Repairing Futures

Admissible Continuation, Historical Reconstruction, and Diffusion-Based Physical Simulation

Flyxion

June 2026

Abstract

Conventional simulators assume complete state descriptions and unique future trajectories. Both assumptions fail in practice: mass, friction, restitution, and contact geometry are typically unobserved, making future prediction the problem of constructing admissible continuations rather than computing unique outcomes. This paper argues that PHYSIFORMER [1], a diffusion transformer generating entire mesh trajectories in a single denoising pass, is more naturally described by repair theory and admissibility theory than by the language of mechanics or generative modelling. The denoising process is an iterative repair operator projecting corrupted histories onto the manifold of physically admissible trajectories. Initial conditions function as witnesses: they do not store the future but contain sufficient structure to reconstruct a sample from the admissible continuation set. The distribution over generated trajectories is a distribution over admissible futures, not over noise. We formalize this interpretation through the Admissibility Compression Principle — physical histories occupy a manifold of dimension far smaller than the ambient trajectory space, making manifold learning easier than explicit simulation — and develop extensions including admissibility fields, repair-aware training objectives, an optionality field measuring how many admissible futures remain open, and an RSVP field-theoretic generalization from which trajectories emerge as derived quantities.

1 The Incomplete State

Every physical simulation begins with the same idealization. The system has a state. That state contains all information relevant to future evolution. Given the state at time t, the dynamics determine the state at time $t + \Delta t$. The resulting picture is Markovian, deterministic, and mathematically elegant. It also fails to describe any realistic system.

The failure is not incidental. It is structural. A complete state description for a mechanical system requires not only positions and velocities but every material parameter, every contact condition, every coefficient of friction and restitution, the precise geometry of all contact surfaces, the distribution of internal stresses, and whatever additional variables govern the specific material response under the transformations the system will undergo. No measurement procedure provides all of these. A robot manipulating an object does not know its mass precisely. A character controller in a physics engine does not know the exact elasticity of the floor. An archaeologist reconstructing the trajectory of a falling vessel does not know its center of mass. In practice, the state is always partially specified, and the gap between the specified portion and the complete state description is not small.

The standard response to this situation is to treat the gap as noise. Unknown parameters become random variables. The dynamics acquire a stochastic component. The unique future trajectory becomes a distribution over trajectories. This response preserves the Markovian architecture by enlarging the probability model, but it conceals the more fundamental point. The distribution is not primarily an expression of epistemic uncertainty about a unique underlying trajectory. It is an expression of the multiplicity of physically admissible trajectories consistent with what is known. The difference is significant because it changes what the model is doing. A model of epistemic uncertainty about a unique future is tracking ignorance. A model of admissible continuation is tracking possibility.

The distinction appears clearly in the structure of PHYSIFORMER. The paper notes that the model does not receive friction coefficients, restitution coefficients, or mass distributions as conditioning variables. These are left unspecified, to be determined during learning. The trained model then produces diverse trajectories from the same initial conditions. The paper describes this as capturing uncertainty in the learned dynamics. But the more precise description is that the model has learned to sample from the set of trajectories consistent with the given initial geometry and velocity under physically plausible values of the unobserved parameters. The distribution is not noise. It is the admissible continuation set.

2 Histories Before States

The difficulty with state-centric simulation runs deeper than incomplete observation. Even setting aside the partial specification problem, there is a prior question about what the appropriate primitive object of physical description actually is. States are not observed. Trajectories are. A camera records sequences of frames. A sensor array records sequences of readings. A simulation produces sequences of configurations. The state at a single instant is a theoretical abstraction inferred from the trajectory, not a directly encountered object. The history is prior.

This priority is not merely epistemic. It is computational. Reconstruction problems naturally live on histories rather than isolated states. An archaeologist reconstructing the history of a site does not reconstruct a single state; the site's current condition is the single state, and it is directly available. What is reconstructed is the trajectory that produced the current condition from some earlier one. A developmental biologist tracing a cell lineage does not reconstruct the present molecular state; that is measured directly. What is reconstructed is the history of divisions and differentiation events that connects the ancestor to the observed descendant. A structural engineer diagnosing a failure does not reconstruct the failed state; the failure is present. What is reconstructed is the loading history that caused the failure. In each case, the trajectory is the object of interest, and the current state is a projection from it.

This observation motivates a shift in mathematical ontology. Let X be a state space. The standard object of classical simulation is a map $s : \{0, 1, \ldots, T\} \to X$ understood as a time series of states, with the generator being the local transition rule $s(t) \to s(t + 1)$. The history-centric formulation instead treats the entire trajectory $H = (s_0, s_1, \ldots, s_T) \in X^{T+1}$ as the primary object, with individual states appearing only as projections $s_t = \pi_t(H)$. The generator in this formulation is not a local rule but an operator on the space of admissible histories.

Definition 2.1 (History Space)**.** Let X be a state space and $T \geq 1$ a time horizon. The history space is $= X^{T+1}$, equipped with the product topology. Individual states are projections $\pi_t :\to X$ defined by $\pi_t(H) = s_t$.

Definition 2.2 (Admissible History)**.** Let Φ be a set of physical constraints. A history $H \in$ is admissible relative to Φ if it satisfies every constraint in Φ. The set of admissible histories is

$$\mathcal{A}(\Phi) = \{H \in: H \text{ satisfies every } \phi \in \Phi\}.$$

The admissible set $\mathcal{A}(\Phi)$ is the central object of the theory. A perfect physics simulator computes the unique element of $\mathcal{A}(\Phi)$ consistent with fully specified initial conditions. A partial-observation system selects from among the elements of $\mathcal{A}(\Phi)$ consistent with the available evidence. The transition from simulation to prediction under partial information is precisely the transition from singleton reconstruction to sampling from a constrained manifold.

The following theorem makes precise the sense in which PHYSIFORMER represents a strictly more general computational paradigm than classical simulation.

