

# Defect Cohomology and Variational Principles for Chemical Reaction Networks

A Comprehensive Formally Verified Framework

Paolo Vella

paolovella1993@gmail.com

January 24, 2026

## Abstract

We present a comprehensive, formally verified framework for chemical reaction network theory (CRNT). Starting from the Onsager–Rayleigh dissipation principle, we develop the variational characterization of steady-state fluxes via the functional  $F(J) = \frac{1}{2}\langle J, J \rangle_{W^{-1}} - \langle \omega, J \rangle$ . We systematically extend to: the deficiency zero and one theorems; higher deficiency networks ( $\delta \geq 2$ ) via the Deficiency-Two Algorithm and Species-Reaction graphs; **cohomological deficiency theory** establishing  $\delta = \dim(H^1)$  for the CRN chain complex; multi-stability and bifurcation theory; oscillations via Hopf bifurcation and the Routh–Hurwitz criterion; reaction-diffusion systems with Turing pattern formation; stochastic CRNs via the Chemical Master Equation; and feedback control theory including antithetic integral control.

The entire framework—comprising **7211 lines** of Lean 4 code with **210 theorems** and **zero unproven assertions or custom axioms**—has been formally verified using the Mathlib library. We demonstrate the theory on examples ranging from the simple  $n$ -cycle to the complete TCA (Krebs) cycle with 16 species and 10 reactions. This constitutes the first machine-checked formalization of comprehensive chemical reaction network theory.

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	Main Contributions . . . . .	3
1.2	Scope and Regimes . . . . .	4
1.3	Related Work . . . . .	4
<b>2</b>	<b>Graph Laplacian and Hodge Decomposition</b>	<b>4</b>
<b>3</b>	<b>The Onsager–Rayleigh Functional</b>	<b>4</b>
<b>4</b>	<b>Optimality and Uniqueness Theorems</b>	<b>4</b>
<b>5</b>	<b>Chemical Reaction Network Theory</b>	<b>5</b>
5.1	Species and Complexes . . . . .	5
5.2	Deficiency . . . . .	5
5.3	Mass-Action Kinetics . . . . .	5
5.4	Deficiency Zero Theorem . . . . .	5
<b>6</b>	<b>Deficiency One Theorem</b>	<b>5</b>
<b>7</b>	<b>Higher Deficiency Networks</b>	<b>5</b>
7.1	Species-Reaction Graph . . . . .	6
7.2	Deficiency Two Algorithm . . . . .	6

<b>8 Cohomological Deficiency Theory</b>	<b>6</b>
8.1 The CRN Chain Complex . . . . .	6
8.2 Cycle and Coboundary Spaces . . . . .	6
8.3 The Main Theorem . . . . .	6
8.4 Physical Interpretation . . . . .	7
8.5 Connection to Onsager–Rayleigh . . . . .	7
8.6 Examples . . . . .	7
<b>9 Persistence and Permanence</b>	<b>7</b>
9.1 Omega-Limit Sets . . . . .	7
9.2 Siphons . . . . .	8
<b>10 Multistability</b>	<b>8</b>
<b>11 Oscillations</b>	<b>8</b>
11.1 Routh–Hurwitz Criterion . . . . .	8
<b>12 Reaction-Diffusion Systems</b>	<b>8</b>
12.1 Turing Instability . . . . .	9
12.2 Pattern Formation . . . . .	9
<b>13 Stochastic Chemical Reaction Networks</b>	<b>9</b>
<b>14 Control Theory</b>	<b>9</b>
<b>15 Examples</b>	<b>10</b>
15.1 The $n$ -Cycle . . . . .	10
15.2 Michaelis–Menten Enzyme Kinetics . . . . .	10
15.3 Glycolysis Pathway . . . . .	10
15.4 TCA Cycle . . . . .	10
<b>16 Discussion</b>	<b>10</b>
16.1 Summary . . . . .	10
16.2 Formal Verification . . . . .	11
16.3 Future Work . . . . .	11
<b>A Lean 4 Formalization</b>	<b>12</b>
A.1 Build Instructions . . . . .	12
A.2 File Structure . . . . .	12
A.3 Correspondence Table . . . . .	12
A.4 Key Assumptions . . . . .	12

# 1 Introduction

Chemical reaction networks (CRNs) provide the mathematical foundation for systems biology, from metabolic engineering to synthetic biology. The fundamental questions—existence, uniqueness, and stability of steady states—have been studied extensively since the pioneering work of Feinberg, Horn, and Jackson [5, 8, 6].

