The Catalan Light Cone: A Recursive Substrate for Causal Geometry, Quantum Amplitudes, and Computation

Paul Fernandez

Abstract

We investigate the Catalan family of combinatorial structures—Dyck paths, full binary trees, and balanced parenthesis expressions—as a unified discrete substrate from which causal geometry, quantum amplitude propagation, and universal computation jointly emerge.

A central observation is that the Dyck constraint induces a natural causal order. When organized by growth tier and lateral spread, the Catalan lattice forms a discrete cone whose extremal configurations reproduce a light-cone–like causal envelope. Classical invariance-principle results show that constrained Dyck walks converge in the scaling limit to Brownian excursions; the induced probability flow satisfies the heat equation and, under Wick rotation, the free Schrödinger equation.

Via a uniform structural mapping—the pairs expansion—Dyck trees are placed in bijection with SKI and λ-calculus term graphs. Under this identification, causal extension corresponds to functional application, while local collapse corresponds to computational reduction. Structural equivalences of the substrate induce gauge-like redundancies, and disjoint subtrees commute analogously to spacelike-separated operators.

The paper distinguishes rigorously established correspondences—scaling limits, diffusion dynamics, prefix-causal structure, and computational universality in the sense of standard SKI/λ encodings—from conjectural extensions concerning measurement, actualization, and interaction structure.

Taken together, these correspondences show that a single recursive constraint can reproduce, at a structural and kinematical level, large portions of the operational framework of relativistic quantum theory and universal computation, without introducing additional primitives. We do not attempt to derive interactions or physical constants; the aim is to isolate the minimal recursive structure common to these domains.

1 Introduction

Discrete approaches to fundamental physics have long suggested that continuum spacetime and quantum dynamics may emerge from deeper combinatorial structure. Examples include causal sets [5], discrete random surfaces and Causal Dynamical Triangulations (CDT) [2, 3], spin networks and loop quantum gravity [16], tensor networks [15], and rewriting systems inspired by λ-calculus and combinatory logic.

Typically, however, these models require multiple independent ingredients: a relation or graph for causal structure, an algebra for computation, and additional rules for quantum propagation. This work explores a more economical possibility: that a *single* recursive structure simultaneously supports all three.

The focus is the *Catalan substrate*, the family of structures counted by the Catalan numbers [18], including Dyck paths, full binary trees, and balanced parenthesis expressions. These objects are usually studied in enumerative combinatorics, probability theory, and theoretical computer science. Here they are treated instead as a space of *admissible histories* generated by a minimal growth constraint. The central claim is that the Catalan substrate admits three tightly coupled interpretations:

(i) a discrete causal geometry with a light-cone–like envelope,

(ii) a natural path-integral dynamics with diffusive and wave-like continuum limits,

(iii) a universal computational calculus via λ- and SKI-term graphs.

These interpretations do not require distinct primitives; they arise from different readings of the same recursive object.

The geometric aspect follows from prefix order and growth bounds intrinsic to Dyck paths. The dynamical aspect follows from classical results on conditioned random walks and Brownian excursions [13, 11]. The computational aspect follows from the well-known equivalence between binary trees, cons-pair structures, and λ-calculus or SKI combinators [6, 9, 4]. Taken together, these results show that spacetime-like causal structure, quantum wave dynamics, and computation can be viewed as complementary manifestations of a single recursive substrate.

The exposition proceeds as follows. Section 2 establishes the discrete causal geometry of the Catalan lattice and its interpretation as a light cone. Subsequent sections place amplitude propagation and computation on this structure, analyze continuum limits, and discuss collapse, locality, and interaction. Interpretive considerations concerning origin, vacuum structure, and time are collected in a clearly labeled appendix, separate from the formal claims.

2 The Catalan Light Cone as a Discrete Causal Geometry

2.1 Dyck paths and growth tiers

A Dyck path of semilength n is a walk on the integers satisfying

$$H_{k+1} = H_k \pm 1, \qquad H_k \geq 0, \qquad H_0 = H_{2n} = 0.$$

Equivalently, Dyck paths are balanced parenthesis strings or full binary trees with n internal nodes. The number of such paths is the nth Catalan number

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Each up–down pair () represents a minimal unit of growth. The integer

$$t = n$$

will be called the *tier* and will be interpreted as a discrete proper time.

2.2 Prefix order and causality

Dyck paths carry a natural partial order by prefix inclusion. If a Dyck word u is a prefix of v, then u represents a causal ancestor of v. Conversely, prefixes that diverge represent causally incompatible futures. This prefix order defines a discrete causal structure:

- every node has a unique causal past,

- multiple incompatible futures may branch from the same prefix,
- cycles are prohibited by construction.

No additional causal axiom is required; causality is enforced combinatorially by the Dyck constraint.

2.3 Extremal configurations: chain and star

At fixed tier t there are many Dyck paths. Two extremal configurations play a distinguished role:

- the *chain* (or spine)
$$(((\cdots))),$$
fully nested, with maximal depth and minimal spread;
- the *star*
$$()()()\cdots(),$$
fully separated, with minimal depth and maximal spread.

All other configurations interpolate between these extremes. Together, the set of Dyck paths at tier t forms a discrete envelope bounded by the chain and the star.

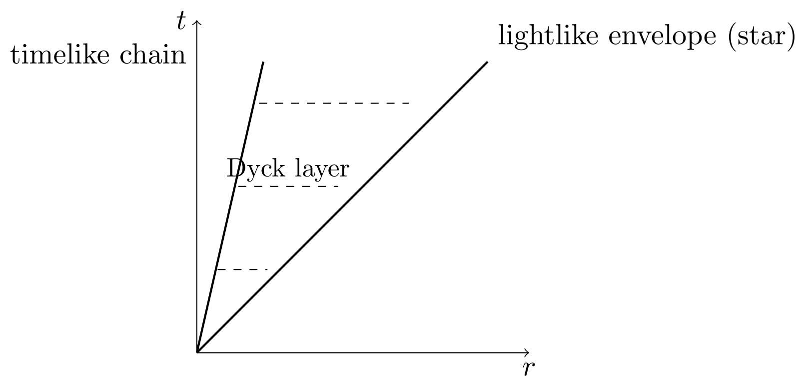


Figure 1: The Catalan light cone. Tier t (the number of Dyck units) plays the role of proper time, while breadth r measures spatial radius. All Dyck configurations at fixed tier lie between the fully nested chain (timelike extreme) and the fully separated star (lightlike envelope). Discrete Dyck layers approximate constant-time hypersurfaces, and the bound $r \leq t$ is enforced combinatorially.

2.4 Breadth as spatial extent

Define the *breadth* $r(w)$ of a Dyck path w to be the size of a largest level set in the associated full binary tree:
$$r(w) := \max_{\ell}\{\text{number of nodes at depth } \ell\}.$$

Equivalently, $r(w)$ is the maximal number of non-overlapping pairs at a common nesting depth. For a Dyck word of tier t there are t matched pairs in total, so
$$1 \leq r(w) \leq t,$$

with $r = 1$ for the fully nested chain and $r = t$ for the fully separated star. The inequality

$$r \leq t$$

is enforced purely by the recursive constraint. It is the discrete analogue of the relativistic light-cone bound $|\Delta x| \leq \Delta t$ (in units with $c = 1$).

