

# COMPLEMENTARY HETEROGENEITY

*A Generating Function Theory of  
Self-Organizing Agent Economies*

Connor Smirl

Department of Economics, Tufts University

February 2026

WORKING PAPER

## Abstract

Autonomous AI agents are entering capital markets, yet no formal theory describes the dynamics of an economy where the marginal market participant is a machine. This paper provides one. The central mathematical object is the CES free energy  $\Phi = -\log F$ —the Hamiltonian of the production technology in a port-Hamiltonian formulation. Its convexity governs the system at every scale: within each level (the CES Triple Role), between levels (the activation threshold), and across timescales (the hierarchical ceiling). One function, five results.

The same convexity forces the network architecture. CES geometry requires aggregate coupling (component-level information is filtered out), directed feed-forward structure (bidirectional coupling cannot bifurcate), and nearest-neighbor topology (long-range coupling equilibrates on the slow manifold). The architecture is derived, not assumed.

The system operates at four scales—hardware (semiconductor learning curves), network (mesh formation), capability (autocatalytic training), and finance (settlement feedback)—governed by a master reproduction number: the spectral radius of a next-generation matrix. The system can be globally super-threshold while every individual

level is sub-threshold. Each scale's growth is bounded by the scale above, formalized as a slow manifold cascade. The long-run growth rate equals the frontier training rate.

Three classes of economic results follow. First, a welfare loss decomposition identifies the most institutionally rigid layer as the binding welfare constraint—not the most visible disequilibrium. Second, a damping cancellation theorem shows that local regulatory tightening has zero marginal welfare effect; reform must target the upstream layer. Third, a computable transition duration from observable drift rates yields a calibrated prediction of approximately 8 years at Wright's Law rates. Together, these yield a directed policy framework: increase gain elasticities, reform upstream, and the CES geometry handles the rest.

**Keywords:** CES aggregation, port-Hamiltonian systems, next-generation matrix, autonomous agents, hierarchical dynamics, welfare decomposition, Lyapunov stability, slow manifold, canard bifurcation

**JEL:** C62, D85, E44, L14, O33, O41

## 1. Introduction

Autonomous AI agents are entering capital markets. They process information at machine speed, optimize portfolios continuously, arbitrage mispricings in milliseconds, and settle transactions through programmable stablecoin infrastructure. No formal theory describes the dynamics of an economy where the marginal market participant is a machine. This paper provides one.

The central mathematical object is the CES free energy  $\Phi = -\log F$ . In the port-Hamiltonian formulation developed below,  $\Phi$  is the Hamiltonian—the internal energy of the production technology. Its convexity governs the system at every scale: within each level, the curvature parameter  $K = (1 - \rho)(J - 1)/J$  controls the CES Triple Role (superadditivity, correlation robustness, strategic independence); between levels, it shapes the activation threshold; and across timescales, it determines the hierarchical ceiling. One function, five results.

The same convexity forces the network architecture. The CES geometry requires aggregate coupling—component-level information is filtered out on the slow manifold, with  $F_n$  as a sufficient statistic up to  $O(\varepsilon)$  corrections. It requires directed feed-forward structure, because passive bidirectional coupling yields unconditionally stable systems incapable of bifurcation. And it requires nearest-neighbor topology, because long-range coupling equilibrates on the slow manifold under timescale separation. The architecture is derived, not assumed.

The system operates at four scales: hardware (semiconductor learning curves), network (mesh formation), capability (autocatalytic training), and finance (settlement feedback). A master reproduction number—the spectral radius of a next-generation matrix  $\mathbf{K}$ —governs cross-scale activation. The characteristic polynomial  $p(\lambda) = \prod_i(d_i - \lambda) - P_{\text{cycle}}$  admits a one-paragraph proof via permutation expansion. The system can be globally super-threshold ( $\rho(\mathbf{K}) > 1$ ) while every individual level is sub-threshold ( $d_n < 1$  for all  $n$ ), when the cycle product  $P_{\text{cycle}}^{1/N} > 1 - \max_i d_i$ : the cross-level amplification compensates for sub-threshold individual levels.

Each scale’s growth is bounded by the scale above, formalized as a slow manifold cascade. Successive equilibration of levels from fastest to slowest yields ceiling functions  $h_4, h_3, h_2$ , each bounding its level by a function of the parent. The long-run growth rate equals the frontier training rate  $g_Z$ . The Baumol bottleneck and the Triffin squeeze are the same mathematical object—a slow manifold constraint—at adjacent layers in the hierarchy.

The theory produces three classes of economic results. First, a welfare loss decomposition: under power-law gains with uniform damping,  $V = (P_{\text{cycle}}/\sigma J) \sum_n (1/\beta_n) g(F_n/\bar{F}_n)$ . The contribution of level  $n$  to welfare loss is inversely proportional to its gain elasticity  $\beta_n$ .

The binding welfare constraint is the most institutionally *rigid* layer, not the most *visible* disequilibrium. Second, a damping cancellation theorem: the Lyapunov dissipation rate  $c_n \sigma_n$  at level  $n$  is independent of  $\sigma_n$  itself. Increasing local damping speeds convergence but lowers equilibrium output; the effects exactly cancel. To accelerate adjustment at level  $n$ , increase  $\beta_n$  or reduce  $\sigma_{n-1}$ —not  $\sigma_n$ . Third, a computable transition duration  $\Delta t = \pi / \sqrt{|a| \varepsilon_{\text{drift}}}$  from the canard at the transcritical bifurcation. At Wright's Law semiconductor improvement rates, this yields an approximately 8-year transition window. Together, these yield a directed policy framework: increase gain elasticities, reform upstream, and the CES geometry handles the rest.

The paper proceeds as follows. Section 2 establishes the CES free energy  $\Phi$  as the Hamiltonian and derives the within-level eigenstructure. Section 3 proves the CES Triple Role Theorem. Section 4 proves the Port Topology Theorem and the Moduli Space characterization. Sections 5–7 apply the framework to the four economic layers: mesh equilibrium, autocatalytic growth, and settlement feedback. Section 8 proves the Master  $R_0$  and the Eigenstructure Bridge. Section 9 derives the hierarchical ceiling and canard transition dynamics. Section 10 develops the welfare decomposition and policy implications. Section 11 presents eleven empirical predictions. Section 12 states the mathematical and empirical limitations. Section 13 concludes.

## 2. The CES Free Energy

### 2.1 Setup and Notation

For  $J \geq 2$  components at level  $n$  of an  $N$ -level hierarchy, the **CES aggregate** with equal weights is

$$F_n(\mathbf{x}_n) = \left( \frac{1}{J} \sum_{j=1}^J x_{nj}^\rho \right)^{1/\rho}, \quad \mathbf{x}_n = (x_{n1}, \dots, x_{nJ}) \in \mathbb{R}_+^J \quad (1)$$

where  $\rho < 1$ ,  $\rho \neq 0$ , is the substitution parameter and  $\sigma_{\text{sub}} = 1/(1 - \rho)$  is the elasticity of substitution. The **CES free energy** (Hamiltonian) is  $\Phi = -\sum_{n=1}^N \log F_n$ .

The hierarchical dynamics evolve by

$$\varepsilon_n \dot{x}_{nj} = T_n(\mathbf{x}_{n-1}) \cdot \frac{\partial F_n}{\partial x_{nj}} - \sigma_n x_{nj}, \quad n = 1, \dots, N, \quad j = 1, \dots, J \quad (2)$$

where  $T_n = \phi_n(F_{n-1}(\mathbf{x}_{n-1}))$  is the feed-forward transmission from level  $n - 1$  to level  $n$ ,  $\sigma_n > 0$  is the damping rate, and  $\varepsilon_n > 0$  is the characteristic timescale.

**Standing Assumptions.** Throughout: (1)  $\rho < 1$ ,  $\rho \neq 0$ ; (2)  $J \geq 2$  components per level,

Table 1: Derived objects and notation.

| Symbol             | Definition  | Name                             |
|--------------------|---|----------------------------------|
| $K$                | $(1 - \rho)(J - 1)/J$                               | Within-level curvature parameter |
| $\mathbf{K}$       | Next-generation matrix (bold)                       | NGM of the reduced system        |
| $\rho(\mathbf{K})$ | Spectral radius of $\mathbf{K}$                     | Threshold parameter              |
| $\Phi$             | $-\sum_n \log F_n$                                  | CES free energy (Hamiltonian)    |
| $V$                | $\sum_n c_n D_{KL}(\mathbf{x}_n \  \mathbf{x}_n^*)$ | Storage function (Lyapunov)      |
| $W$                | $\text{diag}(W_{11}, \dots, W_{NN})$                | Supply rate (Bridge) matrix      |
| $P_{\text{cycle}}$ | $\prod_n k_{n+1,n}$                                 | Cycle product of NGM             |

$N \geq 2$  levels; (3) timescale separation  $\varepsilon_1 \gg \varepsilon_2 \gg \dots \gg \varepsilon_N > 0$ ; (4) damping  $\sigma_n > 0$  for all  $n$ ; (5) gain functions  $\phi_n : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  are  $C^1$  with  $\phi_n(0) = 0$  and  $\phi'_n > 0$ .

*Remark 2.1* (General weights). The equal-weight assumption simplifies the main development. With weights  $a_j > 0$  summing to 1, the cost-minimizing point has effective shares  $p_j = a_j^{1/(1-\rho)}$ , and all results generalize. The curvature parameter acquires a weight-dispersion factor controlled by the secular equation  $\sum_j (w_j - \mu)^{-1} = 0$  with  $w_j = a_j^{-1/(1-\rho)}$ , which has exactly  $J - 1$  roots, one in each interval  $(w_{(k)}, w_{(k+1)})$ . General-weight extensions are stated as remarks following the equal-weight proofs.

## 2.2 The Curvature Lemma

**Proposition 2.2** (Equal marginal products). *At the symmetric point  $x_{nj} = \bar{x}$  for all  $j$ , where  $F_n = \bar{x}$ , the gradient of  $F_n$  is  $\nabla F_n(\bar{x} \mathbf{1}) = (1/J) \mathbf{1}$ . In particular, all marginal products are equal, and the tangent space to the isoquant  $\{F_n = c\}$  at the symmetric point is  $T = \mathbf{1}^\perp = \{\mathbf{v} \in \mathbb{R}^J : \sum_j v_j = 0\}$ .*

*Proof.* The partial derivative is  $\partial F_n / \partial x_{nj} = (1/J) x_{nj}^{\rho-1} F_n^{1-\rho}$ . At the symmetric point,  $x_{nj} = \bar{x}$  and  $F_n = (\frac{1}{J} \cdot J \cdot \bar{x}^\rho)^{1/\rho} = \bar{x}$ , so  $\partial F_n / \partial x_{nj} = (1/J) \bar{x}^{\rho-1} \bar{x}^{1-\rho} = 1/J$ .  $\square$

This is the structural fact underlying the framework: at the symmetric allocation, the CES aggregate treats all components identically regardless of the substitution parameter.

