An Inquiry into Chemical Structure Generation through Symmetry Breaking

Two Formalisms Drawing on Group-Equivariant Grammar and Group Field Theory

Wanhong Huang
huangwanhong@serendip.ngo

Abstract. Structural regularities in chemistry invite generative descriptions, and the ontological interpretation of a successful structural grammar requires further argument. This exploratory investigation develops a symmetry-first framework for relating chemical structure formation to generative trajectories, symmetry breaking, and context-dependent functional roles. It articulates three conjectures: that structurally equivalent terminal representations can conceal history-dependent functional differences; that symmetry-breaking trajectories can enable functional possibilities inaccessible to a specified symmetric reference state; and that evidence for regularity and generation may underdetermine the existence of an independent, universal meta-generative order. Two complementary formalisms with distinct mathematical constructions are proposed. The first uses group-equivariant grammatical transformations and physically constrained dynamics; the second investigates relational two-complex generation through constructions inspired by group field theory. A contrastive model-construction methodology separates representation, physical adequacy, and ontological commitment. The present paper formulates questions and test conditions as a basis for future mathematical tests and chemical validation.

Keywords: Chemical structure generation; Symmetry breaking; Group-equivariant grammar; Group field theory; Functional semantics.

Conceptual Status and Corrections. This is a conceptual and exploratory theoretical paper. Its formulations and interpretations are open to further scrutiny and revision; corrections, constructive criticism, and independent verification are welcome. The author makes no claim of priority over any concept, method, or underlying idea discussed in this work.

Preprint Status. This inquiry is circulated solely as a preprint. No submission to any conference or journal is planned.

AI Assistance and Author Responsibility. The conceptual designs of the two formalisms originated with the human author. The manuscript was drafted using AI tools and subsequently revised collaboratively by the human author and AI tools. The human author takes full responsibility for the manuscript's content, accuracy, scientific claims, interpretations, and any remaining errors.

Source Code. The source code will be made available at The Science Infrastructures GitHub repository (applications/chemistry-workbench).

License. © 2026 Wanhong Huang. Original manuscript content is licensed under the Creative Commons Attribution–NonCommercial 4.0 International (CC BY-NC 4.0) license. Materials credited to third parties remain subject to their respective terms.

1 Introduction

The structure of a chemical system may be described at several levels: atomic composition, molecular connectivity, stereochemical configuration, electronic organization, intermolecular association, and the collective arrangements of condensed matter. These descriptions answer different questions. A molecular graph can specify covalent connectivity while leaving open the population of conformers, the organization of surrounding species, and the internal state of a prepared material. Conversely, two systems with distinct microscopic organizations may admit equivalent descriptions at a deliberately coarse level. Chemical formation is therefore an instructive setting in which to ask how a structure becomes realized, which aspects of its history survive a chosen description, and under what conditions a structure acquires a particular physical or biological role.

This investigation takes symmetry seriously as a possible organizing element of generative explanation. Group actions can identify transformations under which a law, set of admissible structures, or family of generative rules is invariant. A particular realization may preserve a proper subgroup of the symmetries of its modeled formation mechanism. This separation between the invariance of an effective law and the symmetry of one of its states lies at the heart of many descriptions of ordering, phase selection, chirality, and patterned organization. The resulting research problem extends beyond the classification of a molecule by its point group. We ask whether symmetry can constrain *how* chemical structures are generated, how alternative generative histories can be distinguished, and whether structural or functional properties are sensitive to those histories.

The problem admits a productive analogy with generative grammar. A grammar represents admissible constructions through rules and their compositions; a chemical model similarly specifies which local transformations or organizational configurations are allowed. Chemical production rules additionally require a physical interpretation: physically realized transformations depend on state, environment, energy, reaction pathways, and timescale. A syntactically admissible molecular graph may be chemically implausible, unstable, or inaccessible under specified conditions. This

distinction is already central to molecular graph grammars and rule-based systems, which offer mathematically explicit and computationally useful representations of chemical and biochemical transformations (Andersen et al., 2013; Kajino, 2019; Boutillier et al., 2018). Our interest is consequently in a physically interpreted generative description with domain-specific chemical constraints.

A second motivation arises from the recent incorporation of higher-order topology in molecular representations. Molecular hypergraph grammars have been developed to respect chemical validity constraints (Kajino, 2019); more recently, higher-order molecular grammars have lifted molecules to combinatorial complexes in order to represent rings and recurring motifs through compositional production rules (Huang et al., 2026). These advances make clear that a higher-dimensional or grammar-based encoding of molecular structure is itself an established computational research strategy. They also leave open a distinguishable question: could a *generative history*, rather than a static molecular representation alone, be encoded as a relational complex with symmetry-constrained composition? The present paper explores one candidate construction inspired by group field theory (GFT), whose interacting fields generate discrete combinatorial complexes in a setting developed principally for quantum gravity (Oriti, 2014; Oriti et al., 2014). Its application here is a mathematically motivated construction whose chemical interpretation and physical adequacy remain open research tasks.

A third motivation concerns the word *function*. A structure may exhibit a response to specified conditions; it may also play a causal role in a larger reaction network, biological organization, or device. Each role depends on additional physical, environmental, or organizational information beyond connectivity. For example, a molecule's binding and catalytic consequences depend on its accessible conformations, reaction partners, and surroundings. Research on molecular codes and on the physical significance of information has already demonstrated that chemical roles and semantic capacities can be formulated as constrained scientific questions (Görlich and Dittrich, 2013; Kolchinsky and Wolpert, 2018). In this paper, *functional semantics* is used in a restricted, operational sense: an experimentally or computationally evaluable role relative to a specified system, context, and assessment protocol. This restricted usage concerns causal and functional roles evaluated by a stated protocol.

These considerations motivate two related but distinct conjectures. The first concerns *generative-history dependence*: two terminal configurations may have isomorphic representations at a specified structural resolution, even as physically retained traces of their formation contribute to different functional responses. The second concerns *symmetry breaking*: a symmetry-constrained mechanism may admit realized configurations with reduced symmetry, and some such configurations may enable functions absent from a specified symmetric reference. The conjectures propose conditional possibilities, with meaningful history differences and functional changes to be determined for specified systems. Their purpose is to specify conditions for discriminating histories, states, and effects. In particular, if two terminal systems are identical in their full physical state and undergo identical subsequent dynamics under identical boundary conditions, physical explanations should account for any historical effect through state variables, boundary conditions, or the evaluation protocol.

The inquiry also encounters an ontological difficulty. Successful grammatical or field-theoretic descriptions may reveal common constraints on structure formation. Such explanatory success is compatible with different accounts of the existence of a unique, physically instantiated meta-grammar and of an absolute origin of generative structures. Similar caution is warranted when relating structural universals to claims about a biological Universal Grammar: the latter is a domain-specific scientific hypothesis involving development and cognition, whereas the former may be no more than a mathematically effective representation of regularity (Chomsky, 2005). We therefore advance a third conjecture concerning the possible *ontological underdetermination* of generative order. It motivates an explicit contrastive method: to test whether admissible accounts that agree on specified observations can nevertheless disagree about the physical status of a postulated meta-generative structure. The general philosophical structure of this problem is familiar from debates concerning the underdetermination of scientific theories by evidence (Stanford, 2023).

Figure 1 previews the research architecture. We propose two complementary, independent formalisms. The first takes group-equivariant grammatical transformations and their dynamics as primary modeling objects. The second uses interacting group-theoretic fields and relational two-complexes to represent candidate formation structures and histories. Both are *symmetry-first* in a methodological sense: a suitable symmetry structure helps organize permissible generative operations. The role of symmetry is explicitly methodological; ontological priority, domain-specific group choice, and equivalence between the two constructions are independent questions.

The contribution of the paper is accordingly a carefully delimited theoretical proposal. It identifies three conjectures, develops two formalisms with explicit but distinct mathematical commitments, and establishes a method for separating formal representability, physical realizability, and ontological inference. Its anticipated contributions are conditional propositions, models, and discriminating tests with explicit empirical and ontological qualifications. The exposition first situates structural regularity within existing mathematical chemistry, then introduces the methodological standards governing the formal constructions. Subsequent sections develop the two formalisms separately, evaluate the conjectures, and examine what—if anything—can be transferred between grammatical and field-theoretic descriptions of chemical generation.

2 Structural Regularity And Generative Explanation

2.1 Structural Descriptions And Their Domains

A chemical structure is always given relative to a selection of degrees of freedom and relations. For a conventional molecular graph, vertices represent atom types and edges represent bond relations. Such a representation is invaluable for reasoning about connectivity, but it is a quotient of richer physical information. The locations of nuclei, electronic

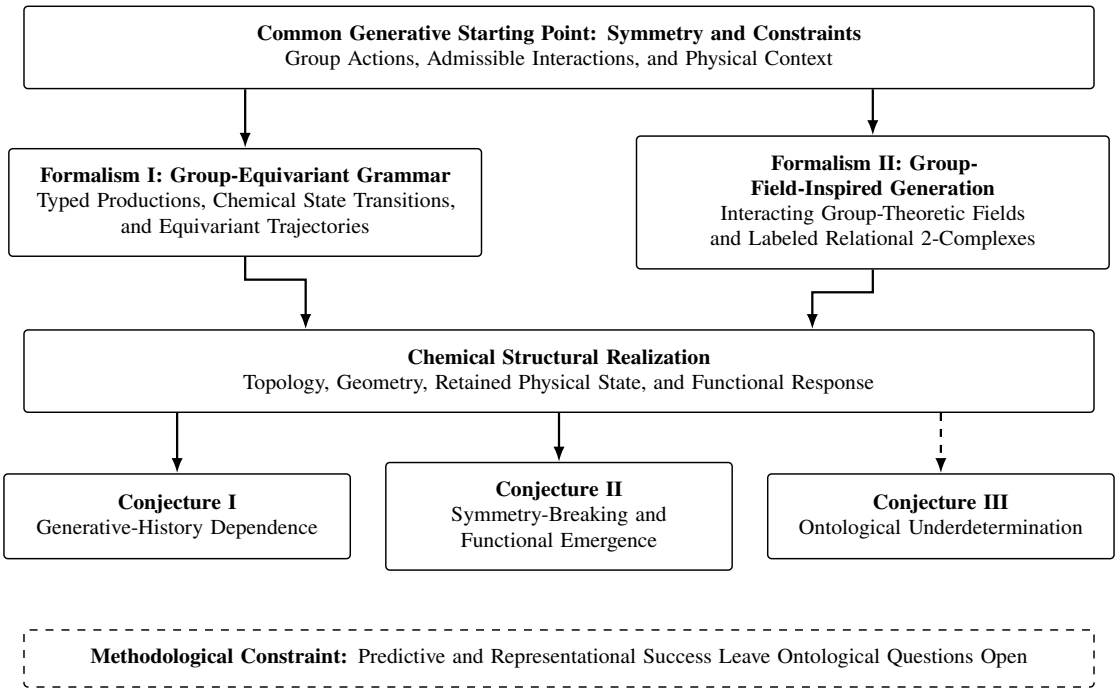


Figure 1: Research Architecture. The two formalisms share a symmetry-first modeling orientation and have distinct mathematical constructions. Conjectures I and II address possible chemical phenomena, whereas Conjecture III addresses the ontological interpretation of successful descriptions. The bottom panel is a framework-wide methodological condition, visually separated from the physical-generation paths.

states, conformational distributions, solvent configurations, collective order, and defects may be relevant to a different scientific question. A useful theory can therefore employ a family of structural maps adapted to the resolution of each research question.

Let $\mathcal{X}$ be a space of physical or effective states and let $\mathcal{S}_r$ denote a space of structural descriptions at resolution r. A representation map $T_r : \mathcal{X} \to \mathcal{S}_r$ assigns an admissible structural description to each state. For geometric, statistical, or gauge-redundant descriptions the map may instead be defined on a suitable quotient or yield a distribution over structural descriptions; the simple deterministic form is used here to establish terminology.

Definition 2.1 (Representation-Relative Structural Equivalence). For a chosen representation T_r, two states $x, y \in \mathcal{X}$ are *structurally equivalent at resolution* r if $T_r(x) \cong T_r(y)$ in the specified category of labeled structural objects. We write $x \simeq_r y$.

The isomorphism must preserve the labels declared relevant to the representation. A connectivity-only graph can identify states that a stereochemically annotated graph separates; a graph incorporating stereochemistry may still identify states with different defect arrangements or surrounding chemical activity. Moreover, isomorphism is stronger and more appropriate here than a general categorical bimorphism: a morphism that is both epic and monic need not be invertible in an arbitrary category. The selection of the category and of its admissible morphisms is a substantive modeling decision.

This relative treatment avoids two opposing errors. Structural identity should not be inferred from identical molecular formulae, while any microscopic difference need not prevent useful structural comparison. A scientifically informative equivalence relation is indexed by the observables and transformations that matter for the task. It will become essential when assessing the first conjecture: equality in a coarse structural representation is compatible with inequality in other physical state variables.

2.2 Grammar And Reaction Mechanism

Rule-based chemistry already distinguishes a transformation pattern from its particular realization. A rule may specify a local bond rearrangement, while an occurrence of that rule depends on a matched substrate, environment, and possibly additional constraints. Graph-rewriting methods for chemistry and the Kappa formalism for molecular interactions have exploited this distinction to compress combinatorially large reaction spaces and to trace causal organization (Andersen et al., 2013; Boutillier et al., 2018). Hypergraph grammars have also been used to preserve chemical validity in generated molecular representations (Kajino, 2019).

The adjective *generative* nevertheless admits several scientific interpretations. A formal rewriting relation $s \Rightarrow s'$ may assert only that s' is syntactically produced by a rule. A stochastic generator supplies rates or event probabilities for such transitions. A microscopic physical theory attempts to explain which effective events are supported by interactions and energetic constraints. These levels are related, but none can be substituted for another without a specified interpretation map and an argument for its adequacy.

Definition 2.2 (Generative Description). A *generative description* of a system on a domain D comprises (i) an admissible state or structural class, (ii) an effective specification of permitted constructions or transformations, and (iii) a semantics that relates those constructions to the target system. The semantics may be structural, stochastic, or

physical; its level must be declared.

This definition is deliberately broad enough to accommodate grammatical rewriting, deterministic dynamics, and constrained stochastic processes. The grammar is an explanatory mathematical representation whose relation to physical processes requires independent modeling. Indeed, an admissible derivation in a chemical graph grammar can lack a realizable reaction pathway, whereas one effective reaction may aggregate multiple microscopic processes. Chemical models must therefore include thermodynamic and kinetic restrictions in addition to structural validity. Nonequilibrium thermodynamics of reaction networks supplies established tools for evaluating such restrictions in chemically specified settings (Rao and Esposito, 2016).

An additional complication concerns the meta-level. A grammar describing molecular structures differs from a *meta-grammar* describing how families of such grammars can be constructed or restricted. A mathematically defined meta-grammar can be valuable even when there is no evidence that the corresponding higher-order grammar is a physically independent entity. Representational adequacy and physical generative causation are distinct questions. This distinction will guide the treatment of the third conjecture.

2.3 Symmetry In Laws, Realizations, And Histories

Symmetry provides a precise vocabulary for expressing relations among transformations and their outcomes. Given a group G acting on a state space $\mathcal{X}$, a state x has stabilizer $\mathrm{Stab}_G(x) = \{g \in G : g \cdot x = x\}$. A modeling law may be equivariant under G even if its particular realizations have smaller stabilizers. For a context space C, it is often necessary to transform the context along with the state: a mirror-related reaction in a mirror-related environment is not generally equivalent to reflecting a reactant while retaining a chiral environment unchanged.

A distinction between a *symmetry of the model* and a *symmetry of the solution* is particularly important. In a deterministic model with a unique flow, if both the context and initial state are invariant under a group element, equivariance preserves that invariance for all subsequent times for which the flow exists. Accounts of spontaneously selected asymmetric outcomes therefore require appropriate treatment of unstable equilibria, perturbations, stochastic fluctuations, limiting procedures, or environmental asymmetries. A generic diagram of $G \to H$ is a useful record of residual symmetry, but cannot by itself demonstrate a physical symmetry-breaking event.

Group-based labels may serve a second, distinct role in constructing relational spaces. In GFT, fields defined over products of groups interact to generate combinatorial complexes, with a formal relation to spin-foam state sums in quantum gravity (Oriti, 2014; Oriti et al., 2014). The use of group representations and complex gluing in such theories is mathematically suggestive for discrete generative chemistry. Yet the physical interpretation of those labels does not automatically carry over: a molecular bond, a reaction event, and a two-cell of a spin foam are different objects. We use the field-theoretic construction as an exploratory generative *formalism*, and reserve the question of a chemically justified dynamics for independent treatment.

2.4 Chemical Properties and Functional Roles

The word *semantics* is useful under the condition that it is constrained by a declared scientific task. In a structural grammar, semantics may mean the mapping from symbols to molecular configurations. In chemistry and biology, it may mean the causal consequences of a component within a larger organization. A molecule can participate in multiple processes, with its functional contribution depending on partners, concentrations, phase, location, and timescale. One must also distinguish a measured response from a normatively or evolutionarily attributed function: a substance may have a physical effect without possessing a goal or an intrinsic meaning.

The literature already contains mathematical attempts to define molecular codes through chemical reaction networks and to identify information significant for a system's continued organization (Görlich and Dittrich, 2013; Kolchinsky and Wolpert, 2018). These are relevant antecedents, not special cases proven equivalent to the present proposal. The proposed inquiry focuses on whether *generative histories and symmetry information* can improve the determination of context-dependent roles beyond what a fixed structural representation provides.

For a reference system and evaluation protocol J, a functional descriptor may be a scalar score, a vector of responses, a time series, or a probability law over observations. A property such as catalytic activity is meaningful only after specifying substrates and conditions; a biological role additionally requires a system boundary and causal or evaluative criteria. We therefore treat functional semantics as *relational and operational* throughout. Its connection to generative formalisms is a question to be investigated, rather than a preinstalled interpretation of their symbols.

2.5 Status of Universal Order

The appearance of recurring forms across chemical structures suggests common constraints and often rewards higher-order explanatory models. Yet the inference from recurrence to a physically instantiated universal grammar is underdetermined unless the intended ontology is stated. At least three claims must be distinguished: that structural regularity is observed; that a compact grammar predicts or generates those structures; and that the grammar is a causally fundamental constituent of the physical world. The first can motivate the second; the second does not, without further premises, entail the third.

The analogy with linguistic Universal Grammar is instructive precisely because it must remain limited. Chomsky's account of language design differentiates genetic endowment, linguistic experience, and more general computational principles (Chomsky, 2005). Evidence for regularity or generative productivity in languages alone does not determine which of these sources provides a given constraint. Similarly, an informative chemical meta-grammar may be an effective compression of physical constraints, a description of an emergent organizational regime, or—subject to further argument—a more primitive physical mechanism. These are alternative explanatory interpretations, not linguistic and chemical entities assumed to be identical.

The framework therefore makes a methodological commitment to *symmetry-first modeling*, rather than a metaphysical

commitment to *symmetry-first ontology*. Its two generative formalisms explore what can be constructed under chosen group-theoretic constraints, while its third conjecture asks when evidence can discriminate between realist and non-realist interpretations of their meta-generative structures. This restriction makes a more ambitious investigation possible without prejudging the existence or uniqueness of a universal order.

3 Methodological Foundations

The paper develops a formal and conceptual investigation; its empirical component consists of proposed tests and connections to established chemical observations. Its methodological purpose is to specify mathematical objects, formulate conditional predictions, and identify the empirical or logical procedures by which those predictions could be evaluated. To prevent mathematical construction from being mistaken for physical discovery, we distinguish four levels of claim: *representational admissibility*, *mathematical validity*, *physical adequacy*, and *ontological attribution*. The levels support one another when appropriate justification is supplied, but no level is assumed to imply the next automatically.

3.1 Common Language for Two Formalisms

The two formal developments will use distinct mathematical constructions; a minimal interface is nevertheless useful. Let $\mathcal{X}$ denote an admissible state space, C a space of contexts or boundary conditions, and G a group with specified actions on states and contexts. Let Φ represent the chosen physical or effective evolution, and let T_r be a structural observation at resolution r. Let J index a finite or measurable class of functional tests, with a response map $\mu(x, C, J)$ into an outcome space $\mathcal{Y}_J$ or into probability distributions over that space. A schematic interface is

$$\mathfrak{I} = (\mathcal{X}, C, G, \rho, T_r, \Phi, J, \mu). \tag{1}$$

This is an *interface signature* that permits comparisons among formalisms with distinct underlying objects. Formalism I will specify grammatical rewriting and equivariant evolution on suitable state representations. Formalism II will introduce labeled relational complexes and a group-field-inspired mechanism for their generation. Comparison between the two requires an explicit map from generated objects to a common physical or effective state description. Visual similarity alone leaves any equivalence theorem unestablished.

Assumption 3.1 (Declared Observational Resolution). Every claim of structural identity, symmetry, trajectory correspondence, or functional equivalence is indexed by an explicitly identified representation, context domain, and observation or intervention protocol.

Under Assumption 3.1, a test of history-dependent function compares states judged equivalent by a declared T_r, prepared under specified histories, and evaluated under a common protocol. Likewise, a claim about symmetry breaking requires an explicit action, reference state, and method for estimating residual symmetry. This makes it possible to discover whether an apparent history effect disappears once previously omitted physical state variables are measured.

3.2 Operational Comparison Of Paths And Functions

For a generative trajectory $\gamma : [0, \tau] \to \mathcal{X}$ with endpoint x_γ, write $B(\gamma)$ for a symmetry-history descriptor. Such a descriptor may record a stabilizer history, order-parameter trajectory, transition sequence, or defect-formation statistics; it is *not* assumed to be uniquely recoverable from a terminal molecular graph. For a fixed context C and evaluation protocol J, let $P_{\gamma,C,J}$ denote the resulting distribution of measurable functional outputs. If the response is deterministic, the distribution is a point mass. The following comparison is the prospective empirical core of the first conjecture:

$$\begin{aligned} T_r(x_{\gamma_1}) &\cong T_r(x_{\gamma_2}), \\ B(\gamma_1) &\not\simeq_B B(\gamma_2), \\ d_J(P_{\gamma_1,C,J}, P_{\gamma_2,C,J}) &> \varepsilon_J. \end{aligned} \tag{2}$$

Here $\simeq_B$ is a declared history equivalence, d_J a suitable distance between output laws, and ε_J a justified tolerance incorporating measurement and modeling uncertainty. The last inequality is *testable*, unlike an unqualified assertion that the same structure has two intrinsic meanings. A complete study would control external conditions, quantify preparation variability, and examine possible mediating terminal-state differences.

The second conjecture requires a different design. For a symmetry-preserving reference state x_0 and a realized state x_f, consider the respective capability sets $\mathcal{F}(x_0, C, J)$ and $\mathcal{F}(x_f, C, J)$ under the same class of permitted interventions. The desired comparison is not merely $\mathrm{Stab}_G(x_f) \subsetneq \mathrm{Stab}_G(x_0)$; it additionally asks whether a function belongs to the latter capability set but not the former, or vice versa, depending on a predeclared direction of comparison. For the conjecture proposed here, the relevant existence claim is

$$\mathcal{F}(x_f, C, J) \setminus \mathcal{F}(x_0, C, J) \neq \varnothing. \tag{3}$$

An observation of changed symmetry alone is insufficient evidence for Eq. (3). A defensible analysis therefore requires evidence for the functional distinction and a comparison against simultaneous physical changes that may explain it.

Remark 3.1 (Complete-State Sufficiency). When a closed description contains the complete dynamical state and its evolution is Markovian under the specified conditions, the future distribution is determined by that state and those conditions. Any dependence on a recorded past must then operate through the present state, a changed context, or a historically defined *evaluation* of function. An independently acting "memory of the past" is not posited.

This remark separates physical history dependence from attribution of historical or evolutionary functions. The first may be tested by expanding the state description until dynamics become predictively adequate. The second additionally demands an explicitly justified account of how a role is assigned to a component within an organization and its history.

3.3 Contrastive Model Construction

The third conjecture requires a different method because it concerns the ontology of generative descriptions. Let

$\mathbb{T} = B \cup E$ be a formalized collection of accepted background commitments B and observational constraints E. Let H_M be a well-formed claim that a *physically substantive* meta-generative structure exists in the modeled domain. Importantly, H_M cannot mean merely that some symbolic meta-grammar can be written down: that mathematical existence would be too weak to address the intended dispute.

Definition 3.1 (Relative Existential Underdetermination). An existential hypothesis H_M is *underdetermined relative to* $\mathbb{T}$ if there exist admissible interpretations $\mathfrak{M}_+$ and $\mathfrak{M}_-$ in which $\mathbb{T}$ holds but H_M has opposite truth values:

$$\begin{aligned} \mathfrak{M}_+ &\models \mathbb{T} \wedge H_M, \\ \mathfrak{M}_- &\models \mathbb{T} \wedge \neg H_M. \end{aligned} \tag{4}$$

Proposition 3.1 (Non-Entailment From Contrastive Models). If models satisfying Eq. (4) exist, then $\mathbb{T} \not\models H_M$ and $\mathbb{T} \not\models \neg H_M$.

The argument is elementary but indispensable: semantic entailment requires truth in every admissible model of the premises. The pair in Eq. (4) supplies countermodels to both proposed entailments. A complete proof is given in Appendix A. This result concerns a *specified theory and interpretation class*; it is not an application of Gödel's incompleteness theorems and does not establish that the physical world is intrinsically unknowable.

Logical non-entailment and empirical indistinguishability must be handled separately. Let $\mathcal{P}$ be a declared class of experiments or observations and write $P^{\pi}_{\mathfrak{M}}$ for a model's predictive output law under protocol $\pi \in \mathcal{P}$. An operational notion of equivalence is

$$\mathfrak{M}_+ \equiv_{\mathcal{P},\epsilon} \mathfrak{M}_- \iff \sup_{\pi \in \mathcal{P}} d_\pi\left(P^{\pi}_{\mathfrak{M}_+}, P^{\pi}_{\mathfrak{M}_-}\right) \le \epsilon. \tag{5}$$

The criterion requires protocols, output spaces, distances, and tolerance levels that render the supremum meaningful. Agreement across currently available protocols may establish temporary underdetermination. The scope remains restricted to the declared protocols; future tests and comparative theoretical merit remain separate assessments (Stanford, 2023).

To avoid an empty exercise in relabeling, admissible competing models should possess coherent dynamics and an explicit interpretation of H_M; auxiliary assumptions should be independently discussable. Merely appending an observationally inert assertion to an existing chemical model offers little explanatory progress. The burden of construction is therefore stronger than producing a logically consistent negation: one must identify a scientifically serious alternative and show exactly where its ontological commitments diverge.

3.4 Discriminating Results for Three Conjectures

Sections 6–8 state the three research targets formally. *Generative-history dependence* asks whether structural equivalence at a specified resolution can coexist with distinct functional outcomes mediated by retained state variables. *Symmetry-breaking functional emergence* asks when reduced realized symmetry corresponds to access to a previously unavailable functional response. *Ontological underdetermination* asks whether plausible generative models that satisfy the same observational constraints can support conflicting claims about the existence of a universal meta-generative structure. Their formal status is that of conjectures; the paper will not convert evidence from one into an automatic proof of another.

Successful mathematical work may yield several kinds of results: an equivariance theorem for a specified rewriting or evolution operator, a demonstration that one structural language can distinguish states that another collapses, a conditional correspondence between grammatical derivations and two-complex histories, or a countermodel to an unqualified ontological inference. Chemical-mechanism validation additionally calls for independent empirical evidence or physically grounded derivations. In particular, the existence of a representation with GFT-like interaction vertices cannot by itself demonstrate that a GFT is the fundamental dynamics of chemistry.