This paper presents a *comprehensive* and *formally verified* treatment of CRNT. By “comprehensive,” we mean coverage of:

- Deficiency theory (zero, one, and higher)
- Cohomological deficiency theory ( $\delta = \dim H^1$ )
- Dynamical behavior (persistence, multistability, oscillations)
- Spatial extension (reaction-diffusion, Turing patterns)
- Stochastic formulation (Chemical Master Equation)
- Control theory (feedback, robustness)

By “formally verified,” we mean that every definition, theorem, and proof has been machine-checked in the Lean 4 proof assistant, yielding 7211 lines of code with 210 theorems and zero unproven assertions.

## 1.1 Main Contributions

- (1) **Variational Framework** (Sections 2–4): Onsager–Rayleigh functional, Hodge decomposition, optimality and uniqueness.
- (2) **Deficiency Theory** (Sections 5–7): Complete treatment of  $\delta = 0$ ,  $\delta = 1$ , and  $\delta \geq 2$  including the Deficiency-Two Algorithm and concordance.
- (3) **Cohomological Deficiency** (Section 8): Chain complex formulation,  $\delta = \dim(H^1)$ , obstruction theory, variational duality.
- (4) **Persistence and Permanence** (Section 9): Omega-limit characterization, siphon theory, global attractors.
- (5) **Multistability** (Section 10): Bifurcation conditions, sign conditions for multiple equilibria, injectivity.
- (6) **Oscillations** (Section 11): Hopf bifurcation theorem, limit cycles, Routh–Hurwitz stability criterion.
- (7) **Reaction-Diffusion** (Section 12): Turing instability, pattern formation, traveling waves.
- (8) **Stochastic CRNs** (Section 13): Chemical Master Equation, product-form distributions, fluctuation-dissipation.
- (9) **Control Theory** (Section 14): Antithetic integral feedback, robust perfect adaptation.
- (10) **Major Examples** (Section 15):  $n$ -cycle, Michaelis–Menten, glycolysis, and the complete TCA cycle.
- (11) **Formal Verification** (Appendix A): Complete Lean 4 formalization with correspondence table.

## 1.2 Scope and Regimes

The Onsager–Rayleigh variational principle governs linear response near detailed balance. The CRNT deficiency theory provides exact structural results for nonlinear mass-action kinetics. These perspectives are complementary: the variational framework identifies the optimal flux structure (projection onto  $\ker B$ ), while deficiency theory guarantees existence and uniqueness of the steady state. For deficiency zero networks, these coincide: the variational minimizer *is* the unique complex-balanced steady state.

## 1.3 Related Work

The connection between network thermodynamics and graph theory has a long history, from Kirchhoff's laws [9] to modern stochastic thermodynamics [16]. Our variational approach builds on Maas [11] and Mielke [12]. The CRNT foundations were laid by Horn, Jackson, and Feinberg [8, 6], with recent advances by Anderson, Craciun, and Kurtz [1].

## 2 Graph Laplacian and Hodge Decomposition

Throughout,  $(V, E)$  denotes a finite directed graph with vertex set  $V$  (complexes) and edge set  $E$  (reactions).

**Definition 2.1** (Incidence Matrix). The incidence matrix  $B \in \mathbb{R}^{V \times E}$  has  $B_{ve} = +1$  if  $e$  enters  $v$ ,  $B_{ve} = -1$  if  $e$  leaves  $v$ , and 0 otherwise.

**Definition 2.2** (Weighted Graph Laplacian). For positive edge weights  $w : E \rightarrow \mathbb{R}_{>0}$ , the weighted graph Laplacian is  $L = BWB^\top$ .

**Theorem 2.3** (Kernel of Laplacian). *For a connected graph,  $\ker(L) = \mathbb{R} \cdot \mathbf{1}$ .*

**Theorem 2.4** (Hodge Decomposition). *Every edge function  $\omega : E \rightarrow \mathbb{R}$  decomposes uniquely as  $\omega = \omega_{\text{harm}} + \omega_{\text{exact}}$ , where  $\omega_{\text{harm}} \in \ker(BW)$  and  $\omega_{\text{exact}} \in \text{im}(B^\top)$  are  $W^{-1}$ -orthogonal.*

**Definition 2.5** (Laplacian Inverse). The Laplacian inverse  $L^+$  satisfies  $LL^+L = L$ ,  $L^+\mathbf{1} = 0$ , and  $(L^+)^\top = L^+$ .