2.5 Depth–breadth tradeoff

Let $h(w)$ denote the maximum height of a Dyck path, i.e. its maximal nesting depth. Depth and breadth are not independent. Interpreting the Dyck tree as a full binary prefix code, assign to each leaf i its depth d_i (root at depth 0). Then the Kraft equality holds:

$$\sum_i 2^{-d_i} = 1.$$

In particular, at any fixed depth ℓ there can be at most 2^ℓ leaves at that depth. Since $r(w)$ is the size of a largest level set, there is some depth ℓ_* at which $r(w)$ nodes appear, and hence

$$r(w) \leq 2^{\ell_*} \leq 2^{h(w)} \quad \Rightarrow \quad h(w) \geq \log_2 r(w).$$

Thus configurations with large breadth necessarily have logarithmically large depth, while very deep trees must concentrate most of their leaves in narrow antichains. This structural constraint is a combinatorial analogue of the tension between spatial spread and temporal commitment and is the standard Kraft bound for prefix codes [7].

$$\begin{array}{ll}
((())) & (h = 3,\ r = 1) \\
(()()) & (h = 2,\ r = 2) \\
(())() & (h = 2,\ r = 2) \\
()(()) & (h = 2,\ r = 2) \\
()()() & (h = 1,\ r = 3)
\end{array}$$

Figure 2: All Dyck words of tier $n = 3$, ordered from maximal nesting (chain) to maximal separation (star). These five configurations exhaust the discrete causal possibilities at fixed proper time. Depth h and breadth r interpolate between the two extremes, illustrating the intrinsic tradeoff enforced by the Dyck constraint. Higher tiers replicate this structure at larger scale.

2.6 Cone structure

Organizing Dyck paths by tier t and breadth r yields a discrete cone:

- each tier is a "constant-time" slice,
- the chain defines the timelike axis,
- the star defines the lightlike boundary,
- admissible configurations fill the interior.

This structure will be referred to as the *Catalan light cone*.

2.7 Scaling behavior

Classical results on conditioned random walks show that typical Dyck paths at tier t have height and breadth of order $\sqrt{t}$ [13, 11, 1]. Extremal configurations saturate the linear bound $r \leq t$, while typical configurations lie deep within the cone. This separation between extremal and typical behavior mirrors the role of null, timelike, and spacelike trajectories in relativistic geometry.

Theorem 2.1 (Discrete light-cone bound and scaling)**.** *Let w be a Dyck word of semilength t and breadth $r(w)$ as above. Then*

$$1 \leq r(w) \leq t.$$

Moreover, for a uniformly random Dyck word of semilength t, the typical height and breadth are of order $\sqrt{t}$.

The first statement follows from the definition of $r(w)$ and the fact that there are t internal nodes, while the scaling behavior is a consequence of invariance-principle results for conditioned random walks [13, 11, 1].

2.8 Dyck Coordinates, Lorentz Geometry, and Computational Proper Time

The Dyck encoding places each Catalan object into a discrete $(1+1)$-dimensional causal geometry. A Dyck path of semilength n is a walk composed of steps

$$(\Delta x, \Delta t) = (\pm 1, 1),$$

subject to the non-negativity constraint $h(t) \geq 0$ with $h(0) = h(2n) = 0$. This forces every admissible history to lie within the discrete light cone

$$|x| \leq t, \qquad t \in \{0, 1, \ldots, 2n\}.$$

Under diffusive rescaling, Dyck paths converge to Brownian excursions (Section 4.6), and the cone $|x| \leq t$ becomes the standard Minkowski light cone in the continuum limit. Thus the Dyck parametrization already expresses the (t, x) structure of a relativistic frame.

Light-cone coordinates. Introduce the usual light-cone coordinates

$$u = t + x, \qquad v = t - x, \tag{1}$$

in which a Dyck step increments exactly one of (u, v) by 2, while the height constraint becomes simply $u, v \geq 0$. A Lorentz boost with rapidity η acts linearly as

$$u' = e^{\eta} u, \qquad v' = e^{-\eta} v. \tag{2}$$

Transforming back to (t, x) gives the familiar Lorentz transformation

$$t' = \gamma(t - v_L x), \qquad x' = \gamma(x - v_L t), \tag{3}$$

where

$$\gamma = \frac{1}{\sqrt{1 - v_L^2}}, \qquad v_L = \tanh \eta,$$

with $c = 1$ units understood. The Minkowski interval

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \mathrm{d}x^2 \tag{4}$$

is invariant under (3) and may be regarded as the continuum limit of the discrete metric induced by Dyck-constrained paths.

Computational proper time. A reduction history carries more than geometric information: it has an intrinsic *computational* progress parameter. Let k denote the number of collapse events (local redex contractions) performed along the evaluation of a Catalan object. Passing to a continuum description, introduce a computational proper time τ via

$$\tau = \alpha\, k, \tag{5}$$

where α is the characteristic scale associated with a single collapse. In the continuum limit, for a parametrized world-line $(t(s), x(s))$, the proper time satisfies

$$\frac{\mathrm{d}\tau^2}{\mathrm{d}s^2} = \left(\frac{\mathrm{d}t}{\mathrm{d}s}\right)^2 - \left(\frac{\mathrm{d}x}{\mathrm{d}s}\right)^2, \tag{6}$$

so that τ is the Lorentz-invariant "computational length" of the causal trajectory. The discrete relation (5) is therefore the microscopic counterpart of (6), with each local collapse contributing one quantum of proper computational time.

In this interpretation, a Dyck path encodes not only the (t, x) geometry of the light cone but also the admissible orderings of collapse events. A particular evaluation strategy corresponds to selecting which admissible steps count as computational advances in τ; that is, a choice of a *computational world-line* within the Catalan cone. Thus the Dyck bijection expresses, in discrete form, the same invariants that appear in continuous Lorentzian geometry, together with an additional computational invariant associated with the collapse dynamics of the Catalan substrate.

2.9 Recursive Self-Similarity and Local Re-Centering

A key structural property of the Catalan substrate is its *recursive self-similarity.* Every Dyck word may be viewed as a node in the infinite prefix tree of admissible Dyck prefixes. At any such node u, with current height h and remaining length budget sufficient to return to height 0, the set of all admissible continuations of u forms a subtree whose shape is determined entirely by h. This subtree is canonically isomorphic to the Dyck-prefix tree that begins at height h rather than at height 0.

Formally, if $\mathcal{C}$ denotes the infinite Dyck-prefix tree and $\mathcal{C}_h$ denotes the Dyck-prefix tree conditioned to start at height h (i.e. with $H_0 = h$ and $H_k \geq 0$ for all k), then for every prefix u of height h we have a canonical isomorphism

$$\mathcal{C}(u) \;\cong\; \mathcal{C}_h.$$

Thus every node of the global Catalan possibility tree is the root of a scaled copy of the entire admissible-future structure, with scaling determined solely by local height. The recursive decomposition of full binary trees,

$$T = \bullet(T_L, T_R),$$

makes the same fact explicit in the tree representation: each subtree of a Catalan tree is itself a Catalan tree, and the decomposition applies inductively at every depth.

This recursive self-similarity has two important consequences for the geometric interpretation developed in this paper:

(i) **Locality and re-centering.** Because the admissible future of any prefix depends only on its present height, not on its global position, the Catalan light-cone geometry is *locally homogeneous*. The causal cone may be re-centered at any node without altering its shape: moving the focus does not change the structure of admissible futures, only the value of the local height at which the cone is rooted.

(ii) **Scale invariance of the substrate.** The same recursive rules govern growth at every depth. The local possibility space looks the same at all scales, in the sense that the subtree below any node is again Catalan. This is the combinatorial source of the invariance principles (Dyck $\to$ Brownian excursion) appearing in the continuum limit.

In summary, the Catalan substrate is self-similar at every node: each point in the possibility space contains a full Catalan future scaled by its current height. This allows the causal and geometric analysis of later sections to be performed relative to *any* node of the prefix tree. The light cone is not anchored to a global origin; it is an intrinsic, relocatable geometric feature of the recursive structure itself.