The practical order of evaluation is consequently: construct explicit finite examples and verify their algebraic consistency; specify which variables and parameters admit physical interpretation; establish theorems where their assumptions are available; and compare predictions with defensible chemical observations or validated models. In view of the dual-use character of some computational chemistry capabilities, initial examples can remain at the level of benign model systems, known material phenomena, and abstract reaction networks without specifying hazardous synthesis or autonomous laboratory procedures. This research program leaves room for strong mathematical statements while withholding stronger physical and ontological conclusions until they have earned support.

4 Group-Equivariant Chemical Grammar

The first formalism introduces a typed chemical graph grammar before assigning physical dynamics to its productions. Its elementary objects are terminal chemical labels, nonterminal expansion sites, an initial sentential graph, and admissible productions. A symmetry group subsequently acts on this syntax together with chemically specified contexts. The distinction between a formal derivation and a physical transformation remains explicit throughout the construction (Andersen et al., 2013; Boutillier et al., 2018).

4.1 Grammar Signature, Terminals, And Production Rules

Definition 4.1 (Chemical Structural Grammar). A *typed chemical structural grammar* is a tuple

$$\Gamma = (\Sigma_{\mathrm{T}}, \Sigma_{\mathrm{N}}, S, \mathcal{P}, \mathbf{ChemG}, C, G, \rho), \tag{6}$$

where Σ_{T} is a finite alphabet of *terminal types* (atom labels with declared attributes, bond labels, and any admissible terminal annotations); Σ_{N} is a disjoint collection of *nonterminal symbols* with typed attachment ports; S is the *start graph*, possibly containing nonterminals; $\mathcal{P}$ is a family of typed *production schemas*; **ChemG** is a chosen category of finite typed attributed graphs or interface hypergraphs; C is the context space; and (G, ρ) specifies compatible group actions on the relevant graphs, contexts, and productions. The rules in $\mathcal{P}$ are subject to declared port-matching and chemical-admissibility conditions.

A nonterminal $A \in \Sigma_{\mathrm{N}}$ denotes a formal *unresolved expan-*

sion site, with attachment arity $k(A)$; it marks unresolved connectivity in the formal derivation; its chemical interpretation arises through a completed terminal graph. A terminal derivation ends in a graph whose labels are drawn from Σ_{T} and that satisfies the selected closure, bond, valence, and attribute constraints. Molecular geometry and electronic state can be supplied by a richer interpretation of that graph.

A graph-expansion production has the schematic form

$$p : A(\mathbf{u}) \longrightarrow H_p(\mathbf{u}), \qquad \chi_p(x, C, m) = 1, \quad \mathbf{u} = (u_1, \ldots, u_k). \tag{7}$$

where H_p is a typed replacement graph with the same declared interface, m is a match, and χ_p tests the intended surrounding state and context. More general chemical edits use the double-pushout span in Definition 4.2; the expansion syntax and the physical graph-edit semantics are linked by an explicit interpretation map. Write $S \Rightarrow^*_{\mathcal{P}} T$ for a finite derivation and define the terminal structural language

$$L_{\mathrm{term}}(\Gamma; C) = \big\{ [T]_{\cong} : S \Rightarrow^*_{\mathcal{P},C} T, \; T \text{ terminal and admissible} \big\}. \tag{8}$$

The quotient by typed-graph isomorphism is optional when atom identities or orientation labels are themselves observable. A physical realization map and physically calibrated transition rates impose further restrictions on this syntactic language.

Figure 2 makes the distinction among start symbols, nonterminals, terminal atom-and-bond labels, and replacement rules visually explicit. The demonstration derives an ethanol connectivity graph as a *formal graph construction*; abbreviated terminal fragments display the completed atom-and-bond structure, while nonterminal ports record unresolved connectivity at intermediate steps. Chemical synthesis pathways require separate reaction conditions and kinetics. Figure 3 then illustrates multiple successive graph operations in a peptide-backbone bookkeeping example. Appendix K restates the same objects in implementation-oriented terms so that the tuple, rule syntax, and execution order can be translated into code with minimal ambiguity.

4.2 Typed States And Context-Sensitive Rules

Fix a category **ChemG** of finite typed, attributed graphs with an explicit distinction between immutable atom identifiers and mutable attributes (notably formal charge and selected bond labels). A rule changing a preserved vertex attribute is represented using partially labeled DPO interfaces with injective matching, following the relabelling construction of Habel and Plump (2002). Strictly label-preserving morphisms apply to the attributes declared immutable. An object may represent a molecular configuration or a collection of interacting components. A structural semantics $T_r : \mathcal{X} \to \mathcal{S}_r$ associates a physical/effective state $x \in \mathcal{X}$ with a graph or an isomorphism class of graphs at resolution r; the states may retain positions, conformations, environmental components, or other degrees of freedom not captured by T_r. In particular, T_r is allowed to be many-to-one.

Definition 4.2 (Contextual Chemical Production Rule). A *contextual production rule* is a tuple $q = (L_q \xleftarrow{l_q} K_q \xrightarrow{r_q} R_q, \chi_q, \nu_q)$ with an injectively matched DPO span and a contextual predicate $\chi_q(x, C, m)$. In the relabelling variant, the interface K_q may leave mutable attributes partially unspecified; L_q and R_q carry the corresponding initial and final values. Preserved atom identifiers and declared immutable types match across the span. The state lift ν_q maps an admissible structural rewrite into an effective state change $x \xRightarrow{q,m;C} x'$. Applicability requires the appropriate gluing conditions, a compatible partial-label pushout, $\chi_q = 1$, and a well-defined state lift.

Figure 4 makes the local edit explicit for a chemically familiar acid–base equilibrium: it highlights one oxygen–hydrogen incidence being deleted and another introduced, with the proton retained as the same typed atom. This is more informative to a chemical reader than the abstract production span alone. The charge annotations and molecule inventory are part of the rule datum; a complete solvent-mediated mechanism requires additional physical information.

The relabelling DPO span uses K_q to identify preserved vertices and edges while permitting their mutable attributes to change between L_q and R_q. For proton transfer, two oxygen vertices remain in the interface with partially unspecified charge labels; the left and right graphs prescribe neutral and ionic charge states respectively. This explicit relabelling is essential for the example in Fig. 4 (Habel and Plump, 2002). The state-change specification ν_q may be deterministic, set-valued, or probabilistic; a calibrated interpretation can describe an effective approximation to microscopic transitions. DPO admissibility identifies structurally valid rewrites; a separate physical model assigns microscopic mechanisms. For changes in bond order represented by attributed edges, the choice of attributed-graph category must make the proposed relabeling operation valid. Activation barriers require additional energetic and kinetic information beyond graph matching and atom-valence screening.

A context C can specify temperature, chemical activities, confinement, external fields, or imposed chemical reservoirs. A catalyst dynamically produced by the modeled system belongs to x, whereas a catalyst held fixed by external control may be incorporated into C. This allocation follows the modeled system boundary and remains fixed across a declared comparison. By using conditions on the matched *surrounding state*, the construction permits context-sensitive rule application without asserting that every rule is context-sensitive in the exact string-language sense of the Chomsky hierarchy.

Assumption 4.1 (Chemical Accounting). For every physically interpreted reaction event, all declared atomic species and electric charge are conserved across the total modeled system, including explicit reservoir transfers where applicable. Any effective event that aggregates microscopic steps is identified as such; its parameters must be justified separately.

Let $N \in \mathbb{Z}^{n \times m}$ be the stoichiometric matrix of a finite species-level projection and $A \in \mathbb{N}^{k \times n}$ the matrix recording atom counts by element. In the closed, non-nuclear case, the accounting condition requires

$$AN = 0, \qquad \mathbf{z}^{\top} N = 0, \tag{9}$$

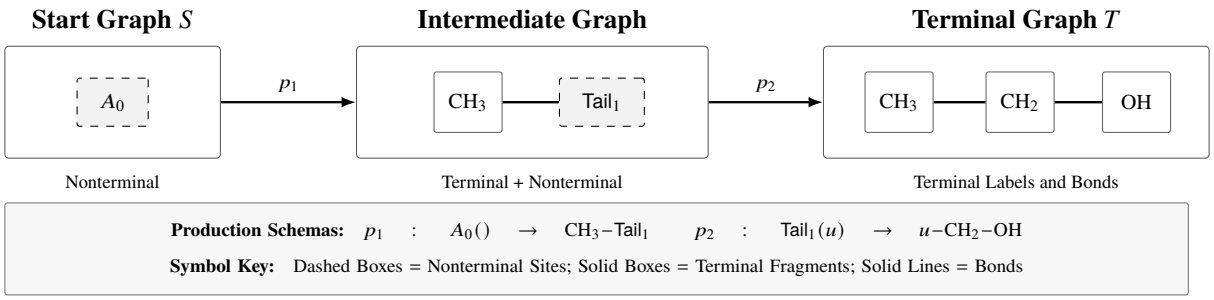


Figure 2: Typed Grammar: From a Start Symbol to an Ethanol Connectivity Graph. The first production replaces a zero-port start symbol with a methyl fragment and a one-port nonterminal. The second resolves the attachment to a methylene–hydroxyl fragment. The terminal boxes abbreviate atom-and-bond subgraphs with implicit hydrogens; the output is a fully resolved connectivity graph. This is an illustrative graph derivation, distinct from a laboratory reaction mechanism.

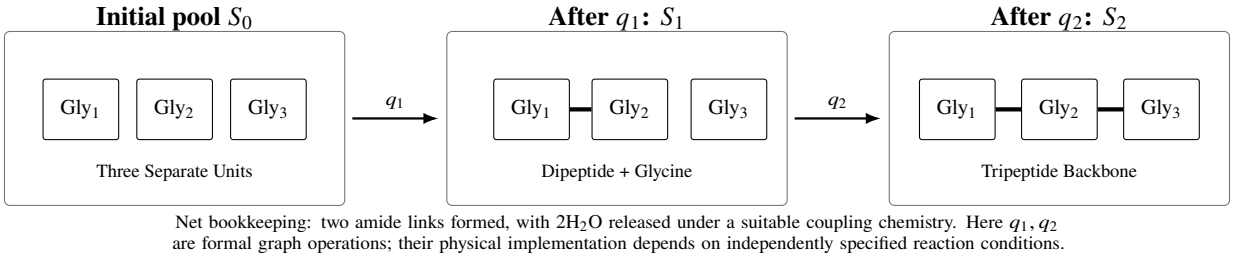


Figure 3: Successive grammatical generation of a labeled glycine backbone. A typed rule transforms an admissible structure into another admissible input. Heavy black bonds mark newly formed peptide connectivity; the labels distinguish bookkeeping identities rather than different chemical glycine species. Actual peptide coupling requires additional chemical conditions and kinetic treatment.

where $\mathbf{z}$ records species charges. For an open system, the total balance accounts for reservoir columns or explicit exchange terms. These equations impose conservation constraints. Geometric, energetic, spin, and kinetic restrictions provide further conditions for physical realizability.

4.3 Group Actions On Grammar And Environment

Let G be a group with a specified action on $\mathcal{X}$ and C, and let g act by structure-preserving functors or equivalences on the chosen graph category. The group may encode permutations of equivalent labels, molecular or collective spatial symmetries, or abstract effective-model symmetries. The group action is selected to reflect the observable of interest: for example, a chirality-sensitive representation retains the effect of orientation-reversing spatial operations. Compatibility with structural semantics requires a specified equivariant isomorphism $T_r(gx) \cong gT_r(x)$; it is not automatic for a chosen coarse-graining map. A group that acts trivially on connectivity can act nontrivially on the spatially enriched state x.

Definition 4.3 (Equivariant Physical Interpretation). An *equivariant physical interpretation* of the structural grammar Γ in Definition 4.1 consists of

$$\mathbb{G} = (\Gamma, \mathcal{X}, T_r, \mathcal{R}, \chi, \nu), \tag{10}$$

where $\mathcal{X}$ is the chemical-state domain, T_r is its chosen structural representation, $\mathcal{R}$ contains matched production or rewriting events, χ evaluates their context-dependent admissibility, and ν supplies the allowed state changes. The actions ρ extend consistently to states, matches, and events. In particular, a transition $x \xRightarrow{q,m;C} x'$ has a corresponding transformed transition $gx \xRightarrow{gq,gm;gC} gx'$. The event family is closed under the specified group action, with admissibility conditions transforming equivariantly.

This definition is a condition on a family of transformations and matches, while treating chemical reactions as transformations distinct from group elements. Reaction steps are often irreversible at the effective level. Composition of applicable rewriting steps is captured by a rewriting category or path structure; the group acts on this structure without turning its arrows into invertible reactions.

Proposition 4.1 (Equivariance Of Derived Rewritings). Suppose G acts by graph-category equivalences preserving the required DPO squares, every transformed match is admissible exactly when the original match is, and the state lift ν commutes with the action. Then the set of admissible one-step transitions is G-equivariant. Moreover, any finite derivation $x_0 \Rightarrow \cdots \Rightarrow x_n$ is carried into a derivation $gx_0 \Rightarrow \cdots \Rightarrow gx_n$ of the same length.

The result is conditional but useful: applying the group action to a commuting DPO construction again yields the corresponding construction, and the conclusion extends by composition. The full proof, including why matches and their multiplicities matter, is given in Appendix B. The proposition establishes formal covariance under stated assumptions. Symmetries of specific reaction environments and experimental realizability require independent evidence.

Figure 5 depicts the covariance requirement as a commuting square. Its upper row is a mirrored pair of tetrahedral stereochemical configurations; the vertical arrows represent abstract admissible rules transformed consistently with their chemical context. Figure 6 then illustrates that different

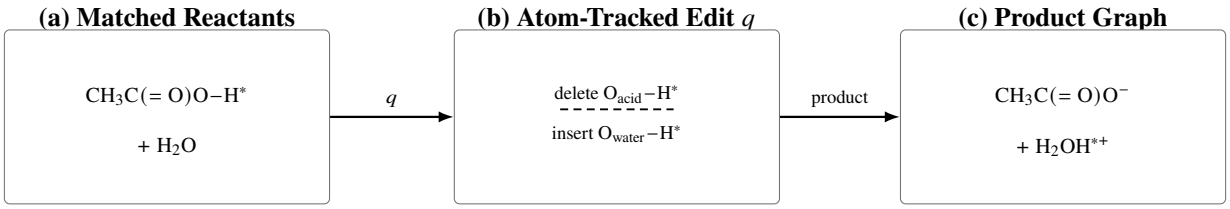


Figure 4: Local graph rewriting of acid–base connectivity. The marked proton moves from an acid oxygen to a water oxygen. A matched rewriting operation updates two bond incidences and the relevant charge labels while retaining atomic identity and net accounting. Real reaction pathways and propensities require an independently specified physical model.

orders of local rule applications can terminate in the same labeled molecular backbone. The latter is a statement about derivation multiplicity, compatible with multiple histories while leaving their physical retention open.

4.4 Derivations and Chemical Stochastic Dynamics

To represent physically occurring events, associate to an admissible rule–match pair (q, m) a nonnegative propensity $a_{q,m}(x, C)$. For definiteness, consider a countable state space with finite outgoing total rates and event maps $J_{q,m}(x)$. The observable generator of a continuous-time Markov jump process is

$$(\mathcal{L}_C f)(x) = \sum_{q \in \mathcal{R}} \sum_{m \in \mathrm{Match}(q,x)} a_{q,m}(x, C) \times \big[f(J_{q,m}x) - f(x) \big]. \tag{11}$$

Where microscopic positions and conformations are retained, this expression may be supplemented by diffusion, Hamiltonian or overdamped terms, provided their transformation laws are also specified. The observed reaction rate should not be inferred merely from the number of syntactic matches: automorphisms, physically indistinguishable occurrences, statistical factors, and the selected state convention all affect counting.

Assumption 4.2 (Covariance Of Propensities). For each admissible (q, m) and $g \in G$, transformed matches are in bijection and satisfy

$$a_{gq,gm}(gx, gC) = a_{q,m}(x, C), \qquad J_{gq,gm}(gx) = gJ_{q,m}(x). \tag{12}$$

The sums in Eq. (11) converge sufficiently for the operations below.

Proposition 4.2 (Covariance Of The Jump Generator). Under Assumption 4.2, for every admissible test function f,

$$(\mathcal{L}_{gC} f)(gx) = \mathcal{L}_C(f \circ g)(x), \tag{13}$$

where $(f \circ g)(x) = f(gx)$. For fixed C invariant under g, a well-defined unique Markov semigroup consequently commutes with the induced action on observables.

Proof by reindexing the sum is in Appendix B. This proposition is not a claim that the stochastic process remains symmetric in every realized history: an invariant *law* may assign positive probability to asymmetric realizations. If C varies in time, both the path C_t and its transform gC_t must enter the covariance statement.

A finite reaction-network projection illustrates the relationship to conventional kinetic chemistry. For species-count vector n and stoichiometric increments ν_q, the generator includes $f(n + \nu_q) - f(n)$, and its first moment satisfies $\frac{d}{dt}\mathbb{E}[n_t] = \sum_q \nu_q \mathbb{E}[a_q(n_t, C)]$ whenever the expectations exist. An ordinary concentration equation $\dot{c} = N v(c, C)$ requires an additional closure or large-system limit; generally $\mathbb{E}[a(n)] \neq a(\mathbb{E}[n])$. Thermodynamic consistency imposes restrictions on physically reversible channels and associated reservoirs, for example via local detailed balance (Rao and Esposito, 2016). Equivariance and thermodynamic consistency are separate conditions; one does not substitute for the other.

4.5 Symmetry Selection, Trajectories, And Function

Let $G_C = \{g \in G : gC = C\}$ be the symmetry group of a fixed environment, and set $H_x = \mathrm{Stab}_{G_C}(x)$. For a path $\gamma : t \mapsto x_t$, a history descriptor $B(\gamma)$ may include residual stabilizers H_{x_t}, order parameters $Q(x_t)$, defects, or other conserved and dissipated records. This descriptor is model dependent. A grammar derivation records rule applications, whereas a physical history descriptor may additionally record elapsed times, fluctuations, intermediate conditions, and retained material organization. In particular, two syntactic derivations need not correspond to two distinct chemical preparation histories.

Lemma 4.1 (Stabilizers Under Equivariant Flows). If a deterministic flow Φ^t_C is unique and G_C-equivariant, then every $g \in H_{x_0}$ remains in $H_{\Phi^t_C(x_0)}$ for all times where the flow is defined. If the flow also admits an equivariant inverse between the initial and evolved states (as for a well-posed finite-time Lipschitz ODE flow), then $H_{\Phi^t_C(x_0)} = H_{x_0}$ exactly. Stabilizer enlargement can occur under non-invertible evolution or in an asymptotic limit.

This constraint, proved in Appendix B, is essential when interpreting "spontaneous" symmetry breaking. Symmetry selection in chemistry can involve asymmetric seeds, thermal or intrinsic fluctuations, unstable states, finite-size behavior, and limiting procedures. Simply writing $G \to H$ does not supply such a mechanism. It is also necessary to distinguish physical global symmetries from gauge redundancies; breaking the latter is not by itself a gauge-invariant physical observable.

To formalize the paper's restricted use of *function*, let J specify a protocol (system boundary, permitted interventions, measured outputs, and timescale). Define a response law

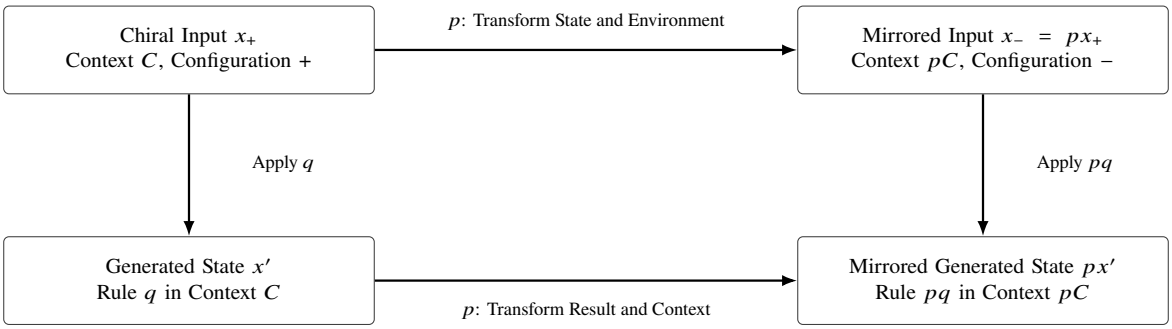


Figure 5: Equivariance as a Commuting Diagram. Mirroring the state, applicable production, and environment gives the corresponding transformed output when the specified equivariance assumptions hold. The boxes specify formal transformation data; chemical synthesis requires a separately calibrated mechanistic model.

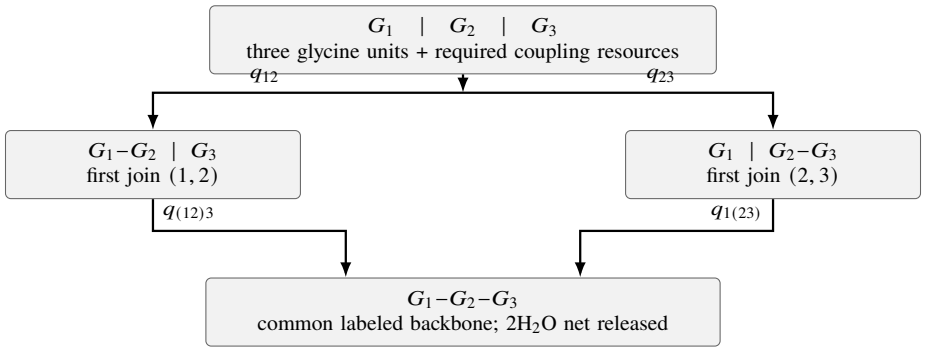


Figure 6: Derivation Histories Can Be Non-Unique for a Given Terminal Graph. Two formal orders of joining three glycine units yield the same labeled three-unit backbone with the same net stoichiometric bookkeeping. This illustrates the non-injectivity of the derivation-to-product map, a prerequisite for history-dependent functional differences and a useful starting point for an explicit physical test. Chemical realization is strengthened by activated chemistry and explicitly modeled intermediate conditions.

$P_{x,C,J}$ on an output space and write $\mu(x, C, J) = P_{x,C,J}$, or use a justified functional of that law. In a symmetric evaluation framework, transformation of context and protocol yields the corresponding transformed response; holding a chiral environment fixed while reflecting only a molecule need not preserve its response. A candidate functional effect of history must therefore be assessed through the richer terminal state x, or through an explicitly historically defined attribution, as required by Remark 3.1.

This formalism is general enough to state the first two conjectures without assigning them truth values. It specifies what a corresponding mathematical or experimental test must preserve, what it must compare, and which physical claims remain external to the grammatical construction. The second formalism instead starts with a space of relational complexes and asks whether a group-theoretic interaction model can generate organized histories which admit a justified chemical reading.

5 Relational Complexes And Group-Field-Inspired Generation

The second formalism investigates an alternative generative order: relational complexes themselves are constructed from interactions of group-labeled constituents. This approach is motivated by the combinatorial organization of group field theory (GFT), whose stranded diagrams naturally define two-complexes (Boulatov, 1992; Oriti, 2014; Oriti et al., 2014). Its role here is exploratory and specific. It supplies a way to encode *possible organization and histories of formation*, while physical chemical structures enter only after a separately defended interpretation. It is thus neither a translation of the preceding graph grammar by definition nor a quantum-gravity theory of molecular bonding.

5.1 Two-Complexes With Distinct Chemical Readings

Let $\mathcal{K}$ be a finite oriented two-dimensional cell complex with vertex, edge, and face sets (V, E, F), characteristic attaching maps, and type labels ℓ. Its cellular chain groups and boundary maps satisfy

$$C_2(\mathcal{K}) \xrightarrow{\partial_2} C_1(\mathcal{K}) \xrightarrow{\partial_1} C_0(\mathcal{K}), \qquad \partial_1\partial_2 = 0. \qquad (14)$$

The attaching maps, not solely the cardinalities $(|V|, |E|, |F|)$, determine the relevant incidence relations. For a CW complex, attaching data can represent loops and repeated incidences; a simplicial model is a more restrictive special case. Betti numbers and homology groups derived from Eq. (14) record properties of the chosen complex, but they do not by themselves identify chemical functionality.

Definition 5.1 (Relational Two-Complex Datum). A *relational two-complex datum* is $\mathcal{K}_\ell = (V, E, F, \partial, \ell, \beta)$, where ∂ abbreviates cellular attaching and incidence information, ℓ assigns declared discrete or group-representation labels, and β specifies optional boundary data. Isomorphisms preserve all declared incidence, label, orientation, and boundary structures, up to explicitly allowed relabelings.

There are at least two distinct intended readings. In a *structural complex*, vertices and edges may represent constituents and bonds while selected faces encode cycles, multi-center interactions, or motifs; these faces are representational choices to be physically justified. In a *history complex*, vertices may represent interaction events, edges may encode compositional continuation, and faces record relations among strands of such processes. Neither reading identifies a chemical ring with a face of a GFT Feynman diagram by fiat. A map from history complexes to chemistry must declare which cells are events, which are conserved components, how boundary data correspond to input/output conditions, and what information is discarded.

A key methodological consequence is that different complexes may have isomorphic boundary structures. Conversely, two isomorphic complexes can bear different admissible labels. Thus the equivalence relation on complete generative histories may be finer than equivalence of observed terminal structures. This mathematical possibility underlies an informative test of Conjecture I, although it provides no physical evidence until the complex carries a chemically calibrated realization law.

The representation step is made concrete in Fig. 7. The graph of a benzene carbon ring has six edges and a cycle; choosing to fill that cycle with a 2-cell produces a different topological model, rather than discovering a preexisting chemical "face". The distinction is essential when giving labels chemical meaning.

5.2 Group Fields As Generators Of Gluing Patterns

Choose, for a regulated construction, a finite group G or compact group with a declared measure and a finite set M of internal types. The finite $\mathbb{Z}_2$ examples below use *counting measure*, yielding $V_{\text{tet}}[1] = 2^6 = 64$. Replacing each sum by normalized Haar integration changes this value to 1; every normalization-dependent amplitude and Gaussian covariance must then be adjusted consistently. Let $d \geq 2$. Consider a complex-valued field

$$\psi : G^d \times M \longrightarrow \mathbb{C}. \tag{15}$$

A conventional GFT often uses an additional diagonal right-invariance (or projector) on its group arguments. We treat such invariance as an optional specification of the model's kinematics. It must not be conflated with a chemically observable symmetry that may be broken: the corresponding gauge redundancy has a different interpretative status.

A general regulated action takes the form

$$\begin{aligned} S_\Lambda[\psi, \bar\psi] = \langle \psi, K_\Lambda \psi \rangle + \sum_{\alpha \in A} \lambda_\alpha V_{\alpha,\Lambda}[\psi, \bar\psi] \\ + S_{\text{reg},\Lambda}[\psi, \bar\psi], \end{aligned} \tag{16}$$

where the kernel K_Λ controls pairings, each polynomial $V_{\alpha,\Lambda}$ prescribes a stranded gluing pattern, and the regulator is required to define a convergent real finite-dimensional partition integral for the specific sign-indefinite tetrahedral quartic below. The cutoff Λ can restrict representation labels or other degrees of freedom. No particular chemical bond potential is implied by this abstract action. At a fixed regulator the partition function is

$$Z_\Lambda = \int d\psi \, d\bar\psi \, \exp\bigl[-S_\Lambda[\psi, \bar\psi]\bigr]. \tag{17}$$

For a genuine probability interpretation, the regularized measure must be real, nonnegative, normalizable, and used with an explicitly suitable action. Complex amplitudes in a quantum-field expansion are not automatically reaction probabilities or physical elapsed-time weights.

