Intelligent Economics

A foundation for the last economy

Emad Mostaque

May 2026

Abstract

Economics began as *oikonomia*, the discipline of provisioning a community for the flourishing of its members. The reframing that has governed the field for a century, that economics is the science of allocating scarce means among alternative ends, set that original subject aside. It could do so because the producing community and the welfare-bearing community were the same. They are no longer. As the cost and capability of non-biological agents pass those of people across a widening set of tasks, the old question returns, and this paper proposes a foundation for the discipline that can answer it.

The foundation is a single forced structure. Bounded comparison against a reference, sustained over time and held to consistency, admits one kinematics: the choice distribution $\rho \propto \mu\, e^{V/\tau}$, its score, and a canonical relaxation to its rest point. This is the object that already appears, under other names, in rational inattention, soft reinforcement learning, variational free energy, and reinforcement learning from human feedback. It recovers central classical equations of economics exactly, as the special case in which its reference is flattened or its temperature sent to zero, extends to strategic interaction, agreement, and information aggregation by coupling agents through the inputs the chain leaves free, and identifies further fields structurally. Its distinctive economic content is one identification: the social levels of the reference measure μ are the community's *doxa*, the unspoken background of expectation that constitutes a shared world. Institutions are maintained concentrations of μ; deviation from them carries the same KL cost that is thermodynamic in the physical case; and AI alignment is doxic alignment as much as value alignment.

Written as one forced object differing only in parameters, the human and the machine make the labour transition a single moving inequality in cost and capability. The framework posits no term that would keep humans in the production story, and reserves no coordinate the human occupies that a machine cannot. With machine cost falling and capability rising, the human place in the economy can no longer be assumed to run through production at all. What remains is the reference measure itself, which turns out to be two things at once: the last scarce factor, and the condition of the community's viability, since the dynamics stays alive only while the population's distribution stays connected. The historical sequence of binding constraints, land then labour then capital then intelligence, is, in this framework, one equation peeled term by term. It ends at the reference, the outermost term, where doxa becomes the last contested factor and the constitutional question, on what basis a community provisions members who need not produce, is not the aftermath of the transition but its final inversion.

Contents

1 The problem: a discipline that set its own subject aside

The word *oikonomia* is Aristotle's. *Oikos*, household; *nomos*, governing principle. Together they name the practical discipline by which a community arranges its material life so that its members can live well. Aristotle distinguished it carefully from *chrematistike*, the art of acquiring wealth as an end in itself: one was answerable to the question of what a good human life consists in, the other only to itself. Adam Smith wrote about the wealth of nations because nations were assemblages of households whose flourishing was the implicit subject of analysis. Marshall, in his *Principles*, opened by defining economics as the study of "mankind in the ordinary business of life" and ended his life worrying that the field had drifted too far from the welfare of ordinary people [Marshall, 1890]. The original referent was stable for two millennia.

In 1932 Lionel Robbins published *An Essay on the Nature and Significance of Economic Science* [Robbins, 1932], in which he proposed that economics was "the science which studies human behaviour as a relationship between ends and scarce means which have alternative uses". The reframing was conscious. By removing the purposive content, treating "ends" as given and declining to evaluate them, Robbins could make the discipline a formal theory of choice under constraint, applicable wherever scarcity and alternative uses appeared. He knew he was setting something aside, and he thought the technical traction the move bought was worth the loss. For most of the twentieth century it was. General equilibrium theory [Arrow and Debreu, 1954], game theory, mechanism design, and the apparatus of modern macroeconomics all emerged from the reframed discipline. The question of what flourishing consists in was treated as upstream of economics, the proper business of ethics or political philosophy or simply individuals' own preferences.

The reframing worked because the implicit answer to the set-aside question stayed stable. Production and consumption happened within the same community. The labour a person performed transmitted, through wages, into the consumption that sustained their household, which sustained the community whose flourishing the discipline was implicitly serving. The chain was sociological and not part of any formal model, but it held. Economists could decline to ask what economics was for because the answer was visible in the economy itself: it was for the people who were producing it. That implicit answer is now dissolving.

In every prior wave of automation, displaced workers found new employment in sectors it had not yet reached or had newly created; the standard rebuttal to displacement worries acquired the name lump-of-labour fallacy. This time the pattern is different in a way that is visible in the cost curves. For a growing set of tasks the fully loaded cost of a human knowledge worker, including salary, benefits, office space, management overhead, training, and the friction of ordinary human variability, exceeds the cost of an AI system performing the same task at comparable or higher quality.