**Proposition 2.3** (CES Hessian). *The Hessian of  $F_n$  at the symmetric point  $x_{nj} = \bar{x}$  is*

$$\nabla^2 F_n = \frac{(1 - \rho)}{J^2 \bar{x}} [\mathbf{1} \mathbf{1}^T - J I] \quad (3)$$

*with eigenvalues: 0 on  $\mathbf{1}$  (multiplicity 1), by Euler's theorem for degree-1 homogeneous functions; and  $-(1 - \rho)/(J \bar{x})$  on  $\mathbf{1}^\perp$  (multiplicity  $J - 1$ ).*

*Proof.* The general CES Hessian entry is

$$\frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{(1-\rho)}{F} \frac{\partial F}{\partial x_i} \frac{\partial F}{\partial x_j} - \delta_{ij} \frac{(1-\rho)}{x_j} \frac{\partial F}{\partial x_j}.$$

At the symmetric point,  $\partial_j F = 1/J$ ,  $F = \bar{x}$ ,  $x_j = \bar{x}$ :

$$\frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{(1-\rho)}{\bar{x}} \cdot \frac{1}{J^2} - \delta_{ij} \frac{(1-\rho)}{\bar{x}} \cdot \frac{1}{J} = \frac{(1-\rho)}{J^2 \bar{x}} (1 - J \delta_{ij}).$$

The matrix  $\mathbf{1}\mathbf{1}^T - JI$  has eigenvector  $\mathbf{1}$  with eigenvalue  $J - J = 0$ , and every  $\mathbf{v} \perp \mathbf{1}$  is an eigenvector with eigenvalue  $0 - J = -J$ . Multiplying by  $(1-\rho)/(J^2 \bar{x})$  gives the stated eigenvalues. The zero eigenvalue on  $\mathbf{1}$  also follows from Euler's theorem:  $\nabla^2 F \cdot \mathbf{x} = 0$  for degree-1 homogeneous  $F$ , and  $\mathbf{x} = \bar{x} \mathbf{1}$  at the symmetric point.  $\square$

**Definition 2.4** (Curvature parameter). *The curvature parameter of the equal-weight CES aggregate with  $J$  components is*

$$K = (1-\rho) \frac{J-1}{J}. \quad (4)$$

$K$  is the effective curvature parameter. It controls every subsequent result. Properties: (i)  $K > 0$  for all  $\rho < 1$ ; (ii)  $K$  increases monotonically with  $(1-\rho)$ ; (iii)  $K \rightarrow \infty$  as  $\rho \rightarrow -\infty$  (Leontief limit); (iv)  $K \rightarrow 0$  as  $\rho \rightarrow 1^-$  (perfect substitutes).

**Proposition 2.5** (Isoquant curvature). *At the symmetric point on  $\{F_n = c\}$ , all  $J-1$  principal curvatures of the isoquant are equal:*

$$\kappa^* = \frac{(1-\rho)}{c\sqrt{J}} = \frac{K\sqrt{J}}{c(J-1)}. \quad (5)$$

*Proof.* The normal curvature in tangent direction  $\mathbf{v} \in \mathbf{1}^\perp$  is  $\kappa(\mathbf{v}) = -\mathbf{v}^T \nabla^2 F \mathbf{v} / (\|\nabla F\| \cdot \|\mathbf{v}\|^2)$ . By Proposition 2.3,  $\mathbf{v}^T \nabla^2 F \mathbf{v} = -(1-\rho)/(J\bar{x}) \cdot \|\mathbf{v}\|^2$  for any  $\mathbf{v} \in \mathbf{1}^\perp$ . By Proposition 2.2,  $\|\nabla F\| = \sqrt{J}/J$ . Thus  $\kappa(\mathbf{v}) = (1-\rho)/(J\bar{x}) \cdot J/\sqrt{J} = (1-\rho)/(\bar{x}\sqrt{J})$ . At the isoquant  $\{F = c\}$ ,  $\bar{x} = c$ , giving the result. The curvature is independent of  $\mathbf{v}$ : the isoquant has uniform curvature at the symmetric point.  $\square$

*Remark 2.6* (General weights). With weights  $a_j > 0$  summing to 1, the principal curvatures are no longer degenerate. They are determined by the constrained eigenvalues  $\mu_1 < \dots < \mu_{J-1}$  of the weighted inverse-share matrix  $\text{diag}(1/p_1, \dots, 1/p_J)$  restricted to  $\mathbf{1}^\perp$ , via the secular equation with  $J-1$  roots. The generalized curvature parameter is  $K(\rho, \mathbf{a}) = (1-\rho)(J-1)\Phi^{1/\rho} R_{\min}/J$ , where  $R_{\min} = \mu_1$  is the smallest root. At equal weights,  $R_{\min} = J^{\sigma_{\text{sub}}}$ ,  $\Phi^{1/\rho} = J^{-\sigma_{\text{sub}}}$ , and  $K$  reduces to  $(1-\rho)(J-1)/J$ .

### 2.3 Within-Level Eigenstructure

Two matrices arise at each level, with different eigenstructures. Both are correct; they measure different things.

**Object A:**  $\nabla^2\Phi_n$  (Hessian of the free energy). This is the *geometric* object. Anisotropy ratio:  $(1 - \rho)$ .

**Proposition 2.7** (Free energy Hessian). *The Hessian of  $\Phi_n = -\log F_n$  at the symmetric point is*

$$\nabla^2\Phi_n = \frac{1}{J^2\bar{x}^2} [(1 - \rho) J I + \rho \mathbf{1}\mathbf{1}^T] \quad (6)$$

with eigenvalues  $(1 - \rho)/(J\bar{x}^2)$  on  $\mathbf{1}^\perp$  (multiplicity  $J - 1$ ) and  $1/(J\bar{x}^2)$  on  $\mathbf{1}$  (multiplicity 1). The anisotropy ratio is  $\lambda_\perp/\lambda_\parallel = (1 - \rho)$ .

*Proof.* By the chain rule,  $\nabla^2(-\log F) = -\nabla^2F/F + \nabla F \nabla F^T/F^2$ . Substituting Propositions 2.2 and 2.3 with  $F = \bar{x}$ :

$$\begin{aligned} \nabla^2\Phi_n &= \frac{(1 - \rho)}{J^2\bar{x}^2} [J I - \mathbf{1}\mathbf{1}^T] + \frac{1}{J^2\bar{x}^2} \mathbf{1}\mathbf{1}^T \\ &= \frac{1}{J^2\bar{x}^2} [(1 - \rho) J I + \rho \mathbf{1}\mathbf{1}^T]. \end{aligned}$$

On  $\mathbf{1}^\perp$ : eigenvalue  $(1 - \rho) J/(J^2\bar{x}^2) = (1 - \rho)/(J\bar{x}^2)$ . On  $\mathbf{1}$ : eigenvalue  $(1 - \rho) J/(J^2\bar{x}^2) + \rho J/(J^2\bar{x}^2) = 1/(J\bar{x}^2)$ .  $\square$

**Object B:**  $Df_n$  (Jacobian of the dynamics). This is the *dynamical* object. Anisotropy ratio:  $(2 - \rho) = 1 + (1 - \rho)$ . The extra 1 is damping.

**Proposition 2.8** (Within-level Jacobian). *At the symmetric equilibrium, where  $T_n = \sigma_n J \bar{x}$  (from the equilibrium condition  $T_n \cdot \partial F_n / \partial x_{nj} = \sigma_n x_{nj}$ ), the within-level Jacobian is*

$$Df_n = \frac{\sigma_n}{\varepsilon_n} \left[ \frac{(1 - \rho)}{J} \mathbf{1}\mathbf{1}^T - (2 - \rho) I \right] \quad (7)$$

with eigenvalues:  $-\sigma_n/\varepsilon_n$  on  $\mathbf{1}$  (aggregate mode, multiplicity 1);  $-\sigma_n(2 - \rho)/\varepsilon_n$  on  $\mathbf{1}^\perp$  (diversity modes, multiplicity  $J - 1$ ).

*Proof.* From the dynamics (2), the within-level Jacobian is  $Df_n = (1/\varepsilon_n)[T_n \nabla^2 F_n - \sigma_n I]$ .

Substituting  $T_n = \sigma_n J \bar{x}$  and the Hessian from Proposition 2.3:

$$\begin{aligned} Df_n &= \frac{1}{\varepsilon_n} \left[ \sigma_n J \bar{x} \cdot \frac{(1-\rho)}{J^2 \bar{x}} (\mathbf{1}\mathbf{1}^T - J I) - \sigma_n I \right] \\ &= \frac{\sigma_n}{\varepsilon_n} \left[ \frac{(1-\rho)}{J} \mathbf{1}\mathbf{1}^T - (1-\rho) I - I \right] \\ &= \frac{\sigma_n}{\varepsilon_n} \left[ \frac{(1-\rho)}{J} \mathbf{1}\mathbf{1}^T - (2-\rho) I \right]. \end{aligned}$$

On  $\mathbf{1}$ :  $\sigma_n/\varepsilon_n \cdot [(1-\rho) - (2-\rho)] = -\sigma_n/\varepsilon_n$ . On  $\mathbf{1}^\perp$ :  $\sigma_n/\varepsilon_n \cdot [0 - (2-\rho)] = -\sigma_n(2-\rho)/\varepsilon_n$ .  $\square$

**Corollary 2.9** (Within-level timescale separation). *The diversity modes decay faster than the aggregate mode by the factor  $(2-\rho) > 1$  for all  $\rho < 1$ :*

$$\frac{\tau_{\text{div}}}{\tau_{\text{agg}}} = \frac{1}{2-\rho}. \quad (8)$$

Equivalently, the dynamical anisotropy satisfies  $(2-\rho) = 1 + KJ/(J-1)$ . At  $\rho = 0$  (Cobb-Douglas): diversity is  $2\times$  faster. At  $\rho = -1$ : diversity is  $3\times$  faster. As  $\rho \rightarrow -\infty$  (Leontief): diversity is infinitely faster—components lock together instantaneously.

The two ratios  $(1-\rho)$  and  $(2-\rho)$  measure different things and both are correct. The Hessian ratio is geometric (curvature). The Jacobian ratio is dynamical (curvature + damping).

## 2.4 The Sufficient Statistic

**Theorem 2.10** (Aggregate coupling—sufficient statistic). *Under the standing assumptions,  $F_n$  is a sufficient statistic for level  $n$ 's state on the slow manifold, up to  $O(\varepsilon)$  corrections.*

*Proof.* The argument has three steps.

*Step 1: Equilibrium uniqueness.* At equilibrium,  $(T_n/J) x_{nj}^{\rho-1} F_n^{1-\rho} = \sigma_n x_{nj}$ , so  $x_{nj}^{\rho-2} = T_n F_n^{1-\rho}/(J \sigma_n)$ . The right side is independent of  $j$ . Since  $\rho < 1$  implies  $\rho-2 < 0$ , the map  $x \mapsto x^{\rho-2}$  is injective on  $\mathbb{R}_+$ , so  $x_{nj} = \bar{x}_n$  for all  $j$ .

*Step 2: Normal hyperbolicity.* Define the critical manifold  $\mathcal{M}_n = \{x_{nj} = F_n/J^{1/\rho} \text{ for all } j\}$ . By Proposition 2.8, the transverse eigenvalues (on  $\mathbf{1}^\perp$ ) are  $-\sigma_n(2-\rho)/\varepsilon_n$  and the tangential eigenvalue (on  $\mathbf{1}$ ) is  $-\sigma_n/\varepsilon_n$ . The spectral gap is

$$\frac{\sigma_n(2-\rho)}{\varepsilon_n} - \frac{\sigma_n}{\varepsilon_n} = \frac{\sigma_n(1-\rho)}{\varepsilon_n} > 0 \quad \text{for all } \rho < 1.$$

*Step 3: Fenichel persistence.* By Fenichel's geometric singular perturbation theorem [6], the normally hyperbolic critical manifold  $\mathcal{M}_n$  persists as a locally invariant slow manifold  $\mathcal{M}_n^\varepsilon$  within  $O(\varepsilon)$  of  $\mathcal{M}_n$ , smoothly parameterized by  $F_n$ . On  $\mathcal{M}_n^\varepsilon$ , the within-level state is determined by  $F_n$  up to  $O(\varepsilon)$  corrections. Consequently,  $F_n$  is a sufficient statistic for level  $n$ 's state on the slow manifold: any coupling through individual components  $x_{nj}$  projects onto coupling through  $F_n$ .  $\square$

The  $(2 - \rho)$  low-pass filter (Corollary 2.9) provides the dynamical mechanism: any signal component in  $\mathbf{1}^\perp$  injected at level  $n$  is suppressed by the factor  $1/(2 - \rho)$  relative to the aggregate component.

## 2.5 Preview: The Port-Hamiltonian Structure

The system admits a port-Hamiltonian (pH) representation with dissipation:

$$\dot{\mathbf{x}} = [\mathcal{J}(\mathbf{x}) - \mathcal{R}(\mathbf{x})]\nabla H(\mathbf{x}) + \mathcal{G}(\mathbf{x})\mathbf{u} \quad (9)$$

where  $H = \Phi$  is the Hamiltonian,  $\mathcal{R} = \text{diag}(\sigma_n I_J)$  is the dissipation matrix (positive semidefinite),  $\mathcal{J}$  encodes the directed coupling with lower-triangular block structure, and  $\mathcal{G}$  encodes the exogenous input.