## 3 The Onsager–Rayleigh Functional

**Definition 3.1** (Onsager–Rayleigh Functional).

$$F(J) = \frac{1}{2}\langle J, J \rangle_{W^{-1}} - \langle \omega, J \rangle = \frac{1}{2} \sum_{e \in E} \frac{J_e^2}{w_e} - \sum_{e \in E} \omega_e J_e.$$

Physical interpretation:  $J_e$  is flux,  $w_e$  is conductance,  $\omega_e$  is driving force,  $\frac{1}{2}\langle J, J \rangle_{W^{-1}}$  is dissipation,  $\langle \omega, J \rangle$  is power input.

**Definition 3.2** (Optimal Flux).  $J^* = W\pi(\omega) = W(\omega - B^\top L^+ BW\omega)$ .

**Proposition 3.3.**  $BJ^* = 0$ , i.e.,  $J^* \in \ker(B)$ .

## 4 Optimality and Uniqueness Theorems

**Theorem 4.1** (KKT Stationarity). *There exists  $\lambda : V \rightarrow \mathbb{R}$  such that  $J_e^*/w_e = \omega_e + \sum_v B_{ve}\lambda_v$ .*

**Theorem 4.2** (Optimality). *For any  $J \in \ker(B)$ ,  $F(J^*) \leq F(J)$ .*

**Corollary 4.3** (Uniqueness). *If  $F(J) = F(J^*)$  for  $J \in \ker(B)$ , then  $J = J^*$ .*

**Theorem 4.4** (Quadratic Expansion). *For any  $h \in \ker(B)$ ,  $F(J^* + h) - F(J^*) = \frac{1}{2}\langle h, h \rangle_{W^{-1}}$ .*

**Corollary 4.5** (Lyapunov Characterization).  *$V(J) = F(J) - F(J^*)$  satisfies  $V(J) \geq 0$  with equality iff  $J = J^*$ .*

## 5 Chemical Reaction Network Theory

### 5.1 Species and Complexes

**Definition 5.1** (Chemical Reaction Network). A CRN consists of species  $\mathcal{S}$ , complexes  $V$ , reactions  $E$ , incidence matrix  $B$ , and complex composition matrix  $Y \in \mathbb{R}^{\mathcal{S} \times V}$ .

**Definition 5.2** (Stoichiometric Matrix).  $N = YB \in \mathbb{R}^{\mathcal{S} \times E}$  gives net species changes per reaction.

### 5.2 Deficiency

**Definition 5.3** (CRNT Deficiency).  $\delta = n - \ell - \text{rank}(N)$ , where  $n = |V|$ ,  $\ell$  = linkage classes.

*Remark 5.4* (Graph vs. CRNT Deficiency). Graph deficiency  $|V| - \ell - \text{rank}(B) = 0$  for connected graphs. CRNT deficiency uses  $\text{rank}(N) = \text{rank}(YB) \leq \text{rank}(B)$ , so  $\delta \geq 0$ .

### 5.3 Mass-Action Kinetics

**Definition 5.5** (Mass-Action Rate).  $v_e(c) = k_e \prod_s c_s^{Y_{sy}}$ .

**Definition 5.6** (Affinity).  $A_e(c) = \ln(k_e^+ / k_e^-) - \sum_s N_{se} \ln c_s$ .

**Theorem 5.7** (Cycle Affinity Independence). *For a stoichiometric cycle,  $A_{\text{cycle}} = \sum_i \ln(k_{e_i}^+ / k_{e_i}^-)$  is concentration-independent.*

### 5.4 Deficiency Zero Theorem

**Theorem 5.8** (Deficiency Zero Equilibrium Existence). *For  $\delta = 0$  with weak reversibility, there exists a positive complex-balanced equilibrium, unique in each stoichiometric compatibility class.*

## 6 Deficiency One Theorem

**Definition 6.1** (Linkage Class Deficiency). For linkage class  $\ell_i$ , the local deficiency is  $\delta_i = n_i - 1 - \text{rank}(N_i)$ .