Theorem 2.3 (Historical Simulation Theorem)**.** *Let $\mathcal{S}$ be a state space and $= \bigcup_{T \geq 0} \mathcal{S}^T$ the associated history space. Every state-transition simulator $F : \mathcal{S} \to \mathcal{S}$ induces a history generator $\widetilde{F} : \mathcal{S} \to$″ by the orbit map $\widetilde{F}(s_0) = (s_0, F(s_0), F^2(s_0), \ldots)$. However, not every history generator factors through a local transition operator. A history generator $G : \mathcal{S} \to$″ is non-Markovian if there exist initial conditions $s, s' \in \mathcal{S}$ and times t such that*

$$\pi_t(G(s)) = \pi_t(G(s')) \quad \textit{but} \quad \pi_{t+1}(G(s)) \neq \pi_{t+1}(G(s')).$$

A non-Markovian history generator cannot be represented as $\widetilde{F}$ for any local F.

Proof. If $G = \widetilde{F}$ for some local F, then $\pi_{t+1}(G(s)) = F(\pi_t(G(s)))$ for all s. If $\pi_t(G(s)) = \pi_t(G(s'))$, then $\pi_{t+1}(G(s)) = F(\pi_t(G(s))) = F(\pi_t(G(s'))) = \pi_{t+1}(G(s'))$, contradicting the non-Markovian hypothesis. □

Remark 2.4. PHYSIFORMER empirically demonstrates that useful physical prediction can be performed directly in″ without explicit construction of a local F. The model G_θ maps initial conditions (X_0, V_0, m) to distributions over trajectories $(X_1, \ldots, X_T)$, and the distribution is non-Markovian in the sense that identical states at an intermediate time can produce different continuations depending on the global trajectory context. This is not a failure but a feature: it is what allows the model to maintain global constraints such as rigidity, which state-local prediction cannot enforce.

The critical observation about PHYSIFORMER is that it operates exactly in this mode. Given (X_0, V_0) and material type, with no further specification of the physical parameters, it produces samples from a distribution that concentrates on the admissible histories consistent with those initial conditions. This is why it generates diverse trajectories rather than a unique one: the initial conditions underdetermine the future, and the model has learned the shape of that underdetermination.

3 Admissible Continuation Sets

When initial conditions do not uniquely determine a trajectory, the appropriate object of study is the set of continuations that remain physically plausible. This set has a natural geometric structure.

Definition 3.1 (Admissible Continuation Set)**.** Given an initial condition $w = (X_0, V_0)$, the admissible continuation set is

$$\mathcal{A}(w) = \{H \in : \pi_0(H) = X_0,\ \dot{\pi}_0(H) = V_0,\ H \in \mathcal{A}(\Phi)\}.$$

This is the set of all physically admissible histories whose initial position and velocity match the given condition.

The admissible continuation set is never a singleton in any realistic setting. Even holding positions and velocities fixed at $t = 0$, the space of physically consistent futures spans a positive-dimensional manifold because the unobserved parameters (mass, friction, elasticity, contact geometry) can vary continuously over ranges consistent with the physical constraint set Φ. The manifold structure of this set determines the character of the prediction problem.

A natural measure of the size of the admissible continuation set is its logarithmic volume, which plays the role of a continuation entropy.

Definition 3.2 (Continuation Entropy)**.** Let μ be a reference measure on˝. The continuation entropy at initial condition w is

$$S_C(w) = \log \mu(\mathcal{A}(w)).$$

High continuation entropy means that many physically plausible futures remain open. Low continuation entropy means that the initial condition strongly constrains the future. A perfectly specified initial condition would have $S_C(w) = 0$, corresponding to a unique trajectory. The continuation entropy of a realistic mechanical system is positive and depends sensitively on the degree of contact and the material regime.

The distribution learned by PHYSIFORMER approximates the uniform distribution over $\mathcal{A}(w)$ subject to the implicit prior induced by the training data. The diversity of generated trajectories from the same initial conditions reflects the continuation entropy of the scene. Scenes with simple, non-interacting dynamics have low continuation entropy and nearly deterministic outputs. Scenes with collisions and material heterogeneity have high continuation entropy and diverse outputs.

The paper's observation that contact-angle differences in collisions produce divergent yet physically plausible trajectories is precisely the empirical signature of a positive-entropy continuation set.

Remark 3.3. The continuation entropy framework clarifies why mean-squared-error against a single ground-truth trajectory is an imperfect evaluation metric for this class of model. A system with high continuation entropy can generate a trajectory that is far from the ground-truth trajectory in coordinate space while remaining fully physically admissible. The paper acknowledges this explicitly, noting that MSE alone does not indicate physical plausibility, and uses rigidity preservation and momentum drift as complementary metrics. From the continuation perspective, the appropriate evaluation would measure whether the generated trajectory lies within $\mathcal{A}(w)$ rather than how close it is to a particular sample from $\mathcal{A}(w)$.

The admissible continuation set has a geometric interpretation as a submanifold of history space. Physical constraints are typically smooth and overdetermined: they impose a family of equations on the trajectory that cut down the dimension of˝ to a lower-dimensional submanifold. The constraint manifold $\mathcal{M} \subset$˝ is the set of all histories satisfying the physical laws, and $\mathcal{A}(w)$ is the fiber of $\mathcal{M}$ over the initial condition w.

Theorem 3.4 (Manifold Structure of Admissible Continuations)**.** *Suppose Φ consists of smooth equality constraints on˝ that are transverse to the initial-condition subspace $\pi_0^{-1}(X_0) \cap \dot{\pi}_0^{-1}(V_0)$. Then $\mathcal{A}(w)$ is a smooth submanifold of˝.*

Proof. By transversality, the constraint equations cut out a smooth submanifold $\mathcal{M}$ of˝. The initial-condition fiber $\{H : \pi_0(H) = X_0, \dot{\pi}_0(H) = V_0\}$ is a closed affine subspace of˝. Their intersection $\mathcal{A}(w)$ is smooth by the preimage theorem applied to the restriction of the constraint map to the initial-condition fiber. □

This theorem establishes the appropriate target for a trajectory model: not a single point in˝ but a distribution over a submanifold. A generative model that places its probability mass on $\mathcal{A}(w)$ is physically faithful regardless of which element of $\mathcal{A}(w)$ it generates.

The dimension of the admissible submanifold relative to the ambient history space is the key quantity determining why trajectory-space learning is tractable at all.