Call this gap Π_H, the profitability of human labour relative to its closest non-biological substitute. For most of the twentieth century Π_H was positive and large across nearly the whole economy. Through the early twenty-first century it has been falling across a widening set of task categories, toward the point where a human employee creates more cost than benefit on the relevant accounting. The labour-market mechanism that has transmitted production into consumption does not survive $\Pi_H < 0$. When the productive entities are no longer the same kind as the welfare-bearing community members, the chain that linked them breaks. Markets allocate among alternatives; they do not provision a community whose membership is no longer coextensive with the entities producing for it. Redistribution policy handles the immediate poverty problem but leaves the underlying question untouched. Who is the community whose flourishing the system is for, when producers and welfare-bearers are different sets of entities? These are the questions the older reframing set aside. They are the questions the transition forces back to the centre.

This paper proposes that economics, to address them, needs to be relocated to a substrate where they can be asked in formal terms. The substrate is available, and it is forced. Bounded comparison against a reference, sustained over time and held to consistency, admits one kinematics, derived in full in Section 3 from a short chain of uniqueness arguments: the choice distribution $\rho \propto \mu\, e^{V/\tau}$, dissipative and finite-temperature, structured by score dynamics and a KL-regularised objective. The same object has been independently arrived at in rational inattention [Sims, 2003, Caplin et al., 2019], in soft reinforcement learning [Levine, 2018, Haarnoja et al., 2018], in variational free energy [Friston, 2010], and in reinforcement learning from human feedback [Christiano et al., 2017, Ouyang et al., 2022]. Each is the same structure under a different name, in a different field, applied to a different state space. What remains, and what this paper does, is identify what the structure becomes when its state space is the configurations of valued action of a human community.

The connection to the expelled question is this: a community provisions itself for flourishing by sustaining the shared background of expectation against which its members act, and that background is precisely the reference measure the structure carries. Provisioning a community will turn out to be the maintenance of its reference; the apparatus of bounded choice is the apparatus of communal provisioning, once the reference is read as what a community holds in common. The next section shows why that apparatus is forced; the one after builds it; the one after that gives the reference its name.

On what this paper does not do, and on how it is to be judged. This paper is the formal core of a larger argument set out for a general readership in *The Last Economy* [Mostaque, 2026a], which develops the same framework under the name *Intelligence Theory*: the proposal that persistence under the second law selects for systems that build accurate, cheap predictive models of their environment, of which economics is the first application. Where the book names, this paper derives: its Intelligence Lagrangian is the action $L = H + C + K$ assembled in Section 3, and its Sorter's Law the principle that a persistent system minimises it. The book carries the historical, institutional, and human case; this paper supplies the structure beneath it. It does not propose policy: the constructive question of how a community should provision itself under the transition is left open here, posed but not answered. It does not give every subfield derivation; it works the cases that bear on the deductive line and treats the rest as inheritance. Its standing rests on the soundness of the uniqueness chain and the exactness of its recoveries. Where it recovers an existing equation, the equation's empirical record is its own; where it makes a genuinely new structural claim, for which there is no predecessor to inherit from, the support is internal consistency and the convergence of independent fields onto the same object. The paper marks which claims are which.

2 Why predictive choice is forced

The chain of the next section begins from three properties of an economic agent: that it compares configurations by value, that its capacity to do so is bounded, and that it acts persistently over time. A reader is entitled to ask why those three and not others. The answer is that they are not assumptions about agents so much as descriptions of what persists, and this section gives the argument, because the framework is stronger if its premises are consequences than if they are posits. The argument is the one *The Last Economy* [Mostaque, 2026a] develops at length under the name Intelligence Theory; it is summarised here only as far as the formal core requires.

Begin with the second law. An ordered structure, left alone, decays: the configurations that preserve its order are vastly outnumbered by those that do not, so undirected change erodes it. Persistence is therefore the thing to be explained, not the default. A structure that persists in a changing environment is doing work against this erosion, and among structures competing for the same finite resources, those that anticipate the environment rather than merely react to it spend less to persist. A system that predicts what is coming can position itself in advance; a system that cannot must pay the full cost of every disturbance after the fact. Over many rounds, the one that maintains an accurate and inexpensive model of its surroundings spends less, and the one that spends less is the one still present later. This is an observation joined to a selection argument, not a theorem, and it is the single empirical premise the framework rests on: persistence under the second law selects for systems that hold accurate, cheap predictive models of their environment. *The Last Economy* makes the case for it in full; the present paper takes it as its one load-bearing assumption about the world and marks it as such.

The three properties the chain needs follow from this. A system selected for prediction acts as though it ranks configurations, preferring those its model marks as conducive to persistence over those it marks as threatening; this preferring is what the chain calls comparison by value, and the value function V is its formal trace. The prediction is not free. Every irreversible step of it has a thermodynamic cost, bounded below by the physics derived in Section 6, so the system's modelling capacity is finite and its resolution of the best configuration is never exact; this is the bounded capacity the chain needs. And the whole argument is a claim about what survives across time, so persistence is not an added stipulation but the frame the other two are read inside. Valued comparison, bounded capacity, and action over time are thus what a persistent predictive system looks like from the outside. That is why they, rather than some other triple, are the right premises for an economics: an economy is a population of such systems provisioning themselves against decay.