**Theorem 6.2** (Deficiency One Existence). *For a mass-action CRN with  $\delta = 1$ , weak reversibility,  $\delta_i \leq 1$  for all  $i$ , and  $\sum_i \delta_i = 1$ , there exists a positive steady state in each compatibility class.*

**Theorem 6.3** (Deficiency One Uniqueness). *Under the Deficiency One Algorithm conditions, the positive steady state is unique.*

## 7 Higher Deficiency Networks

For networks with  $\delta \geq 2$ , we employ advanced structural tools.

## 7.1 Species-Reaction Graph

**Definition 7.1** (SR-Graph). The species-reaction graph has vertices  $\mathcal{S} \cup \mathcal{E}$  with edges connecting species to reactions that produce or consume them.

**Definition 7.2** (Concordance). A CRN is *concordant* if there is no sign pattern  $\sigma \in \{-1, 0, +1\}^E$  with  $N\sigma = 0$  and  $\sigma$  sign-compatible with stoichiometry.

**Theorem 7.3** (Concordance and Injectivity). *A concordant CRN has at most one equilibrium in each stoichiometric compatibility class.*

## 7.2 Deficiency Two Algorithm

**Definition 7.4** (D2A Conditions). A network satisfies D2A conditions if certain linear inequalities on the stoichiometric coefficients are satisfied.

**Theorem 7.5** (D2A Existence). *Under D2A conditions with weak reversibility, existence of positive equilibria is guaranteed despite  $\delta \geq 2$ .*

# 8 Cohomological Deficiency Theory

We now present a cohomological interpretation of deficiency that reveals its deep mathematical structure.

## 8.1 The CRN Chain Complex

**Definition 8.1** (CRN Chain Complex). A chemical reaction network induces a chain complex:

$$0 \longrightarrow \mathbb{R}^E \xrightarrow{B^\top} \mathbb{R}^V \xrightarrow{Y} \mathbb{R}^{\mathcal{S}} \longrightarrow 0$$

where  $B^\top$  is the transpose of the incidence matrix and  $Y$  is the complex composition matrix.

**Proposition 8.2** (Composition).  $Y \circ B^\top = N^\top$ , where  $N = YB$  is the stoichiometric matrix.

## 8.2 Cycle and Coboundary Spaces

**Definition 8.3** (CycleSpace). The *CycleSpace* is  $\ker(Y) \subseteq \mathbb{R}^V$ —vectors in the complex space that are invisible to species.

**Definition 8.4** (CoboundarySpace). The *CoboundarySpace* is  $\text{im}(B^\top) \subseteq \mathbb{R}^V$ —vectors that arise from flux distributions.

**Definition 8.5** (DefectSpace). The *DefectSpace* is  $\ker(Y) \cap \text{im}(B^\top)$ —the first cohomology group  $H^1$  of the chain complex.

## 8.3 The Main Theorem

**Theorem 8.6** (Cohomological Deficiency). *The classical CRNT deficiency equals the dimension of the DefectSpace:*

$$\delta = n - \ell - \text{rank}(N) = \dim(\ker(Y) \cap \text{im}(B^\top)) = \dim(H^1).$$

**Corollary 8.7** (Exactness Characterization). *The chain complex is exact at  $\mathbb{R}^V$  if and only if  $\delta = 0$ .*

## 8.4 Physical Interpretation

**Definition 8.8** (Obstruction). An element  $c \in \text{DefectSpace}$  is an *obstruction* to the existence of a steady state: it represents a complex-space vector that arises from fluxes but is invisible to the species dynamics.

**Theorem 8.9** (Obstruction Theory). *For  $\delta > 0$ , there exist nonzero vectors  $c \in \ker(Y) \cap \text{im}(B^\top)$ . These correspond to “hidden cycles” in the reaction network that create additional constraints on steady-state existence.*

## 8.5 Connection to Onsager–Rayleigh

**Theorem 8.10** (Variational Duality). *Let  $J^*$  be the optimal flux from the Onsager–Rayleigh functional. The Lagrange multipliers  $\mu : V \rightarrow \mathbb{R}$  from the KKT conditions satisfy:*

- (i)  $\mu \in \text{CoboundarySpace}^\perp$  (orthogonal to image of  $B^\top$ )
- (ii) For  $\delta = 0$ :  $\mu$  is uniquely determined by the stoichiometry
- (iii) For  $\delta > 0$ : The DefectSpace creates a  $\delta$ -dimensional family of valid multipliers