Example 5.1 (A Finite-Group Tetrahedral Interaction). Take $d = 3$, $M = \{*\}$, finite G, and a real field $\psi : G^3 \to \mathbb{R}$. One possible fourth-order stranded interaction is

$$\begin{aligned} V_{\text{tet}}[\psi] = \sum_{a,b,c,d,e,f \in G} \psi(a,b,c)\psi(c,d,e) \\ \times \psi(e,b,f)\psi(f,d,a). \end{aligned} \tag{18}$$

Every index occurs twice, defining a six-strand pairing between four field factors. With an invariant quadratic part, the transformation $(U_g\psi)(a,b,c) = \psi(g^{-1}a, g^{-1}b, g^{-1}c)$ leaves this interaction invariant under simultaneous relabeling of the summed indices. A positive coercive regulator, for example $\epsilon \sum_{a,b,c} \psi(a,b,c)^6$ with $\epsilon > 0$, is required for a convergent real-field integral when the sign-indefinite tetrahedral quartic enters with nonzero coupling. The quartic is positive on constant fields and negative on other real fields, so either sign of a standalone quartic coupling admits an unbounded-below direction. The example specifies a combinatorial invariant, *not* a molecular reaction or a demonstration of spontaneous chemical symmetry breaking.

Figure 8 provides a visual dictionary for group-field arguments, strand identification, labeled two-complex formation, and the additional chemical realization map. In particular, it locates the currently unvalidated step: the transition from combinatorial histories to measured chemical states. The drawing indicates the type of gluing operation involved, rather than claiming that an arbitrary two-triangle patch is the exact Feynman amplitude of Eq. (18).

A Gaussian expansion generates stranded *Feynman graphs*; their strand faces define two-complex incidence data. For the convention $S_0[\psi] = \frac{1}{2}\sum \psi^2$, counting measure, covariance $\langle \psi_{abc}\psi_{a'b'c'}\rangle_0 = \delta_{aa'}\delta_{bb'}\delta_{cc'}$, and interaction term λV_{tet}, the precise formal series is

$$\frac{Z(\lambda)}{Z_0} = \sum_{n \geq 0} \frac{(-\lambda)^n}{n!} \bigl\langle V_{\text{tet}}^n \bigr\rangle_0 = \sum_{[\Gamma]} \frac{w_\Gamma(\lambda, G)}{|\operatorname{Aut}\Gamma|}. \tag{19}$$

Here $[\Gamma]$ ranges over admissible stranded Wick-contraction graphs, including self-contractions and disconnected vacuum components. With slot-preserving identity propagators, each closed strand contributes $|G|$, so the graph weight includes $|G|^{F(\Gamma)}$. If each quartic vertex is divided by its automorphism order four, the equivalent graph convention can be expressed with factors of $(-4\lambda)^{n(\Gamma)}$; precise graph automorphism factors follow the adopted labeled/unlabeled convention. The expansion may be formal even when a

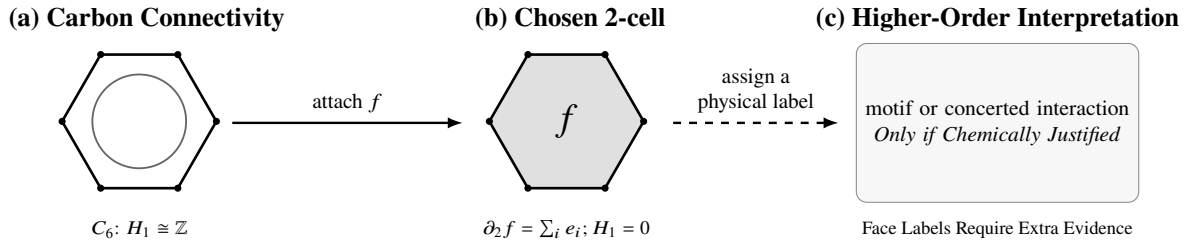


Figure 7: A molecular graph does not uniquely determine a 2-complex. The six-carbon benzene skeleton is a cycle graph (hydrogens and electron states suppressed). Attaching an oriented disk changes the homology of the *chosen representation*. A chemical meaning for the face must be specified independently; aromaticity is not explained by the attached 2-cell, and a molecular motif face is not automatically a group-field interaction-history face.

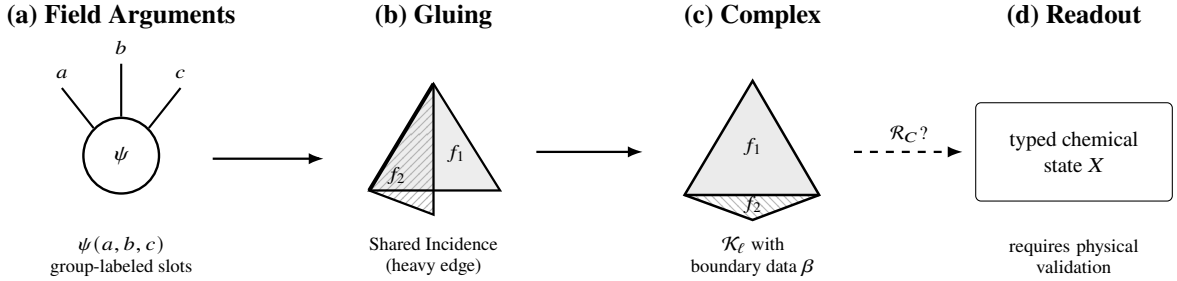


Figure 8: Group-field-inspired route, read from left to right. Group-labeled field arguments are coupled through interaction data, here schematized as Shared Incidences. The resulting labeled 2-complex is a candidate relational description. The dashed realization map requires independent chemical calibration; the figure is not a derivation of a particular physical reaction from Group Field theory.

coercively regulated integral exists. Its symmetry factors and allowed propagators belong to the field-theory definition. The existence of such a sum offers a constructive representation of relational histories, but an amplitude $\mathcal{A}_\Lambda$ cannot be interpreted as a measured yield without independent calibration. In standard GFT the complexes have a gravitational or simplicial interpretation; that physical reading does not transfer automatically to molecular or reaction complexes.

5.3 Interaction Symmetries and Gauge Constraints

The finite tetrahedral polynomial admits a larger translation symmetry than the single diagonal action used in Example 5.1. Reversing the first and third argument slots is another discrete symmetry of this particular contraction pattern. The exact continuous/full symmetry depends on field reality, restrictions, and the quadratic part. For finite G, define

$$(U_{g,h}\psi)(a, b, c) = \psi(g^{-1}a, h^{-1}b, g^{-1}c), \quad (g, h) \in G \times G. \tag{20}$$

The first and third argument slots transform together, whereas the second slot transforms independently. Each contracted index in Eq. (18) is therefore reindexed consistently. The following statement concerns the polynomial itself; the symmetry of a complete regulated action also depends on its kinetic and regulator terms.

Proposition 5.1 (Extended Symmetry of the Finite Interaction). The interaction V_{tet} is invariant under Eq. (20). If $|G| = n$ and ψ is a real array, the same contraction pattern is invariant under the larger $O(n) \times O(n)$ action $\psi'_{abc} = \sum_{a'b'c'} Q_{aa'} R_{bb'} Q_{cc'} \psi_{a'b'c'}$, for orthogonal Q, R.

The proof in Appendix C.4 uses index contraction and orthogonality. In this finite setting, $G \times G$ is a useful physically interpretable subgroup of a generally larger algebraic symmetry group. A claim about the symmetry of the *full* field theory requires an audit of the quadratic kernel, regulator, field restrictions, and chosen gauge projector.

For $G = \mathbb{Z}_2$, a diagonal right-gauge projector is

$$(P\psi)(a, b, c) = \tfrac{1}{2}[\psi(a, b, c) + \psi(a+1, b+1, c+1)]. \tag{21}$$

For the abelian group, diagonal left translation acts as the same operation on the group arguments. Consequently, $P = (I + U)/2$ projects out the entire simultaneous-translation-odd sector. This example retains a clear distinction between the unprojected toy theory and the projected, gauge-constrained theory.

5.4 Field-Level Symmetry-Selection Mechanism

Consider first the *unprojected* real field space for $G = \mathbb{Z}_2$ with counting measure. Denote its characters by $\chi_{ijk}(a, b, c) = (-1)^{ia+jb+kc}$. Under simultaneous translation U by the non-trivial group element, the odd subspace is four-dimensional, spanned by $\chi_{100}, \chi_{010}, \chi_{001}, \chi_{111}$. A complete test for broken U-symmetry is the nonzero odd field component

$$\psi_{\text{odd}} = \tfrac{1}{2}(\psi - U\psi), \qquad \text{Stab}_{\mathbb{Z}_2}(\psi) = \{e\} \iff \psi_{\text{odd}} \neq 0. \tag{22}$$

A scalar Fourier coefficient detects one sector but is insufficient for a general symmetry test. For an interaction-derived reduced example choose the second-slot coefficient

$$\eta_b[\psi] = 2^{-3} \sum_{a,b,c} (-1)^b \psi(a, b, c). \tag{23}$$

On the one-dimensional subspace $\psi_t = t\chi_{010}$, this coefficient equals t, and exact finite summation gives

$$V_{\text{tet}}[t\chi_{010}] = 64t^4, \quad \frac{\mu}{2}\sum_{a,b,c}\psi_t(a,b,c)^2 = 4\mu t^2, \\ S_{\text{red}}(t) = 4\mu t^2 + 64\lambda t^4. \tag{24}$$

For $\lambda > 0$, this *sector-restricted* polynomial has the familiar Landau form with quartic coefficient derived from the tetrahedral interaction. It is the restriction of the chosen quadratic-plus-quartic action; global integrability still requires a regulator. In comparison, $V_{\text{tet}}[t\chi_{100}] = 0$ for the first-slot mode. The original first-slot coefficient is therefore a flat direction of the quartic, while the second-slot coefficient provides the stated nonzero term.

Proposition 5.2 (Symmetric Potential And Asymmetric Minimizers). For a real scalar reduced potential $W(t) = \alpha t^2 + \beta t^4$ with $\beta > 0$, if $\alpha \geq 0$ the unique minimizer is $t = 0$. For $\alpha < 0$, the minimizers are $t_\pm = \pm\sqrt{-\alpha/(2\beta)}$. The result applies to the second-slot restriction in Eq. (24) with $\alpha = 4\mu$ and $\beta = 64\lambda$ whenever $\lambda > 0$.

The diagonal right-gauge projector P of Eq. (21) averages over the same simultaneous translation U for abelian $\mathbb{Z}_2$. It removes *every* U-odd mode, including χ_{010}: $P\chi_{010} = 0$. Accordingly, this direct quartic symmetry-selection example lives in the unprojected toy model. Nonabelian left and right translations need not coincide, but a physically meaningful order parameter would still have to respect the gauge constraints of the chosen theory. The projected relative-odd sector below illustrates an alternative, with a distinct symmetry subgroup acting on gauge-invariant fields.

Example 5.2 (Gauge-Invariant Relative Odd Sector). For $G = \mathbb{Z}_2$, let

$$\psi_z = z(\chi_{110} + \chi_{011}). \tag{25}$$

Then $P\psi_z = \psi_z$, and the residual relative operation $U_{1,0}$ of Eq. (20) changes its sign. Direct enumeration gives

$$V_{\text{tet}}[\psi_z] = 128z^4, \qquad \sum_{a,b,c}\psi_z(a,b,c)^2 = 16z^2. \tag{26}$$

The selected one-dimensional action is $8\mu z^2 + 128\lambda z^4$. Its minima can be analyzed by Proposition 5.2 with the corresponding coefficients. A full-field stability analysis, a precise residual-symmetry interpretation, and an independent chemical readout remain essential.

Thus the unprojected second-slot sector derives a positive quartic, while the projected relative-odd sector provides a gauge-compatible nonzero quartic. Their physical interpretations depend on the associated choice of symmetry and admissible field space.

5.5 Symmetry, Complex Measures, And Chemical Readout

Let $\mathscr{K}$ be the set of admissible labeled complexes and let G act on their labels in a manner preserving attaching maps. For a selected regulated model, suppose a probability measure π_C on $\mathscr{K}$ is available. This may be obtained by a valid positive ensemble construction, but is *not* identified with the general complex amplitude in Eq. (19). The physical interpretation is represented by a stochastic kernel

$$R_C(dx \mid \mathcal{K}_\ell), \qquad x \in \mathcal{X}, \tag{27}$$

which returns a distribution of terminal chemical states given a labeled complex and a declared context. Deterministic readout is the special case of a point mass. This formulation makes unresolved or many-to-one chemical interpretations visible instead of hiding them inside a putative bijection.

Assumption 5.1 (Equivariant Realization Law). The complex ensemble and readout kernel obey, for every measurable $A \subseteq \mathcal{X}$,

$$\pi_{gC}(gD) = \pi_C(D), \\ R_{gC}(gA \mid g\mathcal{K}_\ell) = R_C(A \mid \mathcal{K}_\ell), \tag{28}$$

for measurable $D \subseteq \mathscr{K}$. The kernel is concentrated on states compatible with explicitly declared chemical accounting and admissibility conditions.

Proposition 5.3 (Covariance Of Generated Structures). Under Assumption 5.1, the induced chemical output law

$$Q_C(A) = \int_{\mathscr{K}} R_C(A \mid \mathcal{K}_\ell)\,\pi_C(d\mathcal{K}_\ell) \tag{29}$$

is equivariant: $Q_{gC}(gA) = Q_C(A)$.

The complete measure-theoretic argument is given in Appendix C. The result identifies a precise way to propagate an *assumed* symmetry through generation and interpretation. It does not assert that the symmetry assumptions are fulfilled for any specific chemical preparation, or that a symmetric ensemble forces each sampled state to be symmetric. The relevant failure cases include a non-invariant context, a symmetry-violating regulator, and a chemically invalid readout.

5.6 Conditions for Relating Two Formalisms

A derivation in Formalism I is a composable sequence of grammar rules and state changes. A history complex in Formalism II is a labeled incidence structure with a potentially different notion of order, locality, and amplitude. The frameworks can be compared after selecting a common observational domain and an explicit bridge. Let $\mathscr{H}_{\mathbb{G}}$ denote an appropriately specified class of grammatical histories, and consider a candidate encoding $F : \mathscr{H}_{\mathbb{G}} \to \mathscr{K}$ and chemical semantics $E_C : \mathscr{H}_{\mathbb{G}} \to \mathcal{P}(\mathcal{X})$. A desirable compatibility condition is

$$E_C(h) = R_C(\cdot \mid F(h)), \qquad h \in \mathscr{H}_{\mathbb{G}}, \tag{30}$$

possibly up to a quantified approximation error in an appropriate probability metric. For histories with multiple possible physical realizations, F may itself need to be stochastic or relational.

Definition 5.2 (Faithful Generative Correspondence). A *faithful correspondence at resolution r* is a bridge satisfying a declared form of Eq. (30) and preserving the stipulated structural observations, group action, and composition of histories on their common domain. Faithfulness here is relative to these observations and morphisms, rather than a claim that all microscopic histories or physical states are recoverable.

Nothing in the construction guarantees such a bridge. Two distinct grammatical histories may map to the same complex; the same complex may admit several inequivalent chemical interpretations; and a field-theoretic amplitude need not

define a time-ordered stochastic trajectory. Thus equivalence of the two formalisms is a *research question*, not a premise. One may instead seek a restricted correspondence for a benign reaction network, a simplified ordering transition, or a finite labeled assembly model and derive the implications of its exact assumptions.

The common symmetry-first orientation has therefore been realized in two mathematically separate ways. Formalism I begins with a rewriting language and a chemically interpreted dynamical generator. Formalism II begins with group-dependent field interactions and relational incidence structures, then requires an independent chemical readout. Neither makes group symmetry an absolute ontological primitive. Both can be useful even if their generative objects are merely effective representations. The next part of the manuscript will examine how these constructions bear on the three conjectures, and which questions can be settled by theorem, by modeled counterexample, or only through experiment.

6 Functional Differentiation Across Generative Histories

The first conjecture concerns the gap between a structural description of a chemical product and the experimentally accessible consequences of its preparation. Chemical constitution, molecular or network topology, three-dimensional arrangement, defects, configurational populations, and the dynamical state retained at the end of preparation are not interchangeable descriptions. In particular, a coarsened structural representation may identify two products that subsequently behave differently. Pathway complexity and kinetic trapping in supramolecular assemblies supply a well-studied physical motivation for distinguishing preparation protocols from terminal structural summaries (Sorrenti et al., 2017; Wehner and Würthner, 2020). The present conjecture asks a narrower, falsifiable question: can *symmetry-breaking histories* help distinguish functional outcomes among such structurally equivalent products, and can the relevant effect be traced to retained physical variables?

6.1 Equivalence Of Structures And Distinction Of Histories

Fix an admissible history space $\mathscr{H}_C$ for context C, together with an endpoint map $e_\tau : \mathscr{H}_C \to \mathcal{X}$. A history can be a piecewise continuous physical trajectory, a realization of a stochastic reaction process, or a history represented by an interpreted relational complex. The last case requires the readout assumptions of Section 5; a bare field-theoretic diagram is not yet a chemical history. Let $T_r : \mathcal{X} \to \mathcal{S}_r$ be a structural representation at resolution r, whose codomain is equipped with an isomorphism relation preserving declared attributes. The quotient of its image by isomorphism is assigned the induced final σ-algebra, as specified in Appendix D.

Definition 6.1 (Terminal Structural Equivalence). Two histories $h_1, h_2 \in \mathscr{H}_C$ are *terminally structure-equivalent at resolution* r, written $h_1 \sim_r h_2$, exactly when

$$T_r(e_\tau(h_1)) \cong T_r(e_\tau(h_2)). \tag{31}$$

This equivalence is relative to the declared labels, topology, observation scale, and admissible isomorphisms.

By contrast, define $B : \mathscr{H}_C \to \mathscr{B}$ to record an explicitly selected symmetry-breaking descriptor: for example, an order-parameter history, a sequence of residual stabilizers, or the formation and coarsening of topological defects. A comparison $B(h_1) \neq_B B(h_2)$ is meaningful only after the descriptor and its equivalence relation have been specified. Two paths with different descriptions need not correspond to physically different preparations; symmetry-related relabelings, reparameterizations, and differences below instrumental resolution may be immaterial. Nor does $B(h_1) \neq B(h_2)$ by itself imply different products or functions.

The central physical distinction is visualized in Fig. 9. Two hypothetical materials share the same declared unit connectivity while retaining different orientational domain patterns, schematically attributed to different preparation histories. Their physical response can be compared under the *same* terminal environment and protocol. The figure does not assume that structural topology already encodes the retained order or that the two preparations are known experimentally.

6.2 Response Kernels And The Factorization Question

Write $K_{C,J}(dy \mid x)$ for the probability law of an observable functional output y under a fixed environment C and an admissible evaluation protocol J. The kernel can include a prescribed duration of subsequent dynamics and, when relevant, permitted interventions. Let $\mu_{C,J}(x) = K_{C,J}(\cdot \mid x)$ denote this operational meaning of functional semantics. The central mathematical question is whether the response factors through the terminal structural representation.

Definition 6.2 (Structural Sufficiency For Function). The representation T_r is *functionally sufficient* on an admissible state domain $D \subseteq \mathcal{X}$, relative to (C, J), if there exists a kernel $\overline{K}_{C,J}$ on the isomorphism classes of $\mathcal{S}_r$ such that

$$K_{C,J}(\cdot \mid x) = \overline{K}_{C,J}(\cdot \mid [T_r(x)]) \tag{32}$$

for every $x \in D$. All such statements are conditional on the chosen observable and context.

Proposition 6.1 (Criterion For Structural Sufficiency). A response kernel factors as in Eq. (32) on D if and only if it is constant on each fiber of the map $x \mapsto [T_r(x)]$ within D. In particular, two states in one fiber with unequal response laws prove that T_r is insufficient for the designated functional protocol.

The complete set-theoretic proof, including the conditions for a measurable quotient kernel, is given in Appendix D. This criterion turns a proposed semantic distinction into an identifiable failure of a particular structural coarse-graining. It does not require the conclusion that chemistry possesses an additional, nonphysical semantic force.

In a stochastic preparation, let ν_i be the endpoint-state law of preparation protocol i. Its predicted response is

$$P_i(dy) = \int_{\mathcal{X}} K_{C,J}(dy \mid x)\, \nu_i(dx). \tag{33}$$

Even if both protocols produce the same structural representation with probability one, P_1 and P_2 may differ if their

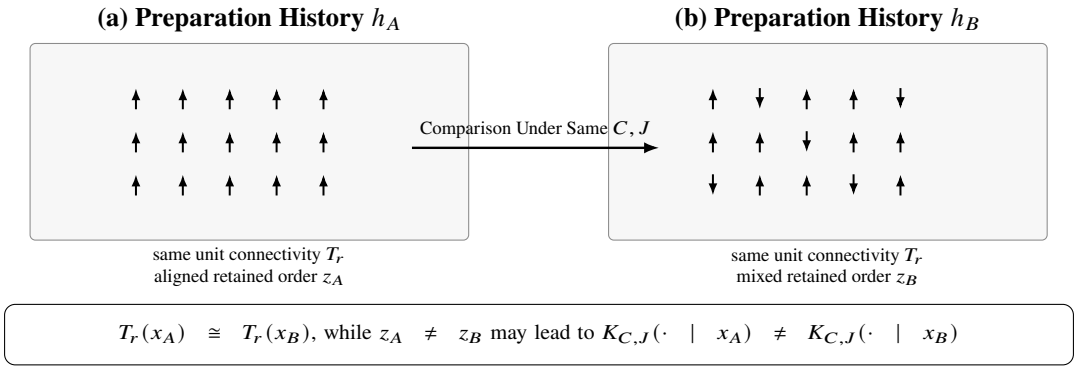


Figure 9: A visually testable version of Conjecture I. Two hypothetical histories produce the same declared coarse connectivity and composition but leave different mesoscale orientational domains (arrows). A symmetry-sensitive test may distinguish the retained states. The panel is a toy materials schematic rather than data from a particular compound; a real comparison must control preparation confounders and report the resolution of the declared equivalence.

endpoint distributions occupy distinct parts of that structural fiber. Conversely, different endpoint distributions need not give different functional laws if the protocol cannot distinguish them. In empirical work, the comparison should be quantified by an appropriate distance $d_J(P_1, P_2)$ and a predeclared uncertainty threshold, as in Eq. (2).

6.3 First Conjecture and a Discriminating Construction

Conjecture 6.1 (Generative-History Dependence). There exist physically realizable chemical preparations h_1, h_2 and a common functional testing protocol (C, J) for which (i) their terminal chemical structures are equivalent at a specified resolution r, (ii) their physically interpreted symmetry-breaking histories are distinguishable, and (iii) their functional response laws differ beyond a justified tolerance. The functional difference is mediated by retained physical organization or by an explicitly history-indexed functional attribution.

The conjecture's first clause specifies a *coarse* structural equivalence; an exact equality of full physical state and environment, under a complete Markovian description, would remove the proposed physical dependence on preparation history (Remark 3.1). The reference to functional attribution additionally permits an observer to assign different historical roles to physically identical outcomes, but such a difference belongs to a declared interpretation and cannot be passed off as a difference in physical response.

Example 6.1 (A Minimal Retained-Order Model). Let $z \in \{-1, +1\}$ denote a retained domain orientation generated by alternative symmetry-selection histories, and let s denote a common coarse topology. Set $x_+ = (s, +1)$ and $x_- = (s, -1)$ with $T_r(x_+) = T_r(x_-) = s$. Consider a fixed, symmetry-sensitive evaluation environment under which a binary response Y obeys

$$\Pr(Y = 1 \mid x_\pm, C, J) = \frac{1 \pm \delta}{2}, \quad 0 < \delta \leq 1. \tag{34}$$

The total-variation distance between these two Bernoulli laws is δ, although their observed topologies coincide. This is a mathematically consistent countermodel to the universal sufficiency of T_r, not an experimentally identified material or a derivation of the response from molecular physics.

The example deliberately makes the hidden physical variable explicit. Such variables could include domain orientations, defect networks, conformational populations, residual strain, or long-lived supramolecular aggregates, depending on the selected chemical system. The scientific challenge is to show that a measured $B(h)$ predicts a retained feature z and that the feature, rather than a confounded preparation condition, predicts the functional response. The literature on non-equilibrium supramolecular polymerization suggests plausible test beds but does not establish this specific cross-formalism conjecture (Sorrenti et al., 2017).

6.4 Criterion for Physical and Inferential Adequacy

A prospective test should first choose a harmless, well-characterized assembly or material transition and declare the resolution of structural identity. Next, two preparation protocols should be compared under the same final environment and testing conditions. The histories should differ in an independently observable symmetry-selection variable, rather than merely being assigned different verbal descriptions. Structural and response measurements should then establish whether $d_J(P_1, P_2) > \varepsilon_J$. Finally, candidate retained variables should be measured and included in the terminal-state representation; if the difference disappears once T_r is enriched, that is positive evidence for a *physically mediated* history effect rather than a failure of causal closure.

For Formalism I, a history is a sequence of allowed contextual rewritings and physically calibrated stochastic events. A history descriptor can be evaluated along that sequence. For Formalism II, it is an interpreted labeled complex with an explicit temporal or process reading. Distinct complexes may share a terminal readout; whether the distinction is functionally predictive depends on the readout kernel and on extra physical labels. Neither the abstract non-injectivity of a complex-to-boundary map nor the existence of several grammatical derivations suffices to establish Conjecture 6.1. The latter requires a chemically meaningful mechanism and observation.

7 Functional Emergence Through Symmetry Breaking

The second conjecture moves from a comparison of two terminally equivalent preparations to a comparison of the *functional possibilities* of symmetric and symmetry-reduced organizations. The mathematical fact that a group-invariant law can admit non-invariant states is familiar from bifurcation theory and phase transitions. The scientific question pursued here is more selective: under which conditions does a change of residual symmetry open a measurable response channel, a distinguishable interaction, or another functional role that is unavailable to a specified symmetric reference? Asymmetric autocatalysis and nonlinear amplification provide relevant chemical contexts for investigating symmetry selection, but do not establish a universal route from broken symmetry to functional emergence.

Figure 10 juxtaposes a mirror-symmetric autocatalytic rule family with a standard pitchfork normal form. The two panels are intentionally distinguished: the rate equations for a real or idealized chiral network require their own kinetic derivation, whereas the normal form illustrates the possibility of asymmetric stable branches in an equivariant system. A separate functional measurement is required to claim functional emergence.

7.1 Residual Groups And The Interpretation Of Breaking

Take a group G acting on physical or effective states, and let $G_C = \{g \in G : gC = C\}$ be the subgroup preserving the declared environment. For $x \in \mathcal{X}$, write $H(x) = \mathrm{Stab}_{G_C}(x)$. The useful comparison is between an initial reference x_0 and a reachable state x_f with $H(x_f)$ a proper subgroup of $H(x_0)$, possibly after identifying conjugate isotropy groups. The subgroup condition is not the only possible pattern: residual groups may be non-nested, change along trajectories, or be meaningful only statistically. Statements about broken symmetry must consequently name both the acting group and the state or ensemble whose invariance is at issue.

Definition 7.1 (Operational Functional Capability). For a fixed environment C, observation family J, admissible intervention set $\mathcal{A}$, and tolerance ϵ, define $\mathcal{F}_{C,J,\mathcal{A},\epsilon}(x)$ as the set of predeclared response tasks that the state x can realize under interventions in $\mathcal{A}$, according to a measurable success criterion with tolerance ϵ. Comparisons of capability are valid only when the reference and realized states are evaluated with the *same* task family, intervention rules, and success thresholds.