The system is *not* a gradient flow ( $\mathcal{J} \neq 0$  is a topological obstruction). A storage function  $V = \sum_n c_n D_{KL}(\mathbf{x}_n \| \mathbf{x}_n^*)$  exists, constructed via the graph-theoretic method of Li, Shuai, and van den Driessche [1]. The **Bridge equation**

$$\nabla^2 \Phi|_{\text{slow}} = W^{-1} \cdot \nabla^2 V \quad (10)$$

relates the two through the supply rate matrix  $W$  with entries  $W_{nn} = P_{\text{cycle}}/(|\sigma_n| \varepsilon_{T_n})$ . The interpretation:  $\Phi$  is the technology (what the economy can do);  $V$  is the welfare loss (how far it is from efficiency);  $W$  is the institutional structure (how efficiently it adjusts). The full proof is in Section 8.

## 3. The CES Triple Role Theorem

The curvature parameter  $K$  from Section 2.2 simultaneously controls three properties of the CES aggregate: superadditivity, correlation robustness, and strategic independence. These are the same geometric fact—the curvature of the CES isoquant—viewed from three perspectives.

### 3.1 Superadditivity

**Theorem 3.1** (Superadditivity). *For all  $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^J \setminus \{\mathbf{0}\}$ :*

$$F(\mathbf{x} + \mathbf{y}) \geq F(\mathbf{x}) + F(\mathbf{y}) \quad (11)$$

with equality if and only if  $\mathbf{x} \propto \mathbf{y}$ . The superadditivity gap satisfies

$$F(\mathbf{x} + \mathbf{y}) - F(\mathbf{x}) - F(\mathbf{y}) \geq \frac{K}{4c} \cdot \frac{\sqrt{J}}{J-1} \cdot \min(F(\mathbf{x}), F(\mathbf{y})) \cdot d_{\mathcal{I}}(\hat{\mathbf{x}}, \hat{\mathbf{y}})^2 \quad (12)$$

where  $\hat{\mathbf{x}} = \mathbf{x}/F(\mathbf{x})$ ,  $\hat{\mathbf{y}} = \mathbf{y}/F(\mathbf{y})$  are projections onto the unit isoquant,  $d_{\mathcal{I}}$  is geodesic distance on  $\mathcal{I}_1$ , and the bound holds locally near the symmetric point.

*Proof.* Step 1 (Qualitative—from concavity and homogeneity alone). Write

$$\frac{\mathbf{x} + \mathbf{y}}{F(\mathbf{x}) + F(\mathbf{y})} = \alpha \hat{\mathbf{x}} + (1 - \alpha) \hat{\mathbf{y}}, \quad \alpha = \frac{F(\mathbf{x})}{F(\mathbf{x}) + F(\mathbf{y})}.$$

By degree-1 homogeneity,  $F(\mathbf{x} + \mathbf{y}) = (F(\mathbf{x}) + F(\mathbf{y})) \cdot F(\alpha \hat{\mathbf{x}} + (1 - \alpha) \hat{\mathbf{y}})$ . Since  $F(\hat{\mathbf{x}}) = F(\hat{\mathbf{y}}) = 1$  and  $F$  is concave:

$$F(\alpha \hat{\mathbf{x}} + (1 - \alpha) \hat{\mathbf{y}}) \geq \alpha F(\hat{\mathbf{x}}) + (1 - \alpha) F(\hat{\mathbf{y}}) = 1.$$

So  $F(\mathbf{x} + \mathbf{y}) \geq F(\mathbf{x}) + F(\mathbf{y})$ , with equality iff  $\hat{\mathbf{x}} = \hat{\mathbf{y}}$ , i.e.,  $\mathbf{x} \propto \mathbf{y}$ . This uses only concavity and degree-1 homogeneity.

Step 2 (Quantitative—from curvature comparison). The point  $\alpha \hat{\mathbf{x}} + (1 - \alpha) \hat{\mathbf{y}}$  lies on the chord of the isoquant  $\mathcal{I}_1$ . By Proposition 2.5, the isoquant has uniform positive curvature  $\kappa^* = K\sqrt{J}/[c(J-1)]$  at the symmetric point. The curvature comparison theorem for convex hypersurfaces (Toponogov comparison applied to 2-plane sections through  $\mathbf{x}^*/c$ ; do Carmo [12], Prop. 3.1) gives, for  $\hat{\mathbf{x}}, \hat{\mathbf{y}}$  in a geodesic neighborhood of  $\mathbf{x}^*/c$  with geodesic distance  $d$ :

$$F(\alpha \hat{\mathbf{x}} + (1 - \alpha) \hat{\mathbf{y}}) \geq 1 + \frac{\kappa_{\min}}{2} \alpha(1 - \alpha) d^2 + O(d^4).$$

Substituting  $\kappa_{\min} = K\sqrt{J}/[c(J-1)]$  and using  $\alpha(1 - \alpha) \geq \min(\alpha, 1 - \alpha)/2$ , with  $\min(\alpha, 1 - \alpha) \cdot (F(\mathbf{x}) + F(\mathbf{y})) = \min(F(\mathbf{x}), F(\mathbf{y}))$ , yields the quantitative bound.  $\square$

### 3.2 Correlation Robustness

**Theorem 3.2** (Correlation robustness). *Let  $\mathbf{X} = (X_1, \dots, X_J)$  be random with  $\mathbb{E}[X_j] = x_j^*$  (the symmetric allocation) and equicorrelation covariance  $\Sigma = \tau^2[(1 - r)I + r \mathbf{1}\mathbf{1}^T]$  with*

$r \geq 0$ . Let  $\gamma_* = \tau/c$  be the coefficient of variation. Then the effective dimension  $d_{\text{eff}} = J^2 g^2 \tau^2 / \text{Var}[F(\mathbf{X})]$  satisfies, to second order in  $\gamma_*$ :

$$d_{\text{eff}} \geq \frac{J}{1 + r(J-1)} + \frac{K^2 \gamma_*^2}{2} \cdot \frac{J(J-1)(1-r)}{[1+r(J-1)]^2}. \quad (13)$$

The first term is the linear baseline (achievable by any linear aggregate). The second is the **curvature bonus**, proportional to  $K^2$  and increasing in idiosyncratic variation  $(1-r)$ .

*Proof.* Expand  $Y = F(\mathbf{X})$  around  $\mathbf{x}^*$ . Let  $\boldsymbol{\epsilon} = \mathbf{X} - \mathbf{x}^*$ . The linear term  $Y_1 = (1/J) \mathbf{1} \cdot \boldsymbol{\epsilon}$  depends only on the common mode  $\bar{\epsilon}$ , with  $\text{Var}[Y_1] = g^2 \tau^2 J [1 + r(J-1)]$  where  $g = 1/J$  at equal weights.

The quadratic term  $Y_2 = \frac{1}{2} \boldsymbol{\epsilon}^T \nabla^2 F \boldsymbol{\epsilon}$  captures idiosyncratic variation through the CES Hessian. Decomposing  $\boldsymbol{\epsilon} = \bar{\epsilon} \mathbf{1} + \boldsymbol{\eta}$  with  $\mathbf{1} \cdot \boldsymbol{\eta} = 0$ :  $Y_2 = -(1-\rho)/(2Jc) \|\boldsymbol{\eta}\|^2$  at equal weights, depending purely on the idiosyncratic norm.

The variance of  $Y_2$  under equicorrelation, via the Isserlis theorem for Gaussian  $\boldsymbol{\epsilon}$ , satisfies

$$\text{Var}[Y_2]^{\text{idio}} \geq \frac{g^2 K^2 J^2}{2(J-1)c^2} \cdot \tau^4 (1-r)^2$$

using  $(J-1)R_{\min}^2 \leq \sum_{k=1}^{J-1} \mu_k^2$  and the identity  $(1-\rho)^2 \Phi^{2/\rho} (J-1)R_{\min}^2 = K^2 J^2 / (J-1)$ .

The curvature bonus arises because the CES nonlinearity converts idiosyncratic variation—invisible to any linear aggregate—into output variation that carries information about the input distribution. By the Cramér-Rao bound, the Fisher information about the mean level  $\mu$  carried by  $Y_2$  is  $\mathcal{I}_2 \geq (\partial_\mu \mathbb{E}[Y_2])^2 / \text{Var}[Y_2]$ . Combining the linear and curvature information channels yields the stated bound.  $\square$

*Remark 3.3.* For strict complements ( $\rho < 0$ ) with bounded  $\gamma_*$ :  $K > (J-1)/J$  at equal weights, so  $K^2 \gamma_*^2 J$  grows linearly in  $J$ , and  $\bar{r} \rightarrow 1$ —nearly perfect correlation becomes tolerable.

### 3.3 Strategic Independence

**Theorem 3.4** (Strategic independence). *For  $J$  strategic agents controlling  $x_j \geq 0$ , any coalition  $S \subseteq [J]$  with  $|S| = k \geq 2$  has manipulation gain  $\Delta(S) \leq 0$ . For any redistribution  $\boldsymbol{\delta}_S$  with  $\sum_{j \in S} \delta_j = 0$ :*

$$\Delta(S) \leq -\frac{K_S}{2k(k-1)c} \cdot \frac{k}{J} \cdot \frac{\|\boldsymbol{\delta}_S\|^2}{c} \leq 0 \quad (14)$$

where  $K_S = (1-\rho)(k-1)\Phi^{1/\rho} R_{\min,S}/k$  is the coalition curvature parameter.

*Proof.* *Step 1 (Qualitative—from convexity of the game).* The characteristic function  $v(S) = \max_{\mathbf{x}_S \geq 0} F(\mathbf{x}_S, \mathbf{0}_{-S})$  defines a convex cooperative game (Shapley [19]), since  $F$  is concave. The Shapley value lies in the core, and no coalition can profitably deviate. This holds for all weight vectors, without curvature computation.

*Step 2 (Quantitative—from the constrained Rayleigh quotient).* A coalition redistribution  $\boldsymbol{\delta}_S$  with  $\sum_{j \in S} \delta_j = 0$  changes output by  $\Delta F = \frac{1}{2} \boldsymbol{\delta}_S^T H_{SS} \boldsymbol{\delta}_S + O(\|\boldsymbol{\delta}\|^3)$ . For  $\boldsymbol{\delta}$  with  $\sum_{j \in S} \delta_j = 0$ , the Hessian quadratic form satisfies

$$\boldsymbol{\delta}_S^T H_{SS} \boldsymbol{\delta}_S = -\frac{(1-\rho) g \Phi^{1/\rho}}{c} \sum_{j \in S} \frac{\delta_j^2}{p_j} \leq -\frac{(1-\rho) g \Phi^{1/\rho}}{c} \cdot R_{\min, S} \cdot \|\boldsymbol{\delta}_S\|^2$$

using the constrained Rayleigh quotient on  $S$ . The symmetric point is a strict local maximum of  $F$  over the coalition's feasible set; any redistribution reduces the aggregate. The quadratic loss, expressed through  $K_S$ , gives the stated bound.  $\square$

*Remark 3.5 (Equal weights).* At equal weights,  $K_S = (1-\rho)(k-1)/k$  for all coalitions of size  $k$ , independent of which components are in  $S$ .

*Remark 3.6 (General weights).* With general weights, the coalition curvature  $K_S$  is computed via the secular equation restricted to  $S$ . The interlacing property ensures  $K_S > 0$  for all coalitions of size  $k \geq 2$ , all  $\rho < 1$ , all weight vectors.

### 3.4 The Unified Perspective

$K$  enters linearly in Theorems 3.1 and 3.4 (first-order curvature effects from the Hessian) and quadratically in Theorem 3.2 (second-order effect: the variance of a Hessian quadratic form). The three roles are the same geometric fact—the strictly positive curvature of the CES isoquant at the cost-minimizing point—viewed from three perspectives:

- **Aggregation theory** (Theorem 3.1): Curvature forces convex combinations of diverse points above the level set.
- **Information theory** (Theorem 3.2): Curvature creates a nonlinear channel through which correlated inputs map to distinct output regions, extracting idiosyncratic information.
- **Game theory** (Theorem 3.4): Curvature penalizes deviations from the balanced allocation, making the cost-minimizing bundle a Nash equilibrium.