## 8.6 Examples

**Example 8.11** (Triangle Network). For the 3-cycle  $A \rightarrow B \rightarrow C \rightarrow A$ :

- $n = 3, \ell = 1, \text{rank}(N) = 2$
- $\delta = 3 - 1 - 2 = 0$
- DefectSpace =  $\{0\}$  (exact)

**Example 8.12** (Deficiency One Network). For  $A \rightarrow 2A \rightarrow 3A$  over single species  $A$ :

- $n = 3, \ell = 1, \text{rank}(N) = 1$
- $\delta = 3 - 1 - 1 = 1$
- DefectSpace =  $\text{span}\{(-1, 2, -1)\}$  (1-dimensional)

## 9 Persistence and Permanence

**Definition 9.1** (Persistence). A CRN is *persistent* if  $\liminf_{t \rightarrow \infty} c_s(t) > 0$  for all species  $s$  from positive initial conditions.

**Definition 9.2** (Permanence). A CRN is *permanent* if there exist  $0 < m < M$  bounding all trajectories uniformly.

**Theorem 9.3** (Deficiency Zero Persistence). *A mass-action CRN with  $\delta = 0$ , weak reversibility, and single linkage class is persistent.*

### 9.1 Omega-Limit Sets

**Definition 9.4** (Omega-Limit Set).  $\omega(c_0) = \overline{\bigcap_{T>0} \{c(t) : t > T\}}$ .

**Theorem 9.5** (Omega-Limit Characterization). *For persistent, bounded trajectories:  $\omega(c_0) \neq \emptyset$ ,  $\omega(c_0) \subseteq \mathbb{R}_{>0}^S$ , and for  $\delta = 0$  with weak reversibility,  $\omega(c_0)$  consists of equilibria.*

## 9.2 Siphons

**Definition 9.6** (Siphon).  $Z \subseteq \mathcal{S}$  is a siphon if every reaction producing a species in  $Z$  also consumes one.

**Theorem 9.7** (Global Attractor). *For  $\delta = 0$ , weak reversibility, and single linkage class, each compatibility class has a unique global attractor.*

## 10 Multistability

**Definition 10.1** (Multiple Equilibria). A CRN exhibits *multistability* if there exist multiple positive equilibria in a single stoichiometric compatibility class.

**Theorem 10.2** (Sign Condition for Injectivity). *If the Jacobian  $\partial f / \partial c$  has a sign pattern precluding positive real eigenvalues, the system is injective (at most one equilibrium).*

**Definition 10.3** (Saddle-Node Bifurcation). A saddle-node bifurcation occurs when two equilibria collide and annihilate as a parameter varies.

**Theorem 10.4** (Bifurcation Conditions). *Multistability requires: (i)  $\delta \geq 1$ , (ii) appropriate sign structure in the Jacobian, (iii) sufficient nonlinearity.*

## 11 Oscillations

**Definition 11.1** (Hopf Bifurcation). A Hopf bifurcation occurs when a pair of complex conjugate eigenvalues of the Jacobian crosses the imaginary axis.

**Theorem 11.2** (Hopf Bifurcation Theorem). *If at parameter  $\mu = \mu_c$ : (i) the Jacobian has eigenvalues  $\pm i\omega$ , (ii) the transversality condition  $\frac{d}{d\mu} \text{Re}(\lambda)|_{\mu_c} \neq 0$  holds, then a family of periodic orbits bifurcates from the equilibrium.*

### 11.1 Routh–Hurwitz Criterion

**Theorem 11.3** (Routh–Hurwitz). *The characteristic polynomial  $p(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_n$  has all roots with negative real parts iff the Hurwitz determinants  $H_1, \dots, H_n$  are all positive.*

**Definition 11.4** (Limit Cycle). A limit cycle is an isolated periodic orbit.

**Theorem 11.5** (Existence of Limit Cycles). *Supercritical Hopf bifurcations produce stable limit cycles; subcritical produce unstable ones.*

## 12 Reaction-Diffusion Systems

**Definition 12.1** (Reaction-Diffusion Equation).

$$\frac{\partial c}{\partial t} = D\nabla^2 c + f(c),$$

where  $D = \text{diag}(D_s)$  contains diffusion coefficients and  $f(c)$  is the reaction term.