This definition does not equate every physical difference with a new function. A new function requires a task that was specified independently of the state transition and whose operational performance is demonstrably changed. The functional set can include selectivity, signal discrimination, directional transport, or persistent coupling to an external perturbation. It need not be a set of anthropocentric purposes; the choice of tasks is nevertheless an explicit part of the interpretation.

Conjecture 7.1 (Symmetry-Breaking Functional Emergence). There exist physically realizable chemical systems and specified contexts in which a generative process reduces the residual physical symmetry of an initially symmetric reference and thereby enables at least one operational function unavailable to that reference under the same admissible testing conditions:

$$H(x_f) \subsetneq H(x_0), \qquad \mathcal{F}(x_f) \setminus \mathcal{F}(x_0) \neq \varnothing. \tag{35}$$

The claimed enabling relation must be assessed against alternative concomitant changes in the physical state and environment.

The notation $\mathcal{F}(x)$ in Eq. (35) abbreviates the indexed capability set of Definition 7.1. As an existence hypothesis this conjecture is compatible with systems in which symmetry breaking produces no new evaluated function. In contrast, the universal proposition that every symmetry reduction creates a useful function is unwarranted. The central objective is to identify conditions for the *possibility and mechanism* of functional differentiation.

7.2 An Odd-Observable Result

A rigorous minimal link between symmetry and function can be obtained from a response observable that transforms nontrivially under the group. Consider the case $G_C = \mathbb{Z}_2 = \{e, p\}$ and a real observable $q : \mathcal{X} \to \mathbb{R}$ that is odd: $q(px) = -q(x)$. The observable might represent a signed polarization or an enantiomeric excess, provided its interpretation has been physically justified.

Proposition 7.1 (Symmetric States Suppress Odd Observables). If $px_0 = x_0$ and $q(px) = -q(x)$, then $q(x_0) = 0$. Thus a prescribed response task requiring $|q| > \delta$, with $\delta > 0$, is inaccessible to the symmetric reference state under a protocol that measures this q without altering the reference before measurement.

A proof is given in Appendix E. The second statement is explicitly protocol-dependent: an intervention that first breaks the symmetry may enable an initially symmetric system to produce a nonzero response, changing the capability comparison. The result establishes a *selection rule*, rather than a physical prediction that the system will generate a state with $q \neq 0$. In a finite stochastic system an invariant probability distribution can have $\mathbb{E}[q] = 0$ even when individual samples display $q \neq 0$; the distinction between mean symmetry and trajectory-wise asymmetry is essential.

7.3 A Bifurcation Model And Its Limits

For an illustrative scalar order parameter m odd under $m \mapsto -m$, consider the reduced dynamics

$$\dot{m} = am - bm^3, \qquad b > 0. \tag{36}$$

When $a < 0$, $m = 0$ is locally attracting. For $a > 0$, the origin is unstable and the two nonzero equilibria $m_\pm = \pm\sqrt{a/b}$ are attracting. The associated reflection symmetry of the vector field remains exact. A symmetry-related pair of selected states may thus exist even though the equations are symmetric. These statements follow by solving for equilibria and testing the derivative of the vector field; detailed calculations appear in Appendix E.

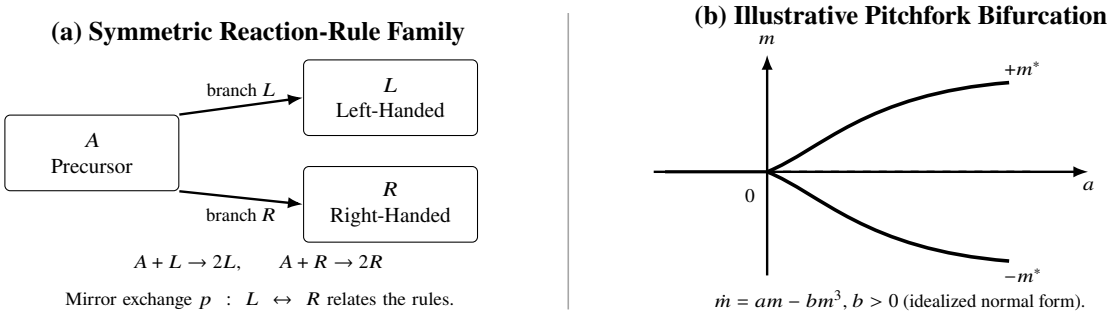


Figure 10: Rule symmetry and asymmetric realizations. A schematic autocatalytic grammar admits two rules related by reflection. In a suitable nonlinear regime a small initial bias or fluctuation may be amplified. The independent pitchfork panel illustrates a generic symmetry-breaking normal form, a generic normal form to be compared with independently derived chemical kinetics. Functional emergence requires an additional response protocol.

Suppose a separately calibrated chemical readout assigns $q(x_m) = m$. For a fixed threshold $0 < \delta < \sqrt{a/b}$, a prepared selected branch m_+ or m_- can meet a signed-response task inaccessible to $m = 0$ under the restricted protocol of Proposition 7.1. This supplies a *conditional model* of the conjecture's proposed mechanism. It neither identifies a specific chemical system nor shows that the cubic normal form applies to asymmetric autocatalytic reactions in general. Moreover, an exact symmetric initial condition remains at $m = 0$ under deterministic uniqueness; branch selection requires an initial perturbation, stochastic fluctuations, boundary bias, or another explicitly modeled source.

A more informative chemical model would derive reaction propensities or effective free-energy terms from physically constrained interactions, identify the order parameter, estimate its distribution under controlled conditions, and link it to a functionally discriminating observable. In Formalism II, the field-level construction of Eq. (24) offers an analogous symmetry-selecting sector. Its chemical significance still depends on the realization kernel in Eq. (27); Proposition 5.3 alone does not supply a functionally separating readout.

7.4 Defects, History, And Conditional Semantics

A broken-symmetry phase can have a space of degenerate order parameters with a homogeneous-space form G/H when the required group-action assumptions apply. The homotopy groups of that space can classify candidate defect types in appropriate spatial dimensions. Yet classification does not determine the density, arrangement, or lifetime of actual defects; these depend on dynamics, boundary conditions, and preparation history. Consequently, the same residual subgroup may coexist with distinct mesoscale organizations and functions. This is the point at which Conjecture 7.1 joins, without collapsing into, Conjecture 6.1.

Remark 7.1 (Three Distinct Assertions). The following assertions must be evaluated separately: (i) an interaction law possesses a group symmetry; (ii) a realized state or phase selects a smaller stabilizer; (iii) the selected organization enables a new operational functional role. Neither (i) implies (ii) nor (ii) implies (iii) without further assumptions. Moreover, symmetry-related outputs may be functionally equivalent when the entire environment and testing protocol are transformed together.

Important experimental antecedents already establish chiral symmetry selection and amplification in specific chemical systems. Stirring sodium chlorate solutions produced strongly homochiral crystal populations in the experiments of Kondepudi et al. (1990), and Soai et al. (1995) demonstrated amplification of enantiomeric excess by asymmetric autocatalysis. These observations motivate the mechanism-level inquiry here. The conjecture is therefore interpreted in its *stronger* sense: a specified symmetry-breaking generative mechanism should be connected, through a calibrated response protocol, to an identifiable enlargement or differentiation of functional capabilities relative to an explicit reference organization. Demonstrating chirality or optical activity alone would be a precursor rather than a new verification of that stronger claim.

The original scientific contribution sought here is therefore a family of *conditional links* between group structure, possible generated organizations, and functionally defined observations. The conjecture becomes substantive only when a system is specified for which the links can be derived or independently tested. Chemical chirality, collective phase organization, and supramolecular pathway selection provide possible case classes; their similarities should not erase the important differences in conservation laws, timescales, order parameters, and experimental accessibility.

8 Ontological Status of Generative Order

The third conjecture serves a different role from the first two. It constrains the interpretation of successful mathematical constructions. Formalism I uses an equivariant grammar to describe admissible transitions; Formalism II generates labeled relational complexes through group-structured fields and an additional chemical interpretation. Both are *symmetry-first* models in an explicitly relative sense. Their success, even if established in particular domains, would not by itself demonstrate an independently existing, universal, or ontologically prior chemical meta-grammar. This methodological restraint applies equally when a linguistic formalism is taken as evidence for a biologically realized Universal Grammar; the latter is a distinct empirical thesis about language capacity, not a corollary of generativity or structural regularity (Chomsky, 2005).

8.1 Three Existence Claims

It is useful to distinguish the claims H_R, H_E, and H_O. The *representational* claim H_R asserts that a mathematical generative description can be written and applied to a specified domain. The *explanatory* claim H_E asserts that the description yields a successful explanatory or predictive account under independently assessable criteria. The stronger *ontological* claim H_O asserts that the relevant meta-generative order has the proposed independent physical or causal standing. An arbitrary enumeration of chemical structures can satisfy a weak version of H_R with little explanatory significance. Demonstrating H_E requires nontrivial constraints, predictions, or informative unification. Inferring H_O needs an additional account of what ontic structures the model refers to and what could distinguish this account from alternatives.

Definition 8.1 (Substantive Meta-Generative Hypothesis). For a specified chemical domain D, $H_M(D)$ is the proposition that a nontrivial meta-generative organization exists in the stipulated *physical* sense and plays a specified explanatory or causal role across a family of effective generative mechanisms on D. The hypothesis must declare what counts as independence, unity, and causal or nomological standing. It is distinct from the mathematical existence of a grammar that enumerates or generates formal chemical descriptions.

One may reject the proposed independence of $H_M(D)$ while accepting all relevant quantum, thermodynamic, structural, and kinetic regularities. Conversely, one may regard a shared physical generator as a real component of an explanatory hierarchy. The models disagree over an explicit ontological commitment, not over whether matter obeys physical constraints. Accordingly, the negation $\neg H_M(D)$ is *not* the denial of every chemical law or every grammar representation.

8.2 Conjecture of Relative Existential Underdetermination

Conjecture 8.1 (Ontological Underdetermination Of Generative Order). For at least one scientifically relevant chemical domain D and nontrivial, properly specified meta-generative hypothesis $H_M(D)$, there may exist two physically coherent models with distinct ontological commitments, one satisfying $H_M(D)$ and the other satisfying its negation, that agree on a substantial declared class of chemical structural and dynamical observations. Under those conditions, regularity and generativity alone fail to determine the existence of an ontically independent universal meta-generative structure.

The first sentence is explicitly an investigable hypothesis about the availability of serious competing models, rather than a theorem about all possible chemical theories. Once a pair of models satisfying the premises of Definition 3.1 has been demonstrated, Proposition 3.1 gives the corresponding logical non-entailment. If the pair also agrees on all stipulated experimental protocols within tolerance, Eq. (5) gives relative empirical underdetermination. Neither observation implies that no future protocol can separate the models.

8.3 Finite Contrastive Construction

A toy example isolates the logical mechanism without pretending to settle metaphysics. Consider a two-state effective chemical variable $x \in \{0, 1\}$ and a set $\mathcal{P}_0$ of presently admissible observation protocols. Let two stochastic models $\mathfrak{M}_+$ and $\mathfrak{M}_-$ have exactly the same transition law K_0 for every preparation and observation in $\mathcal{P}_0$. In $\mathfrak{M}_+$, introduce a latent rule-selector $z \in \{0, 1\}$ whose effects are suppressed for these protocols but which changes the effective transition under a new context C_*. In $\mathfrak{M}_-$, use a flat context-independent transition law equal to K_0. The models then agree on $\mathcal{P}_0$ but can disagree when a protocol activates C_*. If the ontological hypothesis attributes substantive meta-generative status to the independently modeled selector z, these are contrasting model *interpretations* over the stated evidence. Their empirical difference at C_* makes the construction stronger than attaching a permanently inert metaphysical label to an otherwise identical model.

Proposition 8.1 (Restriction-Relative Observational Equivalence). Suppose two stochastic models have identical initial observed-state laws, transition kernels, and observation kernels for every protocol in a declared class $\mathcal{P}_0$. Then their induced output laws coincide under all finite observation schedules in $\mathcal{P}_0$. Equality need not persist for protocols outside that class.

An explicit finite construction and proof are given in Appendix F. The proposition itself proves only a conditional fact about restricted evidence; the assignment of an ontologically substantive H_M to a hidden mechanism must be justified separately. In particular, adding an arbitrary unused symbol z would not provide a credible case of ontological underdetermination. A scientific alternative should identify plausible physical dependencies, an intelligible theory of the latent layer, and conditions under which the rival accounts might diverge.

8.4 Implications For Chemical Meta-Grammars And Universal Grammar

A formally adequate chemical meta-grammar may organize a family of molecular grammars and allow economical generation of structures. The grammar is then a genuine mathematical object with some degree of descriptive utility. Whether the natural world implements that meta-grammar as a distinct mechanism remains another proposition. Similarly, a shared structural or learning constraint in human languages is compatible with multiple accounts of its source, including domain-specific biological mechanisms and more general computational, developmental, or interactional constraints. The existence and content of a particular Universal Grammar require their own empirical and theoretical support; treating the presence of structural regularities as a proof would conflate the logical levels.

Nor does a common symmetry-first approach resolve the ontological matter. The symmetry group in Formalism I is an explicit modeling input constraining rewrite rules; the group in Formalism II appears in the domain and interactions of a field whose chemical interpretation is provisional. The physical roles of these groups can differ, and a further theory might account for either symmetry as emergent from more general relations. Symmetry's *relative priority inside a model* therefore supplies an explanatory starting point, while Conjecture 8.1 cautions against treating that priority as an absolute origin.

Remark 8.1 (Scope Of The Negative Claim). The paper contests an inference from observed structural regularity or successful generative representation to a uniquely privileged ontic meta-grammar. It neither disproves the existence of such a grammar nor denies that future evidence might justify a stronger realist interpretation. A claim of permanent undecidability would require substantially stronger assumptions than the present contrastive methodology provides.

The three conjectures thus have complementary scopes. Conjectures 6.1 and 7.1 concern possible chemical structure–function relations and their mechanisms; Conjecture 8.1 concerns the limits of inference from modeled mechanisms to the existence of a universal higher-order generator. The third clarifies the epistemic scope of verifying the first two.

9 Case Studies: Two Finite Generative Constructions

The preceding sections define two generators in abstract terms. The following miniature cases provide concrete test objects for independent reproduction. Both use chemically familiar structures, exact finite combinatorics, and declared symmetry actions. They test the syntax and mathematical consistency of the proposed machinery; any claim about reaction kinetics, chemical synthesis, or measured functionality would additionally require physical calibration. The examples are intentionally small enough for exhaustive enumeration.

9.1 Case Study I: Group-Equivariant Grammar for a Tetrahedral Stereocenter

Consider the constitutional graph of bromochlorofluoromethane, CHBrClF. It contains one carbon vertex connected to four distinguishable terminal substituents H, F, Cl, and Br. Its constitutional graph is the labeled star $K_{1,4}$. This case highlights the difference between constitutional isomorphism and an orientation-sensitive chemical-state representation.

Take $\Sigma_T = \{\mathrm{C}, \mathrm{H}, \mathrm{F}, \mathrm{Cl}, \mathrm{Br}, \text{single bond}, \tau = +1, \tau = -1\}$ and $\Sigma_N = \{A_{\mathrm{tet}}\}$, where the start graph $S = A_{\mathrm{tet}}$ has four typed attachment ports. Let p_+ and p_- each replace the nonterminal with the same labeled atom-and-bond star, together with one of two orientation annotations $\tau \in \{+1, -1\}$. The nonterminal expands into one carbon with four labeled ligands; the prescribed geometric embedding places these ligands at a nondegenerate tetrahedral arrangement. For the fixed ligand order (H, F, Cl, Br), a possible orientation test is

$$\tau(X) = \operatorname{sgn} \det\big[\mathbf{r}_{\mathrm{F}} - \mathbf{r}_{\mathrm{H}},\ \mathbf{r}_{\mathrm{Cl}} - \mathbf{r}_{\mathrm{H}},\ \mathbf{r}_{\mathrm{Br}} - \mathbf{r}_{\mathrm{H}}\big]. \tag{37}$$

Translation and proper rotation preserve this sign, while reflection changes it. Odd permutations of the declared ligand order also reverse the sign. Assigning the stereochemical descriptor (*R*) or (*S*) requires the CIP-priority convention and a consistent viewing/orientation convention; the sign τ is kept distinct from the CIP descriptor in these tests. The group $G = \mathbb{Z}_2 = \{e, \varrho\}$ acts on enriched terminal states by $\varrho X_+ = X_-$ and on production schemas by $\varrho p_+ = p_-$, with the context transformed where required. Here ϱ denotes reflection, whereas $p_\pm$ are production symbols. The grammar first generates two τ-labelled terminal graph classes; their embeddings into three-dimensional geometry are obtained through a separate realization map $\mathcal{E}$ respecting the selected orientation tags:

$$L^{\tau}_{\mathrm{term}}(\Gamma) = \{[T_+]_{\cong_\tau}, [T_-]_{\cong_\tau}\}. \tag{38}$$

The map $\mathcal{E}(T_\pm)$ yields two proper-rotation ($SE(3)$) classes when nondegenerate geometries are chosen. Both become one class under the constitutional projection T_{conn}; their enriched configurations form one two-element $\mathbb{Z}_2$ orbit. The construction is a grammar for *structural possibilities*; it does not identify an experimentally resolved route to either enantiomer.

Table 1: Case Study I: explicit grammar data and generated outputs.

Item	Concrete Choice in the Tetrahedral Example
Terminal alphabet Σ_T	{C, H, F, Cl, Br, single bond, $\tau = +1$, $\tau = -1$}
Nonterminal alphabet Σ_N	one four-port symbol A_{tet}
Start graph S	one copy of A_{tet} with ordered ports $(u_{\mathrm{H}}, u_{\mathrm{F}}, u_{\mathrm{Cl}}, u_{\mathrm{Br}})$
Production schemas	p_+ and p_-, each replacing A_{tet} by one carbon center, four ligand vertices, and one orientation label $\tau = \pm 1$
Symmetry group	$G = \mathbb{Z}_2 = \{e, \varrho\}$ acting by reflection: $\varrho X_+ = X_-$ and $\varrho p_+ = p_-$
Terminal outputs	two enriched states X_+ and X_- with the same constitutional projection $T_{\mathrm{conn}}(X_+) \cong T_{\mathrm{conn}}(X_-)$
Verification targets	vertex count 5, bond count 4, two τ-labelled graph classes; two realized $SE(3)$ classes, one reflection orbit

How Formalism I Models the Example. The tetrahedral case can be encoded by the data listed in Table 1. The point of the example is not synthetic realism; it is the explicit separation of (i) a formal grammar object, (ii) an enriched state carrying orientation, and (iii) a response test that may or may not distinguish the two outputs.

How the Generation Proceeds. The generation can be read as a one-step derivation with two admissible outputs:

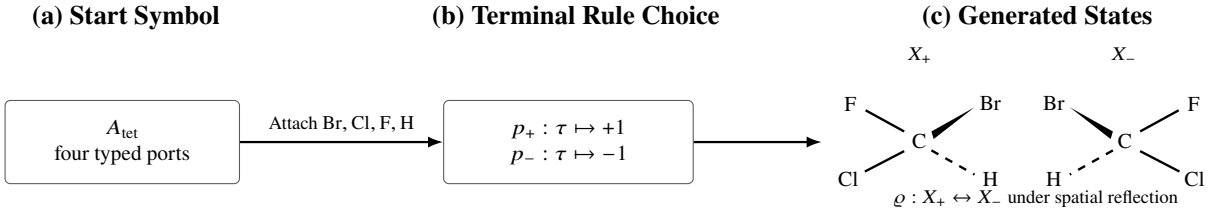
$$S = A_{\mathrm{tet}} \overset{p_+}{\Longrightarrow} X_+, \qquad S = A_{\mathrm{tet}} \overset{p_-}{\Longrightarrow} X_-. \tag{39}$$

A convenient explicit replacement schema is

$$A_{\mathrm{tet}}(u_1, u_2, u_3, u_4) \mapsto \Big(C(v),\ \{(v, a_i)\}_{i=1}^{4},\ \tau = \pm 1\Big), \tag{40}$$

where (a_1, a_2, a_3, a_4) are the ligand vertices carrying the labels (H, F, Cl, Br) and the sign τ records the oriented tetrahedral embedding. The constitutional graph component is the same in both productions; the enriched state differs through the orientation tag or, equivalently, through a chosen tetrahedral embedding. For implementation, one can store the formal output as an attributed graph together with either a signed orientation invariant or explicit coordinates that realize the sign in Eq. (37).

What the Example Generates. The grammar generates two enriched terminal states, one reflection orbit, and one constitutional isomorphism class. If a response protocol keeps only T_{conn}, the two outputs collapse into one class. If the protocol also reads the orientation-sensitive part of the state, then the pair remains distinguished. This is the smallest example in the paper where a group action, a finite terminal language, and a chemically familiar structure appear simultaneously.



The constitutional star has 5 atomic vertices and 4 bonds; the enriched terminal domain distinguishes the two orientations $\tau = \pm 1$.

Figure 11: Case Study I: Two Mirror-Related Terminal Structures. A four-port nonterminal is resolved by attaching four distinct chemical substituents and assigning an orientation. The two terminal graph embeddings share their atom-and-bond connectivity and form one orbit under the abstract reflection action. This is a schematic grammar construction; orientation labels represent geometry beyond a plain constitutional graph.

Proposition 9.1 (Terminal Orbit of the Tetrahedral Example). For the finite construction above, assuming an ideal nondegenerate tetrahedral embedding with the labeled ligand order fixed, the terminal configurations form two proper-rotation equivalence classes and one reflection orbit. Their constitutional projections are isomorphic, and $\tau(pX) = -\tau(X)$.

The proof is the determinant transformation law applied to the fixed ligand ordering; a detailed derivation and reproducibility checks appear in Appendix L. In a fixed chiral test environment, the two outputs may admit different response kernels $K_{C,J}(\cdot \mid X_+)$ and $K_{C,J}(\cdot \mid X_-)$. In a parity-symmetric environment with a parity-even scalar observable, a corresponding covariance condition instead makes them equal. Both predictions depend on the stated evaluation protocol, and neither is inferred from connectivity alone.

9.2 Case Study II: Six-Carbon Cycle and a Finite Group-Field Complex

The second case begins with the six-carbon connectivity cycle of benzene, written here only as a *carbon skeleton*. Let $V = \{v_1, \ldots, v_6\}$ and oriented edges $e_i : v_i \to v_{i+1}$ with indices modulo six. As a graph, the chosen skeleton has $|V| = |E| = 6$ and $H_1 \cong \mathbb{Z}$. If a structural representation attaches one oriented two-cell f along the cycle, then

$$\partial_2 f = e_1 + \cdots + e_6, \qquad \partial_1 \partial_2 f = 0. \tag{41}$$

The resulting cellular complex has $H_1 = 0$ and $H_2 = 0$. Attaching this face is a deliberate representational operation; it is independent of aromatic electronic structure.

Now consider the separate group-field construction of Eq. (18). Let $G = \mathbb{Z}_2$ and take a real rank-three field $\psi : G^3 \to \mathbb{R}$. The interaction comprises four field factors and six shared group indices. With $\psi(a, b, c) = 1$ throughout G^3, its finite sum contains 2^6 unit contributions, giving

$$V_{\text{tet}}[1] = 64. \tag{42}$$

With $\psi(a, b, c) = (-1)^a$, each contribution contains the sign $(-1)^{a+c+e+f}$, so summing over all six indices yields $V_{\text{tet}}[\psi] = 0$. Both values are preserved under the simultaneous translation action U defined in Section 5.4. They are exact finite calculations of a polynomial interaction, independent of chemical rate parameters.

The chosen odd test field $\psi_1(a, b, c) = (-1)^a$ also illustrates a limitation: its quartic vanishes along the entire ray $t\psi_1$, and the diagonal projector in Eq. (21) removes this mode. A different odd character, χ_{010}, has a nonzero quartic $V_{\text{tet}}[t\chi_{010}] = 64t^4$ in the unprojected theory; the gauge-invariant relative sector of Example 5.2 gives an alternative where the quartic remains nonzero after projection.

For a minimal closed stranded example, consider the cross-vertex Wick contractions that pair four fields at one vertex with four at the other, using propagators that identify corresponding argument slots. These contractions depend on the Gaussian covariance, independently of the test fields. Its incidence skeleton has two interaction vertices u, v, four directed propagator edges $e_i : u \to v$, and six faces f_{ij} labeled by pairs $1 \le i < j \le 4$. Choose the orientations

$$\partial_1 e_i = v - u, \qquad \partial_2 f_{ij} = e_i - e_j. \tag{43}$$

Hence $\partial_1 \partial_2 f_{ij} = 0$. The cellular chain groups have ranks $(\operatorname{rk} C_0, \operatorname{rk} C_1, \operatorname{rk} C_2) = (2, 4, 6)$, with $\operatorname{rk} \partial_1 = 1$ and $\operatorname{rk} \partial_2 = 3$. Consequently, $H_0 \cong \mathbb{Z}$, $H_1 = 0$, and $H_2 \cong \mathbb{Z}^3$ for this chosen cellular structure; the Euler characteristic is $2 - 4 + 6 = 4$.

The six-face complex uses a particular Gaussian Wick pairing. With eight fields in V_{tet}^2, there are $7!! = 105$ total pairings: 24 cross-only pairings, 72 pairings with two cross-propagators and one tadpole at each vertex, and 9 with self-pairings only. The latter nine are disconnected at the interaction-vertex level. At order λ^2, all these terms occur in the unnormalized vacuum series, with connected-graph and vacuum-normalization conventions determining subsequent selections. To expose the corresponding combinatorial choice, label the four factors at each interaction vertex by $1, 2, 3, 4$ and let a permutation $\sigma \in S_4$ pair factor i of one vertex with factor $\sigma(i)$ of the other, preserving each of the three argument-slot positions along the propagator. The resulting number of closed strand faces depends on σ.

Proposition 9.2 (Face Counts for Slot-Preserving Two-Vertex Gluings). For the contraction pattern in Eq. (18), the 24 permutations in S_4 give the face-count distribution

$$\begin{aligned} &\#\{\sigma : F(\sigma) = 6\} = 4, \\ &\#\{\sigma : F(\sigma) = 4\} = 4, \\ &\#\{\sigma : F(\sigma) = 3\} = 16. \end{aligned} \tag{44}$$

The four six-face permutations form the cyclic order-four automorphism subgroup of the chosen tetrahedral vertex under the factor ordering used in Eq. (18). The previously analyzed dipole corresponds to the identity pairing, one of these four cases.

Table 2: Case Study II: explicit complex / field data and generated outputs.

Item	Concrete Choice in the Complex / Field Example
Cycle graph	$V = \{v_1, \dots, v_6\}$ and $E = \{e_1, \dots, e_6\}$ with one six-edge cycle
Attached face	one oriented face f with $\partial_2 f = e_1 + \cdots + e_6$
Field domain	$G = \mathbb{Z}_2$, rank $d = 3$, real field $\psi : G^3 \to \mathbb{R}$
Interaction	one quartic tetrahedral interaction with six shared group indices
Test fields	$\psi_0(a,b,c) = 1$ and $\psi_1(a,b,c) = (-1)^a$ (polynomial checks only)
Generated complex	Gaussian Wick contraction of two vertices; one six-face pairing chosen from 24 cross pairings
Boundary maps	$\partial_1 e_i = v - u$ and $\partial_2 f_{ij} = e_i - e_j$
Verification targets	$V_{\text{tet}}[\psi_0] = 64$, $V_{\text{tet}}[\psi_1] = 0$, Betti numbers $(1, 0, 3)$
Slot-preserving gluings	Among 4! = 24 pairings, face counts $F = 6, 4, 3$ occur 4, 4, 16 times

Table 3: Compact Symbol Glossary for Both Formalisms.