For  $\rho = 1$  (linear aggregation,  $K = 0$ ): the isoquant is flat. Gap = 0, curvature bonus = 0, manipulation penalty = 0. All three properties vanish simultaneously.

The CES Triple Role applies to any system with CES aggregation. The results are not specific to the four-level hierarchy developed below; they hold for any CES production function, any number of levels, and any application where complementary heterogeneous agents are aggregated with constant elasticity of substitution.

## 4. The Port Topology Theorem and Moduli Space

The CES curvature constrains the network topology of the hierarchical system. Four claims collapse what could be an arbitrary directed graph with vector-valued coupling functions on  $\mathbb{R}^{NJ}$  into a specific nearest-neighbor chain with scalar aggregate coupling.

### 4.1 Statement

**Theorem 4.1** (CES-Forced Topology). *Under the standing assumptions of Section 2.1, the hierarchical CES system has the following topological properties:*

- (i) (Aggregate coupling) *Each level communicates with other levels only through its aggregate  $F_n$ . That is, the critical manifold at level  $n$  is parameterized by  $F_n$  alone.*
- (ii) (Directed coupling) *The between-level coupling is necessarily feed-forward (non-reciprocal). Any bidirectional coupling—whether power-preserving or merely passive—yields an unconditionally stable system incapable of bifurcation.*
- (iii) (Port alignment) *The port direction is  $\mathbf{b}_n \propto \mathbf{1} = \nabla F_n / \|\nabla F_n\|^2$  at the symmetric equilibrium. The port gain functions  $\phi_n$  are free parameters, not determined by  $\rho$ .*
- (iv) (Nearest-neighbor topology) *Under the timescale separation assumption, long-range coupling is dynamically equivalent to nearest-neighbor coupling with modified exogenous inputs.*

### 4.2 Proof of (i): Aggregate Coupling

This is Theorem 2.10, proved in Section 2.4: equilibrium uniqueness  $\rightarrow$  normal hyperbolicity  $\rightarrow$  Fenichel persistence.

### 4.3 Proof of (ii): Directed Coupling

Consider a two-level system on the slow manifold (scalar dynamics per level).

*Step 1: Power-preserving bidirectional coupling.* Suppose levels 1 and 2 are coupled bidirectionally with port powers summing to zero. The power-preservation constraint forces linear coupling  $\phi(F_1) = cF_1$  and  $\psi(F_2) = cF_2$  for some constant  $c > 0$ . The Jacobian is then

$$\mathcal{J}_{\text{bidir}} = \begin{pmatrix} -\sigma_1 & -c/J \\ c/J & -\sigma_2 \end{pmatrix}$$

with eigenvalues  $-(\sigma_1 + \sigma_2)/2 \pm \sqrt{(\sigma_1 - \sigma_2)^2/4 - c^2/J^2}$ . Whether the discriminant is positive (two real negative eigenvalues) or negative (complex pair with real part  $-(\sigma_1 + \sigma_2)/2 < 0$ ), both eigenvalues have strictly negative real part. The system is unconditionally stable for all  $c, \sigma_1, \sigma_2 > 0$ .

*Step 2: Extension to passive bidirectional coupling.* Any passive bidirectional coupling satisfies  $\dot{V}_{\text{coupling}} \leq 0$  by definition. This contributes a negative-semidefinite term to the effective Jacobian, which can only strengthen stability. Therefore any bidirectional coupling—whether power-preserving or merely passive—yields an unconditionally stable system.

*Step 3: Necessity of directed coupling.* The bifurcation at  $\rho(\mathbf{K}) = 1$  requires that the spectral radius of the next-generation matrix reach 1, requiring net energy injection through the hierarchy. Power-preserving bidirectional coupling contributes zero net energy; passive coupling contributes negative net energy. Neither can produce  $\rho(\mathbf{K}) = 1$ . Therefore the CES geometry, combined with the requirement for nontrivial dynamics, forces the between-level coupling to be non-reciprocal with an external energy source.  $\square$

#### 4.4 Proof of (iii): Port Alignment and Gain

*Step 1: Port direction is forced.* At a symmetric equilibrium  $\mathbf{x}_n^* = \bar{x}\mathbf{1}$ , the equilibrium condition requires  $\mathbf{b}_n \propto \mathbf{1}$ . Furthermore,  $\nabla F_n = (1/J)\mathbf{1} \propto \mathbf{1}$  at the symmetric point (Proposition 2.2), so  $\mathbf{b}_n = \nabla F_n$  is the natural CES-compatible port direction.

*Step 2: Asymmetric ports are penalized.* For  $\mathbf{b}_n \not\propto \mathbf{1}$ , the equilibrium  $\mathbf{x}_n^* \propto \mathbf{b}_n$  is asymmetric. By Jensen's inequality applied to the concave CES function ( $\rho < 1$ ),  $F(\mathbf{x}) \leq F(\bar{x}\mathbf{1})$  for any  $\mathbf{x}$  with  $\sum x_j = J\bar{x}$ . Asymmetric ports produce less aggregate output per unit input.

*Step 3: Gain function is free.* At equilibrium,  $\phi_n(\bar{F}_{n-1}) = \sigma_n J \bar{F}_n$ . For power-law gains  $\phi_n(z) = a_n z^{\beta_n}$ , the exponents  $\beta_n$  are free parameters not determined by  $\rho$ . The coefficients  $a_n$  adjust to satisfy the equilibrium condition but depend on  $\sigma_n$ ,  $J$ , and the cascade—not on  $\rho$  alone.

Geometry constrains direction; physics constrains gain.  $\square$

## 4.5 Proof of (iv): Nearest-Neighbor Topology

Consider a three-level system with long-range coupling from level 1 to level 3. Level 1, being fastest ( $\varepsilon_1 \ll \varepsilon_2$ ), equilibrates to  $F_1^* = \beta_1/(\sigma_1 J)$ , an algebraic constant on the slow manifold. The long-range coupling  $\phi_{31}(F_1^*) = \phi_{31}(\beta_1/(\sigma_1 J)) \equiv \tilde{\beta}_3$  becomes a constant, absorbed into the exogenous input to level 3. The effective dynamics become identical to a nearest-neighbor system with modified exogenous input.

The Jacobian of the reduced system is lower-triangular—*independent* of the long-range coupling strength. Long-range coupling affects the equilibrium location but not the dynamics or stability near that equilibrium.

The argument generalizes by induction to  $N$  levels: after all levels faster than level  $m$  equilibrate, any coupling  $\phi_{nm}(F_m)$  with  $m$  fast becomes  $\phi_{nm}(F_m^*) = \text{const}$ , absorbed into a redefined exogenous input. The effective topology is nearest-neighbor.

This result requires the timescale separation of Standing Assumption (3).  $\square$

## 4.6 The Moduli Space Theorem

**Theorem 4.2** (Structural Determination). *Fix the CES parameter  $\rho < 1$  and the structural integers  $(J, N)$ . The hierarchical CES system is determined as follows.*

Qualitative invariants determined by  $\rho$  (and  $J, N$ ):

- (1) *Within-level eigenstructure and curvature  $K = (1 - \rho)(J - 1)/J$ .*
- (2) *Coupling topology: aggregate, directed, nearest-neighbor, and port-aligned.*
- (3) *Existence of a bifurcation threshold at  $\rho(\mathbf{K}) = 1$ .*
- (4) *Superadditivity, correlation robustness, and strategic independence, with bounds controlled by  $K$ .*
- (5) *The Eigenstructure Bridge relating the free-energy geometry to the Lyapunov geometry.*

Free parameters (quantitative degrees of freedom):

- (1) *Timescales  $(\varepsilon_1, \dots, \varepsilon_N) \in \mathbb{R}_{++}^N$  with the ordering  $\varepsilon_1 \gg \dots \gg \varepsilon_N$ .*
- (2) *Damping rates  $(\sigma_1, \dots, \sigma_N) \in \mathbb{R}_{++}^N$ .*
- (3) *Gain functions  $\phi_n : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ,  $C^1$ ,  $\phi_n(0) = 0$ ,  $\phi'_n > 0$ .*

*The free parameters determine the equilibrium cascade  $\{\bar{F}_n\}$ , the Lyapunov weights  $\{c_n\}$ , the Bridge matrix  $W$ , and the convergence rates. The qualitative dynamics—topology, threshold structure, stability mechanism, and the triple role of curvature—are invariant across the moduli space.*

*Proof.* Parts 1–8 of the framework collectively establish each qualitative invariant. The free parameters are identified by Theorem 4.1(iii) (gain functions free), Corollary 2.9 (timescales enter only through ratios), and the equilibrium conditions  $\phi_n(\bar{F}_{n-1}) = \sigma_n J \bar{F}_n$  (which relate gains, damping, and equilibria without constraining  $\phi_n$ 's functional form). The Bridge matrix  $W_{nn} = P_{\text{cycle}}/(|\sigma_n| \varepsilon_{T_n})$  depends on  $\sigma_n$  and the elasticity of  $\phi_n$ , both of which are free.  $\square$

The gain functions are not a gap in the theory—they *are* the economic content. They encode:  $\phi_1$  = learning curve shape (semiconductor economics),  $\phi_2$  = network recruitment (platform economics),  $\phi_3$  = training efficiency (AI capability research),  $\phi_4$  = settlement demand (monetary economics). The CES geometry provides the architecture; the gain functions are where the discipline-specific economics lives.

This is a model selection result: degrees of freedom collapse from an arbitrary directed graph on  $\mathbb{R}^{NJ}$  to a nearest-neighbor chain with scalar coupling and free gain functions. The circuit theory analogy is apt: component values (resistors, capacitors) are free; circuit topology is constrained by Kirchhoff's laws.

## 4.7 Significance

Theorem 4.1 and Theorem 4.2 together say: you cannot build an arbitrary dynamical system on CES aggregates. The convexity of  $\Phi$  constrains the architecture, and the remaining freedom is modular—change the semiconductor economics ( $\phi_1$ ) without touching the settlement economics ( $\phi_4$ ). What the CES geometry determines (topology, threshold, stability) is robust. What it leaves free (timescales, damping, gain functions) is where the economics lives.

## 5. The Mesh Equilibrium

This section applies the framework to the second level of the hierarchy: the self-organizing network of heterogeneous specialized AI agents.

### 5.1 Setup

After the crossing point ( $\rho(\mathbf{K}) > 1$ ), heterogeneous AI agents with diverse capabilities self-organize into a mesh network. The mesh's degree distribution follows a power law with

exponent  $\gamma \leq 3$  (Barabási and Albert [20]), ensuring a vanishing epidemic threshold for information propagation (Pastor-Satorras and Vespignani [21]).

## 5.2 Network Formation and the Diversity Premium

The percolation threshold vanishes for  $\gamma \leq 3$ , yielding an epidemic threshold  $R_0^{\text{mesh}}$  via the mean-field approximation on the scale-free topology. By Theorem 3.1, the CES superadditivity gap quantifies the diversity premium: combining agents with different capability profiles produces more aggregate capability than the sum of their individual contributions. The premium is proportional to  $K$  and to the squared geodesic distance between agents' capability profiles on the unit isoquant.

By Theorem 4.1(i),  $F_n$  is a sufficient statistic for the level's state on the slow manifold. Individual agent capabilities are filtered out, up to  $O(\varepsilon)$  corrections on fast timescales.

## 5.3 The Crossing Condition

The critical mass  $N^*$  for the mesh equilibrium is finite and decreasing in  $K$ : higher complementarity (lower  $\rho$ ) requires fewer agents to reach the crossing point. The adoption dynamics are logistic:  $\dot{F}_2 = \beta(F_1) \cdot F_2 \cdot (1 - F_2/N^*(F_1)) - \mu F_2$ , where  $\beta(F_1)$  is the adoption rate (increasing in hardware capability) and  $N^*(F_1)$  is the carrying capacity.

