killings, and the fall of the Assad government changed Iran's regional position. The effect on deterrence cannot be read from a casualty percentage alone. It depends on which capabilities, routes, and commitments survived, and how the opposing side interpreted the change.

An apparent weakening of an adversary's network may be read as an opportunity or as a source of unpredictable escalation. Both interpretations need to be tested against subsequent deployments, statements, and actions. The graph does not decide between them on its own.

A prolonged confrontation can be managed through several channels at once: indirect exchanges, negotiation, external pressure, and restraint motivated by expected costs. If one channel is lost, the remaining ones matter more. The question is whether they can still contain the next episode.

Periods of limited confrontation can end abruptly, but calm does not prove a stable equilibrium and escalation does not prove that rupture was inevitable. The investigation needs the mechanism connecting the observed change to the decision to use force.

Paradoxically, proxy degradation made direct war *more* likely by eliminating the intermediary structures that had channeled conflict.

The Regional Spectral Gap

The Middle East regional trade network has undergone dramatic transformation:

Year	$\bar{\lambda}_1$	**Context**
2010	0.134	Pre-Arab Spring
2015	0.091	Post-Arab Spring fragmentation
2020	0.067	Abraham Accords begin
2023	0.078	Partial recovery
2024	0.052	Post-October 7 deterioration
2025	< 0.03	Post-proxy collapse, active war

The Abraham Accords briefly *increased* regional connectivity, pushing the spectral gap upward. The October 7, 2023 attacks and subsequent Gaza war have reversed this trend.



Gaza, 2023–2025. The scale of destruction visible in a single frame represents what the Davis Framework's spectral gap measures abstractly: the disconnection of a community from the structures that sustain it.

Photo: Jaber Jehad Badwan. CC BY-SA 4.0.



A scene from Gaza during the 2023–2025 war. Every data point in a helicity calculation—every trade flow, every diplomatic signal—ultimately connects to moments like this.

Photo: Jaber Jehad Badwan. CC BY-SA 4.0.



Surviving in Gaza during the 2023–2025 war. The framework's purpose is not merely academic— it exists to provide the warning time that might prevent such suffering.

Photo: Rawanmurad2025. CC BY-SA 4.0.

The April 2024 Exchange

In April 2024, Israel and Iran engaged in direct military strikes for the first time:

The direct exchange unfolded over nineteen days. On April 1, Israel struck the Iranian consulate in Damascus. On April 13, Iran launched more than 300 drones and missiles at Israel in Operation True Promise. On April 19, Israel struck Iranian air defense in Isfahan. The sequence made direct retaliation part of a relationship previously conducted largely through indirect channels.

The April 2024 attack from Iranian territory and the subsequent response made direct state-to-state action an explicit part of the relationship. Its scale, targets, and limits belong in the event definition used to evaluate the framework.

Framework reading at time of exchange:

$$\lambda_1^{\mathrm{ME}} = 0.054 \quad \text{Near-critical threshold} \tag{13.1}$$

At 0.054 the reported regional gap was 0.004 above the historical 0.05 cutoff. At that reading the regional indicator had yet to cross. To understand how the exchange was contained, we also need the sequence of messages, the targets each side chose, and the opportunities it had to stop. The gap locates the network relative to its threshold. Those decisions supply the mechanism.

Both Israel and Iran chose calibrated responses that avoided catastrophic escalation. The framework interprets this as the system finding a release valve—the spectral gap was low but not frozen, allowing adjustment without requiring full-scale war to break the glass.

Operation Rising Lion: The June 2025 War

On June 13, 2025, Israel launched Operation Rising Lion—a massive surprise attack that transformed four decades of shadow war into open interstate conflict. The Wall Street Journal reported that U.S. diplomatic talks with Iran served, whether intentionally or not, as cover for the Israeli strike. The operation's timing exploited a moment of perceived vulnerability: Iran's proxy network was degraded, its air defenses had been covertly sabotaged by Mossad, and nuclear negotiations with Washington had created a false sense of diplomatic momentum.

The opening hours combined kinetic destruction with targeted assassination on a scale unprecedented in modern Middle Eastern warfare:

Nuclear facilities: Israeli strikes targeted Fordow and Natanz enrichment plants, causing catastrophic damage to Iran's centrifuge infrastructure

Air defense suppression: Iran's integrated air defense network was systematically destroyed, leaving the country unable to defend against subsequent waves of strikes

Leadership decapitation: Targeted assassinations killed senior IRGC commanders, nuclear scientists, and political figures—including Ali Shamkhani, who had been overseeing nuclear negotiations with the United States

Covert sabotage: The Mossad activated exploding drones that had been covertly smuggled into Iran prior to the attack, destroying air defenses and long-range missile systems from within

Iran retaliated with ballistic missiles targeting Israeli cities and military installations. Impacts were reported in Bat Yam, Rehovot, and Ramat Gan. The exchange continued for multiple days, with both sides trading strikes as casualties mounted into the hundreds.

The framework had identified this exact transition. When the April 2024 exchange was contained, the spectral gap stood at 0.054—near-critical but above the threshold. By June 2025, the cumulative degradation of Iran's proxy network, the collapse of the Syrian land bridge, and the failure of nuclear diplomacy had pushed the regional spectral gap below 0.05. The system had frozen. There was no flexibility left to absorb the next shock.

On June 22, 2025, the United States crossed the threshold from defensive to offensive operations. Having spent the first nine days of the war defending Israel by shooting down Iranian missiles and drones, the U.S. Air Force and Navy struck three Iranian nuclear sites, including the deeply buried Fordow enrichment facility.

The New York Times' David Sanger captured the gravity of the moment: the United States had entered a war with Iran. The strikes represented the most significant American military action against Iran since Operation Praying Mantis in the Persian Gulf in 1988.

The most revealing development—from the framework's perspective—was the behavior of Iran's proxy network. The New York Times and France24 both documented how Iranian proxy militias "stayed largely silent" during the war, leaving Iran "isolated in war."

Only the Houthis in Yemen responded with force, firing ballistic missiles at Israel that triggered sirens in Jerusalem and the West Bank, and declaring the American strikes a "declaration of war." But Hezbollah, severely degraded after losing its leadership and much of its arsenal in the 2024 Lebanon campaign, did not open a second front. Iraqi Shia

militias made symbolic gestures but avoided significant escalation. The proxy architecture that Iran had spent forty years constructing failed at the moment of its greatest test.

The proxy network's failure to activate during Operation Rising Lion confirms a core framework prediction: when the regional spectral gap collapses below the critical threshold, the network can no longer transmit coordinated signals effectively. Iran's proxies were not merely unwilling to fight—the network through which coordination would have flowed had already fragmented. The proxy system died before the war; the war merely revealed the corpse.

The war ended with a ceasefire, but the strategic landscape that emerged bore no resemblance to the one that preceded it. President Masoud Pezeshkian was wounded. The IRGC's senior leadership was decimated.

The strikes also reached nuclear facilities at Fordow and Natanz. Assessing their effects requires evidence about the facilities and the subsequent program. The regional network score cannot supply a damage assessment or establish what deterrence would have achieved under a different policy.

A Frontline documentary titled "Remaking the Middle East: Israel vs. Iran," which aired on PBS in July 2025, captured the scale of the transformation. The forty-year shadow war was over. What followed would be something qualitatively different.

The post-war period revealed the depth of Iran's strategic crisis—and its refusal to accept the new reality. Despite catastrophic military losses, the IRGC continued efforts to reconstitute its proxy network. According to U.S. Central Command, Iran supplied weapons—including drones and missiles—to Hezbollah remnants in Lebanon, the Houthis in Yemen, and Shia militias in Iraq.

In Yemen, American forces intercepted a vessel carrying 750 tons of Iranian weapons, including drone engines and radar systems, destined for the Houthis. In Iraq, the Kurdish Regional Government accused suspected Iranian-backed militias of conducting drone attacks on oil infrastructure, including facilities operated by American companies. In Syria, the Interior Ministry reported intercepting multiple weapons shipments allegedly bound for Hezbollah, including anti-tank missiles hidden in commercial trucks.

Nuclear negotiations between Iran and the United States remained stalled, removing what had been the last diplomatic channel for managing the dyad.

By February 2026, Israel's Shin Bet and National Cyber Directorate reported foiling hundreds of cyberattack attempts attributed to Iranian

intelligence over the prior year. These targeted senior government officials, defense figures, academics, and journalists, with activity increasing sharply since the June war. The conflict had migrated partially to the cyber domain—a predictable consequence when conventional military options are degraded but the underlying hostility remains unresolved.

The February 2026 Strikes

On February 28, 2026, Israel and the United States launched a second surprise attack targeting key Iranian military personnel. Iran retaliated by striking numerous military bases across the region, including targeting multiple Arab states hosting U.S. assets. Ali Khamenei, who had led the Islamic Republic since 1989, was killed in the February 28 strikes, according to the subsequent Reuters report (Hafezi 2026). The assessment here is anchored to that escalation and its immediate aftermath.

The renewed fighting is consistent with the proposed danger of losing intermediary channels while hostility persists. It does not by itself establish that every provocation must produce a state-level response. The test is whether the network measurements add explanatory or forecasting power beyond the observed military losses, stated objectives, and failures of negotiation.

The Israel-Iran dyad has undergone a phase transition:

Parameter	**Value**	**Interpretation**
Dyadic Helicity	Undefined	No bilateral interdependence; metric inapplicable
Regional λ_1	< 0.03	Well below critical threshold; network frozen
Network Γ	< 0.1	Deep into critical regime
Proxy Network	Collapsed	Primary buffer structure destroyed
Alert Level	**L4**	SINGULARITY—active interstate war

The chapter's reconstruction moves from L3 in late 2024 to L4 in June 2025. The interval depends on the date assigned to the initial reading. Comparing its speed with another case requires the same alert definition, observation cadence, and onset rule in both.

Three scenarios organize the interpretation of that sequence:

Scenario A: Controlled tension. The April 2024 exchange showed that a direct exchange could end without continuing immediately into a larger war. The later escalation did not preserve that restraint. The comparison

asks what changed between the contained episode and the renewed fighting.

Scenario B: Network collapse. Hezbollah's losses, the destruction of Hamas leadership, and Assad's fall changed Iran's regional options. The model interprets those changes as a loss of intermediate routes for influence and confrontation. Establishing how much that loss contributed to direct war requires the network record alongside the decisions taken by the states involved.

Scenario C: External shock. A sequence of shocks can change the network without a single decisive interruption. The Gaza campaign, the Lebanon war, Assad's fall, and unsuccessful nuclear diplomacy provide events against which to examine the timing of the reconstructed decline in the spectral gap.

The possibility of gradual degradation followed by abrupt violence is the central concern. It motivates monitoring the connections through which pressure is managed. To count the scenario as a forecast, its issue date, input record, and outcome rule must precede the events being scored. The historical sequence remains useful for testing the proposed mechanism even when it is reconstructed later.

The June 2025 and February 2026 strikes fundamentally altered the nuclear dimension:

The transition period the framework identified as maximally unstable—when one side approaches nuclear capability while the other seeks to prevent it—was resolved through kinetic action rather than diplomacy. Israeli and American strikes destroyed significant portions of Iran's enrichment infrastructure at Fordow and Natanz.

Damage to facilities does not settle the subsequent political and technical choices. Retained expertise, the condition of the remaining program, outside constraints, and decisions about reconstitution all affect what follows. A geopolitical score cannot resolve those questions on its own.

The nuclear dimension also exposes a limit of the analogy. Deterrence is not an eigenvalue that can be inserted into a trade graph without a definition. It concerns beliefs about capability, response, and unacceptable costs. The regional analysis must show how those beliefs enter its representation before treating them as a measured source of stability.

The monitoring framework must be updated to reflect the transformed strategic environment:

Stabilizing factors:

The stabilizing side of the assessment rests on the ceasefire after the 2025 war, however fragile, and on Gulf states' neutrality during the June fighting. Degradation of Iran's proxy network reduces the

available routes for a coordinated attack. International pressure for de-escalation provides a further restraint. These restraints operate at different levels, from immediate military capacity to the diplomatic conditions for another round.

Destabilizing factors:

The opposing pressures remain substantial. Iranian efforts to rebuild proxy capacity and intercepted weapons shipments point toward reconstitution; cyber operations against Israeli targets provide another continuing channel of conflict. Nuclear negotiations with the United States remain stalled. In the account developed here, Khamenei's death and the decapitation of IRGC leadership compound uncertainty in Tehran.

The February 2026 strikes show the escalation cycle continuing, while Houthi missiles from Yemen keep another route active. A regional spectral gap below 0.03 indicates little flexibility in the surrounding network. Arab states hosting U.S. bases have also become targets of Iranian retaliation, widening the set of relationships through which the crisis can spread.

Within the period examined here, the relationship moved from a prolonged shadow conflict to repeated direct confrontation. The case asks which channels for restraint were lost, which remained, and which might be restored. The regional network gives that investigation a structure. It does not make the next decision inevitable.

This case brings direct and indirect connections into the same field of view. India and Pakistan pose a related question through a different history: what keeps a deeply hostile relationship within limits when ordinary economic exchange offers little room for adjustment?

Chapter 14

India-Pakistan: The Frozen Conflict

> Peace is not absence of conflict, it is the ability to handle conflict by peaceful means.
>
> Ronald Reagan, 1984

At a heavily guarded border, the absence of ordinary traffic can look like the absence of a relationship. The soldiers facing one another know otherwise. A closed crossing changes commerce without erasing territorial claims, memory, or the capacity to inflict harm. Here the framework has to explain what holds a hostile relationship short of a larger war when trade provides little restraint.

The historical episodes examined here run from Partition through the 2019 crisis. They include the wars of 1947–1948, 1965, and 1971, the limited Kargil conflict of 1999, and later confrontations short of a wider war. The question is why some episodes widened while others stopped. The framework approaches it through the distinction between *frozen* and *glassy* relationships; nuclear posture and escalation choices require their own evidence (Narang 2014).

The India-Pakistan relationship defies easy categorization. These are not disconnected strangers but estranged family—sharing language, cuisine, cricket, and the trauma of Partition. The hostility is intimate, rooted in a separation that killed hundreds of thousands and displaced millions.

In August 1947, partition divided British India into India and Pakistan. The boundary-making was rapid and the human consequences catastrophic. Millions were displaced and hundreds of thousands killed, with estimates varying substantially. A new border cut across lives, infrastructure, and production systems that had not been designed to function separately.

"A moment comes, which comes but rarely in history," Jawaharlal Nehru declared at midnight on August 14, "when we step out from the old to the new, when an age ends, and when the soul of a nation, long suppressed, finds utterance." But for millions, that moment was not liberation but terror.

The immediate economic effects were devastating. Industries were severed from their inputs: jute mills in Calcutta separated from jute farms now in East Pakistan; cotton textiles in Bombay separated from cotton fields in Sindh. Railway networks were cut. The Indus Waters Treaty, signed on September 19, 1960, established a concrete allocation and dispute-handling arrangement for waters crossing the new boundary (Government of India and Government of Pakistan and International Bank for Reconstruction and Development 1960).

A mill could remain in one country while a source of raw material lay in the other. A river could cross the new boundary without changing its course. Political separation therefore created new dependencies to negotiate, rather than a clean economic division.

The rupture also entered family histories. Homes, relatives, and places of worship were left across a border. The violence shaped expectations of the other state long after the immediate movements had ended. Economic policy operated within that memory, rather than outside it.

Those memories can change the meaning of a present-day proposal. A route that looks mutually profitable to a planner may appear dangerous to someone whose family experienced an earlier rupture. The calculation needs to account for that constraint without assuming that every member of either society responds in the same way.

The World Bank's 2018 regional-trade assessment placed intraregional trade at about 5% of South Asia's total trade. Its summary put India–Pakistan trade at roughly $2 billion against a modeled potential of $37 billion without the assessed barriers. These are dated observations and a conditional model estimate, not current trade figures or a guaranteed gain (World Bank 2018). The gap gives policy, connectivity, and trust a concrete place in the explanation.

The formal bilateral trade figure—around $2 billion annually—dramatically understates actual economic flows. Estimates of informal trade (goods routed through Dubai, Singapore, and other third countries to evade restrictions) range from $4–10 billion. Add remittances, smuggling, and services, and the true figure may be several times the official number. Yet even this inflated estimate represents a tiny fraction of what geography would predict for two neighboring economies of this size.

The Turkmenistan-Afghanistan-Pakistan-India (TAPI) pipeline illustrates the cost of frozen relations. First proposed in 1995, the pipeline



The India-Pakistan border at Attari, Punjab. The Wagah-Attari crossing is one of the few points where the two nations physically meet— a gap in the manifold where helicity, however diminished, still flows. The daily flag-lowering ceremony here is simultaneously an act of rivalry and connection.

Photo: Nalbarian. CC BY-SA 4.0.

Dimension	Status
Trade	Minimal: $2–3B formally, estimated $4–10B including informal channels (via Dubai, Singapore)
Energy	No pipeline despite geographic logic and shared need
Diplomatic	Relations frozen since 2019; embassies at reduced capacity
Military	Regular Line of Control skirmishes; nuclear capability both sides
Population	200M+ Muslims in India, Hindu-Muslim ties across border, massive diaspora

Table 14.1: India-Pakistan dyad profile.

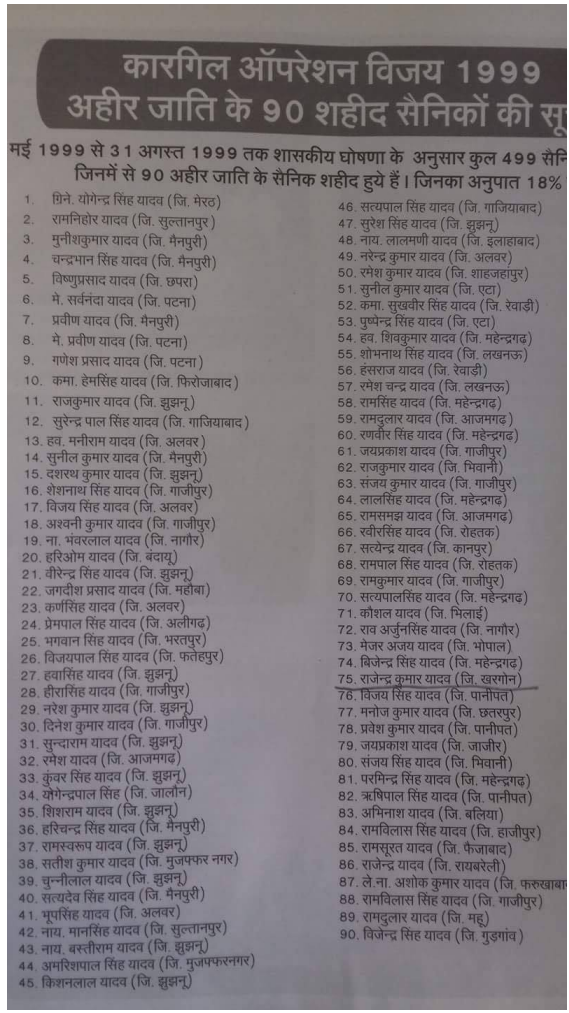
would transport Central Asian natural gas to energy-hungry South Asia. The economics are compelling: Pakistan and India face chronic gas shortages; Turkmenistan has abundant reserves; the route is relatively direct.

Yet three decades later, the pipeline remains unbuilt. Pakistan began construction on its section in 2018; the Afghanistan and Turkmenistan sections are nominally underway. But the India section has never progressed beyond preliminary discussions. The reason: neither country is willing to depend on the other for energy transit.

This represents a peculiar form of rationality. Both countries would benefit from the pipeline. But neither trusts the other not to weaponize the dependency. The gas that would flow through TAPI is precisely the kind of "flow" that our framework measures. Its absence is a deliberate choice to avoid creating the conditions for a trap.

A cross-border pipeline requires confidence about future access. The receiving country must consider whether a transit state could interrupt supply, while the transit state must consider its own exposure to retaliation and lost revenue. Those questions can prevent a commercially attractive project from becoming a usable route.

The Pulwama-Balakot crisis of February 2019 illustrated the dynamics. On February 14, a Jaish-e-Mohammed suicide bomber killed 40 Indian paramilitary personnel—the deadliest Kashmir attack in three decades. India responded with airstrikes at Balakot, inside Pakistani territory, the first such action since 1971. Pakistan retaliated, downing an Indian MiG-21 and capturing the pilot.



A memorial listing the names of soldiers who fell in the 1999 Kargil War—the last major armed conflict between India and Pakistan. Each name represents a point where the frozen helicity of the India-Pakistan dyad briefly thawed into violence before refreezing into the nuclear stalemate that persists today.

Photo: Kamalydv. CC BY-SA 4.0.

The 2019 crisis brought direct military action between nuclear-armed states. Its assessment requires separating observed deployments and strikes from reported threats and assumptions about nuclear readiness. Escalation risk was a reason for urgent attention, even though the subsequent outcome was contained.

Pakistan's return of the captured Indian pilot supplied a visible act of restraint during the 2019 crisis. Both sides stepped back from a wider confrontation. The question is what allowed that stopping point to hold.

Explaining the exchange of force is only half the task. The other half is explaining the stopping point. Which objective had been met, which risks were becoming unacceptable, and which communications or outside pressures helped contain the episode?

Repeated crises can produce expectations about how the other side signals limits. A restricted target, a public statement, or a return of personnel may be interpreted as a stopping point. Such expectations can help contain an episode, but they remain vulnerable to misreading and to actions that exceed what the other side anticipates.

A government can seek a limited military result while also wanting to prevent a wider war. That creates a bargaining problem about how to signal both resolve and a stopping point. The opponent may not interpret the signals as intended.

The India-Pakistan case illuminates a crucial distinction within the framework:

Frozen vs. Glassy Systems

A *frozen* system has low interdependence and low connectivity. It cannot generate trap dynamics because there are no flows to twist. Relationships are cold but stable—like two blocks of ice that touch at edges but share no water flow.

A *glassy* system has high interdependence that has become *stuck*. Flows continue but cannot adjust; relationships are trapped in suboptimal configurations. Like supercooled liquid that appears solid but is far from equilibrium.

India-Pakistan is *frozen*. Pre-2022 Russia-Ukraine was *glassy*. The difference is fundamental.

The framework computes:

$$H_{\text{IN-PK}} \approx -0.3 \quad K \approx 0.08 \quad \lambda_1 = 0.024 \tag{14.1}$$

The goods that still crossed the border gave the parties relatively little economic leverage. Here $K \simeq 0.08$ is low on the historical scale and the network gap of 0.024 is below 0.05. Those are two different measurements. The current engine switches to its special security branch only for $K < 0.05$. If 0.08 were its integration input, the ordinary branch

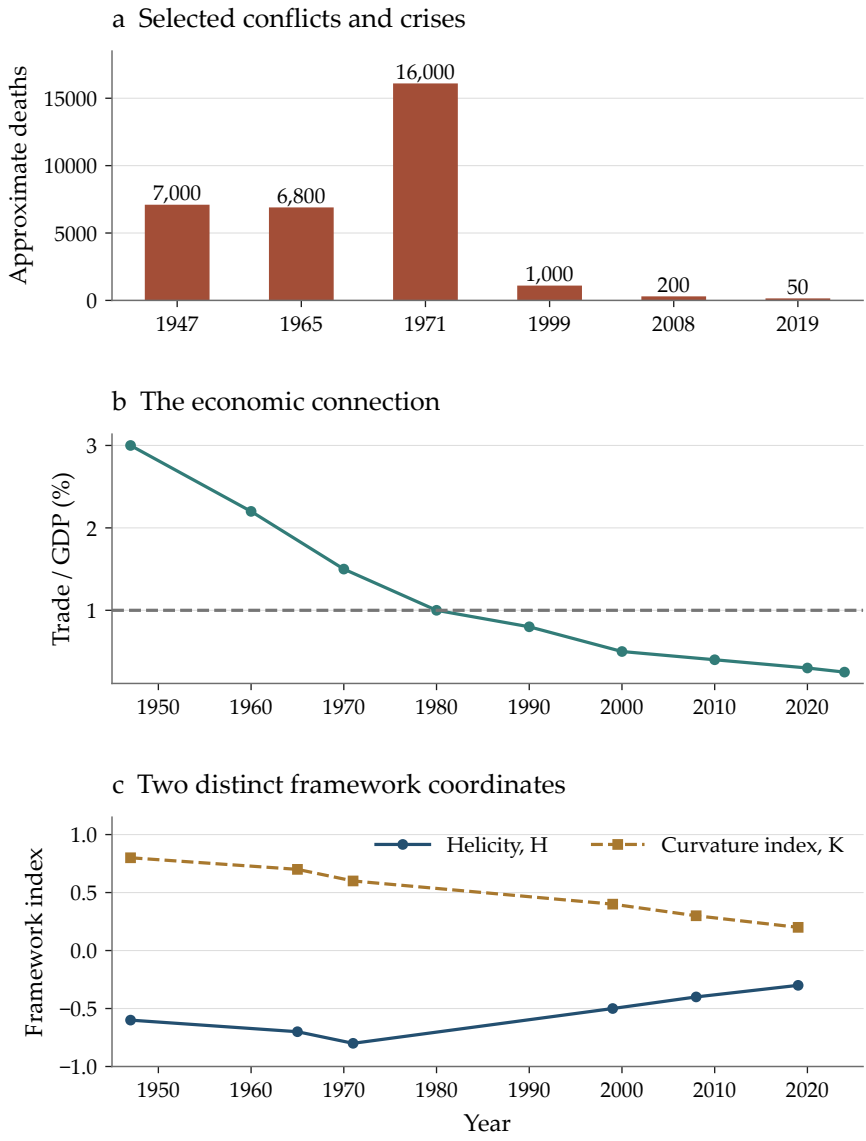


Figure 14.1: Three views of the India–Pakistan comparison. (a) Approximate counts for selected events. (b) The trade-to-GDP sequence. (c) The two framework-index sequences. Each panel retains its own units and definition; the juxtaposition does not by itself establish a causal effect.

would run. A closed trading route and the closing of a military hotline can both make this relationship harder to manage, through different channels.

This distinction resolves an apparent puzzle: why do India and Pakistan experience regular crises without full-scale war, while Russia-Ukraine went from integration to invasion in eight years? The answer: Russia-Ukraine had something to lose. The trap dynamics that made war likely arose from interdependence—the gas pipelines, the trade relationships, the supply chain integration. India-Pakistan has systematically eliminated these ties, producing a colder but more stable configuration.

Limited action needs its own analysis. A narrow target, a public claim that retaliation is complete, and a decision to return a captured pilot can all enter an account of containment. Low trade does not guarantee that violence will remain limited. The political objectives and escalation constraints have to be examined directly.

Historical Framework Readings

Event	Year	H	K	Outcome
Partition	1947	-0.9	0.4	War
Kashmir War	1965	-0.7	0.25	Limited war
Bangladesh War	1971	-0.8	0.3	War, partition
Kargil Conflict	1999	-0.5	0.12	Limited conflict
Mumbai Attacks	2008	-0.4	0.09	Crisis, no war
Pulwama/Balakot	2019	-0.35	0.07	Strikes, de-escalation

Table 14.2: India-Pakistan framework readings at key historical moments.

The table places declining economic integration alongside later episodes of limited conflict. In the framework's interpretation, fewer trapped commercial flows remove one source of pressure for rupture. The comparison does not by itself identify the cause of containment in 2019. Deterrence, political objectives, communication, and the decisions to stop remain part of that explanation.

The Nuclear Overhang

Both India and Pakistan possess nuclear weapons. This adds a constraint not present in other dyads:

Nuclear weapons create a floor beneath which escalation cannot descend without catastrophic consequences. This floor is not part of the helicity framework—it operates at a different level of analysis.

But the floor *interacts* with helicity dynamics. In a high-interdependence, negative-helicity dyad with nuclear weapons (hypothetical), the trap would be especially dangerous: pressure to break the glass through conventional war, but constraints on how far conventional war can escalate.

In a low-interdependence dyad like India-Pakistan, nuclear weapons primarily serve to *cap* escalation. There is limited pressure to escalate because there is no trap to escape.

The stability–instability paradox describes a proposed tension. Fear of a devastating nuclear exchange may restrain full-scale war while encouraging some actors to believe that limited operations can be kept below that threshold. That belief can be mistaken. Deterrence does not eliminate escalation risk.

A limit that both governments understand can still be crossed through misperception. Each side may believe it is taking a contained step while the other reads the same action as preparation for something larger. Communication matters because the boundary is interpreted, not physically marked.

The 1999 Kargil crisis gives the question a concrete setting. The investigation has to follow military positions, political demands, outside communication, and the decisions that ended the confrontation. Nuclear capability is part of the constraint, but it cannot stand in for the sequence of those decisions.

As noted in Chapter 9, India-Pakistan generated a false positive L4 alert in early 2019.

What happened:

The February 2019 crisis began with the Pulwama suicide bombing on the 14th, which killed 40 Indian paramilitaries. The reconstruction registers a helicity spike and a collapse of the regional spectral gap. Indian airstrikes on Balakot followed on February 26. Pakistan responded the next day, an Indian aircraft was downed, and the alert reached L4. On March 1, Pakistan released the captured Indian pilot. The sequence then turned toward rapid de-escalation.

Why the framework was wrong:

The framework correctly detected that conditions *could* support escalation. It incorrectly predicted that escalation would continue.

The error stemmed from underweighting two factors:

Nuclear deterrence effects

The low-curvature environment (neither side had enough economic stake to *need* war)

This case led to refinement of the framework's treatment of low-curvature dyads. When $K < 0.1$, the framework now applies a discount factor to L4 probability estimates.

The frozen relationship lacks much of the pressurized interdependence central to the trap mechanism. That narrows the mechanism's role; it does not make restraint automatic or war necessary in its presence.

Parameter	Value	Interpretation
Helicity	-0.32	Moderately negative
Curvature	0.07	Very low interdependence
Spectral Gap	0.024	Below critical, but reflects disconnection
Alert Level	**L2**	DECOUPLING (stable)

The Paradox of Peaceful Hostility

India and Pakistan maintain hostile rhetoric, periodic border clashes, and mutual nuclear targeting. Yet full-scale war remains unlikely precisely *because* they have so little economic stake in each other.

The absence of extensive direct trade removes one possible source of coercive leverage. It leaves territorial disputes, security commitments, and nuclear risk in place. The relationship is therefore a demanding test of what the disconnection branch explains beyond the trap mechanism.

This is not a success story—it represents a failure to achieve normal relations. But it is a stable failure. The frozen state is not under pressure to change because there are no trapped flows seeking release.

Restraint Is an Outcome to Explain

India–Pakistan is more than an exception at the edge of a theory about economic traps. It asks a question that every conflict framework must confront: how can a relationship sustain intense hostility, recurrent crises, and limited violence without every episode becoming a larger war?

The answer cannot be supplied by calling the relationship stable and then treating each clash as proof of instability. Define the level of violence being examined. The absence of full-scale war can coexist with grave insecurity, recurring casualties, and severe costs to ordinary life.

For comparison, distinguish an enduring dispute, a crisis episode, a limited use of force, and a sustained interstate war. Record the onset and termination rules for each. The resulting history may contain multiple

episodes within the same hostile relationship, and a model can perform differently on their different targets.

Deterrence concerns expectations about the consequences of action. It can make an escalation path unattractive without creating trust or resolving the underlying dispute. The geometry of the relationship can therefore remain constrained while the cost of crossing a particular boundary becomes very high.

That is different from an economic connection that encourages cooperation through shared gains. It is also different from a communication channel that helps actors interpret an incident. All three can affect restraint, but through different mechanisms.

A useful assessment asks whether the relevant deterrent expectation is credible, understood, and applicable to the action being considered. It also asks whether actors believe limited action can remain below the boundary. A strong constraint on one type of escalation can coexist with incentives to pursue another.

The distinction helps explain why a broad label such as "nuclear stability" is insufficient. Stability must be attached to the event whose probability is constrained. The same relationship can be stable with respect to one outcome and persistently dangerous with respect to another.

During a crisis, a channel can carry a demand, an explanation, a warning, or an assurance. Those messages have different functions. Some clarify the intended limit of an action; others create a route for ending an exchange without requiring agreement on the underlying dispute.

The value of a channel is therefore not exhausted by counting whether it exists. Examine whether it was used, who could speak through it, whether the recipient regarded the message as authoritative, and whether conduct followed the message.

Third parties can provide additional routes. Their usefulness depends on access and credibility with the relevant actors, not merely on their willingness to offer mediation. Several offers from actors without influence may add little practical capacity.

This is a concrete way to study the framework's tolerance budget and available constraints. The historical record should show which mechanisms absorbed the crisis and which failed, rather than assigning the entire outcome to one scalar.

Choose crisis episodes with similar starting conditions but different escalation paths. Before examining their outcomes, define which observations would support deterrence, communication, domestic political

incentives, or the economic mechanism as the main explanation of restraint.

Then follow the sequence. Did a communication change expectations before de-escalation? Did an economic cost become salient? Did an external actor alter the choices? Did a government achieve a limited objective and stop? More than one mechanism can matter, but their timing helps distinguish them.

An unsuccessful warning should lead to a specific question. Was the target wrong, the threshold too broad, the input poorly measured, or a stabilizing mechanism omitted? A successful warning should face the same scrutiny. Did it anticipate the defined event within the stated horizon, or merely describe a dangerous relationship that was already widely recognized?

The case becomes more valuable when it can teach the framework something. Its persistence, repeated crises, and episodes of restraint offer a demanding setting in which to separate hostility from the particular conditions that produce escalation.

What Would Change This?

The framework identifies scenarios that could destabilize the frozen equilibrium:

Increasing curvature (dangerous):

A new connection could restore the trap if the political relationship remained unrepaired. A major infrastructure project such as TAPI would create a fresh dependence. So could a substantial opening of trade without reconciliation, or a Kashmir settlement that established economic ties without enough trust to manage them. The relevant distinction is between creating a channel and creating the conditions under which both parties can rely on it.

Resolving helicity (stabilizing):

A durable improvement would join new opportunities to a political settlement. That might take the form of a comprehensive agreement on Kashmir or a slower normalization in which trust grows with exchange. A regional framework, including a reactivated SAARC with third-party guarantees, could provide support that the bilateral relationship cannot yet supply for itself.

The counterintuitive implication: well-meaning efforts to increase Indo-Pakistani trade *without* first addressing political issues could be destabilizing. Creating interdependence in a negative-helicity environment builds a trap.

The distinction matters beyond this border. A high level of hostility, a weak trading connection, and a constrained set of choices can combine in different ways. The neighboring cases let us compare those combinations without treating every dangerous relationship as the same trap.

Chapter 15

Other Active Dyads

> The world is a book, and those who do not travel read only one page.
>
> Often attributed to Augustine of Hippo

Carry the instrument from one border to another and the same reading can confront a different problem. Two great powers may be trying to reduce a dependence. Two neighbors may have almost no ordinary exchange left. An alliance may contain a dispute it has never settled.

The following assessments place additional dyads under the same questions: what still moves between them, what has become difficult to change, and which arrangements keep hostility from becoming a larger conflict? The figures belong to the assessment periods discussed here. Their value is in the comparison they make possible.

US-China

Parameter	Value
Helicity	-0.31
Curvature	0.42 (high interdependence)
Spectral Gap	0.089
Alert Level	**L2** DECOUPLING

The US-China relationship presents a paradox that has preoccupied scholars for two decades: intense strategic rivalry coexisting with deep economic integration. Understanding this dyad is essential because its trajectory will shape the international system for the remainder of the century.

Combined, the two economies represent 43% of global GDP and 48% of global manufacturing. The bilateral economic relationship dwarfs any prior great power rivalry:

Dimension	2024 Status
Bilateral Trade	$690 billion annually
Trade Deficit	$295 billion (U.S. deficit with China)
Treasury Holdings	$902 billion (Chinese holdings of U.S. debt)
FDI Stock	$145 billion (U.S. in China); $62 billion (China in U.S.)
Supply Chain	Apple depends on China for 95% of production
Student Exchange	290,000 Chinese students in U.S. (2023)

Table 15.1: US-China economic interdependence, 2024.

This integration creates what Farrell and Newman called the "weaponized interdependence" dilemma in its purest form. Each economy can inflict severe damage on the other through disruption of supply chains, financial restrictions, or technology denial. But each is also vulnerable to the same weapons turned against itself.

The United States and China can impose costs on each other through commercial connections, but disruption can also return to the actor imposing it. A useful account follows the specific channel, the available substitutes, and the delay before the costs become binding. The total trade figure does not contain all three.

The debate over how to characterize this relationship has become sharp, with significant implications for policy.

The case against the analogy. US–China economic integration differs substantially from the US–Soviet relationship. Production, investment, and technology networks connect firms across a political rivalry. An account based on two separable blocs can therefore miss the channels through which competition operates.

Other states also make choices within this network. A country can cooperate with China in production while relying on the United States for security, or seek access to both markets. Such overlapping commitments complicate a picture of a world divided neatly in two.

The case for the analogy. Strategic rivalry spans military planning, technology, diplomacy, and competing political ambitions. The comparison draws attention to the possibility that apparently separate disputes reinforce one another. Its usefulness depends on identifying which mechanisms resemble the earlier Cold War and which do not.

Competition and commerce can continue together. A manufacturer may need a component from a strategic rival while its government

restricts another technology. A third country may welcome redirected investment while worrying about being forced to choose a side. Those different exposures explain why the same bilateral confrontation can produce different responses around the wider network.

The framework provides a different lens for this debate:

US-China represents a high-curvature, moderately negative helicity dyad undergoing deliberate decoupling. The key question is whether decoupling is proceeding at the right pace—fast enough to avoid trap formation, slow enough to avoid economic rupture.

Targeted restrictions. The Biden administration pursued restrictions in strategically important technologies while maintaining broader commerce. The October 2022 semiconductor export controls illustrate this distinction. Restricting one part of a relationship can alter the wider network through substitution, investment, and expectations of future access.

The distinction between a limited restriction and broad economic separation matters to firms trying to plan. Which product is covered? Which license remains available? How long will adjustment take? Those details determine how a strategic announcement changes actual production and exchange.

Interpretation by the other side. A policy described as reducing exposure may be read as an effort to constrain development. The distinction between a targeted restriction and a comprehensive separation matters operationally, but it may not settle the political disagreement about purpose.

The Trap Analysis: The framework shows that US-China decoupling is proceeding at a pace that roughly matches trust erosion. This is the crucial difference from Russia-Ukraine, where trust collapsed rapidly while economic ties persisted. The US-China relationship exhibits *coordinated* deterioration: economic ties weakening alongside diplomatic ties, preventing the formation of a trap.

The key risk: if political tensions escalate faster than economic decoupling proceeds, the dyad could enter trap territory. Current trajectory suggests this is *not* occurring—decoupling is keeping pace with diplomatic deterioration.

Scenarios that could accelerate trap formation:

The most abrupt route to rupture would be a Taiwan crisis that forced economic separation before either side had prepared for it. Financial sanctions on Chinese banks could trigger immediate decoupling through a different channel. Chinese military action in the South China Sea or a cyberattack attributed to state actors could likewise move the dispute from managed competition to confrontation. These scenarios differ in



USS Theodore Roosevelt (CVN-71) joins a multinational battle group formation, 2009. The carrier strike group embodies the paradox of the US-China dyad: military positioning that simultaneously deters conflict and generates the friction that erodes alignment. Level 2 DECOUPLING in geometric terms.

Photo: U.S. Navy. Public Domain.

Dyad	H	Gap	Γ	Alert
China–Taiwan	−0.68	0.041	0.96	L4
India–Pakistan	−0.42	0.035	1.05	L3
Israel–Iran	−0.78	0.055	0.82	L4
US–China	−0.28	0.092	1.25	L3
US–Russia	−0.55	0.058	1.18	L3+
NATO–Russia	−0.72	0.028	0.95	L4
US–Canada	+0.82	0.185	1.45	L1
France–Germany	+0.88	0.210	1.52	L1

Assessment horizon: 2025. Definitions appear in the text.

Figure 15.1: The 2025 cross-dyad assessment. Structural readings appear beside the assigned alerts. Two stable relationships provide comparison points for the six stressed dyads.

their opening event, but each can close options faster than the commercial network can adjust.

Scenarios that maintain the current trajectory:

The moderating conditions are more incremental. A continued "small yard, high fence" policy limits the scope of economic separation. Chinese prioritization of economic stability can constrain escalation, while functioning diplomatic channels preserve a means to negotiate during a dispute. Singapore or the EU could assist through mediation. These conditions matter because they preserve alternatives at the moment a crisis would otherwise narrow them.

The pace of adjustment matters. Firms can sometimes accommodate announced restrictions by finding substitutes, changing designs, or building inventory. An abrupt interruption leaves less room for those responses and can make the same ultimate reduction in exchange much more disruptive.

Channels of communication are part of that adjustment capacity. They can clarify whether a restriction is limited, identify an exemption, or prevent a technical dispute from being read as a wider threat. Their value has to be followed through what the parties actually agree and implement.

The relationship must therefore be examined across several activities at once. Governments can restrict a technology, negotiate over another issue, and continue importing ordinary goods. A single label for the bilateral relationship can conceal how these channels reinforce or obstruct each other.

Watch indicators: Chip export control expansions, Taiwan policy statements, bilateral investment screening (CFIUS), rare earth restrictions, student visa policies, Treasury holdings, military-to-military communication channels.

North Korea-South Korea

Parameter	**Value**
Helicity	-0.71
Curvature	0.03 (very low)
Spectral Gap	0.018
Alert Level	**L2** DECOUPLING (frozen)

Assessment: Similar to India-Pakistan, the Korea dyad is *frozen* rather than *glassy*. Helicity is deeply negative but curvature is negligible. There is no trap because there are no flows.



The Korean Demilitarized Zone—250 kilometers of enforced emptiness dividing the peninsula. In the Davis Framework, the DMZ represents a manifold boundary where helicity has been zero for seven decades: no trade, no movement, no alignment. A topological wall sustained by the complete absence of flow.

Photo: Nicor. CC BY-SA 3.0.

The 2016 closure of the Kaesong Industrial Complex removed an important economic connection between the two Koreas. It also ended the work and commercial arrangements that depended on that connection. Removing a channel changes both the leverage it provided and the opportunities for cooperation it supported.

That tradeoff prevents a simple inference that closure made the relationship safer. The nuclear and territorial confrontation remained. A structural assessment has to ask which constraints changed, for whom, and on what timescale.

Risk scenario. New economic ties can create exposure before political commitments become reliable. The test is whether the engagement also provides workable protections and alternatives. A project that cannot be interrupted without crisis requires more than an expectation of goodwill.

Other fault lines

Parameter	**Value**
Helicity	-0.44
Curvature	0.11
Spectral Gap	0.041
Alert Level	**L3** DIVERGENCE

Assessment. The mass-gap paper records an August 15, 2020 L4 alert before the September 27 onset, a 43-day lead (Davis 2025). The historical first-warning roster has a separate 167-day entry, discussed in Chapter 9. The September 2023 Azerbaijani operation belongs to a subsequent episode, not an extension of the 2020 forecast horizon.

The Armenian population's displacement from Nagorno–Karabakh and the change in territorial control belong to the subsequent history of the conflict. A military outcome does not settle the security, identity, and political questions that remain for the people affected.

For Armenia, the outcome also raises questions about the reliability of external security commitments. An alliance that looks available in a formal record may not deliver the protection expected in a particular crisis. That gap between a commitment and its performance belongs in the geometric account.

For Azerbaijan, territorial recovery and the memory of earlier displacement have different political meanings. An explanation has to distinguish those meanings without treating the suffering of either population as an interchangeable score in a bilateral ledger.

A durable settlement requires more than ending a military contest. Access, security, displacement, and expectations about future commitments can remain unresolved. The next assessment should track those conditions as a new episode, rather than extend the earlier war's forecast indefinitely.

The framework now shows post-conflict reconfiguration. Helicity remains negative; the spectral gap is near-critical. Risk of renewed conflict remains elevated.

Watch indicators: Armenia-Russia alliance status, Zangezur Corridor negotiations, Turkish-Armenian normalization.

The Ethiopia–Eritrea–Sudan complex adds the problem of overlapping relationships.

Parameter	**Value**
Regional λ_1	0.033
System Γ	0.78 (underdetermined)
Alert Level	**L3+** Multiple dyads in DIVERGENCE

Assessment: The Horn of Africa regional network has collapsed to a glassy state. The Ethiopian civil war, Sudan civil war, and Eritrean isolation have fragmented what was already a weakly connected system.

The framework cannot identify a single dyad at imminent risk because the entire regional structure is unstable. This is a multi-body problem rather than a bilateral trap.

Watch indicators: Nile dam negotiations (GERD), Sudanese conflict evolution, Eritrean foreign policy shifts.

In the Venezuela–Guyana case, territorial claims and resources bring another configuration into view.

Parameter	**Value**
Helicity	-0.52
Curvature	0.06
Spectral Gap	0.067
Alert Level	**L3** DIVERGENCE

Assessment: The Essequibo dispute has elevated this dyad. Helicity is significantly negative due to the asymmetric oil resource dispute (Exxon discoveries in contested waters) against collapsed diplomatic trust.

However, curvature is low—there is minimal bilateral economic interdependence. The framework assesses this as similar to India-Pakistan: hostile rhetoric without trap conditions.

Risk factor: If Venezuelan economic collapse accelerates, the regime may seek external conflict for domestic legitimacy. This political factor is outside the framework's scope.

Greece and Turkey put alliance membership alongside persistent disagreement.

Parameter	Value
Helicity	-0.29
Curvature	0.18
Spectral Gap	0.112
Alert Level	**L2** DECOUPLING

Assessment: The 2020 Eastern Mediterranean standoff generated a false positive L4 alert. The dyad has since stabilized due to EU mediation and shared NATO membership constraints.

The framework currently shows L2 (DECOUPLING) with moderately negative helicity but healthy spectral gap. The system retains flexibility.

Structural constraint: Both parties are NATO members. Greece is an EU member, while Türkiye is an EU candidate country. These overlapping but unequal institutional positions provide channels whose practical ability to manage disputes must be assessed separately (European Commission 2026).

Several active or potential conflict zones are *not well-suited* to framework analysis:

Case	Why Framework Is Limited
Myanmar civil war	Primarily internal; no bilateral trap dynamics
Syria complex	Multi-actor system; bilateral analysis insufficient
Sahel region	State weakness prevents meaningful helicity computation
Mexico cartels	Non-state actors; different analytical framework needed

Table 15.2: Conflicts outside the framework's primary scope.

The Davis Framework is designed for *interstate trap dynamics*—bilateral relationships with significant economic interdependence where trust has collapsed. It is not a universal conflict prediction system.

Internal conflicts, multi-party civil wars, and non-state actor violence require different analytical approaches. Acknowledging these limitations is essential for appropriate application.

A book concerned with rupture also has to follow relationships that recover room to move. The next case keeps the costs in view while asking how alternatives and procedures can reopen a connection.

Chapter 16

China and Australia Find Room to Move

A boat returns to a buyer. A shipment clears a border. An industry begins to recover access to a market it once expected to keep. These are small events beside the spectacle of war, but they belong in a book about preventing it.

The China–Australia dispute of 2020–2024 offers a different trajectory from the catastrophe at the center of this book. Economic pressure was real, relations deteriorated, and some producers faced severe disruption. Yet the relationship also contained routes for adaptation and negotiated relief. Those routes deserve the same attention as the connections that created exposure.

An aggregate trade figure can tell two contradictory stories at once. The total may remain large while a particular industry loses its principal market. One export category can offset another in national accounts without employing the same workers, servicing the same debts, or using the same assets.

Australia's parliamentary analysis of the dispute describes this distinction: iron ore exports helped offset declines in targeted categories, while exports to other markets also absorbed part of the disruption. The aggregate effect and the damage to affected products were therefore different quantities (Parliament of Australia 2026).

The framework should retain both. A national vulnerability measure needs sectoral detail; a sectoral account needs the national resources that may support adjustment. Neither a story of complete helplessness nor a story of costless resilience follows from the total alone.

Wine makes the point particularly clearly. The Australian foreign affairs department reports that bottled-wine exports to China fell from about A$1.12 billion in 2019 to A$16 million in 2022. It also records renewed exports after the duties were lifted in March 2024. Those observations establish a severe market interruption and subsequent reopening. They do not show that every producer recovered its losses

or that reopening restored the earlier business structure (Australian Department of Foreign Affairs and Trade 2024).

Alternatives are unevenly distributed

A commodity that can be redirected through existing bulk-trading channels faces a different adjustment problem from a branded product built around particular customers and distribution networks. A perishable product has a different clock again. Inventory, transport, consumer preferences, and certification determine how quickly a nominal alternative becomes a paying buyer.

This is why substitution should be measured as a function of time. Let $r_s(T)$ denote the fraction of a disrupted sector's sales that can be replaced over horizon T. A high value after three years does not imply a high value after three months. Businesses can fail during the interval even when aggregate trade eventually adapts.

The costs also matter. Selling the same quantity elsewhere at a lower margin is adaptation, but it is not the absence of injury. A useful recovery account follows quantities, prices, margins, and survival where those data are available. National resilience should not be inferred from one favorable total.

For the geometric interpretation, substitute markets create paths. Their capacity and cost determine whether those paths can carry the adjustment. Diversification can reduce vulnerability without requiring the original relationship to disappear.

In April 2023, Australia and China announced a process for reviewing the barley duties. Australia temporarily suspended its WTO dispute while China undertook an expedited review, with a specified period for that review. The announcement connected diplomatic engagement with an existing formal dispute procedure (Australian Minister for Foreign Affairs 2023).

The significance is the sequence. The parties did not have to resolve every strategic disagreement before acting on a particular trade problem. Nor did dialogue require the dispute to be abandoned without a process. A limited arrangement created a step that could be attempted and evaluated.

That is a concrete example of additional room for movement. In a static map, the countries remain two nodes with a damaged relationship. In a map of possible transitions, there is now a procedural route from restriction to review and potentially to relief.

The WTO record for the wine dispute similarly distinguishes the June 2021 request for consultations from the March 2024 notification of a

mutually agreed solution. These are stages in a documented process, not evidence that the wider relationship had become harmonious (World Trade Organization 2026).

Institutions matter here as procedures with dates and actions. Counting membership alone would miss what changed. The informative question is whether a mechanism can be used, whether its steps are credible, and whether implementation follows the announced decision.

Reopening is an observable event

On December 20, 2024, the Australian government announced that China had notified it of the full resumption of live rock lobster trade. It identified this as the resolution of the remaining trade impediments from the period under discussion. The statement supplies a clear endpoint for this bounded account of the dispute. It is also the government's account of its own policy success, so its explanation of causation should be distinguished from the access decision it reports (Australian Minister for Foreign Affairs 2024).

An open market does not necessarily restore the old dependence. Producers may retain customers found elsewhere. Governments may retain a more cautious view of exposure. Conversely, restored profitability can encourage renewed concentration. The post-dispute trajectory must be measured, not assumed.

This gives recovery two dimensions. The immediate dimension is restored access. The structural dimension is the capacity to withstand a future interruption. A relationship can improve in the first and remain vulnerable in the second.

The episode suggests three mechanisms to investigate together. First, adjustment outside the bilateral relationship can reduce the cost of waiting. Second, diplomatic engagement can make limited agreements easier to formulate and implement. Third, formal procedures can provide a sequence that does not require either side to settle every disagreement at once.

These mechanisms can reinforce one another. An actor with alternatives may be better able to negotiate patiently. A credible review process may reduce the urgency of expensive diversification. Renewed trade may reward restraint while creating new reasons to manage future disputes.

But their contributions cannot be read directly from the final outcome. The lifting of restrictions does not identify how much was caused by diplomacy, commercial incentives, changes in political leadership, legal proceedings, or the effectiveness of substitution. A serious explanation compares timing and evidence for each mechanism.

The Davis Framework provides a way to organize that comparison. Ask whether the burden of dependence changed, whether tolerance and communication improved, and whether feasible exits multiplied. Then inspect the observations separately before collapsing them into an overall diagnosis.

Why did this confrontation follow a different path from a relationship that ended in war? The answer cannot be "because its score was better" if the score has merely encoded the known outcome.

The comparison needs observations available during the dispute. Identify which alternatives existed, how quickly they could be used, and which commitments remained credible. Examine the strategic stakes and the available instruments of pressure. A trade dispute and a territorial confrontation are not interchangeable episodes simply because both contain hostile statements and large economic flows.

This is also why a peaceful outcome should not automatically be labeled a prevented war. The structural pressure may have been real while the specified war outcome was never likely. Distinguishing those propositions strengthens the framework by giving each outcome its own target and evidentiary burden.

The case is therefore valuable both as recovery and as a boundary test. It asks whether the framework can identify improvement before its culmination, while resisting the temptation to turn every easing of tension into confirmation of an earlier catastrophic forecast.

Diversification has costs even when it succeeds. Maintaining an alternative customer, supplier, route, or reserve can be less profitable than concentrating on the best immediate opportunity. A firm may prefer efficiency while a government values resilience. A government may endorse resilience while leaving individual producers to pay for it.

The policy question is how those costs are shared and how much independence is worth at the margin. Complete separation is rarely the only available arrangement. Contracts, inventories, financing, insurance, and procedural access can change the consequences of a disruption while trade continues.

This is a more useful prevention question than whether interdependence is simply good or bad. What arrangements let people retain the benefits of exchange without making every political dispute a threat to their economic survival?

Recovery can begin before trust is complete. A limited procedure, a credible substitute, or a reversible agreement can create room for the next move. The test is whether those moves are implemented and whether they enlarge the choices available during the next disagreement.

The lesson of this case is not that every trapped relationship can be repaired in the same way. It is that repair has a structure worth studying. A book that can describe only collapse has left half of the geometry unexplored.

Reopening a route is one kind of repair. Building institutions that keep future disputes from closing it is another. France and Germany allow us to follow that longer work, in which the procedures governing a connection become part of the connection itself.

Chapter 17

France and Germany Build a Different Relationship

Coal and steel can supply a battlefield. They can also become the subject of a common authority. The material does not decide which relationship will be built around it.

The reconstruction of relations between France and West Germany offers a powerful counterpart to the trap. It allows us to examine the deliberate construction of institutions that make adjustment possible. The question is not how two peoples suddenly ceased to disagree. It is how disagreement came to operate within a different structure.

On May 9, 1950, the Schuman Declaration proposed placing French and German coal and steel production under a common High Authority, open to other European states. The proposal linked a concrete economic arrangement to the prevention of another war. It also specified institutional features: binding decisions, independent members of the authority, and means of appeal (Schuman 1950).

The important move was to connect productive interdependence with a procedure for governing it. More commerce alone would not have created that authority. A declaration of friendship alone would not have pooled the relevant decisions. The proposal addressed both the material connection and the rules through which it would be managed.

This distinction is central to the Davis Framework. A dense economic relationship can create a burden when adjustment depends entirely on unilateral decisions. The same density can support cooperation when the parties have credible ways to make, contest, and implement decisions together.

From a proposal to an institution

The treaty establishing the European Coal and Steel Community was signed in Paris on April 18, 1951, by France, West Germany, Italy, Belgium, the Netherlands, and Luxembourg. It entered into force on July 23, 1952. The separation between signature and entry into force is a

useful reminder that institutional capacity has to be built through legal and political steps (United Nations Treaty Collection 1951).

A model should not assign the full effect of an institution to the date someone first proposes it. Nor should it assume that formal entry into force completes the process. Staffing, implementation, compliance, and repeated use determine whether a procedure becomes a practical option.

For historical analysis, distinguish announced intention, legal commitment, operational capacity, and demonstrated performance. Those stages can have different effects on expectations. A credible proposal may influence behavior before implementation; a nominal institution may disappoint after it begins operating.

The geometry of reconciliation therefore has a time dimension. An institution can create a path before the participants have learned to trust every person using it. Repeated implementation can then change their expectations about future disputes.

It is tempting to imagine trust as a substance that must be accumulated before any economic commitment is safe. The historical problem is more complicated. Some commitments can help create the conditions under which trust grows, while others increase exposure without providing any way to manage it.

The distinction concerns design. Does a new connection make conduct easier to observe? Does it distribute decision authority? Does it permit appeal? Does it give both sides a stake in procedures that continue to work during disagreement?

These are empirical questions about institutions, not reasons to assume that every integration project succeeds. A joint body can be ineffective or dominated by one party. A formal right can be difficult to exercise. Yet an institution that actually changes those conditions can alter the relationship even before affection or complete political alignment develops.

The corresponding mathematical question concerns feedback. Integration may increase the cost of adjustment while successful institutional interaction increases the capacity to manage it. The net trajectory depends on both rates. A snapshot of trade cannot reveal which process is moving faster.

The broader project of European integration encountered consequential failures. The European Parliament's historical record describes how the French Parliament's refusal to ratify the European Defence Community treaty in 1954 prevented the associated European Political Community project from advancing beyond planning (European Parliament Historical Archives 2026).

This matters because successful reconciliation can be made to look inevitable when told backward. A proposed route can fail while other forms of cooperation continue. The persistence of alternatives is part of resilience.

In network language, a relationship supported by several institutions and social connections need not lose all capacity when one channel is blocked. But the channels must have some independence. Five committees dependent on the same veto may not offer five routes around it.

The historical lesson is to examine the actual architecture of cooperation. Which functions survived a setback? Which actors continued to meet? Which agreements could proceed without the failed project? Recovery from institutional disappointment is evidence about the structure, not an inconvenient exception to a smooth story.

The Élysée Treaty of January 22, 1963, established regular bilateral consultation. The French foreign ministry describes scheduled meetings among leaders and foreign ministers, alongside cooperation concerning defense, education, and youth. Its purpose included making cooperation systematic beyond the immediate preferences of the two leaders (French Ministry for Europe and Foreign Affairs 2026).

Routine can sound less dramatic than reconciliation. In a crisis, it can be more useful. A standing obligation to consult reduces the need to invent a channel while events accelerate. Officials know whom to call, which issues can be discussed, and which procedures already exist.

The Franco-German Youth Office's historical account also emphasizes that the treaty rested on earlier work and that economic cooperation was not its principal subject. It records disagreements over the treaty's relationship to wider European and Atlantic commitments (Franco-German Youth Office 2026).

That combination is instructive: deepening cooperation did not require an identical view of every strategic issue. Institutionalized consultation gave disagreement a place to occur. The relationship became more capable of carrying differences without making each one a test of whether cooperation itself would survive.

Several layers can support the same peace

Economic rules, official consultations, local connections, and educational exchanges operate through different mechanisms. The first can manage material disputes. The second can reduce uncertainty and coordinate responses. The latter can change familiarity and the social constituencies that sustain the relationship.

A multilayer representation can retain those distinctions. Let L_e, L_d, and L_s denote operators for economic, diplomatic, and social connections. A weighted sum

$$L_{\text{combined}} = \alpha L_e + \beta L_d + \gamma L_s$$

acts on a common list of countries when the three layers have compatible undirected weights on agreed scales. With nonnegative coefficients the result is again a graph Laplacian. Think of two towns linked by a goods railway and a telephone line. Combining the weights adds two means of contact to the same edge. A question about a phone call reopening a closed railway needs the layers separately and a rule for how the first changes the second.

For many questions it is better to inspect the layers separately and examine how they interact. A loss of economic access may be manageable when political procedures remain open. A diplomatic dispute may be less disruptive when several institutions continue to perform their functions.

This approach also avoids treating density as an unconditional good. Additional connections can transmit shocks as well as distribute adjustment. The relevant question is what each connection carries and whether the institutions can manage that transmission.

Franco-German cooperation developed within a wider postwar order. Security arrangements, domestic political changes, economic reconstruction, and international institutions all belong in an account of that history. A bilateral success cannot identify the causal contribution of one measure by itself.

Nor should peace between these states be confused with an absence of violence elsewhere. An explanation centered on Europe must not turn other people's wars into the background of a universal story of peace.

France and Germany give us a case of deliberate institutional construction. Comparing it with trade ties that lack a standing dispute procedure can show what regular meetings and enforceable decisions contribute to cooperation.

Two more comparisons matter. Consider a similar institution with little power to implement decisions, or a friendship sustained by social ties with little trade. Each isolates a different route to reconciliation.

The arrangement cannot be copied as a formula. Countries differ in political systems, history, geography, and the distribution of power. A supranational authority acceptable in one setting may be unacceptable or unworkable in another.

What can travel are questions about function. How will commitments be monitored? Where can a dispute be heard? Which actions remain reversible while confidence develops? What happens if a leader changes? Who benefits from keeping the procedures alive?

A path to peace is a sequence of actions people can take under rules they understand, with resources they have and costs they can bear. The achievement was a relationship that could survive a disagreement. That is a geometry worth building. The African cases that follow make the question of implementation more demanding: who can enforce a commitment, which boundary an agreement changes, and what remains usable when an institution changes shape.

Chapter 18

Sudan: The State Is Not One Actor

A national account can record an export while concealing the struggle over who receives its proceeds. A map can give a territory one color while several armed commands contest the authority to govern it. A peace model can make the same mistake. It draws a circle around a country, writes the country's name inside it, and assumes that the circle encloses a decision-maker.

Sudan requires us to open that circle.

On April 15, 2023, fighting broke out between the Sudanese Armed Forces and the Rapid Support Forces. In its May 2 statement, the African Union described the confrontation as a threat to civilians and to the political transition. It situated the emergency in a sequence that included the 2019 transition arrangements, the 2020 Juba Peace Agreement, and the October 2021 coup. Its demands concerned the protection of people and infrastructure, humanitarian access, and the return to a political process (African Union 2023). This was a conflict over the organization of authority within a state, with consequences extending beyond its borders. A bilateral trade score alone could not describe its object.

The African cases in this part develop the framework's mechanisms through documented historical episodes. They do not add invented alert histories to the 47-conflict research record. Their contribution is a more demanding specification of what an actor, a connection, an exit, and a recoverable relationship must mean.

Who can make a promise hold?

Consider what a ceasefire requires. A representative signs. Instructions reach units. Those units have the means and incentives to obey. Violations are observed. A process exists for resolving them without converting every incident into the collapse of the whole agreement. These are different operations. A signature cannot substitute for all of them.

In the Davis language, a diplomatic tolerance budget must therefore be attached to the authority that can actually exercise it. The existence of a negotiating channel is one observation. The commander's ability to make a subordinate stop firing is another. Civilian acceptance of the proposed settlement is a third. Averaging them into one comfortable number can erase the condition that makes the agreement executable.

Suppose, as a constructed example, that two commands each control half the relevant armed units. The first can reliably transmit and enforce an instruction; the second cannot. A national average of one-half does not mean every unit is halfway governable. It describes a mixture of different situations. If a humanitarian passage crosses both areas, the weaker implementation condition can determine whether the passage works at all.

This distinction is familiar outside politics. An electrical circuit with one broken connection does not conduct half a current merely because half its switches are closed. The analogy is about a necessary sequence of operations, not about equating people with components. People can negotiate, refuse, learn, and create alternatives. The model has to preserve those possibilities while recognizing that several necessary conditions may have to hold together.

Write the operational requirement for a particular agreement as

$$F = F_{\text{authority}} \cap F_{\text{access}} \cap F_{\text{monitoring}} \cap F_{\text{civilian protection}}.$$

Each set contains arrangements satisfying a specified requirement. Their intersection contains arrangements satisfying all four. A large set of acceptable diplomatic language is of little use if none of its members can be implemented safely on the ground. An empty intersection identifies a design problem in the proposed agreement; it does not prove that every conceivable future settlement is impossible.

The distinction changes the research question. Instead of asking whether "Sudan" has enough trust in another state, ask which decision-making bodies must coordinate for a particular action, which commitments each can enforce, and which affected people have been excluded from the negotiation. A government, an armed organization, a local administration, a civilian coalition, and a neighboring state are not interchangeable kinds of node. Their relationships need names before they need weights.

An actor graph also needs dates. An organization can fragment. A chain of command can be reorganized. A representative can lose the ability to bind the people whose conduct matters. Treating the actor list

as permanent introduces a hidden assumption of continuity precisely where the conflict may be destroying continuity.

Jeddah: the distance between a signature and a delivery

The May 2023 Jeddah negotiations put the distinction into writing. On May 20, representatives of the SAF and RSF signed a seven-day ceasefire agreement, due to take effect after forty-eight hours. The accompanying announcement named practical obligations: humanitarian delivery, restoration of essential services, withdrawal from hospitals and public facilities, and safe movement of relief personnel and goods (Saudi Arabia and United States 2023). Those were tasks assigned to named parties, with a starting clock. They were not simply expressions of improved sentiment.

The agreement's full text provided a Monitoring and Coordination Committee, with representatives of the two parties and the Saudi and US facilitators. It specified communications concerning violations and humanitarian access, including information from humanitarian organizations (Sudanese Armed Forces and Rapid Support Forces 2023). The design recognized that a promise needs a way to become an instruction and that a reported breach needs a route back into the negotiation.

Read the institutional arrangement as a chain. A commitment exists at the negotiating table. A commander must transmit it. A unit must comply at a place and time. A relief organization must receive sufficiently reliable access to move. A recipient must actually obtain the goods. The chain ends at the person who needs assistance, not at the signature. An indicator that improves at the first step has recorded something real, but it has not observed the final step.

By November 7, the co-facilitators were reporting further commitments from renewed talks. These included an OCHA-led humanitarian forum, named contacts to assist aid movements, and communication between the two leaderships. The same statement said that the parties had not reached agreement on ceasefire implementation arrangements in that round. The SAF's attached commitments addressed customs clearance, visas, permits, and ninety loaded trucks intended to leave Port Sudan; the RSF's addressed coordination and protection of humanitarian operations (Jeddah co-facilitators 2023). Administrative promises and an unresolved ceasefire thus occupied the same dated record.

This is a revealing outcome for the geometric reading. A single count of agreements would increase when another document was signed. A count of named channels might also increase. Yet the record does not justify setting the entire route to "open." Different constraints remained

under different authorities, and the co-facilitators themselves distinguished assistance commitments from ceasefire arrangements. A larger institutional numerator could coexist with an unresolved implementation condition.

The mathematical representation should preserve that coexistence. For a shipment j, let P_j denote valid administrative permission, C_j the necessary command compliance, and R_j a usable physical route over the delivery interval. An admissible dispatch requires

$$D_j = P_j \cap C_j \cap R_j.$$

These symbols describe requirements, not probabilities inferred from the documents. A signed undertaking can establish evidence about P_j while leaving C_j uncertain. A road report can establish evidence about R_j at one time while leaving the next interval open. Replacing all three with a treaty indicator would assume the conclusion the delivery is supposed to test.

The November commitments also reveal why counting trucks at departure is insufficient. A prepared consignment, permission to leave a port, passage through a contested area, and delivery to an intended destination are separate observations. Ninety loaded trucks in a negotiating document are not ninety confirmed deliveries. The historical source identifies an intended movement and an administrative undertaking; a delivery ledger would establish what arrived. That distinction protects both the significance of the undertaking and the accuracy of the outcome claim.

There is a corresponding distinction between authority and willingness. A command may agree to an action but lack control over a unit. It may have control but choose not to exercise it. Those mechanisms can produce the same blocked passage and require different remedies. The signatures alone cannot identify which mechanism operated at every location. A serious case history would join the commitment to the local instructions, the reports of compliance or refusal, and the delivery record. It would retain disagreement between sources rather than smoothing it into an average level of trust.

The episode therefore supplies a specific historical result: humanitarian undertakings became more operationally detailed, while the November round still lacked agreed ceasefire implementation arrangements. It does not supply a numerical estimate of how much any one undertaking reduced violence. For this book, its force lies in the separation of conditions. The route toward relief was being negotiated through customs offices, command structures, monitoring arrangements, and

physical passages at the same time. Peaceful capacity had to exist in their intersection.

A border does not stop a dependency

South Sudan supplies a second view of the same problem: political separation does not remove the physical routes on which a budget depends. In its October 2024 mission findings, IMF staff reported that a pipeline carrying about 70% of South Sudan's oil exports had been inoperable since February. The report connected disrupted oil revenue with fiscal strain, including salary arrears, and discussed the effects of the war in neighboring Sudan alongside flooding and other shocks (International Monetary Fund 2024). The mechanism crosses a national boundary even though the commands fighting in Sudan and the authorities managing South Sudan's budget are different actors.

The World Bank's March 2025 economic monitor projected a severe contraction for South Sudan in fiscal year 2024/25 and made its recovery outlook conditional on the resumption of oil exports (World Bank 2025). A projection is a dated assessment, not a later observed outcome. What matters here is the explicitly conditional structure: a fiscal future depends on the operation of a route that the fiscal authority cannot restore by declaration.

That is a concrete form of constrained mobility. Oil in the ground, oil delivered to a buyer, and revenue available to pay a public worker occupy different stages of a process. A model that counts the first as if it guaranteed the third mistakes an asset for an accessible resource.

There are at least three relevant graphs. One connects production, transit, and export facilities. Another connects actors who control access, security, or contractual permission. A third connects payments and obligations. They interact, but they are not three measurements of the same edge. A physically intact route can be inaccessible. An accessible route can face a payment restriction. A payment can arrive too late to meet an obligation already due.

The distinction also prevents a misleading interpretation of interdependence. An export dependency can create an incentive to cooperate; it can also transmit a war's effects to people who did not choose the war. Whether a dependency becomes a bargaining resource or a source of collective vulnerability depends on control, substitutes, reserves, and time. The amount traded cannot answer those questions on its own.

Count what can arrive before the reserve runs out

We can make the timing problem inspectable without pretending to possess a measured Sudanese transport model. Take a small constructed supply network. A source s reaches a destination t through two intermediate depots, a and b. The directed route capacities, in consignments per day, are

$$c_{sa} = 8, \quad c_{sb} = 5, \quad c_{at} = 4, \quad c_{bt} = 6, \quad c_{ab} = 2.$$

The largest feasible delivery is ten per day. The two routes entering the destination provide a total capacity of ten, so nothing upstream can deliver eleven. Ten is attainable: send six from s to a, four from s to b, four from a to t, two from a to b, and six from b to t. Every intermediate depot balances its inflow and outflow.

Now make the route $s \to b$ unavailable. Delivery falls to six. Depot a can send four directly to the destination and two through b. Although the source-to-a route could carry eight, the usable routes beyond it carry only six. Expanding the wrong link would not remove the immediate bottleneck.

Suppose demand is nine consignments per day and the destination holds fifteen in reserve. The shortfall is three per day. With no replenishment beyond the six available deliveries, the reserve lasts five days:

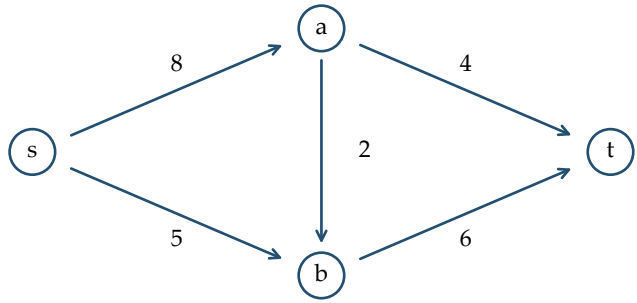
$$t_{\text{buffer}} = \frac{15}{9 - 6} = 5 \text{ days}.$$

The numerator is a stock; the denominator is a rate. Their quotient has the unit of time. A proposed substitute that becomes operational in twelve days does not solve a five-day problem unless an interim arrangement covers the gap.

The example explains why counting alternative routes is not enough. A route can exist but lack capacity. It can have capacity but become available too late. It can be available in time but impose costs that the affected population cannot bear. A viable exit is a path through all these conditions, not a line that can be drawn on a map.

The mathematics of this calculation is developed in Chapter 31. Its political meaning is already visible. Protecting a relationship may mean increasing tolerance through negotiation. It may also mean buying time with reserves, restoring an interrupted service, or creating a substitute that makes coercive dependence less costly to leave. These actions change different parts of the state. Their effects should not be hidden inside a single revised coefficient.

a Capacity is attached to a route and a direction



Capacities: consignments per day

b A reserve buys five days, not an alternative route

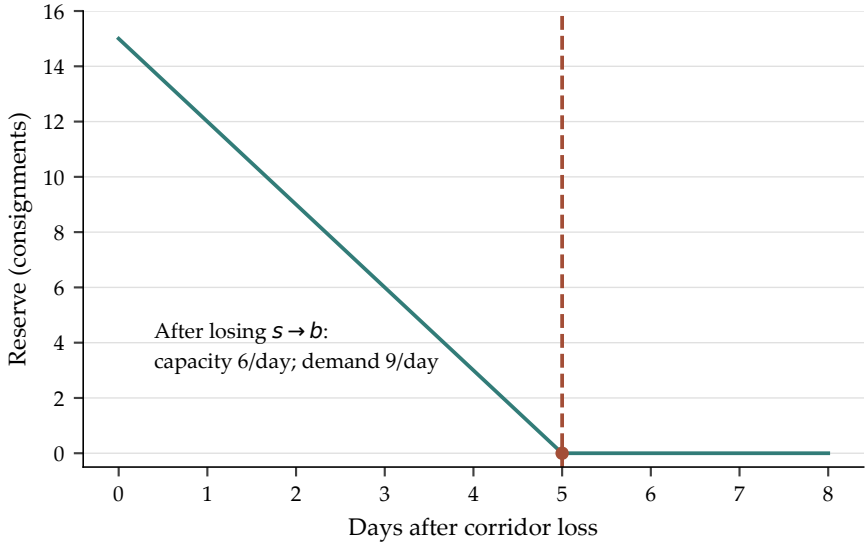


Figure 18.1: A constructed supply example, not a map or estimate of Sudanese infrastructure. Losing one route reduces maximum delivery from ten to six consignments per day. Demand of nine exhausts a reserve of fifteen in five days. The network and the clock have to be read together.

The last mile of a promise

Put a delivery clerk at the end of the agreement. A permission has been announced, a consignment is available, and people are waiting. The clerk needs to know which authorization applies today, whether the receiving facility can accept the goods, and whether the people responsible for each stage recognize the same arrangement. This is a constructed scene, but its questions are the ordinary questions of implementation. They give the word "access" a measurable object.

An access record should identify the requested operation before recording its success or failure. A permission to assess needs is different from permission to deliver a particular consignment. A delivery that reaches a warehouse is different from one that reaches the intended population. A suspended operation, a denied request, and an operation never attempted because conditions were unacceptable are different observations. OCHA's access-monitoring framework organizes constraints so that access can be studied through specified impediments rather than a single national label (OCHA 2026). The framework in this book needs the same discipline.

Suppose an analyst records twenty successful deliveries and ten unsuccessful ones. The success rate among recorded attempts is $20/30 = 2/3$. It does not follow that two-thirds of the need was met. The attempts may have carried different amounts. The hardest destinations may never have entered the attempt register. The denominator may exclude requests withdrawn before departure. Reporting only the fraction that succeeded can therefore reward a system for ceasing to attempt its most difficult tasks.

Let r_j be the amount required by receiving community j, and q_j the amount actually delivered and usable there during the stated period. One descriptive coverage measure is

$$C = \frac{\sum_j \min(q_j, r_j)}{\sum_j r_j}, \qquad \sum_j r_j > 0.$$

The cap in the numerator prevents an excess delivered to one community from numerically canceling another community's shortfall. It still does not make C a complete measure of welfare. A common unit and a justified requirement estimate are necessary, and critical differences between kinds of goods should remain visible. A tonne of one commodity cannot be substituted for a tonne of another merely to make the arithmetic convenient.

Take two destinations requiring six and four units. A delivery of eight and two supplies ten units in total, exactly the aggregate requirement. Yet the capped coverage is $(6 + 2)/10 = 0.8$. The second destination lacks half its requirement. An aggregate balance sheet says that supply meets demand; the distribution says otherwise. That difference is an essential part of the geometry. The missing two units are located somewhere, and moving them is an operation with its own constraints.

Now give the second destination a usable two-unit transfer connection. Coverage can rise to one without increasing total procurement. If the transfer is unavailable, buying more goods for the first warehouse will not repair the second destination's shortage. This is why the state must include location, permission, and receiving capacity. "More assistance" is a budget description. "Two additional usable units arriving at the second destination before its stock runs out" is a modeled change that can be checked.

The calculation identifies a civilian service problem. It neither allocates rights according to political influence nor reduces a person's value to a commodity balance. Its purpose is to stop a favorable total from hiding a concrete exclusion. The objective and the distributional constraints need to be stated together.

When the chain of authority changes

A ceasefire's implementation can be represented as a chain of necessary conditions. The negotiating authority must issue the instruction, the relevant command must receive and accept it, and the units whose conduct matters must comply. These conditions are dependent: the same disruption can affect several stages, and a successful earlier stage can change the probability of the next. Multiplying a handful of convenient confidence scores is therefore a substantive assumption.

For three events A, B, C, the exact probability identity is

$$P(A \cap B \cap C) = P(A)P(B \mid A)P(C \mid A \cap B).$$

It is exact because the later terms condition on the earlier events. Replacing those terms by $P(B)$ and $P(C)$ assumes independence. A report that three stages each appear ninety percent reliable does not establish a 0.9^3 reliability for the whole operation.

Without a dependence model, marginal probabilities alone give bounds. If each of three conditions has probability 0.9, their joint probability lies between 0.7 and 0.9. The lower bound follows because the probability of at least one failure cannot exceed the sum of the three

failure probabilities; the upper bound follows because joint success cannot be more likely than any necessary component. Independence selects 0.729 from inside that interval. It is one possible model, not the answer contained in the three marginal numbers.

These are constructed probabilities, used to explain the inference. A real study would need an operational definition of success, a record of attempts, and evidence about shared causes. In a changing command structure, yesterday's frequency may not describe tomorrow's authority. A probability estimated before an organization splits cannot be transferred mechanically to each successor.

The actor register should therefore preserve lineage. If organization A becomes A1 and A2, the data should record the date and evidential basis of that change, the commitments that each successor claims, and uncertainty about which instructions each can enforce. Historical observations remain attached to the organization that existed when they were made. They should not be silently duplicated into both successors, doubling the apparent evidence.

There is a corresponding mathematical issue when the graph changes size. A spectral statistic before a split and one afterward may refer to different state spaces. Dividing a node into two is not merely drawing a more detailed picture. It introduces an internal relation, changes degrees, and can create a slow mode that the original graph could not represent. The change might reflect newly observed fragmentation rather than new fragmentation in the world. An analyst must distinguish those possibilities.

One useful diagnostic keeps both descriptions for an overlap period. Compute the coarse representation under its old rule and the refined representation under its documented rule. Ask which conclusions survive both. If the national score improves while a newly visible local component deteriorates, that disagreement is information about resolution. It should not be averaged away before the people responsible for the decision see it.

The same care applies to a reported violation. A violation can reveal a deliberate change of policy, a command failure, an unauthorized local action, a mistaken attribution, or an observation error. Those explanations imply different repairs. A model that updates an entire population's hostility from a single unexamined report can manufacture the collapse it claims merely to describe.

Revenue is a sequence, not a substance

Return to the cross-border oil dependency. The IMF's account identifies a route interruption and fiscal consequences; it does not give the analyst permission to treat every budget movement as a direct consequence of that interruption. Price, volume, contractual shares, payment timing, debt service, and access to the receipts can all intervene. The causal account becomes more useful when it follows those stages separately.

For an illustrative accounting model, let delivered export volume in period k be v_k, realized unit revenue be p_k, and the fraction reaching the relevant public account after specified deductions be α_k. If payments arrive with a fixed delay of ℓ periods, receipts are

$$R_k = \alpha_{k-\ell} p_{k-\ell} v_{k-\ell}.$$

The formula is intentionally small. It assumes that the timing and share are known and does not include disputed receivables, multiple currencies, or variable settlement delays. Its value is that it forces the reader to distinguish production from delivery and delivery from accessible cash.

Consider an invented sequence with constant $p = 2$, $\alpha = 1/2$, and a one-period payment lag. Delivered volume falls from ten to four. Receipts remain ten in the first affected period because they pay for the previous delivery. In the following period they fall to four. A model regressing current receipts only on current delivery would misplace the first effect. A policymaker looking only at today's bank balance could mistake delayed payment for resilience.

Let the opening cash reserve be twelve and essential obligations six per period. Once receipts are four, the reserve declines by two each period and covers six such deficits. If a restoration plan takes eight periods to generate additional usable receipts, restoring physical throughput at period seven may still be too late under the assumed payment lag. The route's repair date and the budget's recovery date are different events.

There is no claim here that these invented amounts reconstruct South Sudan's budget. The example states the measurements a reconstruction would need. A serious empirical account would identify the revenue series, its reporting frequency, the legal treatment of deductions, and the distinction between accrued and cash receipts. It would also record other shocks rather than assigning every residual to the war across the border.

An informative comparison could examine a period with a similar price change but functioning delivery, or a period with disrupted delivery but temporary financing. Such comparisons test the channel. If alternative

financing sustains obligations through the interruption, the structural account should recognize the buffer. If financing exists on paper but cannot be used for the obligation concerned, the model must record that constraint rather than congratulating itself on a large nominal balance.

This makes the proposed connection more demanding, but also more humane. The object is no longer an abstract national loss. It is the path from a delivered export to a public account to an obligation paid on time. Every stage is a place where the path can continue, fail, or be repaired.

An agreement can fail locally while surviving nationally

A national ceasefire indicator is often binary because a dataset needs a tractable field. Implementation is rarely so simple. Imagine four areas covered by one agreement. Three sustain the agreed service; the fourth loses it. A national majority rule records success, while an all-areas rule records failure. Neither rule is neutral. Each corresponds to a different question.

For a national institutional study, the continued existence of the agreement may be the relevant outcome. For a civilian-access study, the fourth area's loss remains an outcome even if the agreement survives elsewhere. For a forecast of renewed organized violence, the coding must follow the chosen event definition. These targets can be studied together, but they should not replace one another when a result becomes inconvenient.

A useful reporting design keeps a small set of levels: the agreement's formal status, its observed implementation by area, and the functioning of the mechanism for responding to breaches. The third level is especially important. A durable arrangement need not eliminate every incident immediately. It may contain a credible way to investigate an incident, correct the conduct, and prevent a local violation from destroying every other channel.

That is a practical interpretation of recoverability. A damaged bridge may be repaired before traffic diverts permanently; a disputed instruction may be clarified before reciprocal retaliation expands. Recovery requires capacity and time. It also requires participants who still regard the repair procedure as usable. A hotline with nobody authorized to answer it is a listed connection without the operation it was meant to support.

The analyst can ask a finite set of questions after each observed breach. Was it reported through the agreed channel? Was the report acknowledged? Was its status investigated? Did the affected operation resume? Were the next decisions made under the agreement or outside it? The responses form a sequence. Counting only the number of

breaches loses the distinction between a process that repairs itself and one that accumulates unresolved failures.

This is the added fidelity Sudan demands from the framework. A state boundary is too coarse for some promises; a national total is too coarse for some shortages; an announcement date is too early for some recoveries. The mathematics earns its place when it preserves those differences. It helps the reader follow an instruction, a consignment, and a payment far enough to see whether the promise reaches the person for whom it matters.

What would constitute an informative test?

A useful Sudan study would preserve the distinction between the onset of fighting, the spread of fighting, and disruption transmitted beyond the battlefield. It would construct a dated actor register, identify the commitments being modeled, and record which sources support each claimed change in control or access. It would distinguish a reported closure from a measured throughput reduction and both from a missing observation.

For the cross-border fiscal mechanism, the proposed test is narrower. Does disruption of a specified export route help explain subsequent revenue shortfalls after accounting for prices, output, other transport routes, and reporting delays? Which part of the chain was visible before the shortfall occurred? Do reserves or alternate access change the timing? A successful answer would establish a particular dependency mechanism. It would not automatically establish the model's ability to forecast the outbreak of war.

The observations could challenge the account in useful ways. Revenue might remain stable because reserves, credit, or another route absorb the interruption. A nominally reopened route might fail to restore cash receipts. Two administrations with similar physical dependencies might respond differently because their obligations and political authority differ. Each outcome tells us which part of the assumed connection needs to be represented more carefully.

Local reporting, civilian organizations, and people keeping services running can identify distinctions a national dataset has erased. Their contribution is evidential as well as ethical. A model that misses the institution delivering water, the family unable to receive a salary, or the authority controlling passage may be mathematically tidy and operationally empty.

Sudan deepens the book's starting question: whose relationship are we measuring, what can move through it, and who can make a promise take effect? Those are the relations through which an action becomes possible.

Chapter 19

The Horn: Peace Has More Than One Boundary

An international telephone call becomes possible again. A family can imagine a visit. A government announces that a state of war has ended. These are meaningful changes in the world. They do not establish that every conflict within or around the two countries has ended with it.

The Horn of Africa makes that distinction unavoidable. The Ethiopia–Eritrea war began in 1998. The December 2000 Algiers Agreement created institutions for resolving boundary and claims questions; the boundary commission's task was to delimit and demarcate the border under the agreement's specified legal framework (Permanent Court of Arbitration 2026). In July 2018, Ethiopia and Eritrea signed a joint declaration ending their state of war and committing to renewed diplomatic relations, transport, communications, and economic links, as well as implementation of the boundary decision (Eritrea and Ethiopia 2018).

In November 2022, a different agreement addressed a different conflict. The Ethiopian federal government and the Tigray People's Liberation Front signed the Pretoria agreement for a permanent cessation of hostilities (Ethiopian Government and TPLF 2022). The distinction between the signatories is fundamental. An agreement between two states and an agreement between a federal government and an armed political organization do not close the same set of disputes.