Given an agent that compares by value with bounded capacity and acts over time, together with the consistency requirements stated as they are used, the structure of its choice is unique.

3 The forced structure, applied

A persistent intelligent system, a person, a firm, a community, a learning machine, pays three irreducible costs as it acts. It pays for being wrong: the value it forgoes when its choices miss what it cares about. It pays for the model it must hold to choose well: the complexity of departing from what it already expects. And it pays for changing: the work of moving from one configuration to another as it acts and learns. Write these three costs as H, C, and K. The claim of this section, and the foundation of the framework, is that the system minimises their sum,

$$L = H + C + K,$$

and that the consistent action of bounded valued choice forces this form: the value and complexity costs are forced to their specific form by requirements no alternative survives, and the kinetic cost is the one the canonical geometry induces. The state space is the configurations of valued action of a community; the value function V ranks them; the reference measure μ is what the system brings before valuing; the temperature τ is the rate at which information is paid for.

Theorem (structure of bounded-rational choice). Take an entity that faces configurations x on a state space $\mathcal{X}$, evaluates them by a value function $V : \mathcal{X} \to \mathbb{R}$, has bounded information-processing capacity, and acts persistently over time. This is the input, together with the consistency requirements the chain states as it uses them (non-smuggling at Link 1, independence at Link 2) and the technical conditions stated where they bind (measurability and normalisability at Link 2, standard regularity at Link 4). Then, with a reference measure μ on $\mathcal{X}$ and a temperature $\tau > 0$ as the only free parameters, each structure below is forced by this input, none chosen save the minimality that names the canonical relaxation:

the choice distribution

$$\rho(x) \;\propto\; \mu(x)\,\exp\bigl(V(x)/\tau\bigr);$$

the score

$$\nabla \log \rho(x) \;=\; \nabla V(x)/\tau \;+\; \nabla \log \mu(x);$$

and the variational form, in which ρ^* is the unique maximiser of $\int V\rho - \tau D_{\mathrm{KL}}(\rho\|\mu)$. The canonical dynamics realising ρ^* as its unique stationary distribution is the overdamped Langevin diffusion $dX_t = \nabla \log \rho^*(X_t)\,dt + \sqrt{2}\,dW_t$, which ascends the score; other dynamics share this stationary law, and this one is canonical in the sense of minimal: the simplest reversible diffusion whose drift is the score and whose stationary measure is the distribution.

The structure has two regimes in the temperature τ. The contractive regime ($\tau > 0$) is dissipative and bounded-rational: the dynamics on which real economic life runs, and the structure underlying rational inattention, soft reinforcement learning, variational free energy, and reinforcement learning from human feedback. The zero-temperature limit ($\tau \to 0$) is the perfect-rationality idealisation in which the choice distribution concentrates on the μ-supported maximizers of V: the structure underlying general equilibrium theory and rational expectations. The second is the boundary of the first, recovered by sending the price of information to zero.

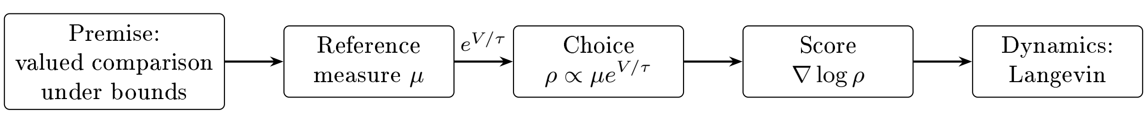

Figure 1: The uniqueness chain for bounded-rational choice. The premise forces that there is a reference measure μ over configurations; the value functional weights it exponentially at temperature τ; the score follows by differentiation; the dynamics is the canonical evolution with this score as drift and $\rho^* \propto \mu e^{V/\tau}$ as stationary distribution. What the chain forces is this form; the content of μ is supplied by the domain.

3.1 The premise

The three conditions Section 2 identified as what persistence selects for, valued comparison, bounded capacity, and action over time, now force a unique structure.

An economic agent, individual or collective, faces a space $\mathcal{X}$ of possible configurations: actions to take, allocations to choose, plans to commit to. Configurations are not equally good. A value function $V : \mathcal{X} \to \mathbb{R}$ ranks them, where V may be subjective utility, social welfare, expected profit, or any monotone aggregate of the things the agent or community is trying to bring about. The agent has bounded information-processing capacity: it cannot evaluate every configuration with infinite precision, cannot compute optimal responses instantaneously, and cannot store unbounded state. The agent acts persistently: choices are made over time, and the distribution of choices made yesterday conditions the distribution available tomorrow. These three conditions, valued comparison, bounded capacity, and persistence over time, are the premise.