## 5.4 Specialization and Knowledge Diffusion

Task differentiation follows the Bonabeau-Theraulaz-Deneubourg [22] threshold model adapted to the mesh's scale-free topology. Knowledge diffusion is governed by the graph Laplacian on the scale-free network. Both processes operate on fast timescales relative to the aggregate  $F_2$ , consistent with the sufficient statistic reduction.

# 6. Autocatalytic Growth

## 6.1 The Autocatalytic Core

When capability becomes a dynamical variable, three mechanisms make growth endogenous: training agents improve other agents (autocatalytic capability growth), operation generates training data (self-referential learning), and the mesh modifies its own composition (endogenous variety expansion). The autocatalytic threshold  $N_{\text{auto}}$  above which the mesh contains a self-sustaining improvement core—a Reflexively Autocatalytic and Food-generated (RAF) set in the sense of Hordijk and Steel [47]—scales as  $N_{\text{auto}} = O(\ln |\mathcal{R}|/\beta_t)$ .

## 6.2 Growth Dynamics

The effective production multiplier including autocatalytic feedback is  $\varphi_{\text{eff}} = \phi_0 / (1 - \beta_{\text{auto}} \cdot \phi_0)$ . Three regimes emerge with sharp boundaries: convergence to a ceiling  $C_{\max}$  when  $\varphi_{\text{eff}} < 1$  and variety is bounded (the most likely near-term regime), exponential growth when autocatalytic coupling pushes  $\varphi_{\text{eff}}$  to unity and variety expands endogenously, and finite-time singularity when  $\varphi_{\text{eff}} > 1$  with no saturation and no model collapse (conditions unlikely to hold simultaneously).

## 6.3 Collapse Protection

By Theorem 3.2, the effective dimension satisfies  $d_{\text{eff}} \geq J / (1 + r(J-1)) + O(K^2)$  when  $r < \bar{r}$ . The CES parameter  $\rho < 1$  does double duty: it governs both the diversity premium for capability aggregation and the diversity protection against model collapse. The curvature of  $\Phi$  prevents information collapse: the same force that makes the mesh capable also makes it robust to self-referential training degradation (Shumailov et al. [49]).

## 6.4 The Baumol Bottleneck

As the mesh automates progressively more inference tasks, the remaining non-automated task—frontier model training—becomes the binding constraint. Mesh growth converges to  $g_Z$  asymptotically. This is the first instance of the hierarchical ceiling derived in Section 9.

# 7. The Settlement Feedback

## 7.1 Settlement Demand

The mesh requires a programmable settlement layer for routing compensation. Dollar stablecoins, backed by US Treasuries, provide a 6.4 percentage point cost advantage over fiat payment rails. As mesh operations scale, settlement demand grows faster than inference demand: the autocatalytic operations layer adds training agent compensation, data marketplace transactions, and variety expansion incentives.

## 7.2 Market Microstructure

By Theorem 3.4, manipulation gain for any coalition  $S$  satisfies  $\Delta(S) \leq 0$ , with the penalty proportional to  $K_S$ . The CES structure makes diversity modes decay  $(2 - \rho)$  times faster than aggregate modes (Corollary 2.9), so manipulation signals are suppressed faster than aggregate price signals—a testable prediction about market dynamics.

### 7.3 Monetary Policy Degradation

As the fraction  $\phi$  of capital managed by autonomous agents increases, monetary policy tools degrade in order. Forward guidance degrades first (it depends on information processing delay, which mesh agents eliminate). Quantitative easing degrades second (it depends on arbitrage speed, which mesh agents improve). Financial repression degrades last but most sharply (it depends on captive savings, which collapse discontinuously when stablecoin access crosses a threshold  $S_{\text{crit}}$ ). The surviving channels—interest rate and lender-of-last-resort—operate through real economy dynamics rather than market frictions. The Brunnermeier-Sannikov [33] volatility paradox applies: low exogenous volatility from mesh efficiency may breed endogenous instability from monetary policy ineffectiveness.

### 7.4 Dollarization and the Triffin Squeeze

Stablecoin access lowers the Uribe [37] bifurcation thresholds  $\bar{\pi}(S)$  and  $\underline{\pi}(S)$ : currencies previously stable become vulnerable as the mesh's settlement infrastructure grows. The Farhi-Maggiori [40] safety zone shrinks as  $\phi$  increases ( $d\bar{b}/d\phi < 0$ ), because mesh agents destroy the information-insensitivity that defines safe assets. Within the instability zone, crisis character shifts from sunspot-driven to fundamentals-driven.

The modern Triffin contradiction: stablecoin demand pushes Treasury supply  $b$  upward while mesh participation makes the safety boundary  $\bar{b}(\phi)$  lower. The squeeze is self-reinforcing when  $\dot{b} > 0$  and  $\dot{\bar{b}} < 0$  simultaneously.

### 7.5 The Coupled ODE System

Four equations on the slow manifold, valid to  $O(\varepsilon)$ :

$$\dot{\phi} = \gamma_\phi \cdot \phi(1 - \phi) \cdot [\mu_\phi(S, \eta) - r_\phi] \quad (15)$$

$$\dot{S} = \gamma_S \cdot S \cdot [g_{\text{mesh}}(\phi) + g_{\text{dollar}}(S, b) - \delta_S] \quad (16)$$

$$\dot{b} = \gamma_b \cdot [d(b, \eta) + s_{\text{coin}}(S) - \tau(b)] \quad (17)$$

$$\dot{\eta} = \mu_\eta(\phi) \cdot \eta - \sigma_\eta^2(\phi, S) \cdot \eta(1 - \eta) - \ell(b, \eta) \quad (18)$$

Three equilibrium classes: low-mesh (current system approximately), high-mesh (weak monetary policy but market discipline substitutes), and an unstable intermediate separatrix. During the bifurcation at  $\rho(\mathbf{K}) = 1$ , the approximation is least reliable—see Section 9.4.

## 7.6 The Synthetic Gold Standard

The high-mesh equilibrium constrains sovereign fiscal policy through real-time market discipline—a “synthetic gold standard” that emerges from the model rather than being assumed. The constraint is embedded in the information infrastructure: a government can default on its debt, but it cannot prevent mesh agents from pricing the default risk. Whether this constraint is desirable is a normative question beyond this paper’s scope.

## 8. The Master $R_0$ and the Eigenstructure Bridge

This is the dynamical core.

### 8.1 The Reduced System

**Proposition 8.1** (Reduced dynamics). *On the slow manifold  $\mathcal{M}^\varepsilon = \prod_n \mathcal{M}_n^\varepsilon$ , the aggregate dynamics are*

$$\varepsilon_n \dot{F}_n = \phi_n(F_{n-1})/J - \sigma_n F_n, \quad n = 1, \dots, N \quad (19)$$

with the convention that  $\phi_1$  depends on  $F_N$  (cyclic) or on an exogenous input  $\beta_1$  (open).

*Proof.* By Theorem 2.10, the slow manifold at each level is parameterized by  $F_n$ , with within-level state  $x_{nj} = F_n/J^{1/\rho}$  for all  $j$  up to  $O(\varepsilon)$  corrections. Project the dynamics onto the aggregate:  $\dot{F}_n = \nabla F_n \cdot \dot{\mathbf{x}}_n = (1/J)\mathbf{1} \cdot \dot{\mathbf{x}}_n$ . Summing (2) over  $j$ :  $\varepsilon_n J \dot{F}_n = T_n - \sigma_n J F_n$ , using  $\sum_j \partial F_n / \partial x_{nj} = 1$  (Euler’s theorem) and  $\sum_j x_{nj} = J F_n$  (symmetric allocation). Thus  $\varepsilon_n \dot{F}_n = T_n/J - \sigma_n F_n = \phi_n(F_{n-1})/J - \sigma_n F_n$ .  $\square$

### 8.2 The Next-Generation Matrix

At the nontrivial equilibrium, the Jacobian decomposes as  $J_{\text{agg}} = T + \Sigma$ , where  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_N)$  with  $\sigma_n < 0$  encodes damping and  $T$  encodes amplification. The **next-generation matrix** is

$$\mathbf{K} = -T\Sigma^{-1} \quad (20)$$

with entries  $K_{nn} = d_n = T_{nn}/|\sigma_n|$  (within-level reproduction) and  $K_{n,n-1} = k_{n,n-1} = T_{n,n-1}/|\sigma_{n-1}|$  (cross-level amplification).

### 8.3 Characteristic Polynomial

**Theorem 8.2** (NGM characteristic polynomial). *For a cyclic  $N$ -level system with diagonal entries  $d_1, \dots, d_N$ , nearest-neighbor coupling  $k_{n+1,n}$ , and cycle closure  $k_{1N}$ , the characteristic*

polynomial of the NGM is

$$p(\lambda) = \prod_{i=1}^N (d_i - \lambda) - P_{\text{cycle}} \quad (21)$$

where  $P_{\text{cycle}} = \prod_n k_{n+1,n}$  is the cycle product.

*Proof.* In the Leibniz expansion  $\det(\mathbf{K} - \lambda I) = \sum_{\pi \in S_N} \text{sgn}(\pi) \prod_i (\mathbf{K} - \lambda I)_{i,\pi(i)}$ , a permutation  $\pi$  contributes a nonzero product only if every factor  $(\mathbf{K} - \lambda I)_{i,\pi(i)}$  is nonzero. The matrix  $\mathbf{K} - \lambda I$  has diagonal entries  $(d_i - \lambda)$ , subdiagonal entries  $k_{n+1,n}$ , corner entry  $k_{1N}$ , and all others zero.

Only two permutations have all nonzero factors: (a) The identity  $\pi = \text{id}$ , contributing  $\prod_i (d_i - \lambda)$ . (b) The full  $N$ -cycle  $\pi = (1 \ 2 \ \cdots \ N)$ , which selects the subdiagonal entries  $k_{i,i-1}$  for  $i = 2, \dots, N$  and the corner entry  $k_{1N}$ . Its sign is  $(-1)^{N-1}$  and its product is  $P_{\text{cycle}}$ .

Any other permutation must include at least one factor from a zero entry. For  $N = 4$ :  $(-1)^{N-1} = -1$ , giving  $p(\lambda) = \prod(d_i - \lambda) - P_{\text{cycle}}$ .  $\square$

**Economic corollary.**  $P_{\text{cycle}}$  is a geometric mean of cross-level couplings. The sensitivity  $\partial \rho / \partial k_{ij} \propto P_{\text{cycle}} / k_{ij}$  is largest for the smallest  $k_{ij}$ . The system's activation is bottlenecked by its weakest cross-level link. Directed investment thesis: invest in the weakest coupling, not the strongest.

#### 8.4 Spectral Threshold

**Corollary 8.3** (Spectral radius). *Equal diagonals ( $d_n = d$ ):  $\rho(\mathbf{K}) = d + P_{\text{cycle}}^{1/N}$ . Zero diagonals:  $\rho(\mathbf{K}) = P_{\text{cycle}}^{1/N}$ . General:  $\rho(\mathbf{K}) > \max_i d_i$  (Perron-Frobenius) and  $\rho(\mathbf{K}) \leq \max_i \sum_j K_{ij}$  (row-sum bound).*

**Theorem 8.4** (Bifurcation threshold). *The nontrivial equilibrium exists and is locally asymptotically stable if and only if  $\rho(\mathbf{K}) > 1$ . The bifurcation at  $\rho(\mathbf{K}) = 1$  is a transcritical bifurcation exchanging stability between the trivial and nontrivial equilibria. The system can be globally super-threshold ( $\rho(\mathbf{K}) > 1$ ) from individually sub-threshold levels ( $d_n < 1$  for all  $n$ ) when  $P_{\text{cycle}}^{1/N} > 1 - \max_i d_i$ .*

Where curvature enters: the CES curvature parameter  $K$  enters through  $k_{32}$  involving the CES marginal product. At the symmetric allocation,  $\partial F_{\text{CES}} / \partial F_2 = 1/J$ . Away from the symmetric point, this marginal product depends on  $\rho$  through  $F^{1-\rho} x_j^{\rho-1}$ , and  $K$  controls its sensitivity to reallocation.