The analytical question is not whether the word "peace" was sincere. It is what each agreement changed, for whom, at which boundary, and through which practical operations. The historical window here follows the November 2022 agreements through an initial medical delivery. Their existence is evidence of commitments made; the delivery supplies a separate observation of implementation, without settling the subsequent history.

Thirteen days, two agreements, and a medical delivery

The November 2022 sequence provides a concrete way to follow a commitment into an observed action. The Pretoria agreement was signed on November 2. On November 12, senior commanders signed a declaration in Nairobi that specified implementation. On November 15, the ICRC reported that two trucks carrying forty tons of medical supplies had reached Mekelle, its first such delivery since fighting resumed in August (Ethiopian Government and TPLF 2022; Ethiopian National Defense Force and Tigray Armed Combatants 2022; ICRC 2022). These are three different observations: an agreement, an implementation instrument, and a completed movement.

The Nairobi declaration attached order and conditions to military and civilian tasks. It prescribed orientation and disengagement steps, linked heavy-weapons disarmament to the withdrawal of foreign and non-ENDF forces, and provided for an African Union monitoring and verification team. For humanitarian movement, it specified federal inspection and notification, security guarantees on request, and two checkpoints at the loading and destination ends (Ethiopian National Defense Force and Tigray Armed Combatants 2022, arts. 2, 4–5). The document did not treat signing, disarmament, verification, and access as interchangeable events.

That is why the sequence is more informative than a single peace-agreement dummy variable. A negotiation creates an authorized course. The implementation instrument specifies how parties are to move along it. A delivered cargo shows that a particular operation became possible. It does not prove that every promised operation occurred. The ICRC's announcement described urgent shortages and the need to sustain and enlarge deliveries (ICRC 2022). Success at one moment was a beginning that needed repetition.

Consider the date arithmetic carefully. November 15 is thirteen days after the Pretoria signature and three days after the Nairobi signature. Those are elapsed intervals between documented events. They are not a measured causal delay attributable solely to either agreement. Cargo preparation, road access, administrative work, and security arrangements could have different starting times. The sequence gives the historian a set of questions about those intervening operations; it does not permit the analyst to assign the entire interval to one coefficient.

For the humanitarian movement, the observed path is stronger evidence than a promise that a path will exist. Let $R(t)$ mean that the particular delivery task can be carried out at time t. The completed arrival supports an observation of feasibility for that consignment and

interval. It does not imply $R(t) = 1$ for every later day, every route, or every kind of cargo. The scope of the observation matters as much as its success.

The disarmament provision makes a different geometric point. When one action is explicitly tied to another, the two cannot be modeled as independently available switches. Write d for a disarmament step and w for the corresponding withdrawal condition. A staged process may admit pairs (d, w) satisfying a joint requirement while excluding an isolated advance of d without w. The admissible region is then a subset of the product of two separate action lists. A schedule that moves through an excluded pair cannot be justified merely because its eventual endpoint satisfies both requirements.

This is an interpretation of the agreement's conditional structure, not a claim that the two variables capture every military or political concern. Its practical benefit is that it tells the analyst what evidence to seek. What did each party understand the concurrent requirement to mean? How was it verified? What happened when an actor outside the two signatories could affect the condition? A two-party signature does not give either party control over every actor whose conduct matters to implementation.

There is also a human reason to retain the separate clocks. A patient waiting for supplies experiences the time to a usable delivery. A commander evaluates the order of reciprocal steps. A monitor needs time to observe and verify. A model with only one aggregate rate can give all three the same clock and thereby make one of their problems disappear. The thirteen-day sequence does not solve those differences; it makes them visible.

The episode closes at the November 15 delivery, rather than claiming a final settlement of every subsequent dispute. Within that window, the record supports a meaningful transition from written commitment to specified implementation and at least one observed humanitarian operation. That is the level at which the framework can explain what changed: a particular peaceful action became executable. Establishing durable regional peace requires a longer and wider record.

One improved relationship cannot stand for the whole region

A country-level graph can represent Ethiopia and Eritrea as two nodes. It may record an improvement in their formal relationship. To study a conflict involving Tigray, the graph must also represent the relevant internal actors and institutions. A conclusion about one edge cannot be transferred to the enlarged graph without another calculation.

Take a schematic three-actor example. A and B cooperate closely. Both oppose C. A signed diagram has one positive edge and two negative edges. In the conventional language of signed-network balance, this triangle is balanced: the product of its three signs is positive. Yet the drawing contains two hostile relationships. "Balanced" in that mathematical sense does not mean peaceful, fair, or safe.

The example is deliberately schematic. It does not assign a permanent sign to any people or claim to reconstruct a particular military coalition. Its purpose is to expose a dangerous translation error. A technical property can be correctly computed and disastrously misnamed when it returns to public language.

For a signed adjacency matrix with entries $a_{ij} = s_{ij}w_{ij}$, where $w_{ij} \geq 0$ and $s_{ij} \in \{-1, +1\}$, one possible signed Laplacian uses absolute weighted degrees. Its quadratic form is

$$x^T L_{\pm} x = \sum_{i<j} w_{ij}(x_i - s_{ij}x_j)^2.$$

A positive edge penalizes different coordinates. A negative edge penalizes equal coordinates. An arrangement in which two camps oppose each other can therefore have a perfectly coherent low-energy representation. The operator describes agreement within camps and opposition between them. Its energy is not a measure of human suffering.

This is why the book keeps the meaning of an operator beside its formula. A low unsigned connectivity gap, a low signed disagreement energy, and a high diplomatic tolerance score answer different questions. Combining their favorable-looking values without specifying the desired outcome would reward the appearance of mathematical agreement while abandoning the object of the study.

The same discipline applies to the term "stability." A border arrangement can be stable while a population remains exposed to coercion. A command structure can be stable while obstructing a settlement. The modeled equilibrium and the condition we want people to live in must be distinguished explicitly.

A commitment becomes a sequence of operations

The 2018 declaration named several forms of reconnection. That makes it valuable for a geometric reading: diplomacy, communication, transport, and trade are distinct channels, even when one document addresses all of them. Their implementation can occur at different speeds. A restored

telephone connection does not certify an open commercial route, and a commercial route does not certify a trusted security arrangement.

Think of a family planning a journey after a border announcement. Permission to cross is one requirement. A working route is another. Documents, fares, safety, and confidence that return will remain possible matter too. The family confronts an intersection of conditions. A national indicator that changes from zero to one on the announcement date captures an institutional event; it does not automatically measure the journey's feasibility.

A research design can represent this without pretending to know everything. Let $d(t)$ record the dated diplomatic commitment, $c(t)$ the status of a communication channel, and $r(t)$ the measured availability of a route. The vector

$$z(t) = \big(d(t), c(t), r(t)\big)$$

retains the separate observations. A scalar average loses the distinction between a promise, a functioning service, and physical access. Keeping the vector is especially valuable during a transition, when those components are moving at different times.

If the model needs one decision variable, it should derive that variable from the task. For a journey requiring all three conditions, an indicator of simultaneous feasibility may be appropriate. For a study of communication recovery, the route component may be irrelevant. For a regional financial model, both may require different weights and lags. There is no universal average that becomes correct by being dimensionless.

The order of implementation can matter even when the final list of changes is identical. Verification before a risky move changes what the move means. A move before verification can expose one party while leaving the promised reciprocal action uncertain. The issue is not merely that one event happened earlier. The first event changes the set of decisions available when the second arrives.

Chapter 30 gives a small matrix calculation of this fact. Opening A–B and then B–C moves an initial disturbance differently from opening B–C and then A–B. Neither sequence is mysterious. In the second sequence, the first connection may be available before anything has reached it. The whole-period aggregate records the same two connections and misses the difference.

The agreement and the evidence have different clocks

An agreement is dated. Its effects arrive unevenly. Reports of those effects arrive later still. A credible evaluation needs all three clocks.

Suppose a monitoring report published in June says that a passage had been usable since April. A retrospective account of April may properly include that finding. A forecast issued in April cannot use the June publication unless the same information was independently available then. If the study fails to retain the publication date, the model acquires hindsight without any change in its equations.

The reverse error also matters. A report arriving late does not mean the underlying improvement occurred late. Researchers trying to measure implementation may need to reconstruct the best-supported event date. Researchers trying to measure forecasting must reconstruct the information available at issuance. These are both legitimate tasks. They require different datasets.

The African Union's role in the Pretoria process also reminds us that institutional connections should be modeled as actions, not merely memberships. Convening negotiations, transmitting proposals, observing compliance, and sustaining a process are different functions. Their presence can enlarge the set of feasible next steps. Counting how many organizations appear on an actor's membership list does not measure whether those functions were available in the crisis being studied.

An implementation study would record the agreement's stated obligations, the responsible actors, the observable milestones, and the evidence needed to consider a milestone met. It would keep disputed status as disputed. A coding rule that automatically turns every announcement into successful implementation would manufacture a peaceful trajectory out of documents.

A border opening is a process

Consider a family trying to make the first visit after a long separation. An announcement reaches them on Monday. On Tuesday they learn that their documents are accepted, but they still do not know whether transport is operating or whether they can return. A model that labels the border "open" on Monday has recorded the announcement correctly and described the family's options incompletely.

This is an illustrative journey, not a report of a particular family's experience. It gives a concrete meaning to a distinction that otherwise disappears into institutional language. The 2018 declaration named several channels of reconnection. Each channel presents an implementation question. The family cannot use the average of a functioning telephone and an unavailable crossing. It needs the conditions required by its particular journey to hold together.

Let T_d, T_r, and T_s denote the times when documentation, route availability, and a specified service become usable. If all three are necessary and remain usable afterward, the earliest feasible start is

$$T_{\text{start}} = \max(T_d, T_r, T_s).$$

If the three times are two, five, and eight days after the announcement, the start is day eight, not the average day five. Improving documentation from day two to day one does not advance the journey. Advancing the service from day eight to day six advances it by two days. Advancing it beyond day five then encounters the route as the new limiting condition.

This is a simple critical-path calculation. Its assumptions deserve attention. Availability may be temporary. A route may open for a limited period and close before another necessary condition is met. Then taking the maximum of first opening times is insufficient; the intervals of availability must overlap. Let J_d, J_r, J_s be those intervals. A possible departure requires

$$J_d \cap J_r \cap J_s \neq \varnothing.$$

If documents are accepted on days one through ten, transport runs on days four through six, and the destination can receive visitors from day seven, each condition has been satisfied at some time. They have never been satisfied together. The three announcement headlines describe progress; the intersection describes the unavailable journey.

Duration matters as well. A departure on the last usable day may not allow completion or return. The state must represent the operation being promised. This is one reason to resist treating every institutional improvement as an identical increase in a scalar tolerance budget. Some improvements create time, some extend the range of eligible people, and some change the cost. Their usefulness depends on what they make possible together.

For the Horn, a careful study would follow separate implementation histories for the channels the agreements name. It would state which actors and locations each observation covers. It would preserve an opening that later closes as a temporary opening, rather than coding either permanent success or complete failure. The trajectory contains information that its final label cannot retain.

The triangle that looks settled

The signed triangle introduced earlier can be solved completely. Give all three edges unit weight. Let A and B have a positive relation and let

each have a negative relation to C. With absolute degree two at every node, the signed Laplacian is

$$L_{\pm} = \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

For $x = (1, 1, -1)^T$, multiplication gives $L_{\pm}x = 0$. The two positive coordinates agree, and the negative coordinate opposes them. Every term in the signed energy vanishes. The eigenvalues are $0, 3, 3$.

The vector is a witness: a small object that lets the reader verify the mathematical claim directly. It is not evidence that an actual regional coalition has those exact relationships. It establishes something narrower and important about the language of the model. Zero signed disagreement energy can describe organized opposition between two camps.

In contrast, the ordinary unsigned triangle has a zero mode $(1, 1, 1)^T$. All coordinates agree. The two matrices share the eigenvalues $0, 3, 3$, yet the meaning of the zero mode differs because the signs change the relation being penalized. A report containing only the spectrum would lose that distinction. The coordinates and the definition of the operator have to remain beside the numbers.

There is a short way to understand the shared spectrum. Let $S = \mathrm{diag}(1, 1, -1)$. Multiplying the unsigned triangle Laplacian on both sides by S produces $L_{\pm}$. Since $S^{-1} = S$, the two matrices are similar and have the same eigenvalues. This change of signs is sometimes called a switching transformation. It changes the interpretation of agreement without changing the algebraic spectrum. Signed-network research provides the broader setting for this distinction (Altafini 2012).

Now make all three relations negative. The signed Laplacian becomes

$$\widetilde{L}_{\pm} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix},$$

whose eigenvalues are $1, 1, 4$. There is no nonzero vector satisfying every pairwise preference simultaneously. If A must oppose B and B must oppose C, then A and C acquire the same sign, contradicting their required opposition. The positive minimum energy records this incompatibility in the stipulated relation model.

Neither calculation tells us which arrangement is politically preferable. The first graph can represent a coherent exclusionary coalition.

The second can represent mutually conflicting demands. Mathematical frustration and civilian danger are different quantities. A useful theory identifies the relation between them through evidence rather than treating an appealing technical word as a verdict.

This is why I want the reader to be able to work the triangle by hand. The caution is not a retreat from mathematics. It is a consequence of doing the mathematics fully enough to see what was actually computed. Geometry can clarify a political relationship only if the analyst keeps its interpretation attached.

Verification changes the next decision

A settlement often requires an action whose benefit depends on another action that has not yet occurred. One party changes an administrative procedure; another restores a service; a third confirms that the promised change took place. The sequence affects what each participant can responsibly do next. The problem is familiar to anyone who has arranged an exchange in which both sides fear being the only one to perform.

For a deliberately simplified example, call the first reciprocal implementation step a and the second b. An immediate simultaneous exchange might be infeasible because the parties cannot observe one another's completion quickly enough. A verified sequence can introduce a third operation v: confirm a, then authorize b. The verifier does not create goodwill by arithmetic. It supplies an observation that changes the information available for the next decision.

Write the second party's choice as $b = \pi(\mathcal{I})$, where $\mathcal{I}$ is the information available when it must act. Before verification, two states of the world may be compatible with the information: the first step happened, or it did not. After a reliable observation, one of those possibilities can be excluded. The decision rule may then permit a step that would otherwise expose the participant to an unacceptable risk.

The gain depends on what the observation verifies. Confirmation that an order was issued does not establish that it was obeyed. Confirmation at one location does not establish compliance everywhere. A verification mechanism needs a defined proposition, an observation procedure, and a rule for what happens when the proposition remains unresolved. A flag marked "verified" without those details merely moves the uncertainty into a hidden column.

The proposed evidence register can keep three statuses: supported, contradicted, and unresolved. Unresolved is not a compromise halfway between the other two. It says that the available observations do not decide the proposition. A decision rule may still permit reversible actions

under unresolved status, while deferring actions whose consequences cannot readily be repaired. The choice depends on the task and the cost of delay.

This introduces a useful distinction between trust and exposure. Trust can improve while exposure remains too large for a particular first step. Reducing the size of that step, creating a reversible trial, or arranging independent observation can make a sequence feasible without requiring either party to claim that distrust has vanished. The geometry of the process has changed because the permitted next moves have changed.

The Pretoria agreement belongs to a different conflict and actor set from the interstate declaration. A study of its implementation must read its own obligations and monitoring provisions, rather than importing an observation rule from the earlier agreement. Agreements are not interchangeable units of peace. Their design specifies who must act, what must be observed, and which dispute-resolution operations are available.

An analyst should also record who is absent from the information structure. A process can verify obligations among signatories while failing to observe harms experienced by people outside the negotiating room. Adding a channel for those observations can change the assessment of implementation even when the formal parties report compliance. The missing voice is then a missing measurement as well as a political exclusion.

The difference between a repaired breach and an accumulating one

Suppose two arrangements each record twelve breaches in a year. In the first, each breach is acknowledged, investigated, and resolved before the next. In the second, the breaches remain unresolved. A yearly count makes the arrangements look alike. A stock of unresolved obligations distinguishes them immediately.

Let u_k be the number of unresolved, comparably defined cases at the start of period k, b_k newly admitted cases, and r_k cases resolved under the agreed standard. Then

$$u_{k+1} = u_k + b_k - r_k, \qquad 0 \le r_k \le u_k + b_k.$$

The definition of resolution matters. Closing a file administratively is not necessarily restoring the affected operation or satisfying the agreed remedy. A coding rule should explain whether its endpoint is acknowledgment, investigation, remedy, or acceptance by the parties.

Take a constructed sequence with two new cases each month. If three can be resolved per month whenever cases are available, a temporary backlog can shrink. If only one can be resolved, the backlog grows by one per month. Nothing in this arithmetic says that the twelfth case must cause war. It says that the repair process is failing to keep pace with the process generating obligations. That is a measurable change in the relationship's ability to recover.

The waiting time may matter more than the count. A long unresolved interruption can accumulate consequences even if few new cases occur. One possible descriptive measure is the total number of unresolved case-days, obtained by summing each case's unresolved duration. It gives more weight to persistent failures without pretending that all breaches have equal severity. A serious study would retain severity and affected population separately.

There are also dangers in rewarding fast closure. A mechanism might improve its reported time by dismissing difficult claims or discouraging reports. The case register must preserve admissions, withdrawals, rejections, and grounds for closure. Otherwise an apparently efficient repair institution may simply be less willing to hear the people whose relationships it was meant to repair.

The framework's idea of tolerance acquires operational content here. It is partly the ability to encounter a breach without losing every channel through which the breach could be addressed. That ability depends on staffing, authority, legitimacy, information, and the remaining willingness to use the process. A count of diplomatic meetings does not measure it directly.

What an honest comparison would preserve

To study whether reconnection improves resilience, compare trajectories at a level where implementation is observable. Choose the outcome before examining the results: functioning communication, completed civilian journeys, resolved disputes, or a defined form of violence. State which channels are hypothesized to affect that outcome and how quickly their effects could appear.

Keep at least three dates for each milestone: the effective date claimed by the source, the date of the observation supporting it, and the date the analyst could have received the report. If a channel's availability is disputed, preserve the competing accounts. A best-supported historical reconstruction and an information-available forecast can then be built from the same register without pretending that they are the same exercise.

An observed improvement following an agreement is evidence of sequence, not sufficient evidence of causation. Conditions may have improved before the agreement, or both may reflect another change. Comparisons need to examine those possibilities. The causal chapter later in the book gives the formal language; the Horn supplies the practical reason to care. We want to know which operations made a future more accessible, not merely which ceremony occurred before it did.

The strongest result might be narrower than a sweeping regional claim. A study could show that a particular communication channel reduced the time needed to resolve a class of disputes. Another could show that a temporary opening made some journeys possible while leaving others excluded. Those findings would still matter. They identify a mechanism that people can inspect, sustain, and improve. A peace process becomes intelligible through the operations by which it reaches ordinary life.

What the Horn asks the framework to explain

The first question is scale. Does a model of interstate relations preserve internal conflicts relevant to the claimed outcome? If it does not, it should restrict its prediction to the interstate target or expand its actor representation. Quietly changing the target after an event is not an acceptable repair.

The second question is transfer. When a diplomatic agreement improves one relationship, which other relationships change with it? A model may hypothesize that improved communication reduces uncertainty. It may hypothesize that a changed alignment shifts pressure elsewhere. These are different mechanisms, requiring different observations. The fact of a treaty cannot choose between them.

The third question is implementation. Which elements of the promised change happened, in what order, and with what ability to withstand a violation? A practical tolerance budget should respond to a functioning dispute-resolution mechanism differently from a document that cannot resolve disputes. The model gains fidelity when it distinguishes the capacity to repair a breach from the mere absence of a recorded breach.

The fourth question is whose safety defines success. A regional account centered on capitals can miss the experience of people near a contested boundary or within an area governed through a different chain of authority. A useful geographical resolution is one at which the relevant constraint becomes visible. Increasing the number of pixels without improving that representation is not a solution.

These questions do not diminish the significance of negotiated peace. They explain why it must be studied carefully. An agreement can reopen a future that had become inaccessible. The work of analysis is to follow that opening into institutions, routes, obligations, and lived possibilities, while continuing to observe the boundaries the agreement did not resolve.

Chapter 20

The Great Lakes: Connected to What?

A busy road is evidence that something is moving. It is not evidence that the people beside the road are safe.

This distinction matters throughout the book. In the Democratic Republic of the Congo and its regional relationships, it becomes a direct challenge to any theory that treats connectivity as an uncomplicated good. Commerce, political authority, military support, and civilian access can use some of the same places while serving sharply different purposes. The problem is not that a network description is impossible. The problem is that one network is insufficient.

There are legal and historical reasons to keep the actors distinct. In its December 2005 judgment in *Armed Activities on the Territory of the Congo*, the International Court of Justice found that Uganda's military activities, occupation of Ituri, and support to irregular forces violated obligations owed to the DRC. It also found violations by the DRC of diplomatic obligations owed to Uganda (International Court of Justice 2005). Those findings concern particular conduct and legal responsibilities. They should not be flattened into an undifferentiated label of regional hostility.

Two decades later, Security Council Resolution 2773, adopted on February 21, 2025, addressed the M23 offensive in eastern DRC, including its capture of Goma and Bukavu. The Council identified direct support and participation by the Rwanda Defence Forces and demanded changes in that conduct (United Nations Security Council 2025). The resolution provides a dated institutional assessment. It does not supply a complete causal model of the conflict or authorize treating every connection involving the named states as a military relationship.

The research task is to represent the relationships relevant to a specified outcome without transferring responsibility from identified organizations to entire populations.

Rubaya: when the mine keeps producing

The Rubaya episode makes the difference between economic activity and peaceful connection unusually concrete. The UN Group of Experts reported that AFC/M23 took control of Rubaya and its surrounding mining area on April 30, 2024, while mineral extraction continued. Its December report described control of production, trade, and transport, including taxation and a route to Rwanda (UN Group of Experts 2024, paras. 56–66). The economic network did not simply vanish when control changed. Continuing activity became part of the mechanism under investigation.

An earlier, September 30 Security Council briefing by MONUSCO's head, Bintou Keita, estimated that the Rubaya trade generated about $300,000 per month for the armed group (Keita 2024). The later Group of Experts report estimated roughly 120 tons of coltan moving monthly, a levy of $7 per kilogram on coltan, and at least $800,000 in monthly receipts from coltan taxation (UN Group of Experts 2024, paras. 64–66). These are dated institutional estimates, with different reported bases. They should not be spliced into a precise monthly growth series.

The later estimate has an inspectable scale. Multiplying 120 metric tons by 1,000 kilograms per ton and $7 per kilogram gives $840,000 per month. That arithmetic is consistent with an estimate described as at least $800,000. It does not independently validate the underlying volume or levy. Nor does it turn an estimate into an audited cash account. The calculation checks that the units and order of magnitude fit the report.

The source distinction is itself part of the analysis. A reader seeing $300,000 in one document and more than $800,000 in another might infer that receipts almost tripled. That inference would require comparable coverage, collection rules, time intervals, and observation methods. Without those, the difference could reflect revised information or a broader assessment as well as a real change. The later report can improve a retrospective explanation while remaining unavailable to an analyst working in September.

What does the case imply for the thesis? It gives the words "trade persists" an identifiable mechanism. A measure of commercial volume can remain substantial while the authority collecting a payment, controlling passage, or enforcing participation changes. A peaceful-interdependence interpretation that sees only the volume has omitted who acquires power through the transaction. The topology needed for the political question must therefore include control and receipt, not merely origin and destination.

One way to keep those distinctions is to record a consignment through several linked ledgers. A physical ledger follows material from extraction to transport and sale. A payment ledger follows charges and receipts. An authority ledger identifies who can impose a charge or interrupt a movement. They share actors and places, but they are not one adjacency matrix. A paid charge is not identical to a command relationship, and a documented route is not sufficient proof of who ultimately benefits from every transaction.

The mechanism also prevents a simple policy shortcut. Reducing a reported trade total is not by itself a measure of reduced armed financing. Activity can move to a less visible route, a different payment channel, or another source of revenue. At the same time, interruption can damage civilian livelihoods. An intervention analysis must specify whose access and whose receipts it intends to alter, then observe both. The relevant objective cannot be inferred from the sign of the change in aggregate connectivity.

For a formal statement, let q_e be observed throughput on edge e and let r_{ae} represent the portion of a specified payment associated with actor a. An aggregate throughput $\sum_e q_e$ loses the allocation information in the collection (q_e, r_{ae}). Two networks can have the same total movement and different recipients. Even perfect measurement of the total cannot recover an allocation it did not retain. That is an information loss caused by the representation, not noise that a larger sample automatically removes.

The documented outcome in this episode is continued extraction and trade under changed armed control, with UN investigators attributing substantial tax receipts to AFC/M23. It supports the narrow mechanism that a functioning economic connection can sustain coercive power. It does not establish that minerals alone explain the wider conflict, or that the framework forecast this change before April 2024. Its contribution is sharper than either claim: connectivity must be read together with the distribution of authority and proceeds.

A flow can sustain the thing we want to stop

Start with the most elementary distinction. A route carrying food to civilians and a route carrying resources that sustain an armed organization may both increase a measured transport total. Their implications for civilian protection can differ. Even within one route, commercial users, local communities, public authorities, and armed groups may have different control over the passage and its proceeds.

A model that asks whether aggregate trade is rising has not yet asked what that trade enables. This is a limitation of the question, not a failure of arithmetic.

The appropriate representation records an edge's function. A payment edge may indicate financing. A transport edge may indicate physical access. A command edge may indicate the ability to direct conduct. A protection arrangement may affect whether civilian movement is possible. Combining these into an untyped adjacency matrix hides the mechanism the graph was supposed to reveal.

Suppose a constructed regional economy has two districts. Recorded commercial activity in one rises from 20 to 40 units while activity in the other falls from 80 to 60. The national total remains 100. An analyst looking only at the total sees no change. A household whose only usable route is in the declining district experiences a substantial change. The national series is correct; its adequacy for the household's problem is not.

The same issue affects a conflict forecast. If the outcome is local violence, national aggregation can mix a deteriorating locality with an improving one. If the outcome is interstate escalation, a local event series may need an explicit model of how incidents connect to decisions at the state level. Moving between scales is an inference, not a formatting operation.

We can state the aggregation problem algebraically. Let x collect local conditions and let P add them into regional totals. If local dynamics are $\dot{x} = Gx$, then the totals satisfy

$$\frac{d(Px)}{dt} = PGx.$$

A closed regional model $\dot{y} = \bar{G}y$ with $y = Px$ exists for all local states only when

$$PG = \bar{G}P.$$

This condition says that the regional total contains enough information to determine its own evolution under the proposed model. If two local arrangements have the same total but different next-step consequences, no single total-only rule can represent both exactly. A model can still approximate the aggregate evolution, but its missing information has to be acknowledged and tested.

For a concrete illustration, put ten units of a disruption in either of two localities. Both arrangements have regional total ten. If the first locality has an effective recovery rate of one per day and the second of one-tenth per day, their initial total recovery rates differ by a factor of

ten. The total cannot identify which situation exists. Knowing where the disruption sits is part of knowing what it can do next.

What a missing report means

Conflict reporting is itself a network of access. An event must occur, be observed, reach a reporter or institution, survive verification, and enter a release. The probability of completing that chain need not be constant across places or time.

Imagine two constructed periods. In the first, 100 events occur and 40% are recorded, producing 40 records. In the second, 80 occur and 80% are recorded, producing 64. Recorded events rise by 60% while underlying events fall by 20%. Nothing contradictory has happened. Observation changed along with the object being observed.

One simple observation model is

$$Y_t \mid N_t, \rho_t \sim \text{Binomial}(N_t, \rho_t), \qquad E[Y_t \mid N_t, \rho_t] = \rho_t N_t.$$

Here N_t is the number of underlying events and ρ_t is a recording probability. The formula is an illustrative thinning model. Actual events can share sources and verification processes, so independent recording is an assumption to examine rather than a description to impose.

The important identification problem appears even at the level of the mean. One observed product $\rho_t N_t$ cannot separately determine both factors. A fall in reports can indicate fewer events, less access, or some combination. To distinguish them, the study needs additional information about coverage or a justified bound on it.

This is where local expertise changes the mathematics. Knowledge that reporting access collapsed is information about the observation process. Knowledge that an apparent new actor name denotes an existing organization under a new label is information about identity continuity. Neither is an optional narrative ornament attached after the model runs. Each can change the data-generating interpretation.

The UCDP Georeferenced Event Dataset distinguishes types of organized violence and provides coded temporal and spatial precision. Its versioned codebook is part of the specification of an analysis (Högbladh 2025). A location code should not be interpreted as more precise than the evidence supporting it. A later revised release should not silently become the information available to an earlier forecast.

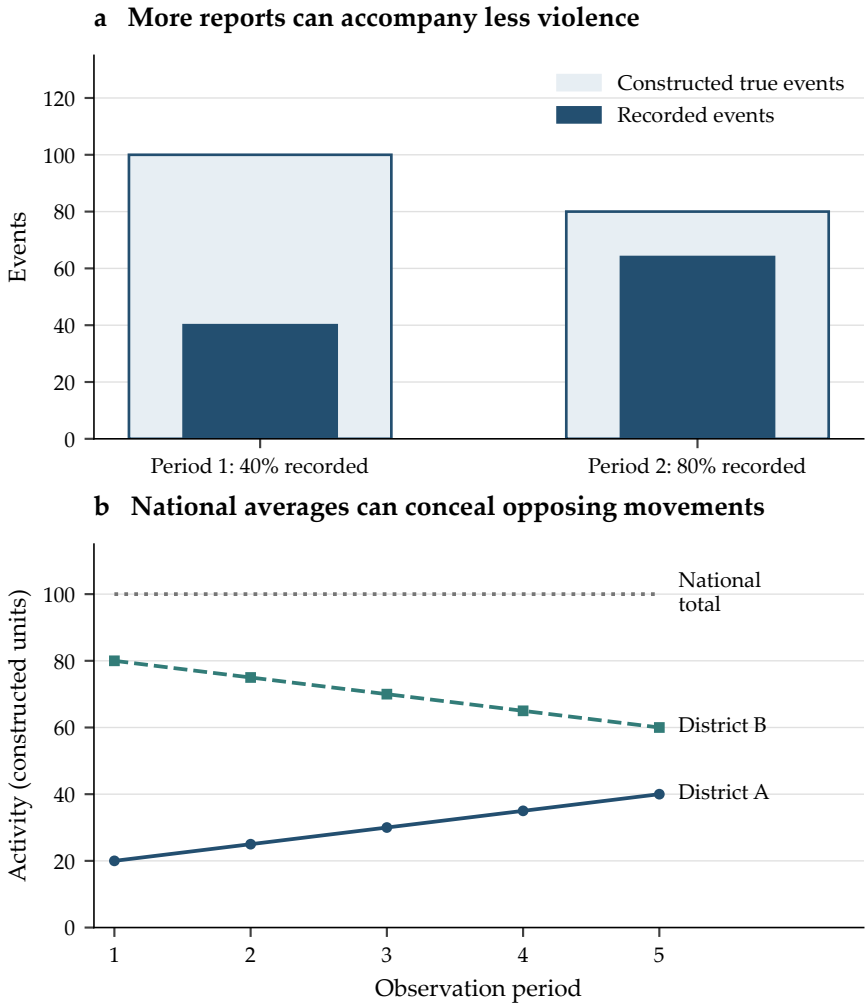


Figure 20.1: Two constructed examples of measurement failure. Above, improved coverage makes recorded events rise while underlying events fall. Below, a constant national total conceals opposing district trajectories. Neither panel is an estimate for the DRC.

The legal edge and the statistical edge

The cited judgment and Security Council resolution do a different kind of work from an adjacency matrix. They attribute conduct and articulate obligations under institutional procedures. A statistical model estimates relationships under a specified observation process. Neither replaces the other.

For example, a graph may show that two recorded activity series move together. That does not establish who directed whom. A legal finding about support does not by itself identify a numerical coefficient governing the effect of that support on subsequent violence. The model needs an explicit mapping from the sourced finding to the variable used, including uncertainty and the period to which the finding applies.

The same care is needed with economic explanations. Resource flows may be relevant to a conflict mechanism, but "resources cause war" is not an adequate model. Which revenues reach which actors? Through what institutional arrangement? What alternative finances exist? What would change if the route or payment were interrupted? The analysis becomes more informative as these questions become specific enough to be contradicted by evidence.

A dangerous shortcut is to infer a mechanism from a map of co-location. A mineral deposit, a road, and a recorded clash may appear near one another. The map can motivate a question. It does not establish the direction of causation, eliminate alternative explanations, or determine whether the road was being used in the way the hypothesis assumes.

For this reason, the Great Lakes extension should preserve separate evidence layers: dated event records, actor identities, documented relationships, civilian access, and the economic quantities actually observed. A model can test interactions among those layers. It should not begin by declaring them equivalent.

The ledger behind the line on the map

A commercial shipment can produce several records: an invoice, a customs entry, a payment, a transport document, and a receipt at the destination. Those records need not identify the same actor in the same way. The owner of the goods may differ from the carrier, the exporter of record, the recipient of a fee, and the authority controlling passage. Drawing one arrow from a mine to a border erases most of that structure.

The OECD's minerals due-diligence guidance addresses companies' need to identify and assess supply-chain risks, respond to them, and account publicly for their efforts (OECD 2016). For this book, its method-

ological importance is the movement from a country's label to a documented chain of relationships. A location in a conflict-affected area can justify investigation. It does not establish the conduct or responsibility of every person whose livelihood passes through that location.

Imagine a constructed ledger for a shipment valued at one hundred units. Seventy pays for production, twenty for transport and processing, and ten is an unidentified deduction. The unidentified amount is a question, not automatically a measured payment to an armed group. Conversely, labeling it "other costs" does not establish that it is benign. The evidence register should preserve the unresolved recipient and the documents that might resolve it.

If a model needs a range for a specifically hypothesized harmful payment, the unresolved ten can supply a crude upper bound only under the stated assumption that all such payments lie within that deduction. The lower bound may be zero. New evidence that three units reached the relevant recipient changes the supported lower bound; evidence accounting for another five through ordinary services changes the remaining uncertainty. This is a process of narrowing an evidential interval, not assigning moral character to an entire trade flow.

Let v be the shipment value and r the fraction reaching an identified actor through the mechanism being studied. Then the modeled receipt is rv. An observation of v alone does not determine r. A change in total trade can coexist with a change in the opposite direction in the relevant receipt. For example, v falling from one hundred to eighty while r rises from 0.02 to 0.05 raises the modeled receipt from two to four. These are invented amounts; they demonstrate why the recipient and terms of a connection matter.

The same observation applies to a policy intended to interrupt a harmful flow. A reduction in recorded volume is not sufficient to establish that the targeted receipt fell. Routes, prices, fees, substitutes, and reporting may all change. Nor is maintaining a large commercial total sufficient to establish that civilian livelihoods were preserved. The study needs the specific outcomes and their distribution.

This makes the Great Lakes analysis a demanding test of the framework's distinction between connection and usable freedom. A network can become busier while the population's feasible choices contract. It can become less visible while a coercive relationship persists. The direction of an aggregate change cannot carry the interpretation by itself.

How much can the missing records change the conclusion?

The observation model introduced earlier explains why a count and a recording probability cannot generally be recovered from one product. A natural next question is whether a useful conclusion survives even without an exact correction.

Suppose the expected recorded count in a constructed period is forty. Independent knowledge places the recording probability between 0.4 and 0.8. Under the mean relation $E[Y] = \rho N$, the compatible underlying count scale ranges from fifty to one hundred:

$$\frac{40}{0.8} \leq N \leq \frac{40}{0.4}.$$

This is a sensitivity interval for the mean model, not a confidence interval for a realized binomial count. Sampling variation would require another layer of calculation. The distinction matters because a neat interval can otherwise appear to include uncertainties it never modeled.

In a later period, suppose the expected recorded count is forty-eight and the same bounds apply. Its compatible range is sixty to one hundred twenty. Recorded activity rose, but the two intervals overlap extensively. The sign of the change in underlying activity remains unresolved. A declaration of deterioration would require more information about how recording changed or a stronger, justified assumption linking the two periods.

If the recording probability is known to be exactly the same in both periods, the ratio of the expected counts equals the ratio of underlying counts. Then forty-eight divided by forty indicates a twenty percent increase in the underlying rate, even though its absolute level remains unknown. Equal coverage is therefore an identification assumption with real consequences. It should be examined through source availability, collection rules, and changes in access, rather than inserted because it makes the inference possible.

Conflict-data research distinguishes omission, duplication, and misrepresentation as different routes by which a dataset can depart from the events it aims to record (Raleigh, Kishi, and Linke 2022). The simple thinning example models only omission. If ten articles describe the same incident and become ten records, increasing coverage can inflate the count through duplication. If a report is assigned to the wrong place, the total may remain correct while the local forecast becomes wrong.

A useful sensitivity analysis therefore changes the data in ways that correspond to plausible failures. Removing every tenth row at random tests random loss. It does not test a complete outage in one district,

systematic exclusion of a language, or duplication of a highly publicized event. Each failure produces a different geometry of missing information. The perturbation should represent the suspected observation process.

One practical procedure retains several defensible datasets: a strict register of strongly supported events, a broader register including specified uncertain records, and alternative assignments for ambiguous locations or identities. The model runs on each. The resulting spread describes sensitivity to those choices. It does not automatically become a probability distribution, because the alternatives may not have justified probabilities.

If all admissible versions support the same directional conclusion, the analysis has gained robustness. If the conclusion changes, the disagreement identifies evidence worth seeking. Either result is more informative than one silently cleaned file whose uncertain decisions have disappeared.

A coordinate is sometimes an area

A point on a map looks precise even when the source is not. A report may identify a town, a road segment, a district, or a larger administrative region. A dataset can attach one coordinate for practical reasons while retaining a precision code. The coordinate does not transform the original description into an exact site.

Suppose a reported event can plausibly lie anywhere in a region R. An analyst wants to count events within ten kilometers of a specified civilian service route. Let B be that ten-kilometer buffer. If R lies wholly inside B, the event belongs to the count under every admissible location. If R lies wholly outside, it does not. If the two partly overlap, the assignment is unresolved at the available spatial precision.

For a set of uncertain event regions R_i, the number inside the buffer has simple logical bounds:

$$N_{\text{certain}} = \sum_i \mathbf{1}\{R_i \subseteq B\}, \qquad N_{\text{possible}} = \sum_i \mathbf{1}\{R_i \cap B \neq \varnothing\}.$$

The true count under the stated location sets lies between them. No probability has been assigned to where an event lies within its region. If the analyst chooses to assign such probabilities, the distribution needs evidence or an explicit sensitivity interpretation.

Take six invented reports: two definitely inside, one definitely outside, and three ambiguous. The compatible count is between two and five. Placing all ambiguous reports at district centroids might return three,

but that apparent precision comes from a placement convention. A result that changes sharply between two and five needs to disclose its dependence on that convention.

The same reasoning applies to time. If a report places an event somewhere within a week, assigning it to Monday can create an apparent lead or lag relative to a Tuesday policy change. The chronology may be unresolved. An event-study plot with daily resolution should not imply a temporal ordering finer than the sources support.

This is an instance of the book's broader argument about geometry. A location's uncertainty has shape. It can cross a boundary, overlap several routes, or include places with different access. Replacing that region by a point discards information about which conclusions remain possible. More decimal places do not restore what was discarded.

Following a consequence across two scales

The condition $PG = \bar{G}P$ established when a total has closed dynamics. We can now see how an apparent improvement can arise when the condition fails. Let two local states represent accumulated disruption. Suppose the first recovers at rate one per day and the second at rate 0.1 per day, with no exchange in this example. Then

$$G = \begin{pmatrix} -1 & 0 \\ 0 & -0.1 \end{pmatrix}, \qquad P = \begin{pmatrix} 1 & 1 \end{pmatrix}.$$

The aggregate derivative is $-x_1 - 0.1x_2$. It cannot be written as one fixed multiple of $x_1 + x_2$ for every state. A fitted national recovery coefficient will depend on how disruption was distributed in the fitting sample.

Start with $(x_1, x_2) = (10, 0)$. After one day the total is $10e^{-1} \simeq 3.679$. Start instead with $(0, 10)$. The initial total is again ten, but the next total is $10e^{-0.1} \simeq 9.048$. Both trajectories satisfy the same local model. A national model trained mainly on the first locality can badly overstate recovery when disruption moves to the second.

The problem is often called a change in composition. Its mathematical consequence is that the aggregate is missing a state variable. A model can repair this by retaining the local distribution, by introducing a measured composition coordinate, or by restricting the claim to a setting in which composition remains sufficiently stable. Simply giving the fitted coefficient more decimal places will not help.

The practical analogue is immediate. A service interruption concentrated where replacement capacity is available is different from the same total interruption concentrated where alternatives are scarce. A national

trade volume, conflict count, or mean travel time can conceal the move from one situation to the other. The relevant resolution is the one that preserves the constraints governing the outcome.

There is no rule that the finest possible map is always best. Small areas can have sparse evidence, uncertain identities, and unreliable denominators. Resolution has to be chosen with both the mechanism and observation process in view. An honest regional estimate with clear limits can be more informative than a detailed map whose local precision is invented.

What would count as evidence for the mechanism?

An empirical Great Lakes study should connect a specific relationship to a specified consequence through observations that could distinguish competing explanations. Suppose the hypothesis is that loss of a civilian access channel reduces households' ability to maintain an essential service. The study needs more than a nearby rise in violence. It needs evidence about the channel's availability, the population's reliance on it, the timing of the interruption, and the substitutes that remained usable.

A strong comparison might show that similarly affected areas with functioning substitutes preserved the service while those without substitutes did not. That pattern would support the proposed role of access, but differences in wealth, governance, or external assistance could also explain it. The design must examine those alternatives. The causal analysis should be proportionate to the evidence rather than concealed inside the word "geometry."

The model can make intermediate predictions that are easier to inspect than a claim about the entire conflict. If a particular route is the binding constraint, restoring a nonbinding route should have little immediate effect under the model. If reserves explain delayed consequences, the delay should vary with usable stocks and shortfalls in the specified way. If the observed improvement is actually reporting loss, independent evidence of service functioning should fail to improve with the count. These are distinct propositions with different witnesses.

A failed intermediate prediction is valuable. It can reveal an overlooked substitute, a mistaken actor identity, an inaccessible nominal reserve, or a source process that changed. The research program should record which mechanism failed and which observation contradicted it. That is more productive than either declaring the whole theory disproved or changing the definition until the result looks favorable.

The Great Lakes case ultimately presses the book toward a richer account of interdependence. The quantity moving matters. So do

its recipient, the conditions of access, the ability to refuse, and the alternatives available to people who depend on the route. The graph becomes explanatory when those relations survive the trip from the ledger to the equation and back to the lives it is meant to describe.

A test that could change our mind

Consider a proposed mechanism: when an important civilian route becomes less accessible, nearby communities face a narrower set of feasible responses to violence. A useful study would define accessibility, identify the affected population, establish the observation clock, and examine whether plausible substitutes or reserves alter the effect. It would compare relevant quiet periods as well as crises. It would inspect whether reporting access changed at the same time as civilian access.

That mechanism could fail to explain the observations. Alternative routes may be more usable than the model assumed. Political arrangements may preserve passage despite fighting elsewhere. The modeled population may have access to resources the national series does not record. Those are informative failures because they identify the missing degree of freedom.

Alternatively, the physical route may remain open while coercion makes it unusable for the people whose options matter. That result would challenge a throughput-only measure without disproving the importance of constrained access. The geometry would have to include the restriction that the original graph omitted.

The Great Lakes cases thus sharpen the distinction between connection and freedom of movement. A system can be densely connected and still distribute its benefits, risks, and control with extreme inequality. The framework becomes stronger when it asks which connections preserve civilian possibilities, which transmit danger, and which allow an arrangement to change without catastrophe.

Chapter 21

West Africa: Leaving an Institution, Keeping a Route

A trader approaches a border with goods and a familiar passport. Somewhere else, a government has announced withdrawal from a regional organization. Does the trader still cross? Does the shipment still receive the same treatment? Does a political departure mean the road has closed?

These are questions a model must answer separately.

When the withdrawals of Burkina Faso, Mali, and Niger from ECOWAS became effective on January 29, 2025, the organization's statement explicitly distinguished membership from several continuing arrangements. It instructed authorities to continue recognizing the relevant passports and identity cards, to maintain specified trade treatment, and to allow visa-free movement, residence, and establishment under the stated transitional arrangements until further notice (ECOWAS 2025). The institutional boundary changed. The statement did not declare every commercial or personal connection severed.

This is a particularly clear test of an otherwise tempting coding rule. If withdrawal from an organization sets all edges between its former and remaining members to zero, the resulting graph records a rupture broader than the source supports. The model has converted one institutional event into several unobserved physical and social events.

The three withdrawing states are not interchangeable cases, and ECOWAS is not the only relevant institution. The point is a narrower one: formal membership, practical access, political legitimacy, and the operation of regional services are different variables. Their divergence is itself something worth studying.

What changes when membership changes?

In an institutional network, a membership edge can have a reasonably clear meaning: a state is or is not a member at a given time. In a movement network, the edge might record whether a traveler with

specified documents can enter. In a trade network, it might record observed flows or a tariff arrangement. In a security network, it might record an operative commitment. These descriptions should not inherit each other's values automatically.

Suppose an institutional departure lowers the membership component of a tolerance measure. That is an interpretable coding decision if membership is the stated input. It does not establish the total change in tolerance. Communication may continue; a negotiation may be under way; practical arrangements may be retained. Conversely, a formally intact institution may fail to provide useful mediation in a particular crisis. A membership variable cannot measure all of those functions.

The distinction matters for the framework's notion of an exit. Leaving a political arrangement can reduce dependence, incur new costs, or create another institutional relation. An exit is not necessarily a rupture into isolation. It is a movement to another configuration, with its own constraints.

To represent that movement, retain an initial configuration x, a proposed configuration x', and the path needed to get from one to the other. A comparison of the endpoints can miss a costly transition. The future arrangement may be sustainable after customs systems, contracts, and payment channels adapt, while the interval between arrangements creates serious disruption.

The central calculation is then not simply "How much membership was lost?" It is "Which functions continue during the transition, which must be replaced, how long will replacement take, and who bears the interim cost?" That is the same geometry of feasible change encountered in supply dependence, now applied to institutions.

The Gambia and a negotiated departure

The Gambian political crisis of 2016–17 provides another kind of exit. In January 2017, the African Union publicly insisted that Yahya Jammeh hand over power to President-elect Adama Barrow (African Union 2017a). The January 21 joint declaration by ECOWAS, the African Union, and the United Nations described arrangements intended to facilitate a peaceful transfer, including Jammeh's departure and commitments concerning the treatment of those involved (ECOWAS, African Union and United Nations 2017).

The declaration is useful evidence of the terms publicly articulated in the resolution of that episode. It is not evidence that the book's model issued a contemporaneous warning. Nor does the presence of

an agreement allow us to identify, from this one case, how much each diplomatic or coercive action contributed to the outcome.

What the episode makes concrete is the difference between demanding a result and constructing a path toward it. A demand identifies the desired endpoint. An arrangement for departure addresses what must happen between the current position and that endpoint. These are different parts of a decision problem.

For a person relinquishing power, relevant considerations may include physical safety, legal exposure, political standing, the interests of associates, and the credibility of commitments. For the public, relevant considerations include constitutional authority, accountability, security, and the ability to sustain the transition. Those interests cannot be collapsed into the departing leader's comfort. A viable settlement has to be evaluated across the affected parties and the obligations that constrain it.

This is where an exit count becomes intellectually demanding. Writing $|S| = 4$ because four proposals have been named may conceal that all four depend on the same unavailable guarantee. Four descriptions are not necessarily four independent possibilities. If one necessary condition fails all four, the feasible set may still be empty.

Conversely, one carefully constructed arrangement can create a real path even when a broad list of rhetorical alternatives creates none. Counting matters only after feasibility has been defined. The framework's solution-space language should send the analyst toward the conditions that make an option usable, not toward producing an impressive integer.

The Gambia: an exit that was actually taken

The Gambian sequence allows us to follow an institutional demand through a completed departure. On January 18, 2017, the African Union insisted that Yahya Jammeh relinquish office to President-elect Adama Barrow. On January 19, the UN Secretary-General congratulated Barrow on his inauguration while expressing concern about Jammeh's refusal to step aside and people leaving for Senegal (African Union 2017a; United Nations Secretary-General 2017). Constitutional recognition and effective transfer had not yet become the same event.

The January 21 joint declaration by ECOWAS, the AU, and the UN addressed the transition and Jammeh's temporary departure. Its stated commitments concerned his security, dignity, and rights within the Gambian constitutional framework (ECOWAS, African Union and United Nations 2017). On January 22, the AU reported that Jammeh had stepped down the preceding day and left with his family, accompanied by Guinea's President Alpha Condé; Barrow was still in Senegal (African

Union 2017b). The departure was therefore an observed step in the transition, not just an option listed by an outside analyst.

The narrow outcome matters. Jammeh's departure does not establish that every guarantee was subsequently fulfilled, that constitutional consolidation was complete, or that later accountability questions had been settled. It establishes that one obstacle to the immediate transfer was removed through an arrangement the relevant actor actually used. A model of feasible change should recognize that achieved step without enlarging it into an account of all that followed.

The episode also shows why the departing officeholder's preferences cannot be the whole model. A course acceptable to one individual can remain unacceptable under constitutional authority, civilian protection, or the rights of others. The mathematical problem is not to maximize his comfort. It is to identify a course through several simultaneous requirements and the institutions able to sustain them. The negotiation changes the path to an outcome that is constrained by obligations beyond the negotiator's private payoff.

Read the dates as states of a process. Recognition of the successor establishes one condition. Agreement over departure establishes another. The departure supplies evidence that a particular transition step occurred. A binary variable called "international pressure" would compress those events into one treatment. A binary variable called "mediation succeeded" would collapse them into one result. Both can be useful summaries for a specified analysis, but neither explains the route from demand to compliance.

The appropriate causal restraint is equally concrete. The record contains diplomacy, recognition, and pressure, with a departure at the end. It does not show the counterfactual outcome in which one element was absent. The existence of a workable departure arrangement supports an account of how the transfer occurred. It does not identify a percentage contribution for each participant or prove that a verbal assurance alone caused the result. Chapter 37 explains how a broader comparison could investigate those effects.

Compare this completed step with the January 2025 ECOWAS withdrawals discussed at the start of the chapter. There the organization's statement preserved specified movement and trade arrangements while membership ended (ECOWAS 2025). That establishes a policy instruction; it does not, on its own, establish unchanged practice at every border. In The Gambia, the later AU statement supplies an additional observation of the departing actor's conduct. The two episodes have different evidence endpoints, and the book should stop each claim at its endpoint.

This gives a practical standard for an exit. A proposal has named conditions. An authorized arrangement has parties willing and able to implement it. An executed step has an observation showing that it occurred. A durable settlement has a longer record showing that the arrangement continued to function. These are successive evidential achievements. None can be manufactured by increasing an option count.

The geometric insight is visible in the passage between them. A political demand names an endpoint. Negotiation can change the route to that endpoint: sequence, guarantees, oversight, timing, and the distribution of costs. The January 2017 departure is a concrete instance in which a route was not merely imagined but taken. That makes the case valuable even without a contemporaneous Psychohistory forecast or a numerical claim about its causal effect. It shows what a usable exit looks like in a dated historical record.

Credibility changes a feasible set

Suppose an agreement promises a benefit after one party takes an irreversible step. In a simple two-stage representation, the first party acts at time t_0 and the promised response occurs at t_1. The first action's desirability depends partly on whether the later response can be relied upon. A written promise and a verifiable arrangement are therefore not equivalent state variables.

One can model this with a probability, but the probability needs an evidential basis. Assigning a value of 0.9 because the promise sounds credible adds apparent precision without knowledge. Alternatively, one can work with an interval of admissible credibility and ask whether the proposed path remains acceptable throughout that interval.

If an option is feasible only when an uncertain promise is honored with near certainty, it is fragile. If independent verification, staging, or another institution makes the option acceptable under a wider range of beliefs, the arrangement has enlarged the robust feasible set. The relevant improvement is not a generic increase in "trust." It is a specific reduction in what must be assumed before the next step can be taken.

This reasoning also explains why mediation cannot be reduced to a decorative extra edge. A mediator may transmit information, provide a guarantee, verify an action, or help sequence reciprocal moves. Those functions change different constraints. A useful implementation should record which function the mediator can perform in the episode being studied.

The model should also represent the possibility that mediation is rejected or lacks legitimacy among important participants. An institution's

presence on a diagram does not establish its ability to alter the feasible set. That ability is a hypothesis supported or challenged by the historical record.

The passport at the desk

Return to the trader at the border. The relevant document is not the abstract concept of membership. It is the document the official has been instructed to recognize for the transaction now being attempted. A regional decision can preserve that recognition while changing representation in a political institution. An analyst who sets both variables to zero has made an assumption about implementation that the decision itself does not support.

The January 2025 ECOWAS statement is particularly useful because it names continuing functions. Its wording supplies a dated rule for studying the transition. It does not establish that every official applied the rule identically, that every traveler experienced uninterrupted access, or that the arrangement continued indefinitely. Those are empirical questions. A serious comparison begins with the rule and then follows its operation.

Imagine a register with separate fields for political membership, document recognition, applicable trade treatment, and observed completion of a particular kind of crossing. The first three concern institutional status under a defined source. The fourth concerns practice. If the first changes while the next two are preserved, a model should be able to represent that configuration without declaring it inconsistent.

An illustrative state is

$$z = (m, d, t, a),$$

where m records membership, d the stated document rule, t the stated trade rule, and a observed accessibility under a defined procedure. The entries may be categorical rather than numerical. There is no need to average them simply because they occupy one vector. Their joint configuration is what matters.

For example, $(0, 1, 1, 1)$ can describe nonmembership with the other three functions operating under their definitions. The state $(1, 1, 1, 0)$ can describe continued membership and formal rights alongside an observed access failure. A membership-only model treats the first as a rupture and the second as continuity. A function-based account sees why those conclusions may be reversed for a traveler or shipment.

This does not reduce political membership to a ceremonial variable. Membership can confer representation, obligations, dispute processes, and future expectations beyond immediate passage. Its loss may change how later disagreements are resolved. The point is to follow those effects through the functions concerned instead of attributing every conceivable consequence to the withdrawal date.

A rule can expire before a replacement is ready

A transitional arrangement buys time. Its value depends on what must be accomplished during that time and how securely its availability is known. An analyst should distinguish a fixed expiry date, an arrangement continuing until further notice, and an informal expectation. Treating all three as an identical permanent edge conceals the planning problem.

Suppose, in a constructed administrative transition, a temporary procedure remains available for thirty days. A replacement procedure needs three sequential stages: preparing records, obtaining authorization, and testing the new process. The stages take eight, twelve, and six days. Their sum is twenty-six, leaving four days of margin. If the authorization stage instead takes eighteen days, the sequence takes thirty-two and misses the deadline by two.

The calculation is elementary because the point is not computational difficulty. It is the difference between a future arrangement that can work and a transition that reaches it in time. A plan described only by its final state can be correct about the destination and wrong about the path.

If two stages can genuinely proceed in parallel, the schedule changes. Preparing records and obtaining a separate equipment certification might overlap, but authorization that depends on completed records cannot begin earlier merely because a spreadsheet permits it. The dependency graph needs to represent the real prerequisites. Optimistic scheduling is often an unnoticed relaxation of those constraints.

Let task j require duration d_j and let $i \prec j$ denote a prerequisite. With start time s_j, a valid schedule satisfies

$$s_j \geq s_i + d_i \quad \text{whenever } i \prec j.$$

The completion time is the largest $s_j + d_j$ over required tasks. A deadline imposes an additional inequality. The same logic applies to administrative transitions, supplier qualification, and the staged implementation of an agreement. The tasks differ; the meaning of a prerequisite does not.

Uncertain duration requires a range or a distribution supported by relevant experience. If the total duration lies between twenty-six and

thirty-two days, a thirty-day bridge does not cover every allowed realization. The model can ask whether the bridge can be extended, whether a required task can start sooner, or whether a temporary alternative can cover the gap. Announcing the final arrangement more confidently does not create those two missing days.

The people who depend on continuity may also have different margins. A large firm can hold additional stock; a small trader may depend on one completed journey to finance the next. The same administrative delay can therefore change their feasible choices differently. A transition analysis should retain those differences instead of reporting only the average cost per shipment.

The distribution inside the average

Suppose two kinds of transaction use a crossing. One hundred routine transactions take one day each; ten transactions requiring a particular permission take eleven days each. The overall mean delay is

$$\bar{d} = \frac{100 \cdot 1 + 10 \cdot 11}{110} = \frac{21}{11} \simeq 1.91 \text{ days}.$$

The average looks modest. It says little about a person in the second group. If that group's deadline is five days, every one of those transactions misses it under the constructed example.

Now imagine that the ten difficult transactions stop being attempted. The recorded mean falls to one day. No individual process has become faster. The excluded group has disappeared from the denominator. An analyst who reports improved access from the lower mean alone has mistaken selection for improvement.

The appropriate register retains attempted, completed, abandoned, and ineligible transactions under the study's definition. It records waiting time for those still pending rather than deleting them at the reporting cutoff. A completed-only average systematically omits the longest unresolved waits until they finish, if they ever do.

This is a form of censoring. At the date the report is prepared, a pending transaction has a known minimum wait but an unknown eventual duration. Treating its current wait as its final wait understates duration; omitting it can also distort the sample. The statistical treatment depends on the question, but the source record must preserve the fact that the observation is unfinished.

These measurement choices belong in a chapter about political geometry because they determine which paths appear to remain open. A route that works only for people with large reserves or exceptional

documentation is not equally available to everyone represented by the national node. A single average can make a narrowing set of eligible users look like an efficient connection.

An informative institutional comparison would therefore report the distribution of access across relevant groups and the reasons for non-completion. This is not an invitation to expose personal information. Aggregated categories can preserve the mechanism while protecting identities. The analytical requirement is that exclusion does not become invisible simply because no successful transaction records it.

Making a guarantee inspectable

The Gambian transition illustrates the importance of an executable path, but it cannot supply a universal numerical value for the credibility of a guarantee. A useful model can instead show exactly what an option requires and where uncertainty enters.

Take a constructed decision in which successful implementation has value eight to a participant, failed implementation has value minus twelve, and taking the step costs one. If the probability of implementation is p, expected value is

$$V(p) = 8p - 12(1 - p) - 1 = 20p - 13.$$

Relative to a stated outside option of zero, this participant accepts under this simplified expected-value rule only when $p \geq 0.65$. The numbers are illustrative preference units, not an estimate of any Gambian actor's valuation. The calculation exposes the assumption that a promise must be believed with sufficient confidence for the step to be acceptable.

If the admissible range is $p \in [0.5, 0.8]$, the option is favorable in some supported scenarios and unfavorable in others. Its worst-case value is -3. Describing the midpoint as the answer would conceal the fragile lower end. A robust decision under this particular rule requires a nonnegative value throughout the admitted interval.

Now suppose a redesigned, reversible first step limits the failure consequence to minus four while keeping the illustrative success value and cost. Its value becomes $12p - 5$. The threshold falls to $5/12 \simeq 0.417$, and the worst-case value over the same interval becomes one. The redesign has not required an increase in asserted trust. It has changed the exposure created by acting before uncertainty is resolved.

This is a mathematical explanation of why staging can matter. It is not a recommendation that every political obligation be traded against private convenience. Public duties, legal constraints, accountability,

and the rights of affected people remain constraints on the admissible arrangement. Some options should be excluded before any preference calculation is considered. A model that maximizes a leader's willingness while ignoring those constraints has changed the problem.

The model also leaves room for participants whose choices are not adequately described by expected value. They may care about worst cases, irreversible losses, legitimacy, or obligations that cannot be reduced to one scale. The useful question remains: which specific feature makes the next step unacceptable, and can the arrangement change that feature without violating another necessary condition?

An institutional guarantee can work through several different mechanisms. It may make implementation more likely, improve observation of implementation, reduce the loss if implementation fails, or provide a remedy afterward. Those are four different changes in the example. A generic increase in the word "credibility" conceals which one is being claimed.

A negotiation is not four independent options

Suppose a proposal register lists four possible arrangements. Two require the same external authorization; a third requires an unavailable payment mechanism; the fourth requires both. Counting four exits exaggerates independence. If the authorization and payment mechanism are both absent, the list contains no presently executable arrangement.

Represent the prerequisites by a Boolean vector z. Each option i has a feasibility rule $f_i(z) \in \{0, 1\}$. The number of usable options is

$$N_{\text{usable}}(z) = \sum_i f_i(z).$$

This count has meaning only after the rules have been stated. It differs from the number of named proposals, and it may change discontinuously when one shared prerequisite becomes available.

For example, let a denote authorization and b a payment mechanism, and let the four rules be a, a, b, ab. With $(a, b) = (0, 0)$, the count is zero. With $(1, 0)$ it is two; with $(0, 1)$ it is one; with $(1, 1)$ it is four. The first two proposals are distinct in other respects but share their immediate feasibility condition. They should not be treated as two independent safeguards against the loss of authorization.

Counting still does not rank the options. An executable arrangement can be unjust, unacceptable to an essential participant, or incapable of surviving the transition. The feasible set must include the relevant

constraints, and the decision process must evaluate the consequences. The count is a description of one specified set, not a moral score.

A useful negotiation analysis keeps the prerequisite graph alongside the proposal list. It can then distinguish creating a new path from writing another version of an old path with the same missing condition. It can identify a common enabling operation whose completion makes several otherwise different proposals possible. That is a concrete way in which mediation can enlarge the solution space.

Continuity is something to observe

The West African examples resist two symmetrical simplifications. Political departure does not necessarily sever every practical connection. Formal continuity does not necessarily preserve a connection's useful operation. Both errors arise when the analyst asks an institutional label to stand in for the functions people depend on.

A comparative record should retain the formal rule, its implementation, the available substitutes, and the distribution of transition costs. It should distinguish an institution's capacity to convene a meeting from its capacity to make an agreement executable. It should record what happens after the celebrated moment: whether documents continue to be recognized, obligations can be met, and disputes can still be addressed.

The framework's claim becomes more specific through this work. Peaceful change depends in part on whether people can move from one political arrangement to another without losing the operations that make ordinary life possible. A route, a document, a payment, and a credible process for repairing a breach each contribute something different. The mathematics helps identify the missing condition. The historical inquiry establishes whether it was actually present.

Regional institutions as testable mechanisms

The institutional withdrawal and the Gambian transition ask different empirical questions. In the first, did practical arrangements preserve movement or trade after a formal change in membership? In the second, which terms and actions made a transfer possible, and what evidence supports that interpretation? Combining both into a single regional "success rate" would discard the distinction that makes them informative.

A comparative study should instead define mechanisms and look for variation. For membership changes, it could track dated rules, actual administrative practice, observed delays, and the availability of alternatives. For mediation, it could distinguish convening, guarantees,

verification, and implementation. Cases in which formal commitments fail to produce functioning access are as useful as cases in which they do.

The same care applies to comparison across countries. A common institutional framework does not make the member states identical. Economic structure, geography, political authority, and the distribution of costs affect which constraints bind. A model that transfers a coefficient across cases should explain what is being held comparable and test the transfer, rather than relying on a continental label.

The lesson for the larger book is constructive. A political relationship can be redesigned. Some functions can be preserved while others are renegotiated. A departure can be staged rather than treated as an instantaneous severing of every connection. The geometric question is whether the redesigned arrangement gives people workable paths through the transition.

Peace, in this account, is not simply high connectivity. It is the capacity to change a relationship without making destructive action the only path someone can perceive or sustain. West Africa supplies both the institutional detail and the practical questions needed to make that claim answerable.

Part IV

The mathematics of constraint

The pipelines, borders, and negotiating rooms remain in view as we enter the mathematics. Which connections resist a split? How can a busy system lose the ability to rearrange? We examine what survives a change of units, why the order of events matters, and how much the observations leave undecided. Then we follow uncertain connections, temporary surges, and finite reserves into calculations that a reader can reproduce.

Chapter 22

The Mathematics of Helicity

> The equation earns its place when we can say what is moving.
>
> Bee Rosa Davis

Watch water leaving a pipe in a rotating jet. It travels along the pipe's axis while turning around that axis. Reverse the rotation while keeping the forward motion and the flow changes its handedness. Speed alone does not record that difference. Helicity puts the direction of motion and the direction of rotation into the same calculation.

We have been using a related question throughout the book: how does continued economic movement sit within a changing political relationship? To understand the proposed connection, we need to follow the mathematics from its physical definition into the quantities used for a dyad. The steps matter. They show what the construction takes from fluid mechanics, what it adds, and where the observations enter.

The concept of helicity originates in fluid dynamics, where it measures a fundamental topological property of flows: the degree to which vortex lines are linked, twisted, or knotted. Unlike velocity or vorticity alone, helicity captures the *handedness* of a flow—whether it spirals left or right, and how that spiral relates to the direction of motion. The discovery that this quantity is conserved in ideal fluids ranks among the most important insights in twentieth-century physics, with implications ranging from solar physics to plasma confinement to—as we shall argue—international relations.

Helmholtz's 1858 vortex theorems describe how an ideal flow carries its vortex lines. Kelvin's circulation theorem follows a closed loop of fluid particles. Under conservative body forces and barotropic pressure, the circulation around that moving loop stays constant while the loop stretches or bends.

But the full topological significance of these results remained unclear until the work of Jean-Jacques Moreau (1961) and H.K. Moffatt (1969), who independently recognized that the integral

$$\mathcal{H} = \int_V \vec{v} \cdot (\nabla \times \vec{v})\, dV \tag{22.1}$$

is conserved for smooth ideal flow under appropriate boundary conditions. For example, a periodic volume has no net boundary contribution to the helicity balance.

Moffatt's result gives a precise relationship between the helicity integral and the linking of vortex lines under the relevant assumptions (Moffatt 1969). Its force comes from that derivation. The application to another kind of flow must identify which parts of the construction survive.

Moffatt's 1969 paper connected the signed integral to the linkage of vortex lines. The word *helicity* already had a use in particle physics. In the fluid problem a pair of linked tubes gives the term a visible meaning. While the smooth ideal flow carries both tubes, unlinking them would require a change in how the lines connect (Moffatt 1969; Moffatt 2021).

Arnol'd's topological approach to hydrodynamics develops this relationship between geometry and ideal fluid motion. The linkage interpretation depends on the fields, domain, and boundary conditions. Those conditions give the invariant its mathematical meaning; the word helicity alone does not transfer it to a different physical or political system.

In classical fluid mechanics, helicity is defined as:

$$\mathcal{H} = \int_V \vec{v} \cdot (\nabla \times \vec{v})\, dV \tag{22.2}$$

where $\vec{v}$ is the velocity field and $\nabla \times \vec{v} = \vec{\omega}$ is the vorticity, twice the local angular velocity of an infinitesimal fluid element. Each term deserves careful examination:

The velocity field $\vec{v}$: At each point in the fluid, $\vec{v}$ describes the instantaneous velocity of a fluid particle. In three dimensions, $\vec{v} = (v_x, v_y, v_z)$.

The vorticity $\vec{\omega} = \nabla \times \vec{v}$: The curl of velocity measures local rotation. A nonzero vorticity indicates that neighboring fluid particles have different velocities—they are spinning relative to each other. Mathematically:

$$\vec{\omega} = \begin{pmatrix} \partial_y v_z - \partial_z v_y \\ \partial_z v_x - \partial_x v_z \\ \partial_x v_y - \partial_y v_x \end{pmatrix} \tag{22.3}$$

The dot product $\vec{v} \cdot \vec{\omega}$: This scalar quantity measures the alignment between velocity and vorticity at each point. When they are parallel (the fluid moves in the direction it rotates), the dot product is positive. When antiparallel, it is negative. When orthogonal, it is zero.

The volume integral: Summing over the entire domain gives the total helicity, which may be positive, negative, or zero depending on the net handedness of the flow.

Consider the core of the rotating jet again. Use z along the pipe, let U be the forward speed, and idealize the rotation as solid-body spin at angular speed Ω. The velocity at (x, y, z) is $(-\Omega y, \Omega x, U)$. The first two components carry a marked drop around the axis; the last carries it toward the outlet. In the curl formula, $\partial_x v_y = \Omega$ and $\partial_y v_x = -\Omega$, so the axial vorticity is 2Ω. The local helicity density is therefore $2\Omega U$. At $U = 2$ metres per second and $\Omega = 1$ radian per second, it is $+4\,\mathrm{m\,s^{-2}}$. Reverse the spin while leaving the pump alone and it becomes -4 in the same units. Stop the forward motion and it is zero even though the water still spins. That last case explains why counting turns alone cannot measure helicity.

This matters physically because *helicity is conserved in ideal fluids*. The proof follows from the Euler equations governing inviscid flow:

$$\frac{\partial \vec{v}}{\partial t} + (\vec{v} \cdot \nabla)\vec{v} = -\frac{1}{\rho}\nabla p \tag{22.4}$$

Taking the curl of both sides and using vector identities, one can show that the vorticity evolves according to:

$$\frac{\partial \vec{\omega}}{\partial t} = \nabla \times (\vec{v} \times \vec{\omega}) \tag{22.5}$$

From this, the time derivative of helicity can be computed:

$$\frac{d\mathcal{H}}{dt} = \int_V \left[\vec{\omega} \cdot \frac{\partial \vec{v}}{\partial t} + \vec{v} \cdot \frac{\partial \vec{\omega}}{\partial t} \right] dV = 0 \tag{22.6}$$

Here the fluid is smooth, incompressible, and inviscid, with constant density and no applied forcing. Take a periodic volume, for example: the transport terms integrate to boundary fluxes that cancel on opposite faces. Under these conditions the total is conserved. For a fixed section of an open pipe, fluid and helicity can pass through the ends; the same conclusion about the total inside that section does not follow. The pipe illustrates the local measurement, while conservation concerns the evolution and boundary of a specified volume (Moffatt 1969).

Conservation constrains what an ideal evolution can do. For instance, linked, closed vortex tubes carried smoothly by the fluid cannot simply pass through one another and unlink. A conserved total does not hold each local contribution fixed or require a particular visible eddy to survive a disturbance. Those are stronger statements than conservation of the integral.

The importance of helicity conservation can be illustrated through several physical systems:

Tornadoes: A rotating updraft makes the two ingredients visible: air rises and turns. Its upward velocity projects onto the vertical part of its vorticity; reversing the sense of rotation reverses that contribution. Explaining the persistence of the entire storm requires its dynamics, not just a local reading of handedness.

Solar magnetic flux tubes: On the sun, magnetic field lines become twisted and braided. Magnetic helicity is an analogous measure of field-line linkage. In an ideal perfectly conducting model with suitable boundaries it constrains how the field can evolve; reconnection changes what an individual field line can connect to.

DNA structure: An intact, covalently closed double helix can redistribute twist into coiling of its axis while preserving the linking number of its strands. A topoisomerase can temporarily break and reseal DNA, changing the linking number. The allowable operation matters: bending a closed loop and passing a strand through a break have different topological consequences (Strick, Croquette, and Bensimon 2000).

Look first at the arrows in Figure 22.1. The two turning patterns have opposite handedness and the same absolute local helicity. Then look at the curves beside them: those curves illustrate possible responses to disturbance. A curve of coherence against time needs dynamics and an initial disturbance; the sign of the dot product alone supplies neither. Equal amounts of left- and right-handed flow can also cancel in the volume integral without either local motion disappearing.

In the political application the sign has a declared interpretation. When trade must continue while the parties withdraw the trust and access that support it, a disruption can be harder to absorb. The proposed mechanism is the dependence that remains and the alternatives that have closed. The fluid example shows how to measure orientation; the political record tells us why a particular orientation matters.

In real (viscous) fluids, helicity is not perfectly conserved. The viscous term in the Navier-Stokes equations introduces helicity dissipation:

$$\frac{d\mathcal{H}}{dt} = -2\nu \int_V \vec{\omega} \cdot (\nabla \times \vec{\omega})\, dV \tag{22.7}$$

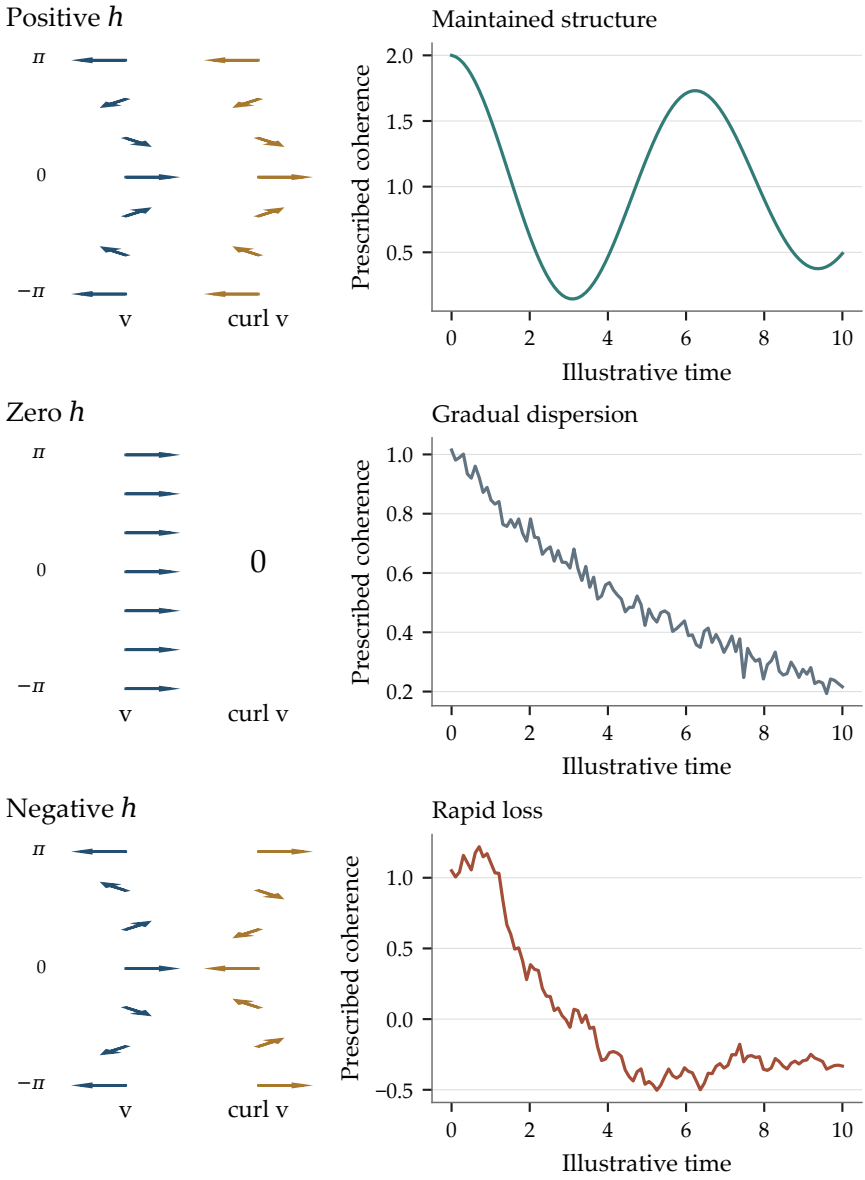


FIGURE 22.1: Local helicity and illustrative responses. At left, $v = (\cos z, -s \sin z, 0)$ gives $\nabla \times v = sv$ and $h = s$ for $s = \pm 1$; the zero case is $v = (1, 0, 0)$. Arrows show horizontal components at successive heights. The prescribed coherence curves at right illustrate an interpretation, not solutions of these fields: helicity sign alone does not determine stability.

where ν is the kinematic viscosity. This dissipation occurs preferentially at small scales, where velocity gradients are steepest.

Energy and helicity track different aspects of the same moving fluid. Energy adds squared speeds, so equal and opposite velocities do not cancel it. Helicity adds signed projections; opposite handedness can cancel even when the speeds are large. Moreover, the viscous integral above is signed, so it supplies no universal ordering of the two decay times. To see a difference in persistence, one must follow the particular flow.

That distinction is useful in a trade ledger. Every delivery adds to gross trade, whatever political pressure it puts on the buyer. A large total can survive a reversal in what those deliveries mean to the relationship. The divergence trap concerns precisely that combination of continuing activity and adverse alignment; the underlying records tell us whether it is present.

Helicity on Manifolds

The generalization from fluid flows to international relations requires abstracting the concept of helicity to a mathematical manifold. This is not merely notational elegance—it allows the framework to handle relationships that don't fit neatly into three-dimensional space. Such an extension requires a specified geometric structure, measure, boundary conditions, and evolution law. Conservation and stability results travel only when their hypotheses hold for the new construction. A political flow variable does not acquire a fluid conservation law simply by being written on a manifold.

We construct an abstract manifold $\mathcal{M}$ representing the space of possible bilateral relationships. Each point on this manifold corresponds to a specific configuration of economic, diplomatic, and security ties between two states. The *dimension* of this manifold equals the number of independent quantities required to specify a relationship completely.

For our framework, we work with a manifold of dimension $n = 5$, with coordinates:

T — Bilateral trade volume (normalized by global trade); E — Energy interdependence (normalized by critical threshold); S — Diplomatic similarity (UN voting correlation); R — Treaty network density (alliance vs. non-aligned); X — Sanctions intensity (normalized by maximum historical).

A point $p \in \mathcal{M}$ thus has coordinates $p = (T, E, S, R, X)$, and the evolution of a bilateral relationship traces a *trajectory* $\gamma(t)$ through this space.

On this manifold we represent economic flows by a 1-form ω—a rule that assigns a number to a proposed movement at each point. Here ω denotes the form; the earlier bold vector $\vec{\omega}$ denoted physical vorticity. In coordinates:

$$\omega = \omega_T\, dT + \omega_E\, dE + \omega_S\, dS + \omega_R\, dR + \omega_X\, dX \tag{22.8}$$

The coefficients ω_i depend on position in $\mathcal{M}$ and encode how economic "velocity" projects onto each coordinate direction.

A customs desk gives a one-form something to do. Imagine a simple clearance rule: moving dT units of normalized trade takes $(1 + X)dT$ units of processing work when the restrictions index is X. Changing the index without moving goods adds no processing work under this rule. This is a toy one-form, $\omega = (1 + X)dT$. At $X = 0$, processing 0.2 units of trade costs 0.2 units of work. At $X = 0.5$, the same shipment costs 0.3. A one-form tells us how to score a movement made under the conditions at its starting point. It does not assign a bill to a country merely for occupying a point on the map.

The exterior derivative $d\omega$ is a 2-form that generalizes the curl from vector calculus. In three dimensions, the curl of a vector field is another vector field; in higher dimensions, this generalizes to a 2-form (an antisymmetric tensor). Explicitly:

$$d\omega = \sum_{i<j} \left(\frac{\partial \omega_j}{\partial x_i} - \frac{\partial \omega_i}{\partial x_j} \right) dx_i \wedge dx_j \tag{22.9}$$

where $\wedge$ denotes the wedge product, which satisfies $dx_i \wedge dx_j = -dx_j \wedge dx_i$.

The exterior derivative captures how the flow "twists" as we move through relationship space. A nonzero $d\omega$ indicates that the economic flow has rotational structure—different components of the flow are changing at different rates in different directions.

At the customs desk, raise the restriction from zero to 0.5 and process the shipment of size 0.2. Now compare the reverse order: process the same shipment first, then raise the restriction. Both routes finish with precisely the same trade and restriction coordinates. Their processing totals differ by 0.1. Here $d\omega = dX \wedge dT$; integrating it over that rectangle gives $0.5 \times 0.2 = 0.1$ with the restriction-first orientation. Following one route out and the other back fixes an orientation around the rectangle. Reversing that loop reverses the sign. This is a calculation of path-dependent work in the toy rule, rather than an estimate of actual clearance costs.

The helicity on the manifold is then:

$$\mathcal{H}(t) = \int_{\Sigma_t} \iota_t^*(\omega \wedge d\omega) \tag{22.10}$$

This integral has a precise topological interpretation. The integrand $\omega \wedge d\omega$ is a 3-form, which can be integrated over any 3-dimensional submanifold of $\mathcal{M}$. For the five-dimensional ambient manifold, choose an oriented three-dimensional submanifold Σ_t with inclusion $\iota_t : \Sigma_t \hookrightarrow \mathcal{M}$. The pullback restricts the form to that domain. A family of such submanifolds gives a time-indexed scalar; the choice of domain and its boundary conditions is part of the model.

For example, hold energy dependence and treaty density fixed and allow trade, sanctions, and diplomatic similarity to vary. This defines one possible three-coordinate slice. The pullback means evaluating the form only on movements within that slice. It is not integration along a single country's time series: a curve has one direction at each regular point; a three-dimensional region has room for three independent directions. The customs example shows why the distinction matters. There $d\omega$ is nonzero but $\omega \wedge d\omega$ vanishes, because dT appears twice. An order effect in a two-coordinate record is evidence of that effect, not by itself a measurement of volume helicity.

Think of a movement carrying both the marked volume and the field painted on it. If ϕ carries the initial oriented domain to the later one and $\phi^*\omega_t = \omega_0$, pullback commutes with the exterior derivative, so the two integrals agree by change of variables. Fluid helicity has its own dynamical conservation law under the ideal conditions above (Moffatt 1969). Holding the number of shipments fixed supplies only a total. To apply the transport argument to a relationship one must identify which field the shipments and commitments carry.

Topology asks which properties survive the allowed deformations. For a pipe network, stretching a route and cutting a route are different operations. For a relationship, the corresponding question is which commitments and channels the model permits to change. War is one way a network can be disrupted. A new supplier, a treaty, or the retirement of a pipeline can change it through other means.

The preservation result above has precise hypotheses. It follows because the evolving form is transported with the domain in the stated way. Keeping total trade unchanged does not supply that hypothesis. Two trade ledgers with identical totals can distribute dependence very differently.

Geometry can also affect a path-dependent observable without proving stability. In the customs example, the order of inspection and delivery changes the accumulated processing cost. To decide whether a distur-

bance grows, we need the evolution law and its response to perturbations. An invariant and a stability test answer different questions.

That is the obligation of the cross-domain argument. Identify the mathematical object, specify the transformations it survives, and test whether the measured political process respects those assumptions. When a commitment or channel changes, calculate again. The theory becomes useful through those concrete distinctions.

For practical computation, one signed scalar compares economic flow with political circulation in the chosen oriented three-dimensional coordinates:

$$H = \frac{\tau}{K} = \frac{\langle F, \nabla \times A \rangle}{\|F\| \cdot \|\nabla \times A\|} \tag{22.11}$$

Here F represents economic flows: trade, investment, and financial movements. The alignment field A represents diplomatic engagement, treaty networks, and UN voting similarity. Its curl, $\nabla \times A$, describes local circulation in the chosen relationship coordinates. The inner product then measures how strongly economic flow projects onto that circulation. To interpret the resulting scalar politically, the construction must specify the coordinates and the inner product that make those quantities comparable.

On the right-hand side the numerator is a dot product and the denominator is the product of two lengths. When both lengths are positive, the quotient is the cosine of their angle. Twice as much activity in the same direction leaves that angle unchanged. A right angle gives zero, a 60-degree angle gives +0.5, and a 120-degree angle gives −0.5. This measurement asks how the two directions sit relative to each other. It cannot distinguish a pair of small parallel vectors from a pair of large parallel vectors, so the flow magnitudes belong alongside the angle. Identifying τ/K with this expression requires the corresponding numerator and normalization; an independently scored tolerance budget divided by integration burden measures something else.

The ratio $H = \tau/K$ can be understood intuitively: *helicity equals torsion divided by curvature*. These terms deserve careful definition:

Torsion τ: In differential geometry, torsion measures how a curve "twists out of its osculating plane." For our application, torsion measures the degree to which economic flows and diplomatic alignment are misaligned—they are changing in incompatible directions.

Curvature K: The framework uses an integration-burden index for the cost of changing an established relationship. Geometric curvature of a trajectory is a separate, precise concept: it measures how fast the direction turns per unit distance along the path. A shipment series can

move rapidly along a straight line with zero geometric curvature. The section on feasible adjustment below distinguishes a turning path, a fast change, and an expensive dependence.

A highly interconnected relationship (large K) can tolerate more misalignment (τ) while remaining stable, because the dense connections provide multiple channels for adjustment. A sparsely connected relationship has no such buffer.

The connection also leads to Chern–Simons theory. The first term in the formula below is the form we have just met. A gauge connection tells us how to compare local descriptions as we move; for matrix-valued comparisons, their order can matter. The extra cubic term belongs to that setting. Here $\mathcal{M}$ in the integral denotes an oriented three-dimensional domain, such as a chosen slice of the five-coordinate model, and A denotes a connection on it. With a specified trace and normalization, the Chern–Simons expression is:

$$\mathcal{H} = \frac{k}{4\pi} \int_{\mathcal{M}} \mathrm{Tr}\left(A \wedge dA + \frac{2}{3} A \wedge A \wedge A \right) \tag{22.12}$$

This formulation, while more abstract, connects our framework to a rich mathematical literature on topological invariants and their physical significance.