Theorem 3.5 (Admissibility Compression Principle)**.** *Let $= \mathbb{R}^{3NT}$ be the ambient trajectory space for N vertices over T time steps, and let $\mathcal{M} \subset$˝ be the admissible manifold* *under a complete set of physical constraints. If* $\dim(\mathcal{M}) = d_{\mathcal{A}}$*, then*

$$d_{\mathcal{A}} \ll 3NT$$

for any realistic mechanical system, and learning a generative model of $\mathcal{M}$ *requires sample complexity scaling with* $d_{\mathcal{A}}$ *rather than with* $3NT$*. Consequently, reconstruction of admissible continuations can be substantially easier than explicit simulation of the underlying differential equations.*

Proof. A trajectory $H \in \mathbb{R}^{3NT}$ must satisfy at minimum the following independent constraint families. Continuity imposes $3N(T-1)$ constraints (velocity compatibility between adjacent frames). Rigid body motion for an object with n vertices imposes $\binom{n}{2}$ distance-preservation constraints per frame. Material constraints, boundary conditions, and momentum conservation impose further independent equations. Each independent constraint reduces the dimension of $\mathcal{M}$ by one. The total number of independent constraints grows at least linearly in N, T, and n, while the ambient dimension grows as $3NT$. In the limit of many vertices and timesteps, $d_{\mathcal{A}}/(3NT) \to 0$. The sample complexity bound follows from the fact that the covering number of $\mathcal{M}$ at resolution ϵ scales as $\epsilon^{-d_{\mathcal{A}}}$, and a generative model requires samples proportional to the covering number. □

Remark 3.6. This theorem explains why PHYSIFORMER succeeds empirically without learning the underlying dynamics. The admissible trajectory manifold is so thin relative to the ambient space $\mathbb{R}^{3NT}$ that a sufficiently expressive model can learn the geometry of coherent histories directly from trajectory examples, without ever representing a force, a differential equation, or a collision operator. The model is learning $\chi_{\mathcal{A}}(H)$, the indicator function of admissibility, and using that learned geometry to project corrupted histories back onto $\mathcal{M}$ during denoising. Physics is easier to imitate than to derive precisely because the space of physically coherent histories is vastly smaller than the space of possible trajectories.

4 Repair as the Fundamental Operation

The denoising process in diffusion models has a natural interpretation in the framework developed above. Beginning from a corrupted or randomly initialized history, the denoiser iteratively adjusts the history to bring it closer to the admissible manifold. At convergence, the output is an element of $\mathcal{A}(w)$. This process is not generation in the usual sense of the word. It is repair.

Definition 4.1 (Repair Operator on History Space)**.** A repair operator is a map $R :\to$″ satisfying the following conditions. First, $R(H) \in \mathcal{A}(w)$ for every $H \not\in$ in the basin of attraction of the operator. Second, R is a contraction with respect to some metric on″ measuring distance to $\mathcal{A}(w)$. Third, R preserves the initial condition: $\pi_0(R(H)) = \pi_0(H) = X_0$ and $\dot{\pi}_0(R(H)) = V_0$.

The connection to diffusion is direct. In the flow-matching formulation used by PHYSIFORMER, training minimizes

$$\mathcal{L}(\theta) = \mathbb{E}_{x,\epsilon} \left\| x_\theta(\tau x + (1-\tau)\epsilon,\, \tau) - x \right\|^2 ,$$

where x is a clean trajectory and ϵ is noise. The network x_θ learns to predict the clean trajectory from any noisy interpolant. During inference, integrating the corresponding ODE from $\tau = 0$ to $\tau = 1$ moves a randomly initialized history toward the data distribution, which concentrates on admissible histories. Each integration step adjusts the corrupted history to reduce its distance to the admissible manifold. The sequence of iterates is exactly a discrete repair sequence in the sense of the definition above.

Theorem 4.2 (Diffusion as Iterative Admissibility Projection)**.** *Let $\mathcal{M} \subset$″ be the admissible manifold and let $d_{\mathcal{M}} :\to \mathbb{R}_{\geq 0}$ denote the geodesic distance to $\mathcal{M}$. Suppose the score function of the trained diffusion model is the negative gradient of $d^2_{\mathcal{M}}$ restricted to each noise level. Then each denoising step is a projected gradient step toward $\mathcal{M}$, and the full denoising trajectory converges to an element of $\mathcal{M}$.*

Proof. Under the stated condition, the denoising velocity field points in the direction of decreasing $d^2_{\mathcal{M}}$, which is the direction toward the nearest element of $\mathcal{M}$. Each Euler step reduces $d_{\mathcal{M}}(H_\tau)$. Since $\mathcal{M}$ is a smooth compact manifold and the step sizes decrease to zero as $\tau \to 1$, the iterates converge to the projection of the initial point onto $\mathcal{M}$. The projection is an element of $\mathcal{M}$, hence admissible. □

Remark 4.3. The assumption that the score function is the gradient of the squared distance to $\mathcal{M}$ is not exactly satisfied in practice, but it identifies what a well-trained diffusion model is approximating. The empirical performance of PHYSIFORMER is evidence that the trained score function is a sufficiently good approximation of this ideal for the trajectory distributions in the training data. The remaining failure modes — interpenetration, orientation discontinuities, and occasional rigidity violations — are precisely the cases where the approximation breaks down and the generated trajectory lies near but not on the admissible manifold.

The repair interpretation also explains why full-trajectory diffusion substantially outperforms autoregressive rollout for rigid bodies, which is the central empirical finding of the PHYSIFORMER paper. The explanation has two layers, one quantitative and one structural.

The quantitative layer concerns the number of constraints that rigidity imposes. For a rigid object with n vertices, rigidity requires the preservation of all pairwise distances:

$$\|X_i(t) - X_j(t)\| = \|X_i(0) - X_j(0)\| \quad \text{for all } i \neq j,\ t \geq 0.$$

This imposes $\binom{n}{2} = O(n^2)$ constraints simultaneously. An autoregressive predictor minimizes local step error $\epsilon_t = \hat{X}_t - X_t$ at each frame. Nothing in the local objective ensures that the $O(n^2)$ pairwise constraints are satisfied. Each step introduces a small violation that is invisible to the local loss. The accumulated error $\sum_t \epsilon_t$ grows until the global rigidity constraint is macroscopically violated, and the rigid body melts. The rigidity loss reported in table 2 of the paper grows by three orders of magnitude between the 10-frame and 49-frame horizons for all autoregressive baselines, confirming this analysis quantitatively.

