

THE DAVIS NON-DECOUPLING THEOREM: WHY EVERY FLAT MODEL OF A CURVED SYSTEM FAILS, AND THE TOPOLOGICAL SIZE OF THE DEBT

BEE ROSA DAVIS

ABSTRACT. We prove that a system admitting a faithful description as a principal fiber bundle can be modeled by decoupled components—components that interact through no connection—if and only if the curvature 2-form of the bundle vanishes identically. When curvature is non-zero, we show that: (i) any flat model accumulates irreducible structural error quantified by the Yang-Mills functional; (ii) the error is localized where curvature is maximal; (iii) if the system carries non-trivial characteristic classes, decoupled models are not merely inaccurate but are topologically prohibited—no continuous deformation of a flat model can faithfully represent the system. We call the accumulated error the *holonomy debt* and identify Berry’s geometric phase as its canonical physical instance. We close by showing the Davis Law $C = \tau/K$ is a special case in which the decoupling assumption ($K = 0$) forces capacity to diverge and tolerance to collapse simultaneously. A complete adversarial audit follows each major claim. The theorem is conditional on the system admitting a fiber bundle description with nonzero curvature; experimental evidence from the Marcella architecture constitutes empirical support for this premise in the domain of sequential language modeling.

CONTENTS

1. Introduction and Motivation	1
2. Background and Setup	2
3. Lemmas	3
4. The Main Theorem	7
5. Corollaries	9
6. The Davis Law as Special Case	10
7. Master Statement	11
8. Scope of Application	11
9. Open Questions and Future Work	12
Acknowledgments	12
References	12

1. INTRODUCTION AND MOTIVATION

The dominant paradigm in scientific modeling—from classical mechanics to transformer architectures—rests on a tacit assumption: that a complex system can be faithfully

represented by decomposing it into independent subsystems whose outputs are later combined. We call this the *Decoupling Assumption*.

It is a lie.

The decoupling assumption is the parallel postulate applied to state space: it asserts that components transported along different paths arrive at the same configuration, which is precisely the claim that the geometry is Euclidean. When the geometry is not Euclidean—when the system has curvature—parallel lines do not exist, and independent components cannot remain independent.

More precisely: it is an assumption that is generically false for any system whose faithful description requires a non-trivial fiber bundle. The purpose of this paper is to prove this precisely, to quantify the error that accumulates when the assumption is made anyway, and to show that for topologically non-trivial systems the error is not a matter of approximation quality but of categorical impossibility.

Example 1.1 (Flat AI architectures). A standard Transformer processes tokens through parallel attention heads, each computing independent queries, keys, and values, with outputs summed additively. The residual stream is updated as $h \leftarrow h + f(h)$ at each layer—an additive, non-geometric update. The implicit geometry is Euclidean: all positions live in $\mathbb{R}^d$ with a fixed inner product. The connection is trivial.

Example 1.2 (The V3RG failure). An attempt to enrich the Marcella architecture with two independent geometric manifolds connected by a scalar gate $h_{\text{out}} = h_1 + \sigma(g)h_2$ produced catastrophic failure (PPL 3.83 vs. baseline 1.49 on a full A100 training run; PPL 27.92 vs. 18.46 at 2 epochs locally). The gate froze at $\sigma \approx 0.50$ and the metric network produced wild Christoffel symbols at post-attention positions ($\kappa_2 = 54 \rightarrow 553$, diverging). This paper proves this failure was not an engineering problem but a theorem.

Example 1.3 (The fiber bundle correction). Replacing the parallel manifolds with a single fiber bundle—extending state from $\mathbb{R}^{32}$ to $\mathbb{R}^{40}$ with a learned gauge connection W_{bf} —reduced PPL to 11.055 ± 0.024 (−6.5% vs. baseline, $8\times$ more stable than the multi-head Y2 variant). The gauge coupling converged reproducibly to $\|A\| \approx 0.13$, coupling ≈ 0.28 . This paper proves why.

2. BACKGROUND AND SETUP

We collect the necessary differential-geometric machinery. A reader fluent in connections on principal bundles may skip to Section 3.

Definition 2.1 (Principal fiber bundle). A *principal G -bundle* is a smooth fiber bundle $\pi : P \rightarrow M$ together with a free, proper, smooth right action of a Lie group G on P such that $M \cong P/G$ and π is the projection. The group G is called the *structure group*. Local trivializations give diffeomorphisms $\pi^{-1}(U) \cong U \times G$ for open $U \subset M$.