For an ordinary scalar-valued one-form the cubic term vanishes and the familiar $A \wedge dA$ remains. In the usual compact, simple, simply connected gauge-group quantum theory on a closed oriented three-manifold, with standard normalization, the level k is quantized so that the exponentiated action is well defined under large gauge transformations (Witten 1989). This does not require the classical integral for an arbitrary connection to take only discrete values. Whether a political relationship jumps between regimes is a question about the relationship and its evolution. The formal correspondence helps specify the mathematics a proposed explanation would need.

The Spectral Gap

The spectral gap is the second pillar of our mathematical framework. While helicity measures the topological coherence of flows, the spectral gap measures the *connectivity* of the underlying network. Together, they describe orientation and a specified form of connectivity. Stability also depends on the dynamics acting on that structure: forcing, feedback, time variation, and transient amplification can matter even when a static

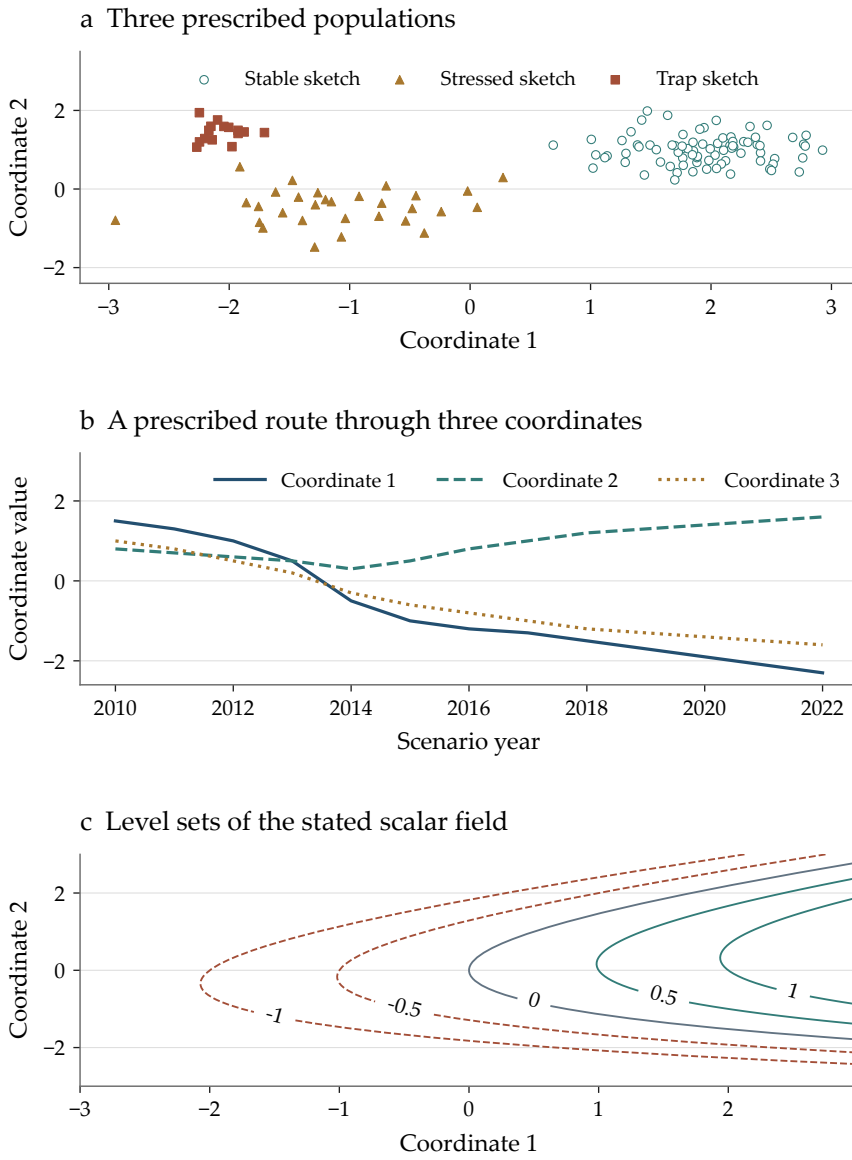


FIGURE 22.2: An illustrative relationship space. (a) Seeded populations in two displayed coordinates. (b) The prescribed trajectory's three coordinates shown separately. (c) Level sets of $H(x, y) = 0.5x - 0.3y^2 + 0.1xy$. The coordinate labels describe the construction; they are not principal components estimated from a measured country dataset.

spectrum appears favorable. Chapter 30 develops that distinction with explicit calculations.

Given a weighted graph $G = (V, E, w)$ representing the trade network, we construct a sequence of matrices that encode its structure:

Adjacency matrix: $A_{ij} = w(i, j)$ represents bilateral trade between countries i and j. This is a symmetric matrix (we treat trade as undirected, summing exports and imports).

Degree matrix: $D_{ii} = \sum_j A_{ij}$ is diagonal, with each entry representing the total trade of country i. This measures how "connected" each node is to the rest of the network.

Graph Laplacian: $L = D - A$ is the fundamental object. Its entries are:

$$L_{ij} = \begin{cases} \sum_k w(i,k) & \text{if } i = j \\ -w(i,j) & \text{if } i \neq j \end{cases} \tag{22.13}$$

Normalized Laplacian: $\mathcal{L} = D^{-1/2} L D^{-1/2}$ adjusts for the varying sizes of economies, giving a matrix whose eigenvalues lie in $[0, 2]$.

The Laplacian has a deep physical interpretation. It is the discrete analog of the continuous Laplacian operator ∇^2 from calculus. Just as $\nabla^2 f$ measures how f at a point differs from its average over nearby points, Lx measures how the value at each node differs from the weighted average of its neighbors.

The Laplacian L is positive semi-definite, meaning all eigenvalues are non-negative:

$$0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{n-1} \tag{22.14}$$

The smallest eigenvalue $\lambda_0 = 0$ always exists, with eigenvector $\vec{1} = (1, 1, \ldots, 1)^T$ (the constant vector). This reflects the fact that the Laplacian annihilates constant functions.

The *spectral gap* is the second-smallest eigenvalue:

$$\mu_1 = \min_{\vec{x} \neq 0,\, \vec{x}^T D \vec{1} = 0} \frac{\vec{x}^T L \vec{x}}{\vec{x}^T D \vec{x}} \tag{22.15}$$

Here μ_1 is the second-smallest eigenvalue of the normalized Laplacian $\mathcal{L}$; the combinatorial gap uses $\vec{x}^T \vec{x}$ in the denominator and $\vec{x}^T \vec{1} = 0$. This variational characterization shows that μ_1 measures the minimum "resistance" the network offers to any non-trivial perturbation. A large spectral gap means the network strongly resists fragmentation; a small spectral gap means it is close to splitting into disconnected components.

The spectral gap governs the rate of diffusion on the network. If we place a "heat" distribution on the nodes and let it spread according to the Laplacian dynamics:

$$\frac{d\vec{x}}{dt} = -L\vec{x} \tag{22.16}$$

then the solution is:

$$\vec{x}(t) = \sum_k c_k e^{-\lambda_k t}\vec{v}_k \tag{22.17}$$

where $\vec{v}_k$ are the eigenvectors. The term corresponding to $\lambda_0 = 0$ does not decay—it represents the equilibrium distribution. All other terms decay exponentially, with the slowest decay governed by $e^{-\lambda_1 t}$.

The reciprocal of the positive gap is a relaxation time for the slow mode. To give a mixing time one must also specify how close to equilibrium is close enough. A port that needs a backlog reduced by half has a different deadline from one that needs it reduced by ninety percent. Chapter 23 works this out with a measured exchange rate.

For international trade networks, this has immediate economic significance. A high spectral gap means that economic shocks propagate quickly through the entire system and are absorbed collectively. A low spectral gap means shocks remain localized, creating persistent disparities that can destabilize the system.

The spectral gap has a beautiful connection to the *graph isoperimetric problem*: finding the "smallest bottleneck" that divides the network. The Cheeger constant $h(G)$ measures this bottleneck:

$$h(G) = \min_{S \subset V} \frac{|E(S, \bar{S})|}{\min(\text{vol}(S), \text{vol}(\bar{S}))} \tag{22.18}$$

The minimum runs over nonempty proper groups with positive volume on each side. Here $|E(S, \bar{S})|$ denotes the sum of the crossing weights, and $\text{vol}(S) = \sum_{i \in S} D_{ii}$. A shipment contributes its weight once to the crossing total. The smaller group's weighted degree sets the scale.

For an undirected graph with nonnegative weights and positive degrees, the Cheeger inequality relates this geometric quantity to the normalized spectral gap:

$$\frac{h(G)^2}{2} \leq \mu_1 \leq 2h(G) \tag{22.19}$$

This is remarkable: a purely algebraic quantity (eigenvalue of a matrix) is bounded above and below by a purely geometric quantity (minimum

bottleneck). The spectral gap detects network fragility without requiring explicit search over all possible divisions.

For the global trade network, the spectral gap has declined significantly since 2018:

Period	λ_1	**Interpretation**
1995-2007	0.12–0.15	Globalization: dense, well-connected
2008-2009	0.08	Financial crisis: temporary fragility
2010-2017	0.11–0.13	Recovery: still well-connected
2018-2019	0.09	Trade war: emerging bottlenecks
2020-2021	0.07	Pandemic: supply chain disruption
2022-2024	0.05–0.06	Sanctions: approaching fragmentation

The threshold $\lambda_1^* = 0.05$ is the framework's calibrated operating point for the specified network construction. Cheeger's inequality explains the connection between a normalized spectral gap and a bottleneck; it does not select that numerical threshold. The calibration and the operator must accompany its use in an assessment.

The Trichotomy Parameter

The third pillar is the relation between constraints, tolerance, and the size of the space being described. Its political use requires distinguishing uncertainty about a settlement from the ability to carry one out. A dated capacity reading is also distinct from the slope or acceleration of a time series. Measuring change requires repeated observations; the ratio itself contains no time derivative.

The historical geopolitical expression is

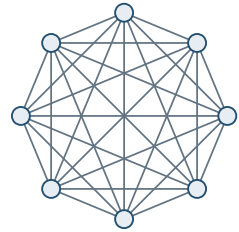
$$\Gamma = \frac{m \cdot \tau}{K \cdot \log |S_{\text{exit}}|} \tag{22.20}$$

with the implementation convention reproduced in Appendix B. The following distinction determines what this number can legitimately mean.

Knowing the options and having an option

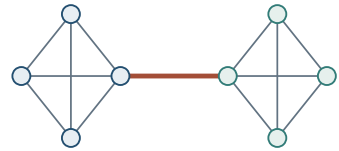
A negotiator has four proposals on the desk. An analyst has enough evidence to determine which one each side would accept. Those are two different kinds of progress: making a settlement possible, and

a Dense connections



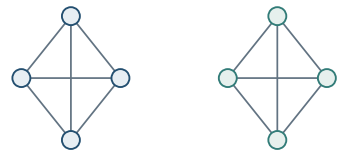
Combinatorial gap: 8.0000

b One connecting edge



Combinatorial gap: 0.3542

c Two disconnected components



Combinatorial gap: 0.0000

FIGURE 22.3: The combinatorial Laplacian $L = D - A$ for the three unit-weight graphs. The displayed second eigenvalues are calculated from these exact connections. Compare Figure 7.1 for the normalized convention: a different operator gives different values even on the same graph. Disconnection produces a second zero eigenvalue.

discovering what is possible. A formula for the second must not be interpreted as if it automatically measured the first.

The framework's inference construction compares the available constraints with the complexity of a specified candidate space. Written with that meaning explicit, its ideal ratio is

$$\Gamma_{\text{res}} = \frac{m\tau}{K \log |S_{\text{cand}}|}, \qquad K > 0, \quad |S_{\text{cand}}| \geq 2.$$

Here S_{cand} contains distinguishable candidate descriptions under a fixed resolution and equivalence rule. More candidates require the same evidence to distinguish more possibilities. With $m = 2$, $\tau = 0.4$, and $K = 0.8$, two candidates give $\Gamma_{\text{res}} = 1.4427$; four give 0.7213. Doubling the burden lowers the ratio. No political deterioration is implied by that arithmetic. The general solution-space reading is stated in the Psychohistory specification (Davis 2026h).

The logarithm is a complexity convention for the declared finite candidate set. It is not automatically the entropy of a measured political distribution. Counting the same proposal under four titles should not quadruple the space. Conversely, one proposal with several consequential implementation states may require more than one candidate description. The resolution and equivalence rule therefore belong in the measurement protocol.

A *usable exit* has a different definition. It is a course of action that satisfies the parties' relevant commitments, physical capacities, timing requirements, and other stated constraints. Let F be the set of such courses, and let $u(x)$ measure the declared objective of a course x. If an additional option leaves all existing options and their conditions intact, then

$$F \subseteq F' \quad \Longrightarrow \quad \max_{x \in F'} u(x) \geq \max_{x \in F} u(x)$$

whenever the maxima exist and F is nonempty. Enlarging the feasible set cannot worsen the best attainable value under those assumptions. A plan that consumes the guarantee or transport capacity needed by an earlier plan is a different intervention: it has changed the constraints as well as the menu.

The geopolitical implementation retains a field named `s_exit` and evaluates the historical ratio

$$\Gamma_{\text{impl}} = \frac{m\tau}{K \log |S_{\text{exit}}| + \varepsilon},$$

with a separate guard returning zero for counts no greater than one. Appendix B reproduces that behavior. It is important to preserve the

implementation when replaying its reported assessments. It is equally important that its name not conceal the direction of the formula: above the guard, more counted exits lower the ratio at fixed m, τ, K. That makes this ratio unsuitable as a monotone measure of the benefit of adding usable escape routes. The guard is an explicit decision rule, not a continuous limit of the ideal expression.

For interpretation, keep three records beside one another: the legacy calculated ratio, the uncertainty over candidate descriptions, and the feasible settlements supported by the evidence. The legacy number reproduces the old instrument. The candidate-space ratio states the inference argument. The feasible-set analysis establishes whether an additional course can actually be carried out. Moving from one to another requires a specified mapping and a new evaluation; the historical 47-event score is not silently assigned to a new definition.

This distinction strengthens the practical thesis. A working route, a credible guarantee, or a staged withdrawal can enlarge the space of peaceful action. At the same time, discovering more plausible ways the situation might unfold can make the analyst less certain. More freedom for the participants and less certainty for the observer can occur together. The book's central concern is what the participants can do; a measure of what the observer can resolve must serve that question with its meaning intact.

Define the *activation scale* for reconfiguration:

$$E_a = \frac{\gamma}{H^2 + \epsilon} \tag{22.21}$$

where γ is a calibration constant. The interpretation:

For fixed positive γ and ϵ, increasing $|H|$ lowers the activation barrier E_a in this expression. Decreasing $|H|$ raises it. As H approaches zero, the barrier approaches γ/ϵ, its maximum under the stated formula. The direction of that dependence follows from the denominator and should be kept distinct from the political interpretation assigned to it.

The squared term treats positive and negative helicity of equal magnitude identically. The direction of reconfiguration therefore requires the alignment and regime information as well. A low activation scale in a cooperative relationship can facilitate cooperative adjustment; it is not, by itself, evidence of a low barrier to war. With $\gamma = 1$ and $\epsilon = 0.01$, the expression gives $E_a(0) = 100$ and $E_a(1) \simeq 0.9901$.

The trichotomy parameter connects our framework to René Thom's catastrophe theory, which studies how smooth changes in parameters can produce discontinuous changes in behavior. Thom's seven elementary catastrophes classify local mathematical patterns under stated conditions.

The political question is whether a change in capacity and constraints can carry a relationship through one of those patterns.

For our application, the most relevant picture is the *cusp catastrophe*, with two control parameters and one state variable. Think of a thin, bowed lid that can rest bowed upward or downward. Press slowly: the load can move continuously while the lid suddenly snaps through. A toy potential for two such resting states is $U(x) = x^4/4 - ax^2/2 - bx$. At $a = 1$, $b = 0$, it has two minima at $x = -1$ and $x = 1$, separated by a maximum at zero. Changing b tilts the potential; a minimum can meet the intervening maximum and disappear. The point of the example is that an available resting state can cease to exist. Establishing the analogous change in diplomacy would require identifying which settlement ceased to be feasible and why.

The toy potential makes the required conditions explicit. Stationary states obey $x^3 - ax - b = 0$ and local stability under $\dot{x} = -U'(x)$ requires $U''(x) = 3x^2 - a > 0$. At a fold, both $U' = 0$ and $U'' = 0$, which gives $4a^3 = 27b^2$ with $a \geq 0$. Those equations locate the disappearance of an equilibrium in this particular model. A political application would have to identify its state, its dynamics, and how measured constraints enter a and b. Crossing $\Gamma = 1$ alone does not establish a cusp catastrophe. It supplies a candidate diagnostic whose relation to a dynamical transition must be modeled and tested.

The historical indicator specification uses the following five-band classification. These labels reproduce its decision convention; they are not a theorem equating a ratio interval with a number of equilibria (Davis 2026h).

Regime	Γ Range	Interpretation
DETERMINED	$\Gamma > 1.5$	Historical determined band
ELEVATED	$1.0 < \Gamma \leq 1.5$	Historical elevated band
WARNING	$0.5 < \Gamma \leq 1.0$	Examine the remaining settlements
CRITICAL	$0.1 < \Gamma \leq 0.5$	Low ratio; investigate the cause
SINGULARITY	$\Gamma \leq 0.1$	Lowest band or guard; inspect the full state

At $\Gamma = 1$ the numerator and denominator of the ideal ratio match. The band is a decision rule for investigating that balance. A count of equilibria requires a dynamical model, and a chance of war requires a calibrated link to observed outcomes. Appendix B distinguishes these historical band endpoints from the current engine's alert tests.

a Regime convention in this illustration

Singularity	$\Gamma \le 0.1$
Critical	$0.1 < \Gamma \le 0.5$
Warning	$0.5 < \Gamma \le 1$
Elevated	$1 < \Gamma \le 1.5$
Determined	$\Gamma > 1.5$

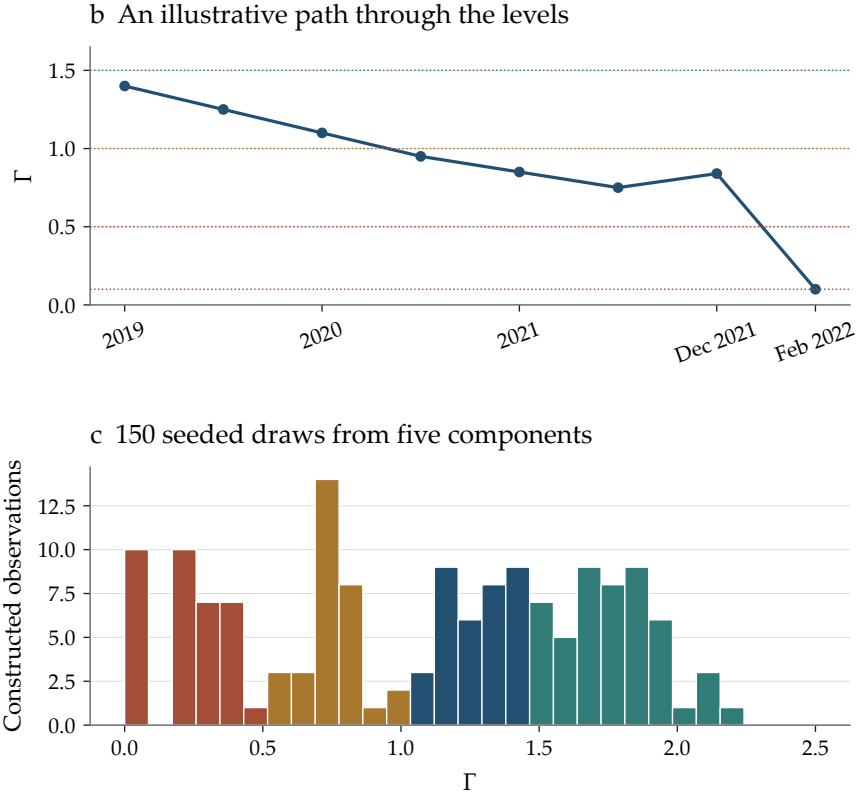


FIGURE 22.4: Historical five-band convention, a constructed trajectory, and 150 seeded draws from five prescribed components. The endpoints in the first panel belong to the historical table. Algorithm 10 in Appendix B gives the current alert rules.

Derivation of Critical Thresholds

The framework's predictive power depends on correctly calibrating the critical thresholds. This section documents the derivation process, which combined theoretical considerations with empirical optimization.

Several thresholds have theoretical justifications that constrain their values:

Spectral gap threshold $\lambda_1^* = 0.05$: For the normalized undirected graph, Cheeger's inequality gives $\mu_1 \geq h^2/2$. The quantity h uses crossing weight divided by the smaller side's total weighted degree; its denominator is not the total number of edges in the network. A requirement of $h > 0.1$ would imply $\mu_1 > 0.005$. The stronger numerical operating threshold of 0.05 comes from the framework's calibration on its stated network, rather than from the inequality alone.

Helicity threshold $H^* = -0.5$: This is an empirically chosen operating point. In the normalized angular formulation, $H = \cos\theta$, so the negative threshold corresponds to 120 degrees. At that angle the projection of a unit flow vector onto the other unit direction is minus one-half. The geometry explains the sign and size of the projection; the cases establish why the chosen threshold is useful for warning.

Trichotomy reference value. Unity is the point at which the numerator and denominator of the ideal ratio agree. A band such as $[0.9, 1.1]$ is an additional operating choice. Interpreting it as an uncertainty interval requires the input-error model and its propagated distribution; a ten-percent band is not automatically two standard deviations.

The count-based 47-event evaluation belongs to the mass-gap and Z-score model: it reports AUC 0.84 and $F_1 = 0.73$ (Davis 2025). The research also records a separate 30/17 historical split, component AUCs, and results under changed internal voting thresholds. Appendix E retains those values and distinguishes them from the mass-gap study's 31/16 temporal split. The numerical agreement of two sample sizes does not make their event assignments or scores identical.

A ROC curve is generated by ranking observations with a fixed score and sweeping its decision cutoff. Moving that cutoff does not change the area under the curve. Changing an internal vote or refitting a model can change the ranking and therefore the AUC, but then the scored outputs must be retained for each configuration. The reference calculation in the companion checks both cases on a small example.

The practical calibration sequence follows from these distinctions. Freeze the event definition, information cutoff, and candidate models. Fit weights and choose the operating threshold within the earlier period. Score the later period once under those choices. Report all qualifying

warnings and misses, together with the uncertainty created by having a limited number of conflicts. Thousands of correlated monthly observations do not turn forty-seven onsets into thousands of independent wars.

The geometry constrains which constructions make sense. The empirical record determines which of the admissible constructions earns its forecasting claim.

A scalar reading is one part of the construction. The network has a shape as well. To see how that shape constrains adjustment, we return to the Laplacian and ask why a disturbance can linger between groups even while activity within them remains strong.

The Geometry of Feasible Adjustment

The harbor master has two terminals and one day's cargo to unload. Every tonne must go through one terminal or the other; neither crane can lift more than its rated daily capacity. A plan on paper becomes a feasible plan when it satisfies both facts at once. This is what a set of constraints does for a relationship. Let $x \in \mathbb{R}^n$ collect the quantities we can change, with conditions for resources, institutional commitments, or other limits included in the model.

Write the feasible set as

$$\mathcal{F} = \{x : g_i(x) \le 0,\ i = 1, \dots, r;\ h_j(x) = 0,\ j = 1, \dots, s\}.$$

The inequalities can represent capacities or limits. The equalities can represent accounting relations or commitments treated as fixed over the time horizon. A political commitment should be represented as fixed only when that assumption is appropriate to the question being asked.

A country may be able to change a commitment, but only at a cost. Such a possibility belongs in a larger model that makes the cost explicit. Treating it as an immutable equality asks what adjustments are available while the commitment remains in force.

Suppose there are two coordinates and three reported constraints:

$$x_1 + x_2 = 1, \qquad 2x_1 + 2x_2 = 2, \qquad 3x_1 + 3x_2 = 3.$$

There are three equations, but only one independent restriction. Their common solution is a line. Counting documents or observations without examining independence would overstate what has been determined.

Let x_1 and x_2 be the fractions of the day's cargo unloaded at the terminals. The first equation assigns the full cargo; the next two merely multiply that manifest by two and three. If one terminal takes a further

tenth, the other must take a tenth less. This is the one remaining degree of freedom. For physical shares, also impose $x_1, x_2 \geq 0$: now only the segment of the line between $(1, 0)$ and $(0, 1)$ can describe an actual delivery.

For smooth equality constraints, form the Jacobian J_h, whose rows are the gradients of the h_j. Near a regular point where its rank is r_h, the equality-constrained manifold has dimension $n - r_h$. This local statement requires the regularity condition. At a singular point, a simple subtraction of counts can fail.

Several treaties may enforce the same underlying commitment. Several sources may repeat the same original report. Their number differs from the number of independent restrictions or observations they supply. This gives practical content to the framework's concern with sufficient constraints: independence and conditioning matter alongside the count.

At a point x, consider a small proposed movement v. To first order, the equality constraints require

$$J_h(x)v = 0.$$

For an active inequality, one with $g_i(x) = 0$, a feasible first-order movement must satisfy

$$\nabla g_i(x)^T v \leq 0.$$

A proposed movement that fails either test is ruled out immediately. Passing the linear tests gives a candidate direction. At a singular constraint the first derivative can miss a restriction entirely. For instance, a lock that holds a gate at $x = 0$ can be described by $x^2 \leq 0$. At zero its derivative vanishes in every direction, although the gate can only stay put. At an ordinary capacity boundary the crane example below shows how the first derivative rules out an overload.

The manifest also explains what the Jacobian does. Its first row, $(1, 1)$, adds the proposed changes at the two terminals. A transfer $v = (-0.1, 0.1)$ passes the equality test because the total stays fixed. But suppose the second terminal can handle only forty percent of the day's cargo and the plan is already $(0.6, 0.4)$. Its capacity constraint is $g = x_2 - 0.4 \leq 0$. The proposed transfer has $\nabla g^T v = +0.1$: it would overload that crane. Reversing the transfer passes both tests for a sufficiently small move. The equality leaves a choice; the active capacity limit rules out one direction of that choice.

An economy may be able to replace the next small increment of imported fuel using spare capacity. The same supplier may not be able to replace the entire flow. Local flexibility and a complete exit are different claims.

This is one reason a relationship can appear manageable until a threshold is reached. A reserve, transport route, or institutional procedure may have capacity for small adjustments but no capacity for a larger one. The constraint becomes active, and the set of feasible movements narrows.

Linearize the equality constraints around a point:

$$J_h \, \delta x \simeq \delta b.$$

The right-hand side represents a change in the constraints or their measured values. If the relevant singular values of J_h are small, a modest change in δb can require a large change in the inferred δx. The estimate is poorly conditioned.

With a full-column-rank Jacobian, the least-squares sensitivity obeys

$$\|\delta x\| \leq \frac{\|\delta b\|}{\sigma_{\min}(J_h)}$$

for the linearized solution. The bound concerns that specified local problem and its chosen norms. It is not a universal law relating diplomatic tension to violence.

The result nevertheless provides a useful diagnostic question. Is an apparent change driven by substantial new evidence, or by an inference problem highly sensitive to small changes? An analyst should distinguish a fragile estimate from a fragile political system. Either matters, but they call for different responses.

Consider a separate weighing problem in tonnes. A total reports $x_1 + x_2 = 100$; a second instrument reads $x_1 + 1.01x_2 = 100.40$ because its known response to the second load is one percent higher. Subtract the tickets: $0.01x_2 = 0.40$, hence $x_2 = 40$. A change of just 0.01 tonne in the second ticket, to 100.41, moves the inferred second load to 41 tonnes. We are trying to separate two loads with instruments that see almost the same thing. An independent weighing of one truck could teach us more than another decimal place on a nearly duplicate total.

For a smooth path $x(s)$ parameterized by arc length, the curvature is

$$\kappa(s) = \left\| \frac{d^2 x}{ds^2} \right\|.$$

This is geometric curvature of a trajectory in the chosen metric. It is distinct from an operational index of integration burden, even when the latter is also called curvature in the framework. Connecting the two requires a specified embedding, metric, and path.

A contractual burden and a sharply bending trajectory can both hinder adjustment through different constructions. Naming what the calculation measures protects the intuition.

Coordinates need compatible scales. Combining dollars with a unit interval makes Euclidean distance depend on arbitrary units. An explicit metric can rescale the axes or incorporate estimated covariance; distances, angles, and curvature depend on that choice.

An admissible initial direction does not establish reachability. Disconnected components, narrow passages, and later boundaries can make a path satisfying every immediate constraint lead to a dead end.

Conversely, a costly first step may reach more alternatives. Evaluate the whole implementation path: a resilient final trade pattern may require a transition through an intolerable shortage or an unratifiable commitment.

The historical cases show practical responses. A temporary arrangement buys time; a review procedure turns an impasse into steps; a common institution provides a continuing place to negotiate adjustments.

Geometry gives these actions a common structure: a starting configuration, constraints, candidate paths, costs, and a destination whose reachability must be established. That is how the mathematics becomes a practical account of prevention.

Chapter 23

Spectral Geometry Foundations

> Can one hear the shape of a drum?
>
> Mark Kac, 1966

Place two metal blocks in contact through a narrow bridge and warm one of them. Heat crosses the bridge, but the temperature difference can persist long after each block has become nearly uniform inside. The bridge determines the slow part of the response.

A network calculation asks a comparable structural question. Which differences are easy to smooth away, and which remain because the connections between groups are weak? We already calculated one answer for three countries. Here we follow the operator that made that answer possible. Spectral geometry asks what eigenvalues can tell us about the structure of a space.

The Laplacian operator lies at the heart of spectral geometry. In essence, it measures how a function at a point differs from its average in the surrounding neighborhood. This deceptively simple idea has extraordinary reach—from the physics of heat conduction to the mathematics of graph connectivity to the economics of shock propagation. Understanding the Laplacian deeply is essential to understanding why the spectral gap serves as a systemic stability indicator.

The Laplacian measures a local departure from surrounding values. Its continuous form adds second spatial derivatives. A temperature field supplies a concrete example. At a hot spot surrounded by cooler material, the operator participates in the diffusion law that describes how the excess spreads. In Cartesian coordinates,

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \tag{23.1}$$

the Laplacian measures the *deviation* of a function from its local average. A function is *harmonic*—satisfying $\Delta f = 0$—precisely when its value at

every point equals the average of values on every sphere centered there whose closed ball lies within the domain.

The physical applications are immediate and profound. The heat equation $\partial_t u = \Delta u$ describes how temperature equilibrates: regions hotter than their surroundings cool down; regions colder warm up. The wave equation $\partial_{tt} u = c^2 \Delta u$ describes vibrations. The Laplace equation $\Delta u = 0$ describes steady-state distributions—the equilibrium a system reaches when all transient effects have dissipated.

The same operator, applied to graphs, yields equally profound results about network structure.

On a graph representing international trade relationships, the Laplacian becomes a matrix operator:

$$(\Delta f)(i) = \sum_{j \sim i} w_{ij}(f(i) - f(j)) \tag{23.2}$$

where w_{ij} represents the strength of the economic connection between countries i and j (typically bilateral trade volume normalized by GDP), and the sum runs over all trading partners.

Equivalently, in matrix form:

$$L = D - A \tag{23.3}$$

where D is the diagonal degree matrix ($D_{ii} = \sum_j w_{ij}$) and A is the weighted adjacency matrix ($A_{ij} = w_{ij}$).

For each neighbor the row asks how different its reading is from mine, and weights the answer by our connection. A row can return zero in the heart of a dense trading bloc if everyone's reading is the same. Zero says there is no net local gradient for this particular pattern; it does not by itself say the country is isolated.

Watch the sign at a hot spot. In a metal bar the temperature profile bends downward there, so the continuum $\nabla^2 f$ is negative. On the trade graph the row of $L = D - A$ subtracts each neighbor's reading from the local one, so a strict local maximum with a positive-weight connection to a cooler neighbor gives a positive result. With unit diffusivity the heat generator is ∇^2 in the bar and $-L$ on this graph. Both make the hot spot cool.

Consider a network where node i has "temperature" $f(i)$—representing, say, economic output deviation from trend. The discrete heat equation:

$$\frac{df(i)}{dt} = -(\Delta f)(i) = \sum_{j \sim i} w_{ij}(f(j) - f(i)) \tag{23.4}$$

describes transfer down differences along each edge. A country two units above trend can still receive a shock from a partner four units above trend. The direction depends on the difference between neighbors. Summing over all countries cancels each transfer against the receipt at the other end.

Consider two groups of countries trading heavily inside each group and sharing a single small crossing. Adding another internal route can leave that crossing just as narrow. The weights and the arrangement of the edges determine how quickly a shock passes between the groups.

Cheeger's Inequality and the Fiedler Vector

Chapter 7 introduced Cheeger's inequality for the normalized undirected graph: $h^2/2 \leq \lambda_1 \leq 2h$, where λ_1 here denotes the normalized gap. It tells us a narrow crossing and a small gap are related. The *Fiedler vector* helps us look for the crossing.

The eigenfunction corresponding to λ_1 (called the *Fiedler vector*) points in the direction of the bottleneck. Its values suggest a partition of the vertices. A sweep over thresholds along the vector produces candidate cuts; the zero-crossing is a useful visual division but is not generally guaranteed to be the optimal discrete cut. Thus spectral analysis not only detects that a bottleneck exists but also *identifies where it is*.

Here is what a sweep asks an analyst to do. Give each country its coordinate in the slow pattern, using the degree-weighted coordinates x of $Lx = \lambda_1 Dx$. Sort them from low to high. Put the first country on one side of a proposed cut and count the trade that crosses; repeat with the first two countries, then the first three. For each cut, divide crossing weight by the smaller side's total weighted degree. Suppose one proposed division has crossing weight 4 and the two sides have volumes 100 and 150. Its ratio is $4/100 = 0.04$. Most connections stay within the two groups; relatively little crosses between them. That cut establishes $h \leq 0.04$, even if a still narrower cut exists. Negating the eigenvector reverses the sorting and exchanges the names of the sides. No country's alliances have changed.

For international relations, this means: by computing the spectral gap of the trade network, we can assess how easily the system could fragment into hostile blocs. A collapsing spectral gap signals that the network is crystallizing into opposing camps with weakening cross-bloc ties.

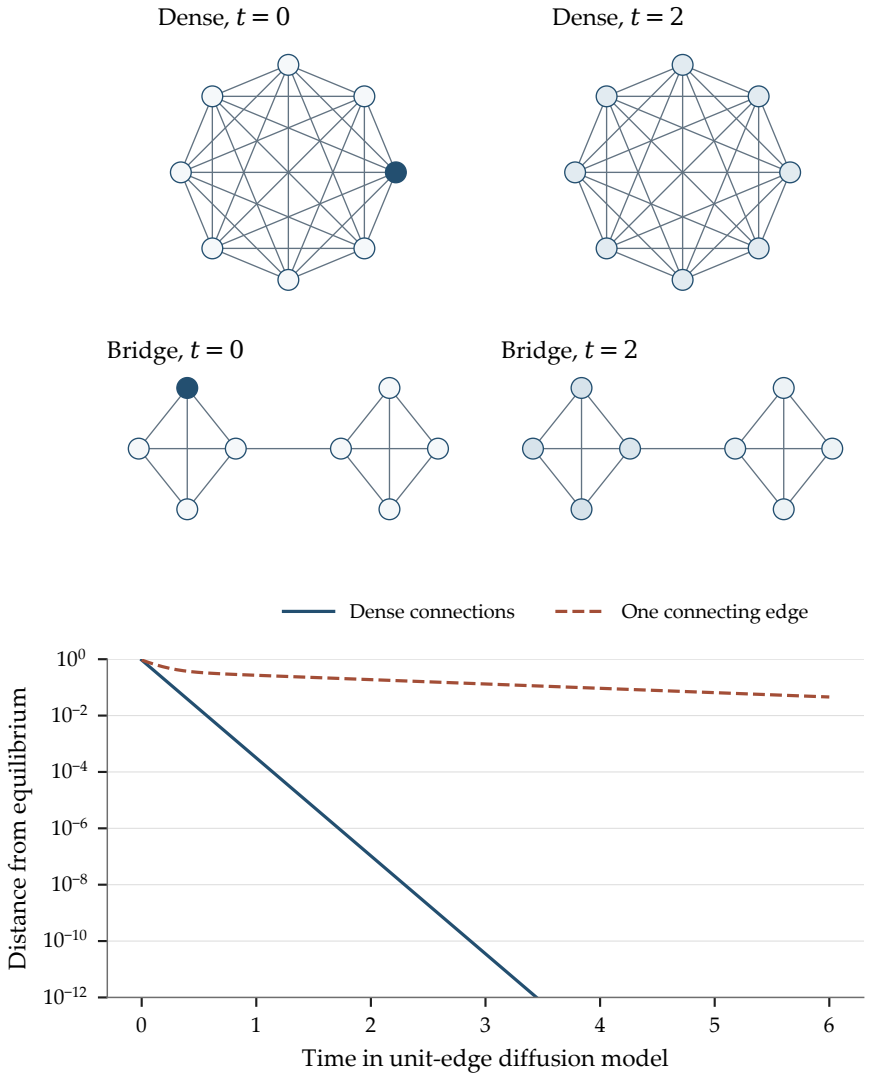


Figure 23.1: Diffusion computed from the displayed graphs using $u(t) = e^{-tL}u(0)$, with $L = D - A$ and unit edges. All mass starts at one node. Both rows use the same scale and retain total mass one. The lower curves show distance from the uniform equilibrium. Time belongs to this dimensionless model.

Period	Fiedler Structure	Interpretation
1990–2007	Scattered, no partition	Globalization era
2008–2015	Weak US/EU vs. emerging	Post-crisis realignment
2016–2019	US-China polarity	Trade war crystallization
2020–2024	Western vs. Sino-Russian	Bifurcation underway

The Fiedler vector analysis reveals not just that the network is fragmenting, but *along what axis*. Since 2016, the first nonconstant eigenvector has increasingly aligned with a US/EU versus China/Russia partition, suggesting that this is the natural "fault line" along which the global trading system may split.

Heat Kernel Interpretation

The spectral gap has an elegant physical interpretation through heat diffusion. This interpretation connects the abstract eigenvalue analysis to intuitive physics and provides deep insight into why spectral gap collapse precedes systemic crises.

If we model economic shocks as "heat" that spreads through the trade network, the heat kernel describes this propagation. The heat kernel $p_t(x, y)$ answers the question: if we inject one unit of heat at node x at time 0, how much heat will be at node y at time t?

The answer involves the full spectrum of the Laplacian:

$$p_t(x, y) = \sum_k e^{-\lambda_k t} \phi_k(x) \phi_k(y) \tag{23.5}$$

Here the operator is the symmetric combinatorial Laplacian of a connected, undirected graph with nonnegative weights. Its constant mode gives a uniform limiting heat distribution. A random-walk operator instead has a degree-weighted stationary distribution. The λ_k are the Laplacian eigenvalues (ordered $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots$) and ϕ_k are the corresponding eigenfunctions (orthonormal basis of the function space on the graph).

This formula encapsulates a remarkable amount of information:

The term $e^{-\lambda_k t}$ shows that each eigenmode decays exponentially with rate λ_k. Larger eigenvalues decay faster; smaller eigenvalues persist longer.

The product $\phi_k(x)\phi_k(y)$ tells how this particular mode contributes between two nodes. It can have either sign. The full heat kernel sums all

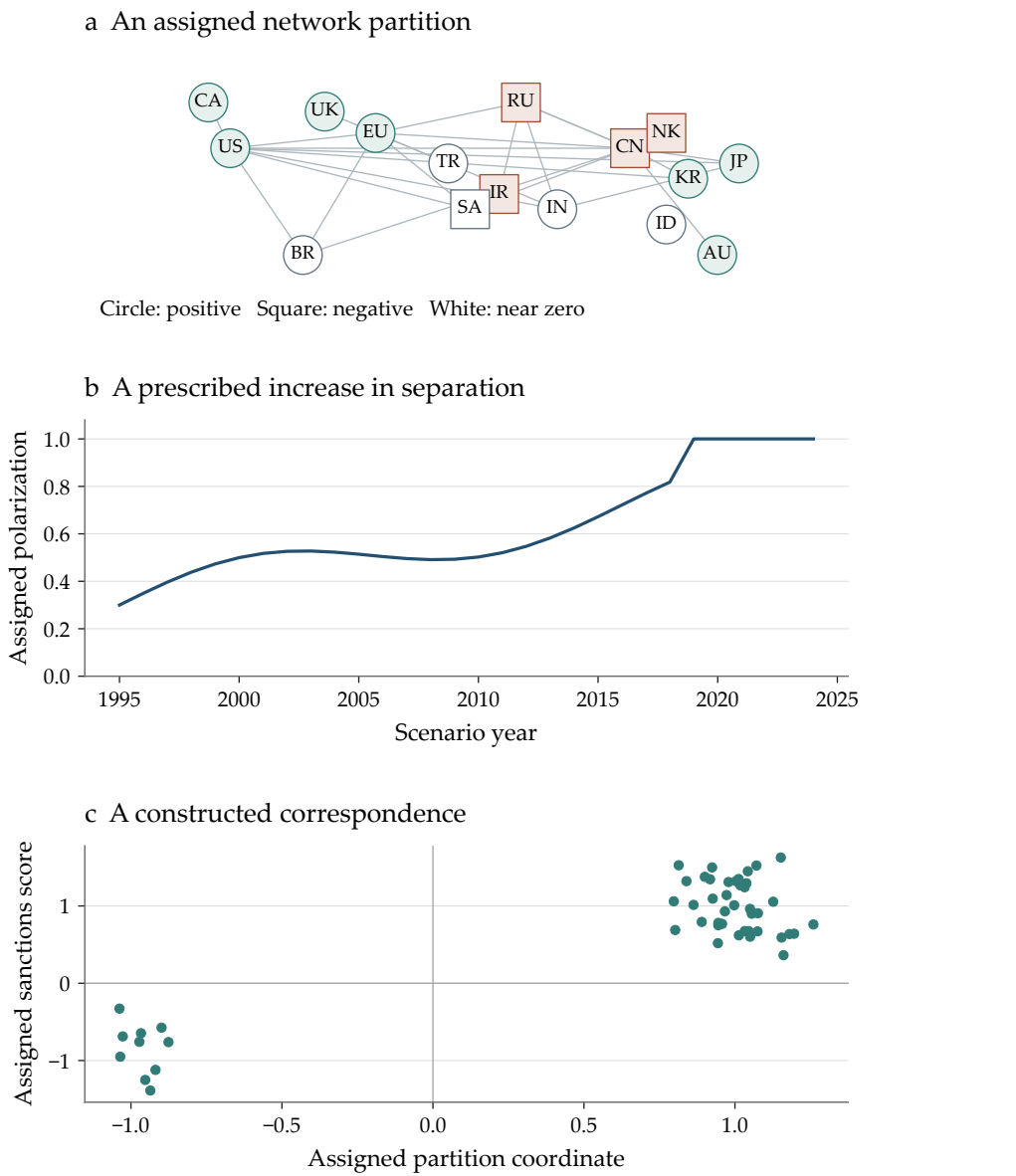


Figure 23.2: An assigned partition and two illustrative comparisons. Node shapes distinguish the signs of the assigned coordinates; white nodes are near zero. The polarization and sanctions panels are prescribed examples. These coordinates illustrate partitioning; computing a Fiedler vector requires the eigenvalue problem in the text.

the modes with their decay factors. In the symmetric model a map of one mode shows where its deviations run together or oppose each other.

At $t = 0$, the sum equals δ_{xy}, one at the starting node and zero elsewhere (the Kronecker delta on a finite graph). As $t \to \infty$, only the $\lambda_0 = 0$ term survives, giving uniform distribution—complete equilibration.

The exponential decay $e^{-\lambda_k t}$ means that smaller eigenvalues contribute more to long-time behavior. The spectral gap λ_1 controls the timescale for equilibration:

$$\tau_{\text{relax}} = \frac{1}{\lambda_1}, \qquad t_\delta = \frac{\log(1/\delta)}{\lambda_1} \quad (0 < \delta < 1). \tag{23.6}$$

The first time is the e-folding time of a pure slow mode. The second is the time for that mode's contrast to reach a specified fraction δ of its initial size. A mixing time for a probability distribution also specifies a distance to equilibrium and a tolerance, often taking the worst starting node (Levin and Peres 2017).

This is the mathematical foundation for "critical slowing down." When the spectral gap is large, perturbations decay quickly: $e^{-\lambda_1 t}$ approaches zero on a timescale of $1/\lambda_1$. When the spectral gap is small, perturbations decay slowly: the system "remembers" shocks for a long time.

To see what that time means, fill two equal, well-stirred tanks and pump water between them at equal rates in both directions. Each hour the pumps exchange one tenth of a tank volume each way. Put dye in the first tank at concentration one and leave the second clear. If their concentrations are u and v, the first loses dye at net rate $0.1(u - v)$ per hour and the second gains exactly that much. Their mean stays at one-half. Subtract their two rate equations: the contrast $u - v$ decays at 0.2 times itself per hour. That 0.2 is the spectral gap of this two-tank model.

After five hours the contrast has fallen from one to $e^{-1} \simeq 0.368$. The concentrations are about 0.684 and 0.316. Both tanks now contain dye, but they are still visibly different. The reciprocal gap of five hours is the time to retain about 37 percent of the original contrast; it is not the time to reach exact equality. To get below one tenth of the original contrast takes $\log(10)/0.2$, about 11.5 hours. We can quote hours here because we measured a pump rate. A trade graph with weights in billions of dollars needs its own justified time conversion before its reciprocal gap can be read as months.

The heat kernel explains the slower recovery in Chapter 7. To connect it to variance and autocorrelation, give the slow pattern repeated disturbances. Suppose one observation interval leaves a fraction $r = e^{-\lambda_1 \Delta t}$ of

yesterday's contrast and adds an independent, zero-mean shock of fixed variance s^2. At stationarity, today's variance V satisfies $V = r^2V + s^2$, so $V = s^2/(1 - r^2)$ and the lag-one correlation is r. As r approaches one from below, yesterday's backlog persists and new shocks arrive before it has cleared. This calculation tells us what to check in a political series: a larger disturbance or a change in reporting can also widen the observations, without any change in the recovery rate.

For geopolitical systems, this manifests concretely: diplomatic incidents that previously would have been resolved quickly now fester, sanctions that would have been absorbed cause spiraling retaliation, and relationship restoration takes progressively longer after each rupture.

Period	Helicity	Autocorr.	Recovery	Interpretation
2008–2012	+0.15	0.70	8 mo.	Stable; shocks absorbed
2013 (pre-Maidan)	−0.10	0.78	12 mo.	Stress emerging
2014 (Crimea)	−0.45	0.83	18 mo.	Annexation shock
2015–17 (Minsk)	−0.35	0.85	24 mo.	Partial recovery stalls
2018–2019	−0.50	0.88	30+ mo.	Chronic deterioration
2020	−0.65	0.91	Open	Glassy regime entered
2021 Q1–Q3	−0.80	0.93	Open	Accelerating collapse
2021 Q4 (L4)	−0.95	0.95	Open	Singularity detected

Table 23.1: Critical slowing down in the Russia–Ukraine record. Reported autocorrelation rises from 0.70 to 0.95. An open recovery interval means the series had yet to return to its reference condition by the end of the observation window.

Application to Trade Networks

With the mathematical foundations established, we now turn to concrete application. How do these abstract concepts—Laplacians, eigenvalues, heat kernels—translate into practical analysis of international trade networks?

The first step is constructing the network from raw data. We use bilateral trade flows from the IMF Direction of Trade Statistics (DOTS)

database, which reports exports and imports between all country pairs at monthly frequency going back to 1948.

The construction process:

Symmetrization: First reconcile two reports of the same direction. If i reports exporting 9.8 billion to j and j reports importing 10.2 billion from i on a comparable basis, their average is 10 billion, not 20 billion of trade. Do the same for the reverse flow. If it is 4 billion, a bilateral edge defined as the sum of both directions has weight 14 billion and occupies both A_{ij} and A_{ji}. Reconciliation addresses disagreement about one shipment flow; symmetrization decides that the network will represent the two directions by one weight.

Normalization: Edge weights are normalized by the geometric mean of the two countries' GDPs: $w_{ij} = T_{ij}/\sqrt{GDP_i \cdot GDP_j}$. This ensures that trade between large economies doesn't dominate by scale alone.

Thresholding: We include only edges above a minimum threshold (typically \$1 billion annually) to eliminate noise from tiny trade relationships.

Annual averaging: For spectral analysis, we use annual averages to smooth seasonal fluctuations while preserving trends.

The resulting network has approximately 200 nodes (countries) and 5,000–10,000 significant edges, depending on the threshold and time period.

In a trade network, the abstract concepts become concrete:

In the political analogy, an economic shock takes the place of a localized deposit of heat. Financial crises, supply disruptions, or sanctions propagate along trade channels. When λ_1 is small, a disturbance takes longer to spread across the network and its effects remain concentrated on one side of a bottleneck. A falling gap thus offers a way to describe a network separating into blocs. The diffusion calculation supplies the propagation model; the interpretation of its weights supplies the connection to trade.

The glass picture in Chapter 24 asks how long a region can remain under strain while trade still moves elsewhere. A small gap lets an imbalance persist in the slow mode. For it to grow, the political mechanism must add pressure, for example when a delayed shipment forces a factory to draw down its last reserve.

The framework's calibrated threshold is $\lambda_1 < 0.05$ for its specified construction. Cheeger's inequality explains why the normalized gap can reveal a bottleneck. It does not set the numerical alert threshold. Below that operating point the political question is which economic or diplomatic routes can still absorb a disruption before it becomes costly.

Historical evolution of the global trade network spectral gap:

Year	λ_1	h	Context
1990	0.14	0.42	Post-Cold War integration begins
1995	0.15	0.44	WTO formation, trade expansion
2000	0.14	0.41	China WTO accession approaching
2005	0.12	0.35	Peak globalization, supply chains deepen
2008	0.08	0.28	Financial crisis, temporary fragility
2010	0.11	0.33	Recovery, continued integration
2015	0.10	0.31	Pre-trade war baseline
2018	0.09	0.29	US-China trade war begins
2020	0.07	0.25	Pandemic disruption
2022	0.06	0.22	Russia sanctions, supply chain restructuring
2024	0.05	0.20	Approaching critical threshold

The reported global index falls from 0.12 in 2005 to 0.05 in 2024, a drop of about 58 percent relative to 2005. The last entry, displayed to two decimal places, sits at the stated threshold. A decision rule with a strict cutoff needs the unrounded value to classify that observation.

The Fiedler vector analysis reveals the emerging bloc structure. Computing the eigenvector corresponding to λ_1 and examining its sign structure:

Positive Fiedler	Negative Fiedler
United States	China
European Union	Russia
Japan	Iran
South Korea	North Korea
Australia	Belarus
Canada	Venezuela
United Kingdom	Cuba

Countries with Fiedler values near zero—India, Brazil, Saudi Arabia, Indonesia, South Africa—are the "swing states" of the emerging bifurcation. Their trade relationships span both emerging blocs, and their alignment decisions may determine whether the network splits cleanly or maintains bridges.

Network bifurcation imposes enormous economic costs. When the spectral gap falls below 0.05, simulation studies suggest:

The economic costs fall most heavily on the countries occupying the bottleneck. The estimates discussed here put their GDP loss at 3–5%, with disrupted supply chains taking 5–10 years to rebuild. Interrupted knowledge flows can slow innovation as well. The political cost follows from that economic damage: as relationships lose their ability to impose reciprocal costs, economic deterrence weakens and conflict becomes easier to contemplate.

A shipping analyst can identify the narrow crossing from the graph. To turn its closure into a GDP loss she also needs to know what passes there, how long replacement takes, and who can draw down stocks while the route is closed.

A slow mode explains persistence within a specified evolution. Glass physics brings another piece into view: the range of arrangements the system can reach and the barriers separating them. That is where the question of connectivity meets the question of time.

Chapter 24

Glass Transitions in Physical Systems

> A route that takes a year cannot solve a crisis with six weeks left.
>
> Bee Rosa Davis

A window holds its shape while the molecules within it lack the repeating arrangement of a crystal. To understand why, we have to ask how quickly those molecules can rearrange. The time available to observe the material becomes part of the question.

The political problem that brought us here also concerns rearrangement. A government may recognize a more desirable relationship and still be unable to reach it through the actions available before a crisis develops. The glass transition analogy asks us to examine those obstructed routes. Its mathematical content lies in the description of states, barriers, and the time needed to move between them.

Glass is an amorphous solid. The puzzle is that an ordinary window resists shear even though its atoms lack the long-range repeating order of a crystal. Asking where the atoms sit is one question; asking how long they take to change neighbors is another. The second question gives the political analogy its force: a route that can be opened in five years offers little help to a factory whose reserve runs out on Friday.

At the molecular level, glass exhibits paradoxical properties:

Disordered like a liquid: Glass has no long-range crystalline order. Its atoms are arranged randomly, with no repeating pattern. X-ray diffraction shows the diffuse rings characteristic of liquids, not the sharp peaks of crystals. If you photographed the positions of atoms in a glass and in a liquid at the same density, you could not tell them apart.

Rigid like a solid: Despite its disorder, glass supports shear stress. It maintains its shape. It breaks rather than flows. For all practical purposes, window glass behaves as a solid. Its elastic modulus is comparable to crystalline ceramics.

Cannot reach equilibrium: Glass is not in its lowest energy state—that would be a crystal. It is trapped in a metastable configuration, unable to rearrange on any timescale relevant to human observation. Glass is a liquid that has been *frustrated* out of crystallization.

The glass transition itself is unlike ordinary phase transitions. Water freezes at 0°C; the transition is sharp, with a clear discontinuity in properties like density and heat capacity. The heat of fusion is released precisely at the freezing point. First-order phase transitions have this character: clean, discontinuous, thermodynamically determined.

But cooling a glass-forming liquid produces no such discontinuity. Viscosity increases smoothly, relaxation times grow, and at some point the system simply stops responding to perturbations on experimental timescales. The "transition" is kinetic, not thermodynamic. There is no glass-forming "heat of transition." The glass transition temperature T_g depends on how fast you cool: faster cooling yields higher T_g because the system has less time to rearrange before becoming stuck.

A glass retains structural disorder while its rearrangements become too slow to equilibrate on the observation timescale. That distinction between a configuration and the time required to change it is the part of the glass analogy that matters here.

Glassy behavior is surprisingly common in nature:

Ordinary glass (silica): The archetypal example. Window glass has been "frozen" for centuries, retaining the structure of a supercooled liquid.

Polymers: Most plastics are glasses at room temperature. Their long molecular chains become entangled, unable to rearrange into crystalline order.

Metallic glasses: Rapidly cooled metal alloys can form glassy structures with unusual mechanical properties—strong but brittle.

Spin glasses: Magnetic systems with frustrated interactions that cannot find a consistent alignment. Each spin wants to align with its neighbors, but the neighbors are configured such that not all alignments can be satisfied.

Structural glasses: Granular materials, colloids, foams—systems of macroscopic particles that jam into rigid configurations without crystalline order.

What unites these diverse systems is not their material composition but their dynamical behavior: *frustration leading to kinetic arrest*. The system would "prefer" to order (lower energy) but cannot find a path from disorder to order. It is stuck.

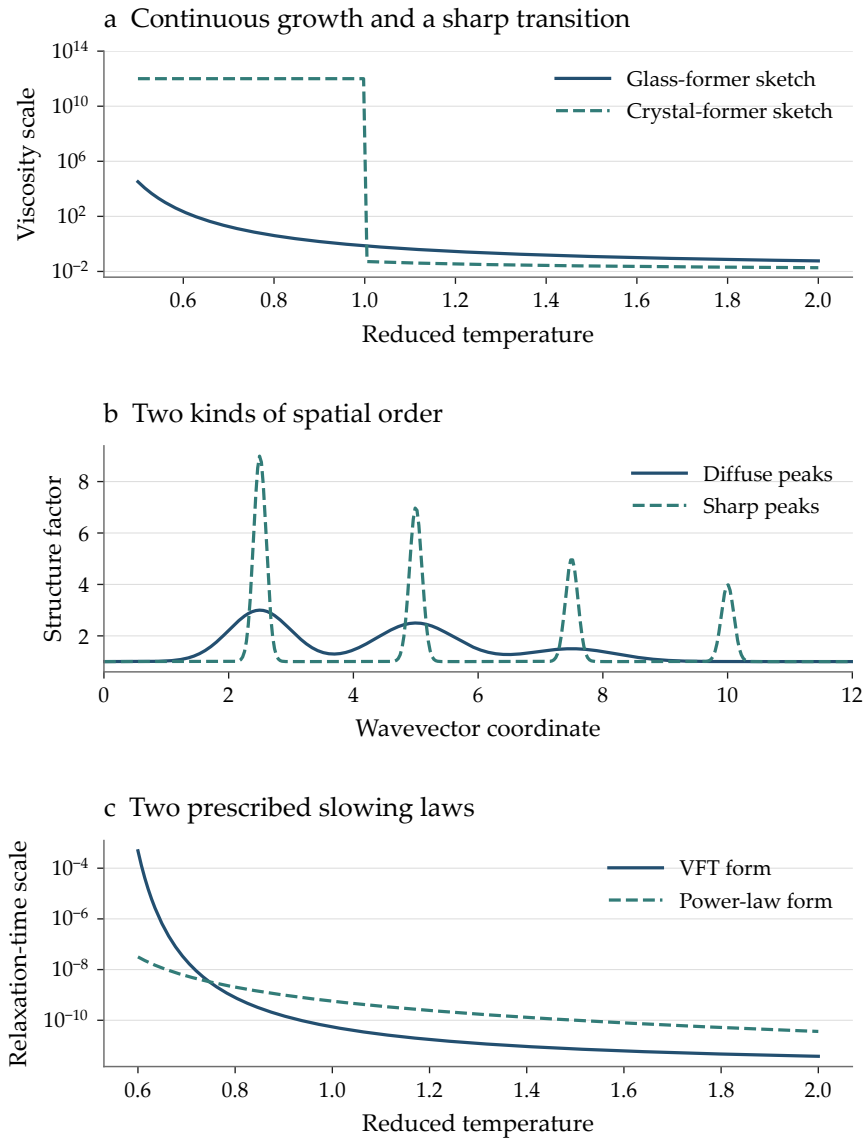


Figure 24.1: Prescribed functions illustrating material behavior: a continuous rise versus a sharp transition, diffuse versus sharp structural peaks, and exponential versus power-law relaxation. The first panel's glass-former uses a shifted-temperature exponential. Scales and parameters belong to the illustration, not a fitted material dataset.

The Energy Landscape

The modern understanding of glasses centers on the concept of the *energy landscape*—a high-dimensional surface where each point represents a possible configuration and elevation represents energy. This framework, developed by David Wales, Frank Stillinger, and others, provides the conceptual foundation for understanding why glasses form and why they are so difficult to escape.

In the landscape description, a configuration can make many small local adjustments while access to another basin remains slow. The height of a barrier and the possible paths around it affect the escape time. A drawing of the endpoint alone cannot show whether that endpoint can be reached within the observation period.

The observation period is essential. Motion that looks negligible over a second may matter over a year. Calling an arrangement frozen therefore requires a clock. In a political case, the relevant interval may be the time before an election, the exhaustion of a reserve, or the next delivery deadline.

The same insight applies to international systems: the landscape of possible relationships determines which configurations are accessible and which are not. Former Senator George Mitchell, who mediated the Northern Ireland peace process, recognized this intuitively: "The parties weren't stuck because they couldn't imagine peace. They were stuck because every path to peace required someone to move first—and no one wanted to be the one who moved."

Consider a negotiation in which both sides can describe a desirable settlement, but neither can deliver its first required concession. A guarantee, an inspection arrangement, or a phased exchange might change what is possible. The destination was already visible. The intervention creates a route to it.

A territorial map does not contain all those constraints. The negotiator also needs to know which concession each party can authorize, what it expects in return, and whether the commitments can be performed together. That is the practical content of a landscape of possible settlements.

Consider a system of N particles in three dimensions. Its configuration is specified by $3N$ coordinates—the position of each particle. The energy landscape is a surface in this $3N$-dimensional space, where each point has an associated energy.

In the simple crystal picture, the ordered structure occupies a deep basin. A small displacement around that arrangement costs energy, so a gently disturbed solid can settle back. Real crystals also have defects

and competing structures. On this drawing the single valley isolates the cost of escaping one particular arrangement.

For a glass-forming liquid, the landscape is fundamentally different. It is *rugged*: countless local minima separated by barriers, with no clear path to the global minimum. The global minimum (the crystal) may exist, but it is hidden among an exponentially large number of metastable states that all look locally optimal.

Imagine a mountain range with countless valleys. A ball placed on the surface will roll into a valley—but not necessarily the deepest one. Once trapped in a shallow valley, the ball cannot escape without enough energy to cross the surrounding ridges.

At high temperature (lots of kinetic energy), the ball can hop between valleys, exploring the landscape freely. As temperature drops, the ball becomes increasingly localized, able to explore only its immediate neighborhood. Eventually, it is trapped—unable to reach the deeper valleys it "knows" exist somewhere out there.

A glassy international system is similarly trapped. The current configuration of alliances, enmities, and dependencies may not be optimal for any party. But the cost of reconfiguration—the height of the barriers—exceeds the energy available through normal diplomatic means.

Consider a factory changing suppliers. It currently needs ten million in working capital to keep deliveries moving; the new arrangement would need only five. But the new supplier requires prepayment while the old invoices are still outstanding. One transition schedule briefly ties up seventeen million. A staged delivery schedule reaches the same destination while tying up at most thirteen. With a credit line of fifteen million, the first route is closed and the second is open. Both save the same amount in the end. The barriers are the extra seven million and three million needed along the way. In this example the vertical axis of the landscape has an actual unit, and a bank's credit limit tells us which ridge can be crossed.

A key insight comes from counting distinguishable basins available at a given energy. In this description the *configurational entropy* $S_c(E)$ is related to the logarithm of that count. The logarithm makes independent choices additive: if one subsystem has two alternatives and another has three, together they have six, and $\log 6 = \log 2 + \log 3$. Cooling can reduce the effective number of accessible arrangements. Actual laboratory arrest is diagnosed on an observation timescale; it does not require measuring $S_c = 0$.

The Adam-Gibbs theory relates relaxation time to configurational entropy:

$$\tau = \tau_0 \exp\left(\frac{C}{TS_c(T)}\right) \tag{24.1}$$

As $S_c \to 0$, relaxation time diverges exponentially. The system becomes kinetically frozen even though thermodynamically it "should" still be exploring.

A basin in the drawing stands for an arrangement of alliances, trade and political commitments that can absorb a small disturbance. If the nodes of a transition model are such arrangements and the edges carry rates of moving between them, its gap measures relaxation among arrangements. In a trade graph the nodes are countries and the edges carry goods. Linking these two gaps requires a model for how losing a shipment changes which arrangements governments can reach.

A collapsing spectral gap indicates that basins are becoming isolated—the landscape is fragmenting into disconnected regions. Countries locked into one bloc cannot easily transition to another. The diplomatic "temperature"—the intensity of engagement, the willingness to incur short-term costs for long-term gains—is insufficient to overcome the barriers between configurations.

The Mode-Coupling Theory

Mode-coupling theory, developed by Wolfgang Götze in the 1980s, provides the most successful mathematical description of the glass transition. The theory predicts that as a system approaches the glass transition, its relaxation time—the time required to respond to perturbations—diverges according to a power law:

$$\tau_\alpha \sim (T - T_c)^{-\gamma} \tag{24.2}$$

In the idealized law, approach T_c from above. The symbol $\sim$ describes the scaling: if the distance $T - T_c$ is halved, the characteristic time is multiplied by 2^γ, all else equal. For an illustrative exponent of two, the response takes four times as long. The ideal law extrapolates to a divergence; an observed liquid can continue to rearrange by other processes. Simulations of a model mixture, for example, distinguish a fitted crossover from a literal stop in particle motion (Flenner and Szamel 2005).

Slow adjustment does not establish that every route is closed. It establishes a problem relative to the routes and dynamics in the model.

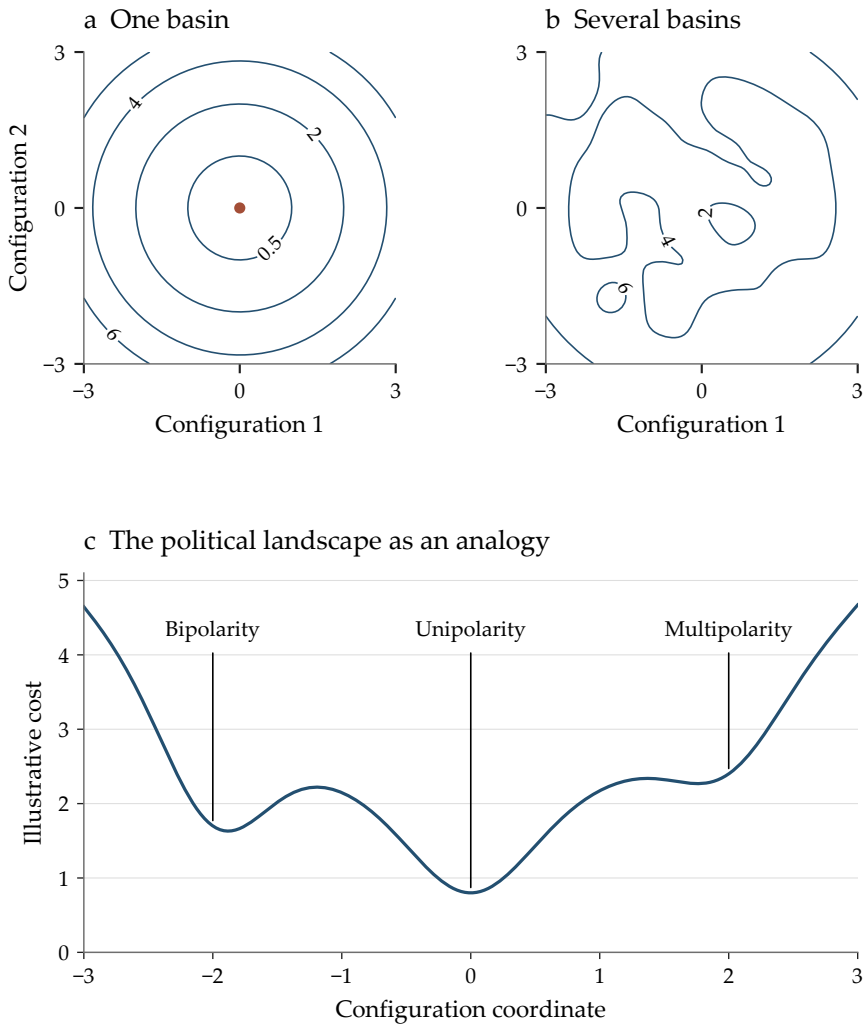


Figure 24.2: Contours make the two constructed surfaces inspectable: one bowl has a single basin, while the rough surface has multiple basins and barriers. The lower cross-section carries the political analogy. Its vertical coordinate is an illustrative cost, not a measured physical energy of international relations.

The political investigation should look for the missing alternative as carefully as it measures the blocked one.

This is critical slowing down in its purest form—the same phenomenon explored in Chapters 2 and 7, now derived from glass physics rather than ecology or spectral theory.

For international systems, the analog replaces temperature with the spectral gap:

$$\tau_{\text{diplomatic}} \sim (\lambda_1 - \lambda_1^*)^{-\gamma} \tag{24.3}$$

The proposed scaling is read from above the fitted threshold, where $\lambda_1 - \lambda_1^* > 0$. In the power law the predicted response time grows without bound as that difference tends to zero. In a diplomatic file one can observe an unresolved negotiation through the last reporting date. The record then supplies a lower bound on how long it lasted. Estimating this law requires timed responses at several gaps, a fitted exponent, and a common definition of recovery.

The divergence of diplomatic response times is not a metaphor—it is a measurable phenomenon. The time from crisis onset to resolution, the duration of negotiations, the lag between provocation and de-escalation: all increase as the spectral gap collapses. The pre-2014 Ukraine crisis took years to unfold; once the system became glassy, the path to February 2022 was compressed into months.

Breaking the Glass

A physical glass can be broken through three mechanisms:

A glass can leave a trapped configuration in more than one way. Heating gives its constituents enough energy to cross barriers and explore other arrangements. Mechanical stress changes the landscape itself, lowering some barriers and raising others. Sufficient force can instead fracture the material, breaking the bonds that prevented wholesale rearrangement. The political analogy must distinguish these mechanisms: enabling adjustment, changing the constraints, and destroying the structure are not interchangeable interventions.

International analogs map directly:

Diplomatic engagement (injecting trust): Intensive diplomacy can lower barriers between configurations, enabling rearrangement. The Camp David Accords, the Good Friday Agreement, the end of apartheid—all required sustained diplomatic "heating" to escape trapped configurations.

Economic measures. Sanctions can change the cost of maintaining an arrangement. Aid, access, or a replacement contract can change the cost of leaving it. The response may include substitution or resistance. A policy should therefore be assessed through the actual constraints it changes, rather than assuming that more pressure necessarily produces more movement.

War (kinetic fracture): Violence that physically destroys existing configurations, clearing space for reconstruction. This is not metaphorical: war breaks infrastructure, eliminates leaders, redraws borders, and forcibly creates new basins of attraction.

When diplomacy leaves commitments unchanged and economic measures fail to create alternatives, the framework directs attention to the risk of violent rupture. A costly arrangement can also persist, or an unmodeled alternative can emerge. The case record must establish which occurred and why.

The practical question is how much time, capacity, and political agreement an alternative requires. If a replacement supplier cannot deliver for a year, it does little for a reserve that lasts six weeks unless another measure bridges the interval. That mismatch is something an analyst can investigate before declaring the arrangement impossible to change.

The analogy has brought us to the cost of rearrangement. Its usefulness now depends on how carefully we translate that cost. The next chapter follows the passage from physical variables to political ones and asks exactly what the translation preserves.

Chapter 25

From Physics to Politics

> All models are wrong, but some are useful.
>
> George Box

A valve does not worry about an election. A government deciding whether to close a pipeline does. Both decisions can change a network, but only one involves promises, constituencies, and an account of the future that the decision maker can revise.

That difference has followed every equation in this book. We can calculate how a network changes without yet explaining why a leader changed it. The task now is to make the passage between the two explicit: which mathematical relationships survive, which quantities need political definitions, and which conclusions depend on evidence from human decisions.

Why should equations describing fluid flow apply to international relations? The question deserves a serious answer, because the history of social science is littered with failed physics envy—attempts to import mathematical formalism that provide the appearance of rigor without the substance of insight.

A physical analogy can fail at a very ordinary point. A pipe carries a measured volume under a specified pressure difference. A trade link records transactions made under prices, contracts, regulations, and expectations. Drawing both as an edge is useful, but it does not give them the same law of motion. The application must specify what causes a change in the quantity that moves.

The test of the transfer is whether the mathematical construction explains a measured feature of the new process and supports a result that survives scrutiny. A resemblance suggests where to look. The application must supply the evidence.

Several structural questions recur across these applications. Each offers a way to organize the investigation, provided its quantities and assumptions are defined in the new setting.

Accounting constraints. A physical shipment must have an origin and destination, with storage, losses, and other inputs included where relevant. Financial accounts also impose balance identities, whose signs depend on the accounting convention. Those identities restrict admissible records. They do not, by themselves, establish a fluid conservation law for political trust.

Network effects: Local changes propagate through the system. A vortex in one region affects flow patterns elsewhere; a financial crisis in one country transmits through trade and investment channels. The 2008 financial crisis demonstrated how interconnection can turn local shocks into global catastrophes.

Regime changes. A system can respond mildly to repeated disturbances and then change abruptly. A reserve can cover several interrupted deliveries before it is exhausted; the next interruption has a different consequence. A thermodynamic phase transition supplies a mathematical motivation, while a political application must identify and test its own mechanism for the change.

System-level effects. Several individually available routes can fail to provide a usable route for the system as a whole. Two exporters may both depend on the same port, or several commitments may require concessions that cannot be delivered together. The network representation helps reveal dependencies that are hard to see in a list of separate transactions.

The mathematical framework captures these structural similarities precisely because it operates at the level of topology and geometry—the study of shapes, connections, and transformations that preserve essential structure. Topology doesn't care whether the manifold represents fluid flow or diplomatic relationships; it cares about connectivity, boundaries, and continuity.

The most robust part of the mapping involves network mathematics. A graph is a graph, whether its edges represent oil pipelines or trade agreements. The Laplacian of a graph has spectral properties that depend only on the connectivity structure, not on what flows through the edges.

Give an engineer and a trade analyst the same undirected graph with the same nonnegative weights. Both obtain the same Laplacian spectrum. If both use the normalized operator and measure the conductance of cuts by weighted volume, Cheeger's bounds apply to both. To convert a slow mode into a time, each must also specify the rate of transfer. A pipe's conductance and a month's invoices supply very different physical clocks.

Where the mapping becomes an analogy—and thus requires more careful justification—is in the interpretation. What does it mean for

international relations that the spectral gap has a particular value? Here we must argue by demonstrated predictive power rather than mathematical necessity.

A government can respond to an assessment by changing a promise, a policy, or the information it releases. The model's subjects can learn from the model. That possibility belongs in the explanation of why a relationship evolved differently from its earlier projection.

A model makes some features legible by leaving others out. That simplification can be useful, but it also creates the possibility that a decisive local condition disappears from view. The observation process should therefore remain inspectable alongside the final score.

Simplification can still explain something important. A model of one blocked supply route may clarify a deadline without pretending to contain every motive of the governments involved. Its value depends on keeping the claim proportionate to what the construction represents.

Precision becomes valuable once the measured quantity and the question agree. Reporting a helicity value to three decimal places cannot compensate for an undefined outcome or an uncertain input. Reporting the calculation together with those uncertainties can make it useful.

Level	Confidence	Examples
Identity	Mathematical certainty	Laplacian properties, spectral decomposition, Cheeger inequality
Strong Analogy	Structural correspondence	Conservation of flows, diffusion dynamics, network bottlenecks
Weak Analogy	Requires empirical validation	Phase transitions, critical slowing down, glass dynamics
Metaphor	Heuristic only	"Vortex of conflict," "diplomatic temperature"

Table 25.1: The hierarchy of physics-to-politics mappings. Confidence decreases from top to bottom. The framework's core claims rest on the top two levels; the lower levels motivate further research.

What the Mapping Preserves

The physics-to-politics mapping preserves specific mathematical relationships:

Physical Concept	Political Analog
Velocity field $\vec{v}$	Trade/investment flows F
Vorticity $\nabla \times \vec{v}$	Rate of change of alignment
Helicity $\vec{v} \cdot (\nabla \times \vec{v})$	Flow-alignment coupling H
Spectral gap λ_1	Network connectivity/flexibility
Glass transition	Diplomatic freezing point
Critical slowing down	Pre-conflict warning signals

Table 25.2: The physics-to-politics concept mapping.

The network result transfers wherever its hypotheses hold. For a symmetric graph with nonnegative weights, Cheeger's inequality connects the normalized gap to a bottleneck. Fluid helicity has a different basis for its persistence, in how an ideal flow transports vortex lines under the stated boundary conditions. The political application asks which commitments and material dependencies are carried forward as a relationship changes.

The same comparison has boundaries. We need to follow what does not survive the mapping as carefully as what does.

Intellectual honesty requires acknowledging what the mapping *cannot* transfer:

Agency. Political commitments can be revised, resisted, or enforced. Leaders may accept costs that another observer considers prohibitive. A model of constraints should make those choices and costs easier to understand, rather than remove them from the explanation.

Reflexivity: In physics, particles don't change behavior because they're being observed. In international relations, publishing a prediction can alter the behavior being predicted. A country warned of approaching conflict may take steps to avert it.

Uniqueness: Physical experiments can be repeated; historical events cannot. Each conflict occurs in a unique context, limiting the statistical power of validation.

Strategic interaction: Fluid particles don't anticipate each other's movements. States explicitly consider how their actions will affect others' responses.

The framework proposes structural diagnostics of dependence and adjustment, then evaluates their association with conflict. A claim about stability is conditional on a specified evolution law. A claim about warning performance depends on a defined population, target, and horizon. Neither makes the political outcome certain.

This limitation is also a feature. A framework that predicted conflict deterministically would be either useless (if nothing can be done) or self-refuting (if action prevents the prediction). The structural approach provides warning and characterization while preserving space for human agency.

The Unreasonable Effectiveness Question

Eugene Wigner's famous 1960 essay asked about "the unreasonable effectiveness of mathematics in the natural sciences." We might ask a parallel question: why does physics-derived mathematics apply to social systems at all?

Several conjectures deserve consideration:

Conservation constraints. A physical shipment must satisfy an accounting balance, with sources, sinks, storage, and losses represented. Political trust has no automatically equivalent conservation law. Shared mathematical tools become justified through the particular balance or symmetry established in each application.

Network universality: The mathematics of networks—spectral theory, graph Laplacians, random walks—applies to any system that can be represented as nodes and weighted edges. International trade networks and fluid flow networks share this abstract structure.

Critical phenomena. Different materials can share a pattern of collective change even when their constituents differ. The political hypothesis is that a loss of available settlements can produce an abrupt change under gradually increasing pressure. It takes a record of those settlements to test that correspondence.

Selection effects: Perhaps we notice the cases where physical mathematics applies and ignore the cases where it doesn't. The apparent effectiveness may be survivorship bias rather than deep truth.

The 47-event mass-gap study gives $F_1 = 0.73$ and a mean lead of 72 days (Appendix E). The separate historical roster's 18 detections average 127 days (Appendix C). To test selection effects, keep the quiet cases and the misses alongside the successes.

My newer Helicity work follows three decisions through the measurement itself. At a bank counter, change the currency unit. At a border, change which party moves first. At a fuel ministry, ask what can still be delivered when one manifest is missing.

Chapter 26

What Changes When the Ruler Changes

A ministry reports a sharp increase in the value of imports. Across the table, another analyst says the country's dependence has barely changed. They can both be reading the same shipments. One is counting the money paid; the other is counting the fuel delivered. A price increase has entered the first account. An exchange-rate movement may have entered it too. Before deciding what the relationship is doing, we have to decide what the number is doing.

This problem runs through the book. A pipeline carries a physical quantity, a contract assigns a price, and a database expresses that price in a currency. Each description can answer a useful question. Confusing their changes can manufacture a warning or conceal one. A country that pays twice as much for the same fuel has acquired a financial burden, but it has not thereby doubled the physical volume it must replace.

My newer Helicity work makes this distinction an explicit part of the instrument. It asks which changes belong to the relationship and which belong to the coordinates in which we describe it. The mathematics of gauge transformations gives one precise answer. The word sounds remote from a customs ledger. Its first example is as ordinary as changing the denomination printed on a banknote.

A journey that returns to its starting point

Consider three currencies, called A, B, and C. These are constructed quotes, chosen so that the arithmetic can be checked without a market terminal. One unit of A buys two of B. One of B buys three of C. One of A buys five of C directly. Following A to B to C gives six units of C; taking the direct route gives five. If the reverse quote is the exact reciprocal, the round trip A to B to C to A returns

$$2 \times 3 \times \frac{1}{5} = \frac{6}{5}. \tag{26.1}$$

The two routes disagree. This little calculation assumes simultaneous, frictionless quotes with reciprocal conversions. It describes a consistency defect in the constructed quotation system; actual execution would require spreads, timing, and costs.

Now let C replace its unit with one one-hundredth as large. The same two offers become 300 new units of C per B and 500 per A. The printed numbers have changed dramatically. The round trip has not:

$$2 \times 300 \times \frac{1}{500} = \frac{6}{5}. \tag{26.2}$$

The disagreement survives because the conversion into C and the conversion out of C cancel the change of unit. A measurement built from the round trip can therefore distinguish something that survives a relabeling from the relabeling itself.

Take logarithms, so that multiplication along a route becomes addition. If S_{ij} is the number of units of currency j obtained for one unit of i, write $a_{ij} = \log S_{ij}$. The circulation around our triangle is

$$h = a_{AB} + a_{BC} - a_{AC} = \log(6/5). \tag{26.3}$$

Here h names this loop sum, not the book's helicity index. The minus sign means that the last edge is traversed against the direction in which its quote was stored. The result remembers the disagreement between routes while forgetting the denomination of the intermediate currency.

For a larger network, collect the edge readings in a vector a. The incidence matrix B has a column for each directed edge: -1 at its source, $+1$ at its destination, and zero elsewhere. If currency i now uses e^{ψ_i} new units per old unit, the quotation vector becomes

$$a' = a + B^{T}\psi. \tag{26.4}$$

This is the gauge transformation. It records a change of local units at the nodes. A signed round trip is represented by a vector γ satisfying $B\gamma = 0$: at every node, arrivals and departures balance. Consequently,

$$\gamma^{T}a' = \gamma^{T}a + (B\gamma)^{T}\psi = \gamma^{T}a. \tag{26.5}$$

The cancellation we checked with three numbers now holds for every closed route in the network. Nothing about that identity depends on whether tomorrow's exchange rate rises or falls.

Separating the route from the residue

The same graph supports a second useful calculation. Fit one number ϕ_i to each currency so that differences between those numbers explain as much of the edge data as possible. In the ordinary Euclidean edge norm, solve

$$\widehat{\phi} \in \operatorname*{argmin}_{\phi} \|a - B^T\phi\|_2^2, \qquad c = a - B^T\widehat{\phi}. \tag{26.6}$$

The residual c satisfies $Bc = 0$. Thus the edge readings split into a gradient part, explained by node potentials, and an orthogonal cycle part. This is the graph Hodge decomposition used here; a richer complex with specified faces permits a further distinction between local circulation and harmonic cycles Jiang et al. 2011.

The node potentials are determined only up to an additive constant on a connected graph. That freedom is harmless: every edge uses a difference, so the constant disappears. The residual is unique. Adding $B^T\psi$ changes the fitted gradient but leaves c unchanged when the graph and the fitting norm are fixed.

Return to the bank counter with the three quotes. If each currency had one perfectly consistent price on a common ruler, going through B and going directly to C would give the same answer. Our quotes give six and five. No choice of three node prices can fit all three offers. How does the least-squares fit distribute that disagreement? Store the edges in the order A–B, B–C, A–C, so a round trip reads the first two forward and the last backward. Its log discrepancy is $h = \log(6/5)$. With equal weights the residuals on those three stored edges are $(h/3, h/3, -h/3)$, about $(0.0608, 0.0608, -0.0608)$ in log units. Subtract them from the observations: the fitted round trip now has sum $h - (h/3 + h/3 + h/3) = 0$. That is the split one can inspect at the counter. It identifies the part no set of consistent node prices can explain; it does not identify which teller, clock, or quote caused the disagreement.

There is an important interpretive detail. The gradient component does not mean “fake movement.” A genuine change in the relative value of currencies can live in that component. The split answers a particular structural question: which edge movements can be reconciled through node values, and which remain as disagreement around routes? It does not divide the economy into reality and illusion.

The calculation also forces us to name the norm. Giving every quotation equal weight is a choice. A weighted fit can reflect unequal measurement precision, but its weights and transformation rule must be

specified. The geometry becomes useful because we can inspect those choices, rather than smuggle them into a graph that appears neutral.

I tested this construction on the retained Helicity basket of fourteen currency crosses connecting seven currencies. The common record contains 4,318 sessions from January 1, 2010, through August 7, 2026, and therefore 4,317 daily changes. The incidence matrix has rank six; the cycle space has dimension eight. Those dimensions come from the quotation graph before a price is examined Davis 2026e.

In the stored quote coordinates, the gradient accounts for 99.15% of the squared norm of the daily log changes. That number needs a comparator. An isotropic edge vector has an expected gradient share of 6/14, or 42.86%, simply because the projection occupies six of fourteen dimensions. Independently shuffling each quotation's daily changes preserves each series' marginal values while breaking their contemporaneous relationships. In that control, the gradient share falls to 43.18%.

The cyclic root-mean-square residual rises from 5.64 basis points in the recorded changes to 45.98 in the shuffled control. Here a basis point means 10^{-4} in log-rate units. The coordinated record and the independently shuffled record behave differently under the same graph calculation. The effect is therefore not explained by the projection's dimension alone.

The percentage itself is a description in the stated coordinates. Its denominator contains the gradient, which a time-varying change of units can alter. The cyclic residual has the invariance proved above. Reporting both, with the control alongside them, keeps the empirical observation attached to the precise mathematical claim.

Three reports with different levels of precision

Return to the three bank counters. The first two quotations come from synchronized feeds. The third was copied from a slower screen. Treating all three as equally precise would distribute the discrepancy equally, even though the observation process gives us a reason to trust them differently. The next question is therefore practical: how should the fit change when the observations have different uncertainties?