The structural layer concerns the fundamental difference between asking *what happens next* and asking *what nearby trajectory is globally coherent*. Autoregressive methods pose the former question at each step. PHYSIFORMER poses the latter question over the entire trajectory simultaneously. The former is local and has no mechanism for enforcing global properties. The latter is global: the denoiser can sense rigidity violations anywhere in the trajectory and correct them because the attention mechanism spans all time steps. Later temporal constraints propagate backward and earlier constraints propagate forward. The future constrains the past. The end constrains the middle. The generated trajectory is jointly admissible rather than locally admissible, and this joint admissibility is what preserves rigidity. The paper's rigidity loss for PHYSIFORMER grows by only a factor of approximately four between 10 and 49 frames, against three orders of magnitude for the baselines.

The repair-theoretic interpretation makes this result a theorem rather than an empirical observation. A repair operator that converges to the admissible manifold will, by definition, produce a trajectory whose rigidity error approaches zero at convergence. An autoregressive operator that minimizes local step error will not in general produce a trajectory on the manifold. The performance gap between the two architectures is a structural consequence of the difference between local optimization and global repair.

Remark 4.4 (Spherepop Formulation)**.** The repair-theoretic account of trajectory generation has a natural expression in the Spherepop computational framework, where histories are the primary computational objects and repair is a primitive operation. In Spherepop terms, the PHYSIFORMER inference procedure can be written as a sequence of four operations: POP the witness $W = (X_0, V_0, m)$ from the environment; EXPAND the admissible continuation set $\mathcal{A}(W)$ by sampling an initial corrupted history $\tilde{H} \sim \mathcal{P}_\sigma$; REFUSE inadmissible histories through repeated repair $\tilde{H} \leftarrow R(\tilde{H})$ until convergence; and COLLAPSE the repaired history $H_{\text{adm}} = R^n(\tilde{H})$ as the model's output trajectory. The denoising loop is the refuse-and-repair step. Generation is a side effect of repair. This formulation makes clear that diffusion is not the fundamental operation; it is an implementation of a more abstract computational pattern in which inadmissible structures are iteratively corrected until they reach a coherent fixed point.

5 Witnesses and the Restorability of Futures

The initial conditions (X_0, V_0) play a distinctive role in this framework. They do not determine the future uniquely, but they do constrain it. They function as witnesses in the sense of restorability theory: a witness is a partial observation from which a targeted class of distinctions remains reconstructible, even though the full state of the system is not observed.

Definition 5.1 (Witness for a Continuation Set)**.** A witness for the continuation set $\mathcal{A}(w)$ is a pair $w = (X_0, V_0)$ such that the conditional distribution $p(H \mid w)$ concentrates on $\mathcal{A}(w)$. A witness is sufficient if $\mathcal{A}(w)$ is non-trivial and bounded in continuation entropy.

The PHYSIFORMER experiments demonstrate that initial vertex positions and velocities constitute a sufficient witness for a wide range of mechanical scenes, including rigid and elastic objects with multiple material types and varied geometries. The model generates diverse but physically plausible trajectories from these witnesses without access to friction, mass, or restitution, because those parameters are not needed to identify the continuation set, only to select among its elements.

This observation connects to a general principle about the relationship between observation and prediction. Classical simulation requires complete specification of the physical state to generate a unique future. Restorability theory replaces this requirement with a weaker one: it suffices to provide enough information to reconstruct the admissible continuation set, not to identify a unique element of it. The future is not stored in the witness. It is recoverable from the witness in the sense that $\mathcal{A}(w)$ is determined by w, even though the specific trajectory that will actually occur is not.

Theorem 5.2 (Witness Sufficiency for Physical Prediction)**.** *Let Φ be a physically complete constraint set and $w = (X_0, V_0)$ be initial conditions. Then w is a sufficient witness in the sense that $\mathcal{A}(w)$ is determined by w and Φ, and every physically admissible trajectory is recoverable from $\mathcal{A}(w)$ by drawing a sample.*

Proof. By definition, $\mathcal{A}(w) = \{H \in \mathcal{A}(\Phi) : \pi_0(H) = X_0, \dot{\pi}_0(H) = V_0\}$. This set is determined entirely by w and Φ. Drawing a sample from a distribution supported on $\mathcal{A}(w)$ produces a physically admissible trajectory whose initial condition matches w. The trajectory is therefore recovered in the sense that it belongs to the admissible continuation set of the witness. □

Not all witnesses are equally informative. Two witnesses may both determine non-trivial admissible continuation sets while differing greatly in how tightly those sets constrain the future. This motivates a quantitative measure of witness quality.

Definition 5.3 (Witness Adequacy Functional)**.** Let $\hat{H}(w)$ denote the trajectory generated by the model from witness w, and let $d : \times \to \mathbb{R}_{\geq 0}$ be a metric on history space. The witness adequacy functional is

$$\Lambda(w) = \sup_{H \in \mathcal{A}(w)} d(H, \hat{H}(w)).$$

A witness w is adequate for the model at tolerance $\epsilon > 0$ if $\Lambda(w) \leq \epsilon$.

The functional $\Lambda(w)$ measures the worst-case distance between any admissible continuation and the model's output. Small $\Lambda(w)$ means the witness strongly constrains the future: all admissible histories are close to what the model generates. Large $\Lambda(w)$ means the witness leaves many admissible continuations open and the model's particular sample represents only a narrow slice of physical possibility. This is a formal notion of ambiguity.

A model could in principle learn to estimate Λ alongside the trajectory, producing pairs (H, Λ) that communicate both a generated future and the degree of underdetermination that future represents. Collision-rich scenes would produce large Λ because contact-angle uncertainty propagates into divergent post-collision trajectories. Free-flight scenes would produce small Λ because few unobserved parameters affect ballistic motion significantly. The witness adequacy functional therefore provides a geometrically grounded account of physical prediction confidence that is richer than any scalar uncertainty estimate.