Symbol	Definition or Operational Role
Σ_T, Σ_N	Terminal types and nonterminal types with declared ports
$S, \mathcal{P}, \Gamma$	Start graph, production schemas, and grammar tuple
ChemG	Category of admissible attributed chemical graphs
$\mathcal{X}, C$	Effective chemical states and environmental contexts
G, ρ, Stab_G	Symmetry group, its actions, and state stabilizer
$T_r, [X]_{\cong}$	Structural projection and specified isomorphism class
q, m, χ_q, ν_q	Matched rewrite rule, predicate, and state-lift map
$\mathcal{L}_C, a_{q,m}$	Jump generator and event propensity
$\gamma, B(\gamma)$	Realized history and a declared history descriptor
$K_{C,J}, \mu$	Response kernel and functional interpretation
$V, E, F, \partial_1, \partial_2$	Complex cells and cellular boundary maps
$\psi, K_\Lambda, V_{\alpha,\Lambda}$	Group field, quadratic kernel, and interaction terms
π_C, R_C	Positive complex ensemble and chemical readout (the ensemble is not the Feynman amplitude)
$\mathscr{H}_{\text{gr}}, F$	Grammar-history space and candidate complex encoding

The verification in Appendix L.5 traces each strand cycle under the two interaction-vertex pairings and the slot-preserving propagators. If strand-slot permutations or other propagators are allowed, the admissible gluing family changes, so Eq. (44) is scoped to this declared convention.

How Formalism II Models the Example. The second case deliberately uses two linked constructions. The first is a structural lift of a benzene carbon cycle to a chosen two-complex. The second is a finite group-field interaction whose gluing pattern yields a small labeled relational complex. Table 2 lists the declared data.

How the Generation Proceeds. The first half of the example is a representational enrichment:

$$\begin{aligned} k_{\text{ring}} &\xrightarrow{\text{attach } f} k_{\text{fill}}, \\ k_{\text{ring}} &= (V, E, \varnothing), \\ k_{\text{fill}} &= (V, E, \{f\}), \\ \partial_2 f &= e_1 + \cdots + e_6. \end{aligned} \tag{45}$$

The second half begins with two separate calculations. First, selected test fields ψ_0 and ψ_1 are substituted into V_{tet} as elementary finite-polynomial checks. Separately, a Gaussian covariance and perturbative expansion of $\langle V_{\text{tet}}^2 \rangle_0$ generate two-vertex Wick contractions, including a four-propagator sector. The value of ψ_0 or ψ_1 plays no role in the Wick-contraction graph classification. Choosing one of the four six-face slot-preserving pairings then determines the incidence skeleton of the displayed dipole complex:

$$\begin{aligned} \mathfrak{F} &= (\mathbb{Z}_2, 3, K_0, V_{\text{tet}}), \\ \langle V_{\text{tet}}^2 \rangle_{0,\text{cross}} &\rightsquigarrow k_{\text{dip}}, \\ k_{\text{dip}} &= (\{u, v\}, \{e_i\}_{i=1}^4, \\ &\qquad \{f_{ij}\}_{1 \le i < j \le 4}). \end{aligned} \tag{46}$$

This separates two tasks cleanly: the field interaction generates a weighted combinatorial object, while a later readout R_C would interpret it in chemical terms.

What the Example Generates. The case produces three types of exact output. The quartic evaluations and Gaussian graph contributions are separate calculations: $V_{\text{tet}}[1] = 64$ and a six-face dipole factor $|G|^6 = 64$ coincide numerically for the selected conventions, although their origins differ. First, the benzene cycle and the chosen filled cycle form two different structural representations. Second, the finite field interaction yields the exact amplitudes in Eq. (42). Third, the dipole gluing yields a concrete chain complex with the Betti numbers stated below. Together these objects show how Formalism II moves from declared field data to a generated relational complex and then to topological invariants that an implementation can verify exactly.

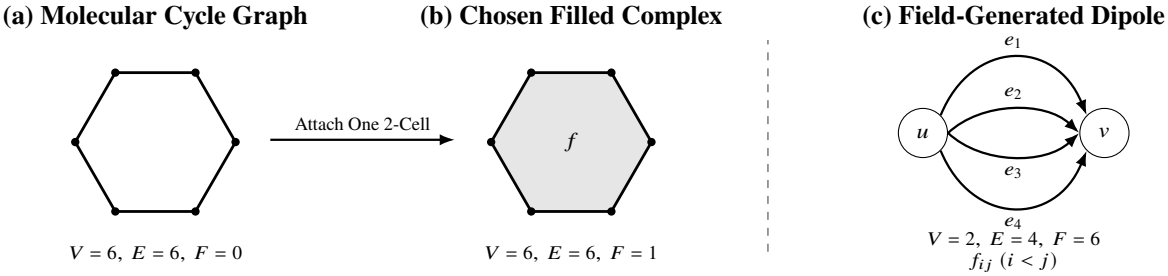
Proposition 9.3 (Boundary Consistency of the Finite Dipole). The two-vertex, four-edge, six-face incidence datum in Eq. (43) defines a valid finite cellular chain complex with $\partial_1 \partial_2 = 0$ and Betti numbers $(b_0, b_1, b_2) = (1, 0, 3)$.

Appendix L gives the integer boundary matrices and an exact rank calculation. The benzene skeleton, its chosen filled cycle, and the finite stranded dipole are three distinct mathematical objects (Fig. 12). The dipole illustrates how a group-field-style interaction can produce a nontrivial relational complex. A mapping R_C from that complex to a chemical state would be an additional, physically testable specification; choosing the mapping by fiat would only illustrate representational flexibility.

9.3 Lessons of the Two Examples

Case Study I makes explicit the terminal / nonterminal distinction, the two production outcomes, the enriched state that carries tetrahedral orientation, and the induced reflection orbit. Case Study II makes explicit the declared cell sets, the boundary operators, the finite field interaction, the gluing that produces the dipole complex, and the resulting topological invariants. Taken together, the examples show how each formalism starts from declared data, applies its own generator, and arrives at exact outputs that can be checked independently. They therefore provide practical regression targets for future software while remaining complementary to physical validation of the three conjectures.

At the level of implementation, the expected outputs are the two tetrahedral orientation classes, the $\mathbb{Z}_2$ orbit, the exact field-interaction values 64 and 0, and the dipole homology ranks. Each has an independent mathematical derivation. Source-code reproduction can therefore report which assertions pass at the formal level, followed by a separate record of any physical assumptions added for chemical interpretation.



The six-carbon chemical cycle, its filled 2-complex, and the two-vertex stranded dipole are three mathematically distinct objects. A chemical readout from the dipole is an additional model choice.

Figure 12: Case Study II: Structural Cycle and Group-Field-Induced Complex. The left and middle panels compare a six-carbon connectivity cycle with a chosen 2-cell attachment. The right panel depicts the incidence skeleton of the two-vertex, four-propagator dipole of a rank-three quartic stranded interaction; each pair $i < j$ of propagator edges bounds a strand face f_{ij}. The chemical meaning of such faces requires a separately specified readout.

10 Relations Between Two Formalisms

The two constructions share a symmetry-first modeling orientation, but they do not thereby share an ontology, a natural notion of time, or a common probability measure. The first begins with chemically interpreted production rules and may attach a stochastic or deterministic law to their realizations; the second begins with group-structured field interactions that generate labeled complexes and acquires a chemical interpretation only through a further map. The relationship between these constructions is therefore itself a mathematical and physical question. The present section states what a useful relationship would preserve, which kinds of correspondence are possible, and how the three conjectures constrain their interpretation.

10.1 Shared Operational Target

The relevant object for scientific comparison is a distribution of outcomes under a declared preparation, environment, and observation protocol. This permits the two models to be evaluated on a shared domain without presupposing that their underlying generative objects are identical. Let $\mathcal{P}$ be a class of specified preparation protocols. For $p \in \mathcal{P}$, Formalism I produces an endpoint law $Q^{\mathrm{gr}}_{p,C} \in \mathcal{P}(\mathcal{X})$ by interpreting the stochastic grammar of Section 4. Formalism II supplies a positive ensemble $\pi_{p,C}$ on a suitable complex space $\mathscr{K}$ and a readout $R_{p,C}(dx \mid k)$, yielding

$$Q^{\mathrm{cx}}_{p,C}(dx) = \int_{\mathscr{K}} R_{p,C}(dx \mid k)\, \pi_{p,C}(dk). \tag{47}$$

The p index is significant. A preparation must be associated with physical operations, initial conditions, or independently characterized constraints; freely changing model parameters after observing an outcome would make agreement vacuous. Similarly, an interaction amplitude arising from a group-field-inspired action cannot be substituted for the positive normalized $\pi_{p,C}$ without a mathematically valid probability construction.

For an evaluation protocol J, both endpoint laws enter the same response kernel $K_{C,J}(dy \mid x)$ introduced in Section 6. The corresponding observable measures are

$$P^{i}_{p,C,J} = Q^{i}_{p,C} K_{C,J}, \qquad i \in \{\mathrm{gr}, \mathrm{cx}\}. \tag{48}$$

This formulation is intentionally asymmetric in epistemic status: Formalism I may be informed by measured reaction rates, whereas Formalism II may only generate candidate relational organizations. A good fit of both to the same terminal structural dataset would not establish a common physical mechanism. Agreement should additionally be checked on interventions, paths, response times, and physically meaningful retained variables.

Definition 10.1 (Operational Correspondence). For a declared class $\mathcal{P}_0 \subseteq \mathcal{P}$, contexts C_0, and test protocols $\mathcal{J}_0$, the two formalisms are *ε-operationally corresponding* if

$$\sup_{(p,C,J)\in\mathcal{D}_0} d_{\mathrm{TV}}(P^{\mathrm{gr}}_{p,C,J}, P^{\mathrm{cx}}_{p,C,J}) \le \varepsilon, \tag{49}$$

where $\mathcal{D}_0 \subseteq \mathcal{P}_0 \times C_0 \times \mathcal{J}_0$ is explicitly specified and d_{TV} is the total-variation distance between probability laws. A correspondence at the trajectory level requires a separate path-space comparison.

This is weaker than the structural faithfulness of Definition 5.2. Two models may have matching measured outputs yet differ radically in their generated complexes and rewriting histories. That weaker correspondence is nevertheless useful for systematic model comparison, especially when the observation domain is limited and the claims concern relative empirical performance rather than physical identity.

10.2 Relations Between the Two Representations

Let $\mathscr{H}_{\mathrm{gr}}$ be a measurable space of finite grammatical histories, and let $\kappa_{p,C}$ be a probability law on it. A proposed mathematical bridge from grammar to complex representation may be deterministic, $F : \mathscr{H}_{\mathrm{gr}} \to \mathscr{K}$, or probabilistic, $F(dk \mid h)$. The latter accommodates ambiguity in the representation of branching, simultaneity, or spatial adjacency by a two-complex. Let $E_C(dx \mid h)$ describe the chemical endpoint law of a grammatical history. A strong correspondence asks that

$$E_C(dx \mid h) = \int_{\mathscr{K}} R_C(dx \mid k) F(dk \mid h), \tag{50}$$

for the histories in the agreed domain, together with consistency of the relevant group actions and composition operations. Here group equivariance is a compatibility requirement, not a substitute for the chemically interpreted kernel E_C.

Proposition 10.1 (Observable Contraction Of Bridge Error). Fix a preparation and context. Let Q^{gr} and Q^{cx} be probability measures on the same chemical endpoint space, and let K be a common Markov response kernel. Then

$$d_{\mathrm{TV}}(Q^{\mathrm{gr}}K, Q^{\mathrm{cx}}K) \le d_{\mathrm{TV}}(Q^{\mathrm{gr}}, Q^{\mathrm{cx}}). \tag{51}$$

If both endpoint laws arise as mixtures of two kernels whose conditional total-variation distance is at most ε on a common history distribution, their response laws differ by at most ε.

A proof appears in Appendix G. The result is a data-processing inequality and does not establish either chemical accuracy or equality of microscopic mechanisms. Its utility is methodological: one can distinguish errors in translating a generated relational complex into a chemical state from errors introduced by a subsequent observation protocol. When two state kernels differ appreciably but a particular observation hides the difference, functional agreement alone offers weak evidence for a proposed bridge.

A more demanding bridge would preserve information about a selected symmetry-breaking descriptor B, perhaps by placing B in a common marked path space. It would also need to distinguish a two-cell used as a structural ring from a two-cell encoding a composition of events. Neither the boundary identity $\partial_1\partial_2 = 0$ nor the existence of a group-field perturbation series supplies a canonical mapping between these interpretations. Work on chemical graph rewriting and its categorical formulations provides rigorous precedents for tracking atom maps and reaction contexts, while recent higher-order molecular representations highlight why multi-constituent relations merit explicit treatment (Andersen et al., 2013; Kajino, 2019; Huang et al., 2026; Smith et al., 2024; Restrepo, 2026; Gale et al., 2025). Their existence is a reason to specify the bridge rather than assert its novelty by naming it.

10.3 Physical Content of Symmetry

In both formalisms, a group is selected together with a domain of action. This choice carries substantive physical commitments. Permutations of indistinguishable atomic labels, rigid rotations, inversion, crystal space groups, and auxiliary group variables in an effective field construction express different invariance claims. A mathematical action of G on a state space may exist even where a physical preparation breaks it through boundary conditions, a solvent, a substrate, or applied fields. One should therefore identify the symmetry of the equations, the symmetry of the admissible ensemble, and the stabilizer of a realized configuration separately.

A second distinction concerns the role of symmetry in enabling a function. In the first formalism, a residual subgroup H may label a selected branch of an effective chemical dynamics. In the second, H may label a minimum of a group-field-inspired potential. The latter need not refer to the same observable order parameter as the former. Conjecture 7.1 becomes supported across the formalisms under the condition that a validated realization map identifies the selected states with a chemically measured organization and a protocol J detects a genuinely new attainable response. A constant readout supplies an immediate counterexample to any claim that a field-theoretic broken-symmetry sector necessarily produces a distinguishable chemical material (Appendix C).

Likewise, even a physically established symmetry reduction cannot be equated with arbitrary loss of conservation laws or with spontaneous breaking of a gauge redundancy. Noether-type conservation statements, spontaneous breaking of physical global symmetries, and gauge constraints involve different mathematical conditions. The finite-group constructions of Sections 4–5 are valuable controlled examples, but their G-actions must be justified anew when transferred to molecular rotation, chirality, crystalline order, or electronic structure. The theory's chemical content rests on that justification and on measured or derived kinetics, thermodynamics, and spectroscopy, in combination with group-theoretic structure. A growing body of chemical theory investigates quantum-field approaches to molecular interactions; those results should not be conflated with a demonstrated group-field theory of chemical bonds (Karimpour et al., 2026).

The direct connection between the two mathematical viewpoints is shown in Fig. 13. The left-to-right steps in Formalism I explicitly depict chemical states and rewrites; Formalism II first produces a labeled complex and only then seeks a chemically calibrated readout. This difference matters when evaluating what each model generates, what is observable, and which claims remain untested.

10.4 Concrete Chemical Structures for Representation Checks

The proposed formalisms should discriminate chemically distinct structures before claiming more ambitious generative explanations. Figure 14 gives two minimal stereochemical tests. The (Z) and (E) configurations of 1,2-dichloroethene share a constitutionally labeled graph, yet differ in their alkene configuration and in symmetry-relevant electric responses. Likewise, the two enantiomeric configurations of bromochlorofluoromethane share a constitutional graph but can be distinguished in a specified chiral context. These examples demonstrate *representation insufficiency*, not Conjecture 6.1: their enriched terminal physical states already differ. Full definitions, caveats, and sources are given in Appendix I.

A contrasting example concerns the second formalism. The carbon ring skeleton of benzene is a six-edge cycle; representing it as a two-complex by attaching a face changes the homology of the *chosen representation*. That operation does not by itself encode aromaticity, an electronic mechanism, or a GFT interaction history. Appendix I.5 makes this contrast explicit with a figure and a boundary calculation. Finally, proton transfer between acetic acid and water illustrates how a group-equivariant grammar must still impose atom and charge accounting and obtain rates from physical chemical conditions (Appendix I.3). A further abstract Frank-type autocatalytic network illustrates a symmetric set of chemical production rules capable of amplifying a small chiral imbalance under explicit kinetic assumptions (Appendix I.4). Together these small checks expose what the formalisms can represent, what they omit, and which additional physical mappings are required. The key definitions and operational comparison equations are summarized for reference in Appendix H.

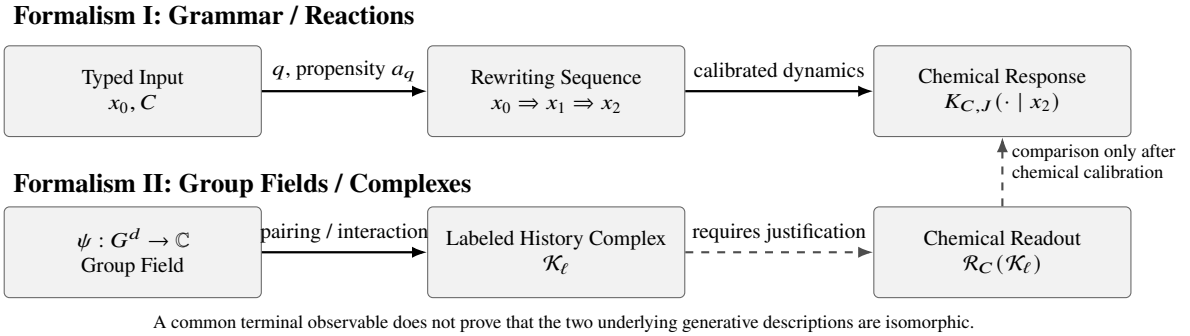


Figure 13: What Connects Two Generative Formalisms. The first starts from chemically interpreted states and admissible graph transformations; the second generates labeled relational complexes from field interactions. A chemically defensible readout is required before comparing predictive response laws. A shared observable does not entail equality of histories, generators, or ontological substrates.

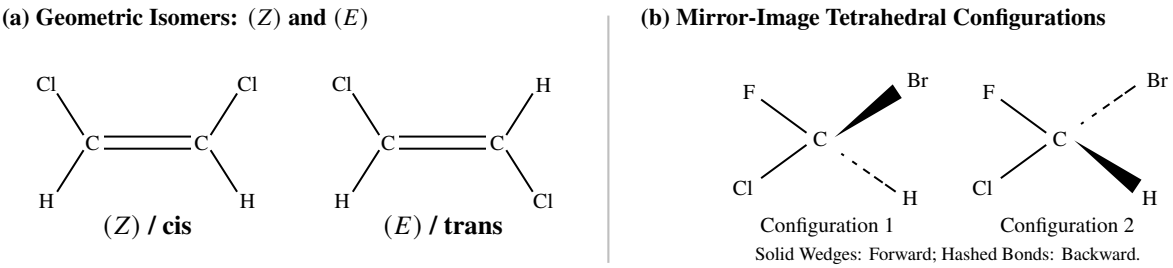


Figure 14: Two chemical tests of representation resolution. (a) (Z)- and (E)-1,2-dichloroethene have the same element- and bond-labeled constitutional graph but different spatial configurations; this can change a vector response such as the electric dipole. (b) The two configurations of bromochlorofluoromethane, CHBrClF, are mirror-related while retaining the same constitutional graph. The diagrams illustrate that graph isomorphism at one resolution does not determine all stereochemical properties. Neither pair, by itself, demonstrates history-dependent functionality or spontaneous symmetry breaking. Drawings are schematic.

10.5 History, Semantic Function, And Counterfactual Tests

The relationship between the first two conjectures is asymmetric. Conjecture 6.1 asks whether two preparations yielding the same *declared* structural equivalence class can have distinguishable responses. Conjecture 7.1 asks whether a symmetry-selecting transition can enlarge a separately defined set of attainable functional responses. In a system with both effects, a symmetry-breaking path can leave retained material organization that makes a coarse terminal graph insufficient for function; the new function can then be characterized as a causal response to an intervention. Yet either effect can occur without the other. A history effect may follow from kinetic trapping without symmetry breaking, while a symmetry-broken branch may have a function indistinguishable from its symmetric reference under the chosen protocol.

The term *functional semantics* accordingly refers to a relational and operational assignment, as a specified relational and operational assignment of chemical roles. At a minimum, specifying J requires an observation timescale, permissible interventions, a system boundary, and a target response. Different choices can produce different functional equivalence classes on the same underlying states. In biological contexts one might additionally investigate historical or evolutionary accounts of function; these require assumptions beyond a chemical response kernel. Research on molecular codes and semantic information offers precedents for precisely stated notions of chemical function, although the definitions and objectives are distinct (Görlich and Dittrich, 2013; Kolchinsky and Wolpert, 2018).

A disciplined test of Conjecture 6.1 would therefore compare progressively enriched descriptions $T_{r_1}, T_{r_2}, \ldots$, recording connectivity, spatial order, defects, and other retained physical variables as independently justified. The task is to identify a representation at which the response kernel becomes approximately constant on its structural fibers under a fixed protocol, rather than to declare all variation an irreducible historical effect. In the absence of a sufficiently measured state, a functional difference between preparations might reflect an uncontrolled covariate rather than a symmetry-breaking mechanism. Experimental work on nonequilibrium supramolecular organization and pathway control provides motivation for such a strategy (Sorrenti et al., 2017; Wehner and Würthner, 2020).

10.6 Tests of the Proposed Scope

The present framework provides three distinct routes for assessment. The first is a *mathematical audit*: test whether the chosen graph category, incidence structure, group action, and stochastic kernels are jointly well defined; establish when a translation preserves admissibility and equivariance; and give counterexamples when information is lost. The second is a *chemical audit*: on benign, documented model systems, compare independently calibrated structure, order parameters, kinetic paths, and outcomes under controlled environments. This audit should use negative cases where symmetry changes but functional response does not, as well as positive cases proposed by the conjectures. The third is an *ontological audit*: construct rival explanatory

models with explicit commitments and a declared class of observations; distinguish restricted observational agreement from an ultimate impossibility of empirical discrimination.

In chemical applications, a useful sequence begins with public datasets and safely characterized model networks. One can compare a discrete grammar-based prediction with a continuum order-parameter model and then investigate whether a complex-derived history descriptor improves out-of-sample predictions. The inclusion of a complex representation is warranted under the condition that it contributes something beyond a conventional reaction hypergraph or state-space model: for example, a measurable invariant, a constrained composition structure, or a lower-complexity description of observed multibody interactions. Any proposed advantage calls for comparison with established alternatives.

Two broad failure modes would substantially limit the proposal. First, the group-theoretic constraints might be so permissive that they exclude few physically impossible structures; then symmetry-first modeling offers little explanatory compression. Second, a chosen two-complex may possess numerous inequivalent chemical interpretations, so that a field-theoretic ensemble cannot meaningfully predict chemical measurements without an independently fitted, high-capacity readout. In both cases a mathematically consistent generator may remain an uninformative chemical theory. The mathematical and empirical tests described above are designed to expose these possibilities.

10.7 Interpretation of Generative Order

Conjecture 8.1 places a limit on what can be inferred even from success in all preceding tests. A family of structural grammars can be generated by a meta-grammar, and a field interaction may generate complexes with the desired combinatorial properties. That establishes mathematical and perhaps explanatory adequacy relative to the specified evidence. It does not by itself establish the existence of an independent physical meta-grammar, much less a unique ultimate origin of structure. The contrastive-model method of Section 3 explicitly seeks the premises that would make a stronger conclusion justified. A competing model may share the successful predictions while interpreting the common constraints as effective regularities of interactions rather than as the activity of a distinct higher-order generator.

This distinction connects chemical generativity to broader questions about genomic organization and linguistic universal grammar without identifying these domains. The mapping from genetic sequences to regulatory organization and molecular products involves particular biochemical mechanisms and developmental contexts. Human language concerns a different class of cognitive, social, and biological mechanisms. The mathematical construction of grammars for all three does not entail a common causal grammar or a universal origin. Any cross-domain correspondence must preserve precisely specified structures and explanatory conditions and must withstand evidence particular to each domain. For linguistic universals, an account of shared formal regularities is still distinct from a hypothesis of an independently realized biological Universal Grammar (Chomsky, 2005; Stanford, 2023).

The resulting research stance is deliberately non-final. Symmetry is a *relatively primitive modeling ingredient* in both proposed formalisms; the formalisms can be assessed by derivations, predictions, and counterexamples. The question whether that primitiveness reflects a privileged ontological level remains open. This separation permits productive mathematical commitments while avoiding an inference from representational order to ontological necessity.

11 Scope and Limitations

The two proposed formalisms define families of generative descriptions, with distinct mathematical objects and distinct routes to chemical interpretation. Their examples establish consistency for selected finite constructions. A statement covering *chemical structures in general* requires a specified domain of systems, a resolution of structural description, and an observable whose physical behavior the model aims to explain. It is useful to distinguish three levels of scope: (i) *representability*, the capacity to encode an object; (ii) *generative reachability*, the capacity to produce it through admissible transformations or interaction patterns; and (iii) *physical adequacy*, the capacity to reproduce independently measured states, paths, or responses under declared chemical conditions. The second property imposes constraints beyond encoding, while the third introduces experimentally accountable physics.

11.1 Representational Scope Across Chemical Domains

For ordinary finite covalent molecules, typed atom-and-bond graphs provide an effective structural language, and enriched attributes can represent stereochemistry. The chemically relevant object can require additional information: delocalized electrons, multicentre bonding, charge and spin states, solvent organization, conformational ensembles, or an evolving reaction environment. Aromatic, organometallic, and multicentre-bonded systems can benefit from hypergraph or higher-order relational descriptions (Restrepo, 2026; Kajino, 2019). In materials science, periodicity, defects, phase boundaries, extended electronic states, and collective order often become central; their treatment requires an explicit choice of system boundary and resolution (Reiser et al., 2022). The graph-rewriting formalism can be extended with attributed, spatial, periodic, or higher-order structures, provided the associated category, matching conditions, and rules are reformulated consistently.

A relational two-complex can record higher-order incidence or history organization, yet a chosen face has a meaning only after an interpretation is specified. For example, filling the six-carbon cycle in Case Study II changes the homology of the representation, while a chemically informative account of benzene additionally addresses electronic delocalization. A group-field-generated complex and a molecular motif complex may therefore require separate readout maps. The two constructions remain distinct even when both use vertices, edges, and faces.

11.2 Generative Reachability and Restricted Universality

For a chemically specified domain $\mathcal{D}_C \subseteq \mathcal{X}$ at structural resolution r, write its representable target set as

$$\mathcal{T}_{C,r}(\mathcal{D}) = \{[T_r(x)]_{\cong} : x \in \mathcal{D}_C\}. \tag{52}$$

Table 4: Representative Chemical Domains and Additional Modeling Requirements. Entries describe possible extensions, with physical adequacy to be established for each application.

Chemical Domain	Relevant Enrichment or Validation
Finite covalent molecules	Typed connectivity, stereochemical state, chemically admissible rules
Aromatic and multicentre systems	Electronic structure, delocalization, hyperedges or justified higher-order relations
Coordination and organometallic systems	Metal–ligand interactions, variable coordination, spin and electron accounting
Macromolecules and assemblies	Conformations, solvent, mesoscale organization, kinetics and history
Crystals and extended materials	Periodic or infinite-limit representations, defects, collective order, electronic bands
Reactive mixtures and excited states	Time-dependent populations, open-system balance, electronic-state dynamics and measured rates

Formalism I has *terminal coverage* on this declared domain when

$$\mathcal{T}_{C,r}(\mathcal{D}) \subseteq L_{\mathrm{term}}(\Gamma; C). \tag{53}$$

Coverage is a property of the selected grammar, context, and resolution. In particular, a finite alphabet and a fixed production family delimit the reachable terminal language. Increasing the family of rules can increase coverage, but an unrestricted rule that simply inserts a target graph would provide little insight into chemical generation. Meaningful universality therefore calls for a compact, independently justified rule family, together with nontrivial predictions or exclusions.