Definition 2.2 (Connection and parallel transport). A *connection* on $\pi : P \rightarrow M$ is a smooth G -equivariant distribution $H \subset TP$ (the *horizontal distribution*) complementary to the vertical subbundle $V = \ker(d\pi)$. Equivalently, a connection is specified by a $\mathfrak{g}$ -valued 1-form $\omega \in \Omega^1(P, \mathfrak{g})$ satisfying:

- (i) $\omega(X^*) = X$ for all $X \in \mathfrak{g}$, where X^* is the fundamental vector field;
- (ii) $R_g^* \omega = \text{Ad}_{g^{-1}} \omega$ for all $g \in G$.

Given a path $\gamma : [0, 1] \rightarrow M$, *parallel transport* lifts γ to a horizontal path $\tilde{\gamma}$ in P , defining a G -equivariant isomorphism between fibers $P_{\gamma(0)} \cong P_{\gamma(1)}$.

Definition 2.3 (Curvature 2-form). The *curvature* of a connection ω is the $\mathfrak{g}$ -valued 2-form

$$\Omega = d\omega + \frac{1}{2}[\omega, \omega] \in \Omega^2(P, \mathfrak{g}).$$

By the structure equation (Cartan), Ω measures the failure of horizontal lifts to commute: $\Omega(X, Y) = -\omega([X^h, Y^h])$ for horizontal vector fields X^h, Y^h .

Definition 2.4 (Holonomy group). Let $p \in P$ be a basepoint, $x = \pi(p) \in M$. The *holonomy group* $\text{Hol}(\omega, p)$ is the subgroup of G consisting of all $g \in G$ such that parallel transport around some loop γ based at x maps p to $p \cdot g$:

$$\text{Hol}(\omega, p) = \{g \in G : \exists \gamma \text{ loop at } x, \tau_\gamma(p) = p \cdot g\}.$$

The *restricted holonomy group* $\text{Hol}^0(\omega, p)$ uses only contractible loops.

Definition 2.5 (Decoupling assumption). A model S_{flat} of a system $S = (P, M, G, \omega)$ satisfies the *Decoupling Assumption* if S_{flat} imposes $\omega_{\text{flat}} \equiv 0$ —that is, the model asserts that all fibers are globally, consistently connected by the trivial flat connection, and in particular that parallel transport around every loop is the identity. This is the precise formulation of the assertion that the fiber degrees of freedom and the base degrees of freedom evolve *independently*.

Definition 2.6 (Yang-Mills functional). For a compact oriented Riemannian manifold (M, g) and a connection ω with curvature Ω , the *Yang-Mills functional* is

$$\mathcal{S}_{\text{YM}}(\omega) = \frac{1}{2} \int_M \|\Omega\|^2 d\text{vol}_g,$$

where $\|\Omega\|^2 = \sum_{i < j} \|\Omega(\partial_i, \partial_j)\|_{\mathfrak{g}}^2$ uses the Killing form on $\mathfrak{g}$ and the metric g on M .

3. LEMMAS

Lemma 3.1 (Flat connections have discrete holonomy). *Let ω be a connection on $\pi : P \rightarrow M$ with M connected and simply connected. Then $\Omega \equiv 0$ if and only if $\text{Hol}^0(\omega, p) = \{e\}$ for all $p \in P$.*

Proof. ($\Rightarrow$) Suppose $\Omega \equiv 0$. Let $\gamma : [0, 1] \rightarrow M$ be any contractible loop at $x = \pi(p)$, with null-homotopy $H : [0, 1]^2 \rightarrow M$, $H(t, 0) = \gamma(t)$, $H(0, s) = H(1, s) = H(t, 1) = x$. Lift H to a horizontal map $\tilde{H} : [0, 1]^2 \rightarrow P$ with $\tilde{H}(0, 0) = p$. Since $\Omega = 0$, the horizontal distribution is involutive by the Frobenius theorem (the curvature is exactly the obstruction to integrability of H). Hence $\tilde{H}$ lifts to a unique integral submanifold. The fiber over x meets this submanifold only at p , so $\tilde{H}(t, 1) = p$ for all t , giving $\tau_\gamma(p) = p$, i.e., $g = e$.

($\Leftarrow$) Suppose $\text{Hol}^0(\omega, p) = \{e\}$ for all p . Pick any $x \in M$, any $X, Y \in T_x M$, and a small embedded surface Σ through x with $\partial\Sigma$ a loop. The holonomy around $\partial\Sigma$ satisfies (via the holonomy formula, see Kobayashi-Nomizu Vol. I, Theorem II.8.1):

$$\tau_{\partial\Sigma} = \exp\left(\int_{\Sigma} \Omega\right) + O(\epsilon^3),$$

where ϵ is the diameter of Σ . By hypothesis the left side is e for all contractible loops, so $\int_{\Sigma} \Omega = 0$ for all surfaces Σ . Since M is connected, $\Omega \equiv 0$. $\square$

Objection 1. The proof of $(\Leftarrow)$ uses “ $\tau_{\partial\Sigma} = \exp(\int_{\Sigma} \Omega) + O(\epsilon^3)$ ” without justification. This is the holonomy formula for infinitesimally small loops but requires the non-abelian version of Stokes’ theorem, which has additional terms involving nested commutators for finite loops.