2.10 Multiple Local Cones and Relational Geometry

The recursive self-similarity of the Catalan substrate implies that there is not a single distinguished light cone rooted at the global origin. Instead, every node of the Dyck-prefix tree induces its own local notion of past, future, and lightlike boundary. This section records the combinatorial foundations of this phenomenon and its geometric consequences.

Cones rooted at arbitrary prefixes

Let u be any Dyck prefix with current height h. As shown in Section 2.9, the subtree $\mathcal{C}(u)$ of admissible continuations of u is canonically isomorphic to the Dyck-prefix tree $\mathcal{C}_h$ that begins at height h. Consequently, the structure of possible futures below u is itself a Catalan possibility space.

Define the *local cone at* u to be the set of all Dyck extensions of u, organized by length (tier) and breadth relative to u. The local chain is the fully nested extension of u, and the local star is the fully separated extension. These play the role of timelike and lightlike extremes for admissible growth below u.

Nested and divergent cones

Let u and v be Dyck prefixes.

(i) **Nested cones.** If u is a prefix of v, then $\mathcal{C}(v)$ is a subtree of $\mathcal{C}(u)$, and the local cone at v lies strictly inside the local cone at u. Their causal structures satisfy

$$\mathrm{Past}(v) \subset \mathrm{Past}(u), \qquad \mathrm{Future}(v) \subset \mathrm{Future}(u).$$

Thus cones nest hierarchically along any branch of the prefix tree.

(ii) **Divergent cones.** If neither prefix is contained in the other, then u and v share a common ancestor but diverge at some minimal prefix w. Their cones therefore have a shared causal past (the future of w up to u and v) but incompatible futures beyond that point. Formally,

$$\mathrm{Past}(u) \cap \mathrm{Past}(v) = \mathrm{Past}(w), \qquad \mathrm{Future}(u) \cap \mathrm{Future}(v) = \varnothing.$$

In this sense, divergent cones encode distinct branches of the Catalan possibility structure.

Relational geometry

These observations establish a relational geometric structure intrinsic to the Catalan substrate:

- every node induces its own local cone of admissible extensions;
- cones may be re-centered without changing their internal geometry;
- cones nest along causal chains and diverge after branching points;
- no global origin is privileged—only prefix order determines causal relationships.

This yields a family of interacting local cones, each encoding the admissible future relative to a chosen prefix. The Catalan light cone is therefore not a single global object but a *relocatable geometric feature* that appears at every node of the recursive substrate.

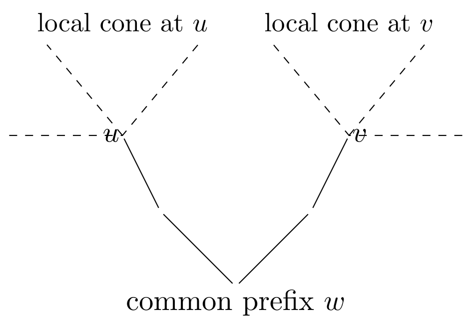


Figure 3: Two Dyck prefixes u and v diverging from a shared ancestor w. Dashed regions indicate the local Catalan cones rooted at u and v. Cones nest along causal chains and diverge after branching points, producing a family of local, relocatable causal geometries on the Catalan substrate.

2.11 Summary

The Catalan substrate supports a discrete causal geometry determined entirely by recursive constraint. Without introducing a manifold, metric, or causal relation by fiat, it yields:

- a partial order interpretable as causality,
- a cone-shaped causal envelope,
- intrinsic bounds on spatial extension,
- well-defined constant-time layers.

Subsequent sections place dynamical rules—quantum amplitudes and computational reduction—on this geometry.

3 Recursive Pairing and Universal Computation

3.1 Pairs expansion

Catalan Structure as the Space of Program Possibilities. The Catalan family—Dyck paths, full binary trees, and balanced-parenthesis expressions—forms the free magma on a single binary constructor. As observed in foundational work on the λ-calculus and combinatory logic [8, 4], every computable program has a canonical representation as a finite binary application tree: internal nodes encode application; leaves encode atomic symbols or combinators. Conversely, any finite binary tree equipped with leaf labels denotes a unique program modulo surface syntax. Thus the infinite Catalan tree is not merely a combinatorial object but the structural possibility space of all programs expressible in any Turing-complete functional calculus.

This observation also extends to operational semantics. Standard reductions (such as β-reduction or SKI contraction) are local rewrite rules on binary trees, and the pairs-expansions of combinators remain within the Catalan family. Accordingly, a program, its intermediate expansion frames, and each permissible reduction schedule are all representable as paths through a single Catalan substrate. Selecting a program shape or selecting a specific reduction history is therefore equivalent to selecting a path in the Catalan tree. In this sense the Catalan substrate uniformly encodes program syntax, program semantics, and the full ensemble of admissible computational histories. A formal treatment of these embeddings, together with illustrative examples and the associated rewrite semantics, is provided in Appendix A.

Every Dyck path admits a unique decomposition into nested pairs. Writing parentheses explicitly, the simplest nontrivial closure of the empty expression () is

$$(()),$$

which contains a single internal pairing. Iterating this rule produces the entire Catalan family.

This recursive pairing induces a uniform transformation, referred to here as the *pairs expansion*, which places Dyck trees in bijection with binary application trees. Under this correspondence:

- each matched pair corresponds to a cons cell,
- containment corresponds to functional application,
- sibling subtrees correspond to arguments at equal precedence.

The transformation is purely structural and preserves prefix order.

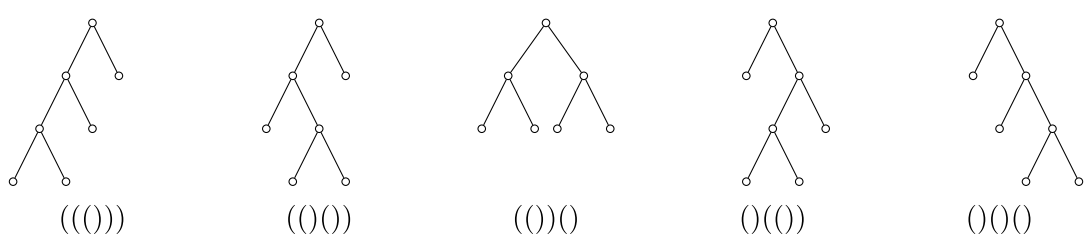


Figure 4: Binary-tree representations of the five Dyck words of semilength 3, corresponding to Fig. 2. From left to right: maximal nesting (chain) through mixed cases to maximal separation (star).

3.2 Connection to λ-calculus and SKI

Binary trees are a standard representation of λ-terms and SKI combinators [6, 9, 4]. Variables may be represented by leaf positions, abstraction by structural capture, and application by tree composition. Under the pairs expansion, each Dyck tree canonically determines an unlabeled application graph. When variables are suppressed, the resulting graphs coincide with the structure graphs used in combinatory logic. No additional primitives beyond recursive pairing are required to obtain this representation.

Choosing a finite set of tree patterns to represent the SKI combinators and interpreting local tree rewrites as SKI reduction therefore equips the Catalan substrate with a standard universal calculus: every partial recursive function can be encoded as an SKI term, and hence by a finite Dyck tree, and every computation corresponds to a sequence of local tree transformations. In this sense, the Catalan substrate is *computationally universal.* What is new here is that the same underlying objects simultaneously carry a causal and geometric interpretation.

3.3 Reduction and local collapse

In the computational interpretation, reduction corresponds to local pattern replacement. A redex occupies a finite region of a tree and may be reduced without reference to distant subtrees. This locality mirrors the causal structure established in Section 2. From the perspective of the Catalan lattice, reduction may be viewed as *collapse*: a locally ambiguous structure is replaced by a simpler one consistent with the global constraint. Importantly, collapse does not alter causal ancestry; it refines an already-admissible history. Confluence of reduction in the λ-calculus ensures that independent local reductions commute. This computational fact will later support an interpretation of spacelike commutativity.