3.2 The chain, in full

The chain is short, and we give it in full, because the economic reader should not have to take the central object on faith; the reader whose interest is in the structure rather than its proof can follow the boxed results and return to the derivations at need. Four links run from the premise to the dynamics, summarised in Figure 1; each forces the next, and at each step the alternatives fail a stated consistency requirement.

Link 1: bounded capacity forces a distribution, not a point. An agent that could evaluate every configuration with unbounded precision would choose a single optimum, a point mass at $\arg\max V$. An agent with bounded capacity does not know which configuration is best. Concentrating on one, as a deterministic bounded heuristic does, adds the assumption that it is the best, which is structure the constraints do not provide: the values the heuristic evaluated do not say the best seen is the best overall, and treating it as such smuggles what has not been learned. The non-smuggling requirement forces the response to reflect the actual uncertainty: a probability distribution ρ over $\mathcal{X}$, weighted by what the agent knows and no more. Nothing here legislates the agent's internal representation, which may be a point estimate, a set of distributions, or anything else. What the premise constrains is choice behaviour sustained over time, and persistent choice under uncertainty realises a single distribution over configurations. ρ is that realised law, the object the next links constrain, whatever internals produce it. Deterministic bounded rules exist; what rules them out is not boundedness but consistency, the same requirement that does the forcing at every later link, since acting on a certainty the information does not contain is the choice-level form of asserting an unearned premise. The requirements are constitutive rather than empirical: they assert nothing about the world but state what consistency means for bounded choice, as Shore and Johnson's axioms state what it means for inference [Shore and Johnson, 1980], and the theorem is that the standard so defined has exactly one realisation. An agent that declines the standard is not a counterexample; it is a different subject, and Link 2 names the branch it exits onto. This is the content of rational inattention [Sims, 2003, Caplin et al., 2019]: the choice is stochastic because attention is scarce and concentrating beyond what scarce attention warrants would assume what has not been earned.

Link 2: independence forces the exponential form. The agent does not begin from nothing. It brings a reference distribution μ, the distribution that would obtain before the current value computation, and the value function V tilts it. Write the tilt as $\rho(x) = \mu(x)\, f(V(x))/Z$: the choice density relative to the reference is some function of value. This form follows from the division of labour the premise already made: μ carries everything the agent brings to the choice, V is the complete current evaluation, so the update of the one by the other can depend on x only through $V(x)$. Any residual x-dependence would be one of three things: prior structure, which belongs in μ; evaluative structure, which belongs in V; or structure contained in neither, which is smuggled, excluded by the same requirement as at Link 1.

One requirement fixes f. When a configuration splits into independent components, $x = (x_1, x_2)$ with the reference factorising, $\mu = \mu_1\mu_2$, and value additive, $V = V_1 + V_2$, an agent that exploits the independence chooses so that the choice over the whole is the product of the choices over the parts,

$\rho = \rho_1\rho_2$, since correlating genuinely independent sub-choices would forgo value for no gain. This requirement, that the choice rule respect independence where it holds, is a consistency requirement, not a premise about the world: an agent that correlates genuinely independent sub-choices wastes its bounded capacity for no gain in value. It forces

$$f(V_1 + V_2) \;=\; f(V_1)\, f(V_2),$$

the Cauchy exponential equation, whose only measurable non-trivial solution is $f(V) = e^{V/\tau}$ for a constant τ. The step is Luce's choice axiom [Luce, 1959], the independence condition that underlies McFadden's logit and the discrete-choice tradition built on it. What is new is the cascade it initiates once its dual face (Link 3) and its dynamics (Link 4) are assembled and the reference μ is given its economic reading. Relaxing the axiom yields different mathematics: replacing KL with a Rényi divergence in the variational objective produces escort distributions (the q-exponentials of Tsallis statistics) rather than the Boltzmann form, so the results that follow are specific to the independence branch, and the axiom is load-bearing. Before Link 2, V records the ordering induced by valued comparison; Link 2 fixes the additive cardinal representative of that ordering on independent product components, up to the scale τ. Hence

$$\rho^*(x) \;=\; \frac{1}{Z}\,\mu(x)\,\exp\bigl(V(x)/\tau\bigr), \qquad Z = \int \mu\, e^{V/\tau},$$

with $\tau > 0$ the one free scale, fixed by normalisability so that higher value is more probable.