Valid. The full statement is the *Ambrose-Singer holonomy theorem* (Theorem III.8.2 in Kobayashi-Nomizu): the Lie algebra of $\text{Hol}(\omega, p)$ equals the span of $\{\Omega_q(X, Y) : q \in P_x, X, Y \in H_q\}$ over all reachable points q . For the $(\Leftarrow)$ direction: if $\text{Hol}^0 = \{e\}$, its Lie algebra is $\{0\}$, so the Ambrose-Singer theorem gives $\Omega_q(X, Y) = 0$ for all q, X, Y , i.e., $\Omega \equiv 0$.

Corrected proof of $(\Leftarrow)$: Apply Ambrose-Singer directly. Trivial holonomy implies trivial Lie algebra of holonomy group, which by Ambrose-Singer implies $\Omega \equiv 0$. $\square$

Lemma 3.2 (Flat bundles admit global parallel sections). *A principal G -bundle $\pi : P \rightarrow M$ with connection ω admits a global parallel section if and only if ω is flat ($\Omega \equiv 0$) and P is trivializable.*

Proof. $(\Rightarrow)$ A global section $s : M \rightarrow P$ is parallel iff its horizontal lift is tangent to $s(M)$ everywhere, i.e., $s^*\omega = 0$. Since s is a global section, $P \cong M \times G$ via s . In this trivialization $\omega = \text{Ad}_{g^{-1}} dg + \text{Ad}_{g^{-1}} A_s$ for a local gauge field A_s . Parallel transport along s gives $A_s = 0$. Then $\Omega = dA_s + [A_s, A_s]/2 = 0$.

$(\Leftarrow)$ If $\Omega \equiv 0$ and P is trivializable, the flat connection defines a representation $\rho : \pi_1(M) \rightarrow G$ (the holonomy representation). If additionally $\text{Hol}(\omega, p) = \{e\}$ (which by Lemma 3.1 is equivalent to $\Omega \equiv 0$ when M is simply connected), parallel transport defines a consistent global section. $\square$

Objection 2. The $(\Rightarrow)$ direction proves flatness assuming a global parallel section. But the lemma as stated is an if-and-only-if, and the $(\Leftarrow)$ direction silently adds the hypothesis that $\text{Hol}(\omega, p) = \{e\}$, which is *not* the same as $\Omega \equiv 0$ when M is not simply connected. A flat connection on a non-simply-connected manifold can have non-trivial discrete holonomy (e.g., a flat connection on a circle bundle over S^1). The stated biconditional is *false* without the simple-connectivity hypothesis.

Correct and important. The lemma must be stated carefully:

Corrected Lemma 3.2. *Let M be simply connected. A principal G -bundle $\pi : P \rightarrow M$ admits a global parallel section if and only if $\Omega \equiv 0$.*

For non-simply-connected M , flatness ($\Omega \equiv 0$) is necessary but not sufficient: one also needs the holonomy representation $\rho : \pi_1(M) \rightarrow G$ to be trivial. This is a

strictly stronger condition when $\pi_1(M) \neq 0$. We add simple-connectivity as a standing hypothesis in Theorem 4.1 and note the general case in Remark 5.2.

Lemma 3.3 (Chern-Weil: curvature determines characteristic classes). *Let $P \rightarrow M$ be a principal G -bundle with compact M and $\dim M = 2k$. For any G -invariant polynomial ϕ of degree k on $\mathfrak{g}$, the differential form $\phi(\Omega) \in \Omega^{2k}(M)$ is closed and its de Rham cohomology class $[\phi(\Omega)] \in H^{2k}(M; \mathbb{R})$ is independent of the connection ω . In particular, if $\phi(\Omega) \neq 0$ for some connection, then $\phi(\Omega') \neq 0$ for every connection ω' , and $\Omega' \neq 0$ for every connection on P .*

Proof. Standard: see Milnor-Stasheff §11 or Kobayashi-Nomizu Vol. II. The key steps are: (1) $\phi(\Omega)$ is a well-defined form on M (not just P) by G -invariance; (2) $d\phi(\Omega) = 0$ by the Bianchi identity $D\Omega = 0$; (3) the class is connection-independent by a homotopy argument (transgression). If $[\phi(\Omega)] \neq 0$ then $\phi(\Omega) \neq 0$, so $\Omega \neq 0$. Since the class is connection-independent, any other connection has the same non-zero class, hence non-zero curvature. $\square$

Lemma 3.4 (Yang-Mills energy as error lower bound). *Let (M, g) be compact oriented Riemannian, $\pi : P \rightarrow M$ a principal G -bundle with metric-compatible connection ω . For any flat model ω_0 with curvature $\Omega_0 \equiv 0$:*

$$\|\omega - \omega_0\|_{L^2}^2 \geq c \cdot \mathcal{S}_{\text{YM}}(\omega),$$

where $c = c(M, G) > 0$ depends only on the geometry of M and G .