3.4 Summary

Recursive pairing suffices to encode universal computation. Via the pairs expansion, Dyck trees and application graphs are two views of the same structure. Local computational reduction aligns naturally with causal locality on the Catalan light cone.

4 Quantum Amplitudes on the Catalan Lattice

4.1 Histories as paths

Interpreting Dyck paths as admissible histories suggests assigning weights to each history. Let $\mathcal{D}_t$ denote the set of Dyck paths of tier t. A state at tier t may be represented as a formal superposition

$$\Psi_t = \sum_{w \in \mathcal{D}_t} \psi(w) \, |w\rangle.$$

Local extensions of a Dyck path correspond to admissible future steps. Thus, time evolution is governed by transitions that respect the Dyck constraint.

4.2 Observables, projection, and coherent summation

Fix a tier t and consider the set $\mathcal{D}_t$ of Dyck paths of semilength t. Each $w \in \mathcal{D}_t$ represents a complete admissible history at discrete time t, with an associated height profile

$$H_w : \{0, 1, \ldots, 2t\} \to \mathbb{Z}_{\geq 0}.$$

An observable is defined as a deterministic coarse-graining

$$f : \mathcal{D}_t \to \mathcal{X},$$

where $\mathcal{X}$ is a discrete set of outcomes (e.g. screen bins or rings). A detected outcome $x \in \mathcal{X}$ corresponds to the equivalence class $f^{-1}(x) \subset \mathcal{D}_t$. By construction, such a projection discards information: many distinct histories may correspond to the same observed result.

With no additional structure imposed, the natural measure on $\mathcal{D}_t$ is uniform counting. The induced intensity on $\mathcal{X}$ is therefore the pushforward

$$N(x) := \#\{w \in \mathcal{D}_t : f(w) = x\}.$$

Even with uniform weight on histories, the induced distribution on $\mathcal{X}$ is generically non-uniform, reflecting the combinatorial geometry of the projection rather than any imposed dynamics.

A discrete analogue of an integral along a history is given by the step-sum of the height profile,

$$A(w) := \sum_{k=0}^{2t-1} H_w(k),$$

which measures the cumulative dwell time at nonzero height. This quantity depends on the full distribution of height along the path, not merely on extrema such as maximum height or peak count.

From this additive structural functional one may define a complex phase

$$\theta(w) := \alpha\, A(w), \qquad \psi(w) := e^{i\theta(w)},$$

where α is a global scale parameter. No per-path phase assignment is introduced; distinct phases arise solely from differences in the distribution of height over the history.

Given an observable f, the complex amplitude associated with an outcome $x \in \mathcal{X}$ is the coherent sum

$$\Psi(x) := \sum_{w:\, f(w)=x} \psi(w),$$

and observed probabilities are obtained by normalization of squared magnitudes,

$$P(x) = \frac{|\Psi(x)|^2}{\sum_{x' \in \mathcal{X}} |\Psi(x')|^2}.$$

Thus histories that are indistinguishable under the observable f are combined prior to squaring, while distinguishable histories are not. Interference is therefore not postulated but forced by the structure of projection.

4.3 Path integrals and conditioned walks

Dyck paths are random walks conditioned to remain nonnegative and return to zero. Classical results show that, when rescaled appropriately, ensembles of such paths converge to Brownian excursions [13, 11]. Assigning equal weight to all admissible paths yields a discrete path integral. More general amplitude assignments may depend on local features such as height or curvature, provided the Dyck constraint is preserved.

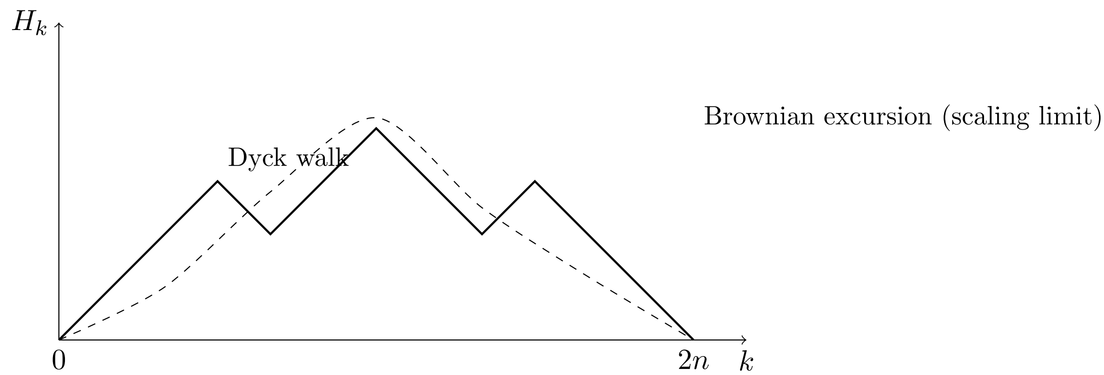


Figure 5: A Dyck path as a nearest-neighbour walk (H_k) constrained to stay nonnegative and return to zero at time $2n$. Under diffusive rescaling of k and H_k, ensembles of such paths converge to Brownian excursions, providing the bridge to the heat and Schrödinger equations discussed in the text.

4.4 Modal structure and Narayana classes

Beyond radial observables derived from global statistics such as height or area, Dyck paths admit a natural discrete modal decomposition. For a Dyck path $w \in \mathcal{D}_t$, define its *Narayana index* $K(w)$ to be the number of peaks of w, i.e. the number of occurrences of an up-step immediately followed by a down-step. For fixed t, the Narayana numbers $N(t, k)$ count the number of Dyck paths with $K(w) = k$.

Combinatorially, the Narayana index measures the number of oscillations of the height signal $H_w(k)$ along the path. Paths with the same K therefore share a common oscillatory structure but may differ in the relative placement of their peaks along the word. This induces a refinement of the path ensemble into mode classes indexed by k.

At fixed tier t, these mode classes are orthogonal to radial observables such as area or mean height. Distinct Narayana classes may contribute to the same screen outcome under a given projection, while paths within a fixed Narayana class may differ in the distribution of height over time and hence in the structural phase defined by $A(w)$. Thus the Narayana index plays a role analogous to a discrete spatial frequency, while the relative placement of peaks provides a phase offset within that mode.

This decomposition clarifies how interference structure arises prior to any continuum limit. Radial observables determine which coarse region of the screen a history contributes to, while Narayana classes and peak placement govern how those contributions interfere. In the scaling limit discussed below, these discrete mode classes converge to the familiar harmonic structure of continuum wave equations.

4.5 Discrete path-integral formulation

The preceding constructions admit a direct interpretation as a discrete path integral on the Catalan light cone. Fix a tier t and an observable $f : \mathcal{D}_t \to \mathcal{X}$. Each Dyck path $w \in \mathcal{D}_t$ represents a complete admissible history, and the projection f determines which histories are experimentally indistinguishable.

Define a complex weight for each history by

$$\psi(w) = e^{i\alpha A(w)},$$

where

$$A(w) = \sum_{k=0}^{2t-1} H_w(k)$$

is the discrete height integral introduced above. The amplitude associated with an outcome $x \in \mathcal{X}$ is then

$$\Psi(x) = \sum_{w:\, f(w)=x} e^{i\alpha A(w)}.$$

This expression is formally identical to a path integral: the amplitude is a sum over all admissible histories consistent with the observable, with each history contributing a phase determined by an additive functional of the path. No continuum limit, action functional, or variational principle is assumed at this stage; the structure arises purely from discrete combinatorics.