## 8.5 The Port-Hamiltonian Structure

**Proposition 8.5** (pH representation). *The hierarchical CES system admits the port-Hamiltonian representation (9) where:  $H = \Phi$  is the Hamiltonian;  $\mathcal{R} = \text{diag}(\sigma_n I_J)$  is the dissipation matrix;  $\mathcal{J}$  encodes directed coupling with lower-triangular block structure and rank-1 cross-level blocks.*

*Three structural consequences:*

- (a)  $\mathcal{J}$  is lower-triangular with nonzero sub-diagonal blocks, so the system is **not** a gradient flow ( $\mathcal{J} \neq 0$ ). This is a topological obstruction: no coordinate transformation can symmetrize a lower-triangular Jacobian.
- (b)  $\mathcal{J}$  has rank-1 cross-level blocks (of the form  $c \cdot \mathbf{1}\mathbf{1}^T$ ), confirming that coupling passes through  $F_n$  (Theorem 4.1(i)).
- (c)  $\mathcal{J}\nabla H \neq 0$ , so the GENERIC degeneracy condition  $L\nabla\Phi = 0$  fails. The directed coupling injects free energy from lower to higher levels. This is not a deficiency: it is the mechanism by which the bifurcation at  $\rho(\mathbf{K}) = 1$  becomes possible.

## 8.6 The Storage Function

**Theorem 8.6** (Graph-theoretic Lyapunov function). *Define*

$$V(\mathbf{x}) = \sum_{n=1}^N c_n \sum_{j=1}^J \left( \frac{x_{nj}}{x_{nj}^*} - 1 - \log \frac{x_{nj}}{x_{nj}^*} \right) = \sum_{n=1}^N c_n D_{KL}(\mathbf{x}_n \| \mathbf{x}_n^*) \quad (22)$$

with tree coefficients  $c_n = P_{\text{cycle}}/k_{n,n-1}$ . Then  $V$  is a Lyapunov function for the hierarchical CES system near the nontrivial equilibrium:  $V \geq 0$  with  $V = 0$  iff  $\mathbf{x} = \mathbf{x}^*$ , and  $\dot{V} \leq 0$ .

*Proof.* Nonnegativity follows from  $g(z) = z - 1 - \log z \geq 0$  with equality iff  $z = 1$ . Along trajectories:

$$\dot{V} = \sum_n c_n \sum_j \left( 1 - \frac{x_{nj}^*}{x_{nj}} \right) f_{nj}(\mathbf{x}).$$

The within-level contributions are

$$\dot{V}_{\text{within}} = - \sum_n c_n \sigma_n \sum_j \frac{(x_{nj} - x_{nj}^*)^2}{x_{nj}} \leq 0.$$

The cross-level contributions cancel by the tree condition on  $c_n$ . This is the Li-Shuai-van den Driessche [1] construction applied to the cycle-graph topology, using the Volterra-Lyapunov

identity  $a - b \log(a/b) \leq a - b + b \log(b/a)$  for each coupling term. The specific coefficients  $c_n = P_{\text{cycle}}/k_{n,n-1}$  are those required for cancellation on the cycle graph. For the 4-level cycle, there is exactly one spanning in-tree per root.  $\square$

## 8.7 The Eigenstructure Bridge

**Theorem 8.7** (Eigenstructure Bridge). *On the slow manifold, the Hessians of  $\Phi$  and  $V$  at the nontrivial equilibrium satisfy*

$$\nabla^2 \Phi|_{\text{slow}} = W^{-1} \cdot \nabla^2 V \quad (23)$$

where  $W = \text{diag}(W_{11}, \dots, W_{NN})$  is the **Bridge matrix** with entries

$$W_{nn} = \frac{P_{\text{cycle}}}{|\sigma_n| \varepsilon_{T_n}} \quad (24)$$

and  $\varepsilon_{T_n} = T'_n(\bar{F}_{n-1})\bar{F}_{n-1}/T_n(\bar{F}_{n-1})$  is the elasticity of the coupling at level  $n$ .

*Proof.* On the slow manifold,  $\Phi|_{\text{slow}} = -\sum_n \log F_n$  and  $V = \sum_n c_n \bar{F}_n g(F_n/\bar{F}_n)$ . Their Hessians at equilibrium are diagonal:

$$(\nabla^2 \Phi|_{\text{slow}})_{nn} = \frac{1}{\bar{F}_n^2}, \quad (\nabla^2 V)_{nn} = \frac{c_n}{\bar{F}_n}.$$

The ratio is  $(\nabla^2 \Phi)_{nn}/(\nabla^2 V)_{nn} = 1/(c_n \bar{F}_n) = W_{nn}^{-1}$ . Expressing  $c_n = P_{\text{cycle}}/k_{n,n-1}$  and  $k_{n,n-1} = T'_n(\bar{F}_{n-1})\bar{F}_{n-1}/|\sigma_{n-1}|$ , together with the equilibrium relation  $T_n(\bar{F}_{n-1}) = |\sigma_n| \bar{F}_n$ :

$$c_n \bar{F}_n = \frac{P_{\text{cycle}}}{k_{n,n-1}} \cdot \bar{F}_n = P_{\text{cycle}} \cdot \frac{|\sigma_{n-1}|}{T'_n \bar{F}_{n-1}} \cdot \bar{F}_n = \frac{P_{\text{cycle}}}{|\sigma_n|} \cdot \frac{T_n}{T'_n \bar{F}_{n-1}} = \frac{P_{\text{cycle}}}{|\sigma_n| \varepsilon_{T_n}}.$$

$\square$

$\Phi$  is the Hamiltonian—it encodes the production technology.  $V$  is the storage function—it measures the welfare loss from disequilibrium.  $W$  is the supply rate matrix—it encodes how efficiently the institutional structure converts energy injection into equilibrium adjustment. The Bridge equation is the passivity inequality of the pH system, not a mysterious coincidence.

The production technology ( $\rho$ ) determines *which* adjustments are fast and which are slow (eigenvectors). The institutional parameters ( $\sigma_n, \phi_n$ ) determine *how fast* (eigenvalues). Different countries or policy regimes have different  $\sigma_n$  and  $\phi_n$ , so they converge at different

rates, but along the same directions. This is a Lucas-critique-compatible statement: the structure is policy-invariant; the dynamics are not.

## 8.8 Special Cases

*Power-law coupling* ( $\phi_n(z) = a_n z^{\beta_n}$ ): The elasticity is constant,  $\varepsilon_{T_n} = \beta_n$ , giving  $W_{nn} = P_{\text{cycle}}/(\beta_n |\sigma_n|)$ .

*Linear coupling* ( $\beta_n = 1$ ) with uniform damping ( $\sigma_n = \sigma$ ):  $W = (P_{\text{cycle}}/\sigma) I$ , so  $\nabla^2 \Phi|_{\text{slow}} = (\sigma/P_{\text{cycle}}) \nabla^2 V$ . This is the only case where the system is “almost” a gradient flow.

# 9. The Hierarchical Ceiling and Transition Dynamics

## 9.1 Timescale Assignment

Table 2: Timescale hierarchy.

| Level             | Domain                     | Timescale           | Ordering                          |
|-------------------|----------------------------|---------------------|-----------------------------------|
| $n = 1$ (slowest) | Hardware (learning curves) | $\varepsilon_1 = 1$ | Reference                         |
| $n = 2$           | Mesh (network formation)   | $\varepsilon_2$     | $\varepsilon_2 \ll 1$             |
| $n = 3$           | Capability (training)      | $\varepsilon_3$     | $\varepsilon_3 \ll \varepsilon_2$ |
| $n = 4$ (fastest) | Settlement (finance)       | $\varepsilon_4$     | $\varepsilon_4 \ll \varepsilon_3$ |

## 9.2 The Slow Manifold Cascade

**Proposition 9.1** (Ceiling functions). *Under the timescale ordering, successive equilibration yields:*

Level 4 (fastest):  $h_4(F_2, F_3) = \bar{S}(F_3)(1 - \nu/\eta(F_3, F_2))$ . Existence requires  $\eta(F_3, F_2) > \nu$ .

**Ceiling:**  $F_4 \leq \bar{S}(F_3)$ .

Level 3:  $h_3(F_2) = (\varphi_{\text{eff}}/\delta_C) \cdot F_{\text{CES}}(F_2)$ . **Ceiling:**  $F_3 \leq (\varphi_{\text{eff}}/\delta_C) \cdot F_{\text{CES}}(N^*(F_1))$ . The CES diversity premium (controlled by  $K$  through Theorem 3.1) enters here: higher  $K$  yields a larger diversity bonus in the capability aggregate.

Level 2:  $h_2(F_1) = N^*(F_1)(1 - \mu/\beta(F_1))$ . Existence requires  $\beta(F_1) > \mu$ . **Ceiling:**  $F_2 \leq N^*(F_1)$ .

Each level is bounded by a function of its parent. The cascade of ceilings  $F_1 \rightarrow h_2 \leq N^* \rightarrow h_3 \leq (\varphi_{\text{eff}}/\delta_C)F_{\text{CES}}(N^*) \rightarrow h_4 \leq \bar{S}(h_3)$  bounds each level by a function of the level below in the timescale hierarchy.

### 9.3 Effective Dynamics and the Baumol Bottleneck

Substituting the cascade into the Level 1 dynamics:

$$\dot{F}_1 = \delta_c \cdot \Psi(F_1)^\alpha \cdot F_1^{\phi_c} - \gamma_c F_1 \quad (25)$$

where  $\Psi(F_1) = I(h_4(h_2(F_1), h_3(h_2(F_1))))$  is the composite feedback function encoding the entire cascade. The long-run growth rate is determined entirely by Level 1—the slowest-adapting sector.

The Baumol bottleneck and the Triffin squeeze are the same mathematical object (a slow manifold constraint) at adjacent layers. Approximation error is bounded by the timescale ratio  $\varepsilon_{n+1}/\varepsilon_n$ .

### 9.4 Transition Dynamics: The Canard

When  $\rho(\mathbf{K})$  crosses 1, the slow manifold loses normal hyperbolicity.

**Proposition 9.2** (Transcritical normal form). *At the bifurcation point  $(\bar{F}_1, \mu^*)$  where  $g = 0$  and  $\partial g/\partial F_1 = 0$ , the dynamics admit the local normal form*

$$\dot{y} = a \epsilon y + b y^2 + O(|y|^3 + |\epsilon|^2) \quad (26)$$

where  $y = F_1 - \bar{F}_1$ ,  $\epsilon = \mu - \mu^*$ , and

$$a = \frac{\partial^2 g}{\partial F_1 \partial \mu} \Big|_{\text{bif}}, \quad b = \frac{1}{2} \frac{\partial^2 g}{\partial F_1^2} \Big|_{\text{bif}}.$$

**Theorem 9.3** (Canard duration). *If the bifurcation parameter drifts linearly at rate  $\varepsilon_{\text{drift}}$ , the crisis duration is*

$$\Delta t_{\text{crisis}} = \frac{\pi}{\sqrt{|a| \varepsilon_{\text{drift}}}} + O\left(\frac{\log(1/\delta)}{\sqrt{|a| \varepsilon_{\text{drift}}}}\right). \quad (27)$$

*If  $\mu = \gamma_c$  (damping drifts down):  $a = -1$ , duration  $= \pi/\sqrt{\varepsilon_{\text{drift}}}$ , independent of all system parameters.*

*If  $\mu = \delta_c$  (efficiency drifts up):  $|a|$  depends on the product of cascade elasticities  $\prod_n \beta_n$ .*

*In both cases, duration scales as  $\varepsilon_{\text{drift}}^{-1/2}$ : halving the drift rate increases transition time by  $\sqrt{2} \approx 41\%$ .*

*Proof sketch.* This is the standard delayed loss of stability result for passage through a transcritical bifurcation (Neishtadt [13, 14]; Berglund and Gentz [15]). In the rescaled variable  $\tau = \sqrt{|a| \varepsilon_{\text{drift}}} t$ , the normal form becomes the Weber equation plus a quadratic perturbation.