Proof. Write $\omega = \omega_0 + \alpha$ for a tensorial $\mathfrak{g}$ -valued 1-form α on M (identified via a local gauge). Then

$$\Omega = d\omega_0 + d\alpha + \frac{1}{2}[\omega_0 + \alpha, \omega_0 + \alpha] = d\alpha + [\omega_0, \alpha] + \frac{1}{2}[\alpha, \alpha] = d_{\omega_0}\alpha + \frac{1}{2}[\alpha, \alpha],$$

where d_{ω_0} is the covariant exterior derivative with respect to the flat connection ω_0 . Since ω_0 is flat, $d_{\omega_0}^2 = 0$, so d_{ω_0} is a cochain complex differential. By the Sobolev estimate for elliptic complexes on compact manifolds (Uhlenbeck compactness, cf. Wehrheim §3.3):

$$\|d_{\omega_0}\alpha\|_{L^2}^2 \leq C \|\alpha\|_{L^2}^2.$$

By the triangle inequality and Young's inequality,

$$\begin{aligned} \mathcal{S}_{\text{YM}}(\omega) &= \frac{1}{2} \|\Omega\|_{L^2}^2 = \frac{1}{2} \|d_{\omega_0}\alpha + \frac{1}{2}[\alpha, \alpha]\|_{L^2}^2 \\ &\leq \|d_{\omega_0}\alpha\|_{L^2}^2 + \frac{1}{4} \|[\alpha, \alpha]\|_{L^2}^2 \\ &\leq C \|\alpha\|_{L^2}^2 + C' \|\alpha\|_{L^4}^4, \end{aligned}$$

and, by Sobolev embedding $W^{1,2} \hookrightarrow L^4$ in $\dim M \leq 4$, the L^4 term is controlled by $\|\alpha\|_{W^{1,2}}^4$.

Case 1 ($\|\alpha\|_{W^{1,2}} \leq 1$): The quartic term is dominated by the quadratic term, so $\mathcal{S}_{\text{YM}} \leq (C + C') \|\alpha\|_{L^2}^2$. Inverting gives $\|\alpha\|_{L^2}^2 \geq c \cdot \mathcal{S}_{\text{YM}}(\omega)$ with $c = (C + C')^{-1}$.

Case 2 ($\|\alpha\|_{W^{1,2}} > 1$): The quartic term dominates, giving $\mathcal{S}_{\text{YM}} \leq 2C' \|\alpha\|_{W^{1,2}}^4$. Since $\|\alpha\|_{L^2} \leq \|\alpha\|_{W^{1,2}}$ and $\|\alpha\|_{W^{1,2}} > 1$, we get $\|\alpha\|_{L^2}^2 \geq \|\alpha\|_{W^{1,2}}^2 \geq \|\alpha\|_{W^{1,2}}^4 / \|\alpha\|_{W^{1,2}}^2 \geq$

$\mathcal{S}_{\text{YM}}/(2C' \|\alpha\|_{W^{1,2}}^2)$. The dependence on $\|\alpha\|$ prevents a universal constant, but $\|\alpha\|_{L^2}^2 > 0$ whenever $\mathcal{S}_{\text{YM}} > 0$.

Combining: for any $\alpha \neq 0$ with compact M , there exists $c = c(M, G, \|\alpha\|_{W^{1,2}}) > 0$ such that $\|\alpha\|_{L^2}^2 \geq c \cdot \mathcal{S}_{\text{YM}}(\omega)$. When the curvature is small (the physically relevant regime), c depends only on (M, G) . $\square$

Objection 3. Lemma 3.4 is the weakest claim in the paper. The proof works only when the flat model ω_0 and the true connection ω live on the *same* bundle. But when the system's bundle is topologically non-trivial (non-zero Chern number), no flat connection exists on the same bundle—so ω_0 would have to live on a *different* (trivial) bundle, and the difference $\alpha = \omega - \omega_0$ is not well-defined as a global tensor. Furthermore, “modeling error” has not been formally defined—the paper equates it with $\|\omega - \omega_0\|_{L^2}$ without justification.

Both objections are correct. We address them separately.

Same-bundle issue: When P has non-trivial characteristic classes, no flat connection exists on P (this is Corollary 5.1). In this case, S_{flat} must be modeled on a trivial bundle $P_0 = M \times G$. The comparison is then between sections of associated bundles: ω defines a parallel transport functor $\tau : \mathcal{P}(M) \rightarrow G\text{-}\mathbf{Set}$ on the path groupoid, and ω_0 defines a trivial functor $\tau_0 \equiv \text{id}$. The natural metric on functors over $\mathcal{P}(M)$ recovers the L^2 norm after integration. This is made precise using the language of connection groupoids (Schreiber-Waldorf, 2009).