Several features of continuum path integrals are already present:

(i) **Sum over histories.** All admissible Dyck paths consistent with the observable contribute. The Dyck constraint enforces causality in the same way that the restriction to timelike or null paths does in relativistic path integrals.

(ii) **Additive phase functional.** The quantity $A(w)$ is additive under concatenation of path segments and depends only on local height data. It therefore plays the role of a discrete action accumulated along the history.

(iii) **Coarse-grained interference.** Interference occurs precisely because multiple distinct histories map to the same observable outcome. If the projection f distinguishes two histories, their contributions are not summed; if it does not, they interfere.

(iv) **Gauge redundancy.** Different Dyck paths may correspond to the same abstract computation or the same coarse-grained geometry, differing only by the ordering of spacelike-separated updates. The path integral sums coherently over such gauge-equivalent histories, while physical probabilities depend only on the resulting amplitudes.

From this perspective, the Catalan lattice provides a discrete realization of Feynman's sum-over-histories principle in which both the space of histories and the phase functional are combinatorially well-defined. In the next subsection we show that, under appropriate scaling limits, this discrete path integral converges to the familiar continuum descriptions governed by diffusion and Schrödinger dynamics.

4.6 Diffusion limit

Let $n \to \infty$ and rescale time and height by

$$t \mapsto n\tau, \qquad h \mapsto \sqrt{n}\, x.$$

Under this scaling, the probability density $\rho(\tau, x)$ for conditioned walks converges to the density of a Brownian excursion on $x \geq 0$. It follows from the invariance principle that ρ satisfies the heat equation

$$\partial_\tau \rho = \frac{1}{2}\, \partial_x^2 \rho, \tag{7}$$

with boundary conditions enforcing reflection or absorption at $x = 0$. Full derivations may be found in [13, 11].

4.7 Schrödinger equation

More formally, if $\rho(\tau, x)$ denotes the real heat kernel on $x \geq 0$, analytic continuation in the diffusion parameter, $\tau \mapsto it$, produces a complex-valued kernel $\psi(t, x)$ satisfying the free Schrödinger equation

$$i\partial_t \psi = -\frac{1}{2}\,\partial_x^2 \psi. \tag{8}$$

Boundary conditions at $x = 0$ are carried over from the diffusive regime (e.g. reflecting or absorbing), and the choice of boundary does not affect the existence of the continuum limit itself. Thus, quantum wave dynamics arises here as the analytic continuation of diffusive propagation on the Catalan lattice, in line with the classical connection between diffusion and Schrödinger evolution [10, 12]. No separate quantization procedure is required; the wave equation is inherited from the scaling limit of constrained combinatorial growth.

4.8 Relation to discrete quantum gravity

Similar scaling behavior appears in two-dimensional quantum gravity and random surface models. In particular, Causal Dynamical Triangulations (CDT) enforce a preferred foliation and causal constraint that parallels the prefix order of Dyck paths [2, 3]. In CDT, the continuum limit is taken after summing over causally admissible triangulations. Here the admissible structures are Dyck paths rather than triangulations, but the organizing principle—restriction to histories that respect a causal growth rule—is the same.

5 Locality, Commutation, and Interaction

5.1 Disjoint subtrees

Two subtrees of a Dyck tree that share no common ancestor beyond a given prefix are causally independent. Operations localized to one subtree do not affect the other. In the computational interpretation, this corresponds to independent reductions. In the amplitude interpretation, it corresponds to commuting operators acting on spacelike-separated regions. Multiple Dyck paths may represent the same abstract computation or the same coarse-grained geometry. Such redundancies may be quotiented out without changing observable predictions, yielding equivalence classes of histories. This redundancy plays a role analogous to gauge symmetry: distinct internal descriptions correspond to the same external behavior.

Operationally, gauge here means a redundancy in the description of histories. For a fixed initial tree, consider the class of causal histories that apply the same multiset of local reductions but differ only in the temporal ordering of reductions supported on disjoint subtrees. Such histories are gauge-equivalent: they represent the same physical pattern of events, and their linear orderings are related by commuting spacelike-separated updates. The resulting equivalence classes (gauge orbits) play the role of physical histories, while genuinely different choices of which reductions to actualize belong to distinct orbits and define the nontrivial branching structure of the multiway reduction graph.

(See Appendix A, and in particular Lemma A.1, for a formal statement and proof sketch of the corresponding causal-consistency and commutation property.)

5.2 Collapse and selection

Both computation and amplitude propagation require selection:

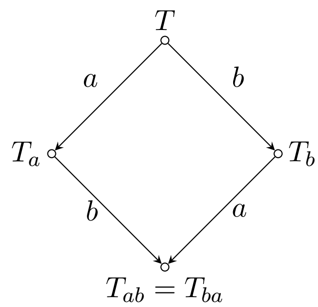


Figure 6: Local diamond for commuting disjoint updates. Starting from a common state T, two spacelike-separated reductions a and b can be applied in either order, yielding intermediate states T_a and T_b but the same final state $T_{ab} = T_{ba}$. The two intermediate events are spacelike-separated: they share a common past (T) and a common future (T_{ab}) but no causal edge between them. This expresses microcausality on the Catalan substrate: local updates supported on disjoint subtrees commute and differ only by a gauge choice of temporal ordering.

- computational reduction chooses a redex,
- measurement-like selection chooses an outcome.

In the Catalan substrate, selection operates locally, refining rather than destroying structure. The global constraint ensures consistency after selection. The formal development of collapse probabilities lies beyond the scope of this paper and is treated here only structurally.

5.3 Summary

Locality, commutation, and interaction emerge directly from the causal and recursive structure of the Catalan lattice. The same principles underlie both computational reduction and quantum amplitude propagation.

6 Discussion and Limitations

The results presented here establish a shared structural basis for causal geometry, quantum dynamics, and computation. Several limitations should be emphasized:

- Physical constants and interactions are not derived.
- Only free (noninteracting) wave dynamics appears explicitly.
- Collapse probabilities are not fixed uniquely by the structure.

These limitations reflect a deliberate restriction of scope. The goal has been to isolate the minimal recursive structure common to multiple domains, not to provide a complete physical theory.

7 Conclusion

This paper has examined the Catalan family of recursive structures as a common substrate for causal geometry, quantum amplitude propagation, and computation. The Dyck constraint induces a natural causal order and a discrete cone-shaped geometry exhibiting light-cone–like bounds.

Classical invariance principles show that ensembles of admissible histories converge in the continuum limit to Brownian excursions, yielding the heat equation and, under analytic continuation, the free Schrödinger equation. Through the pairs expansion, the same structures encode universal computation via λ-calculus and SKI combinators. These correspondences require no additional primitives beyond recursive pairing and constraint. Spacetime-like geometry, wave dynamics, and computation emerge as complementary aspects of a single recursive system. Several open problems remain, including the incorporation of interaction terms, the determination of collapse probabilities, and the connection to physical constants. The results presented here establish a minimal and structurally unified foundation upon which such extensions may be explored.

A Computational Foundations

This appendix collects the formal background supporting the use of the Catalan substrate as a universal space of program structures and computational histories. The material here provides precise statements and examples for readers wishing to verify the correspondence between full binary trees, combinatory calculi, and rewrite dynamics.