The passage through zero eigenvalue creates a delay of  $\pi$  time units in the rescaled variable. Converting back:  $\Delta t = \pi / \sqrt{|a|\varepsilon_{\text{drift}}}$ .  $\square$

**Where  $K$  enters.** The second-order coefficient  $b$  contains  $\partial^2 F_{\text{CES}} / \partial F_2^2|_{\text{sym}} = -K/(J\bar{F}_2)$  by Proposition 2.3.  $K$  controls the sharpness of the transition (through  $b$ ), not its duration (through  $a$ ). Higher complementarity produces a faster snap to the new equilibrium, with less overshooting.

**Economic content.** How long does the low-mesh  $\rightarrow$  high-mesh transition take? Answer:  $O(1/\sqrt{\varepsilon_{\text{drift}}})$ , with the constant controlled by the chain of gain elasticities, not by the CES substitution parameter. If semiconductor costs decline at 15% per year (Wright's Law), the transition duration is approximately  $\pi/\sqrt{0.15} \approx 8$  years.

## 9.5 Dispersion as Leading Indicator

At the bifurcation, the spectral gap  $(1 - \rho)\sigma_n/\varepsilon_n$  closes. Within-level heterogeneity stops being slaved to the aggregate. Prediction: cross-sectional variance of agent performance widens *before* aggregate statistics move. The within-mesh Gini coefficient rises before the crossing and collapses after (as diversity modes re-equilibrate on the new slow manifold).

# 10. Welfare Decomposition and Policy

This section contains the economic implications of the Bridge equation and the three extensions.

## 10.1 The Welfare Loss Decomposition

**Proposition 10.1** (Closed-form Lyapunov under power-law gains). *Suppose all gain functions are power laws:  $\phi_n(z) = a_n z^{\beta_n}$  with  $\beta_n > 0$ . Then the tree coefficients are*

$$c_n = \frac{P_{\text{cycle}} \sigma_{n-1}}{\beta_n \sigma_n J \bar{F}_n} \quad (28)$$

and the Lyapunov function takes the form

$$V = \frac{P_{\text{cycle}}}{J} \sum_{n=1}^N \frac{\sigma_{n-1}}{\beta_n \sigma_n} g\left(\frac{F_n}{\bar{F}_n}\right). \quad (29)$$

Under uniform damping  $\sigma_n = \sigma$ , this simplifies to

$$V = \frac{P_{\text{cycle}}}{\sigma J} \sum_{n=1}^N \frac{1}{\beta_n} g\left(\frac{F_n}{\bar{F}_n}\right) \quad (30)$$

a one-parameter-per-level family indexed by  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_N)$ .

*Proof.* From Theorem 8.6,  $c_n = P_{\text{cycle}}/k_{n,n-1}$  where  $k_{n,n-1} = \phi'_n(\bar{F}_{n-1})\bar{F}_{n-1}/|\sigma_{n-1}|$ . For power-law  $\phi_n(z) = a_n z^{\beta_n}$ :  $\phi'_n(\bar{F}_{n-1}) \cdot \bar{F}_{n-1} = \beta_n \phi_n(\bar{F}_{n-1}) = \beta_n \sigma_n J \bar{F}_n$  (using the equilibrium condition  $\phi_n(\bar{F}_{n-1}) = \sigma_n J \bar{F}_n$ ). Thus  $k_{n,n-1} = \beta_n \sigma_n J \bar{F}_n / \sigma_{n-1}$ , giving  $c_n = P_{\text{cycle}} \sigma_{n-1} / (\beta_n \sigma_n J \bar{F}_n)$ . On the slow manifold,  $c_n D_{KL} = c_n \bar{F}_n g(F_n/\bar{F}_n) = P_{\text{cycle}} \sigma_{n-1} / (\beta_n \sigma_n J) g(F_n/\bar{F}_n)$ .  $\square$

The contribution of level  $n$  to welfare loss is proportional to  $g(F_n/\bar{F}_n)/\beta_n$ . Levels with inelastic gain functions (small  $\beta_n$ ) dominate.

**Economic interpretation.** The binding welfare constraint is the most *rigid* layer, not the most *visible* disequilibrium. If settlement demand is elastic ( $\beta_4$  large) but capability formation is inelastic ( $\beta_3$  small), then capability fragmentation dominates welfare loss even if settlement is farther from equilibrium.

For the current macro debate: the welfare-relevant bottleneck is more likely at the capability or mesh layer (slow-moving training pipelines, regulatory barriers to AI deployment) than at the settlement layer (fast-moving fintech).

## 10.2 The Damping Cancellation

**Proposition 10.2** (Damping-speed tradeoff). *For the reduced system on the slow manifold:*

(i) *The local convergence rate at level  $n$  is  $\sigma_n/\varepsilon_n$ , strictly increasing in  $\sigma_n$ .*

(ii) *The equilibrium output is  $\bar{F}_n = \phi_n(\bar{F}_{n-1})/(\sigma_n J)$ , strictly decreasing in  $\sigma_n$ .*

(iii) *The Lyapunov dissipation rate at level  $n$  near equilibrium is*

$$-\dot{V}_n \approx c_n \sigma_n \frac{(\delta F_n)^2}{\bar{F}_n} = \frac{P_{\text{cycle}} \sigma_{n-1}}{\beta_n J \bar{F}_n} \cdot \frac{(\delta F_n)^2}{\bar{F}_n} \quad (31)$$

*(under power-law gains), which is independent of  $\sigma_n$  itself.*

*Proof.* (i) The eigenvalue of the reduced Jacobian is  $-\sigma_n/\varepsilon_n$  (Proposition 2.8 restricted to the aggregate mode). (ii) Direct from the equilibrium condition. (iii)  $\dot{V}_n = -c_n \sigma_n (F_n - \bar{F}_n)^2 / F_n \approx -c_n \sigma_n (\delta F_n)^2 / \bar{F}_n$ . Substituting  $c_n$  from Proposition 10.1:  $c_n \sigma_n = P_{\text{cycle}} \sigma_{n-1} / (\beta_n J \bar{F}_n)$ , which is independent of  $\sigma_n$ .  $\square$

The Lyapunov dissipation rate at level  $n$  does not depend on  $\sigma_n$  itself. Increasing  $\sigma_n$  speeds local convergence but lowers equilibrium output; the effects exactly cancel. The dissipation depends on  $\sigma_{n-1}$ —the damping of the *upstream* level.

**Theorem 10.3** (Upstream reform principle). *To accelerate welfare-relevant adjustment at level  $n$ : increase the gain elasticity  $\beta_n$  or reduce upstream damping  $\sigma_{n-1}$ —not local damping  $\sigma_n$ .*

The policy chain:

- Fix settlement ( $n = 4$ ): reform capability aggregation ( $\sigma_3$ ) or increase settlement elasticity ( $\beta_4$ ).
- Fix capability ( $n = 3$ ): reform mesh recruitment ( $\sigma_2$ ) or increase training elasticity ( $\beta_3$ ).
- Fix mesh density ( $n = 2$ ): reform hardware investment ( $\sigma_1$ ) or increase recruitment elasticity ( $\beta_2$ ).
- Fix hardware ( $n = 1$ ): reduce  $\gamma_c$  directly (CHIPS Act, etc.).

**Corollary.** Stablecoin regulation ( $\sigma_4$ ) has zero marginal welfare effect. Capability-layer reform ( $\sigma_3$  or  $\beta_4$ ) has positive marginal welfare effect. This is a theorem, not a heuristic.

### 10.3 The Global Welfare Ordering

**Corollary 10.4** (Welfare ordering). *Define the partial order  $\boldsymbol{\beta} \succeq \boldsymbol{\beta}'$  iff  $\beta_n \geq \beta'_n$  for all  $n$ . Then:*

- (i)  $W_{nn}(\boldsymbol{\beta}) \leq W_{nn}(\boldsymbol{\beta}')$  for all  $n$  (the Bridge tightens).
- (ii)  $V(\boldsymbol{\beta}) \leq V(\boldsymbol{\beta}')$  at every non-equilibrium state (welfare loss decreases).

*Proof.* (i)  $W_{nn} = P_{\text{cycle}}/(\beta_n |\sigma_n|)$  under power-law gains, strictly decreasing in  $\beta_n$ . (ii)  $c_n \propto 1/\beta_n$ ; higher  $\beta_n$  gives lower weight  $c_n$ , hence lower  $V$ .  $\square$

Increasing any gain elasticity at any level is unambiguously welfare-improving, regardless of the current state of the economy. Policies that increase the responsiveness of cross-level coupling are always welfare-improving. Policies that flatten response curves are always welfare-reducing.

## 10.4 The Rigidity Ordering

From the Bridge matrix:  $W_{11} > W_{22} > W_{33} > W_{44}$  (hardware stiffest, settlement loosest) when the timescale and damping orderings align. Policy interventions at stiff layers (semiconductor subsidies, export controls) have persistent effects. Interventions at loose layers (stablecoin regulation) have transient effects the system routes around.

## 10.5 The Logistic Fragility Condition

For logistic gain  $\phi_n(z) = r_n z(1 - z/K_n)$ , the elasticity at equilibrium depends on utilization  $u_n = \bar{F}_{n-1}/K_n$ :

$$\varepsilon_{T_n} = \frac{1 - 2u_n}{1 - u_n}.$$

The tree coefficient has a pole at  $u_n = 1/2$ : as the upstream level approaches half its carrying capacity,  $c_n \rightarrow \infty$ . Stability requires  $u_n < 1/2$  (operating below the logistic inflection). At  $u_n > 1/2$ , the elasticity goes negative, the tree coefficient changes sign, and  $V$  ceases to be a Lyapunov function. Operating above the logistic inflection is destabilizing, not merely inefficient.

Prediction: variance of mesh-related indicators spikes when agent density reaches approximately 50% of infrastructure capacity—at the inflection point, not at saturation. Design criterion: engineer carrying capacity so equilibrium utilization stays well below 50%.

# 11. Empirical Predictions

## 11.1 Calibration Inputs

Semiconductor learning curves provide the drift rate  $\varepsilon_{\text{drift}}$ : at Wright's Law rates of approximately 15% annual cost decline, the canard duration formula yields a transition window of order 8 years. The monetary productivity gap (6.4%) anchors the settlement cost advantage. The six-stage country classification maps dollarization vulnerability.

## 11.2 Predictions

**P1–P3: Testing the Triple Role.** (P1) Cross-agent capability profiles on the unit isoquant diverge as the mesh matures, with superadditivity gap proportional to  $K \cdot d_T^2$ . (P2) Model collapse incidence remains below threshold for mesh-trained agents, with  $\alpha_{\text{eff}}$  bounded below by a function of  $K$ . (P3) Coalition manipulation gain in mesh-mediated markets satisfies  $\Delta(S) \leq 0$  with penalty proportional to  $K_S$ .

**P4: Testing the spectral threshold (UPGRADED).** Cross-layer acceleration occurs with delay  $\approx \pi/\sqrt{|a|\varepsilon_{\text{drift}}}$  after the drift parameter crosses the bifurcation threshold. At Wright's Law rates: 6–10 year transition window. *Falsification:* no acceleration by 2035.

**P5–P6: Testing monetary policy degradation.** (P5) Forward guidance effectiveness declines before QE effectiveness, which declines before financial repression collapses. (P6) The duration of market impact from FOMC statements declines as autonomous agent market share grows.

**P7–P8: Testing settlement feedback.** (P7) Stablecoin Treasury holdings exceed 5% of short-duration Treasury supply by 2028. (P8) At least one country group experiences stablecoin-mediated dollarization by 2030.

**P9: Testing the hierarchical ceiling.** The ratio of mesh capability growth to frontier training rate converges:  $\dot{C}_{\text{mesh}}/\dot{C}_{\text{frontier}} \rightarrow 1$ .

**P10: Testing damping cancellation (NEW).** Tightening stablecoin regulation ( $\sigma_4$ ) has no persistent effect on mesh welfare convergence. Capability-layer reforms (reducing  $\sigma_3$  or increasing  $\beta_3$ ) do. *Falsification:* if stablecoin regulation measurably slows or accelerates mesh formation.