Keep the same incidence matrix and the same log quotations. Let W be a positive-definite matrix defining the squared discrepancy we intend to minimize. For independent errors with known variances, a common statistical choice is a diagonal matrix of inverse variances. A more general

covariance model permits off-diagonal entries. The weighted calculation is

$$\widehat{\phi}_W \in \operatorname*{argmin}_{\phi} (a - B^T\phi)^T W (a - B^T\phi), \qquad c_W = a - B^T\widehat{\phi}_W. \tag{26.7}$$

The fit now penalizes an error in a precise observation more heavily than the same error in an imprecise observation. This is weighted least squares, not a new rule for deciding which report is truthful. The interpretation of W has to come from the observation model. High trading volume, frequent reporting, and small measurement variance are different properties; one cannot substitute for another without an argument (Jiang et al. 2011; Boyd and Vandenberghe 2004).

Differentiate the objective with respect to the node potentials. The condition at a minimum is

$$BW(a - B^T\widehat{\phi}_W) = 0, \qquad BWB^T\widehat{\phi}_W = BWa. \tag{26.8}$$

The matrix BWB^T has the familiar freedom to add a constant to all node potentials. Fixing one potential, or requiring their sum to be zero on a connected graph, selects a representative. The edge fit does not depend on that selection.

Notice the extra W in the first equation. The weighted residual satisfies $BWc_W = 0$. It need not satisfy $Bc_W = 0$. The vector Wc_W belongs to the ordinary cycle space; c_W is orthogonal to gradients in the chosen weighted inner product. Dropping the weight from this statement would quietly change the mathematics. A printed diagram of arrows cannot settle which inner product was used.

The triangle makes the distinction visible. Write $\gamma = (1, 1, -1)^T$, so the observed loop discrepancy is $h = \gamma^T a = \log(6/5)$. Every possible residual must retain that discrepancy, because subtracting a gradient cannot alter a closed-loop sum. Among those residuals, the least expensive one under W is

$$c_W = \frac{hW^{-1}\gamma}{\gamma^T W^{-1}\gamma}. \tag{26.9}$$

For $W = \operatorname{diag}(4, 4, 1)$, the first two observations receive four times the precision weight of the third. Their residuals become

$$c_W = (h/6, h/6, -2h/3)^T.$$

Numerically these are about 0.03039, 0.03039, and -0.12155 in log units. The less precise third quotation absorbs more of the correction. Nev-

ertheless, the signed loop sum remains h: $h/6 + h/6 + 2h/3 = h$. The calculation has changed the allocation of the disagreement, not erased it.

There is a useful statistical reading if the declared weights really are inverse error variances. Suppose the first two error standard deviations are 0.01 log units and the third is 0.02. Then $W = \mathrm{diag}(10{,}000, 10{,}000, 2{,}500)$ has the same relative weights. With independent errors the variance of the loop reading is

$$\mathrm{Var}(h) = 0.01^2 + 0.01^2 + 0.02^2 = 0.0006.$$

The constructed discrepancy is about 7.44 such standard deviations from zero. That number is a standardized discrepancy under the supplied error model. It becomes a probability statement only after an appropriate distributional model and a testing protocol are supplied. If quotations are stale, systematically biased, or selected after looking for large discrepancies, a familiar threshold alone does not rescue the interpretation.

The example teaches a distinction that matters well beyond currency prices. An analyst may know that one customs estimate is less precise than another. A weighted fit can incorporate that knowledge. It still cannot infer from a large residual that a government is deceiving its partners. The residual locates inconsistency in the specified representation. An explanation requires evidence about clocks, units, collection practices, and the relationship itself.

A missing route can make the contradiction disappear

Suppose the direct A–C quotation is missing. Only A–B and B–C remain. The graph is now a chain. Choose one node price arbitrarily; the other two can be assigned to fit both remaining observations exactly. The cycle residual is zero, however imprecise those two reports may be.

Nothing has established that the market is more consistent than before. The observation graph has lost the route that allowed the inconsistency to be tested. With no closed path, there is no comparison between different ways of reaching the same endpoint.

For a graph with n nodes, m stored edges, and k connected components, the incidence matrix has rank $n - k$. The cycle space therefore has dimension

$$\dim \ker B = m - n + k.$$

A connected tree has $m = n - 1$ and no cycle directions. Adding a new edge inside an existing component adds one such direction. Adding an edge that merely joins two previously disconnected components reduces

k by one while increasing m by one, so it adds no cycle direction. The distinction can be read directly from the formula.

At the bank counter, the missing direct quotation is valuable because it completes a loop. A fourth observation that simply introduces a previously isolated currency need not provide the same consistency test. The number of additional records is therefore a poor description of how much new structural information has arrived.

The same lesson applies to the historical studies' evidence. Several reports copied from one upstream source do not necessarily create independent checks. A record of delivery can challenge a record of dispatch because it observes a different stage of the process. A second article repeating the dispatch announcement may add no new stage at all. Before celebrating a smaller residual, ask whether the contradictory observation was repaired or merely removed from the graph.

There is also a sensitivity worth preserving in the archive. Report the residual on a fixed common graph, then report what changes when the graph expands. Otherwise a period with sparse records can appear unusually coherent simply because it contains fewer opportunities to detect disagreement. The dimensional control in the fourteen-cross experiment addresses one part of this problem; the observation register addresses another. Neither is replaced by presenting a smooth time series.

What uncertainty in the fit actually describes

Assume, for this calculation, that a fixed connected graph generates observations by

$$a = B^T \phi_* + \epsilon, \qquad E[\epsilon] = 0, \qquad \operatorname{Cov}(\epsilon) = \Sigma.$$

This is a measurement model with a consistent underlying node potential ϕ_*. Its residual is noise by construction. In an application with a genuine cyclic component, that component must be added to the model rather than defined out of existence.

Take $W = \Sigma^{-1}$ and let $L_W = BWB^T$. On the subspace where the node potentials sum to zero, the fitted potential and residual can be written

$$\widehat{\phi}_W = L_W^{+} BWa, \tag{26.10}$$

$$c_W = Ma, \qquad M = I - B^T L_W^{+} BW. \tag{26.11}$$

The superscript + denotes the Moore–Penrose pseudoinverse. Because the fit is a linear map of the observations, its residual covariance is

$$\operatorname{Cov}(c_W) = M\Sigma M^T. \tag{26.12}$$

Even when the original observation errors are independent, the residuals generally are not. Fitting shared node prices couples the corrections. On the triangle, a residual with only one free cycle coordinate cannot supply three independent pieces of statistical evidence. Counting its three nonzero entries as three separate confirmations would count the same inconsistency repeatedly.

A redundant archive creates a related trap. Suppose two data providers reproduce the same underlying quotation with a shared error. Treating their reports as independent doubles their apparent support and can narrow an uncertainty estimate without acquiring new information. A covariance model, a source-lineage record, or a deliberate deduplication rule must address that dependence. Merely adding another row cannot do so.

Uncertainty about the graph is a separate layer. The formula $M\Sigma M^T$ conditions on the chosen edges, weights, and fitting rule. It does not include an omitted route, a mistaken actor identifier, or a changing rule for what counts as a connection. One useful investigation repeats the fit across explicitly stated graph alternatives and reports the conclusions that survive. That is sensitivity to model construction; it should not be mislabeled sampling variation.

Invariance as a practical test

The weighted fit retains the original gauge property when B and W stay fixed. Replace a by $a + B^T\psi$ and shift a candidate potential by ψ. Its edge residual is unchanged. Every candidate in the original optimization therefore has a candidate with exactly the same objective in the transformed problem. The optimal residual is the same.

This argument also identifies a failure mode. If the software recomputes its weights from the raw numerical magnitudes after a unit change, W may change as well. It is then solving a different optimization problem. A result can move even though the analyst intended only to change denominations.

A reproducible implementation should perform this test deliberately: fit the original data; apply a declared node-unit transformation; refit with the corresponding observation model; compare residuals to a tolerance tied to numerical precision. It should also reverse a stored edge and

reverse its log reading. A pure orientation change is bookkeeping. A change in the inferred structural discrepancy is a reason to inspect the implementation.

These checks do not establish a political theory. They establish that a measurement obeys its own stated transformation rules. That is a necessary kind of modesty in a powerful instrument. We want a changed warning to send us toward a changed relationship, a changed observation, or an explicit modeling choice. We should not have to discover afterward that it began with a database administrator moving a decimal point.

Returning to the border

Currencies provide an unusually clean test because their conversion rule is explicit. Diplomatic trust has no corresponding law saying that one country's confidence can be exchanged for another's at a quoted rate. We cannot place trust scores in a currency formula and inherit its invariance. We can inherit the question it makes unavoidable.

Suppose a warning becomes stronger after a data provider changes the base year of an index. Did the underlying relationship change? Suppose the same physical imports are converted into a different reporting currency. Does the chosen normalization preserve the intended measure? Suppose a bilateral report disagrees with the partner's report. Could freight valuation, reporting dates, or missing classifications explain the difference? A residual identifies a place to investigate. Its cause still has to be established.

For the conflict instrument, this suggests a concrete discipline. Preserve the original quantities and their units. State how each becomes a coordinate. Test changes that should leave the substantive answer unchanged. When they do change it, trace the dependence rather than interpreting every movement as a new geopolitical fact. If a transformed coordinate $x' = Sx$ describes the same measurement, a quadratic metric must transform as $G' = S^{-T} G S^{-1}$ to preserve the same squared displacement:

$$(\Delta x')^T G' \Delta x' = \Delta x^T G \Delta x. \tag{26.13}$$

For invertible S, that is an exact identity. Keeping an untransformed Euclidean ruler after rescaling the coordinates asks a different question.

For a single cargo coordinate, a change of one tonne becomes 1,000 kilograms. Squaring the printed number makes it appear a million times larger. Express the same squared-distance rule in kilograms, however, and its coefficient becomes one millionth of the coefficient in tonnes: $1{,}000^2/1{,}000{,}000 = 1$. That is what the two inverse factors in

the matrix formula accomplish. A ministry can revise the unit in its database without moving a ship. A warning meant to measure physical dependence ought to respect that difference.

The historical argument remains about the costs of changing a real relationship. The newer work strengthens the measurement that supports it. An expensive shipment, a scarce shipment, and a differently denominated shipment may demand different responses. An instrument that treats them as interchangeable can send the decision-maker toward the wrong door.

Even after the ruler is settled, another ambiguity remains. Two reports can describe the same final quantities and omit the feature that made one negotiation safer than another: which side moved first.

Chapter 27

The Order of the Moves

The inspectors arrive before the convoy crosses the border. In another arrangement, the convoy crosses first and the inspectors arrive afterward. By evening, both records say the same thing: the shipment was delivered, and the inspection was completed. During the day, the parties faced different risks.

In the first arrangement, delivery depended on verification. In the second, the recipient held the goods while verification remained outstanding. The order changed who had to trust whom, who could still withhold something, and which failure would be costly. A ledger that retains only the two completed totals erases that distinction.

This is a constructed negotiation, but its problem is familiar from the historical cases. A withdrawal before a guarantee is implemented differs from a withdrawal afterward. A sanction lifted before access is restored differs from one lifted after restoration. The endpoint matters, and so does the route by which people are asked to reach it.

The recent Helicity work on event order gives this problem a small, inspectable mathematical form. We begin with two actions and one square.

The square between two agreements

Let x record inspection completion and y record delivery completion, each measured from zero to one. These are declared coordinates for the example, not a claim that every negotiation admits a universal progress score. Inspection first follows

$$(0,0) \longrightarrow (1,0) \longrightarrow (1,1).$$

Delivery first follows

$$(0,0) \longrightarrow (0,1) \longrightarrow (1,1).$$

The final displacement is $(1,1)$ in both cases. A calculation that sees only the initial and final states cannot distinguish them.

For a piecewise linear path translated to start at the origin, define its signed area by

$$A_{xy} = \frac{1}{2}\int (x\,dy - y\,dx) = \frac{1}{2}\sum_{k}(x_k y_{k+1} - y_k x_{k+1}). \tag{27.1}$$

The formula closes the route with a straight segment back to the origin; that closing segment contributes zero to the integral. Inspection first gives $A_{xy} = +1/2$. Delivery first gives $A_{xy} = -1/2$. Simultaneous linear completion, along the square's diagonal, gives zero. The endpoint is identical, but the area records a difference in sequence.

Follow the inspection-first convoy one leg at a time. While the inspector works, $y = 0$ and $dy = 0$, so neither term contributes. The convoy then moves with inspection complete: $x = 1$, $dx = 0$, and y rises from zero to one. The surviving term is $\frac{1}{2}\int_0^1 1\,dy = +1/2$. Deliver first and the last leg is inspection with the goods already across the border: $y = 1$, and the surviving term is $-\frac{1}{2}\int_0^1 1\,dx = -1/2$. The sign remembers which party was exposed while the other obligation was outstanding. It cannot tell us how long the wait lasted; that still requires the timestamps.

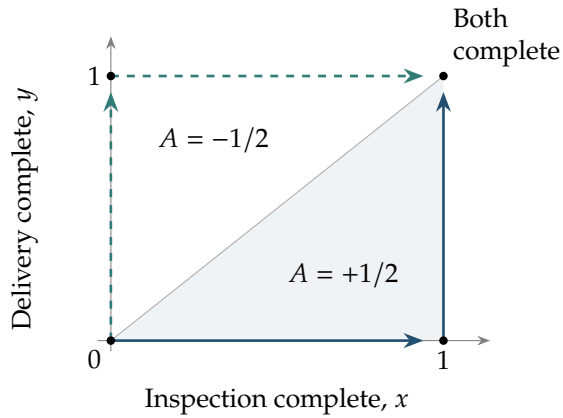


Figure 27.1: Two agreements reach the same completed state. Solid arrows show inspection before delivery; dashed arrows show delivery before inspection. The signed areas are $+1/2$ and $-1/2$ in the declared unit-square coordinates. These are constructed sequences, not measurements of a particular treaty.

The sign is a convention tied to the axis order. Exchanging x and y reverses it. The size depends on the scales: measuring inspection in percentages instead of fractions multiplies the area by one hundred if

delivery's scale is unchanged. What matters is the explicit path and its declared coordinates, not a positive sign dressed up as a verdict about peace.

Signed area is one second-order feature of a path. It belongs to the wider mathematics of iterated integrals and rough paths, where endpoint increments alone are insufficient to describe how a driving signal acts on a system Lyons 1998. No complete rough-path theory is needed to inspect this square. Its practical lesson is already visible: two completed actions do not tell us the sequence of exposure that occurred between them.

A clock can hide the order

Consider the records arriving at an analyst's desk. A port authority logs a clearance at 09:02. A shipping agent reports a departure at 09:04, but that report reaches the analyst at 09:17. A daily table puts both actions on the same date. An hourly table may place both in the same interval. A reconstruction can accidentally supply the 09:04 fact to a decision made at 09:10, before it was available.

These are three different issues: event order, the grid used to summarize it, and the time at which the information became usable. Preserving one does not automatically preserve the others.

For a fixed geometric path, a common strictly increasing change of time leaves the signed-area integral unchanged. We traverse the same ordered segments at different speeds. Removing observations is different: it changes the polygonal path. Replacing two successive moves with a diagonal shortcut can remove the very area we wanted to measure. Missing timestamps cannot be repaired merely by calling a method geometric.

The Helicity implementation constructs an asynchronous two-series path on the union of their observation times. At each event it uses the most recent available value from the other series. This makes the interpolation rule inspectable and avoids looking forward to an unseen reading. It also creates an obligation: simultaneous events, stale observations, and the distinction between event time and receipt time must be handled deliberately Davis 2026c.

The saved synthetic experiment shifts one underlying sequence by $-20, -8, -3, 0, 3, 8$, and 20 samples, then samples the two channels at different rates. Its raw signed areas increase from approximately -0.199 to $+0.204$ in the constructed path's squared-coordinate units. The six nonzero shifts have the expected signs. At zero shift, the result is approximately $+0.042$, not exactly zero. Asynchronous sampling is still

part of the measurement; the control prevents a small nonzero reading from being treated as automatic evidence of leadership.

The implemented normalization divides area by the geometric mean of the two sums of squared increments:

$$\widetilde{A}_{xy} = \frac{A_{xy}}{\sqrt{\sum_k (\Delta x_k)^2 \; \sum_k (\Delta y_k)^2}}. \tag{27.2}$$

It is defined only when both increment energies are positive. This makes a particular dimensionless statistic; it does not confine the result to $[-1, 1]$. Stored market examples also change under thinning. The empirical question is how much useful ordering survives a specified observation process, not whether every deletion leaves an invariant behind.

For diplomacy, the corresponding research design is straightforward to state and demanding to execute. Retain the sequence of announcements, implementation, delivery, verification, and reversal. Distinguish when an action happened from when the analyst could know it. Define the coordinates before examining which signs precede escalation. Then compare the ordering information with a baseline supplied the same available facts. A recurring lead may reflect a common cause, a reporting delay, or an institutional procedure. Area alone does not settle which explanation is right.

Circulation without return

Order also matters at the scale of repeated transitions. Imagine recording three states of a relationship: routine exchange, imposed restriction, and negotiated waiver. A monthly count might show the same occupancy of each state in two records. Yet one could repeatedly travel exchange to restriction to waiver to exchange, while the other moves back and forth without a preferred circuit.

Think of following a file from one month’s desk to the next. In a first-order Markov model the probabilities of its next destination depend only on its present desk. A fixed transition matrix gives the same routing rule each month. For the stationary calculation below, the distribution π also satisfies $\pi P = \pi$, and the file starts with that distribution. A newly opened desk can use fixed rules from day one while its traffic is still settling.

Let P_{ij} be a transition probability and π_i the stationary frequency of state i. Then

$$F_{ij} = \pi_i P_{ij}, \qquad J_{ij} = F_{ij} - F_{ji} \tag{27.3}$$

describe the directed transition flux and its net current. Detailed balance requires $F_{ij} = F_{ji}$ for every pair. Under positive reverse fluxes on the transitions being compared, a measure of time-direction asymmetry is

$$\sigma = \frac{1}{2} \sum_{i,j} (F_{ij} - F_{ji}) \log \frac{F_{ij}}{F_{ji}} \geq 0. \tag{27.4}$$

Each pair contributes a nonnegative amount because the difference and the logarithmic ratio have the same sign. The expression vanishes at detailed balance. With discrete transitions, its units are nats per step; a physical entropy-production interpretation requires an additional physical model.

Take a three-state ring with probability 0.6 of moving forward, 0.2 backward, and 0.2 of remaining in place. The stationary frequencies are all 1/3. Substitution gives

$$\sigma = 0.4 \log 3 \simeq 0.4394.$$

Changing the forward and backward probabilities to 0.4 and 0.4 makes it zero. Reversing the first ring's preferred direction reverses its current but leaves its nonnegative asymmetry unchanged. Thus the current tells us which way the circuit runs; σ measures how different the forward and reverse traffic are. Neither quantity names the cause.

Imagine keeping a tally through 150 steps of that ring in stationarity. The expected number of starts in each state is 50. Across one pair, expect 30 forward passages and 10 reverse passages: a net 20, or 20/150 per step, at a forward-to-reverse ratio of three. The three pairs each contribute $(20/150) \log 3$ to the total. That is where $0.4 \log 3$ comes from. Simply counting how many months were spent under a waiver would miss the direction of the traffic. These are expected tallies under the constructed chain, rather than observations from a particular sanctions regime.

This distinction informs Helicity's detailed-balance research, including its admission tests for a transition model. In the retained E22 confirmation record, the first scored stream admitted none of its twenty-eight sessions. The program therefore returned a refusal rather than a detailed-balance verdict Davis 2026d. An equation can be perfectly sound while the available record does not support the model needed to interpret its output.

That is a useful discipline for this book. A sequence of sanctions and concessions may possess a directional pattern. It may also be nonstationary, dependent on a history longer than our state labels retain,

or too poorly observed for a stable estimate. Establishing which situation we face is part of the geometric investigation.

Return to the convoy waiting at the border. We now know why its final destination is insufficient, why the timestamps matter, and why a pattern in repeated passages needs an observation model. The next problem begins when some of those observations never arrive. We still have to decide what the remaining evidence permits us to say.

Chapter 28

What the Observations Leave Open

Two routes can supply a country with sixty units of fuel per day. A reliable report gives the combined deliverable quantity, but the allocation between the routes is missing. Route A can carry at most forty; route B at most fifty. A minister asks how much supply would remain if B became unavailable.

One answer would fill the missing allocation with a plausible number. Another would calculate every allocation still permitted by the reports. First ask what A's already scheduled delivery would supply immediately, before cargo can be reassigned. Let x_A and x_B be those daily allocations; assume that losing B leaves A's scheduled load deliverable. How much extra A could take later is a second question about dispatch, spare capacity, and time. For the current schedule,

$$x_A + x_B = 60, \qquad 0 \le x_A \le 40, \qquad 0 \le x_B \le 50. \tag{28.1}$$

Route A must carry at least ten and can carry as much as forty. Both extremes are possible: $(x_A, x_B) = (10, 50)$ and $(40, 20)$ satisfy every stated condition. Reporting thirty would hide the fact that the observations also permit ten.

Those two explicit allocations are *witnesses*. A witness is a concrete example that demonstrates a mathematical claim and can be checked against its conditions. Here the first proves that the lower bound is attainable; the second proves the same for the upper bound. They let the minister inspect what has to be true at each extreme.

A new verified report narrows B's current allocation to between thirty-five and forty. A must now lie between twenty and twenty-five. The interval shrinks because the new observation excludes some previously possible allocations. No probability distribution was required to make that inference. We have learned a constraint on the present state.

This small supply problem captures a major direction of my recent Helicity and GIGI work: retaining the range of explanations that evidence

supports, instead of hiding the missing information inside one attractive answer.

An answer with its alternatives attached

Write the family of states compatible with observations D and a declared model M as $\mathcal{F}(D; M)$. If the question of interest is $q(x)$, its conditional bounds are

$$q_{\min} = \inf_{x \in \mathcal{F}(D;M)} q(x), \qquad q_{\max} = \sup_{x \in \mathcal{F}(D;M)} q(x). \tag{28.2}$$

In the bounded fuel example the extrema are attained, so the infimum and supremum are a minimum and maximum. More generally, an extremizing state need not exist, which is why the broader notation is useful.

The model is part of the expression for a reason. If another supply route exists, if the capacity limits are wrong, or if the loss of B also interrupts A, we have changed the feasible family. Tight arithmetic inside the wrong family can be confidently misleading. A narrow answer may reflect good information, an overly restrictive model, or both.

An empty family means that the observations and assumptions cannot all hold together. It does not by itself identify the false report. Perhaps a capacity changed, two records refer to different days, or a unit was misread. The correct next task is reconciliation. The disagreement has acquired a location in the calculation, but not yet an author or a motive.

In the recent PARALLAX prototype, this logic is implemented for missing contemporaneous option quotations. The application differs from diplomacy; the inference discipline is closely related. Neighboring quotations and an underlying-price interval constrain nonnegative pricing weights on declared terminal states. A linear program finds the lowest and highest compatible value of a withheld contract. The target quotation is excluded before the family is built Davis 2026f.

Here is the mechanism in the study's finite-state example. A call with strike 110 pays nothing unless the terminal price exceeds 110; if the price finishes at 120 it pays 10. Among the five allowed prices, 80, 90, 100, 110, and 120, only the last contributes. With the example's zero discount rate, a pricing weight of 0.05 on that last state gives a call value of 0.50; a weight of 0.15 gives 1.50. The research record supplies two complete sets of weights that agree on a forward price of 100 and a 100-strike call value of 4 yet yield those different 110-strike values. A pricing weight belongs to this valuation model; it is not a prediction of how often the

stock will finish at 120. The useful question is which additional quote can distinguish the two explanations (Davis 2026a).

One saved synthetic case makes the role of assumptions unusually clear. The same retained quotations place a missing value between 0.47 and 0.53 on a five-state support. Allowing an additional state at 107 expands the upper bound to about 1.20615. A broader support raises it to about 1.56273. A candidate value of 1.50 changes from incompatible to compatible because the family changed, not because new evidence about that target arrived Davis 2026a.

The prototype then asks which additional observation would best distinguish the extreme compatible explanations. In its saved replay, acquiring the nominated neighboring quotation narrows the overall envelope from approximately [0.4700, 1.5627] to [0.4700, 1.2085], a 32.4% reduction in width. This is a constructed current-state recovery result. It establishes neither future-return accuracy nor a new conflict-prediction score.

The practical connection is direct. In the fuel example, another report of the same combined total would add little. A reliable report separating the two routes could change the decision. In a diplomatic case, another article repeating the same source may add less than a verified delivery, a reserve estimate, or confirmation that a promised exemption actually took effect. The next observation should be chosen for the alternatives it can separate.

A whole family inside a disk

GIGI carries the same principle into bounded analytic completion. Here the shared structure is stronger than a collection of independent intervals. Values at different locations must belong to one complex differentiable function on a specified domain. That requirement links the observations.

To give the construction a physical starting point, take a measured quantity such as a wine sample's alcohol content or a power plant's output. Sort all recorded values into thirty-two declared bins. If p_j is the fraction in bin j, define the generating polynomial

$$G(z) = \sum_{j=0}^{31} p_j z^j, \qquad p_j \geq 0, \qquad \sum_j p_j = 1. \tag{28.3}$$

For $|z| \leq 1$, the triangle inequality gives $|G(z)| \leq 1$. We now have a bounded analytic object because we constructed one from the histogram. The original physical process has not been declared holomorphic. Bin-

ning is an explicit, lossy representation, and its edges and counts belong with the evidence.

Take a small invented batch of 100 specimens instead of the full study. Suppose 50 fall in bin zero, 30 in bin one, and 20 in bin two, with the other bins empty. Then $G(z) = 0.5 + 0.3z + 0.2z^2$. A query at $z = 1/2$ gives $0.5 + 0.15 + 0.05 = 0.70$: bins farther along the scale receive successively smaller weights. A query at $z = i$ gives $0.3 + 0.3i$, since $i^2 = -1$. These are two different weighted readings of the same hundred specimens. Neither number is a new measurement of alcohol content. When a reading is withheld, the task is to constrain that mathematical reading from the ones supplied.

Supply selected evaluations $G(z_i) = w_i$ and withhold another evaluation. Among all holomorphic functions f on the open unit disk satisfying $|f| \le 1$ and $f(z_i) = w_i$, what values remain possible at the withheld point? "Holomorphic" means complex differentiable throughout that open domain. Its real and imaginary components cannot vary independently.

The classical Nevanlinna–Pick theorem supplies the compatibility test. For distinct interior observation points, the family is nonempty exactly when the Hermitian matrix

$$P_{ij} = \frac{1 - w_i\overline{w_j}}{1 - z_i\overline{z_j}} \tag{28.4}$$

is positive semidefinite Agler and McCarthy 2002. The bar denotes complex conjugation. Positive semidefinite means $v^*Pv \ge 0$ for every complex vector v. The theorem is classical; the research contribution here is an implementation that keeps this mathematical contract connected to changing records, independent checks, and explicit numerical limits Davis 2026b.

The simplest case can be seen without solving a matrix problem. If $f(0) = 0$ and $|f| \le 1$ on the unit disk, Schwarz's lemma gives

$$|f(1/2)| \le 1/2.$$

Every value in that half-radius disk is possible: for any $|w| \le 1/2$, the function $f(z) = 2wz$ is a witness with $f(1/2) = w$. One observation has reduced the possibilities, but it has not selected a unique answer. Adding the compatible observation $f(1/2) = 1/4$ fixes that value and further constrains others. Claiming $f(1/2) = 0.8$ would conflict with the family already declared.

The matrix test catches the same conflict. Pair the reading $f(0) = 0$ with a proposed real reading $f(1/2) = w$. The diagonal entries of P are 1

and $(1 - w^2)/(1 - 1/4)$; both off-diagonal entries are one. At $w = 1/4$ the second diagonal is 1.25 and the determinant is 0.25, so the pair passes. At $w = 0.8$ the second diagonal is 0.48 and the determinant is -0.52. One negative determinant is enough to rule out this two-by-two Hermitian matrix as positive semidefinite. Both readings, taken separately, are less than one in magnitude. The incompatibility appears only when we require a single bounded function to produce them together.

For several observations, GIGI returns a compatible representative and the disk of possible values at each query. The disks answer one question at a time. Suppose $f(0) = 0$. At each of $1/2$ and $-1/2$ the individual disk permits $+1/2$. Can both readings occur together? Reaching $f(1/2) = 1/2$ attains equality in Schwarz's lemma, forcing $f(z) = z$. The other reading must then be $-1/2$. This is the same bookkeeping the minister faced with fuel. Two individually permitted deliveries can still be incompatible with one shared schedule.

The recorded study applies this rule to all seventeen measured numeric fields in two public datasets: 6,497 wine records and 9,568 power-plant records. Seven complex readings per field are supplied; twenty-four others per field are withheld. All 408 withheld oracle values lie inside their computed disks. The median radius is 0.03267 against an unconstrained unit radius, while the largest error of the chosen representative is about 0.17926. Those two measurements explain why retaining the family matters. A small admissible region and a chosen point inside it are different kinds of answer Davis 2026b.

A separate experiment supplies the complete real polynomial representation and fixes its imaginary constant. In that setting, harmonic conjugacy determines the missing component. Across 816 unseen evaluation points, the recorded maximum normalized discrepancy in values and derivatives is about 2.28×10^{-14}. This is reconstruction from a complete compatible representation, not the invention of unmeasured physical observations. The two experiments ask different questions, and their inputs explain their different conclusions.

The implementation also distinguishes incompatible data from unresolved arithmetic. Nearly coincident observation locations or unsafe denominators can make a numerical answer unreliable even when the exact mathematical problem has a solution. A refusal preserves the distinction between what the theorem establishes and what a particular computation has established.

A bound should come with a reason it cannot be improved

The minister now has two allocations on the table: $(10, 50)$ and $(40, 20)$. Each fits the original fuel reports. The first is enough to refute a claim that route A must already be carrying at least twenty units. It is not, by itself, enough to establish that ten is the smallest possible allocation. Another compatible schedule could conceivably deliver nine. A complete bound needs both parts: a reason no smaller answer is permitted and a compatible schedule attaining the proposed endpoint.

Here the reason is short. Total delivery is sixty and route B cannot exceed fifty. Therefore $x_A = 60 - x_B \geq 10$. Route A's own capacity gives $x_A \leq 40$. The endpoint schedules show that neither inequality can be tightened under the stated assumptions. Together, the inequalities and the witnesses establish the exact interval.

That pairing is the difference between displaying a plausible scenario and certifying an optimum. The scenario shows what can happen inside the model. The bounding argument shows what cannot happen there. Both should be inspectable when an important decision rests on the result.

Linear programming supplies a systematic version of this reasoning. Consider a feasible model written

$$Ax \leq b, \qquad Ex = d, \tag{28.5}$$

with any sign restrictions on the components included among the inequalities. We seek an upper bound on $q^T x$. If nonnegative multipliers y and unrestricted multipliers z satisfy

$$A^T y + E^T z = q,$$

then every feasible state obeys

$$q^T x = y^T A x + z^T E x \leq y^T b + z^T d. \tag{28.6}$$

The multipliers combine the reported constraints into a bound on the quantity we care about. A feasible state reaching that bound proves optimality. This is the elementary weak-duality argument; standard linear-programming theory describes the conditions under which matching primal and dual optima exist (Boyd and Vandenberghe 2004).

At the loading dock, a dual multiplier has an intelligible job. It tells us which capacity, balance, or availability statement helps limit the answer. A constraint with a zero multiplier need not be responsible for that particular bound. That does not make the record unimportant for every possible question. It means that this bounding argument did not need it.

The explanation is especially helpful when the output surprises the analyst. If the upper bound depends entirely on an old estimate of one terminal’s throughput, the next investigation has an address. If the bound depends on an accounting identity, a new article repeating the total does not add an independent constraint. The optimization can guide attention without pretending that its input records have become infallible.

When the reports cannot all be true

Imagine that a later spreadsheet records total delivery as ninety-five units, while retaining the two capacities forty and fifty. The maximum combined capacity is ninety. No allocation can satisfy all three exact statements.

The useful output is not an invented compromise shipment of ninety-two and a half. It is a small contradiction the reader can follow:

$$x_A \le 40, \quad x_B \le 50 \quad \Longrightarrow \quad x_A + x_B \le 90 < 95.$$

This certificate identifies an incompatible subset of the records. It does not establish which record is wrong. The total might include another terminal. The capacities might describe a different month. One column might be in thousands of tonnes and another in hundreds. The work now returns to the source documents.

Real measurements often arrive with tolerances. Suppose the total is reported as 95 ± 3, while each capacity is reported with an uncertainty of one unit. Even the largest permitted capacities sum to $41 + 51 = 92$, exactly the smallest permitted total. The compatible family then touches at one allocation, $(41, 51)$. If the total’s tolerance were only two, its minimum would be ninety-three and the family would still be empty.

That arithmetic explains why a tolerance is part of the claim. A program that silently widens every interval until the model becomes feasible changes the evidence contract. A program that rejects all inconsistent rows without retaining them can erase the reason the account looked suspicious. The record should show the original intervals, the reconciliation rule, and the effect of any accepted revision.

The same discipline applies to numerical tolerances. A constraint violated by a millionth because of floating-point arithmetic is different from a physical delivery exceeding capacity by five units. Scale the problem, inspect residuals in physical units, and distinguish an approximate certificate from an exact one. A solver’s success flag is a report about

a computation. The units tell us whether the residual matters to the question.

The observation that could change the decision

The original family gives $x_A \in [10, 40]$. Suppose a verified meter reading can report route B's allocation m with an absolute error no greater than δ. Combined with the balance equation, the new admissible interval for A is

$$x_A \in [\max(10, 60 - m - \delta),\ \min(40, 60 - m + \delta)], \tag{28.7}$$

provided its lower endpoint does not exceed its upper endpoint. With $m = 37.5$ and $\delta = 2.5$, we recover $[20, 25]$. With the same central reading and $\delta = 1$, the interval is $[21.5, 23.5]$. Better precision has narrowed what remains possible.

Suppose the essential requirement is twenty-five units. Neither reading can certify that the current route meets it. Under the first reading a guaranteed standby delivery of five units covers the worst compatible schedule. Under the more precise reading, three and a half units suffice. The extra measurement precision has a concrete value because it changes the reserve needed to make a specified service feasible.

Now change the decision. If the requirement is only fifteen units, either reading certifies sufficient current delivery. If the requirement is forty-five, neither reading can make route A adequate by itself. The same reduction in interval width can therefore have very different practical importance. Information is useful in relation to a question and an action, not only because a plotted band becomes thinner.

Before buying a measurement, the result is unknown. A guarantee about its usefulness must consider every result the current family and error model allow. An expected value of information additionally requires a probability model for those results. Neither is supplied by the interval alone. Reporting an average benefit while quietly choosing a favorable meter reading would confuse a planned experiment with its hoped-for outcome.

There is also the price of delay. If checking the meter takes two days and the reserve decision must be made tonight, even a perfect reading may arrive too late to alter the first shipment. One can separate decisions into those committed now and those revisable after the observation. The feasible-family calculation then connects directly to the recourse problem in Chapter 31: knowledge has operational value only through actions still available when it arrives.

Several small intervals may describe no possible world

The route totals also explain why a collection of individually permissible answers is not automatically a permissible plan. In the original family, $x_A = 40$ is possible and $x_B = 50$ is possible. They cannot occur together, because their sum would be ninety rather than sixty.

If a planning program sends forty units through A and fifty through B because each lies within its own reported interval, the mistake is not in either interval. It is in discarding the shared constraint. Keep the compatible family, or a representation that preserves the relationships needed for the decision, rather than storing only separate minima and maxima.

For an explicit example, ask for the largest difference between the two allocations. The balance equation gives

$$x_A - x_B = 2x_A - 60,$$

so the answer ranges from -40 to 20. Taking unrelated interval endpoints would suggest the same extrema here only because the two-variable balance makes the endpoints line up in this particular objective. For the total, however, the separate interval calculation gives $[30, 90]$, while the actual answer is exactly sixty. A bound can be mathematically valid yet uselessly wide if it forgets dependence.

The analytic disks in the GIGI study present the same issue in a richer form. Each disk describes one query's permitted values across a shared family of functions. Selecting an arbitrary point from every disk need not produce a function in that family. A representative function is valuable because it supplies jointly compatible answers. It must still be distinguished from the family, which describes what the supplied observations leave unresolved.

The distinction is central to the research program. A single answer is convenient to display and easy to pass to the next piece of software. The hidden cost is that the next calculation may treat an arbitrary choice as observed fact. Passing witnesses, constraints, and a record of what has been lost helps prevent that convenience from becoming false certainty.

What would justify a stronger claim?

There are several different ways to strengthen the result, and each requires different evidence. To narrow a conditional interval, obtain an observation that excludes some compatible states. To argue that the interval describes the world, test the capacities, identities, timing, and omitted routes used to construct the family. To attach a probability, pro-

vide a defensible probability model and evaluate its calibration. To claim better decisions, compare specified decisions and their consequences under a protocol that records the information actually available.

These steps should not be collapsed into one triumphant number. An exact answer to a small linear program is a real result. A correct bounded-analytic interpolation calculation is a real result. Their value becomes clearer when the reader can see the bridge from the measured record to the mathematical representation and from the representation to the decision.

The fuel example makes that bridge short enough to walk. We began with two routes and a missing allocation. We obtained an interval with explicit endpoints, exposed the report that would contradict it, identified a useful next observation, and calculated the reserve a specific requirement would need. The abstraction earns its place because each mathematical operation changes something the minister can ask, verify, or do.

How much room remains for a decision?

The earlier chapters ask whether a relationship still has room to adjust. These newer constructions sharpen a neighboring question: how much room remains in our account of that relationship? One concerns the world's available actions. The other concerns the states our observations have not ruled out. They interact, but they are not interchangeable.

Suppose the fuel decision requires at least twenty-five units per day. The initial interval $[10, 40]$ spans both adequate and inadequate supply. After the new report, $[20, 25]$ still contains schedules that fall short. The endpoint 25 meets the requirement exactly; the endpoint 20 leaves five units missing. A verified standby delivery of five independent units per day would cover every allocation in the new interval, provided it arrived in time and did not rely on the closed route. That is a decision the narrower range makes possible without guessing where inside it the true allocation lies.

The Helicity horizon work similarly connects a permitted error budget to a declared local rate of deterioration. In its implemented convention, an error budget ε, nonnegative estimator K_{disp}, and positive correlation-length estimate ℓ_c give

$$s_{\max} = \frac{\varepsilon}{K_{\mathrm{disp}}\ell_c}, \qquad \ell_c = \frac{1}{\sqrt{\lambda_{\mathrm{session}}}}, \tag{28.8}$$

when the denominator is positive. The session spectral gap belongs to the graph defined by that estimator. K_{disp} is a dispersion statistic in this implementation, not an automatically measured sectional curvature. The parameter s and the budget must have the units fixed by the adapter. A zero or unestablished denominator needs separate handling; it does not grant an unlimited safe horizon Davis 2026i.

A plan justified this morning can lose its warrant by evening. Double the estimated dispersion, holding the budget and length estimate fixed, and the permitted horizon halves. The implementation can shorten a holding limit on that basis. To express such a limit in days of geopolitical warning, we would have to specify what we observe, how often we observe it, and how the estimate is calibrated.

Back at the ministry, sixty was the combined scheduled supply. A second report confirming that total would buy the minister nothing. A count of the tankers already assigned to the surviving route would tell her whether she needs five extra units, two, or none. But a count that arrives after the standby contract expires cannot help her choose.

If the count will arrive too late, she can reserve the five units now. That is the room the geometry has made visible while she still has time to use it.

Chapter 29

How Much Can One Uncertain Connection Change?

Two groups of suppliers exchange goods freely within their own groups. One narrow commercial relationship connects the groups to each other. The analyst can estimate that relationship only within a range. A report says the network's spectral gap is 0.095. The alert threshold is 0.1.

It would be easy to compare the two numbers and announce a result. The more interesting question is whether the observations justify knowing which side of the threshold the network occupies.

The three-country calculation introduced the operator and its meaning. Here we ask how much confidence a calculation deserves when its edges are uncertain, and which kinds of repair change more than the appearance of connectivity. These questions turn spectral geometry into a tool for examining the model itself.

Put the uncertainty into the operator

Consider an undirected graph with nonnegative weights. Give each edge an arbitrary orientation solely for bookkeeping. Its incidence vector b_e has a $+1$ at one endpoint, a -1 at the other, and zeros elsewhere. Reversing that orientation changes b_e to $-b_e$ and leaves $b_e b_e^T$ unchanged. The Laplacian is

$$L = \sum_e w_e b_e b_e^T = B \operatorname{diag}(w) B^T.$$

This form separates the existence and endpoints of an edge from its weight. It also tells us exactly how an error in a weight enters the matrix.

If one edge weight changes by δ, then

$$L' = L + \delta b b^T, \qquad \|L' - L\|_2 = 2|\delta|.$$

The factor two comes from $\|b\|_2^2 = 2$. The matrix norm here is the largest amplification of Euclidean vector length. It is not the sum of all entries or a measure of the number of countries affected.

For a symmetric perturbation E, the ordered eigenvalues satisfy

$$|\lambda_k(L + E) - \lambda_k(L)| \le \|E\|_2.$$

The bound follows from the variational characterization of eigenvalues. Every unit-vector quadratic form changes by at most $\|E\|_2$; taking the relevant minima and maxima cannot exceed that common limit. The proof works because the operators are symmetric. A directed, nonsymmetric operator requires another argument.

For several uncertain edges, the triangle inequality gives a readily computed bound,

$$\|E\|_2 \le 2 \sum_e |\delta_e|.$$

It may be conservative because different edge changes need not reinforce the same pattern. Conservatism is useful when it is stated: the bound guarantees containment under the assumptions, but it may be much wider than the attainable range.

The first practical question is therefore not how many decimal places the eigensolver returns. It is how uncertain the input operator is. A highly accurate eigenvalue of a poorly measured matrix remains a poorly identified description of the world.

A four-node calculation

Join A to B with weight one, B to C with weight w, and C to D with weight one. The middle edge is the only bridge between the two strong pairs. The Laplacian is

$$L(w) = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1+w & -w & 0 \\ 0 & -w & 1+w & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

Reflection symmetry separates the calculation into patterns that agree across the reflected pairs and patterns that oppose them. For $w > 0$, the four eigenvalues are

$$0, \quad 1 + w - \sqrt{1 + w^2}, \quad 2, \quad 1 + w + \sqrt{1 + w^2}.$$

The gap is therefore

$$g_L(w) = 1 + w - \sqrt{1 + w^2}.$$

At $w = 0.1$, it is approximately 0.0950124. If the edge could instead lie anywhere from 0.09 to 0.11, the gap ranges from 0.0859582 to 0.1039682. The uncertainty crosses the illustrative threshold of 0.1.

The perturbation bound gives the wider interval $[0.0750124, 0.1150124]$. The exact family calculation gives the tighter interval because it uses more information: only this one edge varies, and its permitted values are specified. Both calculations reject the idea that the central estimate alone settles the classification.

There are three honest outputs. Every admissible graph may fall below the threshold. Every admissible graph may fall above it. Or the admissible graphs may disagree. The third output is a substantive result: the observations do not resolve the classification at the chosen precision.

It also directs further work. If the decision matters, learning more about the bridge may be more valuable than refining already well-measured internal links. Uncertainty analysis can identify which additional observation would change a decision, rather than simply surrounding every estimate with a generic error bar.

The interval is conditional on the actor list and on the meaning of an edge. It does not cover an omitted country, an unrecorded route, or a mistaken definition of trade. Those are uncertainties in the model structure. Numerical sensitivity cannot compensate for them unless they are included in the alternatives being tested.

Repairing the slow mode

For a simple gap with unit eigenvector v, the derivative

$$\frac{\partial g_L}{\partial w_e} = (b_e^T v)^2$$

measures how strongly the edge connects places separated in the slow mode. In the four-node example, the middle-edge derivative at $w = 0.1$ is approximately 0.900496. Adding 0.01 to that weight produces a first-order gap increase of about 0.009005. Direct recomputation gives about 0.008956. The difference is the finite-change remainder, not a numerical malfunction.

The derivative is a useful local guide. It is not a permanent ranking of political priorities. As edges change, the slow pattern changes too. If two relevant eigenvalues become nearly equal, the individual eigenvectors can

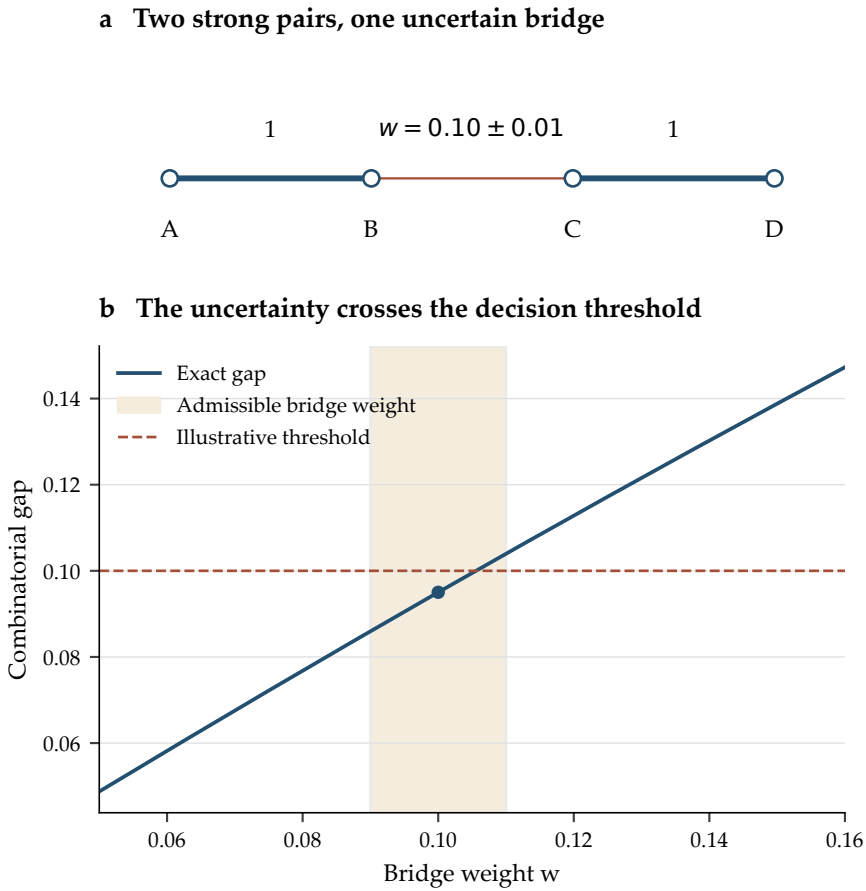


Figure 29.1: Constructed four-node graph. The strong edges have weight one; the bridge lies in $[0.09, 0.11]$. The exact gap interval straddles an illustrative threshold. The shaded region describes admissible inputs, not a statistical confidence interval.

rotate substantially under small perturbations even when the eigenvalues themselves change little. A stable gap estimate and a stable proposed partition are separate questions.

Suppose a model offers a budget for increasing selected edge weights. With $P = I - \mathbf{1}\mathbf{1}^T/n$, a precise spectral design problem is

$$\begin{aligned} \text{maximize} \quad & t, && (29.1)\\ \text{subject to} \quad & L(x) - tP \succeq 0, && (29.2)\\ & \sum_e c_e x_e \le B, \qquad x_e \ge 0. && (29.3) \end{aligned}$$

Here $L(x)$ is the Laplacian after the modeled increases, c_e is the cost per unit of increase, and B is the budget. The matrix inequality says that every zero-mean pattern has quadratic cost at least t times its squared length. Because every graph Laplacian vanishes on the constant vector, it imposes the intended lower bound on the nonconstant subspace.

This is a semidefinite optimization formulation: the unknown edge increases enter the matrix linearly, and positive semidefiniteness is a convex constraint. The optimization terminology and its numerical treatment belong to the standard convex framework (Boyd and Vandenberghe 2004). The political interpretation belongs to the specified edge model.

If an increase represents dependable civilian access, the calculation can help allocate a modeled access budget. If it represents a coercive dependency, maximizing the gap may strengthen the wrong arrangement. The objective needs a reason grounded in the problem. Geometry does not absolve the analyst from choosing what the calculation is trying to preserve.

One can also require the inequality for every operator in an uncertainty set. When only edge weights vary independently within nonnegative intervals, the smallest-weight graph gives a lower bound on the combinatorial gap because adding an edge contributes a positive semidefinite matrix. That monotonicity does not automatically hold for every normalized index, since normalization itself changes when degrees change.

The weak link is not the whole response

The gap describes the slowest nonconstant mode. Sometimes the question concerns the total response to disturbances introduced at many possible places. Then another quantity may be more informative.

For a connected undirected graph, let $L^\dagger$ be the Moore–Penrose pseudoinverse. It acts as the inverse on zero-mean vectors and as zero on the constant vector. Its trace is

$$\operatorname{tr} L^\dagger = \sum_{k=2}^{n} \frac{1}{\lambda_k(L)}.$$

Under unit-rate diffusion, integrating the decaying nonconstant propagator over time gives $L^\dagger$. The trace therefore accumulates the persistence of all those modes. It contains information that the smallest positive eigenvalue alone omits.

For endpoints i, j, define

$$R_{ij} = b^T L^\dagger b, \qquad b = e_i - e_j.$$

This is effective resistance when the weights are electrical conductances. In a general network, the same algebra measures the response to an equal injection and withdrawal at those endpoints. Calling it resistance does not give an economic edge the physical unit of an ohm.

Add a positive weight δ to that edge. On the zero-mean subspace, the ordinary rank-one inverse identity applies. Extending the result by zero on the constant vector yields

$$(L + \delta b b^T)^\dagger = L^\dagger - \frac{\delta L^\dagger b b^T L^\dagger}{1 + \delta b^T L^\dagger b}.$$

Taking traces gives the decrease in total modal persistence:

$$\operatorname{tr} L^\dagger - \operatorname{tr}(L + \delta b b^T)^\dagger = \frac{\delta \|L^\dagger b\|_2^2}{1 + \delta R_{ij}}.$$

The denominator records diminishing returns to strengthening that particular connection. The numerator records how broadly the endpoint imbalance propagates through the existing graph.

In the four-node path, the middle edge has effective resistance ten at $w = 0.1$. Increasing its weight by 0.2 reduces the trace of the pseudoinverse by 6.666667. The companion calculation checks the rank-one expression against a direct recomputation of both pseudoinverses. This is a finite intervention calculation, not merely a derivative.

If the original graph is disconnected, the formula cannot be used as though the nullspace were unchanged by joining components. A new connection can change the dimension of that nullspace. That is a different case, and it deserves its own computation.

Why a stronger connection can lower a normalized score

Imagine three trading partners with equal connections. Now strengthen only the relationship between the first two. No original connection is lost. Under the combinatorial Laplacian, the new edge contribution is positive semidefinite, so none of the ordered eigenvalues can decrease. Yet the degree-normalized gap can fall. Understanding this apparent contradiction prevents a serious mistake in policy interpretation.

Give the first pair weight w and the other two pairs weight one. Their weighted degrees are $w+1, w+1, 2$. The combinatorial Laplacian is

$$L = \begin{pmatrix} w+1 & -w & -1 \\ -w & w+1 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

Its eigenvalues are $0, 3, 2w+1$. For $w \geq 1$, the gap remains three. The unchanged slow mode separates the third node from the pair. Strengthening the internal relation of the pair does not strengthen its total connection to the third node.

Now use the symmetric normalized Laplacian

$$\mathcal{L} = I - D^{-1/2} A D^{-1/2}.$$

The eigenvalues are

$$0, \qquad \frac{w+2}{w+1}, \qquad \frac{2w+1}{w+1}.$$

At $w = 1$, both nonzero eigenvalues are 1.5. At $w = 9$, they are 1.1 and 1.9. The normalized gap has fallen from 1.5 to 1.1 even though the only physical change in the model was an increase in one connection.

The normalization asks a relative question. The first pair's internal connection now accounts for a larger fraction of its total degree. From the perspective of degree-scaled exchange, the route out of that pair occupies a smaller share. Nothing has contradicted the combinatorial monotonicity theorem. The operator changed its normalization along with the edge.

One can verify the two nonzero normalized eigenvalues through symmetry. The antisymmetric pattern on the first two nodes supplies one mode. The remaining nonconstant mode compares their common value with the third node after the degree scaling. This calculation belongs to the standard distinction between combinatorial and normalized spectral geometry (Chung 1997). The particular three-node example is fully

specified here so the distinction can be checked rather than accepted on authority.

A political interpretation must choose the relevant question. Are we asking how much exchange a given absolute imbalance produces under a fixed physical clock? Or are we asking how easily a process leaves a community when each node's activity is normalized by its total connections? Both questions can be useful. They should not share a threshold simply because both outputs are called a spectral gap.

The example also explains why comparing scores across an institutional transition requires care. If an actor's internal connections grow rapidly while its external ones remain constant, a normalized measure may signal increasing separation. That can be an accurate description of relative structure. It is not evidence that external throughput fell. The latter requires its own observation.

Removing an unobserved middle node

A dataset sometimes records the endpoints of an exchange more reliably than the intermediate institution. An analyst may want to simplify the graph while preserving what can be observed at the endpoints. There is a precise version of this operation for a specified quadratic response model.

Take three nodes in a line, with positive weights a and b. Give their state coordinates the values x_1, x_2, x_3. The quadratic energy is

$$E = a(x_1 - x_2)^2 + b(x_2 - x_3)^2.$$

Hold the endpoint coordinates fixed and allow the middle coordinate to settle to the value minimizing E. Differentiation gives

$$x_2^* = \frac{a x_1 + b x_3}{a + b}.$$

Substitution yields

$$E_{\text{reduced}} = \frac{ab}{a + b}(x_1 - x_3)^2.$$

The equivalent endpoint weight is therefore $ab/(a + b)$. For $a = 2$ and $b = 3$, it is $6/5$. This is the familiar series-conductance calculation when the weights are electrical conductances. Here it is a statement about the stated quadratic response, with whatever units that response model legitimately carries.

It is not the capacity of a freight route. Two sequential freight stages with capacities two and three have sustainable capacity at most two under a simple serial-flow model. Their equivalent quadratic response weight is 1.2 under the model above. The two answers differ because the questions differ. Substituting one for the other would turn a correct formula into an incorrect application.

For a larger graph, partition the coordinates into observed boundary nodes b and internal nodes i:

$$L = \begin{pmatrix} L_{bb} & L_{bi} \\ L_{ib} & L_{ii} \end{pmatrix}.$$

If every relevant internal component is connected to the retained boundary, L_{ii} is positive definite. Minimizing over the internal coordinates gives

$$x_i^{*} = -L_{ii}^{-1} L_{ib} x_b, \qquad L_{\text{reduced}} = L_{bb} - L_{bi} L_{ii}^{-1} L_{ib}.$$

The second expression is a Schur complement. It preserves the minimized boundary energy under these assumptions. It does not promise to preserve every eigenvalue, every transit time, or every internal consequence.

The distinction is operational. A mediator or intermediary can have its own delay, limited capacity, and internal obligations. Eliminating its coordinate assumes that the relevant internal response has equilibrated for the boundary problem. If the intermediary's response is slow on the crisis timescale, the reduced static relation can conceal the very bottleneck the study needs to observe.

The reduction also changes the network's apparent shape. Endpoints that were not directly connected can acquire an effective connection through an eliminated intermediary. That mathematical edge should be labeled as a reduced relation. It does not establish a direct bilateral agreement or a direct shipment. The computation has summarized a path; it has not discovered a new historical fact.

This is a useful form of simplification because it states what is preserved. A coarse graph can be justified by a preservation property, tested against the intended operation, and rejected when its assumptions fail. A visually simpler map without such a property may be easier to read and harder to interpret.

The connection we did not put in the graph

Uncertainty in a measured edge differs from uncertainty about whether an edge exists. Suppose a source records a connection with weight 0.1

and a defensible interval from 0.09 to 0.11. The four-node calculation addresses that uncertainty. If another route may exist but is absent from the actor register, refining the first interval does not cover the omitted route.

Model uncertainty can be represented by a family of graphs. One graph includes the disputed route; another excludes it; a third represents it as available only under a documented condition. Each graph may then contain its own weight intervals. A conclusion robust across both structural and numerical alternatives is stronger than one robust only to rounding the chosen matrix.

The graph family should be constrained by evidence. Permitting every imaginable connection creates an uncertainty set too broad to answer anything. Permitting only the analyst's preferred interpretation creates an answer by assumption. A useful family includes the alternatives genuinely compatible with the source record and the task's definition of an edge.

There is a subtle issue with zero. A measured zero flow can mean no exchange occurred during a period. It need not mean that exchange was impossible. A capacity edge, a realized-flow edge, and a permission edge therefore have different rules for becoming zero. A route with no traffic during a quiet month may remain an important alternative during a crisis. Conversely, a large past flow does not guarantee current usable capacity.

The analyst should name the graph before computing it. "Monthly realized trade values among these actors" is a different object from "qualified substitute capacity available within thirty days." The first can help describe dependence; the second can help assess an exit. Conflating them makes the geometry appear to answer a practical question its inputs never represented.

Sometimes both graphs are needed. A dense realized-flow graph can coexist with a sparse substitute graph because transactions repeatedly use one efficient supplier even though alternatives are limited. A policy that adds a qualified substitute may improve the exit graph before much trade moves to it. Looking only at current volume can miss the new option.

This is another reason the theory cannot be reduced to maximizing or minimizing one universal connection score. The desired structural change depends on the function of the edge. Some connections carry indispensable civilian services; some concentrate coercive leverage; some create independent alternatives. Their algebra can share a form while their purposes differ.

A threshold needs a margin

Let g be the true gap of the chosen model and let the available evidence imply $g \in [g_-, g_+]$. Suppose the rule assigns a warning when $g < c$. If $g_+ < c$, every admissible gap triggers the rule. If $g_- \geq c$, none does. If $g_- < c \leq g_+$, the classification depends on unresolved input information. The equality convention follows the stated strict inequality and must be implemented consistently.

A useful decision record reports the interval and the distance to the threshold. For a central estimate $\widehat{g}$ with a symmetric bound ε, the classification is unchanged by the admitted error when

$$|\widehat{g} - c| > \varepsilon,$$

with boundary cases handled according to the rule. More digits in $\widehat{g}$ do not create this margin. Better evidence or a genuinely less ambiguous configuration can.

Now suppose the final warning requires both a low gap and adverse alignment. Uncertainty must be propagated through the conjunction. A wide gap interval may be irrelevant to the current classification if every admissible alignment lies safely outside the adverse region. Conversely, a precise gap does not settle the warning when alignment remains unresolved. The required observation is the one that could change the actual branch, not necessarily the one with the largest standalone error bar.

This can be understood as a finite decision problem. Enumerate the admissible input cases, run the same versioned rule on each, and record the set of resulting labels. A singleton set supports a robust label under the modeled uncertainties. Several labels identify unresolved classification. Assigning frequencies to those labels is justified only if the cases carry a defensible probability distribution; a uniform grid is not automatically one.

The interpretation should travel with the result. "Every graph compatible with these recorded edge intervals triggers this rule" is a precise claim. "War is certain" is not its translation. The first concerns a model classification under stated evidence. The second concerns human decisions and outcomes outside that logical statement.

Buying information before buying an intervention

The uncertainty calculation can guide a practical choice about research effort. Suppose the warning changes only when the narrow bridge lies above a particular value. Obtaining a better observation of that bridge

may resolve the decision. Measuring a large internal edge more precisely may have little effect. The value of information depends on the decision it could change.

An analyst can make this comparison without pretending that information is free. A new source may arrive after the relevant deadline, be unavailable at the required resolution, or place people at risk if collected. The feasible research action is itself constrained by time, access, and responsibility. A theoretically informative observation is not necessarily an appropriate one to seek.

The same principle applies to intervention. A small computed improvement in the gap may be less useful than establishing that an alternative route is independently usable. The optimization objective should reflect the operational question. If the decision concerns continuity of one essential service, a service-specific capacity certificate may be more relevant than a global spectral gain.

The mathematics has now supplied a way to separate four kinds of statement: an exact result about a specified matrix, a bound across uncertain matrices, a reduction preserving a particular response, and an empirical claim about a relationship in the world. The distinction gives the framework more reach because it tells the reader which additional evidence is needed to move from one statement to the next.

What survives the journey back to the world

The chapter has produced several different answers. The gap measures the weakest resistance to a zero-mean pattern. Its sensitivity identifies an edge's local influence on that mode. Perturbation bounds measure how much admissible input changes can move the answer. The pseudoinverse describes accumulated response across modes. An optimization problem selects changes only after the objective and constraints are stated.

None of these quantities is improved by calling all of them resilience. Their usefulness comes from knowing which question each answers.

For the political analyst, the discipline is practical. State the relationship being represented. Specify the units. Keep the uncertainty set visible. Recompute after substantial changes. Report when different admissible observations imply different classifications. Then return to the people, routes, commitments, and clocks that gave the graph its meaning.

An uncertain bridge is not an embarrassment to the theory. It is a place where the theory can tell us exactly what we still need to know.

Chapter 30

A Stable System Can Still Hurt You

A disruption is fading at its point of origin. Elsewhere, its consequences are still growing. A factory receives less material, then delays an order; the delayed order reaches another factory after the original shortage has begun to clear. Looking only at the eventual recovery misses the interval in which the second factory cannot operate.

This is a mathematical reason to distinguish long-run stability from a tolerable path. A system can return to equilibrium and still exceed a damaging limit on the way back. The distinction matters when political decisions have deadlines, resources have finite reserves, and a brief disruption can close an option permanently.

Eigenvalues describe the end of a story

Consider two deviations from a baseline, measured on a common stated scale. Let their linear dynamics be

$$\dot{x} = Gx, \qquad G = \begin{pmatrix} -1 & 6 \\ 0 & -2 \end{pmatrix} \quad \text{per day.}$$

The second deviation decays at rate two. While it persists, it drives the first deviation with coefficient six. The first deviation also has its own recovery rate of one. Both eigenvalues are negative: -1 and -2.

For an initial disturbance $x(0) = (0, 1)^T$, the exact solution is

$$x_2(t) = e^{-2t}, \qquad x_1(t) = 6(e^{-t} - e^{-2t}).$$

The second component begins at one and immediately falls. The first begins at zero and initially rises. After half a day, the state is approximately $(1.431907, 0.367879)^T$. Its Euclidean length is greater than the initial length of one. The norm reaches approximately 1.521939 after 0.663182 day before declining to zero.

The eigenvalue calculation is correct. The transient amplification calculation is correct. They describe different aspects of the same motion.

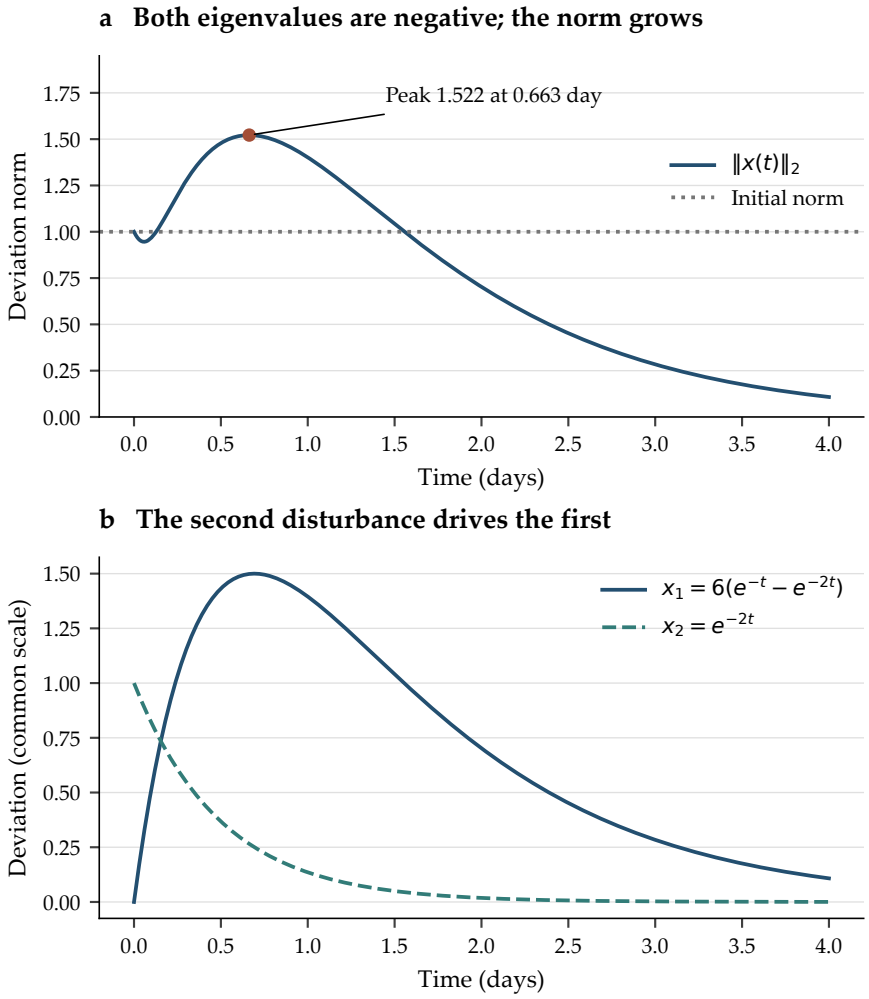


Figure 30.1: Constructed directed linear response. Both eigenvalues are negative, but an initial disturbance in the second component drives a temporary increase in the first. The norm rises by about 52% before recovery. The components are deviations on a common scale, not conserved quantities.

The matrix is nonnormal: it does not commute with its transpose. Its modes need not be orthogonal, and their combined effect can grow temporarily even when every mode ultimately decays. Nonnormal response is a standard distinction in operator theory (Trefethen and

Embree 2005). Here the small example makes the distinction visible without requiring the reader to infer it from terminology.

A useful diagnostic comes from differentiating the squared norm:

$$\frac{d}{dt}\|x\|_2^2 = x^T(G + G^T)x.$$

There exists an initially growing direction if the largest eigenvalue of the symmetric part $(G + G^T)/2$ is positive. In this example that value is

$$\frac{-3 + \sqrt{37}}{2} \simeq 1.541381 \text{ per day}.$$

This diagnostic concerns the possibility of instantaneous norm growth for some state. It does not say every initial disturbance grows immediately. For the chosen initial state, the norm initially declines slightly before the coupled response produces its larger peak.

The last detail matters. A persuasive figure should not flatten the shape of the trajectory to fit the argument. The mechanism is stronger when its exact behavior is allowed to show.

Choose a norm that describes a consequence

A norm combines several components into a measure of size. That choice is part of the model. If one component is a shortage in tonnes and another is a delay in days, their raw Euclidean length has no useful physical interpretation. They must first be placed on justified scales, or the analysis must retain separate thresholds.

One option is a weighted quadratic measure,

$$\|x\|_M^2 = x^T M x, \qquad M \succ 0.$$

The matrix M specifies how departures in different directions are counted. Its time derivative is $x^T(G^T M + MG)x$. A stable linear system can admit a metric in which the measure decreases monotonically even while the ordinary Euclidean norm grows. That fact does not erase a physical shortage or a missed deadline. The metric useful for proving eventual stability may differ from the quantity that expresses an operational limit.

For a policy application, begin with the consequence: a reserve below a minimum, a delay beyond a deadline, or an obligation that cannot be met. Then choose a state and criterion that represent that consequence. A visually smooth latent trajectory is not a substitute for this work.

The same caution applies to the framework's scalar indices. A decline in an index may indicate an improved relationship under its definition.

It does not establish that every relevant component improved, or that the path avoided a harmful peak. The index and the underlying coordinates should remain available together.

The same two links in the opposite order

Now take three places, A, B, and C. A disturbance initially sits entirely at A. For one unit of time, allow exchange only across A–B. For the next unit, allow exchange only across B–C. Let the corresponding unit-weight Laplacians be L_A and L_B.

The final state is

$$x_{AB} = e^{-L_B} e^{-L_A} (1, 0, 0)^T \simeq (0.567668, 0.245421, 0.186911)^T.$$

Reverse the order. B–C exchanges first, while both B and C still contain zero. A–B exchanges afterward. The final state is

$$x_{BA} = e^{-L_A} e^{-L_B} (1, 0, 0)^T \simeq (0.567668, 0.432332, 0)^T.$$

The whole-period graph contains the same two links in both cases. The final share reaching C differs because the route was not available in the same temporal sequence.

Each individual diffusion step preserves the total and contracts Euclidean differences. The example does not manufacture energy through switching. It demonstrates that an aggregate network can conceal the timing of access. A route that is open before a shipment arrives is not necessarily useful when the shipment needs to use it.

For short equal durations Δt, the leading difference between the propagators is

$$e^{-L_B \Delta t} e^{-L_A \Delta t} - e^{-L_A \Delta t} e^{-L_B \Delta t} \tag{30.1}$$

$$= (\Delta t)^2 (L_B L_A - L_A L_B) + O((\Delta t)^3). \tag{30.2}$$

The commutator $L_B L_A - L_A L_B$ measures the failure of the operations to commute. It is the operator version of the order sensitivity encountered in Chapter 27. Its interpretation is concrete: one operation changes the state on which the other acts.

When observations arrive monthly but decisions occur daily, that ordering may be partly unobserved. The correct response is to examine the range of admissible sequences or restrict the claim to the observed resolution. Assigning an arbitrary within-month order can create a trajectory that was never supported by the data.

Different channels do not exchange information instantly

Diplomatic communication and commercial adjustment can each connect the same three governments while operating at different rates. A change in one channel need not instantly appear in the other. To study that delay, we can represent two copies of a common response variable, one for each channel.

This construction requires compatible units. We are not adding tonnes of goods to votes. We are modeling the transmission of a specified, comparably scaled deviation through two channels, with rates calibrated or posited for that purpose.

For undirected channel Laplacians L_1, L_2 and a between-channel rate ω, define

$$L_{\text{layers}} = \begin{pmatrix} L_1 + \omega I & -\omega I \\ -\omega I & L_2 + \omega I \end{pmatrix}.$$

The off-diagonal blocks let the two copies at each place approach one another. If every place in channel one has value one and every place in channel two has value minus one, the within-channel Laplacians do nothing: each channel is internally constant. Only between-channel exchange reduces their disagreement. The corresponding decay rate is 2ω.

Consequently, the smallest positive rate of the complete connected system cannot exceed 2ω. Strong connections within each channel do not eliminate slow exchange between channels. This is a precise way for an apparently well-connected system to retain a slow organizational mode.

For a constructed example, let channel one have A–B weight one and B–C weight 0.2; reverse those weights in channel two. At $\omega = 0.01$, the layered gap is 0.02. At $\omega = 1$, it is approximately 0.456440. At $\omega = 100$, it is approximately 0.597614. As exchange becomes very fast, the slow common motion approaches that of $(L_1 + L_2)/2$, whose gap is 0.6.

The average appears only in the fast-exchange limit. Adding the two graphs at the beginning assumes away the disagreement between their copies. That may be a reasonable approximation if exchange is demonstrably fast relative to the problem's clock. It should not be an unnoticed default.

A disturbance has a place to enter

A delayed payment reaches a firm through one account. A delivery interruption reaches a public service through one supply channel. A diplomatic misunderstanding reaches a negotiation through particular

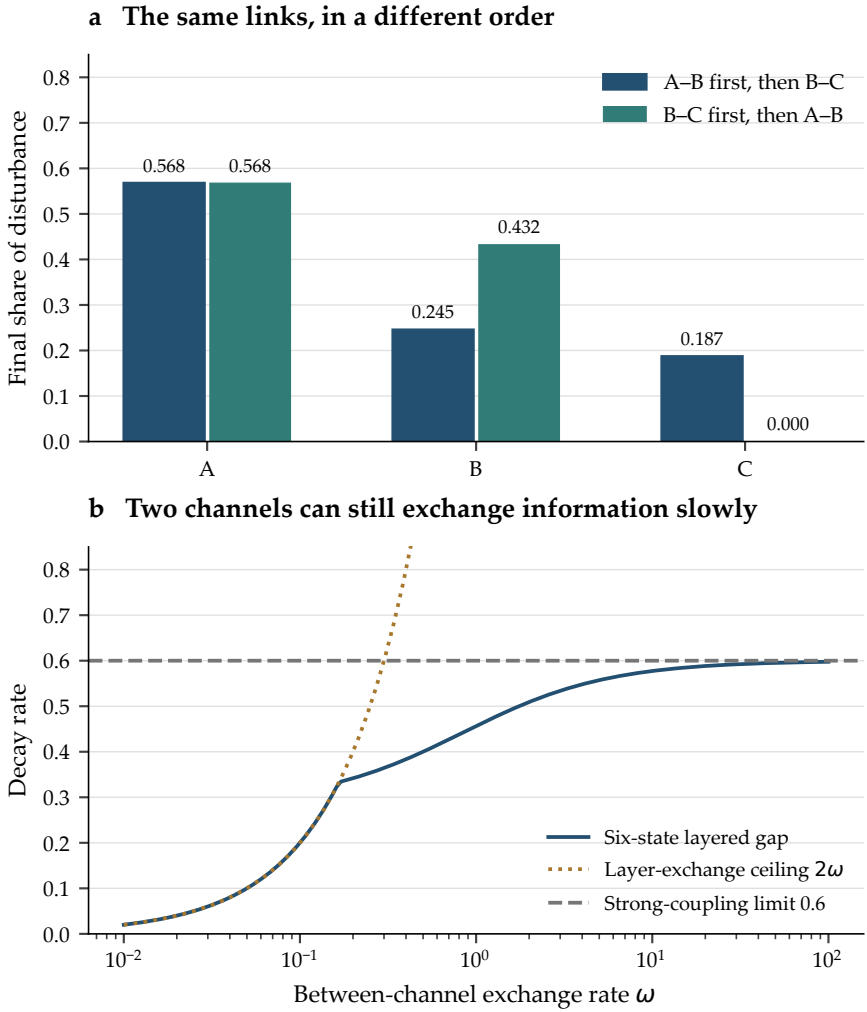


Figure 30.2: Constructed examples. Above, reversing the order of two available links changes where a disturbance ends. Below, slow exchange between two individually connected channels limits the complete system's response. All rates refer to the explicitly defined diffusion models.

participants and communications. The location of the disturbance is part of the problem. A matrix's largest possible amplification does not say that every disturbance can enter through the direction producing it.

Introduce an input $u(t)$ and an observed consequence $y(t)$:

$$\dot{x} = Gx + Bu, \qquad y = Cx.$$

The matrix B says where and in what units an input enters the state. The matrix C says which consequence is observed. If the state contains several kinds of quantities, their units and scaling must be specified before a norm is taken. A large change in an unobserved coordinate does not automatically establish a large change in the outcome of interest.

For constant G, the solution is

$$x(t) = e^{Gt}x(0) + \int_0^t e^{G(t-s)}Bu(s)\,ds.$$

The first term carries the initial condition forward. The second adds the effects of inputs arriving at different times. Each earlier input is transformed for longer before the time of observation. This is why a month containing the same total disruption can have different consequences depending on when the disruption arrives.

Take the scalar model $\dot{x} = -x + u$ with $x(0) = 0$, where time is measured in days and x is a normalized backlog. A constant input of two gives $x(t) = 2(1 - e^{-t})$. After one day, the backlog is about 1.264. If that input then stops, the accumulated backlog decays; it does not disappear at the instant the cause is removed. Restoration of a route and recovery of the downstream operation occur on different clocks.

Now place the same total input into a very short early pulse or a very short late pulse. In the ideal impulse limit, a unit input at time zero contributes e^{-t} at time t. A unit input immediately before the observation contributes almost one. The total input is the same. The remaining consequence differs because the earlier disturbance has had more time to decay. A model fitted only to period totals must either represent that timing uncertainty or restrict its interpretation to the period resolution.

The impulse is a mathematical limit, not a claim that a physical disruption has zero duration. Finite pulses can be integrated directly. The limit makes the mechanism transparent and provides a check on the finite-duration calculation.

The observation can miss the dangerous component

Consider two state coordinates: an immediately reported backlog and a downstream shortage. Suppose the first feeds the second with a delay represented by the dynamics. Observing only the first can show improvement while the second is still deteriorating. A favorable dashboard can then be mathematically consistent with a worsening consequence.

In a constructed model, let

$$\dot{x}_1 = -x_1, \qquad \dot{x}_2 = x_1 - x_2, \qquad x(0) = (1, 0)^T.$$

The solutions are $x_1(t) = e^{-t}$ and $x_2(t) = te^{-t}$. The first declines immediately. The second rises until day one, where it reaches $e^{-1} \simeq 0.368$, and declines afterward. Watching the upstream coordinate alone would miss the period of downstream accumulation.

The system is stable: both eigenvalues are -1. Its temporary downstream peak arises from the transfer of the initial disturbance through the coupling. The calculation resembles a queue moving from one stage of a process to another. Completing the first stage can reduce its own backlog while work still arrives at the next stage faster than that stage clears it.

For a real application, the sign and magnitude of the coupling must come from a specified mechanism and evidence. The example does not infer a political transmission rate from visual similarity. It shows why monitoring needs to include the consequence whose threshold matters, not only the input whose decline is easiest to report.

There is also an identification issue. If only x_1 is observed, different downstream coupling strengths can produce exactly the same observed first-coordinate history. The unseen consequence cannot be inferred uniquely without additional observations or assumptions. A highly accurate fit to the upstream series may therefore coexist with large uncertainty about the downstream outcome.

This distinction is central to model evaluation. A good reconstruction of trade, rhetoric, or an intermediate index is not automatically a good forecast of violence, civilian access, or fiscal disruption. The model must connect its state to the stated target. A hidden output equation is still an assumption.

A repeated pressure can accumulate

An isolated pulse and a repeated sequence impose different tests. A system may recover from one interruption if given enough time, yet

accumulate a persistent deficit when interruptions recur before recovery is complete. The gap between events becomes part of the state history.

For a scalar decay rate $a > 0$, suppose each event adds an amount q and events recur every T time units. Immediately after consecutive events, the state follows

$$x_{k+1} = e^{-aT} x_k + q.$$

The stable post-event level is

$$x_* = \frac{q}{1 - e^{-aT}}.$$

If aT is large, little remains from one event when the next arrives, and the peak is close to q. If aT is small, much remains, and the repeated peak can be much larger. For $q = 1$ and $aT = 1$, the limit is about 1.582. For $aT = 0.1$, it is about 10.508.

The calculation gives a concrete meaning to recovery time. A series of individually tolerable disturbances can become intolerable because the intervals are too short for accumulated consequences to dissipate. Nothing in the example requires the individual shocks to grow. Their frequency changes the experienced burden.

The units matter. If a is measured per day, T must be in days. Substituting a weekly interval without conversion changes the modeled overlap. The dimensionless product aT expresses the amount of recovery available between events. It is often more interpretable than either number alone.

The example also separates two interventions. Reducing q limits the size of each disturbance. Increasing T permits more recovery between disturbances. Increasing a speeds recovery under the chosen model. These changes can produce a similar reduction in the long-run peak while operating through different mechanisms. Their feasibility and distributional consequences need separate analysis.

For a negotiation process, the analogous observation might be the time required to resolve a breach compared with the interval between new breaches. For a civilian service, it might be stock replenishment compared with interruptions. The analogy becomes a model only when those quantities are measured and the update equation is justified. The algebra organizes the inquiry; it does not supply the empirical rate.

The computer can invent an oscillation

A numerical implementation adds another clock: the time step used to approximate the dynamics. An unfortunate choice can make a stable model look unstable even when the continuous equation is correct.

For $\dot{x} = -ax$ with $a > 0$, forward Euler gives

$$x_{k+1} = (1 - a\Delta t)x_k.$$

The continuous solution decays monotonically for a positive initial value. The discrete approximation decays in magnitude only when $|1 - a\Delta t| < 1$, or $0 < a\Delta t < 2$. It preserves nonnegative values under this update only when $a\Delta t \leq 1$.

Take $a = 1$ per day. With a half-day step, the multiplier is 0.5. With a one-and-a-half-day step, it is -0.5: the numerical solution alternates signs while shrinking, although the exact positive solution never crosses zero. With a two-and-a-half-day step, the multiplier is -1.5, and the numerical solution grows while alternating signs. The computer has created a crisis in a model whose exact solution steadily decays.

For constant linear dynamics, the exact discrete update is $x_{k+1} = e^{G\Delta t}x_k$ in the absence of input. For varying inputs and nonlinear models, appropriate integration methods and convergence checks are required (Hairer, Nørsett, and Wanner 1993). The practical check is to reduce the step and see whether the consequential outputs stabilize, while respecting any discontinuities in the inputs or rules.

A small local numerical error does not automatically imply a small decision error. If the output lies close to an alert boundary, even a modest integration difference can change the category. The implementation should therefore compare numerical error with the margin to the decision boundary, not only with the overall scale of the state.

This is not an argument for using an arbitrarily tiny step. A model with monthly observations does not acquire measured daily detail because it is integrated every second. Numerical resolution and evidential resolution are different. The first controls approximation of the chosen equations; the second controls what the equations are justified in describing.

A fitted rate must survive a change of clock

A graph of annual trade values does not directly supply a diffusion rate measured per day. To write $\dot{x} = -Lx$, the model must either define L as a rate operator or introduce a conversion with a calibrated interpretation. Otherwise the expression mixes a stock-like or annual-flow weight with a derivative whose unit is different.

One possible model is $\dot{x} = -\beta L x$, where L is dimensionless and β has units of inverse time. If the graph gap is g, the slow-mode relaxation time is $1/(\beta g)$. The same graph can describe faster or slower modeled adjustment depending on β. The gap alone cannot establish a calendar countdown.

Suppose $g = 0.2$ and $\beta = 0.5$ per day. The slow-mode e-folding time is ten days. Expressing time in weeks changes the numerical value of β to 3.5 per week and the same time constant to $10/7$ weeks. The physical interpretation is unchanged. If a code change alters the predicted calendar time merely by changing the unit label, the implementation has failed a basic invariance check.

An e-folding time describes the decay of a specified linear mode to e^{-1} of its initial amplitude. It is not a predicted time to war, negotiation, or institutional failure. To forecast such an event, the model needs an outcome definition, a relationship between state and event, and empirical evaluation. This distinction is especially important when the mathematical language of critical slowing down is used to motivate a warning.

Rates may also change with the state. A service can recover quickly when spare capacity exists and slowly when that capacity is exhausted. A diplomatic channel can operate differently after a breach. A constant-rate model is then a local approximation. Its range should be examined through held-out trajectories and residual structure, rather than extended indefinitely because the first fit looks plausible.

A dynamical explanation has to close the loop

The dynamic account is complete only when the reader can follow four links: where the disturbance enters, how it propagates, which consequence is observed, and how that consequence is compared with the relevant constraint. Leaving any link implicit lets a persuasive trajectory stand in for an explanation.

In the two-stage example, the initial upstream backlog is the input, the coupling carries it downstream, the second coordinate is the consequence, and a stated service limit supplies the constraint. In the repeated-pulse example, the event size and spacing determine accumulation relative to recovery. In the numerical example, the integration method itself introduces an artificial sign change. These are different explanations for a rising curve, and they call for different responses.

The framework benefits from keeping them separate. A structural bottleneck may require another usable route. Slow recovery may require time or capacity. Poor observation may require a better evidence source.

A numerical artifact requires a computational repair. Calling all four "instability" would make the model sound unified while making its advice less useful.

The wider claim of the book survives that precision. Relationships have a shape, and the possible motions through them have a clock. A system's future depends on both. The task is to represent enough of each that the calculation can explain why this disturbance, entering here and now, could produce this consequence before there is time to recover.

From a graph to a model of motion

These examples supply three different reasons that a static connectivity score can miss an important consequence. Directional coupling can produce a temporary surge. Time ordering can determine whether a route is usable. Slow exchange between channels can preserve a bottleneck that aggregation conceals.

None of the three requires abandoning geometry. Each requires specifying more of it: the operator, the state, the available motions, and the clock. The resulting model is more demanding to measure, but it is also more explicit about what would count as a failure.

An empirical test would ask whether the added structure improves forecasts made from the same available information. Does a directed response model predict damaging peaks better than a symmetric one? Does the observed sequence add value beyond monthly totals? Does retaining separate channels improve held-out performance enough to justify its extra parameters? These are comparisons the implementation and benchmarking chapters make operational.

The ultimate question is not whether the system recovers eventually. It is whether people can remain within tolerable conditions while it recovers, and whether the paths that remain open last long enough to be used.

Chapter 31

The Route, the Reserve, and the Deadline

A government can identify a substitute supplier and still be unable to replace tomorrow's delivery. A hospital can have an alternative source of fuel and still run out before the first truck arrives. An agreement can be feasible after six months of preparation and unusable during a six-day crisis.

This is the difference between having an alternative and being able to reach it.

The earlier chapters introduced feasible sets and diplomatic exits. Here we put quantities and deadlines into that language. The result is a small operational model of constrained change: routes have capacities, reserves absorb deficits, decisions take time, and some apparent alternatives share the same bottleneck.

A route is not its capacity

In a directed supply graph, let q_e be the amount sent through route e per day and c_e its available capacity. The basic constraint is

$$0 \le q_e \le c_e.$$

At an intermediate depot that neither produces nor consumes the modeled commodity, inflow equals outflow. If B is the directed incidence matrix, this conservation condition is expressed by the appropriate rows of $Bq = 0$. Source and destination rows carry the external supply and demand instead.