The exponential is not the Gibbs *choice*; it is the unique density that turns additive value over independent components into multiplicative choice, exactly as the additivity of a local action forces the exponential measure in statistical field theory. This same object, written $\max_\rho \mathbb{E}_\rho[V] - \tau D_{\mathrm{KL}}(\rho\|\mu)$, is the objective of reinforcement learning from human feedback [Christiano et al., 2017, Ouyang et al., 2022] and of soft reinforcement learning [Levine, 2018, Haarnoja et al., 2018], with τ the regularisation weight and μ the reference policy.

Link 3: the information cost is the dual of the same object. The temperature τ has a reading the form makes precise. The distribution ρ^* is exactly the one that, among all distributions achieving a given expected value $\mathbb{E}_\rho[V]$, sits closest to the reference in Kullback–Leibler divergence,

$$D_{\mathrm{KL}}(\rho\|\mu) \;=\; \int \rho\, \log \frac{\rho}{\mu},$$

with τ the multiplier on the value constraint: minimising $-\int V\rho + \tau D_{\mathrm{KL}}(\rho\|\mu)$ returns ρ^* by a one-line Lagrange computation. So the complexity of a choice, the cost of departing from what is expected, is measured in KL, and τ is the price of that departure in units of value. That KL is the right measure of this cost is independently confirmed by the consistency axioms of Shore and Johnson [Shore and Johnson, 1980, Caticha, 2004], which single it out among divergences by the same additivity over independent systems that forced the exponential above; here KL is not a premise but the dual face of the object Link 2 already produced. The free energy $-\int V\rho + \tau D_{\mathrm{KL}}(\rho\|\mu)$ is what the next link descends.

Link 4: persistence forces a relaxation, and the score is all it needs. A static distribution is not yet a dynamics. Configurations change over time, and a persistent agent revises its choices, so the theory needs an evolution that carries an arbitrary distribution toward ρ^*. The canonical evolution is determined by the static structure itself, because the object that drives it is already in hand: the *score*, the gradient of the log-density,

$$\nabla \log \rho^*(x) \;=\; \frac{\nabla V(x)}{\tau} \;+\; \nabla \log \mu(x).$$

An agent that ascends the score follows the overdamped Langevin diffusion

$$dX_t \;=\; \nabla \log \rho^*(X_t)\, dt \;+\; \sqrt{2}\, dW_t,$$

whose unique stationary distribution, on a Euclidean state space under standard regularity (smoothness of the drift, confinement, and ergodicity), is $\rho^* \propto \mu\, e^{V/\tau}$. (In discrete time the Euler–Maruyama step $x_{t+1} = x_t + \varepsilon \nabla \log \rho^*(x_t) + \sqrt{2\varepsilon}\, \xi_t$ approximates this dynamics and converges to it as $\varepsilon \to 0$; it is this discretisation that score-based generative models run in practice. On a finite or discrete state space the diffusion is replaced by a continuous-time Markov chain with Gibbs transition rates, the logit or Glauber dynamics, whose stationary distribution is the same ρ^*; the finite-state recoveries below, logit and quantal response, inherit this rest point without the gradient machinery.) No distribution-space metric need be chosen: the drift is fixed by ρ^*, which Links 1–3 already forced. The kinetic cost of changing, the work of moving from one configuration to another, is the noise-and-step structure of this relaxation; it is real, and it is what persistence adds, but it introduces no new forced object beyond the score.

The dynamics of bounded-rational economic choice and the dynamics by which a diffusion model generates samples are the same equation on different state spaces: both are the relaxation whose rest point is the distribution. The chain closes on itself: the score that defines the dynamics is the gradient of the distribution the static argument already produced.

That is the whole derivation, and it assembles into the action. The two costs of static choice are each forced to a definite form: the potential cost of being wrong is $H = -\int V\rho$, the value forgone under the choice distribution, and the complexity cost of departing from the reference is $C = \tau D_{\mathrm{KL}}(\rho\|\mu)$, the dual face of the exponential established in Links 2–3.

Persistence adds a third, the kinetic cost of changing, K, and the chain pins its informational component too. The reference has a reference *dynamics*: Link 4 applied to μ alone, the relaxation $dX_t = \nabla \log \mu(X_t)\, dt + \sqrt{2}\, dW_t$ of the agent before the value computation; Brownian motion in its place would smuggle a flat doxa. The cost of acting on the value is the divergence of the agent's path law P from this reference path law Q_μ, both started from a shared initial configuration, and no path-space divergence need be chosen: Link 3 already forced KL, and applied to path measures it is evaluated by Girsanov's theorem, under the regularity of Link 4 with Novikov's condition. The drifts differ by exactly $\nabla V/\tau$, the doxic term cancelling, so

$$K_{\text{info}} \;=\; D_{\mathrm{KL}}(P \,\|\, Q_\mu) \;=\; \tfrac{1}{4}\, \mathbb{E}_P \int \left\| \nabla V(X_t)/\tau \right\|^2 dt \,:$$

the cost of changing is the integrated squared value gradient at the price of information; steep landscapes are expensive to act on, flat ones nearly free. The correspondence with generative modelling closes at the level of training as well as sampling: the same path divergence between diffusion processes is the maximum-likelihood training objective of score-based diffusion models [Song et al., 2021], and a value gradient added to a reference drift is the guidance mechanism by which those models are steered. The physical price per unit of this motion remains domain input: the actuation multiplier whose floor Section 6 derives; results below that lean on the physical component, the Coase reading among them, remain identifications at the level of structure, while search and actuation inherit the integrand directly.