The practical significance of this theorem is that it explains why PHYSIFORMER can generalize to unseen geometries, larger object counts, and mixed materials without retraining. The model has learned not a mapping from specific initial conditions to specific trajectories but a mapping from witness structure to admissible continuation distributions. New geometries instantiate a new $\mathcal{A}(w)$ with the same constraint structure Φ, and the model applies the same reconstruction operator. The generalization is to the structure of the continuation set, not to the particular trajectories in the training data.

6 The Kinematics–Mechanics Distinction

The repair-theoretic interpretation clarifies a tension latent in the PHYSIFORMER paper itself. The paper frames its contribution as learning mechanics. The title speaks of simulating mechanics, and the abstract claims that the model approximates how systems behave under the laws of mechanics. The limitations section, however, acknowledges interpenetrations, spurious contacts, and orientation discontinuities arising because no conservation laws or physical constraints are explicitly enforced. This tension has a clean resolution in the admissibility framework.

What PHYSIFORMER learns is admissible kinematics, not mechanics. The distinction is as follows. Mechanics explains why a given trajectory occurs. It specifies the forces and laws that generate the trajectory from the physical state. Kinematics describes which trajectories occur. It characterizes the set of motions consistent with geometric and material constraints. Mechanics implies kinematics, but kinematics does not imply mechanics. A model that learns to sample from the admissible kinematic manifold produces trajectories that look physically plausible because they satisfy the geometric constraints learned from training data, without encoding any causal account of why those constraints hold.

This is precisely the situation of PHYSIFORMER. The model learns the image of the dynamics rather than the dynamics themselves. It learns that rigid objects do not deform, that objects collide rather than pass through each other, and that momentum is approximately conserved, not because it represents any physical law but because training data that violates these patterns is absent. The learned distribution is the empirical distribution over admissible kinematics in the training corpus.

Remark 6.1. This is not a deficiency of PHYSIFORMER in particular. It is a general feature of any model trained solely on trajectory data without explicit physical constraints. The trained model approximates a distribution over admissible histories. Because physically generated data concentrates on the admissible manifold, the learned distribution necessarily does as well. But it is a distribution over a geometric object, not a representation of the causal laws that generate that object. The distinction becomes important at extrapolation: a causal model can reason about counterfactuals and novel force regimes. A kinematic model can only extrapolate to initial conditions covered by its implicit continuation manifold.

An RSVP-theoretic extension would address this distinction directly. In the RSVP framework, dynamics arise from field equations governing a scalar capacity field Φ, a vector transport field v, and a constraint density S. These fields satisfy coupled equations of the form

$$\partial_t \Phi + \nabla \cdot (v\Phi) = -S\Phi, \qquad \partial_t v + (v \cdot \nabla)v = -\nabla \log \Phi + f,$$

where f represents external forcing. Vertex positions are derived as integral curves of the transport field, $\dot{X}(t) = v(X(t), t)$, making geometry an emergent consequence of field dynamics rather than the primary prediction target. A model learning (Φ, v, S) would therefore capture causal structure rather than kinematic structure: the field equations determine why specific trajectories arise rather than merely which trajectories are consistent with training data. The trajectory $X(t)$ becomes a derived quantity $X(t) = \Gamma(\Phi, v, S)$, and learning the generating fields rather than the generated trajectories is the difference between a coordinate-shadow model and a field-world model. PHYSIFORMER is the former; an RSVP-extended model would be the latter. The former is sufficient for prediction within the training distribution. The latter is necessary for counterfactual reasoning and genuine extrapolation to regimes not represented in training data.

Remark 6.2 (Failure Modes as Admissibility Boundary Detection)**.** The failure modes reported by PHYSIFORMER are not arbitrary. The paper lists interpenetrations, spurious contacts, and orientation discontinuities as the primary failure cases. What these share is that they occur precisely where admissibility becomes geometrically difficult to enforce locally. Contact manifolds are singular structures in trajectory space: the admissibility conditions change discontinuously as objects touch or separate, producing high-curvature boundaries in $\mathcal{M}$. Near these boundaries, the gradient of the learned score function points in incoherent directions because small perturbations can move a trajectory to the wrong side of the contact boundary. The denoiser fails not because it has learned the wrong dynamics but because it has reached the high-curvature boundary of the admissibility manifold, where the geometry of coherent histories is locally hard to approximate. This interpretation has a direct design implication: the training objective should weight contact events more heavily, because those events correspond to regions where the admissibility boundary is most sharply defined and most difficult to learn from smooth trajectory data alone.

7 Diffusion Without Gaussian Noise

The repair interpretation suggests a generalization of diffusion that makes repair the fundamental operation rather than denoising as a special case. Standard diffusion perturbs the target with Gaussian noise and trains the network to invert the perturbation. The choice of Gaussian noise is natural for images and latent codes because it corresponds to a well-understood probabilistic model. For trajectory data, however, Gaussian noise is not the most natural perturbation. It corrupts trajectories in coordinate space without any reference to physical admissibility. The perturbed trajectory does not correspond to a physically meaningful inadmissibility; it simply fails to lie on the trajectory manifold in an unstructured way.

A more principled perturbation would be one that corresponds to a specific class of admissibility violation. One natural class is rigidity violation: perturb the trajectory so that pairwise vertex distances are no longer preserved, as if the rigid body had become elastic under the perturbation. Another natural class is momentum violation: add a drift to vertex velocities so that total momentum is no longer conserved. A third class is temporal discontinuity: introduce random gaps in the trajectory so that the position at time $t+1$ is not a smooth continuation of the position at time t. Each of these perturbation classes has a corresponding repair operator: the denoiser would learn to restore rigidity, restore momentum conservation, or smooth temporal continuity as appropriate.

Definition 7.1 (Admissibility-Structured Perturbation)**.** An admissibility-structured perturbation with parameter $\sigma > 0$ is a family of maps $\{\mathcal{P}_\sigma : \to \}_{\sigma>0}$ such that $\mathcal{P}_\sigma(H)$ violates a specified subset of physical constraints in Φ by an amount controlled by σ, while $\mathcal{P}_0(H) = H$ for all H.