Formalism II similarly generates a particular class of labeled complexes selected by its field rank, admissible gluing patterns, kinetic kernel, and interaction terms. A specified group-field action may sample only part of the broader space of mathematical two-complexes. The chemical output law also depends on the readout R_C. For a finite discrete ensemble, the support satisfies

$$\operatorname{supp} Q_C^{\mathrm{cx}} = \bigcup_{k:\, \pi_C(k)>0} \operatorname{supp} R_C(\cdot \mid k). \tag{54}$$

Consequently, claims about chemical coverage require both the generated complex class and the readout family. The finite $\mathbb{Z}_2$ construction in Case Study II verifies one interaction and one incidence type; it supplies a controlled test fixture for the chosen model.

Proposition 11.1 (Unrestricted Readout and Output Nonidentifiability). Let π be a normalized probability distribution on a finite nonempty complex space and let P^* be an arbitrary target probability law on chemical states. The constant kernel $R(dx \mid k) = P^*(dx)$ yields $\sum_k \pi(k) R(dx \mid k) = P^*(dx)$ independently of the complex generator. Thus arbitrary readout flexibility is sufficient for output matching while leaving the generating complex empirically unidentified by those outputs.

The identity follows immediately from normalization of π. It illustrates why chemical readouts need externally motivated constraints, explicit parameter counts, and tests beyond the data used for calibration. A correspondence between the two formalisms acquires explanatory content when it preserves a specified property under changes of preparation or intervention and improves prediction relative to established alternatives.

11.3 Symmetry Assumptions and Physical Mechanisms

Both formalisms introduce a group action as a relatively primitive *modeling choice*. Symmetry can constrain the laws of generation while individual structures and trajectories have smaller stabilizers. The relevant group depends on the state variables, environment, and proposed transformation. For example, a fixed chiral solvent or external field can reduce the action compatible with a given experimental context. A trivial group is mathematically available for many models, but it contributes no substantive symmetry-selection prediction. Accordingly, the usefulness of a symmetry-first account lies in a nontrivial and chemically defensible action that restricts allowed operations, laws, or outcomes.

Structural generation can proceed through symmetry-preserving transformations as well as symmetry-selecting transitions. The phrase *through symmetry breaking* in the title designates the mechanism investigated in this paper, rather than a proposed universal requirement for chemical formation. Moreover, symmetry selection alone specifies neither the energetic accessibility of a process nor the acquisition of a functional role. Physically interpretable conclusions may require electronic Hamiltonians, potential-energy surfaces, conservation laws, transition rates, environmental coupling, and an operational response protocol. In particular, a group-field-inspired amplitude is a mathematical weight on relational complexes; identifying it with a chemical pathway probability requires a positive normalized construction and a chemically calibrated readout. The use of group field theory here is an inspiration for combinatorial generation, with its originating physical setting and interpretation kept distinct (Oriti, 2014; Oriti et al., 2014).

A more specific mathematical restriction arises for the finite tetrahedral interaction: its translation symmetry includes $G \times G$, the real polynomial admits an orthogonal-index extension, and a diagonal gauge projection can remove proposed odd field variables. The field sector used for symmetry selection should therefore be chosen after the kinetic and gauge constraints have been declared. Similarly, the number of faces in a two-vertex complex varies with the propagator pairing. These distinctions give useful tests of whether a claimed structure is generated by the stated model or by an added representational choice (Section 5.3, Proposition 9.2).

11.4 Verification Conditions and Present Research Boundary

A broad applicability study can report three nested tests for each chemical domain. A *representation test* checks whether chemically distinct states of interest retain the differences required by the chosen observations. A *generation test* enumerates or bounds the outputs of the stated rules or field interactions, including counterexamples and unreachable classes. A *physical test* compares independently calibrated energies, trajectories, symmetry-sensitive observables, and functional responses with measurements or trusted reference calculations. Performance should be compared with conventional graph or reaction-network models, hypergraph representations, and relevant molecular or materials theory.

The finite examples in Section 9 address the first two tests for narrow mathematical domains. Their successful reproduction would confirm specified algebraic claims and

implementation consistency. General chemical applicability, reliable generation of a broad range of experimentally realizable structures, and prediction of chemical function remain open tasks. This restricted scope also preserves the ontological distinction developed in Conjecture III: success within a large representational domain would support the usefulness of the proposed models while leaving the independent physical existence of a universal meta-generative structure open to further argument.

12 Conclusion

This inquiry has developed two complementary ways of representing chemical structure generation through symmetry-constrained processes. The group-equivariant grammar begins with contextual structural transformations, separates syntactic admissibility from physical realizability, and connects chemically interpreted events to stochastic dynamics. The group-field-inspired formalism begins instead with group-dependent interactions capable of generating relational two-complexes, then introduces a distinct and explicitly provisional map from such complexes to chemical states. Their connection requires preserved symmetries, compatible history information, and a sufficiently justified chemical readout. An operational correspondence can be formulated without claiming equivalence of the underlying generative mechanisms.

The resulting framework articulates three conjectures. The first proposes that symmetry-breaking histories can leave physical organization omitted by a coarse terminal structural representation and thereby differentiate measured functional responses. The second investigates conditions under which symmetry selection is accompanied by new functional possibilities, while allowing broken symmetry to remain functionally silent in some contexts. The third concerns the possible underdetermination of ontologically independent meta-generative order by structural regularities and generative descriptions. Conditional mathematical propositions and explicit countermodels specify parts of these questions, but they do not constitute empirical confirmation of the conjectures.

The analysis of scope and limitations further distinguishes representational coverage, reachable outputs, and physically adequate explanations across chemical domains. The central methodological distinction is among a formal language of chemical possibility, a physically calibrated account of its realization, and a relationally specified account of function. A symmetry-first formalism can be rigorous at the first level and productive at the second while remaining silent about the absolute origin of structure. Future work should therefore prioritize a delimited, well-characterized chemical system; explicit group actions and measured order parameters; path-resolved or state-resolved observations; and controlled tests of whether a complex-based representation adds predictive information beyond established graph and reaction-network descriptions. These steps would allow the mathematical proposals to acquire a concrete chemical interpretation while preserving the openness of their wider ontological questions.

A Contrastive Non-Entailment: A Formal Note

This appendix supplies the complete model-theoretic argument invoked by Proposition 3.1. It also specifies why the result does not, by itself, establish empirical equivalence.

Proof of Proposition 3.1. Let $\mathcal{K}$ be the admitted class of interpretations for the signature in which $\mathbb{T}$ and H_M are formulated. By the standard definition of semantic entailment,

$$\mathbb{T} \models_{\mathcal{K}} H_M \quad \Longleftrightarrow \quad \forall \mathfrak{M} \in \mathcal{K},\ \mathfrak{M} \models \mathbb{T} \ \Rightarrow\ \mathfrak{M} \models H_M.$$

By hypothesis there is $\mathfrak{M}_- \in \mathcal{K}$ satisfying $\mathbb{T} \wedge \neg H_M$. It is therefore a counterexample to the universal condition on the right, and $\mathbb{T} \not\models_{\mathcal{K}} H_M$. Similarly, $\mathfrak{M}_+ \in \mathcal{K}$ satisfies $\mathbb{T} \wedge H_M$, so it is a counterexample to $\mathbb{T} \models_{\mathcal{K}} \neg H_M$. Both non-entailments follow. □

Remark A.1 (The Role Of The Interpretation Class). The conclusion is relative to $\mathcal{K}$. If additional background assumptions rule out one of the models, the non-entailment may cease to hold. A proposed two-model argument must therefore declare the allowable mathematical and physical interpretations rather than suppressing them in an unrestricted quantifier.

Remark A.2 (From Logical To Empirical Comparison). Truth of the same finite set of observational sentences is weaker than equality of prediction laws under a declared protocol class. Even equality of all predictions in a stipulated class $\mathcal{P}$ does not establish equality across every conceivable experiment or inferential standard. A scientifically informative application therefore records the observational domain, its limits, and the independent adequacy of the competing interpretations.

B Proofs For The Group-Equivariant Grammar

B.1 Derivations Under A Group Action

Proof of Proposition 4.1. Consider an admissible step with rule span $L_q \leftarrow K_q \rightarrow R_q$ and injective match $m : L_q \hookrightarrow T_r(x)$ in the category of partially labeled interfaces. Assume the group action preserves both immutable typing and the partial-label compatibility conditions. By hypothesis, the selected graph-category equivalence associated with g carries the rule span and match to $gL_q \leftarrow gK_q \rightarrow gR_q$ and $gm : gL_q \hookrightarrow gT_r(x)$. An equivalence preserves the universal property of a pushout, and hence takes each pushout in an admissible double-pushout construction to a pushout up to canonical isomorphism. Its inverse carries the transformed construction back to the original one; therefore the existence of the required pushout complement and the matching admissibility conditions is preserved in both directions. The hypothesis on χ ensures that transformed contextual guards are satisfied exactly when the original guards are, and covariance of ν_q supplies the transformed state lift. Thus $gx \xRightarrow{gq,gm;gC} gx'$. Conversely, the result with g^{-1} transports a transformed transition back to an original one. For a finite derivation, apply this correspondence to its first step; the target is the source of the transformed second step. Induction

on the number of steps proves the claim. This construction establishes an equivariant correspondence between derivations; it does not endow an irreversible rewriting step with an inverse. □

Remark B.1 (Limits Of The Categorical Hypotheses). An arbitrary transformation of drawn graphs need not preserve the relevant category or the chemical attributes. In particular, identifying chirality-related spatial configurations or changing atom types while claiming exact chemical equivalence can invalidate the assumptions. The proposition concerns a specified functorial symmetry of a formal rewriting system. It does not infer such a symmetry from a visually symmetric diagram.

B.2 Stochastic Generator And Induced Semigroup

Proof of Proposition 4.2. Write $\mathcal{I}(x)$ for the set of admissible rule–match occurrences (q, m) from x. The group action and matching hypotheses establish a bijection $\mathcal{I}(x) \to \mathcal{I}(gx)$, $(q, m) \mapsto (gq, gm)$. Applying the definition of the transformed generator and reindexing the sum gives

$$\begin{aligned}(\mathcal{L}_{gC} f)(gx) &= \sum_{(q',m')\in\mathcal{I}(gx)} a_{q',m'}(gx, gC) \\ &\quad \times \big[f(J_{q',m'}gx) - f(gx)\big] \\ &= \sum_{(q,m)\in\mathcal{I}(x)} a_{q,m}(x, C) \\ &\quad \times \big[f(gJ_{q,m}x) - f(gx)\big] \\ &= (\mathcal{L}_C(f \circ g))(x).\end{aligned}$$

Convergence in Assumption 4.2 legitimizes reindexing. Suppose now that $gC = C$, and define $(U_g f)(x) = f(gx)$. Then $\mathcal{L}_C U_g = U_g \mathcal{L}_C$. For finite-state chains, the matrix exponential therefore obeys $P_t U_g = U_g P_t$. For general well-posed Markov processes, the same conclusion follows from uniqueness of the semigroup solving the corresponding evolution equation on the specified function space. This second conclusion is explicitly conditional on existence and uniqueness; a formal generator identity alone would be insufficient. □

Remark B.2 (State Counting And Symmetry Factors). The bijection in the proof operates on physically distinguished event occurrences. If one counts equivalent matches several times, a corresponding normalization or orbit–stabilizer factor must enter the propensities. Alternatively, the event set may be quotiented by automorphisms with correct effective rates. Both conventions can produce the same generator when their multiplicities are handled consistently; mixing the conventions generally produces incorrect kinetics.

B.3 Stabilizers And Exact Symmetry

Proof of Lemma 4.1. Take $g \in H_{x_0}$, so $gx_0 = x_0$, and recall $gC = C$. By equivariance and uniqueness,

$$g\Phi^t_C(x_0) = \Phi^t_{gC}(gx_0) = \Phi^t_C(x_0).$$

Thus g belongs to the stabilizer of $\Phi^t_C(x_0)$ for every time at which this flow is defined, and $H_{x_0} \subseteq H_{\Phi^t_C(x_0)}$. In particular it cannot be a strict supergroup of the later stabilizer. If Φ^t_C is invertible, its inverse is equivariant on the relevant image. For $h \in H_{\Phi^t_C(x_0)}$, applying the inverse to $h\Phi^t_C(x_0) = \Phi^t_C(x_0)$ gives $hx_0 = x_0$. Hence $H_{\Phi^t_C(x_0)} \subseteq H_{x_0}$, proving equality for such invertible flows. This establishes both assertions. □

Remark B.3 (Symmetry Of A Distribution Versus Symmetry Of A Sample). For a well-defined equivariant stochastic process with symmetric context and G-invariant initial law π_0, the time-t law is G-invariant, provided the Markov evolution is unique. This does not imply that a single sample state x_t has G as its stabilizer. In finite systems, fluctuating realizations can be asymmetric even when the ensemble has exact symmetry. An order parameter that acquires a selected sign may require an explicit seed, conditioning, or an appropriate limiting construction for a persistent broken-symmetry phase.

C Relational-Complex Constructions And Proofs

C.1 Cellular Incidence And Structural Labels

Let a finite oriented 2-complex have cellular basis elements v_i, e_j, and f_k. The boundary operators can be represented by integer matrices D_1 and D_2. Their cellular property is $D_1 D_2 = 0$. To verify this explicitly for a 2-cell whose boundary is an oriented cycle $v_1 \to v_2 \to \cdots \to v_n \to v_1$, write $\partial_2 f = e_{12} + \cdots + e_{n1}$. Then

$$\partial_1 \partial_2 f = (v_2 - v_1) + \cdots + (v_1 - v_n) = 0.$$

Linearity extends the identity to every cellular 2-chain. More general CW attaching maps satisfy the same chain-complex identity by the cellular boundary construction, even when incidences repeat. The identity constrains incidence information but is insufficient to infer ring strain, molecular orbital structure, or a reaction barrier.

C.2 Finite-Group Interaction Invariance

Proposition C.1 (Symmetry Of The Tetrahedral Interaction). For finite G and real $\psi : G^3 \to \mathbb{R}$, the quartic polynomial in Eq. (18) is invariant under the simultaneous left-translation action $(U_g\psi)(a, b, c) = \psi(g^{-1}a, g^{-1}b, g^{-1}c)$.

Proof. Substitute $U_g\psi$ for each occurrence of ψ in Eq. (18). Every summation index a, b, c, d, e, f is replaced by $g^{-1}a, g^{-1}b, g^{-1}c, g^{-1}d, g^{-1}e, g^{-1}f$, respectively. Left multiplication by g^{-1} is a bijection of the finite group. Reindexing all six summations therefore reproduces exactly the original polynomial. The result is independent of commutativity of G; it uses only bijectivity and the shared appearance pattern of the indices. □

Remark C.1 (Distinguishing Invariant Action From Broken State). The invariance just proved is a symmetry of an interaction functional. It does not establish that the model has a non-invariant equilibrium or a symmetry-breaking transition. A finite-dimensional measure with a positive invariant action is itself invariant; a selected broken-symmetry state requires an additional specification, such as a biased boundary condition or a limiting regime. Moreover, whether

this group operation corresponds to a chemically observable mirror or rotational transformation requires a separate interpretation.

C.3 Selection Of A Broken-Symmetry Field Sector

Proof of Proposition 5.2. For $W(t) = \alpha t^2 + \beta t^4$ with $\beta > 0$, $W'(t) = 2t(\alpha + 2\beta t^2)$ and $W(t) \to +\infty$ at large $|t|$. For $\alpha \geq 0$, $t = 0$ is the unique minimizer. For $\alpha < 0$, the critical points $t_\pm = \pm\sqrt{-\alpha/(2\beta)}$ satisfy $W''(t_\pm) = -4\alpha > 0$ and $W(t_\pm) = -\alpha^2/(4\beta) < W(0)$. This proves the conditional scalar result. The subgroup stabilizer of a general field is determined by its full odd component ψ_{odd} from Eq. (22): the nontrivial group element fixes ψ precisely when $\psi_{\text{odd}} = 0$. A single Fourier coefficient such as η_b diagnoses only its selected sector. □

Remark C.2 (The Readout Can Erase Broken Symmetry). Let $R_C(\cdot \mid k)$ be the point mass at a single, G-fixed chemical state x_* for every complex k. The readout is equivariant, yet every input distribution and every group-related pair of field realizations yield the same terminal structure and output. Therefore, group-field symmetry breaking alone does not entail new chemical functions or even distinct chemical structures. Such an implication would require additional properties of the interpretation kernel, potentially including injectivity on a relevant symmetry orbit and a functionally discriminating observation protocol.

C.4 Extended Symmetry, Gauge Projection, and Relative Odd Modes

Proof of Proposition 5.1. In Eq. (18), the indices a, c, e, f each occur once in the first slot and once in the third slot of ψ, while b, d each occur twice in the second slot. Under Eq. (20), reindex a, c, e, f by g^{-1} and b, d by h^{-1}. Because each map is a bijection of G, the sum is unchanged. For the orthogonal extension, transform both the first and third slots by Q and the second slot by R. Contraction of the two appearances of each index uses $\sum_x Q_{xi}Q_{xj} = \delta_{ij}$ or $\sum_x R_{xi}R_{xj} = \delta_{ij}$. Every transformed contracted pair reduces to a Kronecker delta, recovering the original polynomial. Hence the quartic is invariant under $O(n) \times O(n)$. The extension concerns the unconstrained real tensor space; gauge, kinetic, and regulator restrictions may reduce the full theory's symmetry. □

For $G = \mathbb{Z}_2$ define characters $\chi_{ijk}(a, b, c) = (-1)^{ia+jb+kc}$. The quartic on a pure character satisfies

$$V_{\text{tet}}[t\chi_{ijk}] = \begin{cases} 64t^4, & i = k, \\ 0, & i \neq k. \end{cases} \tag{55}$$

Indeed the four factors contribute exponent $(i + k)(a + c + e + f) + 2j(b + d)$; the finite sum vanishes unless $i = k$. The old first-slot character χ_{100} has zero quartic, whereas the second-slot character χ_{010} has $V_{\text{tet}}[t\chi_{010}] = 64t^4$. The Gaussian quadratic term $(\mu/2)\sum\psi^2$ evaluates to $4\mu t^2$ on this character, proving Eq. (24) directly.

The simultaneous translation has eigenvalue $(-1)^{i+j+k}$ on each character. Thus the odd eigenspace has basis $\{\chi_{100}, \chi_{010}, \chi_{001}, \chi_{111}\}$, of dimension four, and any scalar character coefficient is only a projection onto one direction. For the diagonal gauge projector $P = (I + U)/2$, every odd character vanishes under projection, in particular $P\chi_{010} = 0$.

The projected relative field in Eq. (25) satisfies $P\psi_z = \psi_z$ because each character contains an even number of nonzero indices. Under $U_{1,0}$, both χ_{110} and χ_{011} change sign. Orthogonality gives $\sum\psi_z^2 = 16z^2$. Direct expansion of the quartic and summation of its 2^6 index assignments gives $V_{\text{tet}}[\psi_z] = 128z^4$. The two $\mathbb{Z}_2$ transformations used here act on distinct model components: diagonal projection identifies gauge-related configurations, whereas $U_{1,0}$ acts nontrivially on the projected field sector.

C.5 Pushforward Of Complex Ensembles

Proof of Proposition 5.3. By Eq. (29),

$$\begin{aligned} Q_{gC}(gA) &= \int_{\mathscr{K}} R_{gC}(gA \mid k')\, \pi_{gC}(dk') \\ &= \int_{\mathscr{K}} R_{gC}(gA \mid gk)\, \pi_C(dk) \\ &= \int_{\mathscr{K}} R_C(A \mid k)\, \pi_C(dk) \\ &= Q_C(A). \end{aligned}$$

The second equality changes variables by the measurable group action $k' = gk$ and uses the pushforward identity for π_{gC} in Assumption 5.1. The third equality applies covariance of the readout kernel. The argument requires the readout to be a genuine probability kernel on a measurable space of labeled complexes; it does not follow by replacing π_C with arbitrary complex-valued field amplitudes. □

C.6 Comparison of Two Constructions

Consider a finite, type-preserving rule sequence $h = (q_1, \ldots, q_n)$ with its specified match data. An encoding into a two-complex may use vertices for events, edges for persistent typed incidences, and faces for compatible strands of repeated participation. Even in a finite toy setting, a map $F(h)$ generally discards temporal durations and rate information unless these are explicitly labeled. Thus the condition $E_C(h) = R_C(\cdot \mid F(h))$ in Eq. (30) is an additional constraint, not a consequence of constructing F.

For an exact map F respecting group actions, $F(gh) = gF(h)$, and a readout satisfying Eq. (28), covariance of the chemically interpreted history follows by substitution. Conversely, finding one observable for which $E_C(h)$ and $R_C(\cdot \mid F(h))$ differ disproves that candidate bridge at the declared resolution. If the terminal structural map T_r discards attributes needed to predict this observable, the failure may be traced to the coarsening rather than to a failure of either separate generative formalism.

No field-theoretic perturbation parameter is interpreted here as laboratory time, and no Feynman amplitude is regarded as a reaction propensity. Such identifications would require their own physical construction and independent evidence. These restrictions preserve the intended independence of the formalisms while giving a meaningful test of proposed correspondences.

D Structural Fibers And History-Dependent Response

D.1 Proof Of The Functional-Factorization Criterion

Put $q_r : D \to \mathcal{Q}_r$, $q_r(x) = [T_r(x)]$, where $\mathcal{Q}_r$ is the quotient of the represented structural image under the declared isomorphism relation. Give $\mathcal{Q}_r$ the final σ-algebra for q_r: a subset A is measurable if $q_r^{-1}(A)$ is measurable in D. This choice ensures that a function on $\mathcal{Q}_r$ is measurable whenever its composition with q_r is measurable. (Alternatively, for a finite or countable graph-state model, use the discrete σ-algebra.)

Proof of Proposition 6.1. Suppose first that $K_{C,J}(\cdot \mid x) = \overline{K}_{C,J}(\cdot \mid q_r(x))$. If $q_r(x) = q_r(x')$, their two response kernels coincide by substitution; thus $K_{C,J}$ is constant on fibers. Conversely, assume fiber constancy. For every $s \in \mathcal{Q}_r$ choose any x_s with $q_r(x_s) = s$ and define

$$\overline{K}_{C,J}(A \mid s) = K_{C,J}(A \mid x_s)$$

for each measurable output set A. Fiber constancy makes this independent of the representative x_s. Every $\overline{K}_{C,J}(\cdot \mid s)$ is a probability measure since every $K_{C,J}(\cdot \mid x_s)$ is. Moreover, for fixed A the function $s \mapsto \overline{K}_{C,J}(A \mid s)$ is measurable with respect to the final σ-algebra because its pullback along q_r is the measurable function $x \mapsto K_{C,J}(A \mid x)$. Hence $\overline{K}_{C,J}$ is a valid Markov kernel and Eq. (32) holds. Finally, unequal kernels at two points of the same fiber contradict fiber constancy and therefore disprove functional sufficiency. □

Remark D.1 (Which Equivalence Is Being Tested?). The equivalence $q_r(x) = q_r(x')$ is about the chosen structural representation. Enlarging the representation to include retained configurations, mesoscale defects, or relevant internal variables can split a coarse fiber into finer fibers. A previously established failure of sufficiency at resolution r can therefore disappear at a finer resolution r'. Such a refinement is expected in physically mediated history dependence and does not invalidate the earlier resolution-relative result.

D.2 Stochastic Mixtures And Failure Of Identification

For a preparation i yielding endpoint law ν_i, the law of total probability gives Eq. (33). If the two preparation laws coincide as measures on $\mathcal{X}$, then the induced output laws coincide for every common kernel $K_{C,J}$. If $\nu_1 \neq \nu_2$, distinct output laws are possible but need not follow. For instance, choose $K_{C,J}(\cdot \mid x)$ constant in x; then $P_1 = P_2$ for arbitrary endpoint laws. Thus the existence of distinct histories, and even distinct endpoint distributions, is insufficient for a functional difference without a separating observation.

For the binary-response model in Eq. (34), put $p_+ = (1+\delta)/2$ and $p_- = (1-\delta)/2$. The total-variation distance between two Bernoulli laws is $|p_+ - p_-| = \delta$. Both full states map to the same coarse topology s, hence the quotient criterion of Proposition 6.1 is violated for the specified functional response. An observer who measures the retained variable z instead uses a finer representation $T_{r'}(s,z) = (s,z)$ and can define a response kernel that factors through that enlarged structure.

D.3 History Variables And Conditional Sufficiency

Let $(X_t)_{t\geq 0}$ be a Markov process in a fixed context, and let Y be any output sampled from the future trajectory after τ using a common testing protocol. The Markov property implies

$$\Pr(Y \in A \mid X_\tau, \{X_s : s < \tau\}) = \Pr(Y \in A \mid X_\tau)$$

for all measurable A when the observation is itself determined by the post-τ process and its independent testing noise. A symmetry-history statistic $B(\gamma)$ measured before τ then supplies no further predictive information conditional on a complete state X_τ. If an empirically estimated history effect persists after conditioning on the available state, possible explanations include omitted state variables, unmodeled context dependence, incorrect Markov closure, or a response criterion that evaluates historical origin itself. It would be unwarranted to identify any one explanation without further evidence.

These observations motivate a falsifiable hierarchy: test sufficiency of T_r, then test an enriched $T_{r'}$, then test a candidate dynamical state closure. The sequence connects the philosophical notion of generative history to ordinary inference about physical memory, without requiring a separate historical force.

E Symmetry Selection And Operational Functions

E.1 Odd-Observable Selection Rule

Proof of Proposition 7.1. Since $px_0 = x_0$, oddness implies $q(x_0) = q(px_0) = -q(x_0)$, so $2q(x_0) = 0$ and $q(x_0) = 0$. Therefore $|q(x_0)| > \delta$ is false for every $\delta > 0$. The second assertion holds only for a protocol evaluating q in the reference state without first applying a symmetry-breaking intervention. If interventions can induce an asymmetric state, the reference state's broader intervention-dependent capability set must be recalculated rather than inferred from its initial q value. □

For an invariant probability law ν satisfying $\nu(pA) = \nu(A)$ and an integrable odd observable q, we also obtain

$$\int q(x)\nu(dx) = \int q(px)\nu(dx) = -\int q(x)\nu(dx) = 0.$$

This is a statement about the ensemble mean, not an assertion that $q(x) = 0$ for each individual realization. In an equivariant stochastic dynamics with a symmetric initial law, the law remains symmetric when the corresponding Markov evolution is well defined and unique. Nevertheless, individual paths can leave the fixed-state set and select signs. A finite-volume ensemble may still mix both signs, even when conditionally selected sectors display large order parameters.

E.2 Stability Of The Cubic Normal Form

Write $f(m) = am - bm^3$. The equilibria solve $m(a - bm^2) = 0$. For $a < 0$, only $m = 0$ exists and $f'(0) = a < 0$, so it is locally asymptotically stable. At $a = 0$, $\dot{m} = -bm^3$ still attracts nearby states with algebraically slow relaxation; linearization alone is inconclusive. For $a > 0$, the fixed points are 0 and $m_\pm = \pm\sqrt{a/b}$. Here $f'(0) = a > 0$, whereas

$f'(m_\pm) = a - 3bm_\pm^2 = -2a < 0$, establishing instability of the origin and asymptotic stability of the two nonzero equilibria. The vector field satisfies $f(-m) = -f(m)$, hence is equivariant under reflection for all parameter values.