Their sum is the action of a persistent intelligent system,

$$L \;=\; \underbrace{-\int V\rho}_{H,\ \text{being wrong}} \;+\; \underbrace{\tau D_{\mathrm{KL}}(\rho\|\mu)}_{C,\ \text{model complexity}} \;+\; \underbrace{K}_{\text{changing}}\,.$$

A system choosing once, with no time to pass, minimises only $H + C$, the free energy, whose minimiser is the Gibbs distribution of Link 2, and this is where the static results of the paper live. A system acting over time descends this free energy by the Langevin relaxation, which carries it to that same Gibbs distribution. Static choice is $H + C$; persistent choice adds K; what is forced without qualification is H, C, the score, and the destination ρ^* the relaxation reaches. The paper's results rest on these.

> **The action of intelligence.** The object is the action $L = H + C + K$: three *kinds of cost*, for being wrong, for model complexity, and for changing, summed in a single scalar. This is the foundation from which the rest of the framework descends, the principle *The Last Economy* [Mostaque, 2026a] names Sorter's Law: a persistent system minimises the sum of its predictive error, the complexity of its model, and the cost of updating it. Its static part is the objective minimised in reinforcement learning from human feedback and in active inference, and the Langevin relaxation above is the same sampler used in score-based generative models.

3.3 The measure

The reference measure μ is the distribution over configurations that would obtain in the absence of the value function. In machine learning it is the reference policy the current stage regularises against: the pretrained base for fine-tuning, the supervised checkpoint for RLHF, each stage's μ being the previous stage's output. In rational-inattention models it is the prior before information acquisition; in active inference, the agent's generative model of preferred states. Across all three it carries what the agent brings to the choice before the current value computation. The value function bends the choice distribution away from this reference; the temperature τ controls how sharply.

The economic interpretation of μ is the subject of Section 4. One distinction governs how it should be read. The chain forces the *form* of the structure: that there is a reference measure, that value tilts it exponentially, that the dynamics is gradient flow. It does not force the *content* of the slots: what μ, V, and τ are in a given setting is supplied by the domain, not the theorem. Form claims are forced and carry everywhere; content claims must be earned case by case and can be wrong. The identification of μ with doxa, in the next section, is a content claim, and the reader should hold it to that standard. For the present section it is enough that μ exists, is forced by the structure, and is the free function alongside V and τ that distinguishes one community's economic life from another's.

3.4 The score

The score $\nabla \log \rho(x) = \nabla V(x)/\tau + \nabla \log \mu(x)$ decomposes cleanly into two gradients. One, $\nabla V/\tau$, is the value gradient scaled by inverse temperature: agents move toward higher value, more sharply at low temperature. The other, $\nabla \log \mu(x)$, is the doxic gradient: agents are pulled toward higher concentrations of the reference measure, regardless of value. Real choice is the sum. This is why behaviour systematically departs from pure value-maximisation in directions that look "irrational" to a model that ignores the second term: people do what is valuable and what is expected, and the weighting between the two is set by τ.

The score is also the object that connects the framework to the empirical record of modern machine learning. Score-based diffusion models [Song and Ermon, 2019, Lipman et al., 2023] learn $\nabla \log \rho$ from data and use it to generate samples by integrating the corresponding stochastic or deterministic dynamics. They are training the structure of bounded-rational choice, applied to the state space of natural data. The same operator appears in both because it is the same object: the score of the distribution, deployed in two different domains.

3.5 The variational objective

The dynamics was given in Link 4: agents follow the Langevin relaxation that ascends the score and settles at $\rho^* \propto \mu\, e^{V/\tau}$. The static face of this is a variational principle, and the form in which the rest of the discipline meets the same object.