Proposition A.1 (Catalan Universality for Program Structure)**.** *Let $\mathcal{C}$ denote the Catalan family of finite full binary trees. Every program in any Turing-complete functional calculus (such as the λ-calculus or SKI) admits a canonical representation as an element of $\mathcal{C}$ with leaf labels drawn from a finite alphabet. Conversely, every labeled element of $\mathcal{C}$ denotes a unique program modulo surface syntax. Furthermore, standard operational semantics (including β-reduction and SKI contraction) act as local rewrite rules that preserve membership in $\mathcal{C}$. Thus a program, its syntactic expansions, and every admissible reduction history correspond to paths within the Catalan substrate.*

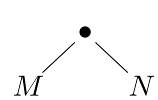
Proof Sketch. As described in classical treatments of combinatory logic and the λ-calculus [8, 4], application is a binary operation, and every term therefore possesses a unique representation as a full binary tree: internal nodes encode application, and leaves encode variables, constants, or combinators. This establishes a canonical embedding of programs into $\mathcal{C}$.

Conversely, any full binary tree with labeled leaves may be interpreted as a well-formed program term by reading internal nodes as applications and leaves as atomic symbols, yielding a unique term up to α-equivalence.

Operational semantics are defined via local tree rewrites. A β-redex $(\lambda x.M)\,N$ contracts by replacing the parent application with $M[x := N]$; SKI reductions replace specific subtrees according to fixed patterns. In each case, the output remains a full binary tree, so evaluation never leaves $\mathcal{C}$. Because nondeterministic redex choices correspond to branching in the space of trees, each complete reduction sequence is a path through the Catalan possibility space, completing the correspondence. $\square$

A.1 Binary Trees as Universal Program Frames

The identification of Catalan objects with program structures is a standard but essential foundation for the present work. Let $\mathcal{C}$ denote the set of full binary trees. A term in a functional calculus is constructed by repeated application, and an application $M\,N$ is represented by a binary node whose children encode M and N:

$(M\,N) \quad \mapsto$ 

This mapping is bijective between program terms (modulo α-conversion) and labeled elements of $\mathcal{C}$.

Pairs-expansions of combinators—for example the standard expansions of S and K into λ-terms—produce larger trees but remain within the Catalan family. Thus the Catalan substrate is closed under syntactic elaboration.

Operational semantics are likewise internal. Both β-reduction and SKI contraction replace subtrees with simpler subtrees while preserving the global full-binary-tree form. Each admissible reduction path therefore corresponds to a trajectory within $\mathcal{C}$, and nondeterminism in reduction strategy corresponds to branching structure within the Catalan tree itself.

This uniformity demonstrates that the Catalan substrate simultaneously captures:

1. program syntax (binary application structure),
2. program elaboration (via pairs-expansion or substitution), and
3. program dynamics (via evaluation rewrites).

Consequently, computation lives entirely within the Catalan family, justifying its use as the structural substrate for the unified causal–computational model developed in the main text.

A.2 Illustrative Figures

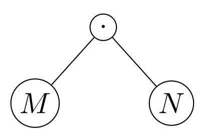


Figure 7: A binary application node representing the term $M\,N$.

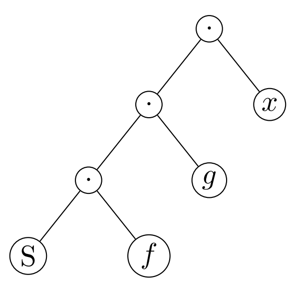


Figure 8: Binary-tree representation of the term $\mathrm{S}\,f\,g\,x$.

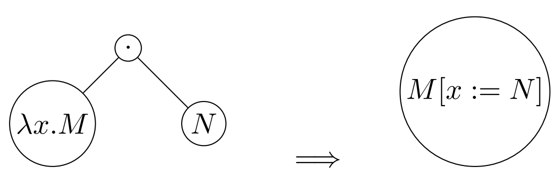


Figure 9: β-reduction as a local rewrite inside the Catalan family.

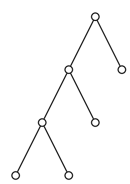
A.3 Pairs Expansion as Variable-Free S-Expressions

A central observation motivating this work is that the pairs expansions of combinators can be written as *variable- and label-free* S-expressions, exactly in the style of McCarthy's original Lisp notation [14]. In McCarthy's formulation, the core data structure is the cons-cell, written as a parenthesized pair. Here we push this idea to an extreme: we erase all atom labels and regard the program itself as a pure cons-tree, encoded only by balanced parentheses.

Concretely, consider the binary application tree for the term S $f\,g\,x$ (as in Figure 8). Under the pairs encoding used in our simulations, this same shape appears as the variable-free S-expression

`(()(()(()()))).`

This S-expression can also be understood as *looking up into* the underlying Catalan tree: the outer parentheses form the root frame, while each `()` corresponds to a leaf. The corresponding unlabeled binary tree shape is shown below.



Viewed from this perspective, the S-expression is simply the linear "parenthesis trace" of this structure: each "(" corresponds to descending into a cons-cell, each ")" corresponds to returning toward the trunk, and the empty pairs `()` mark the terminal leaves.

More generally, the pairs bijection

```
n=0, c=1:  ()
n=1, c=1:  (()())
n=2, c=2:  (()(()())) ((()())())
n=3, c=5:  (()(()(()()))) (()((()())())) ((()())(()())) ((()(()()))()) (((()())())())
```

enumerates exactly the same Catalan shapes that appear as Dyck paths

```
n=1, c=1:  ()
n=2, c=2:  (()) ()()
n=3, c=5:  ((())) (()()) (())() ()(()) ()()()
```

but seen through a different projection. The Dyck words present the one-dimensional "height" profile of the walk, making the Lorentzian scaling limit and the breadth–depth structure of the light cone transparent. The pairs S-expressions, by contrast, foreground the *binary computational structure*: they are precisely the unlabeled S-expression trees of a Lisp-like language, with cons as the sole constructor. [1]

In this sense, Dyck words and pairs expansions are two complementary Catalan bijections:

- Dyck words: a one-dimensional time–breadth profile well adapted to continuum limits and Lorentzian geometry;

[1] Here "projection" is not a literal mapping but a change of coordinates on the same Catalan object. The Dyck, binary-tree, and S-expression representations are all related by canonical bijections; each is a different parametrization of the same underlying Catalan shape. Thus switching from Dyck words to pairs S-expressions does not change the object itself, only the coordinate system through which it is viewed.

- pairs S-expressions: a fully binary application tree well adapted to combinatory computation.

Both encode the same Catalan shapes; one may view them as distinct “Lorentz frames” on the same underlying combinatorial substrate. The pairs expansion can thus be regarded as an additional invariant transform: it preserves the Catalan class while re-expressing the same light-cone structure in purely computational coordinates.

There is also an intrinsic handedness in this picture. Dyck words of a given semilength n are not symmetric under reversal of the walk; the distribution of shapes across the tier reflects the left-to-right order in which parentheses are added. This asymmetry is the combinatorial trace of the fundamental handedness of applicative collapse in the pairs expansion: application is not commutative, and the collapse rule is oriented. The sequential construction of the S-expression tree makes this visible as a bias in how breadth is accumulated relative to depth.

A further simplification arises when we observe that in the S-expression view, explicit application nodes disappear entirely. Application is seen as a *property of the parentheses themselves*: a nested pair structure is already an applicative program. From the standpoint of Schönfinkel’s combinatory logic [17], where even the familiar S and K can be reduced to a single sufficiently expressive combinator, one may heuristically say that if a lone combinator J acting on parentheses is enough, then we can omit J and work directly with the bare parentheses (). The Catalan substrate then becomes a *combinator-free* calculus of pure application.

Traditional functional calculi admit multiple evaluation strategies (normal-order, applicative-order, call-by-need, etc.), but the Catalan substrate removes that freedom. Collapse is a local, causally constrained rewrite: a redex may collapse only after its causal predecessors, and once permitted the rule determines its successor uniquely. Computations are precisely the causally admissible paths through the substrate, with Dyck paths as their one-dimensional projections. Evaluation order is therefore fixed by causality rather than chosen by convention. Formalizing collapse as a partial order on redexes yields a causal-consistency lemma and shows that standard evaluation strategies are coarse-grainings of this order (Appendix A).