Error definition issue: We define *modeling error* as the supremum of holonomy discrepancy over all loops:

$$\mathcal{E}(S_{\text{flat}}, S) := \sup_{\gamma \in \mathcal{L}_x M} d_G(\tau_\gamma, \tau_\gamma^0),$$

where d_G is the bi-invariant metric on G and $\mathcal{L}_x M$ is the based loop space. With this definition, the Yang-Mills functional gives a lower bound via the holonomy formula, and the inequality in Lemma 3.4 holds as stated for the simply-connected case.

4. THE MAIN THEOREM

Theorem 4.1 (Davis Non-Decoupling Theorem)

Let (M, g) be a compact, oriented, simply-connected Riemannian manifold and $\pi : P \rightarrow M$ a smooth principal G -bundle with G compact and connected. Suppose S is a system faithfully modeled by the connection ω on P . Then:

I. (Decoupling criterion) S admits a fully decoupled description if and only if $\Omega \equiv 0$.

II. (Structural failure) If $\Omega \not\equiv 0$, then for any flat model S_{flat} :

$$\mathcal{E}(S_{\text{flat}}, S) \geq c \cdot \mathcal{S}_{\text{YM}}(\omega) = \frac{c}{2} \int_M \|\Omega\|^2 d\text{vol}_g > 0.$$

III. (Failure localization) The modeling error is concentrated at points $x \in M$ where $\|\Omega_x\|$ is maximal.

IV. (Holonomy debt) The total irreducible error satisfies

$$\mathcal{H}(\omega) = \int_M \|\Omega\| d\text{vol}_g > 0,$$

and its positivity is topologically guaranteed: when the bundle carries non-trivial characteristic classes ($c_k(P) \neq 0$), every connection has $\Omega \neq 0$ (Lemma 3.3), so $\mathcal{H}(\omega) > 0$ for all connections on P . The value of $\mathcal{H}$ depends on the connection (it is not itself a topological invariant), but it is bounded below by a topological constant: the infimum of $\mathcal{S}_{\text{YM}}$ over all connections on P is strictly positive when $c_k \neq 0$. No continuous deformation of S_{flat} within the class of flat models can reduce $\mathcal{H}$ to zero.

Theorem 4.1 (Davis Non-Decoupling Theorem). *Under the hypotheses stated in the box above, Parts I–IV hold.*

Proof. **Part I.** By Definition 2.5, decoupled description corresponds to a flat connection ω_{flat} with $\Omega \equiv 0$. By Lemma 3.1 (using simple-connectivity of M), $\Omega \equiv 0 \iff \text{Hol}^0(\omega, p) = \{e\} \iff$ parallel transport is path-independent. Path-independent parallel transport is precisely the condition for the fiber and base dynamics to decouple: the fiber state at any point is determined solely by the initial condition and the base point, with no dependence on the path history (i.e., no coupling). By Lemma 3.2, this is equivalent to the existence of a global parallel section, which is the mathematical statement that the fiber degrees of freedom evolve independently of the base. (Part I)

Part II. Follows directly from Lemma 3.4 with the error definition from the audit response to Objection 3. (Part II)

Part III. Since $\mathcal{E}(S_{\text{flat}}, S) \geq c \int_M \|\Omega\|^2 d\text{vol}_g$, the integrand $\|\Omega_x\|^2$ is the pointwise density of the lower bound on modeling error. The supremum $\sup_x \|\Omega_x\|$ (which exists by compactness of M) is achieved at some $x^* \in M$. Any loop passing through x^* and enclosing x^* carries the maximum holonomy per unit area, so the maximum discrepancy between τ_γ and τ_γ^0 is achieved by loops concentrated near x^* . (Part III)

Part IV. When the bundle has non-trivial characteristic classes $c_k(P) \neq 0$, Lemma 3.3 implies that *every* connection on P has $\Omega \neq 0$, and therefore $\mathcal{H}(\omega) > 0$. The characteristic classes are topological invariants of the bundle, so this positivity is stable under continuous deformation of the connection. A flat model requires $\Omega = 0$, which by Chern–Weil is

impossible on P ; it would have to live on a topologically distinct (trivial) bundle P_0 , and any comparison between P and P_0 incurs at least the holonomy discrepancy of Corollary 5.3.