A repair-based training objective would then minimize

$$\mathcal{L}_{\text{repair}}(\theta) = \mathbb{E}_{H\in\mathcal{A}(w),\,\sigma}\, \|x_\theta(\mathcal{P}_\sigma(H), \sigma) - H\|^2,$$

where x_θ is trained to predict the clean admissible history from the perturbed one. This objective is structurally identical to the standard diffusion objective, but with the perturbation designed to produce specific admissibility violations rather than generic noise. The network would learn repair operators for the specific constraint classes used in the perturbation, rather than learning general-purpose denoising.

The advantage of this formulation is that the learned network's inductive structure aligns with physical structure. A network trained on rigidity violations learns to detect and correct rigidity drift, which is exactly the failure mode that autoregressive models exhibit. A network trained on momentum violations learns momentum conservation as a repair target rather than as an emergent statistical regularity. The physics enters not through explicit constraint enforcement but through the structure of the perturbation and repair objectives.

Theorem 7.2 (Repair-Equivalence of Admissibility-Structured Diffusion)**.** *Suppose the perturbation family $\{\mathcal{P}_\sigma\}$ satisfies the condition that for every $H \in \mathcal{A}(w)$, the perturbed trajectory $\mathcal{P}_\sigma(H)$ has a unique projection onto $\mathcal{A}(w)$ which is H itself, for σ sufficiently small. Then the minimizer of $\mathcal{L}_{repair}$ is a repair operator in the sense of section 4, and the inference-time iteration*

$$H_{k+1} = x_\theta(H_k, \sigma_k)$$

converges to an element of $\mathcal{A}(w)$ as $\sigma_k \to 0$.

Proof. The projection condition ensures that the target H is recoverable from $\mathcal{P}_\sigma(H)$ in a neighborhood of $\mathcal{A}(w)$. The minimizer x_θ^* satisfies $x_\theta^*(\mathcal{P}_\sigma(H), \sigma) = H$ for all $H \in \mathcal{A}(w)$ in expectation, hence maps perturbed histories to admissible ones. The iteration initializes from a perturbed history and applies the repair map repeatedly with decreasing σ_k, each step reducing the distance to $\mathcal{A}(w)$. Convergence follows from the contraction property of the projection and the decay of σ_k. □

This theorem establishes that Gaussian diffusion and admissibility-structured diffusion have the same mathematical structure at the level of repair. The difference is the coordinate system in which repair is performed. Gaussian diffusion repairs coordinate-space corruption. Admissibility-structured diffusion repairs specific physical constraint violations. For trajectory prediction, the latter is more natural because the failures of interest are physical, not coordinate-geometric.

8 Objecthood as Emergent Coherence

A secondary but philosophically significant observation concerns the treatment of object identity in PHYSIFORMER. The paper introduces a factorized attention mechanism that operates separately over time, space, and objects, but without explicit object identifiers. Objects are not primitive entities in the architecture. They are implicitly encoded through the structure of attention over vertex tokens. The paper notes that this choice gives permutation invariance and allows generalization to unseen object counts. It achieves these properties by not treating objects as primitives.

This design choice aligns with a more general principle. Object identity is not a primitive feature of physical scenes but an emergent pattern of coherent persistence. A rigid body is not a thing; it is a pattern of vertex coordinates that maintains constant pairwise distances across time. The objecthood of the body is constituted by the rigidity of its trajectory. An elastic body is not a thing; it is a pattern of vertices that deforms under force but returns to a reference shape. The objecthood of the body is constituted by the elasticity of its trajectory. In both cases, the object is defined by its history, not by any static property of its instantaneous configuration.

This principle can be stated formally. Given a history $H \in$, define a relation on vertices by mutual reconstructibility.

Definition 8.1 (Historical Coherence)**.** Let $H \in$ be a trajectory over vertex set $\{1, \ldots, N\}$. Two vertices i and j are historically coherent, written $i \sim_H j$, if their relative trajectory $\pi_t(H)_i - \pi_t(H)_j$ is recoverable from the trajectory of either vertex alone under the admissible reconstruction family. The historical coherence class of vertex i is

$$[i]_H = \{j : i \sim_H j\}.$$

Proposition 8.2. *The historical coherence relation $\sim_H$ is an equivalence relation, and the coherence classes $[i]_H$ partition the vertex set.*

Proof. Reflexivity: a vertex's trajectory is trivially recoverable from itself. Symmetry: if $\pi_t(H)_i - \pi_t(H)_j$ is recoverable from the trajectory of i, then $\pi_t(H)_j - \pi_t(H)_i$ is also recoverable, and the trajectory of j is obtainable from the trajectory of i together with the relative trajectory, so $j \sim_H i$. Transitivity follows by composition of reconstruction operators. The partition property follows from equivalence. □

The coherence classes $[i]_H$ are the objects of the scene. A rigid body is a coherence class whose relative trajectories are all constant: $\pi_t(H)_i - \pi_t(H)_j$ is independent of t for i, j in the same class. An elastic body is a coherence class whose relative trajectories vary but remain reconstructible from any member's trajectory via the deformation field. Objecthood is therefore not a label but a relational property of the history.

PHYSIFORMER's object-level attention can be understood as discovering these coherence classes without being told where they are. Vertices that belong to the same rigid body will attend to each other across time because their relative positions are invariant and each provides a strong predictive signal for the other. The attention pattern that emerges is the signature of their coherence. The object is not given to the model; it is recognized by the model as a coherence class in the sense of the definition above. When the paper generalizes to unseen object counts, it is not applying a learned object representation to new objects. It is applying a learned coherence detector to configurations of vertices that form new coherent groups under the same historical coherence relation. The generalization follows from the fact that coherence, not object count, is the fundamental structure being represented.