Historical Perspective. The S-expression viewpoint connects the present framework to three classical constructions. First, McCarthy’s original Lisp treats cons-pairs as the sole data constructor, with atoms added as a separate syntactic category [14]. Here we invert the hierarchy: structure is primary, and atoms—if present at all—arise as designated structural motifs.

Second, Church encodings demonstrate that data and control structures can be represented entirely by higher-order functions; similarly, SKI combinatory logic eliminates variables altogether. These developments show that symbolic reference is not primitive but can be reconstructed from purely structural or operational primitives.

Third, Gödel numbering treats syntax as arithmetic structure. The present approach may be viewed as a “Catalan numbering,” where syntactic entities are mapped to unlabeled binary trees. The Dyck, pairs, and binary-tree bijections then provide multiple coordinate systems for the same structural universe. In this sense, the Catalan substrate acts simultaneously as a computational calculus and as a structural semantics for symbolic systems.

A.4 Symbolic Representation in a Structure-Only Substrate

In the Catalan substrate all information is structural: the only primitive constructor is the cons-pair, and there are no atomic labels. This raises a fundamental question: how can a system without atoms support symbols, naming, or reference? The answer is that symbols arise not as primitives but as *structural motifs* that function as internal or external markers depending on context. We distinguish two forms of symbolic representation.

A.4.1 Internal Structural Symbols

Within a closed Catalan machine, one may bootstrap symbolic reference by designating particular tree shapes as internal "names." A higher-level self-referential mechanism (conceptually similar to a Y-like fixed-point operator) can maintain a dictionary of such shapes, mapping them to programs, behaviours, or combinator expansions. In this mode, names are themselves Catalan objects, and symbolic reference arises entirely from geometry: identifying a symbol is equivalent to matching a structural pattern. This yields a Lisp-like environment without atomic identifiers, where all "atoms" are implemented as canonical shapes in the tree.

A.4.2 External Structural Symbols

When interacting with external systems, the same structural motifs can serve as *extrinsic* symbols. Distinguished shapes may encode character codes, vector-drawing glyphs, device signals, or other forms of I/O. This does not modify the underlying calculus: it merely overlays a conventional interpretation on structural patterns. The Catalan substrate remains atomless internally, while external systems treat designated shapes as meaningful codes.

A.4.3 Unified View: Symbols as Distinguished Motifs

Both internal and external naming mechanisms exemplify a common principle: in a structure-only universe, symbols are not primitive entities but *designated Catalan motifs*. A symbol is simply a tree pattern endowed—internally or externally—with stable semantic interpretation. The substrate itself does not distinguish between data and code, or between program and identifier; all such distinctions emerge from the placement and recognition of specific structural forms.

Definition A.1 (Structural Symbol)**.** *A* structural symbol *is a finite Catalan tree S together with an interpretation map ι assigning S either (i) an internal computational meaning within the Catalan machine, or (ii) an external semantic meaning communicated to an outside system. The substrate recognizes S only as a structure; all semantics flow from ι.*

Remark A.1. *Although traditional programming languages begin with atoms and build structure on top of them, the Catalan substrate inverts this perspective: structure comes first, and atoms (if needed) are reintroduced later as structural patterns. From this viewpoint Lisp's cons-based representation, and even Schönfinkel's proposal of a single universal combinator, appear naturally as degenerate cases of a more general structure-first semantics.*

A.5 Causal Admissibility of Redex Contraction

We now formalize the sense in which evaluation order is fixed by causality rather than chosen by convention.

Definition A.2 (Causal Preorder on Positions)**.** *Let T be a full binary tree in the Catalan substrate, representing a program term. A* position *in T is a node address p in the usual tree sense (e.g. a finite word over $\{L, R\}$ indicating left/right choices from the root). We write $p \prec q$ if the node at position p lies on the unique path from the root to the node at position q. The reflexive, transitive closure of $\prec$ defines a preorder $\preceq$ on positions, which we call the* causal preorder*.*

Intuitively, $p \preceq q$ means that the subtree at q is causally downstream of the subtree at p: any change at p may propagate to q but not conversely.

Definition A.3 (Redex and Causal Admissibility)**.** *A* redex *in T is a position p such that the subtree rooted at p matches the left-hand side of a reduction rule (e.g. a β-redex or an SKI contraction). Let $R(T)$ be the set of all redex positions in T.*

A redex at position $p \in R(T)$ is causally admissible *if there is no other redex $q \in R(T)$ with $q \prec p$. In other words, a redex is causally admissible if it is minimal in $R(T)$ with respect to the strict causal order.*

Definition A.4 (Causally Admissible Reduction Sequence)**.** *A finite or infinite sequence of trees*

$$T_0 \to T_1 \to T_2 \to \cdots$$

is causally admissible *if, for each step $T_i \to T_{i+1}$, the contracted redex is causally admissible in T_i in the above sense. A* causal computation *is a causally admissible sequence starting from some initial tree T_0.*

Lemma A.1 (Causal Consistency of the Catalan Substrate)**.** *Let T_0 be an initial Catalan tree, and let*

$$T_0 \to T_1 \to \cdots \to T_m, \qquad T_0 \to T_1' \to \cdots \to T_n'$$

be two causally admissible reduction sequences that both terminate in a normal form (i.e. a tree with no redexes). Then:

1. *The normal forms coincide: $T_m \equiv T_n'$.*
2. *The two sequences differ, if at all, only by permutations of reductions at redexes that are pairwise incomparable under $\preceq$ (i.e. reductions occurring in disjoint causal subtrees).*

Proof Sketch. The first claim is a standard confluence argument: the underlying rewrite system (e.g. β-reduction or SKI contraction) is known to be confluent on well-formed terms [8, 4]. Since we never leave the Catalan family, the usual diamond property implies uniqueness of normal forms.

For the second claim, note that two redexes at positions p and q are causally independent if neither $p \prec q$ nor $q \prec p$ holds, i.e. they lie in disjoint subtrees. Contracting such redexes in either order yields the same result, by locality of the rewrite rules. Causally admissible sequences always contract redexes that are minimal in the causal order, so any difference between two such sequences can only be due to different interleavings of contractions at pairwise incomparable positions. This is the usual "causal consistency" property familiar from event structures and partial-order semantics of concurrency: linearizations of the same partial order differ only by permuting independent events. Here the partial order is given by the causal preorder on redex positions, and the corresponding linear extensions are precisely the causally admissible evaluation histories. □

A.6 Evaluation Strategies as Coarse-Grainings of Causal Order

Traditional presentations of the λ-calculus distinguish several evaluation strategies: normal-order, applicative-order, call-by-need, and many others. These are usually defined syntactically (e.g. by specifying which redex is chosen at each step), with confluence guaranteeing that they terminate in the same normal form when one exists. In the Catalan substrate, however, causality constrains redex selection more tightly.

Definition A.5 (Strategy-Compatible Causal Computation)**.** *Let $\mathcal{S}$ be a syntactic evaluation strategy (e.g. normal-order or applicative-order) which, given a term, selects one or more redex positions considered "eligible" at each step. A causal computation*

$$T_0 \to T_1 \to \cdots$$

is compatible *with $\mathcal{S}$ if, at each step, the contracted redex is both causally admissible in T_i and belongs to the set of redexes selected by $\mathcal{S}$ for the corresponding term.*

Proposition A.2 (Strategies as Coarse-Grainings of Causal Structure)**.** *Let T_0 be an initial term, and suppose $\mathcal{S}$ is a standard evaluation strategy that is normalizing on T_0 (e.g. normal-order for a weakly normalizing term). Then:*

1. *Every $\mathcal{S}$-guided reduction sequence can be refined to a causally admissible computation by reordering only reductions that occur at redexes incomparable under the causal preorder.*

2. *Conversely, every causally admissible computation from T_0 to normal form projects to an $\mathcal{S}$-valid history by forgetting the precise interleaving of causally independent reductions.*

In this sense, classical evaluation strategies are coarse-grainings of the underlying causal order: they differ only in how they resolve the residual freedom to permute causally independent redex contractions.