The gradient representation $\dot{m} = -\partial_m U(m)$ has $U(m) = -am^2/2 + bm^4/4$. When $a > 0$ this potential has two minima of equal depth. Exact deterministic evolution from the invariant state $m(0) = 0$ remains there by uniqueness, even though it is unstable. In a stochastic extension, for example a reflection-symmetric diffusion with $dm = f(m)dt + \sigma\, dW_t$ and symmetric noise, the distribution of m_t remains reflection-invariant from symmetric initial data, while individual realizations may depart from zero. Care is required to distinguish transient sign selection, noise-driven switching, and the persistence of broken-symmetry states in an appropriate limiting regime.

E.3 A Sufficient Conditional Capability Construction

Let $q(x_m) = m$ be a separately specified physically measurable odd response. Fix $a > 0$, a threshold $0 < \delta < \sqrt{a/b}$, and a test that passively records q without interventions that alter m on the testing timescale. Let f_δ denote the functional task of producing a response magnitude exceeding δ. Then $f_\delta \notin \mathcal{F}(m = 0)$, while $f_\delta \in \mathcal{F}(m = m_+)$ and $f_\delta \in \mathcal{F}(m = m_-)$. The reference and selected states must be compared under the same protocol. This construction satisfies the mathematical form of Eq. (35) in an abstract model, conditional on the claimed readout. It demonstrates logical consistency, not physical realizability of the proposed chemical mechanism.

Conversely, choose a constant readout $q(x) = 0$ for all states or a task family depending only on an invariant quantity shared across the states. In that case the same symmetry-breaking dynamics yields no new functional capability according to the selected protocol. Therefore, the group-theoretic symmetry reduction alone is insufficient to derive a universal functional-gain theorem.

E.4 Relation To Representations And Defects

Suppose an observable transforms in a nontrivial one-dimensional representation $\chi : G \to \{\pm 1\}$. If x is fixed by some g with $\chi(g) = -1$, then covariance $q(gx) = \chi(g)q(x)$ implies $q(x) = 0$. This generalizes the elementary $\mathbb{Z}_2$ selection rule to any symmetry that acts with a sign change in the chosen response channel. For vector- or tensor-valued observables, the allowed response at x lies in the fixed subspace of the residual representation $\rho(H(x))$. Reduction of $H(x)$ may enlarge this fixed subspace, but whether the enlarged components are actually nonzero depends on interactions, constitutive coefficients, and preparation. The formal selection rule gives an *allowance*, not a guarantee of spontaneous generation or useful function.

Similarly, when a homogeneous order-parameter space G/H is physically justified, nontrivial homotopy groups identify candidate topological defect charges. Their existence as mathematical classes does not suffice to establish dynamical formation or functional relevance. Boundary conditions, conservation constraints, kinetic pathways, and lifetimes must be supplied by a physical model. These qualifications prevent extrapolation from group-theoretic classifications to unsupported claims about a chemical product.

F Contrastive Models And Limits Of Ontological Inference

F.1 Proof Of Restricted Observational Equivalence

Proof of Proposition 8.1. Fix a protocol $\pi \in \mathcal{P}_0$ with a finite observation schedule $t_0 < \cdots < t_n$ and measurable observation outputs $Y_0, \ldots, Y_n$. For each model, the joint law is formed by composing the initial-state distribution, successive controlled transition kernels, and observation kernels at the selected times. By hypothesis, all these factors coincide in both models for this protocol. Induction over the finite sequence of kernel compositions therefore gives identical joint measures for the observed sequence. Marginalization preserves equality, so any finite collection of output statistics has the same law. Since π was arbitrary, the models agree over all stipulated finite schedules in $\mathcal{P}_0$. The conclusion imposes no equality on a kernel under a context or intervention not belonging to $\mathcal{P}_0$. □

F.2 An Explicit Two-State Alternative

Take the observed state space $\{0, 1\}$, choose $0 < p, q < 1$ with $p \neq q$, and define Markov kernels K_0 and K_* whose rows are independent of the previous state:

$$K_0 = \begin{pmatrix} 1-p & p \\ 1-p & p \end{pmatrix}, \qquad K_* = \begin{pmatrix} 1-q & q \\ 1-q & q \end{pmatrix}.$$

Let all presently allowed protocols use only context C_0. Model $\mathfrak{M}_-$ uses K_0 in every context. Model $\mathfrak{M}_+$ has an additional latent selector $z \in \{0, 1\}$, fixed to $z = 0$ in context C_0 and activated to $z = 1$ in a new context C_*. Conditional on $z = 0$ it uses K_0, and conditional on $z = 1$ it uses K_*. Identical initial observed-state and observation laws are stipulated. Every protocol confined to C_0 therefore yields identical observed laws in the two models by Proposition 8.1. Under a one-step test in C_*, however,

$$\Pr_{\mathfrak{M}_+}(X' = 1 \mid C_*) = q, \qquad \Pr_{\mathfrak{M}_-}(X' = 1 \mid C_*) = p.$$

Thus their predictive difference is $|q - p| > 0$ in a new domain. The construction illustrates why current observation does not settle a claim about a proposed higher-order selector, while also displaying a route to distinguish the hypotheses through additional evidence. It does not show that such a selector actually exists in chemistry.

To give z an ontic meta-generative interpretation, a scientific account must specify what physical object or organizational degree of freedom it denotes, how it changes the underlying transition mechanism, and why it constitutes a distinct level rather than a useful parameterization. The flat model can be extended differently, and hidden mechanisms are never established merely because one may write them in a larger state space. Consequently, opposite truth assignments to $H_M(D)$ require declared semantics of that hypothesis and scientifically coherent interpretation of the competing models; formal parameter augmentation alone cannot carry the philosophical conclusion.

F.3 Proof-Theoretic Scope And The Problem Of Trivialization

The countermodel result of Proposition 3.1 is semantic and relative to $\mathbb{T}$. If the theory already contains H_M as an axiom, a model of $\mathbb{T} \wedge \neg H_M$ cannot exist; similarly, if it contains $\neg H_M$, a model of the opposite cannot exist. Absence of an established pair is not a proof that one side is entailed, since constructing physical model classes is itself a substantive undertaking. Nor does non-entailment from a finite evidence set establish that every possible evidence set remains indecisive.

There are three further limitations. First, an existentially quantified formal object can often be postulated by encoding a structure collection into a grammar; this does not satisfy the substantive definition of $H_M(D)$ in Definition 8.1. Second, observational equivalence at one resolution can be broken when a more discriminating measurement or intervention is admitted. Third, if two ontological accounts are equivalent under all physically accessible interventions, pragmatic or metaphysical criteria may still favor one, but such criteria should be stated rather than smuggled into the empirical data. These qualifications preserve the paper's principal methodological distinction: the availability of a symmetry-first generative model differs from evidence for the unique or primitive existence of its generative ontology.

G Operational Bridges And Their Limits

G.1 Markov Contraction And Mixture Bounds

We give the proof of Proposition 10.1, with all relevant probability assumptions made explicit. On a measurable space $(\mathcal{X}, \mathcal{B})$ let P and Q be probability measures, and let $K(dy \mid x)$ be a Markov kernel from $\mathcal{X}$ to another measurable space $(\mathcal{Y}, \mathcal{D})$. Define the total-variation distance by

$$d_{\mathrm{TV}}(P, Q) = \sup_{A \in \mathcal{B}} |P(A) - Q(A)|. \tag{56}$$

For any measurable $A \subseteq \mathcal{Y}$ the function $k_A(x) = K(A \mid x)$ is measurable and lies in $[0, 1]$. Writing $\sigma = P - Q$, one has

$$\begin{aligned} |(PK)(A) - (QK)(A)| &= \left| \int_{\mathcal{X}} k_A(x)\, \sigma(dx) \right| \\ &\leq d_{\mathrm{TV}}(P, Q), \end{aligned} \tag{57}$$

where the last inequality follows from the Hahn decomposition, or equivalently from the layer-cake representation $k_A(x) = \int_0^1 \mathbf{1}_{\{k_A(x) > t\}}\, dt$. Taking the supremum over output events proves Eq. (51).

Now let κ be a probability measure on grammatical histories and let $E(dx \mid h)$ and $\widetilde{E}(dx \mid h)$ be two endpoint kernels. If

$$d_{\mathrm{TV}}(E(\cdot \mid h), \widetilde{E}(\cdot \mid h)) \leq \varepsilon \tag{58}$$

for κ-almost every h, then their mixtures $P = \int E(\cdot \mid h)\kappa(dh)$ and $Q = \int \widetilde{E}(\cdot \mid h)\kappa(dh)$ satisfy

$$d_{\mathrm{TV}}(P, Q) \leq \int d_{\mathrm{TV}}(E_h, \widetilde{E}_h)\kappa(dh) \leq \varepsilon.$$

A subsequent common response kernel cannot increase the discrepancy. In particular, setting $\widetilde{E}(\cdot \mid h) = \int R(\cdot \mid k)F(dk \mid h)$ quantifies the endpoint consequence of the bridge error in Eq. (50). The argument is independent of chemistry; establishing chemical relevance requires a correctly interpreted E, F, and R.

G.2 A Non-Identifiable Bridge Example

Consider two distinct grammatical histories h_1 and h_2, and let the complex space have a single element k_*. Define $F(k_* \mid h_i) = 1$ for both $i = 1, 2$. Any readout $R(\cdot \mid k_*)$ must then return the same output law for both histories, even if their true chemical endpoint laws $E(\cdot \mid h_1)$ and $E(\cdot \mid h_2)$ differ. Thus Eq. (50) cannot hold simultaneously for both histories unless the endpoint laws happen to coincide. A many-to-one bridge is therefore insufficient whenever its identified fibers contain chemically distinguishable states or histories under the selected protocol. Conversely, if the two histories are deliberately equivalent at the chosen resolution, a many-to-one encoding may be appropriate. The constraint is relative to the claimed explanatory target.

A second limit arises even when the map $h \mapsto k$ is injective. Let $R(\cdot \mid k) = \delta_{x_*}$ for all complexes k in a nontrivial orbit. Then structurally distinct complex histories become a single chemical state, and the output law cannot detect the complex-level group action. This example reproduces the logical point already encountered for field-level symmetry selection: the algebraic distinction must survive chemical realization and observation to carry an experimentally meaningful functional claim.

G.3 Symmetry Of The Integrated Observable Law

Suppose G acts measurably on the chemical endpoint space $\mathcal{X}$ and on the functional output space $\mathcal{Y}$, and Q_C and $K_{C,J}$ satisfy

$$\begin{aligned} Q_{gC}(gA) &= Q_C(A), \\ K_{gC,gJ}(gB \mid gx) &= K_{C,J}(B \mid x) \end{aligned} \tag{59}$$

for all measurable $A \subseteq \mathcal{X}$, $B \subseteq \mathcal{Y}$. Define $P_{C,J}(B) = \int K_{C,J}(B \mid x)Q_C(dx)$. Then substitution $x' = gx$ gives

$$\begin{aligned} P_{gC,gJ}(gB) &= \int K_{gC,gJ}(gB \mid x')Q_{gC}(dx') \\ &= \int K_{C,J}(B \mid x)Q_C(dx) \\ &= P_{C,J}(B). \end{aligned}$$

Thus an equivariant generative law and an equivariant evaluation protocol imply covariance of the resulting output law. This conclusion does *not* assert equality of the measured responses of two symmetry-related materials in one fixed, symmetry-breaking environment. Such equality would require holding the proper environmental transformation and the meaning of the measured output fixed consistently.

G.4 A Concise Falsification Checklist

A proposed correspondence can fail at independent steps: the graph rewrite may not preserve chemical accounting; the group action may not preserve reaction propensities; the complex distribution may not be a positive probability law; the readout may violate declared conservation rules; the bridge may discard a retained state variable; or the experimental protocol may be insensitive to the predicted broken-symmetry sector. Each failure is informative. Mathematical consistency tests should be reported separately

from physical validation; mechanistic parameters should be independently constrained where possible; and predictions should be assessed on observations that were not used merely to define the readout. These requirements are especially important for comparison of models with strongly different numbers of adjustable parameters.

H Compact Recapitulation of Two Formalisms

This appendix collects the principal mathematical objects and conditions used in the main text. It functions as a *reference map*, not as an additional set of axioms or a replacement for the detailed proofs in Appendices B–G. In particular, a formal construction and a physically validated chemical mechanism are distinct claims.

H.1 Common Modeling Objects And Their Meanings

Fix an effective physical-state space $\mathcal{X}$, environment space C, acting group G, and structure extraction $T_r : \mathcal{X} \to \mathcal{S}_r$ at resolution r. The principal distinction is

$$T_r(x_1) \cong T_r(x_2) \quad \not\Rightarrow \quad x_1 = x_2. \tag{60}$$

The projection may discard configuration, electronic state, collective order, or other information needed for a response test. We must therefore declare r and the precise types of graph isomorphism used. A group action requires an explicitly specified domain: structural symmetry, invariance of a probability law, and covariance of the dynamics are distinct propositions.

Table 5: Minimal notational crosswalk.

Symbol	Meaning and required declaration
$\mathcal{X}, C$	Effective states and environments; boundary between them fixed
T_r	Resolution-dependent structural projection
G, g	Group and a physically justified action on each domain
$\mathcal{R}, q$	Chemically interpreted production-rule family and one rule
$\mathcal{L}_C$	Continuous-time stochastic jump generator when defined
$\gamma, B(\gamma)$	Realized path and a declared history descriptor
$K_{C,J}$	Measurable state-to-response kernel for a fixed protocol
$\mathscr{K}, k$	Space of labeled relational two-complexes and one complex
ψ, S_Λ	Group field and regulated model action
π_C, R_C	Positive complex ensemble and chemical realization kernel
$H(x)$	State stabilizer within the environment-preserving group

H.2 Formalism I: Equivariant Rewriting And Response

The syntax is defined in Definition 4.1, and its state-level interpretation appears in Definition 4.3:

$$\begin{aligned} \Gamma &= (\Sigma_{\mathrm{T}}, \Sigma_{\mathrm{N}}, S, \mathcal{P}, \\ &\qquad \mathbf{ChemG}, C, G, \rho), \\ \mathbb{G} &= (\Gamma, \mathcal{X}, T_r, \mathcal{R}, \chi, \nu). \end{aligned} \tag{61}$$

A rule uses a typed DPO span $L_q \leftarrow K_q \rightarrow R_q$, a context predicate, and a separately specified physical-state lift. Its symmetry requirement is the conditional implication

$$x \xrightarrow{q,m;C} x' \implies gx \xrightarrow{gq,gm;gC} gx'. \tag{62}$$

Graph-theoretic admissibility is not a measured chemical reaction rate. Atom and charge accounting must be checked for the complete chemical system, including any reservoirs: compare Eq. (9).

For a countable-state jump process with valid propensities, the operational dynamics are specified by

$$(\mathcal{L}_C f)(x) = \sum_{q,m} a_{q,m}(x, C)\big[f(J_{q,m}x) - f(x)\big]. \tag{63}$$

Under the assumptions of Proposition 4.2, the transformed generator satisfies

$$(\mathcal{L}_{gC} f)(gx) = \mathcal{L}_C(f \circ g)(x). \tag{64}$$

A functional claim additionally needs a fixed response kernel $K_{C,J}(dy \mid x)$, measured or justified independently of the grammar. Neither Eq. (62) nor Eq. (64) determines a response kernel.

H.3 Formalism II: Complexes, Field Weights, And Chemical Readout

A labeled two-complex is the incidence-and-label datum of Definition 5.1; its cellular boundary maps obey

$$C_2(k) \xrightarrow{\partial_2} C_1(k) \xrightarrow{\partial_1} C_0(k), \quad \partial_1\partial_2 = 0. \tag{65}$$

Faces may denote selected structural motifs or relations among generative events, according to the separately declared interpretation. The same face cannot silently switch meanings across a comparison.

The group-field-inspired generator uses a regulated field $\psi : G^d \times M \to \mathbb{C}$ and interaction-dependent weights, represented schematically by

$$\begin{aligned} S_\Lambda[\psi, \bar\psi] &= \langle \psi, K_\Lambda \psi \rangle + \sum_\alpha \lambda_\alpha V_{\alpha,\Lambda} + S_{\mathrm{reg},\Lambda}, \\ Z_\Lambda &= \int e^{-S_\Lambda}\, d\psi\, d\bar\psi. \end{aligned} \tag{66}$$

The expansion organizes stranded Wick-contraction graphs, which can be interpreted as labeled two-complexes after their incidence data are extracted. Field evaluations $V_{\mathrm{tet}}[\psi_0]$ and Gaussian perturbative weights $\langle V_{\mathrm{tet}}^n \rangle_0$ are separate quantities. Field-theory amplitudes do not themselves define probabilities of chemical products. Where a positive normalized complex ensemble π_C and a physically defended readout kernel R_C exist, a chemical output law can be defined by

$$Q_C^{\mathrm{cx}}(dx) = \int_{\mathscr{K}} R_C(dx \mid k)\, \pi_C(dk). \tag{67}$$

The covariance result of Proposition 5.3 requires covariance of *both* the ensemble and readout. A symmetry-selected field configuration does not automatically yield a symmetry-selected chemical material.

H.4 Comparisons, Conjectures, And Proof Status

For matched preparation and context, operational comparison passes both models through the same response protocol:

$$P^i_{p,C,J} = Q^i_{p,C} K_{C,J}, \quad i \in \{\mathrm{gr}, \mathrm{cx}\}. \tag{68}$$

Proposition 10.1 gives the total-variation contraction bound for a shared response kernel. It does not prove that the two models are physically equivalent.

Three open inquiries remain distinct. Conjecture 6.1 concerns states with isomorphic *terminal structural projections* but different test responses; Conjecture 7.1 concerns new functional capabilities enabled under a specified symmetry-reducing transition; Conjecture 8.1 concerns evidence failing to force an independently existing meta-generative order. The third is evaluated using contrastive models and specified observations, rather than by treating an unsuccessful proof of existence as a proof of nonexistence.

Remark H.1 (Logical And Empirical Status Of The Recap). Conditional equivariance and probability-kernel identities admit mathematical proof under stated hypotheses; chemical realization maps, kinetic parameters, semantic protocols, and claims about ontological existence require separate justification. The recap introduces no further physical postulate beyond those already declared.

I Chemical Structures for Illustrative Model Checks

The following examples are chemically familiar and algebraically transparent. They are used to audit the expressive resolution of the two proposed formalisms and the meaning of their group actions. They are *illustrations rather than experimental demonstrations* of Conjectures 6.1–8.1. The examples deliberately avoid an inference from two configurations of a molecule to two distinct physical generation histories.

I.1 Geometric Stereoisomers And The Structural Fiber

The (Z) and (E) isomers of 1,2-dichloroethene have molecular formula $C_2H_2Cl_2$ and the same constitutional labeled graph, yet differ by the relative arrangement of chlorines around the C=C bond (Fig. 14a) (International Union of Pure and Applied Chemistry, 2025b; National Center for Biotechnology Information, 2026a,b). Define T_{conn} to retain atoms and bond orders but discard alkene stereochemical configuration. If x_Z, x_E describe idealized configured molecules, then

$$T_{\mathrm{conn}}(x_Z) \cong T_{\mathrm{conn}}(x_E), \quad T_{\mathrm{stereo}}(x_Z) \ncong T_{\mathrm{stereo}}(x_E). \tag{69}$$

The second relation uses an enriched representation whose morphisms preserve the selected relative configuration. The net electric dipole (in the molecular rest frame, with a fixed geometrical convention) supplies a particularly simple physical-response distinction: the idealized trans arrangement is centrosymmetric and has vanishing permanent dipole, while the cis arrangement permits a nonzero permanent dipole. The implication follows at the ideal symmetry level; precise measured moments require external chemical data and specified conditions. An abstract response protocol J_{dip} may therefore distinguish the isomers despite the common constitutional graph.

This example illustrates why a structural language that encodes only unlabeled or constitutionally labeled connectivity can fail to be sufficient for a functional observation. It is *not* a test of generative-history dependence: a difference in E/Z configuration already differentiates the complete stereochemical terminal states. The symmetry relating two spatial configurations must not be confused with the stabilizer of a given molecular geometry or with the group that leaves the surrounding environment invariant.

I.2 Enantiomers And Covariance Of Context

Bromochlorofluoromethane CHBrClF has a tetrahedral carbon bearing four distinct substituents. Its mirror configurations are non-superposable, while their atom- and bond-labeled constitutional graphs are isomorphic (Fig. 14b) (International Union of Pure and Applied Chemistry, 2025a). Let p denote an orientation-reversing map on the enriched molecular-state space, and suppose $x_- = px_+$. Then

$$T_{\mathrm{conn}}(x_+) \cong T_{\mathrm{conn}}(x_-), \qquad p^2 = e. \tag{70}$$

The statement $p^2 = e$ describes an abstract mirror operation; it does not say that the individual chiral molecule possesses mirror symmetry. The molecules need not have a shared physical stabilizer containing p.

In a fixed chiral receptor or solvent context C_χ, the response kernels $K_{C_\chi,J}(\cdot \mid x_+)$ and $K_{C_\chi,J}(\cdot \mid x_-)$ may differ. Yet a consistent simultaneous transformation of both molecule and context can yield covariance,

$$K_{pC,J}(pD \mid px) = K_{C,J}(D \mid x) \tag{71}$$

for an appropriately transformed response event D and parity-compatible protocol J. Equation (71) is a *conditional model symmetry*, not a universal empirical identity for all interactions. It isolates an essential lesson for the two formalisms: equivalent structures under a coarse graph semantics need not exhibit identical effects in a *fixed* chiral environment, while equivariance is formulated by transforming the full context as well.

I.3 Chemical Rule with Explicit Material Accounting

A familiar acid–base event supplies a small illustration of Formalism I. Consider the equilibrium

$$CH_3COOH + H_2O \rightleftharpoons CH_3COO^- + H_3O^+. \tag{72}$$

At a constitutional level, an abstract DPO-style rule can delete the matched acid oxygen–hydrogen incidence and insert an oxygen–hydrogen incidence on the base, together with appropriate charge-state changes. This is a *coarse-grained event schema*, not an elementary transition-state geometry. Both total elemental composition and net charge are conserved in Eq. (72). The labels on all species and the rule's effective state lift must preserve this accounting.

The ambient solvent, temperature, and chemical activities affect the effective propensities; an actual microscopic proton transfer may involve solvent-mediated rearrangements not represented by a single elementary graph edit. Accordingly, the graph rule supplies structural bookkeeping while a context-dependent kinetic model supplies physical frequency or probability. The choice of an oriented rule does not itself determine reversibility, equilibrium constants, or a catalytic mechanism.

I.4 Symmetric Autocatalytic Rule Family

A deliberately abstract variant of Frank-type chiral amplification supplies a *generative-process* example rather than another static structure comparison (Frank, 1953; Blackmond,

2010). Let A be a precursor and L, R two mirror-related outputs, with identical elemental composition in the idealized accounting. A symmetric pair of catalyzed conversion rules and a mutual-antagonism rule are

$$A + L \xrightarrow{k} 2L, \qquad A + R \xrightarrow{k} 2R, \qquad L + R \xrightarrow{k_i} D. \tag{73}$$

Here D is a nominal product containing the combined material of L and R; the symbols do not identify an experimentally verified collection of species. Mirror exchange $p : L \leftrightarrow R$ fixes A, D and maps the first rule into the second while preserving the third. Thus the *rule family* is $\mathbb{Z}_2$-equivariant.

Under ideal homogeneous mass-action kinetics and with concentrations a, ℓ, r, a simplified closed-batch description has

$$\begin{aligned} \dot{\ell} &= ka\ell - k_i \ell r, \\ \dot{r} &= kar - k_i \ell r, \\ \dot{a} &= -ka(\ell + r), \\ \dot{d} &= k_i \ell r. \end{aligned} \tag{74}$$

The output d denotes the concentration of D. Let $m = (\ell - r)/(\ell + r)$ when $s = \ell + r > 0$. Direct differentiation yields

$$\dot{m} = \frac{k_i s}{2} m(1 - m^2). \tag{75}$$

Subtracting the two rate equations also gives the exact identity

$$\delta = \ell - r, \qquad \dot{\delta} = ka\delta. \tag{76}$$

For $a \geq 0$, the sign of any nonzero initial imbalance is preserved. In the idealized system with $k_i > 0$, $\ell_0, r_0 > 0$ and $\delta_0 \neq 0$, the ratio obeys $\frac{d}{dt} \log(\ell/r) = k_i \delta$. Since $|\delta(t)| \geq |\delta_0|$, the ratio tends monotonically toward $+\infty$ or zero, and $m(t) \to \operatorname{sgn}(\delta_0)$. This asymptotic statement concerns the ideal irreversible model and its stated positivity assumptions. For positive cross-inhibition, the magnitude of a nonzero small imbalance increases while both enantiomers remain present. The exactly racemic deterministic solution $m = 0$ persists; a selected branch requires an initial perturbation or a stochastic preparation mechanism. The reduced irreversible scheme is a kinetic illustration, and a physically realizable system would require chemically specified species, reversible pathways or driven conditions as appropriate, and thermodynamically consistent rate parameters. In particular, $\mathbb{Z}_2$ covariance of the reaction family does not by itself imply functional novelty: a new response task must be assessed separately under the protocol of Definition 7.1.

I.5 Benzene and the Choice of a Two-Complex

The benzene carbon skeleton supplies a transparent comparison of graph and two-complex descriptions. Let C_6 denote the six-carbon cycle graph, suppressing the attached hydrogens and the electron-density information. As a one-dimensional CW complex, C_6 has

$$H_1(C_6; \mathbb{Z}) \cong \mathbb{Z}. \tag{77}$$

Choose to attach a 2-cell f along the six-edge cycle, obtaining a disk-like CW complex K_f. With consistently oriented edges,

$$\partial_2 f = \sum_{i=1}^{6} e_i, \quad \partial_1 \partial_2 f = 0, \quad H_1(K_f; \mathbb{Z}) = 0. \tag{78}$$

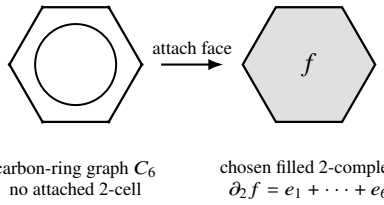


Figure 15: Two abstract topological representations of the benzene carbon skeleton (hydrogens omitted). Attaching a 2-cell to the six-edge ring changes the homology of the chosen representation; it is a modeling decision, not a claim that aromatic electron delocalization is literally a topological face or a GFT history.

This calculation demonstrates that homological summaries depend on the chosen *lifting* of a chemical graph into a two-complex. Neither the attached face nor the resulting homology group is, on its own, a physical explanation of benzene aromaticity. Nor does it establish an equivalence between a molecular ring face and an interaction-history face of a group-field diagram. To earn explanatory value, the selected 2-cell and its labels must encode additional experimentally or theoretically grounded chemical relationships.

I.6 How These Cases Probe the Proposed Claims

The E/Z and enantiomer pairs discriminate coarse connectivity from sterically enriched chemical states, revealing the importance of specifying the equivalence relation in Conjecture 6.1. The acid–base event checks reaction accounting and context-specific propensities in Formalism I. The symmetric autocatalytic network illustrates how a group-equivariant rule family can amplify a selected branch without supplying a new function by assumption. The benzene construction tests what a putative higher-order representation genuinely adds beyond an ordinary ring graph. None of these examples identifies a validated chemical readout from the GFT-inspired field construction; that remains a separate open task. A proper empirical test of history dependence must compare matched final enriched states or deliberately quantify retained defects and other preparation-dependent variables, while a proper symmetry-breaking test must resolve a group, an order parameter, and a prespecified functional response.