The variational objective. The free-energy functional $\mathcal{F}[\rho] = -\int V\rho + \tau D_{\mathrm{KL}}(\rho\|\mu)$ is the variational form of the structure. When the partition function $Z = \int \mu\, e^{V/\tau} < \infty$ and the optimisation runs over $\rho \ll \mu$, its unique minimiser is the Boltzmann choice distribution $\rho^* \propto \mu \exp(V/\tau)$, the rest point of the Langevin relaxation. Written as $\max_\rho \mathbb{E}_\rho[V] - \tau D_{\mathrm{KL}}(\rho\|\mu)$ it is the objective of reinforcement learning from human feedback [Christiano et al., 2017, Ouyang et al., 2022], of soft reinforcement learning [Levine, 2018, Haarnoja et al., 2018], of rational-inattention models [Matějka and McKay, 2015, Caplin et al., 2019], and of active inference [Friston, 2010]. It is the same functional applied to different state spaces and named differently in different disciplines. The correspondence is exact at the optimum as well: the known closed-form solution of the RLHF objective, the policy proportional to the reference policy times the exponentiated reward, is the object $\rho^* \propto \mu\, e^{V/\tau}$ itself, with μ the pretrained base model, V the reward model, and τ the regularisation weight. The aligned systems now entering the economy are trained to this objective, and at its optimum satisfy the framework's equation by construction: the theory of the economy and its newest participants share an equation, and the alignment problem and the economic problem meet in the same two objects, V and μ. Table 1 states the correspondence slot by slot; Table 2 reads the same correspondence the other way, down the generative stack: each engineering artifact, named as the economic object it instantiates.

	This paper (economics)	**Rational inattention**	**Soft RL / control**	**Active inference**	**RLHF**
ρ	choice distribution	stochastic choice	policy π	posterior belief	aligned policy
μ	doxa (social ref.)	prior / default	reference policy	generative prior	pretrained base
V	value function	payoff	reward / Q	log-evidence	reward model
τ	information price	attention cost	entropy weight	precision^{-1}	KL coefficient β
$\rho^* \propto \mu e^{V/\tau}$	choice	logit solution	soft-optimal policy	free-energy min.	DPO/RLHF optimum
$\tau \to 0$ limit	perfect rationality	full attention	greedy / max-Q	precise inference	reward over-optim.

Table 1: One object under five names. Each column is a field that arrived at the structure independently; each row is a slot of the structure; each cell is an identification at the level of the shared functional. The economics of this paper and the training objective of aligned AI systems are the same functional on different state spaces, which is why the alignment problem and the economic problem meet in the same two objects, V and μ.

One property of this optimised functional recurs throughout the paper. By the envelope theorem, the derivative of an optimised value with respect to a parameter is the direct partial derivative alone; the optimiser's own response drops out to first order, because at the optimum the objective is flat in the choice variable. The temperature τ is then the envelope derivative of the optimised expected value with respect to the information constraint, which is what makes it a shadow price; the demand system is the envelope derivative of the optimised log-partition function with respect to

In generative AI	In this framework	Where
Pretrained base model	the reference measure μ, the doxa	§3, §4
Reward model	the value function V	§3
KL coefficient β	the temperature τ, the price of information	§3
RLHF/DPO closed-form optimum	the law $\rho^* \propto \mu\, e^{V/\tau}$	§3
Softmax output at temperature	the choice law on a finite space	§3
Score network $s_\theta \approx \nabla \log \rho$	the score $\nabla V/\tau + \nabla \log \mu$	§3
Denoising sampler (Euler–Maruyama)	the canonical Langevin relaxation	§3
Diffusion training loss (score matching)	the kinetic cost K_{info}, the path KL	§3
Classifier guidance	a value gradient added to a reference drift	§3
Adversarial training and its cycling	the circulating component of non-reciprocal values	§5
Reward over-optimisation	the $\tau \to 0$ limit	Table 1
Training stack (pretraining, tuning, RLHF)	the layers of the doxic hierarchy	§4

Table 2: The generative stack, read as economics. Each row pairs an engineering artifact with the economic object it instantiates, with the identification established at the indicated location. The two columns share their equations because the systems on the left were built by optimising the objectives, and following the gradient fields, that the right column derives.

prices, as Section 5 shows; and the capability of an agent in Section 6 is the optimised value $\mathbb{E}_\rho[V]$ it attains. These are one theorem read three times.

3.6 The two regimes: bounded and perfect rationality

The structure has two regimes in the temperature parameter. At $\tau > 0$ it is contractive: the dynamics is dissipative, the choice distribution has positive entropy, and information is paid for at the rate τ per nat of KL deviation from μ. This is the regime in which real economies, real machine learning systems, real animals, and real human cognition operate. At $\tau \to 0$ the structure collapses to its zero-temperature limit: the choice distribution concentrates on the μ-supported maximizers of V, the KL cost vanishes, the dynamics is dominated by the value gradient and the agent is the perfect optimiser of classical economic theory. The limit is the idealisation; the contractive regime is the empirics.