## 12.1 Turing Instability

**Definition 12.2** (Turing Instability). A homogeneous steady state  $c^*$  is Turing unstable if it is stable to spatially uniform perturbations but unstable to spatially nonuniform perturbations.

**Theorem 12.3** (Turing Conditions). *For a two-species system with Jacobian  $J$  at  $c^*$ :*

$$\text{tr}(J) < 0 \text{ and } \det(J) > 0 \text{ (homogeneous stability),}$$

$$D_2 J_{11} + D_1 J_{22} > 0,$$

$$(D_2 J_{11} + D_1 J_{22})^2 > 4D_1 D_2 \det(J).$$

## 12.2 Pattern Formation

**Definition 12.4** (Critical Wavenumber). The critical wavenumber  $k_c$  satisfies  $\det(J - k_c^2 D) = 0$  at the Turing bifurcation.

**Theorem 12.5** (Traveling Waves). *Under appropriate conditions, reaction-diffusion systems support traveling wave solutions  $c(x, t) = C(x - vt)$  with wave speed  $v$ .*

## 13 Stochastic Chemical Reaction Networks

**Definition 13.1** (State Space). The state is  $n = (n_s)_{s \in S} \in \mathbb{Z}_{\geq 0}^S$ , where  $n_s$  is the molecule count.

**Definition 13.2** (Propensity Function).  $a_e(n) = k_e \prod_s \binom{n_s}{\nu_{e,s}}$ .

**Definition 13.3** (Chemical Master Equation).

$$\frac{dP(n, t)}{dt} = \sum_{e \in E} [a_e(n - \nu_e) P(n - \nu_e, t) - a_e(n) P(n, t)].$$

**Definition 13.4** (Product-Form Distribution).  $\pi(n) = \prod_{s \in S} \frac{c_s^{n_s}}{n_s!} e^{-c_s}$ .

**Theorem 13.5** (Product-Form Stationarity). *For  $\delta = 0$  with weak reversibility, the product-form distribution with parameter  $c^*$  (deterministic equilibrium) is stationary for the CME.*

**Theorem 13.6** (Deterministic Limit). *As volume  $V \rightarrow \infty$  with  $N^V(0)/V \rightarrow c_0$ ,  $N^V(t)/V \xrightarrow{P} c(t)$  (mass-action ODE solution).*

**Theorem 13.7** (Fluctuation-Dissipation).  $\text{Var}(N_s/V) = c_s^*/V + O(1/V^2)$ , connecting to Onsager-Rayleigh structure.

## 14 Control Theory

**Definition 14.1** (Antithetic Integral Feedback). A control motif with species  $Z_1, Z_2$  satisfying:

$$\dot{Z}_1 = \mu - \eta Z_1 Z_2, \tag{1}$$

$$\dot{Z}_2 = \theta X - \eta Z_1 Z_2, \tag{2}$$

where  $X$  is the controlled output and  $\mu, \theta, \eta$  are parameters.

**Theorem 14.2** (Robust Perfect Adaptation). *The antithetic motif achieves  $X^* = \mu/\theta$  at steady state, independent of other system parameters.*

**Definition 14.3** (Robustness). A property is *robust* if it is maintained under parameter perturbations.

**Theorem 14.4** (Structural Robustness). *Integral feedback controllers provide robust perfect adaptation: the steady-state output depends only on the controller parameters, not on the plant.*

**Definition 14.5** (Antithetic Saturation). When  $Z_1$  or  $Z_2$  saturates, the controller loses perfect adaptation but maintains bounded tracking error.

## 15 Examples

### 15.1 The $n$ -Cycle

For vertices  $0, 1, \dots, n - 1$  with edges  $i \rightarrow (i + 1) \pmod{n}$ :

**Theorem 15.1** (Kirchhoff's Theorem).  $\ker(B) = \mathbb{R} \cdot \mathbf{1}$  and  $J^* = \bar{\omega} \cdot \mathbf{1}$  where  $\bar{\omega} = \frac{1}{n} \sum_i \omega_i$ .

### 15.2 Michaelis–Menten Enzyme Kinetics

The mechanism  $E + S \rightleftharpoons ES \rightarrow E + P$  has  $\delta = 0$ .

**Theorem 15.2** (Michaelis–Menten Equation). Under QSSA:  $v = V_{\max} \cdot [S] / (K_m + [S])$  with  $K_m = (k_2 + k_3)/k_1$ .