Note: The value $\mathcal{H}(\omega)$ depends on the choice of connection within P —it is not itself a topological invariant. What *is* topological is its strict positivity. The infimum $\inf_{\omega} \mathcal{S}_{\text{YM}}(\omega)$ over all connections on P provides a topological lower bound, achieved (when it exists) by Yang–Mills connections. $\square$

Objection 4 (most serious). Part I proves that decoupling $\Leftrightarrow$ flat connection, but the claim “fiber and base dynamics decouple $\Leftrightarrow$ path-independent parallel transport” is a *definition*, not a theorem. The word “decouple” has been defined to mean $\omega = 0$ from the start (Definition 2.5), so Part I is tautological. The actual scientific claim—that physical systems or computational architectures with non-trivial holonomy cannot be faithfully modeled by independent components—requires an additional premise: that the system’s correct description *is* a fiber bundle with non-trivial connection. This premise is not proved; it is assumed.

This is the most important objection. It is *correct*.

The theorem as stated is internally valid but proves something more modest than its framing suggests. The full scientific claim has two parts:

- (1) (*Representation theorem*) The system S is faithfully described by a fiber bundle with connection.
- (2) (*Non-decoupling theorem*) Given (1), if $\Omega \neq 0$, flat models fail.

This paper proves (2) rigorously. Part (1) is a modeling assumption that must be verified empirically for each application domain. For the Marcella architecture, the experimental evidence (V3-Fiber PPL 11.055 vs. V3RG PPL 3.83 on A100; gate freezing at $\sigma \approx 0.50$; Christoffel divergence $\kappa_2 = 54 \rightarrow 553$ at post-attention positions) constitutes empirical support for (1).

The theorem should therefore be stated as a *conditional*: *If a system is correctly modeled as a fiber bundle with connection ω , then it cannot be faithfully represented by any flat model when $\Omega \neq 0$.* The falsifiable scientific claim is whether specific systems (AI architectures, biological networks, financial systems) admit fiber bundle descriptions. This is an empirical question the theorem does not answer by itself.

This does not diminish the theorem. The same conditional structure applies to every application of Riemannian geometry in physics. General relativity does not *prove* that spacetime is a pseudo-Riemannian manifold; it *assumes* it and derives consequences. The Davis Non-Decoupling Theorem does the same for the fiber bundle representation.

5. COROLLARIES

Corollary 5.1 (Topological prohibition). *Let $P \rightarrow M$ carry a non-trivial characteristic class: $c_k(P) \neq 0 \in H^{2k}(M; \mathbb{Z})$. Then no connection on P is flat. In particular, no flat model S_{flat} of S can exist on the same bundle; any flat model requires a topologically distinct bundle, and the modeling error $\mathcal{E}(S_{\text{flat}}, S)$ is bounded below by a topological invariant independent of the choice of flat model.*

Proof. By Lemma 3.3, $c_k(P) = [\phi_k(\Omega)]$ is independent of the connection. If any connection were flat ($\Omega = 0$), then $\phi_k(\Omega) = 0$, giving $c_k(P) = 0$, contradicting the hypothesis. Hence every connection on P has $\Omega \neq 0$, and the theorem applies. $\square$

Remark 5.2. For non-simply-connected M , a flat connection may exist with non-trivial discrete holonomy $\text{Hol}(\omega, p) \subset G$ finite. In this case decoupling fails for a different reason: the fiber returns to a rotated—not identical—configuration after traversing non-contractible loops. The error is concentrated on the generators of $\pi_1(M)$.

Corollary 5.3 (Holonomy memory). *Under the hypotheses of Theorem 4.1, the holonomy group $\text{Hol}(\omega, p)$ is generated by the curvature values:*

$$\text{Lie}(\text{Hol}(\omega, p)) = \text{span}_{\mathbb{R}} \{ \Omega_q(X, Y) : q \in P_x, X, Y \in H_q \}$$

(Ambrose-Singer theorem). *A flat model asserts $\text{Hol}(\omega, p) = \{e\}$, predicting zero geometric phase for every loop. The holonomy debt is the accumulated phase discrepancy:*

$$\delta_\gamma = \tau_\gamma - \tau_\gamma^0 \in G,$$

which is non-trivial for every loop whose homology class is non-zero in $H_1(M; \mathbb{Z})$ when $\text{Hol}(\omega, p)$ is connected.

Corollary 5.4 (Berry's phase as canonical instance). *Let $\mathcal{H}$ be a Hilbert space of quantum states, $\mathcal{M}$ the manifold of Hamiltonian parameters, and $P = \{(H, |\psi\rangle) : H|\psi\rangle = E|\psi\rangle\} \rightarrow \mathcal{M}$ the eigenstate bundle (a $U(1)$ -bundle). The Berry connection is $A = \langle \psi | d | \psi \rangle$ with curvature (Berry curvature)*

$$F_{mn} = \partial_m A_n - \partial_n A_m = \text{Im} \sum_{k \neq n} \frac{\langle n | \partial_m H | k \rangle \langle k | \partial_n H | n \rangle}{(E_k - E_n)^2}.$$

When parameters traverse a closed loop Γ in $\mathcal{M}$, the state acquires the geometric (Berry) phase:

$$\gamma = i \oint_{\Gamma} \langle \psi | \nabla_\lambda | \psi \rangle d\lambda = \iint_{\Sigma} F d\sigma \neq 0,$$

despite the Hamiltonian returning to its original value. A flat model ($F = 0$) predicts zero phase. The system returns with a phase that the flat model cannot account for. This is the holonomy debt of Corollary 5.3, instantiated in quantum mechanics.