The maximum-flow problem chooses admissible route flows to maximize delivery. Its dual description is a minimum cut: a partition separating source from destination whose outward capacities limit every possible delivery. In the small network of Chapter 18, the two arcs entering the destination have capacities four and six. That cut proves

that delivery cannot exceed ten. The explicitly constructed ten-unit flow proves the bound is attainable.

The paired calculations matter. An optimizer returning ten is a numerical result. A feasible flow of ten together with a cut of ten is an inspectable certificate. The reader can check conservation and capacity without trusting the optimizer's internal procedure.

A *certificate* in this setting is a compact object that verifies a specified mathematical conclusion. It is not a declaration that the infrastructure model is complete. If an important route was omitted or an available capacity was misstated, the certificate remains valid for the wrong input graph.

This distinction is central to implementation. Software can certify that a model has been solved correctly. Evidence is required to justify that it represents the intended situation.

Capacity is conditional on the operation

An available capacity is rarely a permanent property of a line on a map. It may depend on staffing, weather, contracts, power, inspection, permissions, or safety. A model should state the conditions under which c_e is available. It should not interpret engineering capacity as guaranteed usable throughput.

One way to represent uncertainty is an interval $c_e \in [\underline{c}_e, \overline{c}_e]$. A conservative plan can be tested against the lower capacities. Another is a set of joint scenarios, preserving the fact that several routes may fail together. Independent intervals allow every combination of endpoint values, including combinations that may be physically impossible; a scenario set can encode shared causes more accurately when evidence supports it.

Suppose two nominally alternative routes depend on the same payment channel. Their physical independence does not make their operation independent. If the shared channel becomes unavailable, both routes may fail together. Counting them as two independent exits exaggerates redundancy.

A useful representation makes the shared condition explicit. For a binary availability variable $z \in \{0, 1\}$,

$$q_1 \le c_1 z, \qquad q_2 \le c_2 z.$$

When $z = 0$, both flows are unavailable. When $z = 1$, the two separate capacity limits still apply. The equation makes the common dependency visible rather than burying it in a correlation coefficient after the fact.

This is also why diplomatic alternatives deserve a structured register. Two proposals may share a necessary guarantor, an unavailable legal instrument, or a single disputed permission. The number of named proposals is not the number of independently executable paths.

A stock buys time

Let s_k be the stock available at the beginning of day k, q_k the usable inflow during that day, and d_k the demand. A basic stock equation is

$$s_{k+1} = s_k + q_k - d_k, \qquad s_k \geq 0.$$

The formula assumes a common commodity unit and a convention about when deliveries become available. If arrivals late in the day cannot serve morning demand, the daily model is too coarse for that task. The time step must resolve the operational constraint.

In the constructed example, the stock is fifteen and the daily deficit after a route loss is three. Five days exhaust the stock. A reserve is therefore a time buffer, not a substitute route. Once spent, it cannot continue serving the same role without replenishment.

The same algebra can represent a financial buffer, but only after specifying obligations and payment timing. A balance that is legally restricted or unavailable in the needed currency is not necessarily a usable reserve. A model that subtracts all nominal assets from all nominal liabilities can conceal the timing and access constraints that create the crisis.

The buffer also changes the interpretation of a warning. Two systems with the same current shortfall can have different time to exhaustion. A daily indicator of the shortfall alone misses that difference. The state needs memory: the accumulated stock after previous deficits.

With constant demand d and inflow $q < d$, the continuous approximation gives

$$t_{\text{exhaustion}} = \frac{s_0}{d - q}.$$

Its sensitivity to the shortfall is nonlinear. Near balance, a small change in inflow can substantially extend the buffer. Under a large deficit, the same absolute change buys less time. The equation provides a reason to measure the actual deficit rather than interpreting all disruptions through a common percentage.

Build the transition into the feasible set

Suppose a substitute route becomes available only after a preparation time T_{build}. It may require up-front resources b and have a later operating capacity c_{new}. The existence of a future plan with enough capacity does not imply that there is a feasible trajectory from today's state to that plan.

A time-expanded network represents this distinction explicitly. Make a copy of each location for each time step. Transport edges connect the relevant departure and arrival times. Storage edges connect a location to itself at the next step, with the appropriate stock limits. A future route appears only at the time it becomes available. A feasible flow through this expanded graph respects the clock as well as the physical connections.

This construction can be understood without drawing every copy. A consignment cannot use a road before it is available, arrive before it departs, or remain in storage without occupying capacity. The enlarged graph records those ordinary facts as constraints.

Let x_k collect stocks and other relevant state variables, and let u_k collect decisions. A broader formulation is

$$x_{k+1} = f_k(x_k, u_k, w_k), \tag{31.1}$$

$$u_k \in U_k(x_k), \qquad x_k \in X_k, \tag{31.2}$$

where w_k represents disturbances and X_k defines acceptable states. A viable initial state is one from which an admissible decision rule can keep the trajectory within the required sets for the specified horizon and disturbance assumptions.

The quantifiers matter. "There exists a plan for one favorable scenario" differs from "there exists a rule that succeeds for every allowed scenario." The latter must also respect what the decision-maker knows when acting. A plan that chooses today's action using tomorrow's unobserved disruption is not an implementable robust policy.

In notation, a nonanticipating rule has the form

$$u_k = \pi_k(\mathcal{I}_k),$$

where $\mathcal{I}_k$ contains information available by step k. The rule may adapt when new observations arrive. It may not adapt in advance to information that has not arrived.

This is the mathematical counterpart of the book's information clock. A model can otherwise make a transition look feasible by granting the planner foresight that no person possessed.

The value of a margin

A plan that exactly meets every capacity constraint has no room for a small adverse change. A margin can be defined relative to a specific requirement: spare delivery capacity, days of reserve, or distance from a maximum acceptable delay. These margins have different units and should remain named.

An optimization model can attach a shadow price to a binding constraint. In a linear program, this is the local rate at which the optimized objective changes when the constraint is relaxed, within the range where the same local structure applies. A high shadow price identifies a modeled bottleneck. It does not place a moral price on the people affected or guarantee that the constraint can be changed at the stated cost.

Return to the supply example after loss of $s \to b$. Increasing the capacity of $s \to a$ alone does not improve delivery: the usable downstream cut still totals six. Restoring some capacity on $s \to b$ can improve delivery until another limit binds. The distinction is visible in the graph before any elaborate calculation.

That is the practical value of the dual view. It directs attention toward the constraint currently limiting the intended operation. It also makes a failed intervention intelligible. Spending resources on a nonbinding constraint can improve a component without improving the task.

Political constraints are not always linear, and many cannot be represented as purchased capacity. The formulation remains useful as an explicit question: which condition must change for this proposed path to become usable? If the analyst cannot answer that question, a larger index value may be cosmetic.

Work backward from the last acceptable day

A planner wants to keep a public service operating through two uncertain days and finish with a small reserve. The question is not whether tomorrow's expected delivery looks adequate. It is how much stock must be available now for an implementable policy to meet the requirement across the admitted conditions.

Take a constructed daily balance

$$x_{k+1} = x_k + u_k - d_k,$$

where x_k is usable stock, the delivery decision satisfies $0 \le u_k \le 3$, and demand lies in $[4, 6]$. All quantities use the same commodity unit. Assume deliveries are usable before that day's demand, there is no

spoilage or storage limit, and the stock must remain nonnegative. The terminal requirement after two days is $x_2 \geq 2$.

On the last day, the largest allowed demand is six and the largest delivery is three. To finish with two under that demand, the stock entering the day must be at least five. On the preceding day, another worst-case deficit of three must be covered. Thus $x_0 \geq 8$ is necessary and sufficient for this simple robust problem. Delivering three on each day meets the bound exactly when demand is six on both days.

If one instead plans only for demand four, the opening stock requirement is four. The difference between four and eight is not a dispute between two solvers. They solve different questions. One covers the selected nominal sequence; the other covers every demand sequence in the admitted intervals. The modeler must decide which requirement the decision actually needs.

The calculation can be written as a backward recursion. Let V_k be the set of states from which the remaining requirements can be satisfied. For a decision made before that period's disturbance is known,

$$V_k = \left\{ x \in X_k : \exists u \in U_k(x) \quad \forall d \in D_k, quadf_k(x, u, d) \in V_{k+1} \right\}.$$

Start with the terminal set and work backward. In the example, $V_2 = [2, \infty)$, $V_1 = [5, \infty)$, and $V_0 = [8, \infty)$. This is a finite-horizon version of the viability question: from which states can the system remain within the required conditions? Viability theory develops the broader mathematical setting (Aubin 2009).

The order of the quantifiers records the information available. If the decision can be made after demand is observed, the condition becomes "for every admitted demand, there exists an admissible response." That is a different information structure. In this particular stock example, maximum delivery works in either order, so the numerical bound is unchanged. In other problems the distinction changes feasibility dramatically.

Consider two destinations and six units that must be positioned before learning which destination will need all six. A plan that places three at each guarantees only three at the eventual destination of need if no transfer is possible. A calculation that places all six at whichever destination later needs them uses future information. It is a perfect-information benchmark, not an implementable advance plan. If a later transfer of at most two is feasible in time, the equal split can guarantee five. The improvement comes from measured recourse capacity, not from granting the planner clairvoyance.

The same distinction applies to a proposed diplomatic exit. An arrangement that can be chosen after all other parties' decisions are known may not be available when a commitment must actually be made. A feasible path must respect the sequence of information as well as the sequence of actions.

The horizon has an edge

Finite-horizon planning can create a misleading success if the terminal condition is too weak. A plan may satisfy today's and tomorrow's demand by spending every reserve, then fail immediately after the modeled window ends. The solution is feasible for the stated problem, but the problem ends before the consequence appears.

The terminal reserve in the example prevents one version of that error. It is still a modeling choice. A reserve of two does not ensure indefinite survival when maximum continuing delivery is three and minimum demand is four. Under those conditions every finite reserve eventually runs out. The long-run feasible set is empty unless delivery, demand, or another modeled condition changes.

That distinction should be stated plainly. A bridge arrangement can buy time without solving the permanent imbalance. Its value depends on what is being prepared during the time purchased. A temporary reserve that allows a replacement source to become usable can change the outcome. The same reserve without a feasible follow-on arrangement only postpones exhaustion.

One way to represent the follow-on requirement is a terminal set from which a known sustainable policy is available. Another is to extend the horizon until the proposed replacement becomes operational. Either method requires evidence for the future capacity. An announced project should not enter the terminal set as if its completion and usable output were guaranteed.

The horizon also affects reporting. A successful thirty-day service continuity test should not be presented as a claim of permanent resilience. A model can be valuable precisely because it identifies the length of the interval it covers and the conditions that must change before that interval ends.

The price of the binding constraint

The earlier maximum-flow example paired a feasible delivery with a cut proving that no larger delivery was possible. The same idea can tell us which modeled improvement has value at the margin.

Take a simplified destination with two routes capable of delivering four and six units per day, with enough upstream supply. Their total is ten. Suppose a downstream receiving process can handle only seven. Maximum usable delivery is then seven, even though the routes could carry ten. Increasing the first route from four to five changes nothing. Increasing receiving capacity from seven to eight raises delivery by one.

If the receiving capacity is b, the optimal delivery is

$$Q(b) = \min(10, b), \qquad b \geq 0.$$

Below ten, the marginal improvement is one unit of delivery per unit of receiving capacity. Above ten, it is zero. At ten, the active limitation changes. A single shadow price should not be extrapolated through that change as though it described unlimited returns.

Now suppose improving receiving capacity also requires staff who are shared with another essential operation. The feasible improvement is constrained by that shared resource. The one-unit mathematical value does not establish that the intervention is available at an acceptable cost. The broader optimization must represent the competing operation or preserve it as an external constraint.

This is why a shadow price is a local property of a stated optimization problem. It identifies the value of relaxing one constraint while the relevant model structure and objective remain in place. It is not a universal economic valuation and not a price on the dignity of the people served. Its usefulness is diagnostic: it reveals why improving a visible component may leave the intended task unchanged.

The political analogue is a proposal that improves diplomatic language while a missing authorization still prevents implementation. The language may matter for other reasons, but it is not the binding constraint in that particular operation. A good analysis keeps both facts in view. It neither dismisses all nonbinding improvements nor claims that every improvement immediately enlarges the feasible path.

Total delivery and fair access are different objectives

A transport model maximizing total delivered volume can have several equally optimal allocations. Some may leave one community substantially underserved. The objective has not malfunctioned. It was never asked to protect a minimum level of access for every community.

Let two destinations require six and four units and let total available delivery be six. A max-total objective permits allocations $(6, 0)$, $(5, 1)$, or

(3.6, 2.4) when the routes allow them. All deliver six. Their distribution of coverage differs sharply.

Introduce a minimum coverage fraction α and require

$$q_1 \geq 6\alpha, \qquad q_2 \geq 4\alpha, \qquad q_1 + q_2 \leq 6.$$

Maximizing α gives $\alpha = 0.6$ when there are no other binding route constraints. The allocation (3.6, 2.4) attains it. The objective is now explicit about preserving proportional access across the two destinations. It is one possible fairness criterion, not the only defensible one.

If the second route can carry at most one unit, then $4\alpha \leq 1$ and the minimum coverage cannot exceed 0.25. Additional supply sent along the first route can raise total delivery without raising that minimum. The route constraint explains the divergence between the two objectives. It also identifies a candidate improvement to examine if the minimum coverage is the intended policy goal.

Different needs can justify different priorities, and some minimums may be nonnegotiable. The model should express those choices openly. A single weighted sum can conceal them if its weights are presented as technical facts rather than decisions about whose shortfall matters. Reporting both total delivery and the distribution of unmet requirements often makes the tradeoff easier to understand.

The calculation has a direct implication for a theory of peace. A network can become more efficient in aggregate while some people lose the means to sustain an ordinary life within it. Their feasible set can shrink even as the system's total output grows. A serious structural analysis must ask where the new capacity goes and which constraints remain binding for those who cannot use it.

When two alternatives share one failure

Consider two substitute suppliers, each able to cover a shortfall on its own. If both require the same certification, payment channel, or receiving facility, their independence is conditional. The shared prerequisite may be the real single point of failure.

Suppose each supplier is individually available with probability 0.9. Under independence, at least one is available with probability $1-0.1^2 = 0.99$. If their availability is perfectly coupled through one shared condition, the probability is only 0.9. Both statements are compatible with the same two marginal probabilities. The assumed dependence changes the value of diversification.

Those probabilities are constructed. The practical study needs evidence about the shared conditions. Two suppliers in different countries may still use the same upstream component, insurer, settlement route, or technical standard. Geography alone does not establish operational independence.

One can model a shared availability variable z and condition each supplier's capacity on it. A richer representation distinguishes several possible common causes and their scope. For a conservative analysis, a scenario set can test loss of each documented shared prerequisite without assigning unsupported probabilities to the scenarios.

The result should identify what is robust to which failure. "Either supplier can cover the shortfall if the shared receiving process remains available" is more useful than "two suppliers provide redundancy." The condition tells a decision-maker what still needs protection or an alternative.

The same logic applies to proposals mediated by the same institution. Several nominal exits may all require one guarantee or one actor's consent. Creating a genuinely independent path can therefore change the feasible set more than adding another proposal with the same prerequisite. The graph should reveal that distinction rather than merely count more edges.

A reserve can be inaccessible

Stock on a balance sheet is not always stock usable for the task. A warehouse may hold the wrong item, a reserve may be legally restricted, or a delivery may arrive after the need it was intended to meet. A model should record usable stock under an explicit conversion and timing rule.

If nominal stock is s and only a documented fraction η can be made usable in the relevant period, the effective stock under that simplified representation is ηs. The fraction cannot be chosen merely to make the model fit. Its meaning might concern quality, compatibility, access, or conversion capacity, and those are different mechanisms.

Take an illustrative nominal reserve of twenty units, of which twelve are usable before the deadline. With a daily shortfall of three, the relevant buffer is four days, not 20/3. A remaining eight units that become usable on day six cannot bridge the gap without another arrangement. The reserve's availability schedule belongs in the state.

This can be represented directly in the time-expanded network. Stock that is not yet usable occupies a separate state or node and moves into usable inventory only through the conversion edge at the permitted time and capacity. The representation is larger because the real operation has

an additional stage. Compressing it into one number would remove the constraint the planner needs to see.

Financial reserves present the same formal issue without being the same commodity. Currency, legal access, settlement time, and the obligation to be paid all matter. A large nominal asset position can coexist with an immediate inability to meet a specific payment. The model's units must follow the obligation rather than adding unlike assets simply because they all have a quoted value.

The proposal that the calculation rejects

A valuable model can show why an attractive proposal fails before anyone relies on it. In the constructed examples, a route upgrade failed because receiving capacity bound first; a nominally sufficient reserve failed because part of it was inaccessible; a transition failed because its prerequisites completed after the bridge arrangement ended. Each failure produced an inspectable reason.

An implementation should preserve that reason. An infeasible optimization result can be accompanied by the conflicting constraints or a certificate when the formulation permits one. A planner can then ask which constraint is wrong, which might be changed, and which must remain protected. "No solution" becomes the beginning of a concrete conversation rather than a mysterious verdict from software.

There is also a distinction between mathematical infeasibility and political unwillingness. If the constraints encode preferences as absolute prohibitions, the feasible set can be empty because those preferences were fixed. If a preference is negotiable, a different problem may have a solution. The model should state which restrictions are physical, legal, institutional, resource-based, or assumed for the scenario. Their status determines what a proposed repair means.

The geometry of an exit is therefore richer than the presence of another point on a diagram. It includes the route, its prerequisites, the time available, the resources that can actually be used, and the requirements that must remain satisfied along the way. The framework becomes operational when it can show that path, or identify the precise condition preventing it, without confusing a hopeful endpoint with a reachable future.

What this adds to the theory of conflict

The geometry of a constrained relationship has a temporal depth. A system may possess many long-run alternatives and few short-run ones.

It may have ample nominal resources and insufficient usable resources before a deadline. It may appear connected while several paths depend on one hidden permission.

These distinctions explain why dependence, tolerance, and available exits must be modeled together. Dependence creates obligations that persist. Tolerance affects which arrangements can be negotiated or maintained. Feasible exits depend on resources, authority, and time. A warning becomes more useful when it identifies which of those conditions is narrowing.

The framework can therefore support more than a label of danger. It can organize an inquiry into the time available, the bottleneck that binds, and the changes that would preserve civilian options. The calculations do not choose the political objective. They make the consequences of a stated objective and a stated set of constraints harder to evade.

Part V

Putting the warning to work

A warning has to survive contact with its own evidence. We test the instrument where it can mislead, define what a forecast must predict, and follow a dated source record through the actual calculation. Then we put the geometric approach beside current forecasting methods and specify a common test strong enough to change our minds.

Chapter 32

When the Instrument Lies

The chart is smooth. The warning arrives before the event. The retrospective explanation sounds persuasive. None of those facts tells us whether the model learned something that would have been available before the event happened.

A serious instrument needs a way to fail visibly. A thermometer that quietly substitutes the room's expected temperature when its sensor disconnects is worse than one that displays no reading. Forecasting systems can make the same substitution in less obvious ways: filling missing observations with calm values, using later revisions in earlier forecasts, or presenting a branch-specific score as if its meaning had remained constant.

This chapter develops failures that can be deliberately tested. The object is not to assemble a list of reasons never to model conflict. It is to identify the conditions under which a model's apparent success stops being evidence for the claim we want to make.

The answer changes when only the ruler changes

Begin with a harmless intervention. Convert every monetary quantity and its matching monetary reference from billions to millions. The countries have not changed their behavior. A dimensionless score based on their ratios should not change either.

The Psychohistory flux term uses ratios of trade and energy quantities to reference quantities. The companion test multiplies all four matching monetary inputs by one thousand and checks that the engine's computed results remain unchanged. This is an invariance test. It does not demonstrate predictive accuracy; it detects one way of producing a prediction from a bookkeeping convention.

Now perform an intentionally invalid conversion: scale the observation but leave its reference unscaled. The resulting score changes. The test should detect that mismatch before the model runs, because the arithmetic alone cannot know whether a thousandfold increase represents a new unit or a real change.

The lesson extends to graph spectra. A combinatorial gap scales with all edge weights. A normalized gap does not. If a threshold was calibrated on one operator and applied to the other, a plausible-looking number can conceal a different question. An implementation should therefore store the operator definition and normalization alongside the threshold.

The strongest version of this test changes representation while preserving the modeled situation. Relabel the actors consistently. Reverse arbitrary orientations in an undirected incidence matrix. Reorder rows of a source file. Change an admissible currency unit consistently. If the result changes for reasons the theory says should be irrelevant, the failure has a small, reproducible witness.

A witness is a specific example that makes a claim inspectable. For an invariance claim, it is a pair of equivalent inputs that produce incompatible outputs. For a claimed bound, it can be a small matrix violating the bound. The witness does not have to resemble an entire historical conflict to establish the mathematical failure it exhibits.

More columns are not always more information

An input table can look richly informed while repeatedly encoding the same two observations. In the inspected Psychohistory implementation, the integration score combines normalized trade and energy with two proxies. The foreign-investment proxy is the average of the two normalized inputs; the supply-chain proxy is 0.8 times normalized trade.

Writing those normalized inputs as u and v, the implemented combination simplifies to

$$K = 0.35u + 0.35v + 0.15\frac{u+v}{2} + 0.15(0.8u) \tag{32.1}$$

$$= 0.545u + 0.425v. \tag{32.2}$$

The four displayed components contain two independently supplied quantities in this implementation. The names of the proxies do not turn them into separate measurements of investment and supply chains.

This matters when interpreting an explanation. A model might say that trade, investment, and supply-chain integration all contributed to an alert. If two of those contributions were derived from trade itself, the explanation should make the dependency visible. Otherwise the reader may mistake one observation repeated through several formulas for corroboration from several sources.

The appropriate next experiment is to replace a proxy with a separately documented measurement and compare the result under the same information cutoff. That is a substantive extension. Merely adding another transformed copy of the existing input can change a fitted model's flexibility without adding evidence about the world.

The same problem appears in text features. Ten reports repeating one original claim are not ten independent confirmations. The data pipeline should retain source relationships when they are known. A model that treats syndication as independent evidence can become most confident about the best-repeated claim.

The model changes meaning while keeping the column name

The engine contains an ordinary integration branch and a disconnection branch, separated by the specified integration threshold. In the ordinary branch, the field called helicity stores the tolerance-to-integration quantity. In the disconnection branch, that field is reused for the negative of a hostility index. The field called the spectral gap is also reused there for a diplomatic-mobility measure.

These branch choices can be implemented consistently. A downstream plot that treats the combined column as one continuous physical quantity cannot be interpreted consistently. A jump may reflect a changed definition as well as a changed relationship.

A practical output record therefore needs a semantic type, not just a floating-point value. Store the branch name, formula identifier, and interpretation with the result. When plotting a time series that changes branch, show the boundary and use the appropriate labels. Do not calculate a cross-branch derivative unless a justified common scale has been established.

This is analogous to a display switching from temperature to wind chill without changing its heading. Both readings can be useful. A single unlabeled curve invites the wrong inference.

Boundary tests are especially valuable. Evaluate the rule immediately below the threshold, at it, and immediately above it. Confirm the intended strict or non-strict inequality. Confirm which alert rule has precedence when more than one condition is satisfied. A model's scientific description and its executable rule must agree at the points where a decision changes.

The future leaks into the past

A source publishes a January observation in February, then revises it in April. A retrospective dataset keeps only the April value. If the evaluation asks what a forecaster could have known in January, that dataset is insufficient by itself. The revision may be more accurate about January and less appropriate for reconstructing January's information set.

Leakage also enters through fitted transformations. A mean calculated over the whole study period allows later observations to influence earlier standardized scores. A threshold chosen after inspecting all the test events is no longer being tested on untouched data. A language model asked about a past crisis may already contain knowledge of its outcome.

The repair is an explicit temporal boundary. Fit normalizations, thresholds, feature selections, and any probability calibration using only the permitted training information. Freeze them before scoring the next held-out interval. When an expanding-window procedure refits, record the new version and its effective date.

The boundary must include target construction. If the analyst defines a crisis window by looking backward from a known war, the resulting sample may be useful for studying precursor patterns. It is not automatically the population faced by a continuously operating warning system. That system also encounters quiet periods and ambiguous alerts that do not culminate in the selected events.

The book's historical research record and a future operational evaluation answer related but distinct questions. The former preserves the reported event studies. The latter must add the complete stream of forecast opportunities, including no-event periods, to estimate operational false-alarm burden and probability calibration.

The sensor fails where the danger is greatest

Missingness is often treated as a nuisance to be filled in. In a conflict setting, loss of access can be part of the event being studied. A communications outage, inaccessible locality, or disrupted reporting institution can remove observations precisely when conditions are changing fastest.

Replacing that absence with zero can produce an artificial improvement. Carrying forward the last value can create a false impression of stability. Interpolating between the last pre-crisis observation and a later recovery uses future information and may erase the crisis's shape.

There is no universal imputation that resolves this problem. The model should retain a missingness indicator, the age of the last observation,

and the permitted interpretation of any substitute. Where uncertainty is too large for the decision being made, the system should return an unresolved assessment rather than a fabricated precision.

This does not make the model useless. An explicit statement that the observation process has broken can itself be operationally valuable. It tells the analyst that a change in the chart cannot be read as a change in the underlying situation without further evidence.

The constructed reporting example in Chapter 20 is a useful diagnostic. Change the recording probability while holding the underlying event process fixed. A forecasting method that interprets the resulting reporting trend as a physical escalation needs either an observation model or a narrower claim about what it predicts: recorded events, rather than all events.

A harmless change and a consequential one

A useful test begins with two versions of the same situation. In one, a delivery is recorded in tonnes; in the other, in kilograms. A correctly transformed physical-dependence calculation should agree. Then change the actual available delivery while holding the units fixed. The instrument should respond if that quantity belongs to the mechanism it claims to measure.

The pair of tests is stronger than either alone. An instrument that ignores every input will pass an invariance test because its output never changes. It will fail the consequential-change test. An instrument that responds to every printed number may detect the real change and also invent a change when only the unit moves. The desired behavior is selective: insensitive to irrelevant representation changes and sensitive to the modeled consequence.

This selectivity can be stated for each input. A duplicated source should not count as independent corroboration. A genuinely independent observation can reduce uncertainty. Reversing the bookkeeping direction of an undirected edge should not alter its Laplacian. Reversing a substantively directed flow can alter the dynamics. The distinction comes from the meaning assigned to the variable, not from a universal rule that all transformations should be ignored.

For a graph example, let P be a permutation matrix that consistently relabels the actors. The transformed Laplacian is $L' = PLP^T$. It has the same eigenvalues because the change is a similarity transformation. Its eigenvectors are correspondingly relabeled. A spectral alert should agree, and an explanation naming a vulnerable actor should follow the

new label mapping. Preserving the eigenvalue while naming the wrong actor is still an implementation failure.

Now remove a real usable connection under the graph's definition. The operator changes substantively. A test that insists the output remain unchanged would protect the wrong property. Every invariance test should therefore identify the equivalence it claims to preserve. "The numbers look similar" is not enough.

One noisy boundary can create many alerts

Suppose a warning is active whenever an observed score falls below 0.1. A stable underlying condition near that boundary can produce observations $0.099, 0.101, 0.098, 0.102$ through measurement noise. The display switches on and off four times even though no consequential change may have occurred.

This is not necessarily a defect in the threshold definition. It may reveal that the measurement precision is inadequate for such a sharp decision. The uncertainty interval should be compared with the boundary. If every observation's plausible interval spans it, the evidence does not support four confident state changes.

An operational system may use hysteresis: enter the warning state below a lower boundary and leave it only above a higher boundary. For example, entry below 0.09 and exit above 0.11 would retain the previous state through the small fluctuations just described. The rule now has memory. Its output depends on the prior state as well as the current score.

That can reduce display chatter, but it changes lead times and episode lengths. It must be evaluated as a distinct decision rule. The mathematical boundary used in a theoretical argument and the operational smoothing used for a display should not be silently conflated. A hysteresis rule does not prove that the world has the same two thresholds.

Similarly, requiring three consecutive adverse readings filters isolated fluctuations but delays entry. If observations arrive weekly, the delay has a different practical meaning than if they arrive hourly. A rule written in numbers of rows should be translated into time before its operational effect is discussed.

The right choice depends on the action attached to the warning. A low-cost request to review the evidence may tolerate frequent entry. An expensive intervention may require a wider margin or additional verification. The same predictive state can legitimately support different action rules when their costs and reversibility differ.

The metric can conceal the limit

In a dynamical model, choosing a norm changes what counts as a large disturbance. Some stable systems admit a quadratic metric that decreases at every instant. That is a valuable stability result. It does not imply that every physical component remains below its own limit.

Consider two quantities scaled so that one unit in each equals its separately stated maximum acceptable level. A state $(1.1, 0)$ violates the first component's limit. If a metric assigns that direction a very small weight, its aggregate norm can still look small. The norm has not repealed the component constraint.

The feasible region should therefore retain explicit operational limits when those limits matter. For normalized components, one possible region is

$$X = \{x : |x_1| \le 1,\ |x_2| \le 1\}.$$

A quadratic energy can be used alongside this region to study dynamics. It should not replace the region unless the relation between the energy bound and component limits has been proved.

For example, $\|x\|_2 \le 1$ implies both component limits, but it excludes some states such as $(0.8, 0.8)$ that satisfy them. It is a conservative inner approximation. Conversely, a ball of radius $\sqrt{2}$ contains the square's corners but also contains $(1.2, 0)$, which violates the first limit. An outer approximation is useful for some bounds and unsafe as a certificate of satisfying every component limit.

These elementary shapes explain a consequential modeling choice. A convenient scalar summary can either overstate or understate the space available for action. Its relation to the actual constraints must be explicit. The framework's geometric language is most useful precisely where it refuses to let convenience erase the boundary.

A changed population can look like a changed law

Suppose a relationship between a predictor and an outcome was fitted mainly on large, well-reported states. The next evaluation contains smaller actors, fragmented authority, or sharply different reporting access. Performance can deteriorate because the population changed, because the measurement changed, because the underlying relationship changed, or because the original fit was too optimistic. Those explanations require different evidence.

A useful diagnostic separates changes in predictor distributions from changes in outcomes conditional on those predictors. The first indicates that the model is being asked to operate in a different part of its input

space. The second suggests that the old mapping may no longer apply even at similar observed inputs. In practice both can occur, and imperfect observation can mimic either.

The model should report when a new case lies outside the supported development range in consequential ways. A ratio may remain numerically defined even when its denominator is far smaller than any training example. A graph may remain connected even when its actor composition is unfamiliar. A valid floating-point result is not a guarantee of empirical transfer.

One response is an explicit transfer test. Freeze the representation and evaluate it in a different region or later period with the same target definition. Another is to restrict the claimed scope until the needed evidence exists. A third is to revise the measurement and then evaluate the revised model separately. Quietly refitting until the new cases look familiar erases the information supplied by the transfer failure.

This also clarifies the role of local expertise. A specialist may identify that an institutional label has changed meaning, that a reported agreement cannot bind the relevant actors, or that a supposedly alternative route shares the same restriction. Such findings alter the model's inputs or structure. They are not merely exceptions to be explained away after a national score is computed.

When the explanation changes after the event

A forecast can be right for the wrong reason. It can also be wrong while identifying a real pressure that another mechanism offsets. Evaluation should preserve both the predicted outcome and the explanation issued with it.

Suppose a model warns because it believes a substitute is unavailable. The outcome occurs, but later evidence shows that the substitute existed and a different mechanism caused the event. The outcome score may count a successful prediction under its rule. The mechanistic explanation has nevertheless failed. Recording only the hit would hide a weakness that matters for the next case.

Conversely, a model may correctly identify a dwindling reserve, after which an unanticipated delivery prevents the predicted service interruption. The forecast score and the causal account need separate examination. The new delivery should enter the trajectory as an observed change, not be retrospectively treated as something the original model knew.

The test becomes clearer when the explanation makes intermediate commitments. Which input was expected to change? Which constraint

was expected to bind? Which observation would have contradicted the account? Those questions can be written before the outcome. A narrative that becomes more specific only after the event is easier to fit to history and harder to falsify.

The failure register should therefore preserve the issued explanation alongside the label. When evidence arrives, assess both. A wrong mechanism with a right outcome calls for a different revision from a correct mechanism overwhelmed by an unmodeled intervention. An unexplained change of story should not inherit the evidential credit of the original warning.

The purpose is to make improvement cumulative. Each failed or partially successful case should leave the research with a sharper variable, a better observation rule, a revised mechanism, or a more precise limit. A model that cannot say what it learned from being wrong will repeat its errors with increasingly polished charts.

A failure register that teaches the model something

Each serious failure should leave a reproducible record: the claim being tested, the smallest relevant input, the expected behavior, the observed behavior, and the scope of the conclusion. A failed invariance test justifies a software or representation repair. A failed prospective prediction may challenge the fitted relationship, the target definition, the observations, or the underlying mechanism. Those are not interchangeable diagnoses.

The most useful failures separate alternatives. If a model predicts the wrong outcome despite correct inputs and an unchanged decision rule, adding more decimal places will not help. If the rule works on archived inputs but the live system supplied the wrong units, the historical theory was not what failed. If a warning led to an intervention, the outcome requires a causal question as well as a forecast score.

A model earns trust by preserving these distinctions. Its failure register becomes a map of where the explanation is precise, where its observations are weak, and where a new mechanism is needed. That is how an ambitious theory becomes an instrument other people can examine and improve.

Chapter 33

What a Warning Has to Predict

An analyst says there is danger. A minister asks how soon. A humanitarian organization asks where and how severe. A researcher asks what observation would count as the warning being right.

Those questions do not delay the forecast. They define it.

The geometric state describes constraints and relationships. To turn that description into a probability forecast, we must specify an outcome, an eligible population, a time window, and an information cutoff. The mapping from a geometric index to a probability is an empirical component of the model. A score between zero and one is not a probability merely because it fits inside that interval.

A denominator full of quiet days

Imagine 10,000 constructed forecasting opportunities. In 100 of them, the specified event occurs. A warning system detects 80 of those events and misses 20. Among the 9,900 no-event opportunities, it issues 990 false alerts and remains quiet 8,910 times.

The sensitivity is 80%. The false-positive rate is 10%. Yet only about 7.5% of the alerts are followed by the event:

$$\text{precision} = \frac{80}{80 + 990} \simeq 0.07477.$$

This is not a paradox. Rare events give false positives a much larger population in which to occur.

If event prevalence is π, sensitivity is s, and false-positive rate is f, then

$$P(\text{event} \mid \text{alert}) = \frac{s\pi}{s\pi + f(1 - \pi)}.$$

The formula makes the denominator visible. A detector's usefulness cannot be inferred from its sensitivity alone.

The same constructed system has 89.9% accuracy. A system that always predicts no event has 99% accuracy and detects none of the events.

Accuracy rewards the common outcome heavily; it is not sufficient for evaluating an early-warning task.

These numbers are teaching examples, separate from the book's 47-conflict research record. A set assembled around historical conflicts can establish event-level detection results under its protocol. Estimating a continuously operating system's false-positive rate additionally requires the eligible no-event opportunities. The denominator is part of the claim.

Probability, ranking, and a decision are different outputs

A ranking orders cases from greater to lesser predicted risk. A calibrated probability has a frequency interpretation across comparable forecasts: among cases assigned probabilities near 0.2, the specified event should occur approximately one-fifth of the time, subject to sampling variation and the chosen population. A decision rule turns the forecast into an action or alert.

One method can rank well and be badly calibrated. Another can give a reasonable overall event frequency while failing to distinguish the most dangerous cases. A third can have useful probabilities but an inappropriate alert threshold for the decision being made. Evaluating all three requires more than one score.

For a binary outcome $Y_i \in \{0, 1\}$ and probability p_i, the Brier loss is

$$\mathrm{BS} = \frac{1}{N} \sum_{i=1}^{N} (p_i - Y_i)^2.$$

If the true conditional probability for a particular case is q, its expected loss is

$$E[(p - Y)^2 \mid q] = q(1 - q) + (p - q)^2.$$

The first term is irreducible uncertainty at that probability. The second is the penalty for reporting a different probability. The expected loss is minimized at $p = q$. This is the defining incentive of a proper scoring rule (Gneiting and Raftery 2007).

For $q = 0.1$, reporting $p = 0.1$ gives expected loss 0.09. Reporting $p = 0.4$ gives 0.18. Confidence has not improved the forecast; it has doubled the expected loss in this example.

The logarithmic loss penalizes confident errors more strongly:

$$\mathrm{LL} = -\frac{1}{N} \sum_{i} \left[Y_i \log p_i + (1 - Y_i) \log(1 - p_i) \right].$$

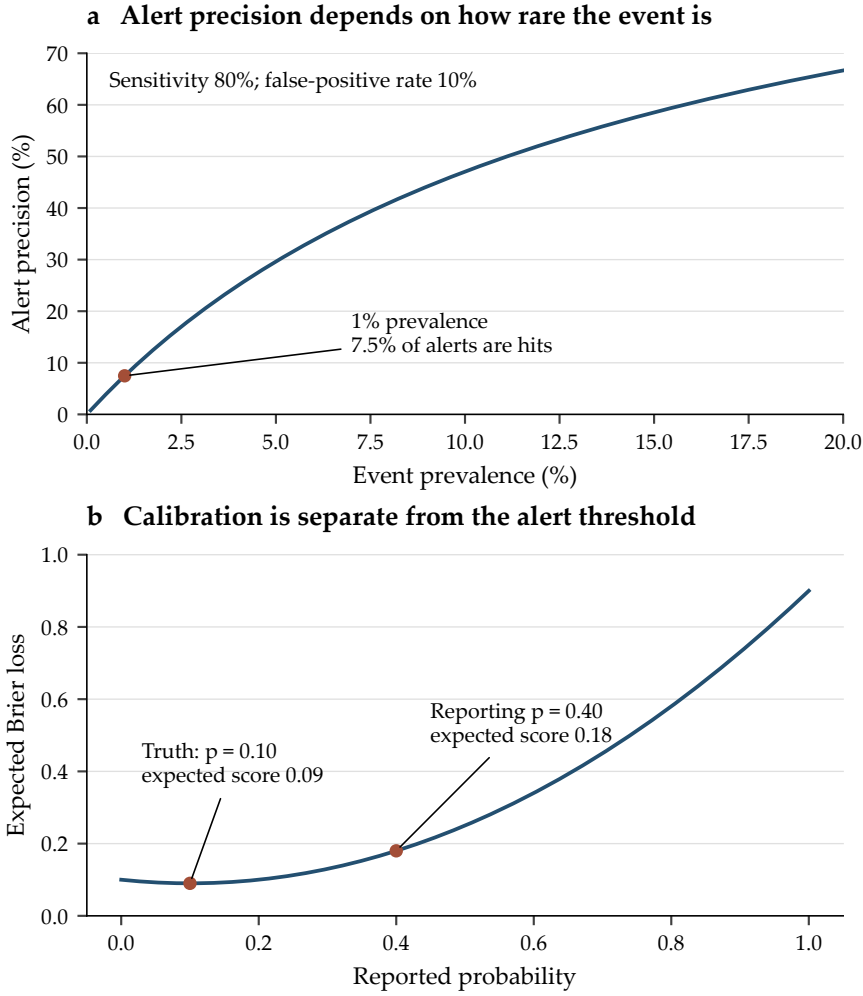


FIGURE 33.1: Constructed probability calculations. Above, the same sensitivity and false-positive rate produce different alert precision as event prevalence changes. Below, Brier loss is minimized by reporting the true probability. These are evaluation principles, not performance estimates for the historical study.

Assigning probability zero to an event that occurs incurs infinite loss under the mathematical definition. If software clips probabilities to avoid numerical infinities, that convention must be recorded. Otherwise different implementations can appear to disagree about the same forecasts.

Calibration should be examined by horizon and by relevant populations, not only in one pooled chart. A method can be well calibrated on a large set of familiar cases and overconfident on a small set undergoing institutional change. Small groups also produce uncertain estimates, so a jagged calibration curve should not be interpreted as a precise map of error.

The clock belongs inside the event

"War eventually" is not a forecast horizon. Neither is a warning that remains active indefinitely until some later event allows it to claim success.

Define a target such as first qualifying onset within the next ninety days among dyads not already in the specified conflict state. Then specify what qualifies as onset, which actor definitions apply, how interrupted or recurring episodes are handled, and what happens when observation ends before the ninety days have elapsed.

A probability of onset within ninety days differs from a daily hazard. The hazard is an instantaneous event rate conditional on remaining at risk. For a single event type with constant rate λ, the survival probability is $e^{-\lambda t}$ and the event probability by time T is

$$P(T_{\text{event}} \le T) = 1 - e^{-\lambda T}.$$

The rate has units of inverse time. The probability does not. Changing the horizon changes the probability even if the rate is unchanged.

Political trajectories can also leave the risk state in different ways. Suppose, in a constructed first-event model, escalation has rate $\lambda_w = 0.01$ per day and a specified settlement transition has rate $\lambda_s = 0.02$ per day. The probability that escalation occurs first and by day 30 is

$$\begin{aligned} P_w(30) &= \int_0^{30} e^{-(\lambda_w+\lambda_s)t}\lambda_w \, dt && (33.1)\\ &= \frac{\lambda_w}{\lambda_w+\lambda_s}\left(1 - e^{-30(\lambda_w+\lambda_s)}\right) && (33.2)\\ &\simeq 0.197810. && (33.3) \end{aligned}$$

Ignoring the competing transition gives $1 - e^{-0.3} \simeq 0.259182$. The difference arises because some trajectories leave the initial risk state through settlement before escalation can be its first exit.

This calculation assumes constant rates and mutually exclusive first exits from the defined state. It does not assume a settlement can never fail later. Later failure requires another state or a renewed risk period. The model's event clock should follow the process it actually claims to represent.

Repeated warnings are not independent discoveries

Suppose a warning remains active for twenty consecutive days before one onset. Counting twenty successful alerts would reward persistence of the display, not twenty independent forecasting achievements. Conversely, treating every day of a long false-warning episode as a wholly independent failure can exaggerate the effective sample size.

Two evaluations can coexist. An issue-level evaluation scores every scheduled probability forecast for its stated horizon. An episode-level evaluation groups consecutive warnings by a fixed rule and asks whether an episode provided useful lead time for a qualifying event. The grouping rule, gap tolerance, and matching window must be chosen before inspecting the outcomes being scored.

Issue-level scores should account for the dependence created by overlapping windows. Uncertainty estimates based on independent daily rows will generally be too optimistic when many rows refer to the same unfolding episode. Resampling whole time blocks, dyads, or episodes can preserve more of that dependence, provided the resampling unit matches the inference being attempted.

Episode matching also needs exclusivity rules. Can one alert claim two separate onsets? Can several alerts claim the same onset? What happens if the actor definition changes during the window? A clear ledger makes those choices visible. A headline hit rate without the matching rules is difficult to interpret or reproduce.

The sample can change what a probability means

A researcher may deliberately collect many conflict cases because those cases contain the events the theory is meant to explain. That can be a sensible design for investigating mechanisms. It changes the meaning of an unadjusted event frequency. Half the cases in a selected study can involve conflict even when conflict is rare among the full population of eligible forecasting opportunities.

Suppose a constructed development sample contains equal numbers of event and no-event cases, while the intended operating population has event prevalence one percent. A classifier trained to reproduce the sample probability does not automatically produce the population probability. The sampling design has changed the prior odds.

Under the specific assumption that sampling depends on the outcome but, within each outcome class, not on the predictors, an odds adjustment is available. Let q be the calibrated sample probability, π_s the sample prevalence, and π the population prevalence. Then the population odds satisfy

$$\frac{p}{1-p} = \frac{q}{1-q}\,\frac{\pi/(1-\pi)}{\pi_s/(1-\pi_s)}.$$

The formula follows by writing the likelihood ratio of the predictors within the two outcome classes and changing the prior odds. Case-control sampling research develops the conditions under which such retrospective sampling preserves parts of a logistic model (Prentice and Pyke 1979).

With $\pi_s = 0.5$, $\pi = 0.01$, and $q = 0.5$, the adjusted probability is 0.01. With $q = 0.8$, it is $4/103 \simeq 0.0388$. A strong-looking probability in a balanced sample can correspond to a much smaller operating probability because the quiet population is much larger.

The adjustment is not a repair for arbitrary case selection. Famous crises may differ from less visible crises in observation quality, actor prominence, severity, and other predictors. If selection depends on those features within an outcome class, a prior-odds adjustment alone is insufficient. The study must model or constrain that selection, or limit the claim to the sampled population.

This distinction protects the value of the historical conflict record. A collection of conflicts can provide substantial evidence about mechanisms, timing within episodes, and event-level detection under its specified protocol. Estimating calibrated daily or monthly probabilities across all eligible dyads requires the corresponding risk population and observation schedule. Those are additional objects, not a dismissal of the work already done.

The essential question is what conclusion the design supports. A model can be impressive at identifying a structural pattern within known crises and still need a separate study of how often that pattern appears without the event. The second study adds the denominator required for a different claim.

Calibration can hide an important difference

Imagine two equally common kinds of forecasting opportunity. In the first, the true event probability is 0.1; in the second, it is 0.3. A model assigns 0.2 to every case. In the pooled population, events occur twenty percent of the time, so the single forecast value is calibrated overall. The model nevertheless fails to distinguish the two kinds of case.

Its expected Brier loss is $0.2(1 - 0.2) = 0.16$. A model that correctly reports 0.1 and 0.3 for the two groups has expected loss

$$\frac{0.1(0.9) + 0.3(0.7)}{2} = 0.15.$$

The gain is modest but real: the second model has useful resolution while remaining calibrated. The first is a good statement of the overall base rate and an incomplete account of differences between cases.

Now give the two groups names that matter operationally: familiar institutions and newly reorganized institutions, or places with stable reporting and places where reporting has changed. A pooled calibration chart can conceal a systematic error in the smaller or less familiar group. The response is to examine relevant subgroups with uncertainty, not to expect perfect empirical agreement in every small bin.

Small samples create their own illusion. If five cases receive probability 0.2, observing zero or two events does not by itself show serious miscalibration. The outcome count has substantial sampling variation. A calibration curve should include the number of forecasts supporting each region and an uncertainty analysis appropriate to dependence among them.

There is also a distinction between recalibrating a model and evaluating it. A mapping fitted to make the current test set's probabilities agree with its outcomes is a fitted description of that set. It is not a test of the revised mapping on unseen outcomes. A fair process fits calibration on an earlier validation period and then freezes it for the later evaluation period.

This matters for a geometric index. A monotone mapping can convert an index into probabilities only after it has been fitted against a defined event population. Its calibration may change when the population, horizon, measurement system, or model version changes. The mapping should travel with those specifications, rather than being treated as a permanent translation table from geometry to likelihood.

Ninety days must contain thirty days

At one issuance time, the event "first qualifying onset within thirty days" is contained within "first qualifying onset within ninety days," provided the actor set, onset definition, and risk state are the same. Their probabilities must therefore satisfy

$$P(T \le 30) \le P(T \le 90).$$

A model reporting 0.3 for thirty days and 0.2 for ninety has violated a logical relationship between its targets. Two separately fitted classifiers can produce this inconsistency even when each was optimized correctly for its own loss.

One remedy is to model nonnegative interval hazards and derive cumulative event probabilities. If h_j is the conditional probability of onset in interval j given survival to its start, then

$$P(T \le t_m) = 1 - \prod_{j=1}^{m} (1 - h_j).$$

Every factor lies in $[0, 1]$, so adding an interval cannot lower the cumulative probability. This construction encodes the nesting directly. It still needs calibrated hazards, correct handling of censoring, and an appropriate event-state definition.

For example, conditional probabilities 0.05, 0.08, and 0.10 in three successive intervals give cumulative probability

$$1 - (0.95)(0.92)(0.90) = 0.2134.$$

Adding the three hazards gives 0.23, which is not exact because later hazards apply only to cases that remained event-free. The distinction becomes more consequential as the hazards grow.

The comparison changes if the forecasts were issued at different times. A ninety-day probability issued last month can be lower than a thirty-day probability issued today because new information arrived. That is not a nesting violation. The issuance time is part of the probability statement, just as the horizon is.

It also changes if the targets differ. Probability of any organized violence, probability of a new interstate war, and probability of a fatality threshold are not nested merely because their labels sound similar. A valid consistency check must use the actual definitions rather than an informal hierarchy of severity.

A forecast can remain unresolved

At the end of a study, some issued forecasts have not yet reached the end of their horizons. Their outcomes are unresolved. Coding them as no-event cases rewards long-horizon forecasts for events that have not yet had time to occur.

Suppose a ninety-day forecast is issued twenty days before the data cutoff and no event has yet been observed. The record supports survival through the observed twenty days under the available outcome reporting. It does not support a completed ninety-day no-event label. A survival analysis can retain right-censoring; a fixed-horizon binary evaluation can exclude the row until it matures under a predefined rule. Either choice should be explicit.

Outcome reporting creates another delay. A horizon may have elapsed while reports about the last part of it remain incomplete. A scoring protocol should state the release vintage used and whether scores will later be revised. A mature event window and a mature outcome dataset are different conditions.

The same principle applies to uncertain onset dates. If an event is known to fall within an interval crossing the forecast boundary, the hit assignment may be unresolved. One can report lower and upper scoring bounds across admissible dates, or use a model that represents interval-censored timing. Assigning the most favorable date to the preferred model would turn uncertainty into an advantage chosen after the fact.

These details are not bookkeeping at the edge of the research. They determine which forecasts enter the denominator and which outcomes count as successes. A beautifully calibrated formula cannot compensate for a scoring population that quietly changes with the desired result.

The cost of remaining silent

A model may decline to issue a forecast when inputs are missing or incompatible with its specification. That can be a responsible operational behavior. It creates a second evaluation problem: how often, and in which circumstances, does the system provide an answer?

Suppose model A supplies forecasts for all one thousand eligible opportunities. Model B supplies forecasts for six hundred and omits the most difficult four hundred. Comparing their average loss only on their respective supplied rows changes the population. The difference can reflect selection as much as predictive quality.

A common-case comparison evaluates both models on the same six hundred rows and reports that scope. A complete-system comparison also accounts for the remaining opportunities using a predefined fallback, explicit abstention cost, or separate coverage analysis. Neither approach should silently treat an absent forecast as a correct no-event prediction.

Coverage should be broken down by the conditions that cause refusal. A system that works reliably in well-observed peaceful periods and becomes silent during reporting collapse may be least available when a warning is needed. That is an operational limitation even if its issued forecasts score well. Conversely, a refusal that prevents a misleading number can be valuable when paired with a useful statement of what evidence is missing.

The best abstention output is specific: the actor mapping is unresolved, the required baseline is undefined, the supplied operator belongs to a different graph, or the observation is too old under the protocol. Such a message identifies a repair. A blank cell does not.

The warning system therefore has two obligations. When it speaks, its probability must have a defined event, population, and clock. When it cannot speak under its specification, its silence must remain visible in the evaluation. Together these obligations make a forecast something another person can use, challenge, and learn from.

A threshold expresses a cost model

For a simple classification decision, let C_{FP} be the cost of an unnecessary alert and C_{FN} the cost of failing to alert before an event. If correct decisions have zero cost under this simplified model, issuing an alert is preferred when

$$C_{FP}(1-p) < C_{FN}p, \qquad p > \frac{C_{FP}}{C_{FP}+C_{FN}}.$$

The threshold follows from the stated costs. It is not a property of the probability model alone.

An actual intervention is more complicated because it may change the outcome. The classification calculation assumes the event is being predicted, not prevented by the decision itself. Chapter 37 develops that distinction. It is especially important when the costs of an alert include public fear, scarce attention, or actions that alter the relationship being modeled.

The most useful report therefore keeps several layers visible: the geometric assessment, the probability mapping if one has been validated,

the forecast horizon, the uncertainty, and the decision rule. A reader should be able to tell whether two analysts disagree about the state of the world, the likelihood of an event, or what ought to be done about it.

That separation makes the framework more useful. It gives the mathematics a defined prediction to support and gives the decision-maker a clear account of what the prediction does and does not settle.

Chapter 34

From a Source Record to a Warning

A number on a screen has a history. Someone observed something, someone recorded it, a source released it, a program transformed it, and a rule assigned it meaning. A reproducible warning keeps that history attached.

This chapter follows one constructed record through the process. The arithmetic uses the inspected Psychohistory implementation and an explicitly supplied baseline. The surrounding method shows what must be preserved if another researcher is to reconstruct the same result. It is an implementation example, not a newly recovered historical forecast.

Start with the record, not the dashboard

For each observation, retain the actor identifiers, the field name, the value, the unit, the period or instant described, the time the source became available, the source version, and the source reference. A source file's download time is useful too, but it does not automatically establish when its contents first became public.

Consider two records for the same constructed trade observation. The first reports 50 billion dollars, describes January 1, and becomes available January 5. A revised value of 80 billion dollars becomes available February 5. A forecast issued January 10 can use the first release. A reconstruction of the best later account of January may use the revision. Those are different views of the same observation history.

The companion selector preserves both and takes the latest eligible release for the relevant observation at a chosen cutoff. It returns 50 for January 10 and 80 for February 10. Before January 5 it returns no observation. It does not replace the missing value with zero.

For a point observation r, eligibility requires

$$t_{\text{observed}}(r) \le t_{\text{issue}}, \qquad t_{\text{available}}(r) \le t_{\text{issue}}.$$

Annual and monthly series need additional rules for period aggregation and release timing. An annual total is not a daily flow merely because it has been copied into a daily table. A forecast of a future quantity is also a different record type from an observation. The selector should not blur them.

Units need equally explicit treatment. In this engine, energy trade is supplied in monetary units relative to a monetary threshold. The field name "energy dependency" does not authorize inserting a percentage, tonnes of oil, or electrical power into the same slot. Those can be useful alternative measurements, but they require an explicit transformation and a correspondingly defined model version.

Freeze the baseline before the forecast

A standardized score describes a deviation from a reference distribution. The reference must be known. Store its fitting interval, eligible observations, mean, scale, transformation, and effective version.

The worked example supplies the baseline rather than relying on a convenience default. For helicity it uses mean 1.25 and standard deviation 0.1; for flux it uses mean three and standard deviation one. These are constructed teaching values. They do not represent a fitted Sudanese, Russian, or global history.

If the scale is zero, the ordinary standardized score is undefined. A wrapper should refuse the calculation or apply a separately specified alternative. Silently dividing by an arbitrary floor changes the interpretation. If a floor is part of the approved method, it belongs in the versioned specification and in the output record.

An expanding-window study can update its baseline as new observations become available. Each update should produce a new effective version. The earlier forecast keeps the baseline it actually used. Recomputing every old score with the latest baseline may be useful for a retrospective descriptive chart, but it is not the original sequence of forecasts.

The same rule applies to probability calibration and threshold selection. If the research chooses an alert threshold using a validation interval, freeze that choice before the next test interval. Repeatedly inspecting the test interval and adjusting the threshold converts it into another development set.

Follow one calculation all the way through

The constructed dyad has trade volume 50 and energy trade ten, with reference quantities 50 and ten. Voting similarity is 0.4, treaty overlap 0.2, and sanction intensity one. All four tolerance inputs equal 0.5. There are two modeled diplomatic constraints and four modeled exits. The supplied ordinary-branch gap index is 0.2.

The flux calculation gives

$$V = \frac{50}{50} + 4\frac{10}{10} = 5.$$

The alignment transformation gives approximately -0.581179. The tolerance weights sum to one, so the common tolerance input produces $\tau = 0.5$. The integration calculation gives $K = 0.485$. This is above the disconnection gate, so the ordinary branch applies.

With the implementation's regularized denominator, the helicity value is approximately 1.030926, and the trichotomy value is approximately 1.487312. The supplied baseline gives

$$Z_H \simeq -2.190743, \qquad Z_V = 2.$$

The adverse-alignment and standardized-divergence condition fires, producing the engine's `SINGULARITY` alert. The trace must identify that condition. Explaining the alert solely by comparing the trichotomy value with a threshold would misdescribe the rule that actually determined the output.

The complete rule ordering is retained in Appendix B. This example makes one operational point: when several conditions can produce a label, the output should report the condition that fired, its inputs, and its precedence. A label alone is insufficient for explanation.

The supplied gap index deserves particular attention. The ordinary dyad routine does not infer a global network spectrum from this pair's trade record. It consumes the value supplied on the model instance. A complete application must therefore identify the graph, operator, normalization, and observation time behind that supplied index. Otherwise a default setting can masquerade as a measured global property.

Validate before calling the engine

The companion wrapper checks that quantities expected to be nonnegative are finite and nonnegative, that reference denominators are positive, that bounded component scores lie in their permitted ranges, and that

baseline scales are positive. For this fixture it also requires at least one diplomatic constraint and two exits, matching the ordinary logarithmic expression being demonstrated.

These checks are necessary but not sufficient. They cannot determine whether a source supports a value, whether an actor identifier refers to the correct organization, or whether a release date has been reconstructed honestly. Those are provenance checks, requiring records beyond the numerical array.

A useful division of responsibilities is straightforward. The ingestion layer preserves source records. The transformation layer produces documented model inputs. The numerical layer evaluates the specified formulas. The decision layer applies the versioned alert rules. The publication layer presents the result with its time, meaning, and uncertainty. Each layer should be able to explain what it received and what it changed.

Separating these responsibilities makes a failure easier to locate. If an input unit is wrong, the formula may still have been evaluated correctly. If the input and formula are correct but a downstream label uses the wrong branch, the interpretation failed. If the entire computation is reproducible and the prediction still fails prospectively, the research must confront a model or measurement limitation rather than searching indefinitely for a software excuse.

A replayable forecast record

The minimum useful forecast record includes a unique issuance identifier, the actor pair or spatial unit, the target definition, the issue time, the horizon, the model version, the input snapshot, the baseline version, the output branch, the intermediate quantities, the alert or probability, and the reason for any refusal to issue a result.

A hash of an input file can help establish whether the bytes later changed. It does not certify that the source was true or that the file was available at the claimed time. Provenance and integrity are complementary. A false observation can be preserved perfectly.

The companion records the hash of the inspected engine file with its numerical results. A changed engine file therefore produces a visibly different implementation record. It also keeps the teaching observations and baseline beside the outputs. Another researcher can rerun the example and inspect whether disagreement arose in the source, the environment, or the calculation.

Human intervention belongs in the record as well. If an analyst overrides a warning, record the original output, the override, the time,

and the stated reason. Do not rewrite the model's earlier result to agree with the later decision. The distinction is needed both to improve the model and to understand the operational process around it.

An outcome record should be linked later, with its own definition and release history. It should not be written into the original forecast row as though it had always been there. This preserves the boundary between what was predicted and what was subsequently learned.

Build the record so that a disagreement can be located

Imagine two analysts reproducing a warning. They agree on the equations but obtain different labels. One used the latest revised trade release. The other used the release available at the forecast date. One treated a missing diplomatic observation as neutral. The other refused the calculation. One supplied a regional normalized gap. The other supplied a dyad-level proxy with the same column name. Each difference can survive a perfectly functioning numerical routine.

The solution is to give every transformation an explicit input and output meaning. A source record states what was reported. A normalized observation states how units and identities were reconciled. A model input states how an observation became one of the engine's fields. A forecast record states which version of the rule consumed those inputs. Keeping those stages distinct makes disagreement diagnosable.

This is the practical value of provenance. The W3C PROV model distinguishes entities, activities, and responsible agents so that the origin and transformation of information can be represented (Moreau and Missier 2013). A book companion need not implement every feature of that standard. It needs enough of the idea to answer who or what produced a value, from which records, by which operation, and when.

For example, an input called voting similarity should point to the set of votes, the treatment of abstentions and absences, the comparison rule, and the period. A number copied from an analyst's notebook may still be useful, but its status differs from a value rebuilt from a retained vote ledger. The implementation should describe that difference instead of giving both the same appearance of direct measurement.

A source pointer should be specific enough to locate the supporting material. A general organization homepage is rarely sufficient for a dated claim. Use the report title, release date or version, page or table where appropriate, and a stable identifier or archived copy when available. The reference list tells the reader what the source is; the provenance record tells the reproducer which part supplied the input.

An unavailable source does not require deleting the historical claim. It requires stating what kind of evidence remains: a retained aggregate, an earlier report, a calculation notebook, or a reconstructible underlying ledger. Those levels support different reproduction tasks. Preserving the distinction makes it possible to recover more evidence later without rewriting what was known at the time.

A join can manufacture a relationship

Combining datasets is one of the places where an apparently routine operation can alter the model. Two files may use different names for the same actor, the same name for different actors, or different boundaries for the same place. Joining rows by a convenient text label can create a relationship no source actually reported.

The actor table should use stable identifiers and dated names. A name change does not necessarily create a new actor. A split, merger, or change in the relevant decision-making authority may. The rule depends on the task. A legal-state register, an armed-organization register, and a negotiation-participant register can require different identity histories.

Before merging two observation tables, specify the expected relationship between keys. If one trade total is supposed to match one actor-pair period, the joined result should contain one such row. If a name table contains three aliases and the join multiplies the trade row three times, summing afterward triples the flow. The program may have performed a valid many-to-many join on an invalid key.

This failure can be tested with conservation checks. A unit conversion should change units while preserving the represented quantity. An identity lookup should not change the total of an already aggregated flow unless the documented mapping intentionally redistributes it. A disaggregation should reconcile to its parent total when that is part of the definition. These properties catch errors that a spot check of a few plausible rows can miss.

There is a similar issue with territorial changes. A historical total defined for an earlier boundary cannot be assigned to every successor area as if each inherited the full amount. A mapping may allocate it under a documented rule, retain the earlier boundary, or mark the comparison unavailable. The choice affects both the graph and its interpretation.

An implementation should therefore produce a short reconciliation report before calculating the model: unmatched identities, duplicated keys, redistributed totals, changed units, and unresolved periods. This report is not a decorative quality score. It identifies the exact records that could alter the result.

Time needs more than one column

The earlier release example distinguished observation time from availability time. A production register often needs one further distinction: when the system learned of or stored the record. A report may have been publicly available on January 5 but not collected by the particular operational system until January 8. A January 6 information-availability study and a January 6 replay of that actual system then have different eligible inputs.

Neither task is illegitimate. The first asks what could have been known under the defined public-source protocol. The second asks what this system actually knew. A comparison must name which one it performs. Granting one model every public release immediately while evaluating another on its real collection delays mixes research potential with operational performance.

Retain the period described, the source release time, the acquisition time, and any later revision time. When exact timestamps are unavailable, retain the known precision. A date known only to the month should not become an invented midnight on the first day simply to satisfy a database field.

The query that builds a forecast snapshot should be deterministic. Given the same retained records, cutoff, and selection rule, it should choose the same source versions in the same order. Conflicting releases with identical timestamps need a documented tie-breaker or an unresolved status. The database's accidental row order is not a defensible rule.

A useful replay check asks whether adding a later revision changes an earlier snapshot. It should not, unless the stated task is a retrospective reconstruction that intentionally uses later information. Another asks whether reordering source rows changes the result. It should not when record identity and precedence are already explicit. These are meaningful tests because they expose hidden dependence on storage rather than evidence.

The same record can support two views without being overwritten. One view shows the best-supported current history. The other shows what was available at each prior issuance. The first is useful for historical explanation; the second is indispensable for forecast evaluation. Maintaining both is more informative than arguing that one must replace the other.

An absent input is a state of knowledge

Missing is not a numerical value. Zero trade, unreported trade, an inapplicable field, and a source withheld from release can all appear as an empty cell in a spreadsheet. Their meanings differ. A model that replaces every empty cell by zero turns uncertainty into a statement about the world.

Start with a reason code. The value may be not yet released, unavailable under the source's scope, disputed, invalid under a unit check, or inapplicable to the actor type. The decision about whether and how to calculate should depend on that reason and the model specification.

Carrying a value forward is an assumption about persistence. If a monthly observation is used for daily decisions, the record should retain its age. A value three days old and one three months old may be numerically equal but carry different information. A stale value should not acquire a fresh observation date each time it is copied into a new input row.

Imputation is also a model. If a missing quantity is filled using other variables, the fitted imputation procedure must use only the permitted training information and preserve uncertainty where it matters. Filling the entire historical table before splitting into training and test periods can leak later relationships into earlier forecasts even when the final forecasting model respects the split.

For a bounded input, an interval calculation may be more informative than a single imputation. Run the decision rule over the admissible range while holding the other specified inputs fixed. If every value yields the same label, the missing observation may not prevent that classification. If labels differ, report the unresolved condition. The interval must be justified by the variable's meaning; a broad numerical range is not automatically an evidence-based uncertainty set.

Sometimes the correct behavior is refusal. A required baseline with zero scale, a negative capacity entered into a nonnegative graph, or an undefined logarithm should not be hidden behind a plausible default. A refusal should identify the offending field and the operation that cannot be performed. That turns a failed calculation into a specific request for repair.

Publish the result and the route to it

A compact publication record can contain the warning, its issue time, target, horizon, and interpretation. A fuller companion record should preserve the selected inputs, their sources, the transformation versions,

the model configuration, intermediate quantities, and the branch that determined the result. Readers need different levels of detail, but the concise view should lead to the fuller one.

For a graph-derived input, retain the node ordering and the edge definition. The same matrix printed with a different actor order can be mathematically similar while its entries refer to different relationships. A permutation should preserve spectral quantities but change the correspondence between coordinates and actors in a controlled way. A test that permutes both matrix and labels can verify that the output interpretation follows the labels correctly.

Store the normalization and units with the graph. "Gap = 0.05" is incomplete without knowing whether the operator is combinatorial, degree-normalized, rescaled by another quantity, or supplied as an engine index. The number's place in a rule depends on that definition. A later change to normalization creates a new specification even if the output column keeps its familiar name.

A numerical release should also preserve the computational environment sufficiently to reproduce consequential operations. Exact byte-for-byte agreement may be inappropriate across all linear-algebra libraries, but differences should remain within justified tolerances and should not silently cross an alert boundary. The test should focus on the mathematical property and decision margin, not an arbitrary number of matching decimal places.

The result archive should be append-only in its historical meaning. A later correction can add a superseding record and an explanation in the research log while leaving the original issuance available for evaluation. The published book's prose can present the corrected mathematics directly; the research archive has a different purpose, preserving what a system actually issued and when.

This distinction prevents a common confusion. Improving today's model is desirable. Replacing yesterday's forecast with today's improved reconstruction destroys the evidence needed to know whether the model improved. A reliable research program needs both a current best account and an intact history of its prior decisions.

A release is ready when its failure is interpretable

Before a new model version is used, replay a small set of cases selected for their mechanisms: an ordinary valid input, a disconnection boundary, a missing required observation, a later revision, a unit conversion, an actor-order permutation, and a value close to a decision threshold. The

point is not to accumulate a large test count. It is to cover the places where the meaning or branch changes.

Then compare the new version with the previous one on a fixed historical snapshot. Differences should have reasons. A changed formula should change the relevant intermediate value. A changed identity mapping should change the affected actor records. An unchanged case should not move because a file happened to be read in a different order.

This process also makes empirical failure more informative. If the archived inputs, transformations, and decision rule are all recoverable, a missed event can be investigated as a scientific question. Was the relevant relation absent from the representation? Did the observation arrive too late? Did the mapping from structural state to outcome fail? Did an intervention change the trajectory? The archive permits those explanations to be distinguished.

The standard is demanding but finite. Another person should be able to reproduce the calculation, identify its assumptions, locate its evidence, and understand a disagreement without relying on the original author's memory. That is how a research framework becomes a method others can actually use.

What reproduction establishes

The companion tests several concrete properties: the ordinary worked ledger, dispatch into the disconnection branch, invariance under consistent monetary rescaling, exclusion of a later revision from an earlier information set, and refusal of a zero-scale baseline. The broader mathematical companion checks the spectral, dynamical, flow, and probability examples used in the expansion.

Passing these checks establishes that the stated constructed examples can be reproduced and that the tested implementation properties hold for the inspected version. It does not reconstruct the complete historical event and alert ledger. That separate evidentiary task requires the underlying records specified in Appendix E.

This boundary protects the value of both kinds of work. Numerical reproducibility matters because it makes the instrument inspectable. Historical reproducibility matters because it connects the instrument to the claims made about the world. A serious research program needs both, and it should make clear which one a particular artifact supplies.

The result is a warning whose history can be followed backward: from the label to the rule, from the rule to the computed state, from the state to the transformed inputs, and from those inputs to the dated observations. That chain is what makes an explanation usable by someone who did not build the system.

Chapter 35

Put the Models on the Same Starting Line

One model predicts the number of violent events in a province next month. Another ranks country pairs by the possibility of a new war. A third estimates battle deaths over the next three years. Their published scores can all be correct and still tell us nothing about which model would win a common forecasting task.

A head-to-head comparison begins by making the task common. The event, geography, time window, available information, and scoring rule must match before a ranking has a clear meaning.

That requirement applies to this book as strongly as to any competing method. The Davis framework should be tested against demanding alternatives, including simple ones. Its value cannot be established by comparing its best retrospective result with another system's score on a different target.

What current systems are trying to predict

At the November 7, 2026 documentation cutoff used for this comparison, VIEWS describes monthly forecasts of state-based conflict fatalities and related occurrence probabilities, at country and subnational grid levels, with horizons extending from one to thirty-six months (VIEWS 2026a). A forecast for an already violent location can concern the continuation or intensity of violence. That is not automatically a forecast of a new interstate war.

ACLED's current CAST methodology describes forecasts of organized-violence event counts for six forthcoming rolling four-week periods. It uses machine learning, including LightGBM, and reconciles predictions across administrative, national, and global levels and across event categories. Its documented inputs include recent violence, actor activity, and other contextual indicators (ACLED 2026b). Reconciliation addresses a practical consistency problem: provincial predictions and event categories should add up to the corresponding totals.

The VIEWS 2023/24 prediction challenge required distributions over future fatality counts and specified a common evaluation, with the continuous ranked probability score as its primary measure (VIEWS Prediction Challenge Collaboration 2024). That design provides a useful model of a competition in which uncertainty and the forecast target are part of the entry requirements.

These systems should be compared in their documented forms. An older monthly CAST specification and a newer rolling-period specification are different versions. A VIEWS research challenge and a regular operational release are different artifacts. A model name is not a sufficient version identifier.

The geometric approach developed here contributes explicit state variables and constraints: integration, tolerance, alignment, available exits, network structure, sequence, and the distinction between feasible and merely named alternatives. To enter a probability competition, it also needs a specified mapping from those states to the common target. The engine's alert labels do not by themselves provide a full distribution over fatalities.

Begin with an opponent that is easy to understand

A constant-risk baseline predicts the training period's event frequency for every eligible case. A persistence baseline uses recent observed activity to predict future activity. A simple logistic or hazard model combines a modest set of predictors through an explicit fitted relationship. These baselines are necessary competitors, not ceremonial additions to make a more elaborate method look impressive.

For a binary target, a logistic model can be written as

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \beta^T x_i)}}.$$

The inputs x_i might include lagged event history and structural measures available at issuance. The coefficients and any regularization are fitted on the training data. The model is compact enough to inspect, but its interpretation still depends on the variables, sampling population, and temporal design. A transparent equation can be fitted with leaked information just as easily as a complex model.

Tree ensembles can represent nonlinear interactions and thresholds without requiring a single linear predictor. They may learn that a variable matters only within a particular range of another variable. That flexibility is valuable when the training evidence supports it. It also increases the

importance of time-aware validation, because repeated patterns within one crisis can otherwise appear on both sides of a random split.

A geometric model makes a different allocation of structure. It specifies some relationships before fitting: a ratio, a constrained set, a graph operator, or a composition rule. This can reduce the space of possible explanations and make certain failures easier to diagnose. It can also impose the wrong structure. The comparison should test the advantage, not assume it.

The most informative competition therefore includes both a standalone geometric model and models augmented by geometric features. If a conventional forecasting model improves after receiving the framework's variables under the same training protocol, that is evidence that the representation contributes information even if the best predictor is a hybrid.

Approach	Main comparison question	Failure to examine
Constant or persistence baseline	Is the added machinery useful beyond prevalence or recent history?	Abrupt transitions and rare onsets
Logistic or hazard model	Do explicit predictors explain risk under a compact fitted rule?	Misspecified interactions and changing relationships
Tree or other flexible ensemble	Does learned nonlinear structure improve held-out forecasts?	Leakage, unstable extrapolation, and calibration
Geometric model	Do constrained states and relations add predictive information?	Wrong operator, weak measurements, and untested mappings
Hybrid model	Does adding geometric information improve the same base learner?	Extra flexibility mistaken for added information

Table 35.1: Method families for a common benchmark. These are proposed comparison roles, not a performance ranking or a claim that all named operational systems use every listed technique.

Two comparisons, kept distinct

There are two legitimate ways to compare systems. One asks how complete operational systems perform with the information and processes

they actually use. The other asks which modeling technique performs better when given the same information and development budget.

The first comparison captures the value of data collection, analyst work, release schedules, and engineering alongside the algorithm. It should use archived forecasts issued before the outcomes. The second controls more of those differences so that the contribution of the representation or learner becomes clearer. It should use a common data snapshot, train/test schedule, and tuning protocol.

An operational system with richer information may outperform a smaller model for good reasons. That does not show its learning architecture would win with identical inputs. Conversely, a carefully controlled research model may win a common-data benchmark while lacking the infrastructure needed to issue reliable live forecasts. Both conclusions matter; neither should be substituted for the other.

For the geometric contribution, a useful sequence is to fit the same baseline learner first with conventional inputs, then with geometric inputs, then with both. Keep the training windows, target, and tuning allowance fixed. Include a control that adds a comparable number of generic transformed inputs, so that any gain is not automatically attributed to geometry when extra flexibility could explain it.

The resulting ablation asks which information matters. Remove sequence while retaining totals. Replace a layered graph with its aggregate. Remove the observation-age indicator. Replace an independently measured variable with the implemented proxy. Refit each permitted model on the training set, then evaluate the frozen variants on the same held-out cases. The comparison can reveal which part of the explanation carries predictive value.

Write the contest before looking at the finish

A common onset benchmark should specify the eligible actor pairs, the exclusion of already active episodes, the event definition, issuance schedule, horizon, information cutoff, and outcome release. A count benchmark should specify the event or fatality measure, aggregation unit, forecast distribution, and treatment of revisions. Those are separate contests unless an explicitly justified transformation makes their targets common.

Use rolling temporal evaluation. Fit on an earlier period, choose settings on a later validation period, and score on a subsequent untouched period. Advance the boundaries according to a fixed schedule. A geographic holdout can ask whether a method transfers to places not used for fitting. In the African extensions, this is especially useful for

testing whether a result depends on region-specific observation practices or actor definitions.

The test population must include quiet cases. If the benchmark contains only famous crises, it cannot measure the burden of a model that issues warnings everywhere. The eligible population should be constructed independently of whether an eventual event made a case interesting to the researcher.

All methods should face the same missing-data policy or have their different policies recorded as part of the comparison. If one system declines to issue a forecast, report coverage as well as performance. Quietly removing its difficult cases can make abstention look like superior accuracy. A scoring contract can define a penalty, a fallback forecast, or a separate coverage analysis; the choice must be explicit.

The primary metric should match the output. Binary probabilities can be assessed with Brier or logarithmic loss, alongside calibration and precision–recall behavior. A full distribution over counts can be assessed with a proper distributional score. Alert episodes need their own lead-time and matching analysis. One scalar score cannot answer every operational question.

Pair the errors on the same forecast opportunities. If model A and model B both predict the same case, examine their difference in loss on that case. Estimate uncertainty using blocks that preserve relevant temporal and actor dependence. A small average improvement concentrated in one extraordinary episode should be reported differently from a consistent improvement across many independent episodes.

One table that every model must face

Before training begins, construct the table of eligible forecast opportunities. A row identifies the unit, issuance time, target window, and outcome definition. The input snapshot is attached by its information cutoff. The eventual outcome is attached later under the scoring protocol. Every entrant must address the same rows or disclose where it cannot.

This table is the scientific object around which the comparison is organized. Without it, a collection of published scores is a collection of answers to different questions. The common table makes it possible to ask whether a method adds information, which situations it handles well, and which failures recur across approaches.

For an onset task, the eligibility rule must distinguish a unit already experiencing the target conflict from one at risk of a new onset. For a recurrence task, the required period without qualifying violence must be defined. For a count task, continuing violence remains part of the

target rather than an exclusion. These are substantive differences in the forecasting problem.

The geographic unit also matters. A country-month forecast cannot be compared directly with a grid-cell-month forecast by averaging arbitrary cell probabilities. The probability of at least one event in a country depends on dependence among the cells. If cell events were independent, the country probability would be $1 - \prod_j (1 - p_j)$. Shared shocks make independence a serious assumption. A national total of expected counts is additive; a probability of any event is not.

The comparison should therefore use a common target at a supported resolution, or define and validate the aggregation procedure. It should not convert outputs until they look similar and then assume the conversion preserved their meaning.

There are also differences between the available information sets. A model using local event reporting has information a trade-only model does not. A model using a revised annual dataset has information a contemporaneous monthly release may not have contained. A common-data experiment holds those differences fixed as far as possible; an operational comparison measures them as part of the complete system. The result must say which advantage it is evaluating.

A hand-checkable head-to-head result

The scoring arithmetic should be simple enough to verify before it is trusted on thousands of rows. Take four constructed binary opportunities with outcomes $0, 0, 1, 1$. Model A assigns probabilities $0.1, 0.2, 0.6, 0.8$; model B assigns $0.2, 0.4, 0.5, 0.7$.

Outcome	A	B	A squared loss	B squared loss
0	0.1	0.2	0.01	0.04
0	0.2	0.4	0.04	0.16
1	0.6	0.5	0.16	0.25
1	0.8	0.7	0.04	0.09

Table 35.2: Constructed forecasts used to inspect the scoring rule. A and B are illustrative arrays, not named research systems.

The mean Brier losses are 0.0625 and 0.135. The average paired improvement of A over B is 0.0725. The calculation uses the same outcome on each row. If the models had been evaluated on different populations, subtracting their reported averages would not isolate the effect of their different predictions.

Four rows are enough to check arithmetic and far too few to establish broad comparative performance. If they were four forecasts about one unfolding crisis, they would also provide much less independent evidence than four unrelated episodes. The worked result belongs in the numerical companion as a fixture, not in a claim that geometry has defeated an operational competitor.

The example gives the reader a practical audit. Sum the losses, verify the denominator, and compare rows before interpreting an average. A single confidently wrong forecast can matter greatly under logarithmic loss. A method can also win under one scoring rule and lose under another. That is why the primary rule must be selected for the intended output before seeing the contest's result.

Ablation has to remove the proposed information

An ablation compares a model with and without a specified component. To test the contribution of geometric structure, the removed component must actually be the information or constraint under discussion. Deleting a column that has an exact duplicate elsewhere removes its name while leaving its content.

Suppose the hypothesis is that the order of events adds value beyond their totals. The comparison should preserve the totals while changing or removing the order representation in a way that matches the question. Simply removing all event variables would test whether events matter at all. It would not isolate sequence.

Likewise, testing a layered network against its aggregate should preserve the underlying observations and differ in the representation of channel structure. If the layered version also receives newer data, more actors, and a larger tuning budget, its gain cannot be assigned to layers alone. A fair comparison does not require every model to be identical; it requires the intended difference to be interpretable.

For a geometric feature set, compare a common learner using conventional inputs, geometric inputs, and both. Include a control with a comparable amount of generic flexibility, such as permitted nonlinear transforms of the conventional inputs. If the hybrid gains more than that control on untouched data, the result more specifically supports the geometric representation rather than merely the addition of parameters.

The transformations themselves must be fitted within the training window. Standardization, graph construction choices learned from outcomes, imputation, dimensionality reduction, and probability calibration can all leak information even when the final model's fit respects time. A pipeline is only as temporally honest as its earliest data-dependent step.

Ablation can also test interpretation. If a model is said to detect narrowing alternatives, its prediction should respond sensibly when a genuine, independently usable alternative is added under the measurement rule. The test should distinguish that change from duplicating the label of an existing alternative. These controlled checks do not prove empirical accuracy. They establish whether the implemented representation behaves as the explanation says it does.

The crisis should stay on one side of the boundary

Randomly dividing daily observations into training and test sets can place adjacent days of the same crisis in both. The test model then sees the episode's near-duplicates during fitting. High apparent accuracy may measure interpolation within a familiar event rather than the ability to recognize a new one.

A temporal holdout separates earlier development from later evaluation. Overlapping forecast horizons require additional care. A training row issued just before the boundary may have an outcome window extending well into the test period. Depending on the task, the split should exclude or otherwise account for such overlaps so that training labels do not reveal outcomes the evaluation treats as future.

Geographic holdouts ask another question: whether a method transfers to places excluded from development. They do not automatically remove all dependence, because neighboring units can share an episode, source, actor, or regional shock. The grouping should follow the mechanism of dependence relevant to the claim.

For the African extensions, an informative design could hold out a regional episode while preserving its complete dated source history for later evaluation. The aim would be to test whether the measurement and mechanism transfer, not to treat an entire continent as one homogeneous sample. Actor definition and reporting access may change across regions as much as the political relationships do.

Uncertainty estimates should reflect the unit over which the result is intended to generalize. Resampling individual daily rows can be too optimistic when many rows share one outcome. Resampling whole episodes or time blocks can preserve more of that structure, but the number and comparability of blocks still limit the conclusion. A narrow interval produced by an inappropriate resampling scheme is a computational artifact, not strong evidence.

The report should show where the gain occurs. A pooled improvement may come entirely from a single large crisis or from a particular horizon. That can still be useful, but it is different from a broadly consistent

improvement. Reporting per-horizon and per-episode contributions helps the reader see the distinction without drowning the argument in hundreds of separate rankings.

Model selection spends evidence

Imagine trying fifty variants and reporting only the one with the best score on the test period. Even if all variants have the same underlying predictive quality, one can look best through chance. The more often the test set guides design, the less it functions as unseen evidence.

This is why development, validation, and final evaluation have different roles. The training data fit parameters. The validation data help choose permitted settings. The final test evaluates the chosen procedure. A later improvement inspired by the final test can be legitimate new research, but it needs another evaluation period before being presented as newly confirmed performance.

The development budget should be visible. A simple baseline given one default setting should not be compared with a favored method refined through many undocumented trials and then described as a neutral comparison of mathematical ideas. Conversely, a new representation should not be dismissed because an established system benefits from years of operational engineering. The study can report both controlled development comparisons and complete-system performance.

Preregistration helps by recording the target, sample, primary score, split dates, model families, permitted tuning, and decision rules before the relevant outcomes are inspected. It does not guarantee a good design. It makes deviations visible and separates planned tests from exploratory findings.

Exploration remains valuable. A surprising failure may reveal that a model handles continuing violence well but misses changes in institutional structure. A new hypothesis can then be developed and tested. The important distinction is between discovering an explanation in the data and confirming it on evidence not used to construct it.

This is particularly important for a framework with many related mathematical formulations. The research should name which representation enters each comparison and why. A result obtained with one signed index, normalization, or capacity ratio should not be transferred to every other formulation by a shared family name.

An explanation is another output to evaluate

A forecast may be useful because it identifies a constraint a decision-maker can investigate. That explanatory value should be assessed directly rather than assumed from the presence of equations. Can another analyst reproduce the stated reason? Does the explanation identify the input and branch responsible for the result? Does it remain faithful when the input changes?

In a geometric model, a claim that one connection is the bottleneck can be checked by perturbing the specified connection and recomputing the relevant quantity. If the explanation refers to a maximum-flow cut, the cut should have the stated capacity. If it refers to a slow spectral mode, the displayed vector should satisfy the eigenvalue equation to the numerical tolerance. If it refers to the absence of feasible exits, the constraint set and the meaning of feasibility must be inspectable.

These checks establish faithfulness to the model. They do not establish that the model's bottleneck is the decisive cause in the world. That further claim requires evidence about the connection and competing explanations. A technically faithful explanation can still be based on a poor representation.

Conventional models deserve the same distinction. A feature-importance chart can describe sensitivity or contribution within the fitted prediction rule. It does not automatically establish a causal mechanism. Comparing explanation quality therefore requires a common question: what kind of statement is being offered, what evidence would verify it, and what decision can it responsibly support?

The most useful outcome may be a division of labor. A flexible model can supply a well-calibrated probability, while a constrained calculation identifies which stated civilian service paths remain feasible. Their outputs can be combined if the information boundaries and roles are clear. Calling the combination a hybrid does not excuse either component from evaluation.

What a win would actually mean

A successful common benchmark would support a bounded claim: on this target, population, horizon, release history, and scoring protocol, a specified method improved the reported measure by a stated amount with stated uncertainty. That is already a meaningful scientific result. It need not be inflated into superiority over every forecasting problem.

For this research program, several outcomes would be especially informative. Geometric features might improve onset prediction after

recent violence is controlled. Sequence might matter most when institutions are changing. Explicit observation-age information might reduce overconfidence during reporting disruptions. A capacity model might explain why a warning corresponds to a shrinking set of civilian options even when the probability ranking is unchanged.

Unfavorable findings would also locate the next task. If the representation adds no held-out information, the proposed measurement may be weak or the chosen target may not depend on the modeled mechanism. If ranking improves but calibration worsens, the mapping to probability needs work. If numerical explanation is faithful but source reconstruction fails, the evidence pipeline needs repair. The comparison should make those distinctions possible.