### 15.3 Glycolysis Pathway

8-species simplified pathway with deficiency zero, demonstrating ATP/ADP conservation.

### 15.4 TCA Cycle

**Example 15.3** (Krebs Cycle). The TCA cycle formalization includes:

- 16 species (acetyl-CoA, citrate, isocitrate,  $\alpha$ -ketoglutarate, succinyl-CoA, succinate, fumarate, malate, oxaloacetate, plus cofactors)
- 10 reactions
- Conservation laws for CoA,  $\text{NAD}^+/\text{NADH}$ ,  $\text{FAD}/\text{FADH}_2$

**Proposition 15.4** (TCA Deficiency). The TCA cycle network has  $\delta = 0$  when all reactions are reversible.

## 16 Discussion

### 16.1 Summary

We have presented a comprehensive framework for chemical reaction network theory:

1. **Variational**: Onsager–Rayleigh functional characterizes optimal fluxes
2. **Algebraic**: Deficiency theory ( $\delta = 0, 1, \geq 2$ ) determines equilibrium structure
3. **Cohomological**:  $\delta = \dim(H^1)$  reveals deep mathematical structure
4. **Dynamical**: Persistence, multistability, oscillations characterize long-term behavior
5. **Spatial**: Reaction-diffusion enables pattern formation
6. **Stochastic**: CME extends to finite-copy-number regimes
7. **Control**: Feedback enables robust regulation

## 16.2 Formal Verification

The entire framework has been formalized in Lean 4:

Metric	Value
Lines of code	7211
Theorems	210
<code>sorry</code> (unproven)	0
Custom axioms	0
Files	24

This provides: (1) correctness guarantees, (2) explicit assumptions, (3) reproducibility via `lake build`, (4) foundation for extensions.

## 16.3 Future Work

1. **Hybrid systems**: Discrete-continuous CRNs
2. **Parameter inference**: Bayesian methods for rate constants
3. **Synthetic biology**: DNA strand displacement, genetic circuits
4. **Enzyme networks**: Integration with CCR framework
5. **Quantum effects**: Coherence in photosynthetic reaction centers

## Acknowledgments

The author thanks the Lean and Mathlib communities for developing the tools that made this formalization possible.

## References

- [1] D. F. Anderson, G. Craciun, and T. G. Kurtz. Product-form stationary distributions for deficiency zero chemical reaction networks. *Bull. Math. Biol.*, 72:1947–1970, 2010.
- [2] D. F. Anderson. A proof of the global attractor conjecture in the single linkage class case. *SIAM J. Appl. Math.*, 71:1487–1508, 2011.
- [3] D. Angeli, P. De Leenheer, and E. D. Sontag. A Petri net approach to persistence in chemical reaction networks. *Math. Biosci.*, 210:598–618, 2007.
- [4] M. Bruna, S. J. Chapman, and M. J. Smith. Model reduction for slow-fast stochastic systems. *SIAM J. Appl. Math.*, 74:1–31, 2014.
- [5] M. Feinberg. Complex balancing in general kinetic systems. *Arch. Rational Mech. Anal.*, 49:187–194, 1972.
- [6] M. Feinberg. Chemical reaction network structure and the stability of complex isothermal reactors—I. *Chem. Eng. Sci.*, 42(10):2229–2268, 1987.
- [7] M. Feinberg. The existence and uniqueness of steady states for a class of chemical reaction networks. *Arch. Rational Mech. Anal.*, 132:311–370, 1995.
- [8] F. Horn and R. Jackson. General mass action kinetics. *Arch. Rational Mech. Anal.*, 47:81–116, 1972.

- [9] G. Kirchhoff. Ueber die Auflösung der Gleichungen. *Ann. Phys.*, 148(12):497–508, 1847.
- [10] T. G. Kurtz. The relationship between stochastic and deterministic models for chemical reactions. *J. Chem. Phys.*, 57:2976–2978, 1972.
- [11] J. Maas. Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.*, 261(8):2250–2292, 2011.
- [12] A. Mielke. A gradient structure for reaction-diffusion systems. *Nonlinearity*, 24(4):1329, 2011.
- [13] L. de Moura and S. Ullrich. The Lean 4 theorem prover. In *CADE*, 2021.
- [14] J. D. Murray. *Mathematical Biology II: Spatial Models and Biomedical Applications*. Springer, 2003.
- [15] L. Onsager. Reciprocal relations in irreversible processes. *Phys. Rev.*, 37(4):405, 1931.
- [16] U. Seifert. Stochastic thermodynamics. *Rep. Prog. Phys.*, 75(12):126001, 2012.
- [17] A. M. Turing. The chemical basis of morphogenesis. *Phil. Trans. R. Soc. B*, 237:37–72, 1952.