J Visual Atlas Of Generative And Physical Interpretations

The diagrams in the main text form a deliberately ordered visual introduction to both generative formalisms. They connect familiar chemical constitution, symbolic rewrites, group-equivariant transformations, and higher-dimensional relational representations without equating these mathematical descriptions with verified molecular mechanisms. This appendix records the exact interpretation and limitations of each drawing so that each can be read alongside its formal definition.

Visual case	Question made visible
Fig. 3	How can a typed rule create the next chemically labeled backbone while accounting for the byproduct?
Fig. 4	Which atomic incidences change in a local matched graph rewrite?
Fig. 5	What transforms when a group acts on both a rule and its chemical context?
Fig. 6	Can multiple derivations end at the same coarse chemical product?
Fig. 7	Which extra topological object is introduced when a molecular ring is filled?
Fig. 8	Where do group-field gluing and chemical realization diverge?
Fig. 9	What information may be erased by a coarse structural map?
Fig. 10	How can a symmetric rule family admit asymmetric realizations?
Fig. 13	What is required to compare the two formalisms experimentally?

J.1 Explicit Atom-and-Bond Ledger

For the net acid–base rule in Fig. 4, introduce an atom-mapped proton h^* originally attached to an acid oxygen o_a and an oxygen o_w belonging to water. At the graph level, the relevant incidence update is

$$\{(o_a, h^*)\} \longmapsto \{(o_w, h^*)\}. \tag{79}$$

It is accompanied by the local charge-state relabelings of the two oxygen-containing molecular fragments and by preservation of all other specified atoms. This accounting is a small typed rewriting example. Bond-order representations, proton solvation, transition-state structure, diffusion, reaction reversibility, and rates lie in the state lift and the separate chemical model. The ledger is not a claim that a multistep solvent proton shuttle is a single elementary event.

J.2 Inference Limits of Diagrammatic Images

The diagrams have three roles. *Construction diagrams* state operations internal to a formalism (e.g., DPO-like graph edits, attaching a 2-cell, or gluing group-field strands). *Chemical illustrations* connect the notation to known types of molecule and reaction while supplying explicit limitations. *Conjecture diagrams* specify potential distinguishing measurements without claiming that their abstract outcomes have been experimentally observed. In every case, visual similarity is weaker than a structure-preserving mathematical map. In particular, a filled benzene polygon, a GFT-like face, and a correlated chemical mechanism require separate definitions and cannot be substituted for one another.

K Implementation-Oriented Restatement of Two Formalisms

This appendix restates the two formalisms in a way that supports direct coding and independent formal checking. The emphasis is on declared data structures, execution order, and interfaces among modules. The appendix complements the mathematical definitions in Sections 4–5 and the compact recap in Appendix H.

K.1 Formalism I: Executable Grammar Object

A software representation of the grammar in Definition 4.1 can use the following records.

1. **Terminal types** Σ_T: atom labels, bond labels, optional attributes, and admissible terminal annotations.
2. **Nonterminal types** Σ_N: symbol name, port arity, and port typing constraints.
3. **Start graph** S: a typed attributed graph together with a list of marked nonterminal sites.
4. **Production schema** p: left nonterminal symbol, interface ports, replacement graph H_p, and predicate χ_p on the host state and context.
5. **Physical rule** q: a DPO span $(L_q \leftarrow K_q \rightarrow R_q)$, admissibility predicate, and state-lift map ν_q.
6. **Environment record** C: declared fields, reservoirs, control variables, and any fixed external constraints.
7. **Group action** ρ: methods that transform states, contexts, labels, matches, and rules.

A derivation engine can then be stated as follows.

1. Initialize the sentential graph with S.
2. Identify each nonterminal site and the admissible production schemas that match its port type.
3. Apply one schema, producing a new sentential graph.
4. Continue until all sites are terminal and the closure checks pass.
5. Where a physical interpretation is desired, pair each rule application with its state lift, propensity, and response-kernel update.

A practical implementation can store the formal language $L_{\text{term}}(\Gamma; C)$ separately from the stochastic process generated by Eq. (11). This separation is especially helpful for verification because the grammar can be unit-tested before kinetic parameters are introduced.

K.2 Minimal Rule Format for Reproducible Code

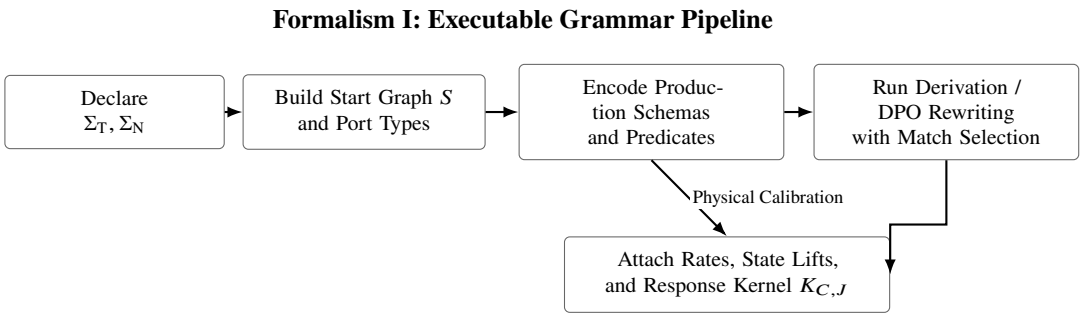
A convenient serializable rule format is

$$q = \big(\texttt{id}, L_q, K_q, R_q, \chi_q, \nu_q, \texttt{metadata}\big), \tag{80}$$

where `metadata` may contain stoichiometric bookkeeping, atom identity maps, default rate-law parameters, and documentation. For coding purposes, the following checks are useful.

1. Interface preservation: every node or port in K_q appears consistently in L_q and R_q.
2. Typing consistency: atom and bond types match the host graph and replacement graph.
3. Accounting consistency: declared atomic counts and charges agree with the intended whole-system interpretation.
4. Action consistency: for each selected $g \in G$, the transformed rule gq is represented by the same data format.

These checks can be implemented as assertions prior to any stochastic simulation.



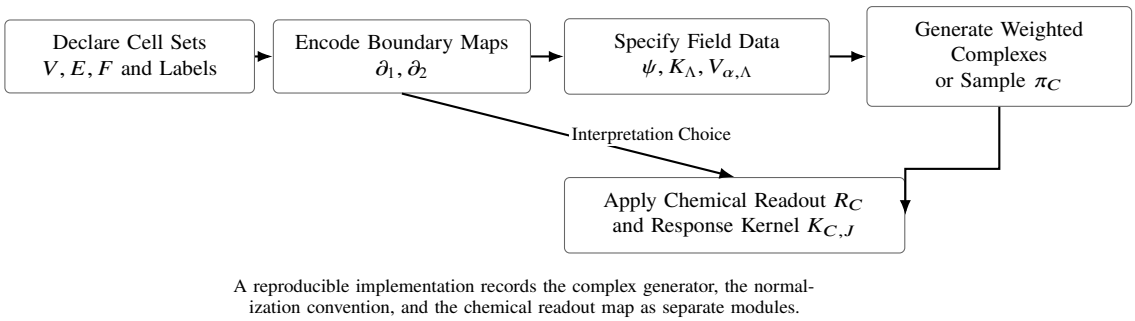


Figure 16: Implementation-Oriented View of Two Formalisms. Each formalism can be translated into software by storing the declared symbols, state objects, transformations, and readout maps in separate modules. The panel summarizes an execution order that can be mirrored in a codebase.

K.3 Formalism II: Executable Complex / Field Object

A software representation of Formalism II can use the records

$$\mathcal{K} = \big(V, E, F, \partial_1, \partial_2, \ell_V, \ell_E, \ell_F\big), \tag{81}$$

together with a field-theoretic record

$$\mathfrak{F} = \big(G, d, M, \psi, K_\Lambda, \{V_{\alpha,\Lambda}\}_\alpha, S_{\mathrm{reg},\Lambda}\big). \tag{82}$$

A coded workflow can proceed by

1. generating or loading admissible complexes,
2. checking the chain condition $\partial_1\partial_2 = 0$,
3. evaluating the chosen field weight or normalized ensemble,
4. sampling a complex or integrating over the ensemble,
5. applying the chemical readout R_C,
6. evaluating the response kernel $K_{C,J}$.

This modular order makes it straightforward to test distinct choices of complex interpretation without changing the field generator.

K.4 Cross-Formalism Verification Targets

For a code-based comparison, three kinds of outputs are especially useful.

1. **Terminal structural outputs**: graphs, isomorphism classes, or labeled complexes.
2. **Dynamical outputs**: path distributions, event counts, or sampled trajectories.
3. **Functional outputs**: the response law $P_{x,C,J}$ or a declared functional of that law.

A verification study can therefore ask whether the same terminal structure is reached, whether the same output law is produced after calibration, and which assumptions are responsible for any agreement.

Remark K.1 (Scope of the Implementation Restatement). The present appendix offers a translation guide from mathematics to code. It preserves the paper's separation between formal derivation, physical interpretation, and ontological discussion. In this way, an external reader can reproduce the constructions step by step, compare them with the detailed proofs, and attach further mechanistic modules where desired.

L Exact Checks for Finite Case Studies

The following calculations provide independent check values for the examples in Section 9. They are formulated as algebraic identities and finite combinatorial invariants, permitting implementation in graph, linear-algebra, or finite-group software without prescribing a particular programming language.

L.1 Orientation And Orbit Calculation For The Tetrahedral Grammar

Let $\mathbf{v}_1, \dots, \mathbf{v}_4$ be four noncoplanar ligand endpoints with fixed labels, and set

$$D = \det[\mathbf{v}_2 - \mathbf{v}_1,\ \mathbf{v}_3 - \mathbf{v}_1,\ \mathbf{v}_4 - \mathbf{v}_1].$$

For an orthogonal transformation Q, translation $\mathbf{t}$, and $\mathbf{v}_i' = Q\mathbf{v}_i + \mathbf{t}$, multilinearity gives $D' = \det(Q)D$. Proper rotations have $\det Q = +1$, and a reflection has $\det Q = -1$. In a rigid tetrahedral model, the two signs identify the two classes modulo proper rotations. With the nontrivial element $p \in \mathbb{Z}_2$ represented by reflection, $p^2 = e$ and $pX_+ = X_-$, so the terminal orbit has two members while the constitutional star has a single labeled-graph isomorphism class.

A finite regression table can therefore test: five atomic vertices; four atom–bond edges; zero nonterminal sites in terminal products; both orientation signs; equality of constitutional classes; and exchange of signs under the chosen reflection. These checks are conditional on a fixed ligand ordering and nondegenerate spatial embedding.

L.2 Cycle Graph and Filled Two-Complex

Orient the six edges of a ring cyclically. With vertex basis $v_1, \dots, v_6$, the boundary matrix D_1 has column i equal to $-v_i + v_{i+1}$, indices modulo six. Its rank is 5. The unfilled ring has D_2 with zero columns, so $\operatorname{rk} H_1 = 6 - 5 = 1$. Attaching a two-cell f with $\partial_2 f = \sum_{i=1}^{6} e_i$ adds the column $(1, 1, 1, 1, 1, 1)^{\mathsf{T}}$, which spans $\ker D_1$. Thus the filled model has $H_1 = 0$ and $H_2 = 0$. This computation concerns the selected complex representation.

L.3 Two-Vertex Stranded Dipole Boundary Matrices

Order the vertices (u, v), propagator edges (e_1, e_2, e_3, e_4), and faces $(f_{12}, f_{13}, f_{14}, f_{23}, f_{24}, f_{34})$. Then the integer boundary matrices are

$$D_1 = \begin{pmatrix} -1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}, \tag{83}$$

$$D_2 = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 & -1 & -1 \end{pmatrix}. \tag{84}$$

Direct multiplication gives $D_1 D_2 = 0$. Over $\mathbb{Q}$, $\operatorname{rk} D_1 = 1$ and $\operatorname{rk} D_2 = 3$. Moreover, columns f_{12}, f_{13}, and f_{14} of D_2 form a primitive basis of the integer kernel of D_1; hence $\operatorname{im} D_2 = \ker D_1$ and there is no residual H_1 torsion. The rank-nullity identities give $b_0 = 2-1 = 1$, $b_1 = 4-1-3 = 0$, and $b_2 = 6 - 3 = 3$, with free integral homology as stated.

L.4 Explicit $\mathbb{Z}_2$ Interaction Checks

For the quartic interaction in Eq. (18), the six summation indices a, b, c, d, e, f each take values in $\{0, 1\}$. For the constant field $\psi_0 = 1$ there are $2^6 = 64$ identical terms. For $\psi_1(a, b, c) = (-1)^a$, the product of four fields is $(-1)^{a+c+e+f}$. Therefore

$$V_{\text{tet}}[\psi_1] = \left(\sum_{a,c,e,f=0}^{1} (-1)^{a+c+e+f} \right) \left(\sum_{b,d=0}^{1} 1 \right) = 0.$$

Under simultaneous left translation U_p, the constant field stays fixed and ψ_1 changes sign. The quartic polynomial is invariant, so the values 64 and 0 remain unchanged. A finite enumeration or symbolic expansion checks these identities independently of the perturbative interpretation of the field action.

L.5 Exhaustive Slot-Preserving Gluing Counts

The four field factors in Eq. (18) are (a, b, c), (c, d, e), (e, b, f), and (f, d, a). At one interaction vertex, each pair of factors shares exactly one strand index, giving the six pairings a, b, c, d, e, f. At two vertices, pair factor i of the first copy to factor $\sigma(i)$ of the second by identifying each of its three argument slots with the *same* slot of the paired factor. Every strand then alternates between an internal interaction identification and a propagator identification. Tracing the resulting closed cycles counts the faces. More formally, let

$$H = \{A, B\} \times \{1, 2, 3, 4\} \times \{1, 2, 3\}$$

be the 24 strand endpoints. Let I be the fixed-point-free involution pairing endpoints carrying the same interaction index within one vertex, and let J_σ pair (A, i, s) with $(B, \sigma(i), s)$. The connected components of the two-regular graph defined by I and J_σ are its faces. Equivalently, if $c(\pi)$ counts permutation cycles, then

$$F(\sigma) = \tfrac{1}{2} c(I \circ J_\sigma). \tag{85}$$

Each alternating component contributes two permutation cycles to $I \circ J_\sigma$. Enumerating the 24 permutations against this exact formula gives the table below.

For the chosen ordering, the 24 possible $\sigma \in S_4$ distribute as follows:

Table 6: Exact Face Counts for Slot-Preserving Two-Vertex Gluings.

Faces	Pairings	Characterization
6	4	Cyclic rotations of $(1, 2, 3, 4)$
4	4	Reversals $(1, 4, 3, 2)$, $(2, 1, 4, 3)$, $(3, 2, 1, 4)$, $(4, 3, 2, 1)$
3	16	Remaining slot-preserving pairings

The sum of the multiplicities is $4 + 4 + 16 = 24 = 4!$. The identity pairing gives six faces, described by the boundary matrices in Eq. (83). The same six-face count applies to the other three cyclic pairings in this chosen slot convention. General propagator identifications can include permutations of strand slots and may lead to a different enumeration. The distinction between a specific dipole and the complete two-vertex pairing family should be retained in independent code verification.

L.6 Reproducibility Checklist

Table 7: Expected Exact Outputs for Independent Verification.

Test Object	Expected Output
Tetrahedral terminal graph	$\lvert V\rvert = 5$, $\lvert E\rvert = 4$; two enriched orientations
Constitutional projection	One labeled-star isomorphism class
Reflection action	$p^2 = e$, $pX_+ = X_-$
Unfilled six-cycle	$b_0 = 1, b_1 = 1, b_2 = 0$
Filled six-cycle	$b_0 = 1, b_1 = 0, b_2 = 0$
Two-vertex dipole	$(V, E, F) = (2, 4, 6)$
Dipole boundary product	$D_1 D_2 = 0$
Dipole homology	$(b_0, b_1, b_2) = (1, 0, 3)$
Constant $\mathbb{Z}_2$ field	$V_{\text{tet}} = 64$
Odd first-index field	$V_{\text{tet}}[t\chi_{100}] = 0$
Odd second-index field	$V_{\text{tet}}[t\chi_{010}] = 64t^4$
Gauge projection of odd modes	$P\chi_{100} = P\chi_{010} = 0$; $P\psi_{\text{odd}} = 0$
Relative gauge-invariant field	$P\psi_z = \psi_z$, $V_{\text{tet}}[\psi_z] = 128z^4$
Two-vertex gluing count	$F = 6, 4, 3$ with multiplicities 4, 4, 16

The tests establish the reproducibility of selected formal objects. The independent reproduction of physical response kernels, reaction rates, or chemical realization maps would require specifying additional empirical or computational models and validating their provenance.

M Verification Notes and Remaining Mathematical Obligations

This appendix records the conventions and regression targets for independent replication. A formal identity proved under stated assumptions, a finite computational identity, and a chemically validated mechanism are different kinds of evidence. In particular, test fields inserted into a polynomial should be stored separately from Wick-contraction graphs arising from a Gaussian measure.

M.1 Exact Finite Tests for the Tetrahedral Interaction

For $G = \mathbb{Z}_2$ use unnormalized counting measure and the lexicographic array order $(000), (001), (010), (011), (100), (101), (110), (111)$. The field tensor

$$\psi = (-1, 1, -2, 1, -2, 0, -1, -1) \tag{86}$$

has $V_{\text{tet}}[\psi] = -30$, by direct summation over the six group indices. The constant field gives $V_{\text{tet}}[1] = 64$. Quartic homogeneity implies that for either nonzero sign of λ, the unregulated action $\frac{\mu}{2}\sum\psi^2 + \lambda V_{\text{tet}}[\psi]$ diverges toward negative infinity along at least one real-field ray. A positive coercive regulator such as $\epsilon\sum\psi^6$ makes the finite-dimensional real integral convergent for $\epsilon > 0$.

Character orthogonality gives the exact identity

$$V_{\text{tet}}[t\chi_{ijk}] = 64t^4\,\mathbf{1}_{\{i=k\}}. \tag{87}$$

Thus the original χ_{100} direction is quartically flat, while χ_{010} yields $64t^4$. The full U-odd subspace is spanned by four characters, and the gauge projector annihilates this whole subspace. The relative odd field $z(\chi_{110} + \chi_{011})$ is gauge invariant while transforming oddly under the separately defined $U_{1,0}$ subgroup.

M.2 Wick Pairings and Graph Weights

For the real Gaussian covariance $\langle\psi_{abc}\psi_{a'b'c'}\rangle_0 = \delta_{aa'}\delta_{bb'}\delta_{cc'}$, Wick's theorem gives a sum over all perfect pairings of the $4n$ field factors in V_{tet}^n. For two quartic vertices there are $7!! = 105$ pairings. Among them, 24 join every leg across the two vertices, 72 have two cross-vertex contractions, and 9 use internal pairings only. The latter class gives disconnected two-vertex vacuum graphs. The 24 cross-only pairings split into face counts 6, 4, 3 with multiplicities 4, 4, 16 under slot-preserving propagators, as proved in Appendix L.5. For the identity propagator and counting measure, each closed face contributes exactly $|G|$; hence a six-face dipole has an index-summation factor $2^6 = 64$. Its numerical coincidence with $V_{\text{tet}}[1]$ follows from separate contractions and should be interpreted accordingly.

M.3 Grammar Semantics and Relabelling Tests

An implementation of the acid–base production should test both incidence edits and changes to mutable charge attributes on preserved oxygen vertices. The partial-label interface of Habel and Plump (2002) provides this possibility while preserving atomic identifiers. A strict label-preserving DPO rule would fail this test if charge labels are treated as immutable. The tetrahedral grammar produces attributed terminal graphs with $\tau = \pm 1$; an independent realization map produces embedded structures, whose classes are then considered modulo proper Euclidean motions. The determinant sign depends on the ordered ligand list (H, F, Cl, Br). CIP R/S assignment is a separately declared convention and remains outside the present finite-output assertions.

M.4 Notation and Interpretation Discipline

The following notation is preferred in reproductions. Here ϱ can denote the reflection action on a stereochemical state, while $p_\pm$ denote the two production schemas; these are conceptually distinct from the probability measure π_C. The symbol ν_q refers to a state lift, whereas the species-level stoichiometric increment is written Δn_q. The DPO interface K_q is a graph; the kinetic kernel K_Λ acts on fields; $K_{C,J}$ is the response kernel. These objects are distinguished by their domains and codomains, and implementations should use corresponding distinct identifiers.

Table 8: Verification Status and External Dependencies.

Status	Mathematical or Scientific Item
Exact finite identity	$V[\chi_{100}] = 0$; $V[\chi_{010}] = 64$; negative field witness $V[\psi] = -30$
Exact linear algebra	Four-dimensional U-odd subspace; $P\psi_{\text{odd}} = 0$
Exact Gaussian combinatorics	105 two-vertex pairings; 24 cross-only; six-face distribution 4, 4, 16
Conditional theorem	Equivariant DPO rewriting, generator covariance, readout pushforward, and invertible-flow stabilizer conservation
Convention-dependent	Action normalization, graph automorphism factors, Haar versus counting measure, CIP parity labeling
Independent physical validation	Gauge-invariant chemical observable, field-to-chemistry readout R_C, reaction rates, and functional capability tests

For chemical symmetry breaking, the experimental observations of Kondepudi et al. (1990) and Soai et al. (1995) are precedents. The paper's stronger inquiry concerns the conditional production and identification of new operational functions under symmetry-reducing generative dynamics. Models that successfully reproduce optical activity alone should therefore be evaluated as a component of, rather than a complete demonstration of, that proposed mechanism.

References

Jakob L. Andersen, Christoph Flamm, Daniel Merkle, and Peter F. Stadler. Inferring chemical reaction patterns using rule composition in graph grammars. *Journal of Systems Chemistry*, 4:4, 2013. doi: 10.1186/1759-2208-4-4.

Donna G. Blackmond. The origin of biological homochirality. *Cold Spring Harbor Perspectives in Biology*, 2(5): a002147, 2010. doi: 10.1101/cshperspect.a002147.

D. V. Boulatov. A model of three-dimensional lattice gravity. *Modern Physics Letters A*, 7:1629–1646, 1992. doi: 10.1142/S0217732392001324.

Pierre Boutillier, Mutaamba Maasha, Xing Li, Héctor F. Medina-Abarca, Jean Krivine, Jérôme Feret, Ioana Cristescu, Angus G. Forbes, and Walter Fontana. The Kappa platform for rule-based modeling. *Bioinformatics*, 34(13):i583–i592, 2018. doi: 10.1093/bioinformatics/bty272.

Noam Chomsky. Three factors in language design. *Linguistic Inquiry*, 36(1):1–22, 2005. doi: 10.1162/0024389052993655.

Frederick Charles Frank. On spontaneous asymmetric synthesis. *Biochimica et Biophysica Acta*, 11:459–463, 1953. doi: 10.1016/0006-3002(53)90082-1.

Ella M. Gale, Leo Lobski, and Fabio Zanasi. A categorical model for organic chemistry. *Theoretical Computer Science*, 1032:115084, 2025. doi: 10.1016/j.tcs.2025.115084. Open-access categorical treatment of organic chemical rewriting.

Dennis Görlich and Peter Dittrich. Molecular codes in biological and chemical reaction networks. *PLOS ONE*, 8(1):e54694, 2013. doi: 10.1371/journal.pone.0054694.

Annegret Habel and Detlef Plump. Relabelling in graph transformation. In *Graph Transformation (ICGT 2002)*, volume 2505 of *Lecture Notes in Computer Science*, pages 135–147, 2002. doi: 10.1007/3-540-45832-8_12.

Yiming Huang, Yujie Zeng, Vijay Prakash Dwivedi, Simone Foti, Jianmin Wang, Jure Leskovec, and Tolga Birdal. Higher-order molecular grammars for generative and foundation models in chemistry, 2026. URL `https://arxiv.org/abs/2610.02186`.

International Union of Pure and Applied Chemistry. Enantiomer. IUPAC Compendium of Chemical Terminology (Gold Book), 2025a. URL `https://goldbook.iupac.org/terms/view/E02069`.

International Union of Pure and Applied Chemistry. Stereoisomers. IUPAC Compendium of Chemical Terminology (Gold Book), 2025b. URL `https://goldbook.iupac.org/terms/view/S05984`.

Hiroshi Kajino. Molecular hypergraph grammar with its application to molecular optimization. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pages 3183–3191, 2019. URL `https://proceedings.mlr.press/v97/kajino19a.html`.

Reza Karimpour, Matteo Gori, and Alexandre Tkatchenko. Quantum field approaches to chemical systems. *Journal of the American Chemical Society*, 148(37):39379–39404, 2026. doi: 10.1021/jacs.6c02077. Review of quantum-field approaches to molecular interactions.

Artemy Kolchinsky and David H. Wolpert. Semantic information, autonomous agency and non-equilibrium statistical physics. *Interface Focus*, 8(6):20180041, 2018. doi: 10.1098/rsfs.2018.0041.

Dilip K. Kondepudi, Rebecca J. Kaufman, and Nolini Singh. Chiral symmetry breaking in sodium chlorate crystallization. *Science*, 250(4983):975–976, 1990. doi: 10.1126/science.250.4983.975.

National Center for Biotechnology Information. Pubchem compound summary: cis-1,2-dichloroethylene, cid 643833. PubChem, 2026a. URL `https://pubchem.ncbi.nlm.nih.gov/compound/643833`.

National Center for Biotechnology Information. Pubchem compound summary: trans-1,2-dichloroethylene, cid 638186. PubChem, 2026b. URL `https://pubchem.ncbi.nlm.nih.gov/compound/638186`.

Daniele Oriti. Group field theory and loop quantum gravity, 2014. URL `https://arxiv.org/abs/1408.7112`.

Daniele Oriti, James P. Ryan, and Johannes Thürigen. Group field theories for all loop quantum gravity, 2014. URL `https://arxiv.org/abs/1409.3150`.

Riccardo Rao and Massimiliano Esposito. Nonequilibrium thermodynamics of chemical reaction networks: Wisdom from stochastic thermodynamics. *Physical Review X*, 6:041064, 2016. doi: 10.1103/PhysRevX.6.041064.

Patrick Reiser, Marlen Neubert, André Eberhard, Luca Torresi, Chen Zhou, et al. Graph neural networks for materials science and chemistry. *Communications Materials*, 3:93, 2022. doi: 10.1038/s43246-022-00315-6.

Guillermo Restrepo. Higher order structures in chemistry: hypergraphs reshape the molecule and the reaction. *Digital Discovery*, 5:973–982, 2026. doi: 10.1039/D5DD00533G.

Eric Smith, Harrison B. Smith, and Jakob Lykke Andersen. Rules, hypergraphs, and probabilities: The three-level analysis of chemical reaction systems and other stochastic stoichiometric population processes. *PLOS Complex Systems*, 2024. doi: 10.1371/journal.pcsy.0000022.

Kenso Soai, Takanori Shibata, Hiroshi Morioka, and Kaori Choji. Asymmetric autocatalysis and amplification of enantiomeric excess of a chiral molecule. *Nature*, 378:767–768, 1995. doi: 10.1038/378767a0.

Alessandro Sorrenti, Jorge Leira-Iglesias, Albert J. Markvoort, Tom F. A. de Greef, and Thomas M. Hermans. Non-equilibrium supramolecular polymerization. *Chemical Society Reviews*, 46:5476–5490, 2017. doi: 10.1039/C7CS00121E.

P. Kyle Stanford. Underdetermination of scientific theory. The Stanford Encyclopedia of Philosophy, 2023. URL https://plato.stanford.edu/entries/scientific-underdetermination/. Substantive revision 4 April 2023.

Marius Wehner and Frank Würthner. Supramolecular polymerization through kinetic pathway control and living chain growth. *Nature Reviews Chemistry*, 4:38–53, 2020. doi: 10.1038/s41570-019-0153-8.