The generalization to unseen geometries is the strongest evidence for this claim. PHYSIFORMER trains on simple convex primitives with at most 88 vertices and successfully generates physically plausible motion for fish, teapots, bunnies, and other complex meshes with hundreds of vertices per object. A model that had learned object-specific dynamics would fail at this task, because the cube and the teapot are different geometric objects. A model that has learned admissibility class invariants succeeds, because the cube and the teapot belong to the same rigid-body admissibility class. They are different shapes but the same coherence structure. Their trajectories satisfy the same distance-preservation constraints, the same momentum equations, and the same collision geometry (up to local curvature effects). The admissibility manifold does not distinguish between geometrically distinct rigid bodies; it only cares whether the history satisfies the underlying coherence relations. PHYSIFORMER's generalization is therefore not object recognition under distribution shift but admissibility class recognition across geometric variation. A cube is not a bunny, but both are rigid, and rigidity is what the model has actually learned.

This picture is consistent with the broader theme of the present paper. Objects, like trajectories, are not the primitive elements of physical description. They are derived from the persistence structure of histories. A trajectory is admissible if it lies on the physical constraint manifold. An object is identified if it constitutes a coherent persistent distinction within that admissible trajectory. Objecthood emerges from admissibility, and admissibility is the fundamental concept.

9 World Models as Admissibility Engines

The trajectory-generation perspective admits a natural generalization that connects to the broader problem of world modelling. A world model is typically described as a system that predicts future observations given present ones. Under the admissibility interpretation, this description is imprecise. A world model need not predict future observations. It must construct admissible continuations of present states, where admissibility is determined by the constraint structure of the domain being modelled.

Definition 9.1 (Admissibility Engine)**.** An admissibility engine for domain $\mathcal{W}$ is a map

$$\mathcal{E} : \mathcal{W}_0 \times \Omega \to$$

″ where $\mathcal{W}_0$ is the space of initial conditions, Ω is a probability space, and the output $\mathcal{E}(w, \omega)$ lies in $\mathcal{A}(w)$ for almost every $\omega \in \Omega$. The engine samples admissible continuations from an implicit distribution over $\mathcal{A}(w)$.

PHYSIFORMER is an admissibility engine for the domain of Newtonian rigid and elastic mechanics. The initial condition $w = (X_0, V_0)$ determines the constraint set $\mathcal{A}(w)$, and the diffusion process samples from this set by iterative repair. The generality of the definition, however, is not restricted to physics. The same structure applies to language: a language model that generates grammatically and semantically coherent continuations from a prompt is an admissibility engine for the domain of natural language, where admissibility is determined by grammatical, semantic, and pragmatic constraints. It applies to planning: a trajectory planner that generates dynamically feasible paths from an initial configuration is an admissibility engine for the domain of robot motion. It applies to simulation in any domain where continuations are constrained by rules: physical laws, grammatical rules, economic laws, institutional rules.

An admissibility engine that generates trajectories $X(t)$ can be extended to generate trajectory-optionality pairs $(X(t), \Omega(x, t))$, where the optionality field measures how many admissible futures remain available at each point in space-time.

Definition 9.2 (Optionality Field)**.** Let μ be a reference measure on″. The optionality field is the map

$$\Omega : \mathcal{S} \times [0, T] \to \mathbb{R}_{\geq 0}, \qquad \Omega(x, t) = \log \mu(\{H \in \mathcal{A} : \pi_t(H) = x\}) \,.$$

High optionality at (x, t) means that many admissible futures pass through position x at time t. Low optionality means the trajectory is strongly constrained at that point.

The optionality field has a natural physical interpretation. During free flight, the admissible continuation set is large because the trajectory is fully determined by the current state and the unobserved parameters have limited influence. The optionality is high. During a collision, the contact geometry, friction, and restitution interact to produce sensitive dependence on unobserved parameters. Small differences in impact angle produce large differences in post-collision trajectories. The optionality is low. The moment of impact is a rapid optionality collapse: a large admissible set contracts suddenly to a much smaller one as the collision event locks in the subsequent trajectory.

Proposition 9.3 (Optionality Monotonicity Under Constraint)**.** *If* $\Phi' \supset \Phi$ (*strictly stronger physical constraints*)*, then* $\Omega'(x,t) \leq \Omega(x,t)$ *for all* (x,t)*, where* Ω' *and* Ω *are the optionality fields under* Φ' *and* Φ *respectively.*

Proof. Stronger constraints reduce the admissible set: $\mathcal{A}(\Phi') \subseteq \mathcal{A}(\Phi)$. The fiber $\{H \in \mathcal{A}(\Phi') : \pi_t(H) = x\}$ is therefore a subset of $\{H \in \mathcal{A}(\Phi) : \pi_t(H) = x\}$, and the logarithmic measure of a subset does not exceed that of the set. □

An admissibility engine extended with the optionality field would produce qualitatively richer outputs than PHYSIFORMER currently does. Instead of only reporting what happens, it would report how constrained the future has become at each moment of the generated trajectory. A scene in which two rigid bodies narrowly miss each other would show a brief optionality spike near the near-miss event, reflecting the sensitivity of the post-near-miss trajectories to the exact geometry. A scene in which a deformable object strikes a rigid surface would show optionality collapse at impact and gradual optionality recovery as the deformation field stabilizes. These patterns are invisible to any model that outputs only vertex trajectories.

The unifying insight is that successful prediction under partial information is not the computation of a unique truth from complete specifications. It is the construction of admissible continuations from incomplete witnesses, together with a characterization of how tightly those witnesses constrain the future. PHYSIFORMER provides the first large-scale empirical confirmation of the former principle in the domain of physical trajectory generation. The optionality field extends it to the latter.

Conclusion

The PHYSIFORMER paper presents itself as an engineering contribution: a better architecture for neural physical simulation. The present paper has argued that it is also a conceptual contribution, confirming at scale a set of theoretical claims about the structure of prediction under partial information.