**P11: Testing the dispersion indicator (NEW).** Cross-sectional variance of AI agent performance metrics widens before aggregate mesh statistics shift. *Falsification:* if aggregate leads dispersion.

## 12. Limitations

### 12.1 Mathematical

1. *Gain functions genuinely free.* Theorem 4.1(iii) establishes this as a proved impossibility: the exponents and coefficients of  $\phi_n$  are not determined by  $\rho$ . Since  $\phi_n$  determines  $\{\bar{F}_n\}$ ,  $\{c_n\}$ , and  $W$ , this freedom propagates through the quantitative predictions.
2. *Timescale separation cannot be eliminated.* Fenichel's theorem requires  $\varepsilon_n \ll \varepsilon_{n-1}$  for each  $n$ . Without it, the slow manifold need not exist, and the reduction from  $NJ$  to  $N$  dimensions is unjustified. The nearest-neighbor topology (Theorem 4.1(iv)) also requires timescale separation.
3. *The system is not a gradient flow.* The lower-triangular  $\mathcal{J}$  is a topological obstruction (Proposition 8.5(a)). No coordinate transformation can make the system a gradient flow while preserving directed coupling.

4. *Local stability only.* Theorem 8.6 proves  $\dot{V} \leq 0$ , establishing local asymptotic stability. Global asymptotic stability requires boundary analysis depending on the specific gain functions.
5. *Symmetric weights for quantitative bounds.* General weights yield results via the secular equation, but bounds are less clean.
6.  *$O(\varepsilon)$  error in the sufficient statistic.* Fast-timescale dynamics and crisis episodes may depend on the within-level distribution.
7. *Canard duration is leading order.* Correction terms involve  $K$  through  $b$ ; amplitude behavior is less precisely bounded.

## 12.2 Empirical

1. Gain elasticities  $(\beta_1, \dots, \beta_4)$  uncalibrated.
2. Damping rates  $(\sigma_1, \dots, \sigma_4)$  uncalibrated.
3. Predictions span 2027–2040.
4. Six-stage classification simplified.
5. Crisis duration estimate requires  $\varepsilon_{\text{drift}}$ , which is itself uncertain.

## 12.3 Frameworks Considered and Rejected

*Mean field games* (Lasry-Lions): agents not exchangeable. The CES structure ( $\rho < 1$ ) ensures non-exchangeability; MFG would average over the heterogeneity that drives both efficiency results and collusion resistance.

*Minsky financial instability hypothesis*: insufficiently formalized. The Brunnermeier-Sannikov framework captures the same insight rigorously.

*Bitcoin maximalism*: the model predicts dollar stablecoins, not Bitcoin. The dollar strengthens as unit of account while the Federal Reserve loses control over dollar-denominated markets.

*Full continuous-time GE*: intractable. The four-ODE deterministic skeleton captures qualitative dynamics; a stochastic extension is deferred.

## 12.4 What the Model Does Not Predict

Smooth versus crisis transition (the canard gives duration, not character). Desirability of the synthetic gold standard. Which governments adapt. Endogenous  $\rho$ . Which specific layer binds first (depends on uncalibrated  $\beta_n$ ).

## 13. Conclusion

One Hamiltonian, one derived architecture, five results, three policy principles.

$\Phi = -\log F$  is the CES free energy—the Hamiltonian of the production technology. Its convexity on each isoquant yields the CES Triple Role (Part I): complementary heterogeneous agents are simultaneously productive, informationally robust, and competitively healthy.  $K = (1 - \rho)(J - 1)/J$  controls all three.

The same convexity forces the network architecture (Part II): aggregate coupling, directed feed-forward, nearest-neighbor chain. The architecture is derived, not assumed.

The spectral threshold  $\rho(\mathbf{K}) = 1$  (Part III) activates the hierarchy. The hierarchical ceiling (Part IV) bounds each layer by its parent. The Lyapunov structure (Part V) connects the technology  $\Phi$  to the welfare loss  $V$  through the institutional supply rate  $W$ .

Three policy principles follow from theorems:

1. *Reform upstream, not locally* (damping cancellation, Proposition 10.2).
2. *Increase gain elasticities at any layer* (global welfare ordering, Corollary 10.4).
3. *The transition takes  $O(1/\sqrt{\varepsilon_{drift}})$* —invest in the weakest cross-level link (Theorem 9.3).

Eleven predictions, spanning 2027–2040, test the theory.

## References

- [1] Li, M. Y., Shuai, Z., and van den Driessche, P. (2010). Global-stability problem for coupled systems of differential equations on networks. *J. Differential Equations* 248, 1–20.
- [2] Korobeinikov, A. (2004). Lyapunov functions and global properties for SIR and SEIR epidemiological models with nonlinear incidence. *Math. Biosci. Eng.* 1, 57–60.
- [3] Shuai, Z., and van den Driessche, P. (2013). Global stability of infectious disease models using Lyapunov functions. *SIAM J. Appl. Math.* 73, 1513–1532.
- [4] Diekmann, O., Heesterbeek, J. A. P., and Metz, J. A. J. (1990). On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* 28, 365–382.
- [5] Van den Driessche, P., and Watmough, J. (2002). Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* 180, 29–48.
- [6] Fenichel, N. (1979). Geometric singular perturbation theory for ordinary differential equations. *J. Differential Equations* 31, 53–98.
- [7] Jones, C. K. R. T. (1995). Geometric singular perturbation theory. In *Dynamical Systems* (Lecture Notes in Mathematics 1609), Springer.
- [8] Smith, H. L. (1995). *Monotone Dynamical Systems*. AMS.
- [9] Hirsch, M. W. (1985). Systems of differential equations which are competitive or cooperative. II. *SIAM J. Math. Anal.* 16, 423–439.
- [10] van der Schaft, A., and Jeltsema, D. (2014). Port-Hamiltonian systems theory: An introductory overview. *Found. Trends Syst. Control* 1(2–3), 173–378.
- [11] Shahshahani, S. (1979). A new mathematical framework for the study of linkage and selection. *Mem. Amer. Math. Soc.* 211.
- [12] do Carmo, M. P. (1992). *Riemannian Geometry*. Birkhäuser.
- [13] Neishtadt, A. I. (1987). Persistence of stability loss for dynamical bifurcations, I. *Differential Equations* 23, 1385–1391.

- [14] Neishtadt, A. I. (1988). Persistence of stability loss for dynamical bifurcations, II. *Differential Equations* 24, 171–176.
- [15] Berglund, N., and Gentz, B. (2006). *Noise-Induced Phenomena in Slow-Fast Dynamical Systems*. Springer.
- [16] Arrow, K. J., Chenery, H. B., Minhas, B. S., and Solow, R. M. (1961). Capital-labor substitution and economic efficiency. *Rev. Econ. Stat.* 43, 225–250.
- [17] Dixit, A. K., and Stiglitz, J. E. (1977). Monopolistic competition and optimum product diversity. *Amer. Econ. Rev.* 67, 297–308.
- [18] Jones, C. I. (2005). The shape of production functions and the direction of technical change. *Quart. J. Econ.* 120, 517–549.
- [19] Shapley, L. S. (1971). Cores of convex games. *Int. J. Game Theory* 1, 11–26.
- [20] Barabási, A.-L., and Albert, R. (1999). Emergence of scaling in random networks. *Science* 286, 509–512.
- [21] Pastor-Satorras, R., and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* 86, 3200–3203.
- [22] Bonabeau, E., Theraulaz, G., and Deneubourg, J.-L. (1996). Quantitative study of the fixed threshold model for the regulation of division of labour. *Proc. R. Soc. Lond. B* 263, 1565–1569.
- [23] Jones, C. I. (1995). R&D-based models of economic growth. *J. Polit. Econ.* 103, 759–784.
- [24] Romer, P. M. (1990). Endogenous technological change. *J. Polit. Econ.* 98, S71–S102.
- [25] Aghion, P., Jones, B. F., and Jones, C. I. (2018). Artificial intelligence and economic growth. In *The Economics of Artificial Intelligence*, University of Chicago Press.
- [26] Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *Amer. Econ. Rev.* 110, 1104–1144.
- [27] Baumol, W. J. (1967). Macroeconomics of unbalanced growth. *Amer. Econ. Rev.* 57, 415–426.
- [28] Grossman, S. J., and Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *Amer. Econ. Rev.* 70, 393–408.

- [29] Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica* 53, 1315–1335.
- [30] Holden, C. W., and Subrahmanyam, A. (1992). Long-lived private information and imperfect competition. *J. Finance* 47, 247–270.
- [31] Duffie, D., Gârleanu, N., and Pedersen, L. H. (2005). Over-the-counter markets. *Econometrica* 73, 1815–1847.
- [32] Dou, W. W., Goldstein, I., and Ji, Y. (2025). AI-powered trading, algorithmic collusion, and market efficiency. Working paper.
- [33] Brunnermeier, M. K., and Sannikov, Y. (2014). A macroeconomic model with a financial sector. *Amer. Econ. Rev.* 104, 379–421.
- [34] Brunnermeier, M. K., and Sannikov, Y. (2016). The I theory of money. Working paper, Princeton.
- [35] Woodford, M. (2003). *Interest and Prices*. Princeton University Press.
- [36] Lucas, R. E. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conf. Ser. Public Policy* 1, 19–46.
- [37] Uribe, M. (1997). Hysteresis in a simple model of currency substitution. *J. Monet. Econ.* 40, 185–202.
- [38] Calvo, G. A. (1998). Capital flows and capital-market crises. *J. Appl. Econ.* 1, 35–54.
- [39] Obstfeld, M. (1996). Models of currency crises with self-fulfilling features. *European Econ. Rev.* 40, 1037–1047.
- [40] Farhi, E., and Maggiori, M. (2018). A model of the international monetary system. *Quart. J. Econ.* 133, 295–355.
- [41] Caballero, R. J., Farhi, E., and Gourinchas, P.-O. (2017). The safe assets shortage conundrum. *J. Econ. Perspect.* 31, 29–46.
- [42] Triffin, R. (1960). *Gold and the Dollar Crisis*. Yale University Press.
- [43] Gorton, G. B. (2017). The history and economics of safe assets. *Ann. Rev. Econ.* 9, 547–586.

- [44] Ahmed, R., and Aldasoro, I. (2025). Stablecoins and the pricing of safe assets. BIS Working Paper.
- [45] Gorton, G. B., et al. (2022). Leverage and stablecoin pegs. NBER Working Paper.
- [46] Kauffman, S. A. (1986). Autocatalytic sets of proteins. *J. Theor. Biol.* 119, 1–24.
- [47] Hordijk, W., and Steel, M. (2004). Detecting autocatalytic, self-sustaining sets in chemical reaction systems. *J. Theor. Biol.* 227, 451–461.
- [48] Jain, S., and Krishna, S. (2001). A model for the emergence of cooperation, interdependence, and structure in evolving networks. *Proc. Natl. Acad. Sci.* 98, 543–547.
- [49] Shumailov, I., et al. (2024). The curse of recursion: Training on generated data makes models forget. *Nature* (forthcoming).
- [50] Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press.
- [51] Jones, C. I. (2015). Pareto and Piketty. *J. Econ. Perspect.* 29, 29–46.
- [52] Gabaix, X., et al. (2016). The dynamics of inequality. *Econometrica* 84, 2071–2111.
- [53] Arthur, W. B. (1989). Competing technologies, increasing returns, and lock-in. *Econ. J.* 99, 116–131.
- [54] Arthur, W. B. (1994). *Increasing Returns and Path Dependence*. University of Michigan Press.
- [55] Katz, M. L., and Shapiro, C. (1985). Network externalities, competition, and compatibility. *Amer. Econ. Rev.* 75, 424–440.
- [56] Diamond, D. W., and Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *J. Polit. Econ.* 91, 401–419.
- [57] Strogatz, S. H. (1994). *Nonlinear Dynamics and Chaos*. Addison-Wesley.
- [58] Kuznetsov, Y. A. (2004). *Elements of Applied Bifurcation Theory*, 3rd ed. Springer.