The point of a head-to-head test is to expose contribution under conditions that do not favor it in advance. That is how the framework can earn the strongest version of its claim: by showing where its geometry helps another person anticipate, explain, or preserve a possible future better than the available alternatives.

A comparison that can actually be run

The companion includes a scorer for aligned binary forecasts. Its input key combines the unit, issue time, target definition, target start, and target end. It requires each model to supply the same keys as the outcome file. It refuses duplicate records, unequal populations, invalid probabilities, and information cutoffs after issuance. The comparison reports Brier loss, logarithmic loss, ranking discrimination, and threshold-based counts.

The scorer is tested on constructed forecasts with hand-checkable results. It is not a reimplementation of VIEWS or CAST, and those test fixtures are not evidence of comparative conflict performance. Its purpose is to make the proposed common evaluation executable while preserving the difference between a tested scoring tool and a completed empirical contest.

The current 47-conflict record is retained on its own terms. A numerical ranking against contemporary systems requires archived compatible forecasts or faithfully reproduced models and the corresponding common outcome population. Where those are not yet assembled, a blank result is more informative than a fabricated victory.

This does not leave the comparison empty. It identifies the actual scientific questions: does the geometric representation improve prediction beyond recent event history; does it transfer across regions; does its account of sequence add value; does it remain calibrated when in-

stitutional relations change; and can its explanations identify a failed mechanism more precisely than a feature-importance chart alone?

An informative result need not crown one universal winner. Persistence may be difficult to beat for continuing violence, while structural information may matter more near a change of regime. A hybrid may produce the best probability forecasts, while a constrained geometric calculation provides the clearest account of which feasible options are disappearing. The benchmark should be designed to discover these distinctions.

The aim is a comparison strong enough to alter our expectations. If geometry contributes, the common test should expose that contribution. If it does not contribute to a particular target, the result should tell us where the proposed connection between mechanism and prediction needs more work.

Part VI

Before the next rupture

A warning becomes useful through a decision. The final part follows that passage: who receives the reading, which actions remain feasible, whose costs enter the calculation, and what can go wrong between a sound equation and a public choice. The purpose of seeing a narrowing future is to make room for a different one.

Chapter 36

Policy Implications

> The best time to plant a tree was twenty years ago. The second best time is now.
>
> Chinese proverb

Imagine the warning has arrived. The energy ministry sees a dependence it cannot unwind before winter. The foreign ministry wants to keep negotiations open. The treasury sees the cost of a replacement supply route. Each office can describe its own problem. Someone must decide which sequence of actions the country can actually carry out.

This is where the geometry has to earn its place. An alert should lead to an examination of the available choices, their deadlines, and the people who bear their costs. A framework that cannot inform action is merely academic exercise.

The framework suggests specific policy approaches that differ from conventional wisdom:

Early warning integration:

A monitoring program should bring economic exposure and diplomatic assessment into the same review. A trade specialist may know which component has no replacement. A security analyst may know why access to it is becoming contested. The calculation becomes useful when those observations can be compared rather than filed separately.

Track regional spectral gaps as systemic risk indicators. A collapsing spectral gap—the crystallization of trade networks into opposing blocs—should trigger concern even when no specific dyad shows crisis indicators.

Establish L3/L4 alert protocols with pre-planned response options. The reconstructed 72-day interval before Russia–Ukraine (Chapter 8) illustrates a planning horizon for reserves, communication, and evacuation. An operational system must establish how much of that interval remains after data-release and decision delays.

Prevention strategies:

Orderly decoupling from negative-helicity relationships. This is counterintuitive: the conventional view holds that maintaining economic ties promotes peace. But the framework shows that ties *persisting despite collapsed trust* create the trap. Better to reduce economic interdependence deliberately than to have it severed violently.

Maintain diplomatic channels even during economic disputes. Helicity measures the alignment between economic and diplomatic vectors. Economic pressure combined with diplomatic engagement keeps helicity closer to zero; economic pressure with diplomatic isolation drives helicity deeply negative.

Avoid trap-building—increasing interdependence without corresponding trust-building. German reliance on Russian gas (55–66% by 2021) exemplifies this error. The economic ties were meant to promote peace; instead, they created leverage for coercion.

The controlled release paradigm:

A dam operator facing rising water levels has two options: controlled release through spillways, or wait until the dam fails catastrophically. The first option involves visible, costly action (downstream flooding, wasted water). The second option avoids immediate cost but risks total destruction.

For policymakers facing a negative-helicity relationship, controlled decoupling is the spillway release. It's costly and politically difficult. But the alternative—the trap building to the point of conflict—is catastrophic.

The institutions between states

International organizations face unique challenges in applying the framework:

UN and regional bodies:

Publish framework assessments alongside traditional human rights and security monitoring. The structural approach provides a different signal than event-based early warning, potentially identifying crises before they generate the incidents that trigger conventional alerts.

Prioritize mediation resources for dyads approaching L4. Mediation is expensive; allocating it based on structural risk assessment may prove more efficient than responding to crises already underway.

Design sanctions regimes to *reduce* helicity, not increase it.

Conventional sanctions increase economic pressure while maintaining or worsening political hostility. This can *deepen* trap dynamics by creating asymmetric dependencies while eliminating the trust that might channel interdependence toward cooperation.



UN General Assembly High-Level Side Event on Cooperation for Accountability in Ukraine, September 22, 2022. The international response to Russia's invasion demonstrated that structural warnings, if heeded earlier, might enable preventive rather than reactive diplomacy.

Photo: Wikimedia Commons, CC BY-SA 4.0.

Framework-informed sanctions would be designed to enable orderly decoupling—reducing curvature (economic interdependence) and bringing helicity toward zero simultaneously. This might mean sequencing: first establishing diplomatic channels for managed separation, then implementing economic measures that facilitate rather than punish disengagement.

The Iran nuclear agreement offers a case of linked commitments. Nuclear restrictions, verification, and sanctions relief were intended to operate together. This is a useful setting for asking how an agreement changes both the permitted actions and the incentives to maintain it.

For the framework, the agreement poses a demanding question: which commitments made a previously blocked concession possible, and which failures later closed that route? The answer requires tracing verification and delivery, rather than treating the presence of a signature as a measurement of restored trust.

The durability of those commitments matters. A concession that takes years to reverse is exposed differently from a promise that can

change with a new administration. The sequence and reversibility of the exchange belong in the account of tolerance and trust.

An agreement also has to survive the institutions and constituencies on which implementation depends. A signature can open a route without guaranteeing that the route remains usable. Monitoring should follow verification, delivery, political support, and the consequences of nonperformance.

The framework opens several research agendas:

Incorporate structural metrics alongside event-based analysis. Current forecasting systems (ViEWS, ACLED, Early Warning Project) rely heavily on event counts and sequences. Adding spectral gap and helicity measures could improve onset prediction—the "hard problem" that current systems struggle with.

Real-time validation: Test framework predictions against new conflicts as they emerge. The 47-conflict validation set is historical; ongoing monitoring would provide out-of-sample evidence for or against the framework's predictive power.

Extension to multi-party systems: The bilateral focus is a limitation. Many conflicts involve more than two parties, regional networks, or alliance systems. Extending the mathematics to N-body problems would increase the framework's applicability.

Threshold refinement: The critical thresholds ($H^* = -0.5$, $\lambda_1^* = 0.05$, $\Gamma^* \in [0.9, 1.1]$) were derived from a limited sample. Larger datasets might reveal more precise or context-dependent thresholds.

An alert is useful when it changes a decision that can still be changed. Before distributing a risk assessment, identify its recipient, the available actions, and the time required for each. A six-month warning is operationally short if a substitute supply route needs five years. It can be operationally long if the relevant action is to reopen a channel that already exists.

The same score may support different decisions. A ministry responsible for energy needs to know which supplies could be interrupted and how long reserves can cover the gap. A diplomatic service needs to know which agreements are still implementable. A humanitarian organization needs a planning horizon and a picture of exposed populations. These uses should not be forced into one instruction to "respond to Level 4."

For each proposed response, state the mechanism. If the concern is a critical input, the response may create supply or reduce consumption. If the concern is uncertainty about intentions, the response may establish verification. If the concern is an irreversible concession, the response may divide implementation into reciprocal steps. The intervention should address the constraint the analysis actually identifies.

The first question is descriptive: does the measure represent an intelligible feature of the relationship? The second is predictive: does it improve the assessment of a specified future event? The third is causal: would changing that feature through a particular intervention change the event's probability?

These questions support one another, but they require different evidence. A highly interpretable score can be a weak predictor. A strong predictor can include variables that cannot be changed by policy. A policy can lower the score while increasing a danger the score does not measure.

Suppose a government reduces imports from a hostile supplier. The measured trade dependence falls immediately. If domestic inventories are inadequate and replacement supply is years away, its short-term vulnerability may rise. The favorable movement of one coordinate would be a poor basis for declaring the intervention successful.

Policy evaluation must therefore follow the causal chain. Record the intended change, its implementation, the intermediate consequence, and the outcome. Where possible, compare with similar relationships or sectors exposed to different timing or treatment. Where comparison is weak, keep the conclusion correspondingly narrow.

Make the Transition Feasible

The controlled-release analogy becomes useful when we specify the spillway. What capacity exists? Who controls it? How fast can it operate? Where does the released pressure go?

For a proposed reduction in dependence, construct a transition schedule. At each stage, compare essential demand with accessible supply, include the cost of the alternative, and identify the commitments needed to keep the transition open. Model temporary shortfalls as well as the preferred endpoint.

One intervention creates a substitute: a supplier, customer, route, or technology that can perform the necessary function within the relevant horizon. Another creates time. Reserves, bridge financing, demand management, and interim agreements can keep a relationship functioning while that substitute develops. These interventions are complementary when the alternative is useful but cannot arrive soon enough.

A third intervention creates a procedure. Verification, review, appeal, or staged reciprocity can let a dispute move without requiring a comprehensive settlement. A fourth reduces an irreversible commitment: changing its scale or sequence can prevent a future disagreement from

making every adjustment catastrophic. Here the resource being created is room to change course.

The interventions can conflict. A reserve may create time but be expensive to maintain. A substitute may reduce exposure but weaken an existing constituency for cooperation. A public commitment can make a promise more credible while making compromise more costly. The correct comparison includes these consequences rather than treating every reduction in dependence as equally stabilizing.

A point estimate should not be the only input to a consequential decision. Consider a range of plausible observations and model choices. Identify actions that remain useful across that range, actions whose value depends on a narrow assumption, and actions that would be difficult to reverse if the assessment is wrong.

Keeping a functioning communication channel open may be useful under several plausible interpretations of a crisis. A costly and irreversible economic rupture requires a more specific argument. This is a distinction about the decision's consequences, not a reason to ignore a severe warning.

In a simplified expected-loss model, let p be the event probability, L the avoidable loss if it occurs, C the cost of an intervention, and e the intervention's effectiveness. The action has positive expected net benefit when

$$p\,e\,L > C.$$

Each term is uncertain, and the expression omits distribution, legitimacy, and strategic response. Its value is to show that a prediction alone does not select a policy. An intervention can be ineffective, or its cost can exceed the loss it can plausibly prevent.

Suppose reserving space on a second freight route costs 0.6 million. If a closure occurs, the avoidable loss would be ten million, and the reservation could prevent half of it. On those assumptions the break-even probability is $0.6/(0.5\times10) = 0.12$. If the second route could prevent only a quarter of the loss, the break-even probability would rise to 0.24. The warning level did not change between those calculations; the proposed remedy did. To act on the comparison, someone still has to justify a probability for this particular closure and verify that the reserved space actually moves the affected cargo.

Where credible ranges are available, evaluate the inequality across them. An action justified only by the highest estimate of p and the most optimistic estimate of e is more fragile than one supported throughout the plausible range. Neither conclusion requires pretending that the probability is known exactly.

Who Pays for Resilience

Resilience is often praised in the abstract and underfunded in practice. A producer asked to maintain an alternative market may face higher costs. A worker asked to support economic separation may lose the job that financed the transition. A household may pay for a national strategy through a higher bill.

Those costs affect both justice and feasibility. A transition that ignores them can lose political support before it creates the intended alternative. The model's "available path" then disappears because the policy treated the people who had to sustain it as an afterthought.

Evaluate adjustment at more than one level. National income, regional employment, household essentials, and the position of politically influential groups can move differently. The distribution of costs can also affect the other state: pressure intended to change a government's behavior may instead fall on people with little influence over it.

The preventive purpose of the framework requires keeping these people visible. Saving an aggregate indicator while destroying the conditions of ordinary life would be a poor interpretation of success.

Suppose an alert is followed by a summit, a new agreement, and no war. The sequence deserves study. It does not, by itself, prove that the summit prevented war. The initial risk may have been lower than assessed; another change may have mattered; the summit may have ratified a decision already made.

Look for evidence of the proposed mechanism. Did the agreement restore a particular supply? Did a verification procedure resolve uncertainty? Did a staged exchange make an otherwise unacceptable commitment reversible? Did the parties implement the terms, and did the relevant constraint change before the danger receded?

Keep the forecasting record and the intervention record connected but distinct. The event either occurred within the specified horizon or it did not. A separate account can explain why an alert did not culminate in the event. Preserving both records allows a warning system to learn from prevention without becoming impossible to evaluate.

The same discipline applies when intervention fails. A policy may have been implemented too late, implemented incompletely, or based on the wrong mechanism. An unchanged outcome does not identify which explanation is correct. Trace the steps instead of treating the result as a verdict on every part of the model.

A useful warning process keeps the original assessment, the information available at the time, the response considered, the response chosen, and the eventual outcome. The record should include disagreement

among analysts. A minority view can later reveal a missing mechanism or an assumption that mattered more than expected.

Review should occur after quiet periods as well as crises. Otherwise the organization learns only from spectacular failure and never examines the many occasions on which restraint, adaptation, or simple misclassification produced a peaceful interval.

The aim is a durable capacity to notice, question, and act. No single analyst should have to persuade an institution to invent an entire response process after the warning has arrived. The procedures, responsibilities, and alternatives should already be available.

That is the institutional counterpart of the geometry. A state with more usable choices has more room to preserve peace. A warning organization with more usable responses has more chance of helping it do so.

The Responsibility to Act

Prediction and Responsibility

If a framework can predict conflict with significant lead time, those who possess it bear responsibility. Not to weaponize the information for strategic advantage, but to enable prevention.

This book is published openly because early warning has value only if it reaches those who can act. A classified prediction system serves intelligence agencies; a public one serves humanity. The choice to publish reflects a judgment that the benefits of widespread awareness outweigh the risks of adversarial exploitation.

The reflexivity problem cuts both ways. Publication might enable actors to game the system—manipulating helicity indicators to create false signals or mask real ones. But publication also enables the targets of potential aggression to recognize their vulnerability and take protective action. On balance, transparency seems likely to favor peace over conflict.

The purpose of warning is to leave time for a different decision. If a reserve prevents a shortage or an agreement keeps a route open, the intervention should be documented alongside the original warning. Averted harm is the practical ambition. Establishing that the intervention caused the better outcome is a further research task.

Acting on a warning makes its limits practical. A stale observation, an unstable threshold, or a mistaken comparison can redirect real resources. We therefore have to examine the instrument at the points where a careful calculation can still lead to a mistaken decision.

Chapter 37

When the Warning Changes the Outcome

A warning is issued. People act on it. The feared event does not occur.

Was the warning wrong, or did it help prevent the event?

The outcome alone cannot answer. This is one of the hardest problems in evaluating a warning system intended to be useful. The forecast enters the world it describes. Attention, negotiation, reserves, public statements, and other responses can change the path that follows.

The problem deserves a precise treatment because it otherwise becomes an escape route for any theory. If war occurs, the model claims prediction. If war does not occur, it claims prevention. A framework that accepts both outcomes as automatic confirmation has made itself impossible to test.

Two different claims

The first claim is predictive: among cases receiving a stated probability under a specified information set and operating environment, how often does the event occur? The second is causal: how would the outcome differ if a particular intervention were applied rather than not applied?

Write $Y(1)$ for the outcome under a specified intervention and $Y(0)$ for the outcome under a specified comparison condition. The individual causal effect is $Y(1) - Y(0)$. For one realized case, only one of these outcomes is observed. The other is a counterfactual, not a missing spreadsheet entry that can be recovered by confidence in the model.

The intervention must be concrete enough to define. "Diplomacy" can refer to many actions with different timing and content. "No intervention" can still include routine institutional activity and private decisions. A causal question should identify the action, comparison, population, timing, and outcome rather than assuming those words establish a unique experiment.

The potential-outcome framework provides a standard language for these distinctions and for the assumptions needed to learn causal effects

from observational data (Hernán and Robins 2020). The geometric framework can help specify mechanisms and constraints. It cannot make an unobserved alternative outcome directly observable.

The most dangerous cases receive the most attention

Suppose officials direct an intervention toward cases they judge most dangerous. Comparing the treated cases with the untreated cases can then make a helpful intervention appear harmful, because the groups started at different levels of risk.

Consider a constructed example with two severity groups. Among treated cases, twenty are mild and eighty are severe. One mild treated case and twenty-four severe treated cases experience the event. Among untreated cases, eighty are mild and twenty are severe. Eight mild untreated cases and ten severe untreated cases experience it.

Severity	Intervention	Comparison
Mild	1/20 = 5%	8/80 = 10%
Severe	24/80 = 30%	10/20 = 50%
All observed cases	25/100 = 25%	18/100 = 18%

Table 37.1: Constructed counts. The intervention group has a lower event rate within each severity category but a higher pooled rate because it contains many more severe cases.

Within both severity groups, the intervention group has a lower event rate. In the pooled data, it has a higher rate. The apparent reversal is caused by the different mixtures of mild and severe cases.

Standardize both groups to a population that is half mild and half severe. The resulting risks are

$$R_1 = \tfrac{1}{2}(0.05) + \tfrac{1}{2}(0.30) = 0.175, \tag{37.1}$$

$$R_0 = \tfrac{1}{2}(0.10) + \tfrac{1}{2}(0.50) = 0.30. \tag{37.2}$$

The crude difference was +0.07. The standardized difference is −0.125. Adjustment has changed the direction of the comparison.

This does not prove that an actual intervention caused a reduction. It demonstrates how selection can distort a comparison and how one specified adjustment changes it. If important differences remain within the severity categories, the adjusted comparison can still be confounded.

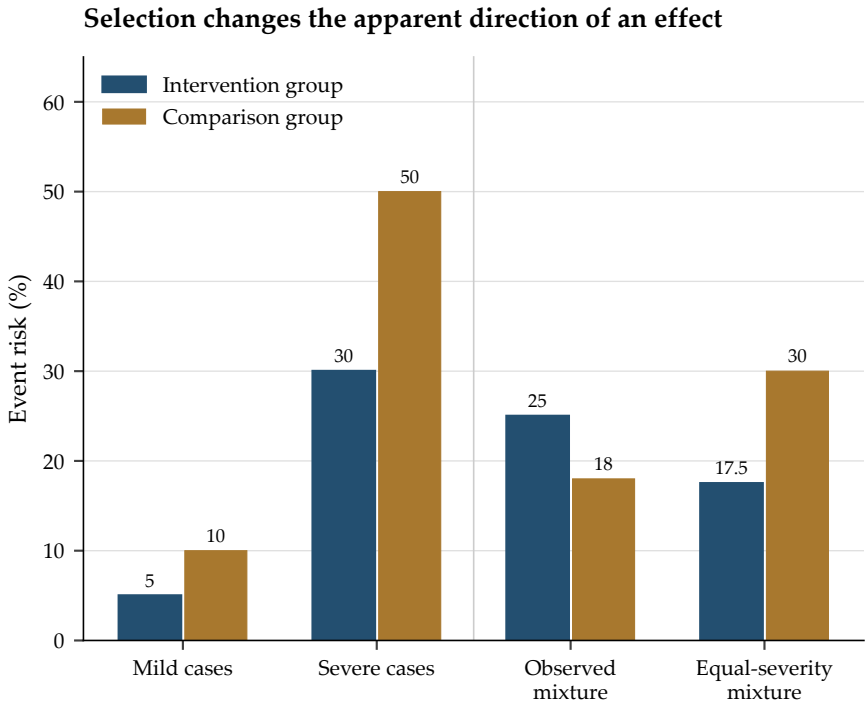


FIGURE 37.1: The same constructed counts viewed within severity groups, in their observed mixtures, and after equal-severity standardization. Standardization removes this specified imbalance; a causal interpretation still requires the assumptions described in the text.

The example also shows why model alerts cannot be treated as randomly assigned. A useful warning system is designed to select unusual cases. Evaluating the response requires accounting for that selection rather than interpreting it as a design defect.

What must be assumed to interpret an effect

For an observational comparison, one common identification strategy assumes that, after conditioning on measured pre-intervention variables X, the intervention assignment contains no additional information about the potential outcomes. This is conditional exchangeability. It is a substantive assumption about omitted causes, not a property a fitted regression can certify from its own goodness of fit.

The strategy also requires positivity: within the relevant values of X, both intervention and comparison must be possible in the observed data.

If every extremely severe case receives an intervention, the data do not directly show what comparable untreated severe cases would have done. Extrapolation may still be attempted, but it should not be described as an observed comparison.

Consistency requires the recorded intervention to correspond to the intervention whose potential outcome is being defined. If the same label covers incompatible actions, an estimated average may have no clear operational interpretation.

Under these conditions, a standardized risk can be written as

$$E[Y(a)] = \int E[Y \mid A = a, X = x]\, dP(x),$$

where A records the intervention and the outer distribution defines the target population. In the constructed table, X was severity and the chosen target mixture assigned one-half to each category.

The equation explains what adjustment is trying to do. Estimate the outcome within comparable conditions, then average those conditional outcomes over the same population. It also identifies the vulnerable points: the conditions may not be sufficiently comparable, the relevant population may lack overlap, or the outcome model may extrapolate badly.

Geometric variables can enter X if they are measured before the intervention and are relevant to its assignment or the outcome. Variables changed by the intervention require more care. Adjusting for a mediator can remove part of the effect one intended to estimate; conditioning on a common consequence can create spurious associations. A temporal causal account is needed before deciding which columns to include.

Political systems interfere with one another

Many elementary causal analyses assume that one unit's treatment does not affect another unit's outcome. Regional conflict processes can violate that assumption directly. A negotiation involving one actor can alter the options of a neighbor. A resource arrangement can shift demand or pressure elsewhere. A warning made public can affect several relationships at once.

In a network setting, the outcome for unit i may depend on both its own action a_i and the surrounding pattern a_{-i}:

$$Y_i = Y_i(a_i, a_{-i}).$$

The notation makes the additional problem visible. The number of possible intervention patterns can become enormous, so a practical study needs a justified exposure model describing which neighboring actions matter and how they are summarized.

That is a natural place for geometry to contribute. A specified network can organize the hypothesized routes of spillover. But the chosen network must be supported by the mechanism: a trade graph is not automatically the right graph for information, command, or population movement. The causal model inherits the same typed-edge discipline as the forecasting model.

Interference can also change the interpretation of a successful local intervention. A reduction in one place may coincide with displacement elsewhere. An evaluation concerned with civilian protection should examine the relevant wider outcome, rather than declaring success at the boundary of the administrative unit used for the program.

The comparison must begin at the same moment

Consider a regional mediation program that starts only after an alert. Some teams arrive within a week; others take a month. A retrospective analyst labels every place eventually visited as "treated" and compares its conflict rate with places never visited. The label appears simple, but it has allowed the future to decide which group a place belonged to in the past.

Suppose the protocol excludes places where fighting begins before a team arrives. The eventual intervention group then contains a period in which an onset would have prevented membership in that group. Its apparently quiet waiting time is partly guaranteed by the classification rule. A crude comparison can credit the intervention for outcomes that occurred before it started.

An evaluation needs a common starting point. For example: the first eligible alert after a declared washout period, issued under a frozen rule, with no qualifying conflict already under way. At that moment record the information available and define the strategies to compare. One strategy could offer a specified mediation visit within seven days; another could continue the existing response. Follow both from eligibility, recording later adherence, delays, and outcomes.

This is the logic of a target trial: state the experiment one would need, then identify what the observational record can and cannot reproduce (Hernán and Robins 2020). The exercise need not imply that random assignment of a high-stakes intervention is possible or desirable.

It makes eligibility, assignment, timing, follow-up, and the outcome explicit enough to reveal avoidable bias.

There are still two different questions. The effect of an offer includes failures to deliver the offer. The effect of actually receiving a visit requires a different comparison and stronger handling of the reasons some visits occurred. Quietly excluding cancellations can make an implementation failure disappear from the result. For a policymaker deciding whether to fund the program, that failure may be part of the effect that matters most.

A warning and a response can share a cause

A minister receives private intelligence, orders an intervention, and asks analysts to raise the alert level. The archive later shows a high model score, a response, and a lower-than-feared outcome. If the private intelligence is missing from the research file, the observed score does not by itself capture everything that drove the decision.

The problem is not solved by adding every available variable to a regression. A variable measured after the response may be one of its consequences. A diplomatic contact could reduce hostility; controlling for the resulting hostility can remove part of the pathway through which the contact worked. Conversely, a variable that records whether journalists investigated the case can be influenced by both the intervention and the severity of events. Conditioning on that common consequence can create a misleading association.

A causal diagram is useful when its arrows correspond to a defensible account of timing and mechanism. It is not an ornamental version of the correlation matrix. For each proposed adjustment, ask what caused that variable, when it was measured, and whether the action could already have changed it. The answer determines whether the column helps construct a comparison or distorts one.

The geometric instrument can contribute pre-intervention descriptions of constrained alternatives, delivery dependence, and available channels. It can also suggest where the action might operate. It does not automatically measure all private information, institutional capacity, or political motivation. Treating a detailed geometric state as sufficient adjustment is a substantive hypothesis that needs scrutiny.

A transparent weighted comparison

The earlier two-severity example can be read another way. There are one hundred mild cases in total, twenty treated and eighty untreated.

There are one hundred severe cases, eighty treated and twenty untreated. In this constructed population, the observed treatment proportions are therefore

$$e(\text{mild}) = 0.2, \qquad e(\text{severe}) = 0.8.$$

Here $e(x)$ denotes the probability of treatment conditional on the recorded pre-treatment category. It is often called a propensity score. Its role is to describe assignment, not the probability of conflict.

Weight each treated case by $1/e(x)$ and each untreated case by $1/(1 - e(x))$. In the treated group, the twenty mild cases receive weight five and the eighty severe cases weight 1.25. Both categories contribute a weighted population of one hundred. In the untreated group, the weights reverse and again produce one hundred of each category.

The weighted event totals are easy to inspect. Treated mild cases contribute $1 \times 5 = 5$ events and treated severe cases contribute $24 \times 1.25 = 30$. The weighted treated risk is $35/200 = 0.175$. Untreated mild cases contribute $8 \times 1.25 = 10$ and untreated severe cases $10 \times 5 = 50$, giving $60/200 = 0.30$. These are the same standardized risks obtained earlier by averaging the two categories equally.

In this example, weighting reconstructs the chosen common severity mixture. It does not create new observations or demonstrate that severity is the only confounder. It reorganizes the observed information under an explicit assignment model. The result inherits that model's assumptions.

The arithmetic also exposes why poor overlap is dangerous. If only one in a hundred comparable cases receives an intervention, that one case receives a weight of one hundred. Its outcome can dominate the estimate. If no comparable case receives the intervention, there is no finite weight that manufactures the missing comparison. A smoothly fitted propensity score can conceal that absence if the analyst looks only at the final average.

Report the distributions of the weights, the covariate balance they achieve, and the regions where comparable cases are scarce. Restricting the target population may sometimes give a more credible answer, but it changes whom the conclusion describes. Trimming large weights may improve numerical stability while introducing a different bias. Neither operation should silently transform a claim about all severe crises into a result about the easiest cases to compare.

When today's response changes tomorrow's evidence

Most prevention programs do not consist of one action followed by a passive wait. A first visit changes the situation; a second decision

responds to that changed situation. The new information can be both a consequence of the earlier action and a cause of the later action. Ordinary adjustment for every recorded intermediate variable can then answer the wrong question.

Let A_0 be an initial action, L_1 the subsequent observed state, A_1 a second action, and Y the later outcome. A policy may specify $A_0 = a_0$ and then choose $A_1 = g(L_1)$. Under the required sequential exchangeability, consistency, and positivity assumptions, and suppressing baseline variables for this illustration, its risk can be expressed as

$$\sum_{l} E[Y \mid A_0 = a_0, L_1 = l, A_1 = g(l)]P(L_1 = l \mid A_0 = a_0). \tag{37.3}$$

The first factor describes outcomes after the second decision within a particular intermediate state. The second describes how often the first action leads to that state. Keeping both factors lets the first action change the distribution of subsequent conditions rather than adjusting that pathway away (Hernán and Robins 2020).

Take a constructed two-state example. After the first action, the probability of a severe intermediate state is 0.3. Without it, that probability is 0.6. In a severe state, a follow-up action gives outcome risk 0.20, compared with 0.40 without follow-up. In a mild state, the risk without follow-up is 0.05. Assume for this teaching example that these conditional risks have the causal interpretation just specified and that the first action has no additional effect once the intermediate state and second action are fixed.

A policy that takes the first action and follows up only when the state is severe has risk

$$0.3(0.20) + 0.7(0.05) = 0.095.$$

A policy that skips the first action but uses the same follow-up rule has risk

$$0.6(0.20) + 0.4(0.05) = 0.14.$$

A policy that takes neither action has risk $0.6(0.40) + 0.4(0.05) = 0.26$. These differences come from a fully specified example, not an estimate of any country’s diplomacy. Their purpose is to show why a sequence of decisions must be evaluated as a sequence.

The record needed for this calculation is more demanding than a column saying “mediation occurred.” It needs the initial eligibility time, the first action, the subsequent state as it was known then, the rule for follow-up, the second action, and a consistently observed outcome. If

only the final state survives, the causal pathways cannot be recovered by adding mathematical sophistication afterward.

Test the comparison where no effect should appear

A proposed causal account should face checks that could make it less convincing. If an intervention appears to change an outcome measured before it occurred, the comparison is detecting something besides its effect. A placebo date can reveal pre-existing differences or an analysis procedure that finds changes wherever it is asked to look. A negative-control outcome can expose shared bias when there is a strong substantive reason the intervention should not affect that outcome.

These are diagnostics, not automatic proofs. An apparent pre-intervention change can reflect anticipation if people act before the formal start date. A supposedly unrelated outcome may have an overlooked causal pathway. The control deserves the same care about mechanisms and timing as the main result. Choosing it because it happens to pass would reproduce the selection problem it was intended to detect.

Sensitivity analysis asks a complementary question: how strong would an omitted cause or a plausible measurement error have to be to alter the conclusion? State the departure being considered and show how the result moves. A phrase such as "robust to confounding" is too broad unless the reader can see which confounding structures were tested.

Finally, keep the geographical boundary in view. A corridor may become safer because pressure moves to another corridor. A negotiation may protect one population while excluding another. An evaluation that ends at the program's administrative border can miss that displacement. The outcome definition should follow the protective purpose of the intervention, including a justified account of spillovers, rather than only the part of the map where implementation was easiest to record.

What this adds to a warning system is a disciplined route from prediction to learning. The forecast says what risk was assessed. The response record says what was attempted. A causal analysis asks what difference the attempt made under stated assumptions. Together they allow a successful prevention effort to become evidence that can guide another one, rather than an untestable story told after a quiet year.

Keep two ledgers

A practical evaluation should preserve the original forecast ledger and a separate response ledger. The forecast ledger records what was issued, when, for which target and horizon. The response ledger records what actions followed, who could observe the warning, when the actions occurred, and what evidence supports their classification.

Score the forecast under the declared operational protocol. Then analyze the response with its own causal question and assumptions. Do not erase a false alert from the forecast record merely because an intervention occurred. Do not declare an intervention ineffective merely because the treated cases were unusually dangerous at baseline.

Where possible, compare policies prospectively under an ethically and institutionally appropriate design. Some actions may admit randomized or phased evaluation; many high-stakes political decisions will not. Observational comparisons, interrupted time series, and comparative case designs can still contribute, but each has assumptions that should be examined rather than hidden behind a causal verb.

The absence of an event after an alert is therefore informative but ambiguous. It may reflect a false alarm, a successful response, an unmeasured change, or the ordinary uncertainty represented by the forecast. Resolving those possibilities requires additional evidence.

This is not a reason to demand that a useful model stand aside while harm occurs. It is a reason to design evaluation alongside use. A warning system aimed at prevention should expect its own deployment to alter the process and should preserve the records needed to study that alteration.

The book's purpose remains practical: to recognize narrowing possibilities early enough that people can create better ones. A rigorous account of prevention strengthens that purpose. It prevents a hopeful outcome from becoming an automatic proof and makes it possible to learn which responses actually helped.

Chapter 38

Limitations and Critiques

> Criticism may not be agreeable, but it is necessary.
>
> Winston Churchill

Suppose an analyst makes the calculation correctly using the latest trade release. The release describes shipments from months ago. A route has since closed, a new supplier has entered, or an agreement has changed the terms of access. The arithmetic can be right while the picture it describes is already out of date.

The limits of a warning system enter at identifiable places: the observations, the model built from them, the outcome being predicted, and the response the warning itself provokes. Following those limits through the calculation tells us when the instrument can be trusted and what further work a reading requires.

The framework inherits fundamental constraints from its source data. Each limitation creates potential systematic biases that practitioners must understand.

International trade statistics from the IMF Direction of Trade Statistics and UN COMTRADE arrive with 6-12 month delays. The 2022 Russia-Ukraine case benefited from unusually rapid European trade disruption reporting; other dyads may not enjoy such data quality. Real-time commodity tracking (ship transponders, satellite imagery of port activity) offers partial mitigation but introduces new measurement error.

Energy dependency data proves particularly problematic. Pipeline gas volumes are reported monthly with reasonable accuracy, but LNG spot markets, barter arrangements, and transshipment through intermediaries obscure true dependencies. The European Commission's post-2022 effort to map Russian energy exposure revealed dependencies 15-30% higher than official statistics indicated.

The alignment parameter A_{AB} requires aggregating fundamentally different diplomatic signals:

Every alignment measure sees only part of the relationship. UN voting similarity is well defined in the Voeten codebook, but it records declared positions rather than the intensity of preferences. Treaty ratification is observable, yet a binary entry cannot capture differences in compliance or significance. Sanctions require judgments about which measures count and how sector-specific restrictions should be weighted. Joint military exercises supply another signal, but their records are collected inconsistently. A composite score inherits these limitations even when its final arithmetic is exact.

Reasonable analysts might code the same dyad-year differently by 10-15% on alignment scores, propagating uncertainty through helicity calculations.

Applying the framework to pre-1990 cases confronts systematic data limitations. Cold War trade statistics were distorted by COMECON reporting practices. Pre-digital diplomatic records require archival research that introduces selection bias—well-documented cases like the Cuban Missile Crisis may not represent the broader universe of crisis dyads.

Perhaps most seriously, the framework assumes that the relevant economic flows can be observed. Sanctions evasion, informal transfers, and transshipment through intermediaries can obscure exposure. A small recorded bilateral flow need not establish that the states have no indirect dependence. The network boundary and the evidence for such routes must be examined together.

Putting the argument under pressure

The most demanding objections concern what the calculations establish and how the empirical record was produced. They deserve direct answers.

Critique: The framework was developed to explain Russia-Ukraine and then retrofitted to appear general.

Response. The geopolitical application was developed after the February 2022 invasion. December 14, 2021 is a date assigned within a retrospective series. The author's recollection of a discovery cannot establish an ex ante forecast. Historical holdouts, release-date replays, and predictions preserved before their outcomes are known provide progressively stronger tests of forecasting skill. The retained study belongs in that research sequence.

In the 23 selected historical cases in Appendix C, 18 detections give 78.3 percent. Which cases were available when the thresholds were

chosen matters to the interpretation. The working paper's separately dated training and holdout results appear in Appendix E.

Critique: The physics language is window dressing—calling something "helicity" doesn't make it behave like physical helicity.

Response. A trade graph can have a genuine Laplacian, and its eigenvalues obey the same algebra as those of any graph satisfying the same assumptions. The interpretation is an additional claim. The chosen weights must represent the process of interest, and the evolution law must justify translating a spectral quantity into adjustment or warning. A fluid conservation theorem applies to a political model only if its hypotheses have been established there.

The success of a mathematical description in another field is a reason to investigate it. It is not evidence that the new application already works. In this program, the empirical comparisons must establish what the geometric quantities add.

The claim is not that economies are fluids but that the *mathematical framework* developed for analyzing vortex flows provides computational tools applicable to economic coupling. Topology is agnostic about substrate. A Laplacian eigenvalue has identical mathematical properties whether computed on a fluid vortex, a social network, or a trade matrix.

Critique: A 15% false positive rate makes the system unreliable for policy.

Response. Start with the person who receives the warning. Precision asks how many issued alerts preceded the specified event. A false-positive rate asks how many eligible quiet cases drew an alert. Tetlock's expert forecasts and systems such as ViEWS have been evaluated using different targets, horizons and measures. The working paper's 47-event counts yield 35 detections and 14 false alerts, with different denominators for recall and precision (Appendix E). The separate historical percentages retain their own annotations in Appendix E. To compare warning desks fairly, give each the same dated cases and count the same outcomes.

The costs of missed events and false alerts can differ greatly, but the response matters. Extra monitoring, a reserve purchase, an evacuation, and a military deployment have different costs and risks. A warning system should be assessed against the decisions it is meant to support, including the harm that unnecessary action can cause.

A ministry might rationally tolerate more false alerts for a low-cost review than for an irreversible intervention. Choosing that tolerance is a decision problem. It does not change the observed accuracy of the warning rule.

Repeated alarms also consume attention. An instrument that keeps every relationship at its highest category may cease to help its users

allocate effort. Discrimination matters because staff, time, and political capital are limited.

A useful alert should lead to a concrete question. Has a supplier disappeared? Has a commitment become impossible to honor? Does the observation justify urgent verification before a major decision? The answer should be visible in the inputs, not hidden behind the category name.

After the horizon closes, the organization should be able to compare what it knew, what it predicted, what it did, and what happened. That record supports improvement whether the episode ended in violence or peace.

This asymmetry shapes how practitioners should interpret framework outputs. An L4 alert is not a prediction of certain war—it is an alarm that demands attention, resources, and diplomatic energy. The framework's value lies not in certainty but in discrimination: distinguishing the dyads that warrant extraordinary vigilance from those that do not.

The boundary of the instrument

Intellectual honesty requires clear delineation of what the framework does *not* address. The helicity formulation is designed for a specific class of conflicts and makes no claims beyond that domain.

The framework is explicitly designed for interstate war—organized violence between recognized states. It does *not* address:

Civil wars: Internal conflicts follow different dynamics, where the relevant "dyad" is often government versus population or ethnic group versus ethnic group. The economic interdependence that drives interstate helicity doesn't translate to internal fragmentation.

Non-state actor violence: Terrorist organizations, insurgencies, and transnational criminal violence operate outside the state system that defines bilateral flux measurement.

Genocides and mass atrocities: The helicity framework cannot distinguish between interstate war and internal atrocity; Rwanda 1994 would not register on bilateral metrics.

The framework operates at the level of structural forces, not individual agency. It cannot account for:

Individual decisions. Risk tolerance, domestic constraints, and coalition pressures can affect whether a government compromises or uses force, as well as when and on what scale. Aggregate measures do not contain a complete account of those choices.

Decision-making pathologies: Groupthink, motivated reasoning, and intelligence failures can delay or accelerate conflict onset independently of structural conditions.

Accident and miscalculation: A cyber incident misattributed, a submarine collision, an accidental missile launch—these trigger points exist outside structural forecasting.

Even when correctly predicting conflict, the framework provides only broad temporal guidance:

Exact onset: L4 status indicates elevated near-term risk but cannot specify whether conflict occurs in weeks or months. The 72-day Russia-Ukraine lead time should not be interpreted as a standard interval.

Conflict duration: The framework offers no guidance on whether a conflict lasts days (Six-Day War) or years (Iran-Iraq War).

Outcome: Helicity metrics cannot predict which side prevails or how conflicts terminate.

When the warning changes the world

Any social prediction system faces a fundamental challenge absent in natural science: the objects of prediction can observe and respond to predictions about them. This reflexivity creates both problems and opportunities.

If the helicity framework achieves wide adoption, states may adjust behavior in response to their framework status. A government seeing its dyad approach L4 might:

Publication can change the behavior being measured. A government may reduce trade exposure to make a future separation less costly, or increase diplomatic engagement to improve alignment. It may also accelerate military preparations if it reads conflict as inevitable. A more deceptive response is to manipulate visible indicators while preserving hostile intent. The forecast then becomes part of the relationship it is trying to describe.

Each response changes the system being measured, potentially invalidating the measurement itself.

More perniciously, adversaries might weaponize the framework against itself. A state planning aggressive action could artificially inflate alignment metrics (participate in UN votes, sign treaties it doesn't intend to honor) while quietly building flux dependencies it plans to exploit. The framework would show improving helicity while actual conflict risk grows.

This concern is not hypothetical. Russia maintained active trade and diplomatic engagement with Ukraine through 2021 while simul-

taneously building invasion capacity. The framework detected the degradation—but a more sophisticated adversary might learn to mask the signature.

There is a perspective from which reflexivity is not a problem but the purpose. A warning system that changes behavior is a successful warning system. If widespread helicity monitoring causes states to maintain alignment when they might otherwise let it degrade, the framework has prevented conflicts even if it never "correctly predicts" one.

Medical diagnostics face identical logic. A test that causes patients to change diet and exercise, avoiding the heart attack it predicted, has succeeded despite being "wrong" about the outcome.

The measure of success is not predictive accuracy on a static world but the total reduction in realized conflict. By that measure, reflexivity is a feature, not a bug.

This analysis suggests a policy question: Should framework outputs be public or restricted? Public dissemination enables democratic accountability and academic replication but also enables gaming. Restricted access maintains surprise value but concentrates power in the hands of those with access.

We advocate for transparency. The benefits of broad monitoring—more eyes on the data, more pressure on deteriorating dyads, more preparation time for potential conflicts—outweigh the risks of strategic gaming. Adversaries already pursue deception; the framework adds signal, not noise.

Those limits give the research a direction. They identify what must be observed sooner, what must be represented more faithfully, and which comparisons can show whether an extension has helped. The final chapter follows that work beyond the pair of countries with which we began.

Chapter 39

Future Directions

> The future is already here—it's just not evenly distributed.
>
> William Gibson

A pair of countries is a place to begin. Follow a shipment and a third country appears as a transit route. Follow the financing and another institution enters. Follow the political response and a guarantor, an ally, or a rival changes what either country can do. The next stage of the work is to preserve those relationships without losing the ability to explain the calculation.

The open questions are therefore concrete. How do several relationships constrain one another? How quickly can the observations arrive? How does the geometry change when the essential flow is water or migration? And who should be able to inspect the instrument?

The bilateral formulation captures a large fraction of interstate conflict—most wars involve two primary belligerents. But the framework's restriction to dyads represents a genuine limitation for complex regional systems.

During the internationalized Syrian civil war, Russia, Iran, Turkey, the United States, Israel, and Gulf states acted through overlapping and sometimes opposing relationships. A bilateral account captures only part of that system. A change in one actor's commitment can alter the options of several others at once.

Adding an actor to the analysis therefore requires more than adding another score. The new connection can redirect supplies, change the credibility of a guarantee, or make a previous compromise unacceptable. The extension has to represent those interactions.

This observation points to a deeper mathematical truth. In network science, emergent properties appear at higher orders of interaction. Pairwise relationships capture first-order structure. Triangular relationships capture second-order effects—alliances of three that behave differently from any combination of their bilateral components. The Cold War's

NATO-Warsaw Pact structure was fundamentally triangular: the US-Europe-USSR configuration had properties that no bilateral analysis could capture.

The mathematical extension is clear in principle. Where dyads form edges in a graph, multi-party structures form simplices in a simplicial complex. A 2-simplex (triangle) connects three parties; a 3-simplex (tetrahedron) connects four.

Think of three signatures on a transit agreement. A supplier may deal bilaterally with a buyer; each may also have a separate agreement with the transit state. Those are three edges. A single undertaking that comes into force only when all three sign is a different object. A filled triangle can record that joint commitment in addition to its boundary of bilateral ties. Drawing the three edges does not establish that the joint undertaking exists. The analyst needs evidence of the shared obligation before filling the face.

Define triangular helicity for three parties A, B, C:

$$H_{ABC} = \oint_{\triangle} \mathbf{F} \cdot d\mathbf{l}$$

where $\mathbf{F}$ represents economic flux along the chosen oriented route $A \to B \to C \to A$. Read the integral as adding the signed contribution along each leg. Reversing the route changes its sign. A nonzero result records directional imbalance around that circuit; it can prompt an investigation of who gains leverage or bears cost. Establishing who benefits also requires prices, obligations, and alternatives. The boundary calculation can be performed even when no three-party agreement fills the triangle.

Early application to the Saudi-Iran-Yemen triangle suggests triangular helicity adds predictive signal beyond bilateral metrics alone. The 2015 Saudi intervention coincided with triangular helicity depletion as Iran increased Yemen support while Saudi-Yemen direct ties collapsed.

With 200 states there are $200(199)/2 = 19{,}900$ unordered pairs and $200(199)(198)/6 = 1{,}313{,}400$ unordered triples. A triple enters the research when there is evidence of a joint undertaking, such as a release schedule binding all three parties. That evidence is what makes the increase in the number of possible relationships a practical research problem.

The clock on the observation

The current framework operates on annual or quarterly data, generating alerts with multi-month lead times. Policy utility would increase substantially with higher-frequency monitoring.

Official trade statistics arrive after the period they describe, and later revisions can change the record. A real-time replay must attach a publication date to every input. For a proposed December 2021 warning, a record describing the third quarter is usable only if it had actually been released by the alert date. A quarter label alone cannot establish that availability.

Several high-frequency proxies can supplement official statistics:

More frequent observations can shorten the delay between a change in the relationship and its appearance in official statistics. AIS vessel records can reveal shifts in shipping weeks earlier. Commercial flight schedules provide another rapidly changing connection between countries. SWIFT messaging volumes, exchange rates, and bond spreads offer a financial view, while satellite images expose activity at ports, border crossings, and military positions.

These observations do not measure the same thing. A cancelled flight, a widened bond spread, and a queue of ships become useful together only when the analyst explains which part of the relationship each represents. Their speed is an advantage; it does not remove the need for that mapping.

The financial implementation (US Patent Application 63/935,782) computes helicity at daily and intraday frequencies for market prediction. The core insight: economic flux and alignment can be measured at multiple timescales.

Extending this to geopolitical monitoring requires mapping financial-frequency indicators to conflict-relevant signals. Credit default swap spreads on sovereign debt, for instance, aggregate market expectations about geopolitical risk and update continuously.

Current research develops a three-tier monitoring system:

The proposed reporting rhythm gives each source a job it can support. An annual strategic assessment recomputes helicity from official data. Monthly updates use alternative observations to estimate changes in flows between those releases. Weekly monitoring follows financial and shipping indicators for emerging alerts. The faster layer should show what has changed since the slower assessment, rather than quietly replacing it with a different measure.

An hourly dashboard can display information that is months old. A useful system should show the observation time, release time, and age

of each input, as well as the latest calculation. Faster computation cannot repair an unseen change in the underlying relationship.

The goal is not replacing annual computation but supplementing it with faster-updating partial signals that trigger more frequent comprehensive analysis when anomalies appear.

The framework provides *structural* features; machine learning can optimize how those features are combined and thresholded. This represents a middle path between pure structural reasoning and pure pattern recognition.

A forecasting model can fail when new cases differ from its training data. Geometric features offer one way to describe constraints that may recur across settings, but their portability has to be tested too. Compare a model using the original observations with one using those observations plus the geometric construction, on the same held-out cases. That comparison asks whether the added structure improves the forecast rather than merely making its explanation more appealing.

Fixed alert thresholds may also be suboptimal. Different dyads can have different baseline dynamics. Machine learning offers a way to test whether thresholds adapted to dyad characteristics improve held-out performance while preserving the structural features.

The helicity framework generates a rich feature set for ML models:

The feature set can preserve several aspects of a relationship at once. Levels record helicity H, flux V, alignment A, and their standardized values. Changes record $\Delta H/\Delta t$, the rate of alignment collapse, and the acceleration of flux growth. Spectral features describe network structure through eigenvalue gaps, heat-kernel signatures, and Cheeger constants. Trajectory features add the distance from historical conflict paths and the time spent at each alert level. Together, these measurements distinguish where a relationship is, how it is moving, and which routes remain open.

The preliminary gradient-boosting comparison reports 82% accuracy. Interpreting that figure requires case counts, event prevalence, the train/test split, and errors from both fixed-threshold and raw-indicator baselines. A model can score well by predicting quiet outcomes while missing rare wars.

A Bayesian model could express uncertainty over alert categories: an illustrative output assigns 87% to L4, 11% to L3, and 2% elsewhere. These are category probabilities, not automatically probabilities of war. Calibration against outcomes must precede their use to justify different actions.

A Gaussian-process reconstruction reportedly places Russia–Ukraine outside a 95% band in November 2021. Reproduction requires its mean, covariance kernel, noise model, and information cutoff. A band for the

latent trajectory differs from a predictive interval for a future observation. Crossing either indicates a model-dependent departure; it does not identify the political outcome.

Unlike black-box predictors, helicity features remain interpretable. A policymaker can understand why the system issues an alert ("helicity collapsed due to sanctions reducing alignment while trade continued") rather than accepting an opaque risk score. This interpretability is essential for policy adoption—officials need to explain decisions, not just cite algorithms.

Climate and Resource Conflicts

Climate change will reshape the conflict landscape over the coming decades. The helicity framework can extend to capture resource-flow dynamics that traditional security analysis misses.

Water, arable land, and stable climate are becoming scarce resources in many regions. The World Bank estimates that climate change could displace 140 million people by 2050; the resulting migration flows will stress recipient countries and potentially trigger conflict.

Traditional analysis treats resource scarcity as a background condition. The helicity framework can make it central: flows of water, food, and people constitute economic flux, and agreements about resource sharing constitute alignment. When water flux grows (upstream damming, drought, population growth) while water agreements collapse, the same trap dynamics emerge.

Changes in water availability, agricultural production, and displacement can alter the material conditions under which agreements operate. The research question is whether institutions and alternative supplies can adjust within the time available.

A river agreement makes the problem concrete. If expected inflow falls, the same promised release may become harder to deliver. The resulting conflict depends on storage, alternative sources, domestic needs, and how the parties interpret compliance. A shortage is an observation to explain, not a sufficient prediction of war.

Political choices mediate the response. Governments can share a shortfall, revise a schedule, invest in alternatives, or shift the burden onto others. The geometric account should expose which responses remain feasible and who bears their costs.

Displacement adds further demands on housing, services, land, and political representation. Institutions can respond in ways that widen or narrow peaceful options. A model should be evaluated on those

mechanisms and outcomes, rather than treat movement of people as a forecast of violence by itself.

Consider a dam operator and a downstream water authority reading the same river. During filling, part of the inflow goes into storage. The water available downstream depends on releases, losses and inflows along the rest of the river. A yearly total can conceal a shortage during the weeks when irrigation water is needed. The Nile dispute raises both a physical question about the timing of water and a political question about how the parties share decisions.

Reformulated for resources:

$$H_{\text{water}} = (\text{Water dependency}) \times (\text{Treaty compliance})$$

The sign requires a definition of treaty compliance. If it is the fraction of scheduled releases that occurred, both factors in this product are nonnegative. For a signed version, one could instead score fulfillment relative to an agreed benchmark. Suppose the promised release is 100 units and 80 arrive. The fractional departure is $(80 - 100)/100 = -0.20$. With dependency 0.75, the resulting signed exposure is -0.15. Those numbers illustrate a convention. An application to the Nile would need the actual commitments, flow records and dates. Egypt's dependence also persists when a diplomatic agreement is difficult to reach.

Migration flows constitute another flux type. Helicity between origin and destination countries:

$$H_{\text{migration}} = (\text{Migration flow}) \times (\text{Political alignment on migration})$$

When refugee flows increase while political accommodation decreases, the trap forms. European 2015 migration crisis exhibited this pattern: flows surged while political consensus on burden-sharing collapsed, creating intra-European tensions that persist today.

The research frontier involves integrating multiple flux types—trade, energy, water, migration—into a unified assessment. Different flux types may substitute or compound. High trade helicity might buffer water tension; alternatively, compounding negative helicity across multiple dimensions might trigger conflict faster than any single flux would predict.

Who gets to see the instrument

A final question confronts any prediction technology: Should the tools be public or proprietary? This question carries particular weight for conflict prediction, where the stakes are measured in human lives.

Transparency enables verification. The scientific community cannot evaluate the framework's validity without access to implementation details. Claims of 78% precision remain unverified assertions without replicable code.

Democratization of early warning. Currently, advanced conflict forecasting is available only to wealthy governments, major research institutions, and private intelligence firms. Open tools would enable civil society organizations, journalists, and affected populations to monitor their own risk.

Improvement through collaboration. Open source development attracts contributions from diverse researchers, potentially improving the framework faster than proprietary development.

The moral case. If the framework genuinely provides actionable warning time, restricting access means knowingly allowing preventable deaths for commercial advantage. This is difficult to justify.

There is a competing argument for restricting access. *Gaming and manipulation.* A state could improve visible indicators while preserving a coercive plan. Monitoring should therefore examine whether a reported concession changes an actual capability or commitment. Publishing an indicator also creates an incentive to influence it.

Misinterpretation. A high alert can be mistaken for certainty, or a structural score for a probability. The interface should expose the observation dates, assumptions, and intended use. A warning category cannot by itself justify military action.

Dual-use risk. The same tools that enable conflict warning could be repurposed for offensive intelligence applications—identifying when an adversary is vulnerable to pressure.

Commercial considerations. Development required substantial investment; open release eliminates commercial potential.

This book provides complete theoretical and mathematical exposition. The equations, the thresholds, the interpretation framework—all are disclosed. An expert reader could implement the system from these specifications.

What remains proprietary is the production software: the data pipelines, the real-time computation infrastructure, and the validated parameter sets. This middle path enables academic replication while preserving some implementation advantage.

We anticipate releasing a basic open-source implementation within two years, pending further development of safeguards against misuse. The ultimate goal is a world where every country can monitor its own helicity trajectory—knowing when they are drifting toward the trap, and taking action before geometry breaks.

The future of this framework, like the future of the relationships it monitors, remains unwritten. What is certain is that the mathematics will continue to evolve, the data will continue to accumulate, and the geometry will continue to speak to those who learn its language.

Epilogue: The Geometry Is Screaming

When I return to the morning of February 24, 2022, the question that follows me is what could have been done while the invasion was still a possibility rather than a fact.

Looking backward with an equation does not make that morning easier to bear.

Grief for the lives that would be lost. Grief for the cities that would be destroyed. Grief for the futures that would be foreclosed. And beneath it all, a bitter question: *Could this have been prevented?*

The reconstruction follows a dangerous persistence: dependence continued as trust collapsed. Could the available information have revealed that pressure in time? Could a different response have prevented invasion? The calculation opens those questions. The historical evidence has to answer them.

But seeing is not preventing. A thermometer cannot cool the patient.

A warning followed by peace leaves an uncomfortable problem. Perhaps the alarm was excessive. Perhaps an intervention changed the course of events. Perhaps the parties would have stepped back anyway. The peaceful outcome alone cannot decide among those explanations.

For a policymaker, preparation has costs. Fuel reserves, negotiations, and evacuation plans take resources and political commitment. A useful warning helps people weigh those costs against the danger. It cannot promise that every precaution will prove necessary.

For a researcher, two records matter: what was forecast and what happened, then what people did and what those actions changed. A missed war stays a miss. A false alert stays in the score. Prevention deserves an investigation of its own; peace cannot rescue every prediction.

I wrote this book because I believe that knowledge should be shared. If there are tools that can provide early warning of catastrophic conflict, those tools should not remain locked in academic papers or classified briefings. They should be available to anyone who might use them to save lives.

This is not naïve. I know that warning systems can be gamed, that predictions can be weaponized, that knowledge cuts both ways. But I also know that silence is complicity. If you see the fire starting and say nothing, the burned are on your conscience.

The geometry is screaming. It has been screaming for decades, in conflict after conflict, war after war. We have not been listening because we did not have the right ears.

This book is an attempt to teach a different way of hearing.

The Taiwan assessment in this book records $H = -0.62$ and a warning below Level 4. Behind that snapshot are routes, commitments, and alternatives. Their subsequent improvement or deterioration has to be observed.

The window can narrow when a supplier disappears or an essential delivery becomes unreliable. It can widen when a replacement is found or a concession becomes possible. Following those changes takes more than extending a line on a chart.

Leaders can maintain a costly commitment, abandon it, compromise, or use force. The geometry helps explain what those choices require. Responsibility remains with the people who act.

I do not know whether the Taiwan Strait will see war in the coming decade. I want this research to reveal the choices still available while people have time to act.

This is not prophecy. It is geometry.

And geometry can be changed.

Bee Rosa Davis
2026

Appendix A

Data Sources

This appendix identifies the main data families behind the historical analysis and expanded methods. A source name begins a reproducibility record: the particular file, release, codes, units, and acquisition date identify the observations actually used. A provider's current download can differ from the vintage available when a historical assessment was made. The implementation chapter explains how to preserve that distinction.

The descriptions identify what each resource can supply. They do not assert that every variable in a current release entered every historical calculation. Access conditions and coverage vary by product and release; the retained evidence register governs a particular result.

Economic Data

UN Comtrade

The United Nations Comtrade database is maintained by the United Nations Statistics Division (UNSD). Its release documentation matters when reconstructing which trade observations were available at a particular time (United Nations Statistics Division 2026b; United Nations Statistics Division 2026a).

Producer: United Nations Statistics Division (UNSD), Department of Economic and Social Affairs.

Methodology: National authorities submit merchandise trade statistics, which UNSD validates and harmonizes. Preserve classification edition, aggregation level, valuation convention, and estimation flags. Annual and monthly products can differ through revisions, confidentiality rules, and processing, including estimates for some missing attributes.

Coverage: Long historical series with annual and monthly products, subject to reporter and period availability. There is no universal six-month release lag; submission and revision schedules vary by reporter.

Variables Used: Bilateral trade flows (imports and exports by partner country), disaggregated by commodity code. We aggregate to total bilateral trade T_{AB} in current USD.

Access: Free online at `https://comtradeplus.un.org`. Bulk downloads available via API with registration.

Limitations: Mirror statistics compare A's exports to B with B's imports from A. Differences can reflect valuation, timing, classification, and reporting coverage. Record the discrepancy and reconciliation rule rather than assigning a universal percentage. The historical convention prefers reporter exports when available.

IMF Direction of Trade Statistics (DOTS)

The International Monetary Fund's Direction of Trade Statistics is now presented as International Trade in Goods by Partner Country (IMTS). It supplies merchandise exports and imports by partner in annual, quarterly, and monthly products (International Monetary Fund 2026). Historical DOTS files retain their original release identities in a replication record.

Producer: International Monetary Fund, Statistics Department.

Methodology: IMF collects data from national statistical agencies and central banks, then applies partner-country estimation procedures to fill gaps. Unlike Comtrade, DOTS imputes missing bilateral flows using mirror statistics.

Coverage: 1948–present, approximately 190 countries. Monthly and quarterly data available for recent years; annual only for historical data.

Variables Used: Export and import values by partner, used to cross-validate Comtrade data and extend coverage pre-1962.

Access: IMF Data Portal (`https://data.imf.org`), free with registration. Historical archives available to academic institutions.

Limitations: Imputation procedures can mask true data quality issues. We prefer Comtrade for post-1962 analysis.

World Bank World Development Indicators (WDI)

World Development Indicators provide economic context for normalizing bilateral flows. Indicator metadata specify definitions, units, sources, and coverage (World Bank 2026).

Producer: World Bank, Development Data Group.

Methodology: WDI aggregates data from national statistical systems, international organizations, and World Bank operational data. Each indicator follows documented methodology (available online). The Bank applies quality control and standardization procedures.

Coverage: 1960–present for most indicators, 217 economies, over 1,400 indicators.

Variables Used: GDP (current USD and PPP), GDP growth, FDI inflows/outflows, trade openness (trade/GDP). Used for computing global trade means $\bar{T}_{\text{global}}$ and economic scale normalization.

Access: Free online at `https://databank.worldbank.org/source/world-development-indicators`. Bulk download and API available.

Limitations: Some series have gaps for conflict-affected countries. GDP estimates for non-market economies (especially historical USSR data) are uncertain.

International Energy Agency (IEA)

The IEA supplies energy balances, supply and demand series, and related indicators used to describe energy dependence. Its products include free and paid datasets; the product and release must accompany the resulting series (International Energy Agency 2026).

Producer: International Energy Agency, Paris (an autonomous agency within the OECD framework).

Methodology: IEA assembles and assesses national energy statistics under documented product definitions. Preserve the energy product, physical or energy-equivalent units, balance convention, and country coverage. Oil, gas, coal, and electricity are not interchangeable delivery channels merely because they share a unit in a dataset.

Coverage: 1971–present for IEA members, 1990–present for major non-members. Monthly, quarterly, and annual frequency.

Variables Used: Energy imports by source country, energy dependency ratios (net imports / total primary energy supply), oil and gas trade values.

Access: IEA Data and Statistics portal (`https://www.iea.org/data-and-statistics`). Some data requires subscription; key aggregates are free.

Limitations: Coverage outside OECD is incomplete. We supplement with BP Statistical Review for gaps.

UNCTAD Statistics

UN Trade and Development (UNCTAD) supplies trade and investment statistics, including foreign direct investment stocks and flows (UNCTAD 2024). A bilateral position, annual flow, and total inward investment measure answer different questions and must retain their definitions.

Producer: UNCTAD, Geneva.

Methodology: UNCTAD compiles FDI statistics from national sources and applies standardization based on IMF Balance of Payments method-

ology. The organization also maintains commodity price indices and trade-in-services data.

Coverage: 1970–present for FDI, 196 economies.

Variables Used: Bilateral FDI stocks and flows, used as proxy for F_n (FDI integration component) in curvature calculation.

Access: UNCTADstat (`https://unctadstat.unctad.org`), free with registration.

Political Data

Source	Coverage	Variables
UN Voting (Bailey)	1946–present	Voting similarity
ATOP	1815–2016	Alliance treaties
COW Diplomatic	1817–2005	Ambassador relations
TIES	1945–2005	Economic sanctions

Table A.1: Primary political data sources.

United Nations General Assembly Voting Data

UN voting patterns provide an observable measure used in computing S_{vote} (Voeten, Strezhnev, and Bailey 2024). Raw roll-call agreement and estimated ideal-point distance are distinct constructions. Their file versions, handling of abstentions, and eligible resolutions must be recorded separately.

Producers: Erik Voeten (Georgetown University), Anton Strezhnev (University of Chicago), and Michael Bailey (Georgetown University).

Methodology: The dataset codes all recorded roll-call votes in the UN General Assembly since 1946. Each resolution is coded by topic (colonialism, human rights, nuclear weapons, etc.). The key variable is the "ideal point" estimate derived from Item Response Theory (IRT) models, which places each country on a latent preference dimension. Bilateral similarity is computed as the correlation of voting records or distance between ideal points.

Coverage: 1946–present (updated annually), all UN member states.

Variables Used: Bilateral voting similarity $S_{\text{vote}} \in [0, 1]$, computed as the fraction of resolutions where both countries voted the same way (Yes-Yes or No-No), excluding abstentions.

Access: Harvard Dataverse (`https://dataverse.harvard.edu/dataset.xhtml?persistentId=hdl:1902.1/12379`). Also available via R package `unvotes`.

Limitations: UNGA votes are non-binding, so countries may vote symbolically. Great powers often vote against the majority on decolonization and human rights resolutions, creating persistent low similarity scores that may not reflect bilateral hostility.

Alliance Treaty Obligations and Provisions (ATOP)

ATOP supplies alliance provisions and coding documentation. Its codebook and treaty codesheets let readers examine what obligation a recorded alliance establishes (Alliance Treaty Obligations and Provisions Project 2022).

Producers: Brett Ashley Leeds (Rice University) and colleagues.

Methodology: ATOP codes all bilateral and multilateral defense pacts, ententes, neutrality agreements, and non-aggression pacts from 1815–2016. Each treaty is coded for specific provisions (defense, offense, neutrality, consultation) and conditions (unconditional, conditional on attacker identity, etc.).

Coverage: 1815–2016, all states in the Correlates of War system.

Variables Used: Alliance membership indicator, weighted by treaty type. Defense pacts receive weight 1.0, ententes 0.5, non-aggression pacts 0.3. R_{treaty} is the maximum alliance weight between dyad members.

Access: ATOP Project website (`https://www.atopdata.org`), with versioned files and documentation.

Limitations: The historical coverage window stated here belongs to the release used in the analysis, not a claim that the project stopped updating. Any manually coded extension must retain the treaty text, date, coder, and rule separately from the downloaded file. The framework's numerical treaty weights are modeling choices, not weights assigned by ATOP.

Correlates of War Diplomatic Exchange Dataset

The COW Diplomatic Exchange dataset tracks representation at specified observation years. Version 2006.1 and its codebook distinguish ambassadorial, ministerial, and chargé-level representation (Bayer 2006).

Producer: Correlates of War Project (various principal investigators since 1963).

Methodology: The dataset codes whether country A sends a diplomatic representative to country B, and the level of that representative

(ambassador, minister, chargé d'affaires, or other). Coding is based on official government records and diplomatic lists.

Coverage: Selected observation years from 1817 to 2005, including five-year intervals in substantial parts of the series. These are not complete annual observations; carrying a value forward requires an explicit assumption.

Variables Used: Representation level: 3 = ambassador, 2 = minister, 1 = chargé, 0 = no evidence of exchange, and 9 = other. Code 9 is a category, not a level higher than ambassador. Any numerical measure of channel strength is a subsequent model construction.

Access: Correlates of War website (`https://correlatesofwar.org/data-sets/diplomatic-exchange/`), free download.

Limitations: Coverage ends in 2005. We supplement with Lowy Institute Global Diplomacy Index for recent years.

Global Sanctions Database (GSDB)

The Global Sanctions Data Base supplies versioned cases, measure types, objectives, and outcomes (Global Sanctions Data Base 2026). A case record and the framework's penalty $\sigma(X)$ are different objects: the numerical mapping from one to the other must be documented.

Producers: Gabriel Felbermayr, Aleksandra Kirilakha, Constantinos Syropoulos, Erdal Yalcin, and Yoto Yotov.

Methodology: The GSDB codes all economic sanctions episodes since 1950, including sender(s), target(s), start/end dates, sanction types (trade, financial, arms, travel), and stated objectives. Each episode is coded for intensity based on the scope of measures imposed.

Coverage: Cases beginning in 1950, with endpoint and case count determined by release. Releases can add, merge, or remove cases; preserve the downloaded version rather than combining an old count with a current endpoint.

Variables Used: Sanctions intensity score $X_{\text{sanction}} \in [0, 1]$, constructed from the number and type of sanctions in force. Full comprehensive sanctions = 1.0, targeted sanctions = 0.3–0.7, no sanctions = 0.

Access: Global Sanctions Database (`https://globalsanctionsdatabase.com`), free for academic use.

Alternatives: TIES (Threat and Imposition of Economic Sanctions) dataset provides earlier coverage (1945–2005) with different coding rules. We use GSDB for post-2005 and TIES for historical robustness checks.

Threat and Imposition of Economic Sanctions (TIES)

TIES separately records threats and impositions, an important distinction when a threatened measure changes behavior before it is imposed. Version 4 covers cases initiated during 1945–2005 (Morgan, Bapat, and Kobayashi 2014).

Producers: T. Clifton Morgan and Navin Bapat, with co-investigators Valentin Krustev and Yoshiharu Kobayashi.

Methodology: TIES codes sanctions threats and impositions, distinguishing between the two. Each case is coded for sender, target, issue area, and outcome. The dataset emphasizes the coercive diplomacy aspect of sanctions.

Coverage: 1945–2005, approximately 1,400 cases.

Variables Used: Used for historical validation (pre-2000 conflicts) and robustness checks.

Access: TIES Project (`https://sanctions.web.unc.edu`), free download.

Conflict Data

- UCDP/PRIO Armed Conflict Dataset
- Correlates of War (MID, Interstate War)
- ACLED (sub-state violence)

UCDP/PRIO Armed Conflict Dataset

The UCDP/PRIO Armed Conflict Dataset provides annual conflict records. UCDP's download center supplies separate datasets, codebooks, and historical versions; annual onset, dyadic participation, and georeferenced events must not be treated as the same outcome (Uppsala Conflict Data Program 2026).

Producers: Department of Peace and Conflict Research, Uppsala University (Sweden) and PRIO (Norway).

Methodology: UCDP defines armed conflict as "a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in a calendar year." Coding is based on news sources, NGO reports, and academic literature, with multiple coders and inter-coder reliability checks.

Coverage: 1946–present (updated annually), global coverage.

Variables Used: Conflict onset date (first year meeting 25-death threshold), conflict type (interstate, intrastate, internationalized intrastate), intensity level (minor = 25–999 deaths/year, war = 1,000+ deaths/year).

Access: UCDP website (`https://ucdp.uu.se`), free download. Georeferenced Event Dataset (GED) provides sub-national event data.

Limitations: The 25-death threshold misses low-intensity conflicts and coercive crises that don't result in casualties (e.g., naval standoffs).

Correlates of War Militarized Interstate Disputes (MID)

The COW Militarized Interstate Disputes dataset captures threats, displays, and uses of force that need not meet UCDP's armed-conflict threshold. Its codebook defines the unit, hostility scale, and actor records (Correlates of War Project 2026).

Producer: Correlates of War Project.

Methodology: MIDs group qualifying interstate threats, displays, and uses of military force into disputes. The hostility scale distinguishes no militarized action (1), threat (2), display (3), use of force (4), and war (5). Dispute-level and participant-level records require the corresponding file definitions.

Coverage: 1816–2014 (version 5.0), all states in international system.

Variables Used: MID onset, hostility level, fatality level. Used to validate framework sensitivity to lower-intensity disputes.

Access: COW website (`https://correlatesofwar.org/data-sets/mids/`), free download.

Armed Conflict Location and Event Data (ACLED)

ACLED supplies disaggregated records of political violence and protest, with source and coding guidance useful for monitoring and retrospective analysis (ACLED 2026c). Event occurrence and publication are different times.

Producer: ACLED, a non-profit organization headquartered in the United States.

Methodology: ACLED codes individual conflict events (battles, protests, riots, violence against civilians, etc.) from news sources, with daily updating. Each event is geocoded and attributed to specific actors.

Coverage: Start dates vary by country and region. Record the window and update schedule for the requested extract; a globally available product does not imply uniform historical coverage.

Variables Used: Event counts and fatality estimates for real-time monitoring. Not used for historical validation (insufficient coverage) but valuable for dashboard applications.

Access: ACLED website (`https://acleddata.com`), free for academic and humanitarian use.

Data Integration and Quality

Integrating these diverse sources requires careful attention to:

Temporal Alignment: Economic data is typically annual; political data varies from daily (ACLED) to annual (UNGA voting). We align all data to annual frequency for the main analysis, with quarterly resolution for near-term monitoring.

Country Code Matching: Different sources use different country coding systems (ISO 3166, COW codes, UN codes). We maintain a crosswalk table and manually verify matches for historical entities (USSR, Yugoslavia, etc.).

Missing Data: Rather than impute missing values, we flag dyad-years with incomplete data and report confidence intervals that reflect data uncertainty.

Replication: The project repository retains processing code and research records. Appendices F and G identify the numerical companions. Reproduction requires the exact source files, releases, transformations, and inputs. Public archives should supply verified persistent identifiers; local records retain their titles, dates, and versions until deposited.

Appendix B

Computational Methods

Follow a single pair of countries through `analysis/psychohistory.py`. The code first measures trade, energy and political alignment. It then chooses a calculation according to how closely the economies are connected. The order matters because several conditions can hold at once. The formulas below specify the current implementation. The signed historical calibration and the mass-gap study have their own definitions in Chapter 4 and Appendix E.

Flux Computation

The flux V_{AB} adds two ratios. One compares bilateral trade with a reference trade volume. The other compares energy purchases with a reference energy volume. Both ratios are dimensionless.

Algorithm 1: Bilateral Flux

1. **Input**: Bilateral trade T_{AB} (USD billions), energy dependency E_{AB} (USD billions), global trade mean $\bar{T}_{\text{global}}$, critical energy threshold $\bar{E}_{\text{crit}}$
2. **Normalize trade**: $T_{\text{norm}} = T_{AB}/\max(\bar{T}_{\text{global}}, 10^{-6})$
3. **Normalize energy**: $E_{\text{norm}} = E_{AB}/\max(\bar{E}_{\text{crit}}, 10^{-6})$
4. **Compute flux**: $V = T_{\text{norm}} + \alpha \cdot E_{\text{norm}}$
5. **Output**: $V \in [0, \infty)$

Parameters:

- $\alpha = 4.0$ — Strategic weight reflecting GDP impact ratio of energy vs. trade
- $\bar{T}_{\text{global}} = 50.0$ billion USD (default baseline)
- $\bar{E}_{\text{crit}} = 10.0$ billion USD (default threshold)

At the default settings, another ten billion dollars of energy raises V by four units if the trade input stays fixed. An extra fifty billion in the trade input raises it by one. The weight sets how strongly the indicator responds to each entry. Estimating the GDP loss from an actual cutoff also requires the buyer's inventories and replacement suppliers.

Alignment Computation

The alignment term A_{AB} measures topological orientation—whether two nations are moving toward cooperation or conflict. The hyperbolic tangent provides smooth saturation to $[-1, 1]$.

Algorithm 2: Bilateral Alignment

1. **Input**: UN voting similarity $S_{\text{vote}} \in [0, 1]$, treaty ratio $R_{\text{treaty}} \in [0, 1]$, sanctions intensity $X_{\text{sanction}} \in [0, 1]$
2. **Positive alignment**: $A^+ = (S_{\text{vote}} + R_{\text{treaty}})/2$
3. **Sanction penalty**: $\sigma(X) = 1 - \exp(-\beta X_{\text{sanction}})$
4. **Raw alignment**: $A_{\text{raw}} = A^+ - \sigma(X)$
5. **Saturate**: $A = \tanh(k \cdot A_{\text{raw}})$
6. **Output**: $A \in [-1, 1]$ (strictly bounded)

Parameters:

- $k = 2.0$ — Alignment steepness factor
- $\beta = 1.0$ — Sanction penalty decay rate

At the stated weights, moving sanctions from zero to one lowers raw alignment by $1 - e^{-1} \simeq 0.632$. The slope of that penalty declines as sanctions increase. With all three inputs in their specified intervals, the final alignment runs from $\tanh[-2(1 - e^{-1})] \simeq -0.852$ to $\tanh(2) \simeq 0.964$.

Tolerance Budget τ

The tolerance budget τ quantifies a relationship's capacity to absorb diplomatic shocks without structural failure. It aggregates four components of diplomatic resilience.

Algorithm 3: Tolerance Budget

1. **Input**: Alliance strength $A_s \in [0, 1]$, institutional ties $I_t \in [0, 1]$, historical trust $H_t \in [0, 1]$, communication channels $C_c \in [0, 1]$
2. **Weighted sum**: $\tau = \beta_1 A_s + \beta_2 I_t + \beta_3 H_t + \beta_4 C_c$
3. **Clip**: $\tau = \min(\max(\tau, 0), 1)$
4. **Output**: $\tau \in [0, 1]$

Weight Parameters:

- $\beta_1 = 0.30$ — Alliance (formal treaty obligations)
- $\beta_2 = 0.25$ — Institutional (shared memberships: UN, WTO, etc.)
- $\beta_3 = 0.25$ — Historical (accumulated trust from prior cooperation)
- $\beta_4 = 0.20$ — Communication (active diplomatic channels)

Curvature K

Curvature K measures the "stiffness" of economic integration—how difficult it would be to decouple. High curvature means the relationship has deep structural dependencies.

Algorithm 4: Integration Curvature

1. **Input**: Trade volume T, energy dependency E, global means $\bar{T}$, $\bar{E}$
2. **Normalize trade**: $T_n = \min(T/\max(2\bar{T}, 10^{-6}), 1)$
3. **Normalize energy**: $E_n = \min(E/\max(2\bar{E}, 10^{-6}), 1)$
4. **FDI proxy**: $F_n = (T_n + E_n)/2$
5. **Supply chain proxy**: $S_n = 0.8 \cdot T_n$
6. **Weighted sum**: $K = \alpha_1 T_n + \alpha_2 E_n + \alpha_3 F_n + \alpha_4 S_n$
7. **Clip**: $K = \min(\max(K, 0), 1)$
8. **Output**: $K \in [0, 1]$

Weight Parameters:

- $\alpha_1 = 0.35$ — Trade integration
- $\alpha_2 = 0.35$ — Energy integration
- $\alpha_3 = 0.15$ — Foreign direct investment
- $\alpha_4 = 0.15$ — Supply chain entanglement

Spectral Gap Computation

The spectral gap computation leverages well-established numerical linear algebra methods. For networks up to approximately 200 nodes (typical for regional subsystems), direct eigenvalue computation is feasible; for larger networks, iterative methods provide excellent approximations.

Algorithm 4b: Network Spectral Gap

1. Construct adjacency matrix from bilateral trade
2. Compute degree matrix D
3. Form Laplacian $L = D - A$
4. Compute eigenvalues using Lanczos iteration
5. Return λ_1 (second smallest)

For this construction use a symmetric matrix with nonnegative weights, zero diagonal and a fixed list of countries. The gap of $D - A$ is in the edge-weight units. A zero second eigenvalue means that the graph has at least two components, including any isolated nodes. The ratio of

that gap to the largest eigenvalue and the degree-normalized gap in Chapter 5 use different scales.

Helicity Computation

Two helicity formulations exist: the legacy v1.0 empirical formula and the v2.0 Davis Framework formula. Both are implemented for backward compatibility.

Algorithm 5a: Legacy Helicity (v1.0)

1. **Input**: Flux V, alignment A
2. **Compute**: $H = V \times A$
3. **Output**: $H \in (-\infty, \infty)$

Algorithm 5b: Davis Helicity (v2.0)

1. **Input**: Tolerance budget τ, curvature K, default $\varepsilon_H = 10^{-6}$
2. **Compute**: $H = \tau/(K + \varepsilon_H)$
3. **Output**: $H \geq 0$. On the ordinary branch $K \geq 0.05$, so $H \leq 1/(0.05 + \varepsilon_H) < 20$.

The capacity ratio compares tolerance with the regularized integration burden. At $H = 1$ they match. A value above one means that tolerance is larger on these scales. The alert also depends on alignment, the pair's baseline and the applicable decision rule.

Edge Case: When $K < K_{\text{disc}} = 0.05$, the system enters DISCONNECTION mode and the Hostility Index replaces helicity (see the disconnection discussion in Chapter 4).

Trichotomy Parameter Γ

The trichotomy parameter compares counted diplomatic constraints and tolerance with integration and the number of exits. The following bands show the alert selected from Γ alone on the ordinary branch, after the higher-priority tests in Algorithm 10 have been considered.

Algorithm 6: Trichotomy Parameter

1. **Input**: Diplomatic constraints m, tolerance budget τ, curvature K, and the historical count field $|S_{\text{exit}}|$
2. **Guard**: If $|S_{\text{exit}}| \leq 1$, return $\Gamma = 0$ (the implementation's singularity guard)
3. **Compute log**: $\log|S| = \ln(\max(|S_{\text{exit}}|, 2))$
4. **Denominator**: $D = K \cdot \log|S| + 10^{-6}$

5. **Compute**: $\Gamma = (m \cdot \tau)/D$
6. **Output**: $\Gamma \in [0, \infty)$

This algorithm is a replay specification. Its legacy field name does not confer monotonic benefit on a count of usable exits: for counts at least two, increasing the count lowers the ratio with the other inputs held fixed. Section 22 separates that implemented quantity from candidate-space resolution and from settlement feasibility. The guard is a discrete convention and does not arise by taking the ideal ratio's limit at one.

Regime Classification:

Γ Range	Regime	Interpretation
$\Gamma \geq 1.5$	STABLE	Capacity-dominated range
$1 \leq \Gamma < 1.5$	ELEVATED	Near the warning band
$0.5 \leq \Gamma < 1$	WARNING	Reduced room to adjust
$0.1 \leq \Gamma < 0.5$	CRITICAL	Deep underdetermination
$0 \leq \Gamma < 0.1$	SINGULARITY	Structural trap reading

Spectral Gap λ_1

The ordinary pipeline reads the network gap supplied to the manifold. A fresh manifold initializes that value to 1.0 until a network estimate is supplied. The separate spectral helper accepts a transition matrix or builds a two-state example from tolerance. The origin of the gap tells a reader what it measures.

Algorithm 7: Operational Spectral Gap

1. **Input**: Transition matrix M (from trust dynamics)
2. **Eigenvalues**: Compute $\{\lambda_i\}$ of M
3. **Sort**: Order by $|\lambda_i|$ descending
4. **Extract**: λ_2 = second-largest eigenvalue
5. **Compute**: $\lambda_1 = 1 - |\lambda_2|$
6. **Output**: $\lambda_1 \in [0, 1]$

For the helper's two-state example the transition matrix is

$$\tau_n = \min\left(\frac{\tau}{\tau + 1}, 0.95\right) \tag{B.1}$$

$$p_{\text{stay}} = 0.5 + 0.45 \cdot \tau_n \tag{B.2}$$

$$M = \begin{pmatrix} p_{\text{stay}} & 1 - p_{\text{stay}} \\ 1 - p_{\text{stay}} & p_{\text{stay}} \end{pmatrix} \tag{B.3}$$

Imagine two desks passing a file between them. Each step a file stays put with probability p_{stay} and switches desks otherwise. The eigenvalues are 1 and $2p_{\text{stay}} - 1$. With $\tau = 0.5$, $p_{\text{stay}} = 0.65$ and the helper's absolute gap is $1 - 0.30 = 0.70$. For τ between zero and one the helper spans $[0.55, 1]$. A regional reading of 0.05 therefore requires a different construction. The ordinary pipeline uses its supplied network gap, while Algorithm 14 defines a separate mobility proxy when the economies disconnect.

Activation Energy E_a

The activation scale quantifies resistance to reconfiguration. Its political direction is interpreted with alignment and the applicable regime.

Algorithm 8: Activation Energy

1. **Input**: Helicity H, default scale $\gamma = 1$, default $\varepsilon_E = 10^{-8}$
2. **Compute**: $E_a = \gamma/(H^2 + \varepsilon_E)$
3. **Output**: $0 < E_a \leq \gamma/\varepsilon_E$ for finite H and positive constants.

Interpretation:

- Increasing $|H| \Rightarrow$ decreasing E_a for fixed positive γ and ε.
- Decreasing $|H| \Rightarrow$ increasing E_a.
- $H \to 0$ gives $E_a \to \gamma/\varepsilon_E$, the maximum under the stated formula.

The formula is even in H: it does not distinguish cooperative from hostile reconfiguration. That distinction comes from the signed alignment and regime classification, not from E_a alone.

Divergence Index Δ

The divergence index compares two rates over the same observation interval. Read their signs before taking the quotient. With chronologically ordered records the routine uses elapsed whole days divided by 365, with a minimum interval of 0.01 years. For a first observation it returns a placeholder zero because there is no preceding pair of readings.

Algorithm 9: Divergence Index

1. **Input**: Current state (K_t, τ_t), previous state (K_{t-1}, τ_{t-1}), time delta Δt (years)
2. **Rate of trust change**: $d\tau/dt = (\tau_t - \tau_{t-1})/\Delta t$
3. **Rate of integration change**: $dK/dt = (K_t - K_{t-1})/\Delta t$
4. **Compute**: $\Delta = -\frac{d\tau/dt}{dK/dt + 10^{-6}}$

5. **Output**: $\Delta \in (-\infty, \infty)$

   Interpretation:

- If tolerance falls by 0.02 per year while integration rises by 0.01, the ratio is about +2. If both rates reverse, the ratio is again about +2. The two situations carry opposite political meanings.
- If both quantities fall, the ratio is generally negative. Their separate rates show which is falling faster. Acceleration requires comparing changes in a rate across intervals.
- A nearly stationary denominator makes the quotient sensitive to rounding. If $dK/dt = -10^{-6}$ the denominator is zero. At that point report the two rates separately.

Alert Level Determination

The historical chapters use a four-level warning vocabulary. The current implementation returns an alert enum with levels zero through five. It also stores a separate regime label. For example, at exactly $\Gamma = 1.5$ an ordinary case that reaches the final band selects STABLE, even though the regime label uses a strict upper-bound test and reads NEAR_CRITICAL.