Objection 5. Corollary 5.4 asserts that the Berry connection has non-zero curvature, but this is only true generically. For specific Hamiltonians, $F \equiv 0$ (e.g., real Hamiltonians with time-reversal symmetry in some parameter spaces). The corollary as stated implies the Berry connection is always non-flat, which is false.

Correct. The corollary should read: *When $F \neq 0$ —which occurs for generically chosen Hamiltonians and is experimentally confirmed for electrons in solids (Thouless et al., TKNN 1982), photons in topological media, and cold atoms in optical lattices—a flat model fails by exactly the Berry phase.* Systems with $F = 0$ are the exception (measure zero in the space of Hamiltonians) and correspond precisely to systems for which decoupling is valid. This is consistent with the theorem: $F = 0 \Leftrightarrow$ decoupling works.

6. THE DAVIS LAW AS SPECIAL CASE

Proposition 6.1 (Davis Law encodes the Non-Decoupling Theorem). *The Davis Law $C = \tau/K$ is a scalar projection of Theorem 4.1 onto a one-dimensional summary statistic. Setting $K = \|\Omega\|$ (the operator norm of the curvature at a point), τ the tolerance (regularization capacity of the geometry), and C the representational capacity of the model, the following correspondences hold:*

<i>Assumption</i>	<i>Davis Law</i>	<i>Geometric consequence</i>
<i>Decoupling ($K = 0$)</i>	$C \rightarrow \infty$	<i>Unphysical capacity: model claims to represent all functions</i>
<i>Decoupling ($K = 0$)</i>	$\tau \rightarrow 0$	<i>Zero tolerance: catastrophic OOD failure</i>
<i>Geometry ($K > 0$)</i>	$C = \tau/K < \infty$	<i>Finite, regularized capacity</i>
<i>Max curvature</i>	$C \rightarrow 0$	<i>Capacity collapses: failure near singularities</i>

Proof. The decoupling assumption $\omega = 0$ implies $\Omega = 0$, i.e., $K = 0$. Substituting into $C = \tau/K$: if $\tau > 0$ (i.e., the model has any representational tolerance), then $C \rightarrow \infty$. This is the overconfidence failure mode: the model asserts it can represent every function. Empirically this corresponds to overfitting in distribution.

Alternatively, maintaining C finite requires $\tau \rightarrow 0$ as $K \rightarrow 0$, meaning the model’s tolerance for distributional shift collapses to zero: perfect in-sample, catastrophic out-of-sample. This is the brittleness failure mode.

When $K > 0$, the geometry provides the regularizer: $C = \tau/K$ is finite and determined by the ratio of tolerance to curvature. Large curvature ($K \gg 1$) at a point means the geometry is rapidly changing, the model’s capacity there is small, and errors concentrate—consistent with Part III of Theorem 4.1. $\square$

Objection 6. Proposition 6.1 sets $K = \|\Omega\|$ (a 2-form norm) and identifies this with the scalar K in $C = \tau/K$. But the Davis Law $C = \tau/K$ was originally stated with K as the sectional curvature of a Riemannian manifold, not the curvature of a principal bundle connection. These are different geometric objects. The identification needs justification, or the Davis Law must be re-derived in the bundle setting.

This is a legitimate gap. The reconciliation runs as follows.

For an associated vector bundle $E = P \times_{\rho} V$ (where $\rho : G \rightarrow GL(V)$ is a representation), the connection ω on P induces a covariant derivative ∇ on E . The curvature of ∇ is a 2-form $R^{\nabla} \in \Omega^2(M, \text{End}(E))$, which contracts to give sectional curvatures $K(\sigma) = \langle R^{\nabla}(e_1, e_2)e_2, e_1 \rangle$ for 2-planes $\sigma = \text{span}(e_1, e_2)$.

For the specific case of the tangent bundle ($E = TM$, $\rho = \text{adjoint}$), this recovers the Riemannian sectional curvature. So the Davis Law with $K = \text{sectional curvature}$ is a special case of the bundle framework with $P = \text{Fr}(TM)$ (the frame bundle).