## A Lean 4 Formalization

Source code: <https://github.com/paolovella/DefectCRN>  
DOI: 10.5281/zenodo.18363235

### A.1 Build Instructions

```
git clone https://github.com/paolovella/DefectCRN.git
cd DefectCRN
git checkout v4.0.0
lake exe cache get
lake build
```

### A.2 File Structure

### A.3 Correspondence Table

### A.4 Key Assumptions

Explicit in Lean:

1. **Finite types:**  $V, E, S$  are Fintype
2. **Positive weights:**  $hw : \forall e, w e > 0$
3. **Incidence property:**  $hBcol : \forall e, \sum_v B v e = 0$
4. **Weak reversibility:** Each linkage class is strongly connected
5. **Positive concentrations:**  $hpos : isPositive c$

File	Description	Lines	Thms
Basic.lean	Onsager–Rayleigh, Hodge	852	38
CRNT.lean	Species, deficiency, mass-action	512	10
DeficiencyOne.lean	$\delta = 1$ theorem	367	4
Persistence.lean	Persistence, permanence	312	8
Stochastic.lean	CME, product-form	241	6
HigherDeficiency.lean	$\delta \geq 2$ , D2A, SR-graph	191	4
Multistability.lean	Bifurcations, sign conditions	238	7
Oscillations.lean	Hopf, limit cycles, Routh–Hurwitz	273	5
ReactionDiffusion.lean	Turing, traveling waves	270	3
Control.lean	Antithetic feedback, robustness	284	7
Cohomology/ChainComplex.lean	CRN chain complex	254	6
Cohomology/Cycles.lean	Cycle/coboundary spaces	307	23
Cohomology/Deficiency.lean	Main theorem: $\delta = \dim(H^1)$	286	15
Cohomology/Obstruction.lean	Physical interpretation	241	10
Cohomology/VariationalDuality.lean	Onsager–Rayleigh connection	258	9
Cohomology/Examples/*.lean	Triangle, MM, Def. One	601	21
Examples/Triangle.lean	3-cycle	319	11
Examples/Cycle.lean	$n$ -cycle	438	8
Examples/MichaelisMenten.lean	Enzyme kinetics	407	11
Examples/Glycolysis.lean	Metabolic pathway	237	2
Examples/TCA.lean	Krebs cycle (16 species)	273	2
<b>Total</b>		<b>7211</b>	<b>210</b>

Table 1: Formalization statistics. Version 4.0.0.

Paper	Lean Theorem	File
Thm. 2.4	hodge_decomp	Basic.lean
Thm. 4.2	onsager_rayleigh_optimal	Basic.lean
Cor. 4.3	onsager_rayleigh_unique	Basic.lean
Thm. 5.7	cycle_affinity_constant	CRNT.lean
Thm. 5.8	deficiency_zero_equilibrium_exists	CRNT.lean
Thm. 6.2	deficiencyOne_existence	DeficiencyOne.lean
Thm. 7.3	concordance_injectivity	HigherDeficiency.lean
Thm. 8.6	deficiency_eq_dim_defect_space	Cohomology/Deficiency.lean
Cor. 8.7	deficiency_zero_iff_exact	Cohomology/Deficiency.lean
Thm. 8.9	defect_is_obstruction	Cohomology/Obstruction.lean
Thm. 8.10	variational_cohomology_duality	Cohomology/VariationalDuality.lean
Thm. 9.3	deficiency_zero_persistence	Persistence.lean
Thm. 9.7	global_attractor_single_linkage	Persistence.lean
Thm. 10.4	multistability_conditions	Multistability.lean
Thm. 11.2	hopf_bifurcation	Oscillations.lean
Thm. 11.3	routh_hurwitz_stability	Oscillations.lean
Thm. 12.3	turing_instability_conditions	ReactionDiffusion.lean
Thm. 13.5	product_form_is_stationary	Stochastic.lean
Thm. 14.2	antithetic_rpa	Control.lean

Table 2: Paper–Lean correspondence (selected).