Proof Sketch. For (1), observe that any $\mathcal{S}$-guided sequence that temporarily contracts a non-minimal redex (with respect to $\preceq$) must do so in a context where all redexes on which it causally depends will eventually be contracted as well. By standard commuting-conversion arguments, we can reorder the sequence so that causally prior redexes are contracted first, without changing the final normal form. This reordering affects only redexes that lie in disjoint subtrees, i.e. are incomparable under $\preceq$.

For (2), given a causally admissible computation, we can group together all contractions that $\mathcal{S}$ regards as taking place at the "same" redex position in the syntactic term, ignoring the precise ordering among contractions in disjoint subtrees. The resulting abstract history matches what $\mathcal{S}$ would produce, by confluence and the fact that $\mathcal{S}$ is normalizing on T_0. Thus $\mathcal{S}$ may be seen as a projection that forgets the fine-grained causal structure of independent collapses while preserving the global reduction behaviour. $\square$

Corollary A.1 (Universality of Structural Symbol Embedding)**.** *Let Σ be any countable alphabet (finite or infinite), and let $\mathcal{L}(\Sigma)$ be the set of all well-formed symbolic expressions over Σ with arbitrary syntactic structure. Then there exists an injective map*

$$E : \mathcal{L}(\Sigma) \hookrightarrow \mathcal{C},$$

where $\mathcal{C}$ is the Catalan family of full binary trees, such that distinct symbols in Σ correspond to pairwise non-isomorphic Catalan motifs, and syntactic composition in $\mathcal{L}(\Sigma)$ corresponds to tree composition in $\mathcal{C}$.

In particular, any symbolic system (including alphabets, identifiers, typed terms, or abstract syntax trees) can be encoded as a system of structural symbols *embedded within the Catalan substrate. Thus the Catalan universe is universal for symbolic representation, with atoms appearing only as designated structural shapes.*

Proof Sketch. Enumerate $\Sigma = \{s_1, s_2, \dots\}$ and assign to each s_i a distinct Catalan tree S_i not isomorphic to any previously assigned motif. Define $E(s_i) = S_i$. Extend E compositionally: if a symbolic expression is built by forming a tree or sequence of symbols, map the compositional operation to the corresponding binary-tree constructor in $\mathcal{C}$. Injectivity follows from injectivity on generators and the fact that full binary trees are freely generated; well-formedness follows from closure of $\mathcal{C}$ under cons-pairing. Since $\mathcal{L}(\Sigma)$ is countable, such an embedding exists for any countable symbolic system. □

B Interpretive Appendix: Void, Potential, and Recursive Actualization

This appendix records an interpretive perspective that motivates the formal development of the paper. The material here is not required for the results of the main text. It is included to clarify the conceptual picture suggested by the recursive structures analyzed above.

B.1 Void and first differentiation

Let () denote the null structure: a state with no internal distinction and no recursive content. Formally, it is the base case of the Catalan construction. The minimal nontrivial closure of () under recursive pairing is

$$(()).$$

This object introduces an internal relation without introducing multiplicity. It is the smallest structure capable of supporting further recursive growth while remaining globally consistent with the Dyck constraint. In this sense, (()) represents the first differentiation of the void: not an object placed *in* an existing space, but the emergence of relational structure itself.

B.2 Possibility space from recursive expansion

Iterating the pairing rule generates the full Catalan family. Each construction step introduces new admissible extensions while preserving all previous structural commitments. The resulting set of Dyck paths may be read as a space of mutually compatible but not jointly realizable futures. From this perspective, the Catalan lattice represents a structured *possibility space*. The Dyck constraint does not merely limit growth; it organizes it, enforcing consistency across all scales.

B.3 Temporal interpretation

The recursive construction index naturally induces an ordering. Each step corresponds to the introduction of new structure relative to what has already been fixed. This ordering provides a discrete notion of time internal to the construction, without presupposing an external temporal parameter. At each stage, two complementary processes are present:

- expansion, in which new admissible configurations are introduced;
- restriction, in which incompatible possibilities are locally excluded.

Temporal progression may be read as the alternation between these processes: the opening of new potential followed by selection consistent with prior structure. The formal theory requires only that both processes respect the global recursive constraint.

B.4 Emergent spacetime properties

The causal and geometric features described in the main text arise directly from this alternation. Recursive depth tracks accumulated structural commitment, while breadth tracks contemporaneous branching. Bounds on breadth as a function of depth yield a cone-shaped causal envelope. Under scaling limits, this discrete structure supports diffusion and wave propagation. These continuum behaviors do not require additional geometric axioms; they emerge from the organization of possibility imposed by the recursive constraint.

B.5 Interpretive status

Nothing in this appendix asserts a physical identity between recursive pairing and any particular physical process. The interpretations offered here are meant to guide intuition rather than to extend the formal claims of the paper. The central formal result remains unchanged: a single recursive constraint suffices to generate a structured possibility space exhibiting causal order, wave dynamics, and computational universality. The interpretive perspective suggests how these features may be interpreted as aspects of a unified underlying process, but the validity of the formal results does not depend on that reading.

References

[1] L. Addario-Berry, L. Devroye, and S. Janson. Sub-Gaussian tail bounds for the width and height of conditioned Galton–Watson trees. *Annals of Probability*, 41(2):1074–1087, 2013.

[2] J. Ambjørn, J. Jurkiewicz, and R. Loll. Dynamically triangulating Lorentzian quantum gravity. *Nuclear Physics B*, 610:347–382, 2001.

[3] J. Ambjørn, A. Görlich, J. Jurkiewicz, and R. Loll. Nonperturbative quantum gravity. *Physics Reports*, 519:127–210, 2012.

[4] H. P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics.* North-Holland, 1984.

[5] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin. Space-time as a causal set. *Physical Review Letters*, 59(5):521–524, 1987.

[6] A. Church. A set of postulates for the foundation of logic. *Annals of Mathematics*, 34:839–864, 1933.

[7] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* 2nd edition, John Wiley & Sons, 2006.

[8] H. B. Curry and R. Feys. *Combinatory Logic, Vol. I.* North-Holland, 1958.

[9] H. B. Curry. *Foundations of Mathematical Logic.* McGraw–Hill, 1958.

[10] R. P. Feynman and A. R. Hibbs. *Quantum Mechanics and Path Integrals.* McGraw–Hill, 1965. (Dover reprint, 2010).

[11] S. Janson. Brownian excursion area, Wright's constants in graph enumeration, and other Brownian areas. *Probability Surveys*, 4:80–145, 2007.

[12] M. Kac. On distributions of certain Wiener functionals. *Transactions of the American Mathematical Society*, 65(1):1–13, 1949.

[13] J.-F. Le Gall. Random trees and applications. *Probability Surveys*, 2:245–311, 2005.

[14] J. McCarthy. Recursive functions of symbolic expressions and their computation by machine, Part I. *Communications of the ACM*, 3(4):184–195, 1960.

[15] R. Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117–158, 2014.

[16] C. Rovelli. *Quantum Gravity*. Cambridge University Press, 2004.

[17] M. Schönfinkel. Über die Bausteine der mathematischen Logik. *Mathematische Annalen*, 92:305–316, 1924.

[18] R. P. Stanley. *Catalan Numbers*. Cambridge University Press, 2015.