Those claims are as follows. State-centric simulation is an idealization of complete knowledge; realistic physical prediction operates over sets of admissible continuations rather than unique trajectories. Histories are prior to states by the Historical Simulation Theorem: not every history generator factors through a local transition operator, and PHYSIFORMER is the first large-scale empirical demonstration of a useful history generator in this non-Markovian class. Diffusion-based trajectory generation is iterative repair on the manifold of admissible histories; the denoiser is a repair operator in the formal sense, and the Gaussian noise is not fundamental but is one implementation of a more general inadmissibility perturbation structure that makes repair the primitive. Initial conditions function as witnesses; the witness adequacy functional $\Lambda(w)$ quantifies how tightly the initial condition constrains the future, and the diversity of PHYSIFORMER's generated trajectories is a direct readout of high Λ. Objecthood is historical coherence rather than a primitive label: the coherence classes $[i]_H$ are the objects, and PHYSIFORMER's generalization to unseen object counts is evidence that the model has learned to detect coherence rather than to represent predetermined objects. The optionality field $\Omega(x, t)$ extends the model's outputs beyond trajectory coordinates to a geometrically grounded characterization of how constrained the future has become at each moment.

Taken together, these claims constitute a single thesis, which is the cleanest way to characterize what PHYSIFORMER does when interpreted through these frameworks:

PHYSIFORMER is a learned repair operator on admissible physical histories.

It is an empirical instance of the larger claim that worlds are not best understood as static states plus transition rules, but as persistent histories whose futures are reconstructed, repaired, constrained, and collapsed from partial witnesses. In Computational Worlds terms, the mesh vertices are the world-state registers, the material labels are permission constraints, and the generated trajectory is the world-history, reconstructed in a single pass rather than ticked forward. In the History Before State program, the primary output is the trajectory H, not any intermediate state, and state access is projection $X_t = \pi_t(H)$. In repair theory, the denoising loop is iterative repair $\tilde{H} \to R(\tilde{H})$ converging to the admissible manifold. In admissibility theory, the distribution $p(H \mid X_0, V_0, m)$ is a distribution over $\mathcal{A}(X_0, V_0, m)$, not over noise. In the Geometry of Witnesses, the initial condition is a witness of bounded adequacy $\Lambda(w)$ from which the continuation space is reconstructed rather than stored. In RSVP, the model currently learns the coordinate shadow of field dynamics; the extension to field prediction would make the causal structure explicit.

The Admissibility Compression Principle connects all of these to a single underlying reason why the approach works: the admissible trajectory manifold $\mathcal{M}$ is vastly lower-dimensional than the ambient history space $\mathbb{R}^{3NT}$. Physical worlds are massively overconstrained. The set of coherent histories is so thin relative to the set of possible histories that a model can learn to detect and repair admissibility violations without ever learning the laws that define admissibility. The generalization to unseen geometries confirms this: a cube and a bunny are different objects but the same admissibility class, and the model has learned the class, not the object. The failure modes at contact boundaries confirm it from the other direction: contacts are the high-curvature regions of $\mathcal{M}$ where the compression breaks down and the manifold is hardest to approximate.

The deepest lesson is therefore not about diffusion models or physical simulation specifically. It is about the structure of coherent continuation in overconstrained worlds:

$$\text{Persistence} \rightarrow \text{Recoverability} \rightarrow \text{Repairability} \rightarrow \text{Prediction.}$$

Prediction works because coherent histories can be repaired from partial witnesses. PHYSIFORMER is a large-scale empirical machine for learning how physical histories repair themselves. The frameworks developed here show what that means precisely, and what the extensions look like.

The conclusion is not that PHYSIFORMER has solved physical simulation. The paper itself is clear about the model's limitations: interpenetrations, orientation discontinuities, and fixed trajectory length indicate that the repair operator does not yet achieve perfect convergence to the admissible manifold. The conclusion is that the architecture implicitly demonstrates a principle that deserves to be explicit: a world model need not represent physical truth. It need only construct admissible futures.

References

[1] Chen, Y., Lan, Y., and Vedaldi, A. PHYSIFORMER: Learning to Simulate Mechanics in World Space. *arXiv preprint*, arXiv:2606.27364, 2026.

[2] Battaglia, P. W., Pascanu, R., Lai, M., Rezende, D. J., and Kavukcuoglu, K. Interaction Networks for Learning about Objects, Relations and Physics. *Advances in Neural Information Processing Systems*, 29, 2016.

[3] Pfaff, T., Fortunato, M., Sanchez-Gonzalez, A., and Battaglia, P. W. Learning Mesh-Based Simulation with Graph Networks. *International Conference on Learning Representations*, 2021.

[4] Sanchez-Gonzalez, A., Godwin, J., Pfaff, T., Ying, R., Leskovec, J., and Battaglia, P. W. Learning to Simulate Complex Physics with Graph Networks. *International Conference on Machine Learning*, 2020.

[5] Shao, Y., Loy, C. C., and Dai, B. Transformer with Implicit Edges for Particle-Based Physics Simulation. *European Conference on Computer Vision*, 2022.

[6] Peebles, W., and Xie, S. Scalable Diffusion Models with Transformers. *International Conference on Computer Vision*, 2023.

[7] Lipman, Y., Chen, R. T. Q., Ben-Hamu, H., Nickel, M., and Le, M. Flow Matching for Generative Modeling. *arXiv preprint*, arXiv:2210.02747, 2022.

[8] Ho, J., Jain, A., and Abbeel, P. Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems*, 33, 6840–6851, 2020.

[9] Todorov, E., Erez, T., and Tassa, Y. MuJoCo: A Physics Engine for Model-Based Control. *International Conference on Intelligent Robots and Systems*, 2012.

[10] Allen, K. R., Rubanova, Y., Lopez-Guevara, T., Whitney, W., Sanchez-Gonzalez, A., Battaglia, P., and Pfaff, T. Learning Rigid Dynamics with Face Interaction Graph Networks. *arXiv preprint*, arXiv:2212.03574, 2022.

[11] Chen, B., Monso, D. M., Du, Y., Simchowitz, M., Tedrake, R., and Sitzmann, V. Diffusion Forcing: Next-Token Prediction Meets Full-Sequence Diffusion. *Advances in Neural Information Processing Systems*, 2024.

[12] Li, Y., Wu, J., Tedrake, R., Tenenbaum, J. B., and Torralba, A. Learning Particle Dynamics for Manipulating Rigid Bodies, Deformable Objects, and Fluids. *International Conference on Learning Representations*, 2019.

[13] Kabsch, W. A Solution for the Best Rotation to Relate Two Sets of Vectors. *Acta Crystallographica*, 32(5), 922–923, 1976.