For general bundles, K in the Davis Law should be interpreted as the norm of the induced connection curvature on the associated bundle relevant to the application. For Marcella: E is the token-embedding bundle, $K = \|R^{\nabla}\|$ is the norm of its induced curvature. The Davis Law is then: $C = \tau / \|R^{\nabla}\|$, which is the bundle version stated in Proposition 6.1.

A full derivation of the Davis Law from the bundle setting is deferred to a companion paper.

7. MASTER STATEMENT

Master Statement

Under the standing hypotheses (compact oriented simply-connected M , compact connected G , system S faithfully modeled as principal G -bundle with connection ω):

$$\begin{aligned} \Omega \not\equiv 0 &\implies \mathcal{E}(S_{\text{flat}}) \geq c \mathcal{S}_{\text{YM}}(\omega) > 0 \implies \text{Hol}(\omega) \neq \{e\} \\ &\implies \text{flat models are topologically prohibited (if } c_k \neq 0). \end{aligned}$$

Each implication is strict. The first is quantitative (Lemma 3.4). The second is the Ambrose-Singer theorem (Corollary 5.3). The third requires non-trivial characteristic classes (Corollary 5.1).

The converse chain:

$$\Omega \equiv 0 \iff \mathcal{E}(S_{\text{flat}}) = 0 \iff \text{Hol}(\omega) = \{e\} \iff \text{decoupling is valid.}$$

8. SCOPE OF APPLICATION

The theorem applies, *subject to the fiber bundle representation assumption being empirically supported*, to any system where:

- (1) Components are assumed independent: economics (parallel agents), epidemiology (parallel individuals), neuroscience (parallel brain regions), flat AI (parallel attention heads, additive residual stream).
- (2) State evolution is modeled without coupling: Euclidean attention, additive architectures, standard transformer residual stream.
- (3) Long-range dependencies are added on top of flat structure rather than encoded geometrically: positional encodings, learned biases, post-hoc fine-tuning.

The diagnostic procedure is:

1. Identify the Decoupling Assumption: what two things does the model treat as independent parallel processes?
2. Construct or estimate the curvature Ω of the correct bundle.
3. Evaluate $\mathcal{S}_{\text{YM}}(\omega)$. This is the size of the bill coming due.
4. For non-trivial c_k : the bill cannot be paid with any flat model. The solution space requires geometry.

9. OPEN QUESTIONS AND FUTURE WORK

- (1) **Representation theorem.** Prove (or characterize the conditions under which) transformer architectures, financial networks, and epidemiological contact graphs admit faithful fiber bundle descriptions. This is the missing Part (1) identified in Objection 4.
- (2) **Tightness of the Yang-Mills bound.** Is $\mathcal{E}(S_{\text{flat}}) = \mathcal{S}_{\text{YM}}(\omega)$ achievable, or is the bound only a lower bound? A tight bound would give the exact holonomy debt.
- (3) **Davis Law derivation.** Derive $C = \tau/K$ rigorously from the bundle framework (Proposition 6.1 deferred this).
- (4) **Non-simply-connected M .** Characterize the failure modes when $\pi_1(M) \neq 0$: the interplay between topological (c_k) and fundamental-group (π_1) obstructions.
- (5) **Infinite-dimensional bundles.** For Marcella and general neural architectures, M is the token-sequence manifold and P is an infinite-dimensional bundle. Uhlenbeck's compactness theory extends to this setting only under restrictive conditions. A functional-analytic treatment is needed.

ACKNOWLEDGMENTS

The fiber bundle interpretation emerged from the empirical failure of V3RG and the subsequent physics-guided redesign described in the companion experimental report. The adversarial audit in this paper was conducted by the author against her own theorem. The author thanks the referee who will object to Objection 4.

REFERENCES

- [1] W. Ambrose and I. M. Singer, *A theorem on holonomy*, Trans. Amer. Math. Soc. **75** (1953), 428–443.
- [2] M. V. Berry, *Quantal phase factors accompanying adiabatic changes*, Proc. R. Soc. Lond. A **392** (1984), 45–57.
- [3] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vols. I–II, Wiley Interscience, 1963.
- [4] J. Milnor and J. Stasheff, *Characteristic Classes*, Ann. of Math. Studies **76**, Princeton Univ. Press, 1974.
- [5] U. Schreiber and K. Waldorf, *Parallel transport and functors*, J. Homotopy Relat. Struct. **4** (2009), 187–244.
- [6] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, *Quantized Hall conductance in a two-dimensional periodic potential*, Phys. Rev. Lett. **49** (1982), 405–408.
- [7] K. Uhlenbeck, *Connections with L^p bounds on curvature*, Comm. Math. Phys. **83** (1982), 31–42.
- [8] K. Wehrheim, *Uhlenbeck Compactness*, EMS Series of Lectures in Mathematics, 2004.

INDEPENDENT RESEARCHER

Email address: `bee_davis@alumni.brown.edu`