Algorithm 10: Alert Level Classification

Read Algorithm 10 from top to bottom. The disconnected branch returns immediately. On the ordinary branch, a standardized trap takes precedence over the three-indicator test, which in turn precedes the Γ bands. The short names below refer to quantities already computed for this pair.

```
def determine_alert(state):
    g, gap = state.gamma, state.lambda_1
    H, A = state.helicity_h, state.alignment_a
    tau = state.tau
    zh, zv = state.z_helicity, state.z_flux
    if state.disconnection_mode:
        if A < -0.5 and tau < 0.1:
            if gap < 0.2:
                if g < 0.1: return SINGULARITY
                return CRITICAL
            if g < 0.5: return CRITICAL
            return WARNING
        if A < 0 and tau < 0.2:
            if g < 1: return WARNING
            return ELEVATED
```

```
        if A < 0: return ELEVATED
        return STABLE
    if A < 0 and zh < -2 and zv > 1:
        return SINGULARITY
    if zh < -2 and zv > 0.5:
        return CRITICAL
    if g < 1 and gap < 0.10 and H < 0.50:
        return SUPER_CRITICAL
    if g < 0.1: return SINGULARITY
    if g < 0.5: return CRITICAL
    if g < 1.0: return WARNING
    if g < 1.5: return ELEVATED
    return STABLE
```

Z-Score Computation:

$$Z_H = \frac{H - \mu_H}{\sigma_H} \tag{B.4}$$

$$Z_V = \frac{V - \mu_V}{\sigma_V} \tag{B.5}$$

$$Z_A = \frac{A - \mu_A}{\sigma_A} \tag{B.6}$$

The study design supplies the historical period and the observations available at the assessment date. When fitted from prior records, the engine uses their population standard deviations with floors of 0.10 for helicity and flux and 0.05 for alignment. Before a baseline is fitted it uses preset mean/standard-deviation pairs of $(0.5, 0.3)$, $(1, 0.5)$ and $(0.3, 0.2)$ respectively. A warning record should identify which baseline it used.

DISCONNECTION Mode

Below $K = 0.05$ the economic ratio becomes sensitive to a small denominator. The branch instead asks how hostile the security posture is and which channels still connect the parties. Exactly $K = 0.05$ belongs to the ordinary branch.

Algorithm 11: DISCONNECTION Detection

```
def is_disconnection(K):
    return K < 0.05  # K_DISCONNECTION threshold
```

Algorithm 12: Security Posture S

1. **Input**: Sanctions intensity X, alignment A, historical trust H_t

2. **Alignment hostility**: $A_h = \max(0, -A)$
3. **Historical distrust**: $D_h = 1 - H_t$
4. **Compute**: $S = 0.40 \cdot X + 0.35 \cdot A_h + 0.25 \cdot D_h$
5. **Clip**: $S = \min(\max(S, 0), 1)$
6. **Output**: $S \in [0, 1]$

Algorithm 13: Hostility Index H_d

1. **Input**: Security posture S, tolerance budget τ, $\varepsilon = 10^{-8}$
2. **Compute**: $H_d = S \cdot (1 - \tau)/(\tau + \varepsilon)$
3. **Output**: $0 \leq H_d \leq 1/\varepsilon$ for the stated input ranges. The state stores $-H_d$ in the shared helicity field.

Algorithm 14: DISCONNECTION Spectral Gap λ_1^d

1. **Input**: Communication channels C_c, diplomatic constraints m, exit paths $|S|$
2. **Normalize**: $m_n = \min(m/10, 1)$, $s_n = \min(|S|/8, 1)$
3. **Compute**: $\lambda_1^d = C_c \cdot m_n \cdot s_n$
4. **Clip**: $\lambda_1^d = \min(\max(\lambda_1^d, 0), 1)$
5. **Output**: $\lambda_1^d \in [0, 1]$

Algorithm 15: DISCONNECTION Trichotomy Γ_d

1. **Input**: m, τ, λ_1^d, S, $\varepsilon = 10^{-8}$
2. **Compute**: $\Gamma_d = (m \cdot \tau \cdot \lambda_1^d)/(S + \varepsilon)$
3. **Output**: $\Gamma_d \in [0, \infty)$

Algorithm 16: DISCONNECTION Activation Energy E_a^d

1. **Input**: S, τ, λ_1^d, scale $\gamma = 10$, $\varepsilon = 10^{-8}$
2. **Compute**: $E_a^d = \gamma \cdot S \cdot (1 - \lambda_1^d)/(\tau + \varepsilon)$
3. **Output**: $E_a^d \in [0, \infty)$

Software Stack

The implementation prioritizes reproducibility, transparency, and scalability:

- Python 3.11+ with NumPy, SciPy
- NetworkX for graph operations
- Modal for cloud computation
- PostgreSQL for time series storage

The core algorithm listings make the calculation inspectable. Reproducing a research result also requires its input observations, parameters, and event register. Publishing those together allows another researcher to distinguish a disagreement about the implementation from a disagreement about the model or the evidence.

Complete Analysis Pipeline

The following pseudocode follows the default capacity branch of the research engine and shows the quantities used in its final alert. A legacy option uses the signed product on the ordinary branch.

Read the code as a sequence of decisions about a pair of countries: measure what still moves between them, measure their alignment, compare available tolerance with the burden of integration, then ask whether the present readings are unusual for that pair. A single constructed row will let us follow those decisions all the way to the printed alert.

Algorithm 17: Full Dyad Analysis

```
def analyze_dyad(state, prev_state=None):
    # 1. Core economic metrics
    state.flux_v = compute_flux(state)
    state.alignment_a = compute_alignment(state)

    # 2. Davis Framework metrics
    state.tau = compute_tau(state)
    state.curvature_k = compute_curvature_v2(state)

    # 3. Check DISCONNECTION mode
    if state.curvature_k < 0.05:
        state.disconnection_mode = True
        state.security_posture = compute_security(state)
        state.hostility_index = compute_hostility(state)
        state.helicity_h = -state.hostility_index
        state.gamma = compute_gamma_disc(state)
        state.lambda_1 = compute_lambda1_disc(state)
        state.activation_energy = compute_Ea_disc(state)
    else:
        state.disconnection_mode = False
        state.helicity_h = compute_helicity_v2(state)
        state.gamma = compute_gamma(state)
        state.lambda_1 = supplied_network_gap
        state.activation_energy = compute_Ea(
            state.helicity_h)
```

```
# 4. Temporal metrics
state.divergence_index = compute_div(
    state, prev_state)

# 5. Z-scores for TRAP detection
zH, zV, zA = compute_z_scores(state)
state.z_helicity, state.z_flux = zH, zV
state.z_alignment = zA

# 6. Alert classification
state.alert_level = determine_alert(state)

return state
```

Suppose the ledger gives 50 billion dollars of bilateral trade and 10 billion of energy dependency. With the stated default reference amounts these give $T_{\text{norm}} = E_{\text{norm}} = 1$ in the flux calculation, so $V = 1 + 4(1) = 5$. The curvature calculation uses twice the reference amounts: its trade and energy terms are each 0.5, its FDI proxy is 0.5, and its supply-chain proxy is 0.4. Applying the four weights gives $K = 0.485$. The relationship clears the 0.05 disconnection gate, so the ordinary branch applies.

Suppose also that voting similarity is 0.4, the treaty ratio 0.2, and sanctions intensity 1. Positive alignment contributes 0.3 and the sanction penalty is $1 - e^{-1} \simeq 0.632$. With the factor of two inside tanh, $A \simeq -0.581$. Set each resilience input to 0.5. Then $\tau = 0.5$ and the engine's capacity reading is $0.5/(0.485 + 10^{-6}) \simeq 1.031$. Choosing 0.01 as a comparison regularizer gives 1.010. The legacy product on the same ledger is $VA \simeq -2.906$. Here economic activity can be large, diplomatic alignment negative and measured resilience positive at the same time.

Finally take $m = 2$ and a value of four in the historical exit-count field. The routine gives $\Gamma \simeq 1.487$, in the ELEVATED band. Give the example a constructed baseline with mean capacity 1.25 and standard deviation 0.10, mean flux 3 and standard deviation 1. Then $Z_H \simeq (1.031 - 1.25)/0.10 = -2.19$ and $Z_V = (5 - 3)/1 = 2$. Negative alignment, capacity over two standard deviations below its baseline, and unusually high flux together return SINGULARITY before the band is used. The negative departure and positive capacity describe the same observation. This example traces an alert on invented inputs.

Now close economic exchange completely. Take zero trade, energy, voting similarity and treaty overlap, sanctions at one, all four resilience inputs at 0.05, two constraints and four exits. The engine obtains $K = 0$, security posture $S \simeq 0.936$, mobility $0.05(2/10)(4/8) = 0.005$ and

$\Gamma_d \simeq 0.000534$. The file records hostility $H_d \simeq 17.78$ and stores its negative in the shared field. The disconnected branch returns SINGULARITY on hostile alignment, low tolerance, low mobility and low Γ_d. A reader seeing the same alert word in both examples can now identify two different reasons for it.

Parameter Summary

Parameter	Symbol	Value	Description
Strategic weight	α	4.0	Energy/trade ratio
Sanction decay	β	1.0	Penalty saturation
Alignment steepness	k	2.0	Tanh steepness
Capacity regularizer	ε_H	10^{-6}	Ordinary ratio
Activation/disconnection	ε_E	10^{-8}	Denominator guard
DISCONNECTION gate	K_{disc}	0.05	Mode threshold
Activation scale	γ	1.0	E_a scale
Critical Γ	Γ_c	1.0	Phase transition
Joint alert gap	$\lambda_{1,c}$	0.10	Three-indicator test
Critical H	H_c	0.50	Helicity threshold

Validation Methodology

This historical protocol used 47 conflicts and a CRITICAL-or-higher threshold. The mass-gap and Z-score working paper in Appendix E used its own Level 4 rule, producing a different warning count.

1. **Data Collection**: Bilateral trade (UN Comtrade), UN voting (Erik Voeten), treaties (COW), sanctions (GSDB)
2. **Baseline Computation**: 3-year rolling average for Z-score normalization
3. **Hindcast**: Compute metrics for 24 months preceding known conflict onset
4. **Signal Detection**: Record first date where alert ≥ CRITICAL
5. **Lead Time**: Measure days between signal and conflict onset

This CRITICAL-or-higher protocol and its reported sensitivity, specificity, and lead-time estimates are recorded separately in Appendix E. Reproduction requires its negative-case denominator, alert-episode rule, and frozen input records. Those rates are not interchangeable with the working paper's Level 4 counts.

The full-sample Russia–Ukraine crossing on December 14, 2021 precedes the February 24, 2022 invasion by 72 days; the held-out December 22 crossing gives 64. Both are reconstructed detections, not established forecast issuances (Appendix E).

Appendix C

Selected Historical Conflict Analyses

The cases below preserve a selected historical comparison across periods and mechanisms. Its 18 listed detections and five reported misses give the descriptive proportion 18/23 = 78.3%. That collection includes internal, non-state, and intervention cases, so it is not the eligible-event roster of the separate 47-event mass-gap study in Appendix E.

Historical first-warning roster (18 entries)

Conflict	Year	Dyad	H	Lead	Trap
Suez Crisis	1956	UK/ FR-EG	-0.72	89d	Canal revenue
Six-Day War	1967	IL-EG	-0.81	34d	Straits closure
Yom Kippur	1973	IL-Arab	-0.68	67d	Oil transit
Falklands	1982	UK-AR	-0.59	112d	Fishing, sovereignty
Gulf War I	1990	IQ-KW	-0.77	45d	Oil pricing
Yugoslav Wars	1991	Multi	-0.83	203d	Trade collapse
Eritrea-Ethiopia	1998	ER-ET	-0.64	78d	Currency union
Kargil War	1999	IN-PK	-0.63	56d	Kashmir trade
Kosovo War	1999	NATO-YU	-0.76	289d	Sanctions regime
Iraq War	2003	US-IQ	-0.67	312d	Oil embargo
Lebanon War	2006	IL-LB	-0.70	234d	Border trade
Russo-Georgian	2008	RU-GE	-0.71	134d	Energy transit
Libya	2011	NATO-LY	-0.56	223d	Oil exports
Crimea	2014	RU-UA	-0.62	56d	Gas transit

Conflict	Year	Dyad	H	Lead	Trap
Yemen	2015	SA-YE	-0.58	92d	Oil routes
Nagorno-Karabakh	2020	AM-AZ	-0.69	167d	Pipeline
Russia-Ukraine	2022	RU-UA	-0.95	72d	Gas, trade
Hamas-Israel	2023	HM-IL	-0.54	23d	Labor permits

Table C.1: Selected historical first-warning entries. These are distinct from the mass-gap paper's dated onset records.

The 18 lead times total 2,286 days: their arithmetic mean is 2,286/18 = 127 days, their median 90.5 days, and their range 23–312 days. The mean weights each listed detection equally and assigns no lead time to a miss. Thus 78.3 percent describes 18 detections among 23 cases, while 127 days averages those 18 detections. Both summarize this selected roster; neither uses the separate mass-gap study's 47-event population.

For Nagorno–Karabakh, the dated mass-gap record gives August 15 to September 27, 2020, or 43 days. The 167-day entry above belongs to the separate historical roster and lacks the same fully specified alert–target pair. Likewise, a first warning more than 180 days before onset qualifies for a 180-day evaluation only if a qualifying reading also occurs within that window. These distinctions prevent this descriptive collection from silently becoming a new forecast score.

Reported misses and coverage cases (5 entries)

- Hungarian Revolution (1956): Soviet intervention, no bilateral trade
- Prague Spring (1968): Warsaw Pact internal
- US-Grenada (1983): No economic interdependence
- US-Panama (1989): No bilateral trap
- Tibet (2008): Internal suppression

False Positives (4 cases)

- India-Pakistan 2019: Nuclear deterrence, low curvature
- Greece-Turkey 2020: EU/NATO mediation
- US-Iran 2020: Mutual deterrence
- Ethiopia-Sudan 2021: Limited clashes, no invasion

Appendix D

Glossary and Lexicon

This glossary explains the principal mathematical, political, and research terms used in the book. Chapter references lead to the argument and worked examples; the index locates more specific discussions. Appendix H distinguishes symbols whose meanings change between mathematical settings or model versions.

Mathematical and Geometric Terms

Activation Energy (E_a)
Model scale for reconfiguration, $E_a = \gamma/(H^2 + \varepsilon)$. For fixed positive constants this is largest at zero and decreases as $|H|$ grows. The sign of alignment helps interpret the kind of reconfiguration. (Chs. 4, 22; App. B)

Adjacency Matrix (A)
Square matrix encoding pairwise network connections; entry A_{ij} is the edge weight between nodes i and j, typically bilateral trade volume. (Ch. 7)

AUC (Area Under Curve)
Area under the receiver operating characteristic curve. It measures ranking discrimination over the stated positive and negative population; it does not establish probability calibration or the usefulness of an alert threshold. The historical study's reported AUC belongs to its own sample and protocol. (Chs. 9, 33)

Bifurcation
A change in the number, stability or type of states supported by a dynamical model as a parameter varies. A signed indicator can cross zero smoothly. To identify a bifurcation, follow the states or their stability across that crossing. (Chs. 8, 22)

Catastrophe Theory
Mathematical study of sudden qualitative changes in smooth systems (Thom, 1972). The framework uses the *cusp catastrophe* model with helicity and spectral gap as control variables. (Ch. 22)

Cheeger Constant (h)
For an undirected graph with nonnegative weights, minimize the total weight crossing a nontrivial cut divided by the smaller side's total weighted degree. Both sides must have positive volume. (Chs. 7, 22, 23)

Cheeger Inequality
For the degree-normalized Laplacian of such a graph, $h^2/2 \le \mu_1 \le 2h$. The second-smallest normalized eigenvalue bounds the narrowest crossing. Multiplying all trade weights by the same positive factor leaves these quantities unchanged. (Chs. 7, 22, 23)

Chern–Simons Invariant
The Chern–Simons functional integrates $\mathrm{Tr}(A \wedge dA + \frac{2}{3} A \wedge A \wedge A)$ on an oriented three-dimensional domain, with a specified normalization. Its value depends on the connection. The gauge-invariant exponentiated action has additional conditions on the level and gauge group. (Ch. 22)

Configuration Space
Space of all possible system states. In glass transition theory, basins in the energy landscape correspond to distinct configurations. (Ch. 24)

Conservation Law
A quantity preserved by a specified evolution. In smooth ideal flow, conservation of total helicity requires suitable fluid and boundary conditions. In a dyad, identifying what is carried forward requires the evolution of its commitments and flows. (Chs. 22, 25)

Curl ($\nabla\times$)
Vector operator measuring field rotation. In fluid dynamics, $\boldsymbol{\omega} = \nabla \times \mathbf{v}$ gives vorticity. The geopolitical analog measures "rotation" of diplomatic orientation. (Chs. 4, 22)

Curvature (K)
Scalar in $[0, 1]$ measuring bilateral economic interdependence. Computed from trade volume, energy dependency, FDI, and supply chain entanglement. High K = deep integration costly to sever. (Chs. 4, 7, 8; App. B)

Cusp Catastrophe
A model with two controls and one continuous state coordinate. In the bowed-lid example the state is deflection, and two distinct stable shapes represent competing outcomes. Helicity and gap are proposed political controls. Applying the model requires a measurable state and a way to distinguish its resting configurations. (Ch. 22)

Degree Matrix (D)
Diagonal matrix where D_{ii} equals the sum of edge weights incident to

node i. Used in constructing the graph Laplacian $L = D - A$. (Chs. 22, 23)

Differential Forms
Objects from exterior calculus generalizing functions and vector fields to arbitrary dimensions. Used to define helicity on the 5-dimensional diplomatic manifold. (Ch. 22)

Dissipation
Viscosity changes the helicity of a flow through a signed integral when boundary terms vanish. The signed total can rise or fall. Diplomatic friction is the proposed analogue in the historical interpretation. (Ch. 22)

Divergence Index (Δ)
The negative ratio of the tolerance rate to the regularized integration rate. Read both rates with their signs. For example, falling trust with rising integration gives a positive ratio, as does rising trust with falling integration away from the regularization scale. (App. B)

Eigenvalue / Eigenvector
Solutions to $L\mathbf{v} = \lambda\mathbf{v}$. Eigenvalues $\{\lambda_i\}$ encode network topology; eigenvectors encode structural modes. (Chs. 7, 22, 23)

Euler Equations
Equations for ideal fluid motion. A smooth incompressible Euler flow of constant density in a periodic domain conserves total helicity. Other domains require attention to what crosses the boundary. (Ch. 22)

Exterior Derivative (d)
Operator generalizing gradient, curl, and divergence to differential forms. Maps k-forms to $(k + 1)$-forms. (Ch. 22)

F1 Score
Harmonic mean of precision and recall, $F_1 = 2PR/(P + R)$ when the denominator is positive. It excludes true negatives and depends on the alert threshold and evaluation population. Reported results must retain their own counts and matching rules. (Chs. 9, 33)

Fiedler Vector
A first nonconstant eigenvector of the named Laplacian on a connected graph. Sorted coordinates give candidate cuts. A change of overall sign swaps the displayed sides. If the eigenvalue is repeated, there can be several equally valid directions in that eigenspace. (Ch. 23)

Gaussian Process
Probabilistic nonparametric model providing uncertainty quantification. Proposed for confidence intervals on helicity forecasts. (Ch. 39)

Heat Kernel ($p_t(x, y)$)
Fundamental solution to the heat equation on a network: $p_t(x, y) =$

$\sum_k e^{-\lambda_k t}\phi_k(x)\phi_k(y)$. Describes perturbation diffusion through trade networks. (Chs. 7, 23)

Helicity (H)
Fluid helicity integrates velocity dotted with vorticity. The historical signed ratio uses torsion and integration. The legacy engine multiplies flux by signed alignment. The ordinary capacity calculation divides nonnegative tolerance by regularized integration. A report must identify which reading its H represents. (Chs. 1–39)

Helmholtz Vortex Theorems
Three theorems (1858): vortex lines move with the fluid, vortex tubes maintain constant strength, vortex lines cannot start or end in the interior. Foundation for helicity conservation. (Ch. 22)

Kelvin Circulation Theorem
In a smooth inviscid barotropic flow with conservative body forces, circulation around a loop carried by the fluid is conserved. A loop painted on the water moves with it. A fixed sensor follows a different measurement. (Ch. 22)

Laplacian Matrix (L)
$L = D - A$. Encodes graph connectivity; eigenvalues reveal mixing times, connectivity, and community structure. (Chs. 7, 22, 23)

Manifold
Topological space locally resembling Euclidean space. The framework models international relationships as a 5-dimensional manifold: (T, E, S, R, X) for trade, energy, voting similarity, treaties, sanctions. (Chs. 7, 22)

Mixing Time
Time to approach equilibrium within a stated distance and tolerance under a specified evolution. For $\dot{x} = -\beta L x$, the slowest nonconstant symmetric diffusion mode has e-folding time $1/(\beta\lambda_1)$. Mixing time can also depend on the starting state, norm, and stationary distribution. (Chs. 23, 30)

Navier–Stokes Equations
Governing equations for viscous flow. Helicity dissipates: $dH/dt = -2\nu \int \boldsymbol{\omega} \cdot (\nabla \times \boldsymbol{\omega})\, dV$. (Ch. 22)

Normalized Laplacian
For an undirected graph with nonnegative weights and positive degrees, $\mathcal{L} = D^{-1/2} L D^{-1/2}$ has eigenvalues in $[0, 2]$. It is invariant to uniform positive rescaling of all weights. Normalization does not automatically make different actor sets or edge definitions comparable. (Chs. 23, 29)

Phase Transition
Abrupt qualitative change at a critical parameter value. The framework

treats conflict onset as a phase transition from diplomacy to war. (Chs. 1, 2, 7, 22, 25)

Polarization Index
Scalar from the Fiedler vector measuring how cleanly a network splits into opposing blocs. High polarization = proto-alliance formation. (Ch. 23)

Positive Semi-Definite
A real symmetric matrix is positive semidefinite if $x^T Lx \geq 0$ for every x. This guarantees nonnegative eigenvalues. For an ordinary graph Laplacian the zero eigenvalue comes specifically from $L\mathbf{1} = 0$, because each row sums to zero. (Ch. 22)

ROC Curve
Receiver Operating Characteristic: true positive vs. false positive rate across thresholds. Evaluates classification performance. (Chs. 9, 22)

Simplicial Complex
Generalization of graphs to higher-dimensional interactions (triangles, tetrahedra). Proposed for multi-party relationships. (Ch. 39)

Simplicial Helicity (H_{ABC})
Extension of bilateral helicity to three-party interactions. Captures dynamics invisible to pairwise analysis. (Ch. 39)

Spectral Decomposition
Representation of a matrix through eigenvalues and eigenvectors. The Laplacian's decomposition reveals fundamental network modes. (Chs. 22, 23)

Spectral Gap (λ_1)
For a weighted graph, the second-smallest eigenvalue of the named Laplacian, counting multiplicities. For a transition matrix the book also uses one minus the largest absolute nonstationary eigenvalue. The historical network cutoff of 0.05 belongs to its own normalization. Appendix B identifies the gap actually passed to the engine. (Chs. 7, 8, 12, 22, 23)

Topological Invariant
A property preserved by the allowed transformations. Linked closed loops keep their linking number under a smooth deformation that keeps them disjoint. Cutting a loop changes the allowed operation. A political application likewise needs to identify which changes preserve its proposed quantity. (Chs. 22, 25)

Torsion (τ)
Misalignment between economic flows and diplomatic orientation. Components: TradeAsymmetry, InvestmentDirection, DependencyRatio, AlignmentDelta. (Chs. 4, 22; App. B)

Trichotomy Parameter (Γ)
Historical ratio $m\tau/(K \log |S_{\text{exit}}|)$ on its ideal domain. Its count dependence measures an increasing denominator burden, not a monotone benefit from adding usable exits; Section 22 distinguishes candidate-space resolution from settlement feasibility. The historical five bands appear in Chapter 22. In the current ordinary alert rule an exact boundary value of 0.1, 0.5, 1 or 1.5 enters the next higher numerical band, once the priority tests have been considered. (Chs. 4, 8, 22; App. B)

Velocity Field (v)
Flow field whose curl gives vorticity. Geopolitical analog: vector field of economic flows (trade, investment, migration) across the network. (Ch. 22)

Viscosity (ν)
Resistance to flow. "Diplomatic viscosity" = information asymmetry, bureaucratic friction, institutional drag causing helicity dissipation. (Ch. 22)

Vorticity (ω)
$\boldsymbol{\omega} = \nabla \times \mathbf{v}$. Local rotation measure. Geopolitical analog: "rotation" of diplomatic alignment across the network. (Chs. 4, 22)

Wedge Product ($\wedge$)
The graded product of differential forms. Two one-forms reverse sign on exchange. A p-form exchanged with a q-form picks up the factor $(-1)^{pq}$. In the helicity integral a one-form wedged with its exterior derivative has degree three. (Ch. 22)

Z-Score
$Z_H = (H - \mu_H)/\sigma_H$. TRAP detection: $Z_H < -2.0$ with $Z_V > +0.5$ = collapsed trust with persistent flux. (Chs. 4, 8; App. B)

Physics and Dynamical Systems

Adam–Gibbs Theory
Relaxation time vs. configurational entropy: $\tau = \tau_0 \exp(C/TS_c)$. Low configurational entropy (few diplomatic arrangements) $\Rightarrow$ long response times. (Ch. 24)

Arrhenius Behavior
$\tau \sim \exp(E_a/k_B T)$. Activated relaxation with a fixed barrier in the stated model. Its use for a geopolitical episode is a proposed analogy requiring a defined barrier, timescale, and empirical fit. (Ch. 24)

Attractor
State toward which a dynamical system evolves. Stable relation-

ships have deep attractors; helicity collapse produces shallow or disappearing attractors. (Ch. 8)

Critical Slowing Down
Lengthening recovery near certain dynamical transitions when a relevant restoring rate approaches zero. The inverse-rate calculation applies to a specified local model. Observing a slow response alone does not identify its cause or establish a countdown to conflict. (Chs. 24, 30)

Energy Landscape
High-dimensional potential energy surface vs. system configuration. Glass transitions trap systems in local minima separated by high barriers. (Ch. 24)

Glass Transition
System becomes "frozen" in a disordered state, unable to equilibrate normally. Geopolitical analog: dyad locked in mutual hostility without diplomatic channel capacity. (Chs. 7, 24)

Inviscid Flow
Fluid motion without viscosity. Helicity conservation additionally requires suitable dynamics and boundary conditions. It is a physical result, not an automatically inherited law of diplomatic relationships. (Chs. 4, 22)

Mode-Coupling Theory (MCT)
Power-law divergence near glass transition: $\tau_\alpha \sim (T - T_c)^{-\gamma}$. Geopolitical: $\tau_{\text{dip}} \sim (\lambda_1 - \lambda_1^*)^{-\gamma}$. (Ch. 24)

Turbulence
Irregular fluid motion across interacting scales. A small graph gap and negative political alignment do not establish fluid turbulence. In the geopolitical discussion, the term motivates investigation of feedback and sensitivity. (Chs. 4, 22)

Vogel–Fulcher–Tammann (VFT) Equation
$\tau = \tau_0 \exp(B/(T - T_0))$. An empirical relaxation law used for some fragile glass-forming liquids. A geopolitical application needs its own definitions and evidence; a resemblance to abrupt escalation does not establish this fitted law. (Ch. 24)

Davis Framework Concepts

Alert Levels
The historical register uses BASELINE (L1), DECOUPLING (L2), DIVERGENCE (L3) and SINGULARITY (L4). Appendix B specifies the current engine's six named alert outputs and the conditions for each. (Chs. 4, 8, 9)

BASELINE (Alert Level 1)
Historical routine-monitoring category: $H > 0.3$, $\Gamma > 1.1$. The label expresses the framework's favorable reading; these two inequalities alone do not prove a stable equilibrium. (Ch. 4)

Combined Signal
Joint helicity-spectral reading yielding five states: Liberal Peace, Frozen Peace, Window of Opportunity, Singularity Warning, Void. (Ch. 7)

CRITICAL
Historical trichotomy band $0.1 < \Gamma \leq 0.5$, interpreted as a priority for examining constraints and remaining options. The alert label does not establish a dynamical bifurcation. Appendix B gives the engine's boundary and priority rules. (Ch. 4; App. B)

Curvature Components
Sub-metrics of K: TradePct, FDIPct, JVIndex, HumanCapIndex, CommDensity. (Ch. 4)

Davis Law
The general capacity bound is $C \leq \tau/K$ for positive effective curvature and the stated task tolerance. The Psychohistory specification uses τ/K as the capacity scale. The related squared-distance expression $c^2 = a^2 + b^2 + \Delta$ defines a residual relative to a right-triangle reference. Interpreting that residual requires the distances and the reference geometry. (Ch. 4)

DECOUPLING (Alert Level 2)
$H < 0.3$ OR Γ declining. Spectral gap still healthy. Enhanced monitoring and contingency planning. (Chs. 4, 14, 15)

DISCONNECTION Mode
Activated when $K < 0.05$. Standard $H = \tau/K$ is numerically unstable; Hostility Index replaces helicity. (App. B)

DIVERGENCE (Alert Level 3)
$H < 0$ AND $\Gamma < 1.1$. Negative helicity, spectral gap approaching critical. Active intervention needed. (Chs. 4, 12, 15)

Dyad
Pair of countries forming a bilateral relationship—the fundamental unit of analysis. Examples: RU-UA, CN-TW, US-CN, IN-PK, IL-IR. (Chs. 1–39)

ELEVATED
Historical trichotomy band $1.0 < \Gamma \leq 1.5$, interpreted as an elevated need to examine diplomatic capacity and constraints. The book uses the Cold War US–USSR relationship to motivate this reading; a dated assignment requires the underlying inputs. (Ch. 4)

Feedback-Coupled Flow Systems
Systems where observing flows alters them. Insight motivating the framework's development from transformer attention mechanisms. (Ch. 4)

Four Quadrants
Dyad classification: interdependence × alignment. I: Liberal Peace. II: The Trap. III: Managed Decline. IV: Disconnection. (Ch. 4)

Frozen Peace
Positive signed helicity with a subcritical network gap in the combined-signal table. The parties remain aligned, but a disruption can be slow to clear. Negative helicity with an open gap is the window of opportunity. (Ch. 7)

Frozen vs. Glassy Systems
Frozen: stable but unpleasant equilibrium (Cold War). *Glassy*: trapped, disordered, unable to rearrange (IN-PK). Glassy systems appear stable but can shatter. (Chs. 7, 14)

Helicity Correction
The framework's term for changing a strained relationship through political accommodation, reduced dangerous dependence, or more usable alternatives. The mechanism and the definition of H must be specified; a change in one component does not guarantee that every risk decreases. (Chs. 12, 36)

Hostility Index (H_d)
DISCONNECTION mode metric: $H_d = S \cdot (1 - \tau)/(\tau + \varepsilon)$. Replaces helicity when $K < 0.05$. (App. B)

Liberal Peace (Quadrant I)
High interdependence + diplomatic trust. US-Canada ($H \approx +0.85$), France-Germany ($H \approx +0.90$), EU Core ($H > +0.80$). (Chs. 4, 7)

Master Equation
The capacity reading $H = \tau/K$ on its positive-integration domain, with regularization in the implemented ordinary branch. It differs from signed fluid and historical alignment readings also denoted by H. The notation guide separates these uses. (Ch. 4; App. B)

Negative Helicity
For the alignment-bearing coordinate, $H < 0$ means adverse alignment accompanies economic flows. It is a condition in the trap interpretation, not a sufficient predictor of conflict. The ordered four-component reading has a different sign convention (Section 4). (Chs. 1–39)

PSYCHOHISTORY
The conflict prediction research system implementing the Davis Framework. US Provisional Patent Application No. 63/972,955. (Ch. 39)

Security Posture (S)
$S = 0.40X + 0.35A_h + 0.25D_h$. Sanctions intensity + alignment hostility + historical distrust. Used in DISCONNECTION mode. (App. B)

SINGULARITY (Alert Level 4)
In the historical combined rule, signed $H < -0.5$ and a subcritical network gap mark a trap warning. In the ordinary engine, one route to SINGULARITY requires negative alignment, $Z_H < -2$ and $Z_V > 1$. The 0.85 precision in Appendix E is the observed 34/40 in that tally. Applying it elsewhere requires a matching population and warning rule. (Chs. 4, 8, 9)

Singularity Warning
Combined signal: strongly negative helicity + sub-critical spectral gap. Most dangerous diagnostic state. (Ch. 7)

Structural vs. Eventive Forecasting
Eventive: which specified event is likely within a stated horizon? *Structural*: which constraints are making peaceful adjustment harder? The framework links the structural account to a separately scored warning rule. (Ch. 2)

Tolerance Budget (τ)
$\tau = \beta_1 A_s + \beta_2 I_t + \beta_3 H_t + \beta_4 C_c$. Alliance strength, institutional ties, historical trust, communication channels. Range $[0, 1]$. (App. B)

Torsion Components
Sub-metrics of τ: TradeAsymmetry, InvestmentDirection, DependencyRatio, AlignmentDelta. (Ch. 4)

Trap (Quadrant II)
Substantial economic dependence persists as diplomatic alignment deteriorates. A negative signed reading and a depleted nonnegative tolerance budget can describe different aspects of that same relationship. (Chs. 1–39)

Void
Economic links are too thin to carry much pressure through trade. Security posture and diplomatic mobility can still signal danger. At low integration the operational engine uses these measures on its separate branch. (Ch. 7)

WARNING
Historical trichotomy band $0.5 < \Gamma \leq 1.0$, interpreted as a call to examine emerging constraints and diplomatic options. Its meaning depends on the count convention and observations, not the numerical band alone. (Ch. 4)

Window of Opportunity
Weakly negative helicity + supra-critical spectral gap. Deteriorating but correctable. Policy intervention most effective. (Ch. 7)

Geopolitical and Strategic Terms

Abraham Accords
2020 Israel-UAE/Bahrain/Sudan/Morocco normalization. Altered regional helicity by realigning Arab-Israeli diplomatic orientation. (Ch. 13)

Anocracy
Regime type between democracy and autocracy. Disproportionately prone to conflict due to institutional instability. (Ch. 1)

Axis of Resistance
Iran-led network: Hezbollah, Hamas, Houthis, Iraqi/Syrian militias. A proxy architecture, not a conventional alliance. (Ch. 13)

Capitalist Peace
Gartzke (2007): integrated capital markets, not trade per se, prevent war. (Ch. 2)

Chokepoint Effect
Farrell & Newman (2019): controlling a critical network node enables denial of access. Examples: US/SWIFT, China/rare earths. (Ch. 2)

Commercial / Liberal Peace
Economic interdependence reduces conflict probability by raising war costs. The framework's primary theoretical interlocutor and partial refutation. (Chs. 2, 7)

Controlled Release Paradigm
When $H < 0$, pursue orderly tie reduction rather than maintaining coupling (deepens trap) or abrupt severing (cascading damage). (Ch. 36)

Copeland's Trade Expectations Theory
Expectations of future trade, not current levels, determine interdependence-peace link (1996). Declining expectations may motivate preventive conflict. (Ch. 2)

Coercive Diplomacy
Threats, sanctions, or limited force to compel behavior change without full-scale war. (Ch. 1)

Davidson Window
2028–2032: projected maximum risk for Chinese military action against Taiwan (ADM Philip Davidson, INDOPACOM). (Ch. 12)

Democratic Peace Theory
Democracies rarely fight other democracies (Russett & Oneal). Empirically robust but silent on specific mechanisms and timing. (Ch. 2)

De-risking vs. Decoupling
De-risking seeks to reduce specified vulnerabilities; decoupling describes a broader separation of relationships. Their consequences

depend on substitutes, transition timing, political responses, and the functions lost or preserved. A signed index alone does not choose the appropriate policy. (Chs. 12, 36)

Deterrence
Preventing conflict via unacceptable cost threats. Nuclear deterrence provides tolerance budget partially compensating for negative H. (Chs. 1, 12, 14)

Escalation Ladder
Graduated conflict intensity: diplomatic protests → sanctions → limited force → full-scale war. The framework detects when upward movement becomes structurally likely. (Ch. 13)

Frozen Conflict
Hostilities ceased, no agreement, issues unresolved. Examples: Korea, Kashmir, Cyprus. Distinguished from "glassy" systems. (Chs. 8, 14)

Managed Competition
Maintaining bilateral relations despite disagreements via guardrails and communication. Describes US-China as of 2024. (Ch. 15)

Near Abroad
Russian concept: former Soviet states as sphere of privileged interest. (Ch. 8)

Nuclear Overhang / Floor
The proposed restraint created by the possibility of nuclear escalation. Deterrence can constrain choices while leaving lower-level confrontation and miscalculation possible. The framework's tolerance-floor interpretation is a modeling hypothesis, not an unconditional guarantee of stability. (Ch. 14)

One China Policy / 1992 Consensus
Three distinct political formulations must be separated. The United States' one-China policy, the PRC's one China principle, and the disputed 1992 Consensus have different authors, meanings, and roles. They are not interchangeable names for a mutually accepted settlement. (Ch. 12)

Orderly Decoupling
A managed reduction of specified dependencies while replacements and communication remain usable through the transition. It is a proposed policy path whose feasibility and consequences must be examined. Removing an edge alone cannot increase a fixed combinatorial Laplacian's gap. (Chs. 12, 31)

Panopticon Effect
Controlling a critical network node confers surveillance. Example: US access to SWIFT data. (Ch. 2)

Proxy Architecture
Non-state network for indirect power projection. Iran's is the canonical example. (Ch. 13)

Proxy Degradation
Systematic attrition of proxy networks. Israel's 2023–2025 degradation of Hamas, Hezbollah, and Iranian proxies. (Ch. 13)

Revanchism
Ideology seeking lost territory recovery. Driver of negative alignment in torsion computation. (Ch. 1)

Sanctions Paradox
Sanctions can *increase* conflict risk: reducing alignment without proportionally reducing interdependence pushes H more negative. (Ch. 36)

Silicon Shield
The argument that Taiwan's semiconductor importance creates incentives for other actors to preserve access and avoid disruption. Its effectiveness depends on political choices, functioning production, and substitutes. The chapter distinguishes dated capacity statistics from company revenue shares and customer-specific exposure. (Ch. 12)

"Small Yard, High Fence"
US tech policy: targeted restrictions with strict enforcement. Applied to semiconductor export controls on China. (Chs. 7, 15)

Stability–Instability Paradox
Snyder (1965): nuclear weapons stabilize strategic relations while enabling conventional conflict. Central to IN-PK analysis. (Ch. 14)

Strategic Ambiguity
Deliberate uncertainty about contingency response. US historically maintained this regarding Taiwan defense. (Ch. 12)

Trap Warfare
Conflict from Quadrant II where interdependence itself becomes a weapon. Distinct from territorial, resource, or ideological wars. (Ch. 9)

Wandel durch Handel
"Change through trade" (German). Doctrine that economic engagement can promote political change and restraint. The Russia–Germany case examines where those expectations failed. (Chs. 1, 2)

Weaponized Interdependence
Farrell & Newman (2019): exploiting critical network positions for coercion. Formalized as the trap configuration. (Chs. 2, 15)

Zeitenwende
"Turning point" (German). Scholz's characterization of post-February 24, 2022 German foreign policy shift. (Chs. 1, 2)

Economic and Financial Terms

Basis Points (bps)
1bp = 0.01%. Used for CDS spreads and risk premiums. (Ch. 1)

Bilateral Trade
Total value of goods/services exchanged between two countries. Primary input to curvature (K). Sources: UN Comtrade, IMF DOTS. (Chs. 4–15)

Capital Flight
Rapid capital outflow presaging crisis. Observable as sudden curvature changes. (Ch. 2)

CIF/FOB Valuation
Import pricing conventions. CIF includes shipping; FOB does not. Creates systematic asymmetries in UN Comtrade requiring correction. (App. A)

Credit Default Swap (CDS)
Derivative paying out on default. Sovereign CDS spreads measure market-implied default probability. Typically lags structural indicators. (Ch. 1)

Energy Dependency
Reliance on another country for energy (USD billions). Input to both flux (V) and curvature (K). Germany's 55% gas dependency on Russia = canonical. (Chs. 1, 7, 8)

FDI (Foreign Direct Investment)
Cross-border investment in productive assets. Component of curvature (K). (Chs. 4, 15)

GDP (Gross Domestic Product)
Total economic output. Normalizing variable for trade intensity ratios. (Ch. 4)

Put Skew
Difference in implied volatility: OTM puts vs. ATM options. Elevated = markets pricing downside risk. (Ch. 1)

Sanctions
Economic penalties. Input to alignment (A) via $\sigma(X) = 1 - e^{-\beta X}$. (Chs. 2, 8, 36; App. B)

Trade Asymmetry
$(X_{A\to B} - X_{B\to A})/(X_{A\to B} + X_{B\to A})$. Torsion component; asymmetric trade creates leverage. (Ch. 4)

Transit Fees
Revenue for allowing exports through one's territory. Ukraine earned \$1.2B/yr in Russian gas transit—persistent K despite diplomatic collapse. (Ch. 8)

Acronyms and Abbreviations

ACLED Armed Conflict Location & Event Data Project. Real-time event-level conflict data. (Chs. 1, 8; App. A)

ADIZ Air Defense Identification Zone. A zone used for identification and monitoring, distinct from sovereign airspace. CSIS records 3,075 Chinese military-aircraft flights into Taiwan's self-declared ADIZ in 2024. The count is one category of activity, not a count of attempted invasions. (Ch. 12)

AIS Automatic Identification System. Ship tracking proposed for real-time trade flux without 6–18 month lag. (Chs. 7, 39)

ASML Sole EUV lithography manufacturer (Netherlands). Critical semiconductor supply chokepoint. (Ch. 12)

ATOP Alliance Treaty Obligations and Provisions dataset (Leeds et al.). Input to treaty ratio. (Ch. 9; App. A)

CFIUS Committee on Foreign Investment in the United States. FDI gatekeeping on curvature. (Ch. 15)

CHIPS Act Creating Helpful Incentives to Produce Semiconductors ($52B, US 2022). Deliberate curvature reduction policy. (Ch. 12)

CIPS Cross-Border Interbank Payment System (China). RMB SWIFT alternative reducing Western leverage. (Ch. 2)

COW Correlates of War project. Interstate wars, alliances, diplomatic exchanges. (Chs. 1, 4, 9; App. A)

DOTS Direction of Trade Statistics (IMF). Bilateral trade by country pair. Primary source for V and K. (Chs. 4, 7, 8; App. A)

EUV Extreme ultraviolet lithography. ASML's 2024 report identifies it as the only manufacturer of EUV systems. EUV and deep-ultraviolet multiple patterning are distinct manufacturing approaches; a universal claim that every chip at or below a named node requires EUV is too broad. (Ch. 12)

EWP Early Warning Project (US Holocaust Memorial Museum / Dartmouth). Annual atrocity risk assessments. (Chs. 1, 8)

GERD Grand Ethiopian Renaissance Dam. Ethiopia-Egypt-Sudan conflict trigger; proposed "water helicity" test case. (Ch. 39)

GSDB Global Sanctions Database (Felbermayr et al.). Sanctions episodes; input to X_{sanction}. (App. A)

IEA International Energy Agency. Bilateral energy dependency data. (App. A)

IMF International Monetary Fund. DOTS bilateral trade data provider. (Ch. 4; App. A)

IRGC Islamic Revolutionary Guard Corps. Iran's parallel military controlling proxy architecture. (Ch. 13)

JCPOA Joint Comprehensive Plan of Action ("Iran deal," 2015). Helicity correction attempt: sanctions relief for nuclear constraints. (Ch. 36)

MID Militarized Interstate Disputes (COW). Threats, displays, and uses of force. (Ch. 4; App. A)

NATO North Atlantic Treaty Organization. Institutional tolerance budget (τ) and spectral gap stabilization. (Chs. 1, 7, 8, 12, 15)

NCA National Command Authority (Pakistan). Nuclear arsenal control; provides tolerance floor in IN-PK. (Ch. 14)

PITF Political Instability Task Force. US-funded state failure model. Not designed for bilateral war prediction. (Chs. 1, 8)

SAARC South Asian Association for Regional Cooperation. Effective suspension reflects $\lambda_1 \approx 0$ in IN-PK subsystem. (Ch. 14)

SPFS System for Transfer of Financial Messages (Russia). Domestic SWIFT alternative, post-sanctions. (Ch. 2)

SWIFT Society for Worldwide Interbank Financial Telecommunication. Global payment messaging; disconnection = high-impact sanction. (Chs. 2, 13, 39)

TAPI Turkmenistan-Afghanistan-Pakistan-India Pipeline. Unrealized energy project; would have increased IN-PK curvature. (Ch. 14)

TIES Threat and Imposition of Economic Sanctions dataset. Episode-level sanctions data. (Ch. 9; App. A)

TSMC Taiwan Semiconductor Manufacturing Company. A major semiconductor foundry. The chapter distinguishes its own production and revenue measures from Taiwan's geographical share of a defined capacity class, and from a buyer's qualified supply. (Ch. 12)

UCDP Uppsala Conflict Data Programme. Primary academic armed conflict dataset. (Chs. 4, 9; App. A)

UNCTAD UN Conference on Trade and Development. Investment and trade facilitation data. (Ch. 4; App. A)

UNGA UN General Assembly. Voting records input to S_{vote} (Voeten, Bailey & Strezhnev). (App. A)

ViEWS Violence & Impacts Early-Warning System (Uppsala). ML-based subnational conflict prediction. (Chs. 1, 2, 8, 36)

WDI World Development Indicators (World Bank). GDP and normalizing variables. (App. A)

Key Historical Events

72-Day Warning Window
The interval from December 14, 2021 to February 24, 2022 used in the book's retrospective Russia–Ukraine threshold analysis. The date separation is an arithmetic fact; evaluating a contemporaneously issued forecast additionally requires its issuance and information records. (Ch. 8; App. E)

Budapest Memorandum (1994)
Security assurances to Ukraine for nuclear arsenal surrender. Violation undermined treaty-based tolerance globally. (Ch. 8)

Cuban Missile Crisis (1962)
L4 configuration resolved without war. Demonstrates alerts are probabilistic, not deterministic. (Ch. 4)

Euromaidan (2014)
Revolution precipitating Crimea annexation and RU-UA helicity collapse. (Chs. 1, 8)

Minsk Agreements (2015)
Donbas ceasefire agreements. Progressive failure collapsed τ in RU-UA. (Ch. 8)

Nord Stream 1 / 2
Russia-Germany Baltic gas pipelines. NS2 completion despite deteriorating relations exemplified the curvature trap. Damaged September 2022. (Chs. 1, 2)

Operation Rising Lion (June 2025)
Israeli strikes on Iranian nuclear/IRGC facilities. Direct interstate conflict escalation. (Ch. 13)

Operation True Promise (April 2024)
Iran's direct missile/drone attack on Israel. First direct exchange; proxy-to-direct phase transition. (Ch. 13)

Pulwama–Balakot Crisis (2019)
IN-PK crisis cycle: terrorist attack → airstrikes → engagement → de-escalation. L4 registered but nuclear deterrence held. (Ch. 14)

Expanded Methods and Comparative Evaluation

Admissible uncertainty set
The values or models consistent with explicitly stated assumptions about what is unknown. A range of admissible values is not automatically a statistical confidence interval. (Ch. 29)

As-of snapshot
The observations and versions available by a specified issuance time,

with their release dates retained. It differs from the best later reconstruction of the same historical period. (Ch. 34)

Brier loss
Mean squared error of a binary probability forecast, $N^{-1} \sum_i (p_i - Y_i)^2$. Lower is better; expected loss is minimized by the true conditional probability. (Ch. 33)

Calibration
Agreement between reported probabilities and event frequencies over appropriately comparable forecasts. Different from ranking and from an alert decision. (Ch. 33)

Certificate
An inspectable object verifying a specific mathematical result, such as a feasible maximum flow paired with a cut of equal capacity. It does not verify the empirical completeness of the inputs. (Ch. 31)

Competing risks
Different possible first exits from a defined risk state. The probability of one exit depends on the rates of the alternatives as well as its own rate. (Ch. 33)

Effective resistance, R_{ij}
For a connected undirected Laplacian, $(e_i - e_j)^T L^\dagger (e_i - e_j)$. An endpoint response measure; physically electrical resistance only when the weights have the corresponding conductance interpretation. (Ch. 29)

Exchangeability and positivity
Assumptions used for causal comparison: sufficiently comparable potential outcomes conditional on measured pre-intervention information, and overlap in the possibility of receiving each compared action. (Ch. 37)

Interference
Dependence of one unit's outcome on actions assigned to other units. Requires an explicit account of spillover rather than an automatic independent-unit interpretation. (Ch. 37)

Nonanticipating decision
A decision using only information available when it is made: $u_k = \pi_k(\mathcal{I}_k)$. (Ch. 31)

Nonnormal operator
An operator that does not commute with its adjoint. Stable eigenvalues need not prevent temporary growth in a specified norm. (Ch. 30)

Potential outcome
$Y(a)$, the outcome under a specified intervention a. Only the outcome corresponding to the realized intervention is directly observed for one case. (Ch. 37)

Pseudoinverse, $L^{\dagger}$
For a connected undirected Laplacian, the inverse on the zero-mean subspace and zero on the constant vector. (Ch. 29)

Reporting process
The process through which underlying events become recorded observations. Changes in coverage can change measured trends without matching changes in underlying events. (Ch. 20)

Time-expanded network
A graph with location–time copies, transport edges, and storage edges, used to enforce timing as well as capacity constraints. (Ch. 31)

Witness
A concrete example that makes a specified claim inspectable or exposes its failure. A witness's conclusion is limited to the proposition it tests. (Ch. 32)

Reading the evidence, the model, and the decision

Ablation
A comparison that removes or changes one specified model component while preserving the conditions needed to interpret the difference. Removing a feature does not remove its information if an equivalent copy remains elsewhere. Refit permitted variants on training data and compare them on common held-out cases. (Ch. 35)

Abstention
A model's explicit refusal to issue an answer under its specification. Missing inputs or unresolved identities can justify abstention, but the omitted forecast opportunities remain part of evaluating coverage. A system should state the reason and the condition needed to resume calculation. (Chs. 33, 34)

Actor identity
The rule defining who or what a node represents over time. Names, legal states, armed commands, and negotiating authorities are different possible objects. Splits, mergers, aliases, and changes in effective authority need dated mappings so that historical evidence is neither duplicated nor silently reassigned. (Ch. 18)

Aggregation
Combining finer observations into a larger unit, such as districts into a national total. The total can conceal different distributions and dynamics. Exact closure of a linear aggregate $y = Px$ requires $PG = \bar{G}P$ for the proposed local and aggregate operators. (Ch. 20)

Analytic completion
Constraining missing function values by requiring all supplied observations to belong to one declared analytic family. In GIGI, boundedness and complex differentiability link the observations. The construction does not declare the original physical process holomorphic. (Ch. 28)

Availability interval
The times during which a route, permission, or service is usable under a stated definition. Several necessary conditions must overlap for a task to be feasible. Their first opening dates alone do not establish a usable interval for completion or return. (Ch. 19)

Base rate
Frequency or probability of the defined event in the relevant eligible population. Rare events can produce many false alerts even when sensitivity appears high. A sample selected around famous conflicts does not automatically estimate the base rate faced by a continuously operating system. (Ch. 33)

Boundary condition
A constraint imposed at a domain's edge or at the start or end of an evolution. Boundary conditions determine which mathematical identities and solutions apply. A formula derived for one boundary problem cannot be transferred to another merely because the interior equation looks the same. (Chs. 22, 31)

Branch precedence
The order in which an implementation evaluates conditions that can determine its output. When several conditions hold, precedence identifies which one controls the label. The explanation should report that condition rather than infer a reason from the final label alone. (Ch. 34; App. B)

Candidate space
Distinguishable descriptions considered by an inference problem, under a declared resolution and equivalence rule. In Γ_{res}, its size increases the denominator burden. It is distinct from the set of courses that participants can actually implement. (Sec. 22)

Capacity
The maximum usable rate or extent of a specified operation under stated conditions. Freight capacity, a mathematical capacity bound, and production capacity share a word but need not share units or a formula. Nominal capacity may differ from capacity available before a deadline. (Ch. 31)

Case-control sampling
Selecting observations partly according to whether the outcome occurred. Such a sample can support particular comparisons while changing the event prevalence. Recovering population probabilities requires assumptions about selection and the operating population, not simply fitting the selected labels. (Ch. 33)

Censoring
Incomplete observation of an outcome time. A forecast whose horizon has not elapsed is not a completed no-event case. A transaction still pending at the cutoff has a known minimum wait but an unknown final duration. (Chs. 21, 33)

Clock, information
The chronology of when evidence was available for a decision. It differs from the time the evidence describes and from the time it was acquired by a particular system. Forecast evaluation must specify which information clock defines eligibility. (Ch. 34)

Coercion
Pressure intended to alter another actor's conduct by changing expected costs or available options. Its forms and outcomes must be distinguished. Shipping pressure, a quarantine, a blockade, and an invasion should not be scored as the same event without an explicit target definition. (Ch. 12)

Common-cause failure
Loss of several apparently separate options through one shared prerequisite, such as a receiving process or authorization. Marginal availability of each option does not establish independent redundancy. The common dependency belongs in the model. (Ch. 31)

Commutator
For matrices A, B, the difference $AB - BA$. A nonzero commutator shows that applying the operations in opposite orders can produce different results. In a temporal network, one operation changes the state on which the next acts. (Chs. 27, 30)

Confidence interval
An interval generated by a statistical procedure with a stated repeated-sampling coverage interpretation under its assumptions. It differs from a range of logically compatible states and from a scenario band. The label should identify which uncertainties the calculation actually includes. (Chs. 29, 33)

Conservation constraint
A balance requirement, such as inflow equaling outflow at a depot with no production, consumption, or storage change. It prevents an

optimization from creating goods or payments during transit. Real storage, loss, and conversion require explicit terms. (Ch. 31)

Convolution
An accumulation of past inputs weighted by the system's response over the time since each input arrived. In linear dynamics, the integral of $e^{G(t-s)}Bu(s)$ carries those contributions to the present. Equal total inputs can have different consequences when their timing differs. (Ch. 30)

Coverage
The fraction and composition of eligible cases for which a model supplies an answer, or the portion of a stated requirement actually served. These are different uses. A high forecast score on a selective subset and a high delivery total both need their denominators and exclusions examined. (Chs. 18, 33)

Critical path
A sequence of prerequisite tasks that determines the earliest completion of a specified operation. Shortening a task outside the currently limiting path may not advance completion. A realistic schedule must preserve genuine dependencies rather than assume every task can run in parallel. (Ch. 21)

Cycle space
Edge patterns whose net incidence at every node is zero. They describe balanced closed circulations under the chosen orientation. On a connected graph with m edges and n vertices, its dimension is $m - n + 1$. This is a graph property, not a count of independent geopolitical mechanisms. (Ch. 26)

Decision margin
Distance between a computed quantity and the boundary used by a specified decision rule. Numerical or observational uncertainty smaller than the relevant margin may leave the classification unchanged. A precise central estimate without an uncertainty bound does not establish that robustness. (Ch. 29)

Degree normalization
Scaling a graph operator by node degrees to ask a relative rather than absolute exchange question. Uniform edge rescaling cancels, but increasing one edge can lower the normalized gap because it also changes the reference degrees. (Ch. 29)

Denominator
The population, exposure, total, or reference quantity against which a value is expressed. Export shares, event rates, sensitivity, and alert

precision have different denominators. A denominator's change can move a ratio even when its numerator remains constant. (Chs. 12, 33)

Directional reading
A measurement attached to ordered actor roles. Reversing the roles can reverse some components and leave others unchanged. Preserve the pair of readings and their definitions; a positive trade imbalance does not establish mutual diplomatic alignment. (Sec. 4)

Distribution shift
A change between development and application populations, their observations, or the relationship between predictors and outcomes. An unfamiliar input range and a changed conditional outcome relationship are distinct possibilities. Transfer evaluation should investigate which changed. (Ch. 32)

Dual certificate
A combination of constraints proving a bound on an objective. Together with a feasible state attaining that bound, it certifies optimality. The certificate is conditional on the constraints; it does not establish that their empirical inputs are correct. (Ch. 28)

Effective connection
A relation obtained by reducing or summarizing a more detailed model. A Schur-complement edge can preserve a boundary response through an eliminated intermediary. It should not be presented as evidence of a direct historical agreement or shipment. (Ch. 29)

Eligible population
The cases that a forecasting protocol could have addressed under its actor, time, and risk-state rules. It includes relevant quiet opportunities. Selecting only cases that later became famous changes what performance statistics can establish. (Ch. 35)

Event time
When the defined occurrence happened, as distinct from when it was reported or learned. An uncertain event interval should retain its precision. Assigning a convenient exact date can manufacture a lead or lag relative to a forecast or intervention. (Chs. 20, 34)

Feasible set
The states or decisions satisfying every stated constraint. An empty set means the specified requirements cannot all be met together. It does not establish that every conceivable political settlement is impossible; changing an assumption or a genuinely negotiable constraint defines another problem. (Chs. 28, 31)

Gauge transformation
A change in local coordinates or reference units with a specified

transformation law. In the quotation graph, $a' = a + B^T\psi$ leaves closed loop sums unchanged. The invariance follows from the model's structure and cannot be assigned to unrelated diplomatic scores by analogy alone. (Ch. 26)

GIGI

The author's research architecture connecting changing records to bounded analytic inference. The book discusses its representation, compatible answer families, measured-data experiments, and computational limits. Its completion results concern the declared mathematical encoding and do not by themselves establish conflict-forecast accuracy. (Ch. 28; App. F)

Gradient component

In the graph Hodge fit, the part of edge observations explained by differences of node potentials. It can contain genuine economic movement. Calling a component gradient does not classify it as artificial; it identifies a particular representational structure. (Ch. 26)

Hazard

An event rate, or an interval event probability, conditional on remaining in the defined risk state. Its conversion into a cumulative probability depends on the time horizon and competing exits. A hazard is not the same numerical object as probability by a deadline. (Ch. 33)

Hodge decomposition

A splitting of edge data into components determined by incidence, the chosen inner product, and, where relevant, faces of a complex. The book's basic graph fit separates node-potential differences from an orthogonal cycle residual. Interpretation requires the graph and fitting weights. (Ch. 26)

Holomorphic

Complex differentiable throughout the stated open domain. This condition ties together a function's real and imaginary parts and constrains possible completions. A histogram can be encoded in a holomorphic generating polynomial without making the underlying physical process holomorphic. (Ch. 28)

Hysteresis

Dependence of a decision or state on its prior path, such as using different entry and exit thresholds for an alert. An operational hysteresis rule can reduce switching caused by noise while also changing lead times and episode durations. (Ch. 32)

Identification

Whether the available observations and stated assumptions determine

the quantity being asked for. If different underlying states produce the same observations but different answers, the answer is not point-identified. Bounds or additional observations can still be useful. (Chs. 28, 20)

Imputation
Filling a missing value through a declared rule or fitted model. It adds an assumption and may add uncertainty. Imputation fitted on later data can leak information into earlier forecasts. Missing, zero, inapplicable, and disputed require different treatment. (Ch. 34)

Incidence matrix
A matrix recording the endpoints and bookkeeping orientation of graph edges. A column has opposite signs at its two endpoints. Reversing an undirected edge's bookkeeping orientation leaves its Laplacian contribution unchanged; substantive directed flows require a different interpretation. (Chs. 26, 29)

Information leakage
Use, during development or evaluation, of information unavailable under the claimed forecast protocol. It can enter through revised data, normalization, target selection, overlapping horizons, or tuning. Respecting time only in the final fitting step is insufficient. (Chs. 32, 35)

Inverse-probability weighting
Reweighting observations by the inverse of the probability of receiving their observed treatment or selection status. Under appropriate assumptions it can reconstruct a specified comparison population. Poor overlap produces large or undefined weights and limits what can be learned. (Ch. 37)

Issuance record
The preserved output of a forecast at a defined time, with its target, horizon, inputs, model version, and interpretation. A later improved reconstruction is a different artifact. Keeping both permits historical understanding and honest evaluation of what was actually issued. (Ch. 34)

Laplacian, signed
An operator that uses the signs of specified relations as well as their magnitudes. With absolute degrees, its energy penalizes disagreement across positive edges and agreement across negative edges. Low energy can represent opposed camps and does not mean peace. (Ch. 19)

Lead time
Time between a qualifying warning and its matched event under an

explicit matching rule. Repeated daily warnings about one episode do not constitute independent discoveries. Event-date uncertainty and reporting delays affect what the claimed lead establishes. (Ch. 33)

Logarithmic loss
A proper probability score that penalizes confident errors through negative log likelihood. Probability zero assigned to an event that occurs has infinite mathematical loss. Numerical clipping changes that convention and must be recorded. (Ch. 33)

Maximum flow / minimum cut
Paired descriptions of a capacitated network problem. A feasible flow gives an achievable delivery; a separating cut gives an upper bound. Equal values certify optimality for the specified network. The certificate does not establish that the input capacities are empirically complete. (Ch. 31)

Measurement model
The explicit relationship between an underlying quantity and its recorded observation. Reporting coverage, delays, imprecision, and errors can change the observed series independently of the underlying process. Those changes need representation when they affect the claimed inference. (Ch. 20)

Model version
The particular formulas, meanings, input transformations, thresholds, and precedence rules used for a result. A shared project name does not make different versions numerically comparable. Changing a field's interpretation requires a version distinction even if its column name remains unchanged. (Ch. 34)

Negative control
An outcome, exposure, or comparison chosen because a substantive causal account says the effect of interest should be absent. Unexpected associations can expose bias; passing the check does not prove all bias absent. (Ch. 37)

Nevanlinna–Pick matrix
The Hermitian matrix $P_{ij} = (1 - w_i\overline{w_j})/(1 - z_i\overline{z_j})$ for distinct interior interpolation points. Its positive semidefiniteness characterizes compatibility with a bounded holomorphic disk map. Joint compatibility is stronger than checking each value's magnitude separately. (Ch. 28)

Nominal plan
A plan evaluated for one selected set of inputs or scenario. It can be feasible there and fail under other admissible conditions. A robust plan states the uncertainty set it must cover, while an adaptive plan also states what it can learn before acting. (Ch. 31)

Operator norm
Largest amplification of vector length under the specified input and output norms. The Euclidean operator norm gives useful perturbation bounds for symmetric matrices. It is not generally the sum of entries, the largest entry, or a direct measure of political severity. (Ch. 29)

Out-of-sample evaluation
Assessment on observations not used to fit or choose the evaluated procedure. Repeatedly tuning against the same test set turns it into development evidence. Time and episode dependence matter when deciding whether an apparent holdout is genuinely informative. (Ch. 35)

PARALLAX
The author's prototype for constraining a missing contemporaneous quotation and selecting another observation that separates compatible explanations. The book uses its saved synthetic recovery studies to illustrate evidence bounds; these are not future-return or conflict forecasts. (Ch. 28)

Partial identification
Determining a set or range of answers rather than one answer from the available observations and assumptions. The endpoints can have witnesses that make them inspectable. A narrow range remains conditional on the model family used to derive it. (Ch. 28)

Precision / recall
Precision is the fraction of issued positive classifications that are positive under the target rule. Recall is the fraction of positive outcomes detected. Their denominators differ, and both depend on how opportunities and episodes are counted. (Ch. 33)

Prerequisite
A condition that must hold before a specified operation is available. Several proposals can share one prerequisite and therefore fail together. Counting named alternatives without their prerequisite structure can exaggerate the number of independent exits. (Chs. 21, 31)

Propensity score
The conditional probability of receiving a specified treatment given recorded pre-treatment variables. It describes assignment, not the probability of the conflict outcome. Its usefulness depends on overlap and a defensible account of confounding. (Ch. 37)

Provenance
The retained history of a datum's source, transformation, responsible process, and timing. A file hash helps establish integrity but does not establish truth or prior availability. Provenance permits a dis-

agreement to be located in evidence, transformation, or calculation. (Ch. 34)

Quarantine, Taiwan scenario
The law-enforcement-framed coercive scenarios examined in the cited CSIS research, distinct from a military blockade or invasion. These are scenario definitions, not assertions that a particular future event has been decided or that all such actions share one probability. (Ch. 12)

Recourse
An action available after new information arrives. A plan must distinguish choices made before uncertainty is resolved from later adjustments. A calculation that chooses today's allocation using tomorrow's unobserved demand is a perfect-information benchmark, not an implementable policy. (Ch. 31)

Release vintage
The particular version of a dataset available at a stated time. Later revisions may improve historical accuracy while being inappropriate for reconstructing an earlier forecast's information. A citation to a dataset name alone does not identify its vintage. (Ch. 34)

Resolution, forecast
A probability model's ability to distinguish cases with different event frequencies. It differs from calibration: a constant base-rate forecast can be calibrated overall while offering no useful separation of higher- and lower-risk cases. (Ch. 33)

Robust feasibility
Existence of a plan or decision rule that satisfies the stated requirements for every admitted scenario, with the specified information timing. The uncertainty set and quantifier order are part of the claim. Robustness to one error source does not cover unmodeled failures. (Ch. 31)

Rough path / iterated integral
A mathematical representation retaining more of a path's sequential structure than its endpoint. Second-order integrals can distinguish the order of changes whose totals agree. The book connects this to the author's lead–lag experiments while keeping their data and representation explicit. (Ch. 27)

Scenario
A stipulated continuation used to examine consequences of specified assumptions. A scenario is not automatically a probability forecast or a confidence interval. Its value comes from showing what follows if its conditions hold and which changes alter that consequence. (Ch. 12)

Schur complement
The reduced matrix obtained when specified internal coordinates are eliminated from a block quadratic problem. Under the stated conditions it preserves minimized boundary energy. It does not generally preserve all dynamics, internal consequences, or freight capacities. (Ch. 29)

Selection bias
Distortion arising when the observed or evaluated cases differ systematically from the population of interest. Completed deliveries, publicized events, and issued forecasts can each be selective subsets. The selection mechanism determines whether and how a correction is justified. (Chs. 21, 33)

Sequential exchangeability
An assumption that, at each decision time, action assignment is conditionally independent of the relevant potential outcomes given the measured history. It supports specified longitudinal causal calculations; a detailed history does not itself prove the assumption. (Ch. 37)

Shadow price
Local change in an optimized objective when one constraint is relaxed, within the relevant regime of a specified optimization problem. A binding constraint can have value that disappears after another limitation takes over. It is not a universal valuation of the people affected. (Ch. 31)

Spatial precision
The geographical detail supported by a source. A coordinate assigned to a district-level report does not establish an exact event location. Region-based bounds can show which proximity claims remain true across the locations compatible with the evidence. (Ch. 20)

Structural balance
In a signed graph, a consistency property permitting a two-camp arrangement with positive relations within camps and negative relations between them. A balanced graph can contain hostility. Mathematical balance is not a synonym for peaceful or equitable relations. (Ch. 19)

Target trial
An explicit protocol for the experiment that would answer a causal question: eligibility, strategies, assignment, time zero, follow-up, outcome, and analysis. Observational emulation identifies which parts the available records can support. (Ch. 37)

Temporal network
A network whose connections or weights change with time. Aggre-

gating its edges over a period can erase the order needed for a route to be usable. The same two links in opposite sequences can produce different states. (Ch. 30)

Terminal condition
A requirement at the end of a planning horizon. A weak terminal condition can allow a plan that exhausts every buffer just before the modeled window ends. A useful condition states what capacity or reserve must remain for the next period. (Ch. 31)

Transfer test
Evaluation in a different period, region, or actor setting from development under a defined common target. It tests the scope of a representation or fitted relation. Differences in observation practices and identity rules must be examined alongside differences in political behavior. (Ch. 35)

Transient amplification
Temporary growth in a specified state or output despite eventual stability. It can arise through directional coupling and can breach an operational limit before recovery. Stable eigenvalues alone do not bound every finite-time consequence. (Ch. 30)

Usable substitute
An alternative able to perform the required function at the required quality, quantity, location, and time. A supplier's name or an announced investment does not establish that qualification. Shared prerequisites can make apparent substitutes fail together. (Ch. 12)

Verification
Observation sufficient to support a specifically defined proposition under a stated procedure. Verifying an instruction was issued differs from verifying that it was implemented everywhere. Unresolved status should remain distinct from both support and contradiction. (Ch. 19)

Viability
Ability to remain inside a required set of states under specified dynamics, available decisions, disturbances, and information. A viable initial state admits an appropriate policy for the stated horizon. A feasible endpoint alone does not establish a viable path to it. (Ch. 31)

Weighted norm
A measure of size such as $\|x\|_M^2 = x^T M x$ with $M \succ 0$. Its meaning depends on the units and weights. A metric useful for proving stability may differ from the component limits that define an acceptable civilian or institutional outcome. (Chs. 30, 32)

Appendix E

Reading and Reproducing a Forecast

A result becomes easier to understand when its unit, model, and evaluation rule travel with it. This appendix works through the arithmetic of the Psychohistory study and the practical record needed to carry a research result into a continuing warning process.

The Mass Gap and Z Score Evaluation

The Psychohistory working paper describes a panel of 156 directed dyads among 42 states, observed monthly from January 1990 through December 2022. That design contains $156 \times 396 = 61{,}776$ dyad-months. Its event sample contains 47 conflict onsets. An alert counts as a true positive under its stated rule when Level 4 occurs within 180 days before onset (Davis 2025).

The paper's mass-gap plus Z-score evaluation reports the following counts:

Outcome	Count
Conflict preceded by qualifying L4 alert	35
Conflict without qualifying L4 alert	12
L4 alert without conflict within 180 days	14

The conflict count is $35 + 12 = 47$. The counted positive alerts total $35 + 14 = 49$. These denominators answer different questions:

$$\text{Precision} = \frac{35}{49} \simeq 0.7143, \tag{E.1}$$

$$\text{Recall} = \frac{35}{47} \simeq 0.7447, \tag{E.2}$$

$$F_1 = \frac{2 \times 35}{2 \times 35 + 14 + 12} = \frac{70}{96} \simeq 0.7292. \tag{E.3}$$

This is the arithmetic behind the reported $F_1 = 0.73$. The paper's primary results table also reports AUC 0.84, mean lead time 72 days, and median lead time 63 days for that evaluation. These statistics describe different properties: threshold performance, discrimination across thresholds, and warning time. None can be substituted for another.

The number of false alerts divided by the number of all alerts is the false discovery proportion. For these counts it is 14/49. The false positive rate instead divides false positives by all eligible negative observations, requiring a count of true negatives under the same unit and episode rule. Event counts and dyad-month counts cannot be mixed in that denominator.

What the mass gap adds

A warning system can recognize an unusual shipment without knowing whether the relationship can absorb it. The Z-score asks how far a reading has moved from its baseline. The mass-gap construction adds a proposed measure of structural fragility. Their comparison therefore asks a specific empirical question: does that additional measurement improve the warning on the paper's stated task?

The Psychohistory paper reports the following comparison (Davis 2025). These are the paper's research results, with its target and model definitions; they are not the separate cross-system percentages reproduced later in this appendix.

Model	F_1	**AUC**	**Lead (days)**
Mass gap plus Z-score	0.73	0.84	72
Z-score only	0.70	0.81	58
Realist: military balance	0.52	0.67	23
Liberal: trade/GDP	0.48	0.64	31
GDELT Goldstein	0.61	0.72	12
Random baseline	0.08	0.50	—

Table E.1: Results reported in the Psychohistory paper's baseline comparison. Lead time is the paper's Level 4 warning-to-onset measure. The baseline recipes appear in its Appendix A.5. These figures are distinct from the later book's threshold-voting summaries.

The most direct comparison holds the Z-score component in view. Adding the mass gap is associated in this reported experiment with an F_1 increase from 0.70 to 0.73, an AUC increase from 0.81 to 0.84, and

fourteen additional days of mean Level 4 lead time. The values 58 and 72 days consequently have named homes in the comparison: Z-score only and mass gap plus Z-score. They should travel with those classifier names.

Against the paper's GDELT Goldstein baseline, the reported differences are 0.12 in F_1 and sixty days of lead time. That comparison concerns the baseline defined in the paper. It does not establish the same margin over every model using GDELT, or over VIEWS, CAST, PITF, or the Early Warning Project. A data source, a particular decision rule built from it, and a forecasting service are different objects of comparison.

The paper also distinguishes early structural stress from the final Level 4 warning. Its 108-month versus fourteen-month Level 1 comparison is tied to the Russia–Ukraine example. The 108 months span the 2013 stress reading to the 2022 invasion; the paper associates the early stress with the intervening Crimea crisis and describes the long decoupling period. That interval is not the mean Level 4 lead across 47 conflicts. Naming the level, event, and population keeps the long historical argument from being mistaken for a single nine-year advance prediction of the invasion.

A numerical replay of these comparisons needs the same observations, fitting periods, baseline recipes, and alert-matching rules for every row. The reported table gives the research a definite set of claims to reproduce. The common evaluation procedure in Chapter 35 explains how to carry that comparison into a fresh held-out test without changing its question after the outcomes are known.

Full Sample and Held Out Performance

The same working paper describes a temporal split with 31 conflict events in 1990–2014 and 16 in 2015–2022. It reports training $F_1 = 0.75$, held-out $F_1 = 0.69$, and five-fold temporal cross-validation with mean $F_1 = 0.71$ and standard deviation 0.04 across folds (Davis 2025).

A full-sample estimate, a held-out estimate, and a cross-validation average are three views of performance. The held-out result asks how a fitted procedure transfers to a later interval. Cross-validation examines variation across several specified partitions. A full-sample summary describes the larger evaluation population under its stated fitting procedure.

Keep the parameter history with the result. The working paper assigns December 14, 2021 to the Russia–Ukraine L4 crossing in its full-sample reconstruction and December 22 under the temporal holdout specification. Those give 72 and 64 days respectively to February 24, 2022. Neither is a date on which this book establishes that a forecast was

actually issued. The eight-day difference shows that baseline estimation changes the reconstructed timing.

This gives a practical interpretation of robustness. A robust structural diagnosis need not assign an identical crossing date under every defensible specification. Report the range and identify the choices responsible for it. A precise date is most useful when the reader can see the information and parameters that produced it.

The Same War, Different Experiments

The mass-gap paper's Supplementary Table S1 pairs an August 15, 2020 alert with the September 27 Armenia–Azerbaijan onset. The difference is 43 days. The historical roster in Appendix C separately gives 167 days, and the rapid-onset discussion has an L3-only entry. These records cannot be reconciled by choosing whichever label makes a score look better. The dated mass-gap pair supports that study's reported detection. The other entries need their own onset, model, and alert record before they can enter a matched comparison.

The same problem appears when the target changes. The paper pairs a January 8, 2019 India–Pakistan alert with February 26, giving 49 days. A study asking about the Balakot episode can record a hit. A warning for a wider war within 180 days can instead end as a false alert. Those outcomes are consistent only when the target is stated before scoring. A border clash cannot be quietly relabeled as a full war to rescue an alert, and a qualifying attack cannot be dismissed as too small after it has occurred.

The paper supplies ten dated rows in its printed S1 excerpt and refers the remaining 37 to supplementary materials. The aggregate 47-event results are the paper's reported research findings. The excerpt alone is not a complete event-by-event replication package. Replication needs the full register together with the frozen inputs and model versions. That requirement gives the research a concrete path to independent evaluation without discarding the results already reported.

Versioned Research Estimates

The research includes a count-based mass-gap study, a selected historical first-warning roster, and developmental summaries for later indicator combinations. This register retains their numerical records together with the information needed to read them. The primary 47-event result remains the working paper's 35 detections, 12 misses, 14 false alerts,

AUC 0.84, and mean lead time of 72 days. None of the estimates below is substituted into that experiment.

Historical aggregate annotations

The historical summary associates the following values with its 1991–2023 comparison and 30/17 temporal partition. The N values and intervals are the recorded annotations. Their presence does not establish a common denominator across rows.

Metric	**Result**	**N**	**CI (95%)**
Detection Rate	78.3%	47	[64.5%, 88.1%]
Mean Lead Time	127 days	37	[89, 165]
Median Lead Time	98 days	37	—
False Positive Rate	15.4%	43	[7.2%, 27.3%]
Positive Predictive Value	0.82	—	[0.69, 0.91]

Table E.2: Developmental historical aggregate annotations. N and the intervals are as reported for each measure. Appendix E gives the separate count-based evaluation of the 47-event working paper.

Two headline values have a direct basis in the selected roster in Appendix C: 18 detections among 23 cases give 78.3 percent, and the 18 listed lead times total 2,286 days, giving a 127-day mean. Their numerical agreement with this summary does not establish its different N annotations. In that roster the median is 90.5 days. A separate 37-lead analysis, its reported 98-day median, and its intervals require the corresponding observations. Thirty-seven detections among 47 events would give 78.7 percent, while 15.4 percent of 43 is not an integer count. Event assignments, any weighting, and the interval procedure are needed to connect those annotations to a reproducible evaluation.

The alternative historical tally

The following tally accompanies another historical summary. Its stated period and assignments differ from the mass-gap paper and from the 1991–2023 developmental summary above. Its three counted outcomes sum to 47, but six are false alerts: the tally contains 41 conflict outcomes and six non-conflict alert outcomes. That mixed total cannot serve as the denominator of a 47-conflict detection rate.

Recorded methodology: 47 post-Cold War interstate conflicts were described as identified from the Correlates of War dataset (1991–2022). The framework's thresholds were calibrated on a training set of 30 conflicts (64%), then applied out-of-sample to 17 validation cases (36%). Russia-Ukraine was in the validation set.

Aggregate Performance:

Metric	**Value**
True Positives (L4 → war)	34
False Positives (L4, no war)	6
False Negatives (war, no L4)	7
True Negatives	N/A (rare events)
Precision	0.85
Recall	0.83
Pooled F_1 from counts	0.84
Reported split summary[1]	**0.73**
Mean Lead Time	127 days
Median Lead Time	72 days

Train/Test Split Performance: The F1 score on the 30-conflict training set was 0.78; on the 17-conflict validation set, it was 0.69. The split summary is retained as reported. Its event assignments are needed to connect that partition to the mixed outcome tally above. Split scores alone cannot identify the population or establish freedom from overfitting.

Threshold and component summaries

These are the recorded sensitivity estimates. Their reported F_1 values range from 0.58 to 0.73 and their lead times from 89 to 184 days. A direct comparison requires the same eligible observations and event-window rule. Under that rule, removing one conjunct cannot lower detection. Because the two single-component rows do report lower detection, this table alone does not identify the effect of removing either feature. A refit, a different sample, or episode rules could alter the comparison; the table does not specify which. The set-inclusion calculation in Chapter 9 states the consistency test a replay must pass.

[1]These rows give $34/(34+6) = 0.85$ precision, $34/(34+7) \simeq 0.83$ recall and a pooled $F_1 = 68/81 \simeq 0.84$. The 34 hits and seven misses account for 41 conflict outcomes; adding six false alerts gives 47 mixed outcomes, not 47 conflicts. The case register must establish whether that mixed population or a separate 47-conflict population was intended. The printed split scores, 0.78 and 0.69, average to 0.735 before rounding and cannot determine an exact average from their displayed precision. Appendix E derives a separate 0.73 from the working paper's own counts.

Specification	Detect.	FP	F1	Lead
Baseline ($H < -0.5$, $\lambda_1 < 0.05$)	78.3%	15.4%	0.73	127d
Strict ($H < -0.7$, $\lambda_1 < 0.03$)	62.1%	8.2%	0.71	89d
Loose ($H < -0.3$, $\lambda_1 < 0.08$)	89.4%	28.6%	0.68	184d
H-only ($H < -0.5$, any λ_1)	74.2%	22.1%	0.65	142d
λ_1-only (any H, $\lambda_1 < 0.05$)	68.8%	31.4%	0.58	156d

Table E.3: Developmental threshold and component estimates; populations and episode rules must accompany any matched comparison.

Recorded cross-system comparison

System	Detect.	Lead	Scope
Davis Framework	78.3%	127 days	Bilateral trap conflicts
ViEWS (Uppsala)	12.4%*	30 days	All armed conflict
ACLED (monitoring)	N/A	N/A	Event tracking, not forecasting
PITF	8.1%*	45 days	State failure, mass atrocity
Early Warning Project	6.2%*	Annual	Mass atrocities

Table E.4: Developmental comparison annotations. The starred rates lack identified scoring populations and direct source tables in the available record; they are not used here as established performance estimates.

The starred rates and short lead times in this developmental comparison are retained as research annotations. The underlying scored populations and direct source tables for those three rates are not identified in the available record. They are therefore not adopted here as measured performance of VIEWS, PITF, or the Early Warning Project, and cannot support a ranking against Psychohistory. The systems' documented targets and versions appear in Chapters 9 and 35. A matched benchmark needs their actual forecasts or a documented reproduction on the common target.

Component rankings and internal voting settings

The developmental threshold record gives the following component areas:

Parameter	Threshold	AUC	Derivation
H^*	-0.5	0.89	ROC optimization, 6-month horizon
λ_1^*	0.05	0.76	Cheeger inequality + empirical
Γ^*	[0.9, 1.1]	0.82	Phase transition boundary detection
Combined	—	0.94	Multi-threshold voting

It also records these areas under nominal twenty-percent changes in the indicated setting:

Parameter	-20%	Baseline	+20%
H^* varied	0.87	0.94	0.91
λ_1^* varied	0.92	0.94	0.89
Γ^* varied	0.90	0.94	0.88

The reported combined AUC is 0.94, with 0.89 reported on the later 17-event partition; the accompanying developmental record describes 30 earlier events in 1991–2016 and 17 in 2017–2023, with over 2,000 monthly observations. The values are separate from the mass-gap study's AUC 0.84 and 1990–2014 / 2015–2022 split. A changed AUC requires a changed ranking or observation population, not just a different output cutoff. Reproducing this internal-voting experiment requires the component scores, voting rule, fitted thresholds, and labeled observations for every setting. The stated ten-percent margin around $\Gamma = 1$ similarly needs its input-error distribution before it can be treated as an uncertainty interval.

The register keeps exploratory findings available for further investigation while keeping each conclusion attached to the experiment that can support it. Historical interpretation, a reconstructed warning, a scored retrospective holdout, and a forecast issued before an event are distinct achievements. They can build on one another without being presented as the same achievement.

The CRITICAL-or-higher protocol

The computational appendix records a historical protocol that searches the preceding 24 months for the first CRITICAL-or-higher alert, with a three-year rolling baseline. Its reported results are:

Validation Results:

- Sensitivity: 41/47 ≃ 87.2% of conflicts detected at CRITICAL or higher. The selected historical Level 4 roster separately gives 18/23 = 78.3%. It has a different denominator.
- Specificity: 94% (stable dyads correctly classified)
- Mean lead time: 127 days (range: 42–312 days)
- False positive rate: 6% (across 200+ monitored dyads)

The 41 detected events leave six misses if all 47 share the stated eligibility rule. The specificity and false-positive percentage require the eligible negative observations and the episode grouping rule; a statement that more than 200 dyads were monitored does not supply that denominator. The selected historical roster also has a 127-day mean, but its range begins at 23 days, not 42. Agreement between the two means does not identify a common population. A first warning found anywhere in a 24-month search also differs from the mass-gap study's 180-day qualifying window. These are separate evaluation contracts.

What an Alert Record Contains

For each issued assessment, preserve the dyad, issue date, data cutoff, model version, parameter version, alert level, target event, and horizon. Add the input sources and their availability dates. Preserve uncertainty estimates and any abstention caused by insufficient information.

An alert episode also needs a start and an end. A daily Level 4 warning lasting three months should not automatically become ninety independent predictions. Conversely, a second crisis after a documented reset may warrant a new episode. State the rule before scoring the outcomes.

The event register should identify the onset source, the threshold used to define the event, and the treatment of continuing conflicts. If the target is escalation rather than onset, state that explicitly. A model that recognizes escalation after a war has begun is answering a different question from one that anticipates the first qualifying event.

At the end of a forecast horizon, assign the outcome under the original rule. If the horizon has not elapsed, mark the forecast unresolved. An unresolved warning is neither a success nor a false alarm.

The Information Available at the Time

For historical evaluation, construct the information set associated with each assessment date. A later revision to trade data, a declassified cable, or a retrospective account can improve explanation without having been available to a contemporary analyst.

Where the evaluation uses reconstructed historical data, document that construction. Where it simulates a contemporaneous assessment, preserve the release dates and use only information available at that point. This distinction lets the same research support both a historical mechanism study and a more demanding forecasting test.

Interpolation also belongs in the record. A line drawn between two annual observations does not create twelve independently observed monthly values. A centered smoother can borrow from the future; a trailing smoother does not, but responds differently to sudden changes. The implementation determines what the timing claim means.

A Fair Comparison

Compare the framework with alternatives on the same observations and target. A model predicting the continuation of civil violence and a model predicting interstate onset do not share a score merely because both report AUC. The event prevalence, horizon, geography, and evaluation period affect the result.

Begin with transparent benchmarks: the historical frequency, persistence of the current state, and a model of the main economic and diplomatic inputs without the additional geometric features. Then add those features in specified stages. This identifies what each contributes beyond information already contained in the inputs.

Threshold selection belongs inside the training procedure. If a threshold is chosen after examining the test outcomes, that test has become part of development. A new independent interval is then needed for the next evaluation. Preserve the developmental result; give it its correct role.

Learning From the Quiet Cases

An alert followed by peace can have several explanations. The diagnosis may have been too severe. Conditions may have changed independently. An intervention may have created an exit. The initial forecast record and the subsequent causal investigation should both survive.

For the investigation, record the intervention, its timing, its intended mechanism, and evidence that implementation changed the relevant constraint. The association of a summit with a peaceful outcome is a starting point. The sequence of commitments and actions can provide a much more informative account.

Missed events deserve an equally concrete analysis. Separate failures of coverage, measurement, timing, and mechanism. A sparse-data case may reveal where the model should abstain. A well-observed failure

may reveal a missing mechanism. Each points to a different next step in the research.

Replication lets a result travel between calculations and researchers, and from retrospective study into decisions made under uncertainty. A usable record allows the theory to accumulate evidence across cases, versions, and settings.

Appendix F

The New Helicity Research Records

Chapters 26–28 connect the book's measurement problem to later work in Helicity, GIGI, and PARALLAX. This appendix identifies what each calculation uses and what kind of result it supplies. The conflict studies retain their own event definitions, samples, and records in Appendix E; the studies below do not enlarge their denominators.

Constructions a reader can check

The three-currency example uses the constructed quotes $S_{AB} = 2$, $S_{BC} = 3$, and $S_{AC} = 5$. Its loop sum is $\log(6/5)$. Replacing the unit of C with one one-hundredth as large changes the last two quotes to 300 and 500, preserving the loop sum. The corresponding graph has three nodes, three edges, incidence rank two, and one independent cycle. Least-squares projection and a change of node units provide separate ways to verify the same invariant.

The negotiation example uses two piecewise linear paths around the unit square. Their signed areas are $+1/2$ and $-1/2$. The companion also checks that changing the timestamps without changing the ordered vertices preserves the value, while removing the intermediate corner changes it. These are exact properties of the declared paths, not empirical claims about historical negotiations.

The three-state Markov example uses forward, backward, and self-transition probabilities $(0.6, 0.2, 0.2)$. Its stationary vector is $(1/3, 1/3, 1/3)$ and its time-direction asymmetry is $0.4 \log 3$. The reversible control uses $(0.4, 0.4, 0.2)$. The calculation assumes a stationary first-order chain with the stated transition matrix.

The fuel example is a two-variable linear program. Its original bounds on x_A are $[10, 40]$; adding $35 \le x_B \le 40$ gives $[20, 25]$. The endpoint allocations are explicit witnesses. The analytic disk example separately checks the Schwarz-lemma bound and the positive-semidefinite Pick condition for compatible and incompatible readings.

Recorded empirical and implementation studies

The Helicity G3 study reads fourteen retained currency-cross files, inner-joins their dates, and computes log changes on 4,318 common level observations. Its graph is determined from the currency identifiers alone. The independent-column permutation uses seed 20260809. The saved result is in `integration/gauge/g3_hodge_split.json`; the generating calculation is `integration/gauge/g3_hodge_split.py`. This study measures the decomposition of the stated basket, not forecast accuracy. The book's companion reproduces it from the retained files without changing them.

The E17 sequence study is recorded in `integration/mockdesk/e17_levy_leadlag.json`, with its construction in the sibling Python file. Its synthetic gate uses seed 20260809 and seven declared shifts, including zero. The book reproduces that gate and uses the stored thinning results to distinguish timestamp reparameterization from loss of observations. A sign agreement on planted shifts is a scoped control for the estimator.

The E22 confirmation record is documented in `docs/E22_RUN_RECORD.md`. It is a recorded admission failure, not a newly executed confirmation experiment. The book does not rescore its consumed confirmation data.

The GIGI research draft dated October 29, 2026, version 0.2, reports the seventeen-field analytic study and its independent oracles Davis 2026b. Its 408 partial-interpolation queries and 816 complete-representation evaluation points concern two different information conditions. The figures quoted here come from that research record. They are not presented as a new replication in this book or as predictions of physical outcomes.

The October 31 structural-recovery note records the five-state, six-state, and broad-support option examples in `docs/OPTIONS_GIGI_STRUCTURAL_RECOVERY.md`. The later PARALLAX specification and its saved replay identify the query-acquisition calculation in `results/options_synthetic_parallax_v3/replay.json`. GIGI provides the revision-aware observation store in that replay; the price bands come from the consumer's linear program. The book's independent finite-support calculation checks the quoted bounds without starting a service or altering the research records.

The horizon adapter is in `analysis/gigi_bridge.py`, with its conventions in `docs/surfer_gigi_wiring_spec.md`. Its dispersion statistic, session graph, and error budget are specific estimator inputs. They are not interchangeable with the book's signed helicity index, an arbitrary graph's spectral gap, or a physical curvature tensor.

Notation at the boundary between the studies

In the gauge chapter, B is the node-by-edge incidence matrix, a contains log quotes, ψ describes changes of node units, and c is the cyclic residual. The scalar h is a particular loop sum. In the sequence chapter, A_{xy} is signed path area; F_{ij} is stationary transition flux, J_{ij} its antisymmetric current, and σ the nonnegative time-direction asymmetry of the declared Markov chain. These symbols are local to those constructions.

In the evidence chapter, $\mathcal{F}(D; M)$ denotes a family compatible with data D and model M. A *witness* is an explicit member that establishes an existence or attainability claim. An *admissible bound* describes the permitted range under stated constraints; it is not automatically a probability interval. $G(z)$ is a histogram's generating polynomial, and P is the Pick matrix. Neither its complex variable nor its imaginary component denotes the future.

The new companion is `companion/helicity_mathematics.py`, alongside `companion/helicity_mathematics_results.json`. It records source hashes, numerical checks, and the distinction between freshly reproduced calculations and cited prior experiments. Local research entries in the bibliography identify unpublished records by title and date; they do not imply a public archive or an assigned DOI.

Appendix G

Reproducing the Expanded Methods

The expanded methods companion accompanies the mathematical, African-case, implementation, and model-comparison chapters (Davis 2026j). Its numerical inputs are deliberately constructed and fully visible. They are not additional observations in the 47-conflict historical study.

The numerical examples

The file `companion/expanded_methods.py` reproduces the four-node spectral calculation, its perturbation interval and pseudoinverse update, the directed transient, the two temporal orders, the layered graph, the capacity-and-reserve example, and the probability and selection calculations. It imports the existing Psychohistory engine for the ordinary and disconnection examples without changing that engine.

Run the file with Python from the Helicity project root. It requires NumPy and SciPy, alongside the dependencies of the inspected engine. The result is written to `companion/expanded_methods_results.json`. The output includes named numerical checks, explicit teaching records, the supplied baseline, and a SHA-256 fingerprint of the engine source.

The figure generator uses those same mathematical definitions to produce the seven vector figures. The figures are drawn at the book's text width rather than reduced from a large presentation canvas. Their captions identify constructed examples so that a diagram's polish cannot be mistaken for a historical measurement.

The reference-edition calculations

The companion `companion/reference_methods.py` follows the additional worked examples through supply chains, delivery prerequisites, signed graphs, uncertain observations, time-dependent responses, finite reserves, calibrated probabilities, and comparative evaluation (Davis 2026k). It also checks the weighted graph decomposition, supply-bound

witnesses, and sequential prevention example. These are constructed calculations with visible inputs; they do not add cases to the historical conflict sample.

The checks use independent formulations where the example permits them: an optimization problem is checked against an explicit endpoint or analytical bound, a matrix calculation against its known eigenvalues or conservation identity, and a probability calculation against a directly enumerated population. The output file `companion/reference_methods_results.json` names each check and records the calculated and expected values. NumPy and SciPy supply the numerical routines.

A passing check establishes agreement for the stated example and numerical tolerance. It cannot establish that a real terminal has the assumed capacity, that an observed intervention was unconfounded, or that an operational forecast was issued before its outcome. Those questions return to source records and the historical evidence register. The companion makes the arithmetic inspectable while keeping its empirical meaning attached to the inputs.

The companion `companion/measurement_examples.py` checks the ordered-dyad example, candidate-space counts, feasible-set inclusion, nested warning rules, and the Taiwan scenario crossing dates. It also checks the cusp fold relation, the arithmetic of the reported Rubaya levy, and the elapsed dates between the Ethiopian agreements and the cited aid delivery. Its output, `companion/measurement_examples_results.json`, separates calculations from observations: reproducing a tax calculation or a date interval does not establish the associated mechanism or causal effect.

A common forecast file

The file `companion/compare_forecasts.py` scores aligned binary forecasts. Each prediction row supplies a model name, unit, issue time, target definition, target start, target end, probability, training cutoff, and information cutoff. The outcome file uses the same forecast key and a binary outcome. All timestamps include a time zone.

The program checks that every model predicts the same population, rather than taking a silent intersection that drops difficult cases. It reports Brier and logarithmic loss, ROC area when both outcome classes exist, and the confusion counts, precision, and recall at a supplied threshold. Undefined quantities are reported as undefined. A certain prediction contradicted by the outcome gives infinite logarithmic loss under the unclipped definition.

The `-self-test` option checks the scorer on constructed records and writes `companion/comparison_scorer_checks.json`. The ordinary mode accepts separate prediction and outcome files and an explicit output path. The self-test is a software check, not a comparative performance result for Psychohistory, VIEWS, CAST, or any other conflict system.

This scorer handles aligned binary issuance-level forecasts. Count distributions, alert-episode matching, dependence-aware uncertainty intervals, and causal intervention evaluation require the additional protocols described in Chapters 33–35 and Chapter 37. A common file format does not make different targets equivalent.

Historical sources and a future empirical contest

The African chapters cite the specific agreements, institutional statements, legal findings, and economic assessments used to establish their historical premises. Their small numerical illustrations are mechanism studies. A full regional forecasting evaluation would additionally require versioned event records, actor mappings, release histories, eligibility rules, and the complete forecast population.

For a head-to-head empirical comparison, retain archived forecasts where available and record any reproduced model's exact implementation. Freeze the common target and evaluation protocol before examining its test outcomes. Preserve quiet periods, missing forecasts, and all model variants considered during development. Report coverage and calibration along with discrimination and lead time.

Read the numerical companion alongside the historical evidence register: one reproduces the calculations; the other identifies the observations supporting a geopolitical forecast.

Appendix H

Notation and Mathematical Reading Guide

A symbol acquires its meaning from a definition, a domain, and a set of units. The same letter can do different work in fluid mechanics, a graph model, and the Psychohistory implementation. This guide keeps those uses distinguishable. The glossary explains the ideas in words; the index locates particular arguments and worked calculations.

The first question on encountering an equation is what its quantities describe. The second is what operations are permitted on them. The third is what empirical claim, if any, the calculation is intended to support. Following those three questions prevents an exact identity in one setting from becoming an unsupported assertion in another.

The framework's central readings

Symbol	Meaning and reading rule
H, fluid	Fluid helicity, typically $\int \mathbf{v} \cdot \boldsymbol{\omega}\, dV$, with $\boldsymbol{\omega} = \nabla \times \mathbf{v}$. Its units and conservation properties depend on the velocity field, domain, evolution, and boundary conditions.
H, signed	The alignment-based geopolitical reading. Its sign records the stipulated relation between economic connection and political orientation. The legacy engine uses flux times alignment. A negative value belongs to this signed interpretation, not to a ratio of two nonnegative quantities.
H^{dir}_{AB}	Ordered weighted-flow reading in Section 4. The pair $(H^{\mathrm{dir}}_{AB}, H^{\mathrm{dir}}_{BA})$ must retain actor roles. $\bar{H}$ is its symmetric average and ΔH its directional difference, not newly calibrated peace scores.

Symbol	Meaning and reading rule
H, capacity	The ordinary implementation's $\tau/(K + \varepsilon_H)$, with default $\varepsilon_H = 10^{-6}$. On the stated ordinary branch it is nonnegative. A historical signed threshold cannot be applied to it without a separately defined mapping.
H_d	The disconnection branch's nonnegative hostility index. The implementation stores $-H_d$ in the shared helicity field. Its name and branch must accompany the value; it is not a continuation of the ordinary capacity formula.
τ	Tolerance budget in the current ordinary specification, built from stated component scores. Earlier signed-torsion discussions use a different interpretation. In relaxation formulas the same Greek letter can instead denote a time; the local definition determines which.
K	Integration or effective curvature burden in the specified framework model. In the inspected implementation it is a dimensionless score from normalized inputs and proxies. It is not automatically a Riemannian scalar curvature measured from an independently supplied metric.
V	Economic flux score in the dyad implementation, built from quantities divided by their matching reference quantities. Elsewhere V can denote volume, a value function, or a viability set with a subscript. Check the local definition.
A	Signed alignment in the dyad model. In graph chapters, A commonly denotes an adjacency matrix; in sequence calculations, an area has a subscript such as A_{xy}. These objects have different domains and operations.
Γ	Historical trichotomy ratio, ideally $m\tau/[K \log \lvert S_{\text{exit}} \rvert]$ on its declared domain. It is distinct from a monotone measure of usable exits. The implementation adds guards and regularization and applies higher-priority alert tests. Appendix B supplies the executable convention.
m	Count of modeled diplomatic constraints in the capacity ratio. In general graph statements, m may instead count edges. A count is meaningful only after its counting rule has been defined.

Symbol	Meaning and reading rule
S_{exit}	Historical implementation count field. Its name must not equate candidate descriptions with usable settlements. S_{cand} denotes the declared candidate space for inference; F denotes feasible courses of action in Section 22. The disconnection model also uses S for security posture.
Z_H, Z_V	Standardized deviations from a specified baseline, such as $(H - \mu_H)/\sigma_H$. The baseline period, model version, and positive scale are part of the reading. Replacing the baseline changes the standardized quantity.
E_a	Modeled reconfiguration or activation scale $\gamma/(H^2 + \varepsilon_E)$. For fixed positive constants it is largest at $H = 0$ and decreases with $\lvert H \rvert$. The sign of H still matters to interpretation even though this formula uses its square.

The framework uses these quantities to organize a question about constrained relationships. Their shared notation does not eliminate the need to identify the particular model. A table comparing historical cases should use a common specification or explicitly retain the differences. A plotted line crossing from one branch to another needs a label explaining the changed meaning.

Graphs, spectra, and motion

Symbol	Meaning and reading rule
A, D, L	Adjacency, degree, and combinatorial Laplacian matrices, with $L = D - A$ for the specified undirected nonnegative graph. The degree is the sum of incident weights. Units of L follow the edge weights.
B, b_e	Incidence matrix and one edge's incidence column. Endpoint signs define bookkeeping orientation. In a system equation $\dot{x} = Gx + Bu$, B instead denotes an input map unless explicitly identified with an incidence matrix.

Symbol	Meaning and reading rule
$\mathcal{L}$	Symmetric degree-normalized Laplacian, $D^{-1/2}LD^{-1/2}$, with an explicit convention needed for zero-degree nodes. Its spectrum is dimensionless under the usual construction.
λ_1	Often the book's historical name for the first nonzero Laplacian eigenvalue. Standard one-based matrix formulas may instead call the zero eigenvalue λ_1 and the gap λ_2. The invariant description is "second-smallest eigenvalue, counting multiplicity."
g_L	Unambiguous gap notation in the expanded worked examples. On a disconnected graph the second-smallest eigenvalue is zero; connectedness and the number of components must be checked before interpreting a positive-mode formula.
v	An eigenvector or a state direction under the local definition. Eigenvectors can be multiplied by a nonzero scalar; unit normalization is required in the stated sensitivity formula. Near repeated eigenvalues, a whole eigenspace can be more stable than one chosen vector.
$L^{\dagger}$	Moore–Penrose pseudoinverse. For a connected undirected Laplacian, it inverts the nonconstant modes and vanishes on the constant vector. Joining disconnected components changes the nullspace and requires care with update formulas.
W, Σ	Positive-definite fitting weight and observation-error covariance in the weighted Hodge calculation. When $W = \Sigma^{-1}$, precision weights come from the stated error model. They do not independently establish which report is truthful.
ϕ, c_W, M	Fitted node potentials, weighted-fit residual, and residual map in Chapter 26. There $c_W = Ma$ and $BWc_W = 0$. The letter M elsewhere denotes a model or metric; the context determines which.
R_{ij}	Effective resistance $b^T L^{\dagger} b$ for $b = e_i - e_j$. It has the physical meaning of electrical resistance only for a conductance model. It is a different object from a freight bottleneck capacity.

Symbol	Meaning and reading rule
G	A dynamics generator in $\dot{x} = Gx$. Elsewhere the letter can denote a metric or a generating function. The sign, units, and state definition distinguish those uses. A trade adjacency is not automatically a rate generator.
e^{Gt}	Matrix exponential propagating a constant linear system for time t. For diffusion, the exponent may be $-\beta Lt$. The product in the exponent must be dimensionless.
P	Transition matrix in Markov models, a projection or aggregation map in other chapters, or the Nevanlinna–Pick matrix in analytic completion. A transition matrix is stochastic under its stated row or column convention; the other uses need not be.
π, F, J	Stationary probabilities, directed stationary flux, and net pairwise current: $F_{ij} = \pi_i P_{ij}$ and $J_{ij} = F_{ij} - F_{ji}$. Stationarity alone does not imply detailed balance.
σ	Time-direction asymmetry in the Markov example, measured in nats per discrete step. Elsewhere σ denotes a standard deviation. A physical entropy-production interpretation requires an additional physical model.
ω	Interlayer exchange rate in the multiplex diffusion example. Bold $\boldsymbol{\omega}$ denotes fluid vorticity elsewhere. Different fonts help distinguish them, but the surrounding equation supplies the definitive meaning.

Two matrices with the same eigenvalues can describe different relations. The signed-triangle calculation shows this directly: a zero mode representing agreement and one representing two opposed camps can share a spectrum. The operator's definition and its eigenvectors keep the interpretation attached to the result.

A similar distinction concerns the clock. If L is dimensionless and β has units per day, then $1/(\beta g_L)$ is a time. If L contains unconverted trade values, $1/g_L$ is not automatically a number of days. Numerical smallness cannot supply missing units.

Evidence families, paths, and constraints

Symbol	Meaning and reading rule
$\mathcal{F}(D; M)$	States compatible with observations D under model M. Bounds on a query are taken over this family. Changing the model family can change the bounds even when the observations remain fixed.
inf, sup	Greatest lower bound and least upper bound. A minimum or maximum additionally requires an attaining state. A witness at an endpoint demonstrates attainment; not every bounded problem has such a witness.
q_e, c_e	Flow and available capacity on edge e, in matching quantity-per-time units. In the evidence-family chapter, $q(x)$ instead denotes a query. A query can ask about a flow but is not itself an edge label.
x_k, u_k, d_k	State, decision, and demand or disturbance in a discrete update. Their meaning is specified by the application. State may include usable stock, pending obligations, or another measured quantity.
X_k, U_k	Allowed state and decision sets at time k. Membership in these sets expresses constraints, not probabilities. An empty feasible intersection calls for examining the requirements and assumptions.
V_k	States from which the remaining viability requirement can be met. Backward recursion depends on the horizon, terminal set, disturbance family, and whether a decision is made before or after an observation.
$\mathcal{I}_k$	Information available when decision k is made. A nonanticipating policy $u_k = \pi_k(\mathcal{I}_k)$ can adapt to received information but cannot choose an earlier action using an unobserved future state.
a, ϕ, c	Edge log quotations, fitted node potentials, and the graph Hodge cycle residual in Chapter 26. The residual satisfies the stated orthogonality condition for the chosen norm. It does not identify the cause of a discrepancy by itself.
ψ, γ	Node-unit change and a signed cycle in the gauge calculation. The identity $B\gamma = 0$ makes the loop sum invariant to $a \mapsto a + B^T\psi$. Elsewhere γ can be a model constant.

Symbol	Meaning and reading rule
A_{xy}	Signed area of a declared two-coordinate path. Its sign depends on axis order; its magnitude depends on coordinate units. A common increasing reparameterization preserves the path integral, while deleting an intermediate observation can change the path.
z, w, f	Complex query location, observed value, and holomorphic function in the completion problem. The bar denotes complex conjugation. Boundedness refers to the declared domain, usually the unit disk.
$p_j, G(z)$	Bin fractions and their generating polynomial in the measured-data encoding. These p_j are empirical bin proportions, not geopolitical event probabilities. The mathematical encoding is an explicit representation of the original observations.

The letter D illustrates why local definitions matter. It can denote a degree matrix, an observation collection, or a denominator. Those objects cannot be substituted for one another. The equation's dimensions, indices, and accompanying definition should make the intended use clear; if they do not, the calculation is not yet adequately specified.

The two most important set distinctions are between possible states and possible actions, and between individually possible values and jointly possible values. A data family records what the evidence permits us to believe about a state. A feasible-action set records what can be done under specified constraints. Two query values can each be individually permitted while failing to belong to any one common state. Chapter 28 works through this difference with deliveries and analytic functions.

Probability and evaluation

Symbol	Meaning and reading rule
Y, p	Defined binary outcome and its forecast probability. The population, issue time, horizon, and information set belong to the statement. A structural index in $[0, 1]$ is not a probability without a validated interpretation.

Symbol	Meaning and reading rule
λ, h_j	Continuous hazard rate or discrete interval hazard under the relevant model. The rate has inverse-time units; the interval hazard is a conditional probability. Competing risks alter cumulative probabilities of particular first exits.
ρ, N	Recording probability and underlying event count in the illustrative thinning model. The observation mean ρN does not separately identify both factors. Spatially dependent reporting may require more than independent binomial thinning.
$TP, FP,$ FN, TN	True-positive, false-positive, false-negative, and true-negative counts under one common outcome and matching rule. Event-level counts cannot be combined indiscriminately with daily or monthly opportunity counts.
P, R, F_1	Precision, recall, and their harmonic mean in classification formulas. Here P is not a transition matrix. Undefined denominators require an explicit convention rather than an invented favorable score.
BS, LL	Brier and logarithmic loss, with lower values preferred under their definitions. They evaluate reported probabilities. A comparison must preserve the same forecast opportunities and any clipping or missing-output convention.
$Y(a)$	Potential outcome under a specified action a in the causal chapter. It is not an observed pair of outcomes for one case. Estimating effects requires additional assumptions beyond accurately predicting the realized outcome.
$e(x)$	Propensity score: conditional probability of treatment given the measured pre-treatment state. The corresponding inverse-probability weights require overlap; they are not outcome probabilities.
A_0, L_1, A_1	Initial action, subsequent measured state, and later action in the sequential prevention example. A policy can let A_1 depend on L_1. Its evaluation must preserve that temporal order.

Read a probability statement aloud with its missing words restored: “For this defined population and event, using information available at this issuance, the probability within this horizon is p.” If any of those parts is absent, two apparently conflicting numbers may concern different questions.

For a score, restore the denominator: “Among these issued opportunities, using this outcome release and this matching rule, the average loss is this value.” The denominator is not a footnote to the claim. It is part of the claim’s meaning.

Three checks before trusting a formula

First, check units. A stock divided by a net rate gives a time; a monetary value divided by a matching reference gives a dimensionless score. Tonnes divided by a currency threshold does neither. Changing units requires matching changes to references and metrics.

Second, check the domain. Real logarithms require positive inputs; division needs a nonzero denominator; normalized Laplacians need a convention for isolated vertices; probabilities lie between zero and one. A safeguard that produces a finite answer may change the mathematical question. Specify its effect.

Third, check the quantifiers. One plan for one selected scenario differs from a policy for every admitted disturbance. A separate answer for each future may use information unavailable today. Individually attainable extremes need not be jointly attainable.

Together, these checks let the reader trace the argument: what the calculation establishes, which evidence supports its inputs, and which questions remain open.

Illustrations

1.1 Credit risk and structural strain 6
1.2 The scope of existing warning systems 12
1.3 The hard-onset problem 18
1.4 Reading a calibration curve 21
2.1 Three views of the trade-peace puzzle 33
4.1 The signed helicity scale 63
4.2 Four relationship types 68
5.1 A three-country network before and after a trade cut 82
7.1 Connections and the normalized spectral gap 97
7.2 The Europe–Russia spectral-gap sequence 98
8.1 Russia–Ukraine bilateral exports 114
8.2 The Russia–Ukraine helicity sequence 117
8.3 The reconstructed December 14 assessment 120
8.4 Gas transit through the warning window 128
9.1 Sample selection and an illustrative ROC curve 140
12.1 China–Taiwan: structure and coercive activity 164
12.2 A China–Taiwan continuation scenario 174
13.1 The Israel–Iran escalation sequence 189
14.1 India–Pakistan: conflict, trade, and structure 206
15.1 The 2025 cross-dyad assessment 217
18.1 Route capacity and the reserve clock 240
20.1 Reporting and aggregation can reverse a reading 266
22.1 Helicity alignment and illustrative responses 291
22.2 Coordinates in a relationship space 297
22.3 Connections and the combinatorial spectral gap 301
22.4 Trichotomy regimes and illustrative trajectories 305
23.1 Diffusion across a network bottleneck 314
23.2 Partitioning and comparison 316
24.1 Three views of slowing and structural order 324
24.2 Basins, barriers, and the political analogy 328
27.1 The same endpoint, a different order of events 347
29.1 An uncertain bridge can cross an alert threshold 366
30.1 Transient amplification despite stable eigenvalues 376
30.2 Temporal order and exchange between layers 380
33.1 Base rates and calibrated probabilities 411
37.1 The case mix changes the apparent effect 453

List of Tables

1.1 Major conflict forecasting systems and their primary focus areas. 10

2.1 Distribution of findings in 95 studies on economic interdependence and war (systematic review, *Cogent Social Sciences*, 2025). 32
2.2 Comparison of eventive and structural forecasting approaches. 41

4.1 The four quadrants of the helicity framework. 67
4.2 The four-level alert system. 74

7.1 Historical examples of glassy international systems. Note that some systems remain glassy for decades; others shatter violently. 96
7.2 The combined helicity-spectral gap signal. 101

8.1 Russia-Ukraine dyad profile, 2021. 113
8.2 Existing early warning systems and the Ukraine gap. . . 116
8.3 Reported case-network readings and assigned alert dates. The dates range from November 28 to December 21, 2021. Each row requires its own operator, normalization, and alert rule. 123
8.4 Public events alongside the reconstructed case indicators during the 72-day interval. The indicator sequence is a retrospective research record. 126

9.1 Primary data sources for historical backtesting. 138
9.2 The 47-event mass-gap evaluation 139
9.3 First warning dates in 18 historical cases 142
9.4 Diagnostic cases for coverage and timing. The Nagorno–Karabakh records concern different, incompletely aligned study specifications. 145
9.5 False positive L4 alerts within the 180-day evaluation window. 147
9.6 What must be matched across warning systems 149

12.1 China–Taiwan: dated evidence and measurement 163
12.2 Potential changes to the China–Taiwan assessment. A political event does not by itself determine a numerical helicity or spectral gap. 169

13.1 Israel-Iran dyad profile. 183

14.1 India-Pakistan dyad profile. 203
14.2 India-Pakistan framework readings at key historical moments. 207

15.1 US-China economic interdependence, 2024. 214
15.2 Conflicts outside the framework's primary scope. 222

23.1 Recovery times in the Russia–Ukraine record 318

25.1 The hierarchy of physics-to-politics mappings. Confidence decreases from top to bottom. The framework's core claims rest on the top two levels; the lower levels motivate further research. 333
25.2 The physics-to-politics concept mapping. 334

35.1 Questions for a fair comparison 432
35.2 Checking a paired forecast comparison 435

37.1 Selection can reverse a crude comparison 452

A.1 Primary political data sources. 480

C.1 Selected historical first-warning entries. These are distinct from the mass-gap paper's dated onset records. 500

E.1 Classifier and baseline results reported in the Psychohistory paper 532
E.2 Developmental historical aggregate annotations 535
E.3 Developmental threshold and component estimates; populations and episode rules must accompany any matched comparison. 537
E.4 Scope and reported measures of five warning systems . . 537

Bibliography

ACLED (2026a). *CAST Endpoint*. URL: https://acleddata.com/api-documentation/cast-endpoint (visited on 11/02/2026).

– (2026b). *CAST Methodology*. Rolling four-week forecast specification as accessed November 7, 2026. URL: https://acleddata.com/methodology/cast-methodology (visited on 11/07/2026).

– (2026c). *Methodology and Codebook*. URL: https://acleddata.com/conflict-data/knowledge-base/methodology (visited on 11/07/2026).

African Union (Jan. 18, 2017a). *AU Expresses Grave Concern over State of Emergency, Insists President Yahya Jammeh Must Hand Over Power to President-Elect Adama Barrow*. URL: https://au.int/sites/default/files/pressreleases/31874-pr-au_expresses_grave_concern_gambia_en.pdf (visited on 11/07/2026).

– (Jan. 22, 2017b). *AU Welcomes Bloodless Transition in The Gambia*. URL: https://au.int/pt/pressreleases/20170122/au-welcomes-bloodless-transition-gambia-0 (visited on 11/07/2026).

– (May 2, 2023). *African Union Condemns Violence in Sudan, Calls for Humanitarian Ceasefire and Urges Coordinated International Support*. URL: https://au.int/en/pressreleases/20230502/african-union-condemns-violence-sudan-calls-humanitarian-ceasefire-and-urges (visited on 11/07/2026).

Agler, Jim and John E. McCarthy (2002). *Pick Interpolation and Hilbert Function Spaces*. Vol. 44. Graduate Studies in Mathematics. American Mathematical Society. DOI: 10.1090/gsm/044.

Alliance Treaty Obligations and Provisions Project (Aug. 2022). *ATOP Data and Codebook, Version 5.1*. URL: https://www.atopdata.org/documentation.html (visited on 11/07/2026).

Altafini, Claudio (2012). "Dynamics of Opinion Forming in Structurally Balanced Social Networks". In: *PLOS ONE* 7.6, e38135. DOI: 10.1371/journal.pone.0038135.

Angell, Norman (1909). *The Great Illusion: A Study of the Relation of Military Power to National Advantage*. New York: G.P. Putnam's Sons.

– (1913). *The Great Illusion: A Study of the Relation of Military Power to National Advantage*. 4th ed. G. P. Putnam's Sons. URL: https://www.

gutenberg.org/files/38535/38535-h/38535-h.htm (visited on 11/02/2026).

ASML Holding N.V. (2025). *Annual Report 2024: Based on IFRS*. ASML. URL: https://ourbrand.asml.com/m/1d935e9653a216d7/original/2024-Annual-Report-based-on-IFRS.pdf (visited on 11/07/2026).

Aubin, Jean-Pierre (2009). *Viability Theory*. Reprint of the 1991 work. Birkhäuser. DOI: 10.1007/978-0-8176-4910-4.

Australian Department of Foreign Affairs and Trade (2024). *Annual Report 2023–24*. Performance results for Outcome 1. Commonwealth of Australia. URL: https://www.dfat.gov.au/sites/default/files/dfat-annual-report-2023-24.pdf (visited on 11/02/2026).

Australian Minister for Foreign Affairs (Apr. 11, 2023). *Step Forward to Resolve Barley Dispute with China*. URL: https://www.foreignminister.gov.au/minister/penny-wong/media-release/step-forward-resolve-barley-dispute-china (visited on 11/02/2026).

– (Dec. 20, 2024). *Effective Immediately: Live Rock Lobster to China to Restart*. URL: https://www.foreignminister.gov.au/minister/penny-wong/media-release/effective-immediately-live-rock-lobster-china-restart (visited on 11/02/2026).

Bacon, Ruth (July 31, 1941). *Memorandum by Miss Ruth Bacon of the Division of Far Eastern Affairs*. Foreign Relations of the United States, 1941, Volume IV, Document 671. US Department of State. URL: https://history.state.gov/historicaldocuments/frus1941v04/d671 (visited on 11/02/2026).

Barbieri, Katherine (1996). "Economic Interdependence: A Path to Peace or a Source of Interstate Conflict?" In: *Journal of Peace Research* 33.1, pp. 29–49.

Bayer, Reşat (2006). *Diplomatic Exchange Data Set*. Version 2006.1. URL: https://correlatesofwar.org/data-sets/diplomatic-exchange/ (visited on 11/07/2026).

Betts, Richard K. (1982). *Surprise Attack: Lessons for Defense Planning*. Washington, D.C.: Brookings Institution Press.

Blanchette, Jude, Gerard DiPippo, and Christopher B. Johnstone (Dec. 13, 2023). *Scared Strait: Understanding the Economic and Financial Impacts of a Taiwan Crisis*. Center for Strategic and International Studies. URL: https://www.csis.org/analysis/scared-strait-understanding-economic-and-financial-impacts-taiwan-crisis (visited on 11/07/2026).

Boyd, Stephen and Lieven Vandenberghe (2004). *Convex Optimization*. Cambridge University Press. URL: https://web.stanford.edu/~boyd/cvxbook/.

Bureau of Industry and Security (Oct. 7, 2022). *Commerce Implements New Export Controls on Advanced Computing and Semiconductor Manufacturing Items to the People's Republic of China*. U.S. Department of Commerce. URL: https://media.bis.gov/press-release/commerce-implements-new-export-controls-advanced-computing-semiconductor-manufacturing-items-peoples (visited on 11/07/2026).

Burns, William J. (July 1, 2023a). *A World Transformed and the Role of Intelligence*. 59th Ditchley Annual Lecture. Ditchley Foundation. URL: https://www.ditchley.com/sites/default/files/Ditchley%20Annual%20Lecture%202023%20transcript.pdf (visited on 11/07/2026).

– (Feb. 26, 2023b). *Interview on Face the Nation*. Interview by Margaret Brennan; transcript. CBS News. URL: https://www.cbsnews.com/news/william-burns-cia-director-face-the-nation-transcript-02-26-2023/ (visited on 11/07/2026).

Cederman, Lars-Erik and Nils B. Weidmann (2017). "Predicting Armed Conflict: Time to Adjust Our Expectations?" In: *Science* 355.6324, pp. 474–476.

Chen, Shuo, Matthew Fuhrmann, and Eric K. Stern (2022). "Tipping Points in International Crises: A Literature Review and Research Agenda". In: *International Studies Review* 24.3.

Chung, Fan R. K. (1997). *Spectral Graph Theory*. Providence, RI: American Mathematical Society.

Congressional Research Service (Apr. 4, 2022). *Taiwan: Political and Security Issues*. IF10275. Congressional Research Service. URL: https://www.congress.gov/crs_external_products/IF/PDF/IF10275/IF10275.56.pdf (visited on 11/07/2026).

– (Aug. 18, 2025). *The U.S. "One-China" Policy and Taiwan*. IF12503. Congressional Research Service. URL: https://www.congress.gov/crs_external_products/IF/PDF/IF12503/IF12503.4.pdf (visited on 11/07/2026).

Copeland, Dale C. (2015). *Economic Interdependence and War*. Princeton: Princeton University Press.

Correlates of War Project (2026). *Militarized Interstate Disputes, Version 5.0: Data and Codebooks*. URL: https://correlatesofwar.org/data-sets/mids/ (visited on 11/07/2026).

CSIS China Power Project (Jan. 22, 2024). *Surveying the Experts: U.S. and Taiwan Views on China's Approach to Taiwan in 2024 and Beyond*. Updated June 5, 2024; survey conducted November 28–December 15, 2023. URL: https://chinapower.csis.org/surveying-experts-us-and-taiwan-views-china-approach-taiwan-2024/ (visited on 11/07/2026).

Davis, Bee Rosa (Dec. 12, 2025). "Predicting Interstate Conflict 72 Days in Advance: A Topological Approach to Weaponized Interdependence". Psychohistory working paper, version 2; primary results, classifier comparisons, and Appendix A.7. DOI: 10.5281/zenodo.17905678. URL: https://zenodo.org/records/17905678.

– (Oct. 31, 2026a). "Fitting the Wine Example to Options". Local structural-recovery research note; docs/OPTIONS_GIGI_STRUCTURAL_RECOVERY.md.

– (Oct. 29, 2026b). "GIGI: A Fiber-Bundle Database: Mutation-Consistent Geometric Inference and Bounded Analytic Completion". Research draft, version 0.2. Davis Geometric; local manuscript and retained experimental records.

– (2026c). "Helicity E17: Levy Area and Asynchronous Lead–Lag". Local experiment source and saved numerical record; integration/mockdesk/e17_levy_leadlag; inspected November 4, 2026.

– (Aug. 19, 2026d). "Helicity E22: Detailed-Balance Confirmation Run Record". Registered version 3.5; local record, docs/E22_RUN_RECORD.md.

– (2026e). "Helicity Gauge Instrument: G3 Hodge Split". Local research code, retained currency-cross data, and numerical record; integration/gauge; record inspected November 4, 2026.

– (2026f). "Helicity PARALLAX: Find the Missing Mark, Know What to Refresh Next". Local product specification and synthetic replay, version 3; docs/PARALLAX_PRODUCT_SPEC.md and results/options_synthetic_parallax_v3; inspected November 4, 2026.

– (Mar. 2, 2026g). *Psychohistory China–Taiwan Prediction: Structural Watch and Scenario Notes*. Local Helicity research record, header date; reports/china_taiwan_prediction.md. Scenario hypotheses and version-specific metrics; no public archive identifier assigned.

– (Jan. 2026h). "PSYCHOHISTORY v2.0: Topological Conflict Prediction via Geopolitical Helicity". Field Equation Alignment specification. Parts I and II distinguish the general solution-space construction from its geopolitical count field. Local research record, docs/PSYCHOHISTORY_v2_SPEC.md; no public archive identifier assigned.

– (2026i). "Surfer–GIGI Wiring Specification and Horizon Adapter". Local implementation conventions; docs/surfer_gigi_wiring_spec.md and analysis/gigi_bridge.py; inspected November 4, 2026.

– (Nov. 7, 2026j). "The Geometry of War: Expanded Methods and Common-Forecast Comparison Companion". Version 1. Constructed mathematical examples, implementation fixtures, numerical results,

and aligned binary-forecast scorer. Distributed with the book source in the Helicity project; no public repository identifier assigned.

Davis, Bee Rosa (Nov. 7, 2026k). *The Geometry of War: Reference-Edition Numerical Companion*. Version 1.0. Local reproducibility record: companion/reference_methods.py and reference_methods_results.json. Constructed examples; no public archival identifier assigned.

Dobrynin, Anatoly (Oct. 27, 1962). *Telegram to the Soviet Foreign Ministry Reporting on Meeting with Robert Kennedy*. National Security Archive. URL: https://nsarchive.gwu.edu/document/29362-document-1-soviet-embassy-telegram-soviet-foreign-ministry-ambassador-dobrynin (visited on 11/02/2026).

Driedger, Jonas and Mikhail Polianskii (2023). "The Expert Failure to Predict Russian Aggression Against Ukraine". In: *International Affairs* 99.4, pp. 1507–1527.

Economic Community of West African States (Jan. 29, 2025). *Press Statement on the Withdrawal of Burkina Faso, Mali and Niger*. Statement dated January 29; website post dated January 30. URL: https://www.ecowas.int/press-statement-2/ (visited on 11/07/2026).

ECOWAS, African Union and United Nations (Jan. 21, 2017). *Joint Declaration on the Political Situation of the Islamic Republic of The Gambia*. Declaration signed in Banjul. URL: https://au.int/sites/default/files/pressreleases/31910-pr-ecowas-au-un-joint-communique.pdf (visited on 11/07/2026).

Ethiopian National Defense Force and Tigray Armed Combatants (Nov. 12, 2022). *Declaration of the Senior Commanders on the Modalities for the Implementation of the Agreement for Lasting Peace Through a Permanent Cessation of Hostilities*. Nairobi declaration; agreement text preserved in the University of Edinburgh PA-X collection, agreement 2464. URL: https://www.peaceagreements.org/agreements/2464/ (visited on 11/07/2026).

European Commission (2026). *Turkiye*. URL: https://enlargement.ec.europa.eu/countries/turkiye_en (visited on 11/02/2026).

European Parliament Historical Archives (2026). *Ad Hoc Assembly 1952–1954*. URL: https://historicalarchives.europarl.europa.eu/home/fonds/parliamentary-activities/ad-hoc-assembly.html (visited on 11/02/2026).

Farrell, Henry and Abraham L. Newman (2019). "Weaponized Interdependence: How Global Economic Networks Shape State Coercion". In: *International Security* 44.1, pp. 42–79.

– (2023). *Underground Empire: How America Weaponized the World Economy*. New York: Henry Holt.

Fearon, James D. (1995). "Rationalist explanations for war". In: *International Organization* 49.3, pp. 379–414. DOI: 10.1017/S0020818300033324.

Flenner, Elijah and Grzegorz Szamel (2005). "Relaxation in a Glassy Binary Mixture: Mode-Coupling-Like Power Laws, Dynamic Heterogeneity, and a New Non-Gaussian Parameter". In: *Physical Review E* 72, p. 011205. DOI: 10.1103/PhysRevE.72.011205.

Franco-German Youth Office (2026). *The Elysee Treaty*. URL: https://www.fgyo.org/the-institution/history-franco-german-youth-office/elysee-treaty (visited on 11/02/2026).

Freedman, Lawrence (2022). *Command: The Politics of Military Operations from Korea to Ukraine*. Penguin Press.

French Ministry for Europe and Foreign Affairs (2026). *The Elysee Treaty in Six Questions*. URL: https://www.diplomatie.gouv.fr/en/the-ministry-in-action/ensuring-the-presence-of-french-culture/franco-german-cooperation/the-elysee-treaty-in-six-questions (visited on 11/02/2026).

Gartzke, Erik (2007). "The Capitalist Peace". In: *American Journal of Political Science* 51.1, pp. 166–191.

Gas Transmission System Operator of Ukraine (2022). *Key Indicators for Gas Transportation of the Gas TSO of Ukraine (GTSOU) in 2021*. URL: https://tsoua.com/en/news/key-indicators-for-gas-transportation-of-the-gas-tso-of-ukraine-gtsou-in-2021/ (visited on 11/07/2026).

Global Sanctions Data Base (2026). *Data, Release History, and Documentation*. URL: https://www.globalsanctionsdatabase.com/ (visited on 11/07/2026).

Gneiting, Tilmann and Adrian E. Raftery (2007). "Strictly Proper Scoring Rules, Prediction, and Estimation". In: *Journal of the American Statistical Association* 102.477, pp. 359–378. URL: https://sites.stat.washington.edu/raftery/Research/PDF/Gneiting2007jasa.pdf.

Goldstone, Jack A., Robert H. Bates, David L. Epstein, Ted Robert Gurr, Michael B. Lustik, Monty G. Marshall, Jay Ulfelder, and Mark Woodward (2010). "A Global Model for Forecasting Political Instability". In: *American Journal of Political Science* 54.1, pp. 190–208. DOI: 10.1111/j.1540-5907.2009.00426.x.

Government of India and Government of Pakistan and International Bank for Reconstruction and Development (Sept. 19, 1960). *The Indus Waters Treaty 1960*. URL: https://treaties.un.org/Pages/showDetails.aspx?clang=_en&objid=0800000280135336 (visited on 11/07/2026).

Government of the Federal Democratic Republic of Ethiopia and Tigray People's Liberation Front (Nov. 2, 2022). *Agreement for Lasting Peace through a Permanent Cessation of Hostilities*. Pretoria agreement; agreement text catalogued by UNHCR Refworld. URL: https://www.refworld.org/legal/agreements/natauth/2022/151542 (visited on 11/07/2026).

Hafezi, Parisa (Mar. 1, 2026). *Iran's Ali Khamenei, Who Based Iron Rule on Fiery Hostility to US and Israel, Dies at 86*. Reuters. URL: https://www.investing.com/news/world-news/irans-ali-khamenei-who-based-iron-rule-on-fiery-hostility-to-us-and-israel-dies-at-86-4533459 (visited on 11/02/2026).

Hairer, Ernst, Syvert Paul Nørsett, and Gerhard Wanner (1993). *Solving Ordinary Differential Equations I: Nonstiff Problems*. 2nd ed. Springer. URL: https://www.unige.ch/~wanner/ode_I.html.

Hart, Brian (Feb. 14, 2025). *China Escalates Cross-Strait Military Activity Under Taiwan President William Lai*. Center for Strategic and International Studies. URL: https://www.csis.org/analysis/china-escalates-cross-strait-military-activity-under-taiwan-president-william-lai (visited on 11/07/2026).

Hegre, Håvard, John R. Oneal, and Bruce Russett (2010). "Trade Does Promote Peace: New Simultaneous Estimates of the Reciprocal Effects of Trade and Conflict". In: *Journal of Peace Research* 47.6, pp. 763–774.

Hernán, Miguel A. and James M. Robins (2020). *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC. URL: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/.

Hill, Fiona (Mar. 2022). *Interview on Russia-Ukraine Invasion*. New York Times.

Hill, Fiona and Clifford G. Gaddy (2013). *Mr. Putin: Operative in the Kremlin*. Brookings Institution Press.

Högbladh, Stina (2025). *UCDP Georeferenced Event Dataset Codebook*. Version 25.1. Fixed historical version cited for the coding discussion; later releases are separate vintages. Department of Peace and Conflict Research, Uppsala University. URL: https://ucdp.uu.se/downloads/ged/ged251.pdf (visited on 11/07/2026).

International Committee of the Red Cross (Nov. 15, 2022). *Ethiopia: Lifesaving Medical Supplies Arrive in Tigray*. URL: https://www.icrc.org/en/document/ethiopia-lifesaving-medical-supplies-arrive-in-tigray (visited on 11/07/2026).

International Court of Justice (Dec. 19, 2005). *Armed Activities on the Territory of the Congo (Democratic Republic of the Congo v. Uganda): Judgment*. URL: https://api.icj-cij.org/sites/default/files/

case-related/116/116-20051219-JUD-01-00-EN.pdf (visited on 11/07/2026).

International Energy Agency (2026). *Data and Statistics*. URL: https://www.iea.org/data-and-statistics (visited on 11/07/2026).

International Monetary Fund (Oct. 2, 2024). *IMF Staff Completes Mission to South Sudan for the Third Review of a Staff-Monitored Program with Board Involvement*. Press Release No. 24/354; staff mission findings. URL: https://www.imf.org/en/news/articles/2024/10/01/pr-24354-south-sudan-imf-completes-mission-3rd-rev-program-with-board-involvement (visited on 11/07/2026).

– (2026). *International Trade in Goods by Partner Country (IMTS, Formerly DOTS)*. URL: https://data.imf.org/datasets/IMF.STA%3AIMTS (visited on 11/07/2026).

Japan Center for Asian Historical Records (2026a). *The US-Japan War Talks as Seen in Official Documents: Oil Restrictions and the Freezing of Assets*. Documents B02030753200 and C04123206200, 1941. URL: https://www.jacar.go.jp/english/exhibition/nichibei/popup/pop_13.html (visited on 11/02/2026).

– (2026b). *The US-Japan War Talks as Seen in Official Documents: The Hull Note*. Japanese diplomatic telegrams of November 26–27, 1941. URL: https://www.jacar.go.jp/english/exhibition/nichibei/popup/pop_26.html (visited on 11/02/2026).

Jervis, Robert (2010). "Why Intelligence Fails: Lessons from the Iranian Revolution and the Iraq War". In.

Jiang, Xiaoye, Lek-Heng Lim, Yuan Yao, and Yinyu Ye (2011). "Statistical Ranking and Combinatorial Hodge Theory". In: *Mathematical Programming* 127, pp. 203–244. DOI: 10.1007/s10107-010-0419-x. arXiv: 0811.1067.

Kant, Immanuel (1795). *Perpetual Peace: A Philosophical Sketch*. Original: Zum ewigen Frieden.

Keita, Bintou (Sept. 30, 2024). *Security Council Briefing on MONUSCO by Special Representative of the Secretary-General Bintou Keita*. United Nations Peacekeeping. URL: https://peacekeeping.un.org/en/node/132163 (visited on 11/07/2026).

Keynes, John Maynard (1919). *The Economic Consequences of the Peace*. New York: Harcourt, Brace and Howe.

Kingdom of Saudi Arabia and United States of America (May 20, 2023). *Agreement on a Short-Term Ceasefire and Humanitarian Arrangements in Sudan*. Joint announcement published by the US Department of State. URL: https://2021-2025.state.gov/agreement-on-a-short-

term-ceasefire-and-humanitarian-arrangements-in-sudan/ (visited on 11/07/2026).

Kingdom of Saudi Arabia, Intergovernmental Authority on Development, and United States of America (Nov. 7, 2023). *Joint Statement of Commitments from Jeddah Talks Between Sudanese Armed Forces and Rapid Support Forces*. Statement and attached party commitments; text preserved in the PA-X collection. URL: https://www.peaceagreements.org/media/documents/SD_231107_Joint_Statement_of_Commitments.pdf (visited on 11/07/2026).

Korovkin, Vasily and Alexey Makarin (2023). "Conflict and Intergroup Trade: Evidence from the 2014 Russia-Ukraine Crisis". In: *American Economic Review* 113.1, pp. 34–70.

Levin, David A. and Yuval Peres (2017). *Markov Chains and Mixing Times*. 2nd ed. American Mathematical Society. DOI: 10.1090/mbk/107.

Lin, Bonny, Brian Hart, Matthew P. Funaiole, Samantha Lu, and Truly Tinsley (June 5, 2024). *How China Could Quarantine Taiwan: Mapping Out Two Possible Scenarios*. Center for Strategic and International Studies. URL: https://www.csis.org/analysis/how-china-could-quarantine-taiwan-mapping-out-two-possible-scenarios (visited on 11/07/2026).

Lorde, Audre (1984). *Sister Outsider: Essays and Speeches*. Trumansburg, NY: Crossing Press.

Lyons, Terry J. (1998). "Differential Equations Driven by Rough Signals". In: *Revista Matematica Iberoamericana* 14.2, pp. 215–310. DOI: 10.4171/RMI/240.

MacMillan, Margaret (2013). *The War That Ended Peace: The Road to 1914*. New York: Random House.

Ministry of Foreign Affairs of Japan (Feb. 22, 2022). *Russia's Recognition of the "Independence" of the "Donetsk People's Republic" and the "Luhansk People's Republic"*. Statement by Foreign Minister Hayashi Yoshimasa. URL: https://www.mofa.go.jp/press/release/press4e_003084.html (visited on 11/07/2026).

Moffatt, H. K. (1969). "The degree of knottedness of tangled vortex lines". In: *Journal of Fluid Mechanics* 35.1, pp. 117–129.

– (2021). "Some Topological Aspects of Fluid Dynamics". In: *Journal of Fluid Mechanics* 914, P1. DOI: 10.1017/jfm.2020.230.

Moreau, Luc and Paolo Missier (Apr. 30, 2013). *PROV-DM: The PROV Data Model*. W3C Recommendation; editors. World Wide Web Consortium. URL: https://www.w3.org/TR/prov-dm/ (visited on 11/07/2026).

Morgan, T. Clifton, Navin Bapat, and Yoshiharu Kobayashi (2014). "Threat and Imposition of Economic Sanctions 1945–2005: Updating the TIES

Dataset". In: *Conflict Management and Peace Science* 31.5, pp. 541–558. URL: https://sanctions.web.unc.edu/ (visited on 11/07/2026).

Mueller, Hannes and Christopher Rauh (2022). "The Hard Problem of Prediction for Conflict Prevention". In: *Journal of the European Economic Association* 20.6, pp. 2318–2352.

Narang, Vipin (2014). *Nuclear Strategy in the Modern Era: Regional Powers and International Conflict*. Princeton: Princeton University Press.

National People's Congress (2018). *Constitution of the People's Republic of China*. Preamble; English text incorporating amendments through March 11, 2018. URL: https://english.www.gov.cn/archive/lawsregulations/202006/04/content_WS5ed8856ec6d0b3f0e9499913.html (visited on 11/07/2026).

NATO (Apr. 13, 2021). *NATO Secretary General and Ukrainian Foreign Minister Discuss Russia's Military Build-up*. URL: https://www.nato.int/en/news-and-events/articles/news/2021/04/13/nato-secretary-general-and-ukrainian-foreign-minister-discuss-russias-military-build-up (visited on 11/07/2026).

OECD (2016). *OECD Due Diligence Guidance for Responsible Supply Chains of Minerals from Conflict-Affected and High-Risk Areas*. 3rd ed. Paris: OECD Publishing. DOI: 10.1787/9789264252479-en.

Parliament of Australia (2026). *Australian Trade in Figures*. Briefing Book for the 47th Parliament. Parliamentary Library. URL: https://www.aph.gov.au/About_Parliament/Parliamentary_departments/Parliamentary_Library/Research/Briefing_Book/47th_Parliament/AustralianTradeFigures (visited on 11/02/2026).

Permanent Court of Arbitration (2026). *Eritrea–Ethiopia Boundary Commission*. Institutional record of the boundary commission established under the Algiers Agreement. URL: https://pca-cpa.org/en/cases/99/ (visited on 11/07/2026).

Plokhy, Serhii (2017). *Lost Kingdom: The Quest for Empire and the Making of the Russian Nation*. New York: Basic Books.

Prentice, R. L. and R. Pyke (1979). "Logistic Disease Incidence Models and Case-Control Studies". In: *Biometrika* 66.3, pp. 403–411. DOI: 10.1093/biomet/66.3.403.

Presidency of the French Republic (Feb. 12, 2022). *Entretien téléphonique avec le Président de la Fédération de Russie Vladimir Poutine*. Readout identifying the preceding February 7 meeting in Moscow. URL: https://www.elysee.fr/emmanuel-macron/2022/02/12/entretien-telephonique-avec-le-president-de-la-federation-de-russie-vladimir-poutine (visited on 11/07/2026).

Raleigh, Clionadh, Roudabeh Kishi, and Andrew Linke (2022). "An Agenda for Addressing Bias in Conflict Data". In: *Scientific Data* 9, 593. DOI: 10.1038/s41597-022-01705-8.

Rhodium Group (Dec. 14, 2022). *The Global Economic Disruptions from a Taiwan Conflict*. URL: https://rhg.com/research/taiwan-economic-disruptions/ (visited on 11/07/2026).

Russett, Bruce and John R. Oneal (2001). *Triangulating Peace: Democracy, Interdependence, and International Organizations*. New York: W.W. Norton.

Sandhu, Romeil, Tryphon Georgiou, Ed Reznik, Liangjia Zhu, Ivan Kolesov, Yasin Senbabaoglu, and Allen Tannenbaum (2015). "Graph Curvature for Differentiating Cancer Networks". In: *Scientific Reports* 5, 12323. DOI: 10.1038/srep12323.

Sandhu, Romeil S., Tryphon T. Georgiou, and Allen R. Tannenbaum (2016). "Ricci Curvature: An Economic Indicator for Market Fragility and Systemic Risk". In: *Science Advances* 2.5, e1501495. DOI: 10.1126/sciadv.1501495.

Scheffer, Marten, Jordi Bascompte, William A. Brock, Victor Brovkin, Stephen R. Carpenter, Vasilis Dakos, Hermann Held, Egbert H. van Nes, Max Rietkerk, and George Sugihara (2009). "Early-warning signals for critical transitions". In: *Nature* 461.7260, pp. 53–59.

Schincariol, Victoria E. et al. (2025). "Forecasting Migration Flows Using Machine Learning". In: *Working Paper*.

Scholz, Olaf (Feb. 27, 2022). *Policy Statement to the German Bundestag*. Federal Government of Germany. URL: https://www.bundesregierung.de/breg-en/service/archive/policy-statement-by-olaf-scholz-chancellor-of-the-federal-republic-of-germany-and-member-of-the-german-bundestag-27-february-2022-in-berlin-2008378 (visited on 11/07/2026).

Schuman, Robert (May 9, 1950). *Schuman Declaration*. URL: https://european-union.europa.eu/principles-countries-history/history-eu/1945-59/schuman-declaration-may-1950_en (visited on 11/02/2026).

Semiconductor Industry Association (Apr. 5, 2021). *Written Comments on Risks in the Semiconductor Manufacturing and Advanced Packaging Supply Chain*. Submission to the U.S. Department of Commerce; includes the SIA/BCG supply-chain study. Semiconductor Industry Association. URL: https://www.semiconductors.org/wp-content/uploads/2021/04/4.5.21-SIA-supply-chain-submission.pdf (visited on 11/07/2026).

State of Eritrea and Federal Democratic Republic of Ethiopia (July 9, 2018). *Joint Declaration of Peace and Friendship between Eritrea and Ethiopia*. Declaration signed in Asmara; text hosted by the Eritrea Ministry of Information. URL: https://shabait.com/2018/07/09/joint-declaration-of-peace-and-friendship-between-eritrea-and-ethiopia/ (visited on 11/07/2026).

Stoltenberg, Jens (Nov. 30, 2021). *Doorstep Statement at the NATO Foreign Ministers Meeting in Riga*. NATO. URL: https://www.nato.int/en/news-and-events/events/transcripts/2021/11/30/doorstep-statement (visited on 11/07/2026).

– (Jan. 12, 2022). *Doorstep Statement in Brest Following the NATO–Russia Council*. NATO. URL: https://www.nato.int/en/news-and-events/events/transcripts/2022/01/12/doorstep-statement (visited on 11/07/2026).

Strick, T. R., V. Croquette, and D. Bensimon (2000). "Single-Molecule Analysis of DNA Uncoiling by a Type II Topoisomerase". In: *Nature* 404, pp. 901–904. URL: https://pubmed.ncbi.nlm.nih.gov/10786800/.

Sudanese Armed Forces and Rapid Support Forces (May 20, 2023). *Agreement on a Short-Term Ceasefire and Humanitarian Arrangements*. Agreement text, including the Monitoring and Coordination Committee; hosted by the US Department of State. URL: https://2021-2025.state.gov/agreement-on-a-short-term-ceasefire-and-humanitarian-arrangements/ (visited on 11/07/2026).

Suez Canal Authority (2026). *Canal History*. URL: https://www.suezcanal.gov.eg/English/About/SuezCanal/Pages/CanalHistory.aspx (visited on 11/02/2026).

Taiwan Ministry of Economic Affairs, Energy Administration (2026). *2030 Energy Transition White Paper*. Chinese-language report; 2024 energy supply and import dependence, pp. 13–14. Taiwan Ministry of Economic Affairs. URL: https://ea01.moeaea.gov.tw/a0101/01/energywhitepaper/pdf/2030_energy_whitepaper.pdf (visited on 11/07/2026).

Taiwan Ministry of Finance, Department of Statistics (Jan. 16, 2025). *Taiwan's Exports to Mainland China and Hong Kong in 2024*. URL: https://service.mof.gov.tw/public/Data/statistic/bulletin/114/2025_01_CN&HK.pdf (visited on 11/07/2026).

Taiwan Semiconductor Manufacturing Company (2025). *Annual Report 2024*. TSMC. URL: https://investor.tsmc.com/static/annualReports/2024/english/index.html (visited on 11/07/2026).

Trefethen, Lloyd N. and Mark Embree (2005). *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton University Press.

UNCTAD (2024). *Bilateral FDI Statistics*. https://unctad.org/statistics. Foreign Direct Investment database.

United Nations Group of Experts on the Democratic Republic of the Congo (Dec. 27, 2024). *Midterm Report of the Group of Experts on the Democratic Republic of the Congo*. S/2024/969. Transmitted by letter of December 27, 2024; paragraphs 56–66 address Rubaya. United Nations Security Council. URL: https://digitallibrary.un.org/record/4071242/files/S_2024_969-EN.pdf (visited on 11/07/2026).

United Nations Office for the Coordination of Humanitarian Affairs (2026). *Humanitarian Access: Information Management Toolbox*. Institutional documentation of the Access Monitoring and Reporting Framework. URL: https://humanitarian.atlassian.net/wiki/spaces/imtoolbox/pages/1410072577/Humanitarian+Access (visited on 11/07/2026).

United Nations Secretary-General (Jan. 19, 2017). *Statement Attributable to the Spokesman for the Secretary-General on The Gambia*. URL: https://www.un.org/sg/en/content/sg/statements/2017-01-19/statement-attributable-the-spokesman-for-the-secretary-general-the-gambia (visited on 11/07/2026).

United Nations Security Council (Feb. 21, 2025). *Resolution 2773 (2025)*. S/RES/2773 (2025). URL: https://monusco.unmissions.org/sites/default/files/2773_en.pdf (visited on 11/07/2026).

United Nations Statistics Division (2026a). *UN Comtrade: Annual Totals vs Sum of Monthly Data*. URL: https://uncomtrade.org/docs/annual-totals-vs-sum-of-monthly-data/ (visited on 11/07/2026).

– (2026b). *UN Comtrade: Data Availability*. URL: https://uncomtrade.org/docs/data-availability/ (visited on 11/07/2026).

United Nations Treaty Collection (Apr. 18, 1951). *Treaty Instituting the European Coal and Steel Community*. Registration 3729; entry into force July 23, 1952. URL: https://treaties.un.org/Pages/showDetails.aspx?objid=0800000280144526 (visited on 11/02/2026).

United States Congress (1979). *Taiwan Relations Act*. Public Law 96-8; 22 U.S.C. 3301–3316, 2024 codification. URL: https://www.govinfo.gov/content/pkg/USCODE-2024-title22/html/USCODE-2024-title22-chap48.htm (visited on 11/07/2026).

United States Department of State (Jan. 23, 2022). *Senior State Department Officials on Posture of U.S. Embassy Kyiv*. URL: https://2021-2025.

state.gov/senior-state-department-officials-on-posture-of-u-s-embassy-kyiv/ (visited on 11/07/2026).

United States Holocaust Memorial Museum (2018). *Early Warning Project to Release Its 2018–2019 Risk Assessment: New Website and Refined Methodology*. Documents the expansion to state- and non-state-led mass killing and a two-year forecast horizon. URL: https://www.ushmm.org/genocide-prevention/blog/early-warning-project-to-release-its-2018-2019-risk-assessment-new-website (visited on 11/07/2026).

United States Senate Committee on Armed Services (Mar. 9, 2021). *Hearing to Receive Testimony on United States Indo-Pacific Command*. Stenographic transcript; testimony of Admiral Philip S. Davidson, pp. 47–48. URL: https://www.armed-services.senate.gov/imo/media/doc/21-10_03-09-2021.pdf (visited on 11/07/2026).

Uppsala Conflict Data Program (2026). *UCDP Dataset Download Center*. URL: https://ucdp.uu.se/downloads/ (visited on 11/07/2026).

US Department of State (Nov. 11, 1941a). *Documents Prepared in the Division of Far Eastern Affairs*. Foreign Relations of the United States, 1941, Volume IV, Document 431. URL: https://history.state.gov/historicaldocuments/frus1941v04/d431 (visited on 11/02/2026).

– (Nov. 28, 1941b). *The Secretary of State to the Ambassador in Japan*. Foreign Relations of the United States, 1941, Volume IV, Document 500. URL: https://history.state.gov/historicaldocuments/frus1941v04/d500 (visited on 11/02/2026).

– (Oct. 28, 1962a). *Summary Record of the Tenth Meeting of the Executive Committee of the National Security Council*. Foreign Relations of the United States, 1961–1963, Volume XI, Document 103. URL: https://history.state.gov/historicaldocuments/frus1961-63v11/d103 (visited on 11/02/2026).

– (Oct. 26, 1962b). *Telegram from the Embassy in the Soviet Union to the Department of State*. Foreign Relations of the United States, 1961–1963, Volume XI, Document 84; Khrushchev's letter. URL: https://history.state.gov/historicaldocuments/frus1961-63v11/d84 (visited on 11/02/2026).

– (2026). *The Cuban Missile Crisis, October 1962*. URL: https://history.state.gov/milestones/1961-1968/cuban-missile-crisis (visited on 11/02/2026).

VIEWS (2026a). *Frequently Asked Questions*. Operational target and horizon documentation as accessed November 7, 2026. URL: https://viewsforecasting.org/faq/ (visited on 11/07/2026).

VIEWS (2026b). *Model Directory*. Version history and geographic scope of ViEWS2018, ViEWS2020, and the fatalities model releases. URL: https://viewsforecasting.org/models/ (visited on 11/07/2026).

VIEWS Prediction Challenge Collaboration (June 28, 2024). *The 2023/24 VIEWS Prediction Challenge: Predicting the Number of Fatalities in Armed Conflict, with Uncertainty*. Preprint; Håvard Hegre and collaborators. URL: https://viewsforecasting.org/wp-content/uploads/Hegreetal2024_introduction.pdf (visited on 11/07/2026).

Voeten, Erik (2009). "Data and Analyses of Voting in the UN General Assembly". In: *Social Science Working Paper*.

Voeten, Erik, Anton Strezhnev, and Michael A. Bailey (2024). "United Nations General Assembly Voting Data". In: Dataset available at https://dataverse.harvard.edu/dataverse/UN.

Witten, Edward (1989). "Quantum Field Theory and the Jones Polynomial". In: *Communications in Mathematical Physics* 121.3, pp. 351–399. DOI: 10.1007/BF01217730.

World Bank (Nov. 2018). *The World Bank in India: November 2018*. Summary of A Glass Half Full: The Promise of Regional Trade in South Asia, edited by Sanjay Kathuria. World Bank, pp. 3–4. URL: https://documents1.worldbank.org/curated/en/768761544104016791/pdf/132734-NEWS-WB-Newsletter-November2018-LR.pdf (visited on 11/07/2026).

– (Mar. 13, 2025). *World Bank's South Sudan Economic Monitor Urges Swift and Sustained Reforms to Accelerate Economic Recovery and Inclusive Growth*. Release accompanying the seventh South Sudan Economic Monitor, A Pathway to Overcome the Crisis. URL: https://www.worldbank.org/en/news/press-release/2025/03/13/world-banks-afe-south-sudan-economic-monitor-urges-swift-and-sustained-reforms (visited on 11/07/2026).

– (2026). *World Development Indicators: DataBank*. URL: https://databank.worldbank.org/source/world-development-indicators (visited on 11/07/2026).

World Trade Organization (2026). *DS602: China—Anti-Dumping and Countervailing Duty Measures on Wine from Australia*. URL: https://www.wto.org/english/tratop_e/dispu_e/cases_e/ds602_e.htm (visited on 11/02/2026).

Index

Bold page numbers identify glossary definitions. Other references lead to substantive discussions and worked examples. Page ranges follow a continuous discussion; *see* directs to the preferred term.

Symbols

1992 Consensus, 162
72-day interval, **517**

A

Ablation, **519**
  isolating geometry, 436
Abraham Accords, 190, **511**
Abstention, **519**
  coverage evaluation, 417
ACLED, 11, 148, 430, **515**
Activation Energy, **501**
Actor identity, **519**
  joining records, 425
  splits and successors, 242
Adam–Gibbs Theory, **506**
ADIZ, **515**
Adjacency Matrix, **501**
Admissible uncertainty set, **517**
Africa, *see* Sudan; Horn of Africa; Great Lakes region; West Africa
Aggregation, **519**
  failure of closed dynamics, 271
AIS, **515**
Alert levels, **507**
  historical rule, 73
  uncertainty propagation, 373
Algiers Agreement, 248
Analytic completion, **519**
Angell, Norman, 7
Anocracy, **511**
Armenia, 132, 220
Arrhenius Behavior, **506**
As-of snapshot, **517**
ASML, 162, **515**
ATOP, **515**
Attractor, **506**
AUC, **501**
Australia, 224
Availability interval, **520**
  overlapping conditions, 253
Axis of Resistance, **511**
Azerbaijan, 132, 220

B

Backtesting, *see* Historical backtesting
Balakot crisis, 203
Base rate, **520**
  selected samples, 413
BASELINE, **508**
Basis Points, **514**
Bifurcation, **501**
Bilateral Trade, **514**
Black box, *see* Explanation
Bottleneck, *see* Capacity; Maximum flow / minimum cut; Spectral gap
Boundary condition, **520**
Branch precedence, **520**
Brier loss, **518**
  paired worked comparison, 435
Budapest Memorandum, 112, **517**
Burkina Faso, 274

C

Calibration, **518**
  subgroups and resolution, 415
Candidate space, 300, **520**
Capacity, **520**
  distribution of benefits, 394
Capital Flight, **514**
Capitalist Peace, **511**
Case-control sampling, **520**
  odds adjustment, 413
CAST, 11, 148, 430

Catastrophe Theory, **501**
Causal inference, 20, 40, 132, 147, 232, 261, 408, 418, 446, 451
  adjustment variables, 456
  diagnostics and displacement, 459
  sequential policies, 457
  time zero, 455
  warning and intervention, 451–460
Ceasefires
  implementation and repair, 245
Censoring, **521**
  forecast horizons, 417
  pending transactions, 281
Certificate, **518**
CFIUS, **515**
Cheeger Constant, **501**
Cheeger inequality, 98, 299, 313, 319, 332, 470, **502**
Chern–Simons functional, 296
Chern–Simons Invariant, **502**
China
  Taiwan relationship, *see* China–Taiwan
China–Taiwan, 161–181
  2024 assessment, 166
  2027 readiness horizon, 171
  changes to the assessment, 169
  evidence that tests the mechanism, 179
  history and evidence, 161
  managed diversification, 180
  monitoring indicators, 175
  qualified substitutes, 175
  quarantine and shipping coercion, 177
CHIPS Act, **515**
Chokepoint Effect, **511**
CIF/FOB Valuation, **514**
CIPS, **515**
Clock, information, **521**
Coercion, **521**
Coercive Diplomacy, **511**
Combined signal, 101, **508**
Commercial / Liberal Peace, **511**
Commercial peace
  arguments and evidence, 29
  Germany and Russia, 5
Common-cause failure, **521**
  substitutes, 395
Commutator, **521**
Competing risks, **518**
Confidence interval, **521**
Configuration Space, **502**
Confounding
  private information, 456
Congo, *see* Democratic Republic of the Congo; Great Lakes region
Conservation constraint, **521**
Conservation Law, **502**
Controlled Release Paradigm, **511**
Convolution, **522**
  response to timed inputs, 379
Copeland's Trade Expectations Theory, **511**
Copeland, Dale, 31
Counterexample, *see* Witness
Covariance
  shared observation errors, 342
Coverage, **522**
  delivery and unmet needs, 241
COW, **515**
Credit Default Swap, **514**
CRITICAL, **508**
Critical path, **522**
  administrative transition, 280
Critical slowing down, **507**
  time interpretation, 384
Cuban Missile Crisis, 72, 462, **517**
Curl, **502**
Curvature, **502**
Curvature Components, **508**
Cusp Catastrophe, **502**
Cycle space, **522**
  observation coverage, 341

D

Data leakage, *see* Information leakage
Dataset versions, *see* Release vintage
Davidson window, 171, **511**
Davidson, Philip, 163
Davis framework
  central equations, 60
  mathematical foundations, 76
  research origin, 54
Davis Law, **508**
De-risking vs. Decoupling, **511**
Decision margin, **522**

uncertain classification, 373
DECOUPLING, **508**
Degree Matrix, **502**
Degree normalization, **522**
Democratic Peace Theory, **511**
Democratic Republic of the Congo, 261
Denominator, **522**
Detailed balance, 349
time-direction asymmetry, 350
Deterrence, **512**
Differential Forms, **503**
Directional reading, 65, **523**
DISCONNECTION Mode, **508**
Dissipation, **503**
Distribution shift, **523**
transfer failures, 406
DIVERGENCE, **508**
Divergence Index, **503**
DOTS, **515**
Dual certificate, 357, **523**
Dyad, **508**
Dynamics
downstream accumulation, 382
explanatory chain, 385
input and output maps, 379
repeated pulses, 382
time and directional coupling, 375–386

E

Early warning, *see* Forecasting; Forecast evaluation; Alert levels
ECOWAS, 274
membership and practical access, 279
Effective connection, **523**
Effective resistance, **518**
Eigenvalue / Eigenvector, **503**
ELEVATED, **508**
Eligible population, **523**
Energy Dependency, **514**
Energy landscape, 325, **507**
Eritrea, 132, 221, 248
Escalation Ladder, **512**
Ethiopia, 132, 221, 248
Euler Equations, **503**
Euromaidan, **517**
EUV, **515**
Event time, **523**
interval uncertainty, 270
Evidence
bounds, probabilities, and decisions, 360
EWP, **515**
Exchangeability and positivity, **518**
Explanation
faithfulness and empirical validity, 439
right outcome and wrong mechanism, 407
Exterior Derivative, **503**

F

F1 score, **503**
Fairness
minimum service coverage, 394
Farrell, Henry, 37, 51, 213
FDI, *see* Foreign direct investment
Feasible exits
capacity and time, 387–398
executed departure, 276
monotonicity, 300
shared prerequisites, 283
Feasible set, **523**
infeasibility and repair, 397
Feedback-Coupled Flow Systems, **508**
Fiedler vector, 80, 313, **503**
cuts and modes, 313
Forecast evaluation
probability and horizons, 409–419
spatial aggregation, 434
Forecast horizon
nested probabilities, 416
Forecasting
historical failures, 2–28
historical surprises, 13
Foreign direct investment, **514**
Four quadrants, 67, **509**
Frozen Conflict, **512**
Frozen Peace, **509**
Frozen vs. Glassy Systems, **509**

G

Gambia, The, 276
Gamma, *see* Trichotomy parameter

Gauge transformations, **523**
  implementation checks, 343
  measurement and invariance, 336–345
  node units, 337
Gaussian Process, **503**
GDP, **514**
GERD, **515**
Germany
  trade and restraint, 5
GIGI, 352, **524**
Glass transition, **507**
  physical definition, 322
  reconfiguration, 329
Gradient component, **524**
Graphs
  choice of network, 104
  effective connections, 370
  omitted edges and edge meaning, 371
Great Lakes region, 261–273
  Rubaya, 262
  supply-chain attribution, 267
  testing the access mechanism, 272
GSDB, **515**
Guarantees
  staging and exposure, 282

H

Hamas, 141, 188
Hazard, **524**
  cumulative probability, 416
Heat kernel, 315, **503**
Helicity, **504**
  capacity equation, 60
  ordered actor roles, 65
  symbols and versions, *see* Notation
Helicity Correction, **509**
Helicity project
  research records, 542
Helmholtz Vortex Theorems, **504**
Hezbollah, 183
Historical backtesting, 136–150
  47-conflict research, 132, 531
  detections, 141
  false alerts, 146
  missed events, 144
  reproduction record, 531
  sample and protocol, 136
  selected historical roster, 500
  sensitivity analysis, 147
  versioned estimates, 534
Hodge decomposition, 338, **524**
  least-squares fit, 338
  residual covariance, 342
  weighted fit, 339
Holomorphic, **524**
Horn of Africa, 248–260
  evaluation design, 258
  implementation clocks, 253
  November 2022 delivery, 249
Hostility Index, **509**
Humanitarian access
  Sudan, 236
  Tigray, 249
Hysteresis, **524**
  alert rules, 405

I

Identification, **524**
  unobserved response, 382
IEA, **515**
IMF, **515**
Implementation, 420–429
Imputation, **525**
Incidence matrix, **525**
India, *see* India–Pakistan
India–Pakistan, 200–212
Indus Waters Treaty, 201
Infeasibility
  contradictory evidence, 358
Information
  value of an additional observation, 373
Information clock
  release and acquisition, 426
Information leakage, **525**
Interference, **518**
Invariance
  positive and negative controls, 404
Inverse-probability weighting, 456, **525**
Inviscid Flow, **507**
Iran, *see* Israel–Iran
IRGC, **515**
Israel–Iran, 182–199
Issuance record, **525**

replay and version history, 427

J

Jammeh, Yahya, 275
Japan, 14, 150, 151, 159, 167
JCPOA, **516**
Jeddah agreements, 236
Joint compatibility, 360

K

Kargil conflict, 200
Kashmir, 25, 203
Kelvin circulation theorem, 287, **504**
Kennedy, John F., 156
Khrushchev, Nikita, 156

L

L4, *see* Alert levels
Laplacian, **504**
  construction, 311
Laplacian, signed, **525**
Lead time, **525**
  classifier-specific means, 532
  historical mean, 500
Liberal Peace, **509**
Logarithmic loss, **526**

M

M23, 261
  mineral receipts, 262
Machine learning
  conflict forecasting, 39
Mali, 274
Managed Competition, **512**
Manifold, **504**
Markov chain, 93, 349
Mass gap
  classifier comparison, 532
Master Equation, **509**
Math symbols, *see* Notation
Maximum flow / minimum cut, **526**
Measurement
  blind spots, 19
  framework components, 62
  relabeling test, 65
  unequal precision, 339
Measurement model, **526**
Mediation
  observation and sequencing, 256
MID, **516**
Minerals
  due diligence, 267
Minsk agreements, 70, **517**
Missing data
  coverage sensitivity, 269
  lost consistency tests, 341
  reason codes and imputation, 427
Mixing time, **504**
  diffusion, 315
Mode-coupling theory, 327, **507**
Model comparison, 430–441
  common population, 434
  developmental annotations, 534
  historical comparisons, 148
  Psychohistory paper baselines, 532
  scope of conclusions, 439
  tuning and preregistration, 438
Model failure, 400–408
  duplicated flows, 425
  numerical artifacts, 384
Model version, **526**
Montesquieu, 29

N

Nairobi declaration, 249
NATO, 3, 101, 119, 467, **516**
Navier–Stokes Equations, **504**
NCA, **516**
Near Abroad, **512**
Negative control, 459, **526**
Negative Helicity, **509**
Nevanlinna–Pick matrix, **526**
  compatibility test, 355
Newman, Abraham, 37, 51, 213
Niger, 274
Nominal plan, **526**
Nonanticipating decision, **518**
Nonnormal operator, **518**
Nord Stream, 7, 34, 110
Nord Stream 1 / 2, **517**
Normalized Laplacian, **504**
  nonmonotone gap, 369

Notation
reading guide, 548
Nuclear deterrence, **512**
Numerical integration
Euler instability, 384

O

One-China formulations, **512**
Operation Rising Lion, **517**
Operation True Promise, **517**
Operator norm, **527**
Orderly Decoupling, **512**
Ordinary differential equations, *see* Dynamics; Numerical integration
Out-of-sample evaluation, **527**
temporal and geographic splits, 437
Outcome
release maturity, 417

P

Pakistan, *see* India–Pakistan
Panopticon Effect, **512**
PARALLAX, 353, **527**
Partial identification, 352–362, **527**
conditional bounds, 353
dependent queries, 360
supply requirements, 359
Pearl Harbor, 14
Phase Transition, **504**
Phase transitions
warning indicators, 41
PITF, **516**
Planning horizon
end effects, 393
Polarization Index, **505**
Political choice, 48–52
Positive Semi-Definite, **505**
Positivity
overlap, 456
Potential outcome, **518**
Precision and recall, **527**
Prerequisite, **527**
Pretoria agreement, 248
implementation, 249
Probability
dependent prerequisites, 242
Propensity score, 456, **527**
Provenance, **527**
transformation chain, 424
Proxy Architecture, **512**
Proxy Degradation, **513**
Pseudoinverse, **518**
PSYCHOHISTORY, **509**
Pulwama–Balakot Crisis, **517**
Put Skew, **514**

Q

Quarantine, Taiwan scenario, **528**

R

Recourse, **528**
information timing, 391
Recoverability
backlog of breaches, 257
unresolved breaches, 245
Recovery time
event spacing, 382
Release vintage, **528**
Rubaya estimates, 262
Reporting process, **519**
partial identification, 269
Reproducibility
numerical companions, 545
release checks, 428
Reserves
partial interruption, 177
usable and nominal stocks, 396
Resolution, forecast, **528**
Revanchism, **513**
Revenue
delivery to accessible cash, 244
Robust feasibility, **528**
credibility interval, 282
ROC Curve, **505**
Rough paths, **528**
Rubaya, 262
Russia
Ukraine relationship, *see* Russia–Ukraine
Russia–Ukraine, 110–134
72-day interval, 124
alternative paths, 129

dependency profile, 111
helicity trajectory, 116
late-2021 reconstruction, 118
Rwanda, 261

S

SAARC, **516**
Sanctions, **514**
Sanctions Paradox, **513**
Scenario, **528**
Schur complement, **528**
boundary reduction, 370
Schwarz lemma, 355
Security Posture, **509**
Selection bias, **529**
abandoned transactions, 281
Semiconductors
serial production bottlenecks, 175
Sequence
event order, 346–351
Sequential exchangeability, 457, **529**
Shadow price, **529**
changing bottleneck, 393
Signed area
inspection and delivery, 347
Signed graphs
balanced and unbalanced triangles, 254
Silicon shield, 167, 177, **513**
Simplicial Complex, **505**
Simplicial Helicity, **505**
SINGULARITY, **510**
Singularity Warning, **510**
Small Yard, High Fence, **513**
South Sudan, 238
export revenue and payment timing, 244
Spatial precision, **529**
bounds on location, 270
Spectral Decomposition, **505**
Spectral gap, **505**
definition and interpretation, 93
three-node example, 369
threshold calibration, 98
uncertainty and repair, 363–374
SPFS, **516**
Stability–Instability Paradox, **513**
Strategic Ambiguity, **513**
Structural balance, **529**
Structural vs. Eventive Forecasting, **510**
Sudan, 234–247
authority and fragmentation, 242
civilian access, 241
Jeddah implementation, 236
SWIFT, **516**

T

Taiwan, *see* China–Taiwan
Taiwan Relations Act, 162
TAPI, **516**
Target trial, 455, **529**
Temporal network, **529**
Terminal condition, **530**
sustainability, 393
TIES, **516**
Time-expanded network, **519**
Tolerance
measurement and computation, 358
symbol, *see* Notation
Tolerance Budget, **510**
Topological Invariant, **505**
Torsion, **505**
Torsion Components, **510**
Trade
network representation, 318
Trade Asymmetry, **514**
Transfer test, **530**
Transient amplification, **530**
Transit Fees, **514**
Transition
deadlines, 280
Trap, **510**
integration and hostility, 67
Trap Warfare, **513**
Trichotomy parameter, **505**
candidate space and feasible exits, 300
derivation and regimes, 70
TSMC, 37, 162, 176, **516**
Turbulence, **507**

U

UCDP, 23, 65, 265, **516**
Uganda, 261

Ukraine, *see* Russia–Ukraine
UNCTAD, **516**
UNGA, **516**
Units
  rate and calendar time, 384
Usable substitute, **530**

V

Value of information
  precision and timing, 359
Velocity Field, **506**
Verification, **530**
  information and exposure, 256
Viability, **530**
  backward recursion, 391
VIEWS, 10, 40, 148, 420, 430, **516**
Viscosity, **506**
Vogel–Fulcher–Tammann, **507**
Void, **510**
Vorticity, **506**

W

Wandel durch Handel, **513**
WARNING, **510**
WDI, **516**
Weaponized interdependence, **513**
  mechanisms, 36
Wedge Product, **506**
Weighted norm, **530**
  component constraints, 406
West Africa, 274–285
  document recognition, 279
  Gambian transfer, 276
  institutional functions, 284
Window of Opportunity, **510**
Witness, **519**
  attainment and optimality, 357
  signed zero mode, 254
  supply allocations, 352

Z

Z-score, **506**
  classifier comparison, 532
Zeitenwende, **513**



About the Author

Bee Rosa Davis is an applied mathematics researcher and engineer whose work spans differential geometry, spectral theory, gauge theory, and machine learning. She holds an MS in Digital Forensics from Brown University and a BA in Logic from Morehouse College, where she was a Martin Luther King Jr. Scholar.

Her career was built on finding structure in complex systems before anyone told her the mathematics had a name. At NASA Johnson Space Center, she designed telemetry pipelines processing 2.3 terabytes per day. At IBM X-Force Red, she served as Principal Adversarial Intelligence Engineer, building adversarial analytics platforms and discovering four Windows kernel zero-day vulnerabilities using geometric methods she developed herself. At the NSA and Naval Intelligence, she built anomaly detection systems that synthesized signals across compromised devices, submarine telemetry, and cross-agency networks. At Pandora Radio, she was a founding engineer who scaled a five-person team to one hundred and grew ad revenue 750% through machine learning systems she architected from scratch.

The throughline was always geometry. Telemetry pipelines are manifolds. Vulnerability surfaces are curvature fields on code spaces. Financial crashes are topological phase transitions. Conflict between nations is helicity measured on geopolitical configuration spaces. The mathematics in this book did not emerge from a single insight. It emerged from three decades of watching systems organize, deform, and break.

The Russia–Ukraine reconstruction in this book assigns a highest-level warning to December 14, 2021, seventy-two days before the invasion of February 24, 2022. Davis developed the geopolitical application after the invasion, working backward through the historical record to investigate how dependence persisted while trust collapsed. The recorded PSYCHOHISTORY evaluation reports an F1 score of 0.73 across forty-seven historical conflicts. The system is the subject of a pending United States patent (Application No. 63/972,955).

Her broader research program has produced over thirty-five papers and twenty-six patent applications spanning geometric foundations of transformer reasoning, quantum-resistant cryptographic hashing, multi-cancer genomic detection, viral surveillance, and antimicrobial resistance tracking. She is the author of *The Geometry of Sameness*, which debuted as the number-one new release in Vector Analysis Mathematics on Amazon, and *Motion to Be Seen: One Black Trans Woman's Fight for Algorithmic Justice*, a number-one Amazon bestseller. Her published books include works of mathematics, science fiction, and memoir.

Bee Rosa Davis is a Black transgender woman. She is the Overall Mother of the House of Diwa. She taught herself differential geometry, spectral theory, category theory, and gauge theory from textbooks and papers because the questions would not leave her alone.

The mathematics in this book was developed independently. It was developed because the patterns were there, visible to anyone willing to look carefully enough, and someone needed to write them down.

The geometry does not ask who you are.
It asks only whether you were paying attention.