The relation between them is a single parameter. As the price of information τ falls to zero, the bounded-rational agent, which samples the value-weighted reference distribution with noise, approaches the perfect-rationality boundary, where the choice distribution concentrates on the value-maximising configurations within the reference's support. Perfect rationality is not a different theory from bounded rationality; it is its zero-temperature boundary. The two are not rivals but the limits of one structure, and the contractive regime is primary because it is where economic life happens, the boundary being reached only in the idealised limit where deliberation is free.

The mainstream and the heterodox. Neoclassical economics, with its rational agents, frictionless markets, and convergent equilibria, is economics at the zero-temperature limit of the structure, $\tau \to 0$. The major heterodox traditions, behavioural economics, rational inattention, complexity economics, agent-based modelling, and the accounts of bounded rationality from Simon onward, are partial descriptions of the contractive regime at $\tau > 0$. The framework does not adjudicate between them; it locates them. Each is doing economics on a different part of the same object: the mainstream on the zero-temperature boundary, the heterodox in the finite-temperature interior. The boundary is a limit of the interior, not an alternative to it.

4 What the framework is: economics as choice under doxa

The structure of the previous section is a substrate. What makes it specifically economic is what fills its slots, and one slot carries the weight: the reference measure μ, the shared reference on which the rest of the paper converges.

4.1 The claim

The foundational claim of intelligent economics. Economics, foundationally, is the dynamics of bounded-rational choice on a value manifold under a community's doxa. The kinematic structure of this dynamics is forced, by the chain of Section 3.2, and has two regimes in the temperature τ. In the contractive regime ($\tau > 0$) the dynamics is dissipative and finite-temperature: the choice distribution is $\rho \propto \mu \exp(V/\tau)$, agents follow the Langevin relaxation that ascends the score and settles at it, and the structure is the same one found in rational inattention, soft reinforcement learning, active inference, and reinforcement learning from human feedback across cognitive science, machine learning, and behavioural economics. In the zero-temperature limit ($\tau \to 0$) it collapses to the classical perfect-rationality case of general equilibrium theory and rational expectations.

The framework's specifically economic content is the identification of the social levels of the reference measure μ with the community's *doxa*: the unspoken background of expectation that constitutes a shared social world, the upper part of a measure constructed hierarchically by successive maximum-entropy projection from raw possibility through symmetry, salience, institutions, and history. Institutions are maintained concentrations of μ. The cost of deviation from doxa is the KL term in the variational objective, the same cost structure that is thermodynamic in the physical-computation case and carries the same formal properties in the social one: deviation is costly, maintenance requires work, and neglect leads to decay. Flourishing in the Aristotelian sense is read, on this identification, as the sustained maintenance of the community's doxic structure through productive activity that reproduces it; this reading is a content claim, not part of the grammar.

4.2 What doxa is

The structure has a reference measure μ that the structure itself does not fill in, and the lower levels of μ are pre-social, fixed by the world rather than by any community: the bare possibility space, the symmetries physical law imposes, the focal points perception makes salient. Its upper levels are different. They carry the economic content, the institutional and historical part, and they are filled by something with a proper name: the community's *doxa*, the unspoken background of expectation, taken-for-granted assumption, and pre-reflective orientation that constitutes a shared social world. The reference measure is the formal object; doxa is what fills its social levels. No other discipline using the same substrate takes the maintenance of a shared social world as its object, and that identification is the paper's distinctive contribution.

The word is Greek. Plato used it for opinion, in contrast to *episteme*, knowledge; the contrast set up most of subsequent epistemology. The use here is the sociological one developed by Pierre Bourdieu [Bourdieu, 1977], where doxa names something more basic than either opinion or knowledge: the structure of the unspoken, the background of assumption against which both opinion and knowledge are formed. Doxa is what nobody argues about because the disagreement would not register. It is what a community takes for granted so deeply that the taking-for-granted is invisible to its members.

Doxa is not orthodoxy. Orthodoxy is what a community consciously affirms; doxa is what the affirmation rests on. Orthodox doctrine in a religious tradition is what the catechism says; doxa is the structure of expectations within which the catechism makes sense as the kind of thing one might assent to. Doxa is not heterodoxy either. Heterodoxy is conscious dissent from orthodoxy; doxa is the substrate both orthodox affirmation and heterodox dissent share. A heterodox thinker disagrees with what the orthodox think; both agree, prereflectively, on what kind of question is at stake and what would count as an answer. That shared agreement is doxa.

Doxa is not preference, not constraint, not knowledge. Preferences are what individuals want; doxa is the background that makes the wanting feel sensible. Constraints are what is physically or institutionally possible; doxa is the background that makes the constraints feel binding rather than arbitrary. Knowledge is what is consciously believed and can be articulated; doxa is the background that makes the articulation cohere. Each of these familiar economic categories presupposes a doxic substrate without naming it. The framework names it.

The mathematical formulation makes the identification precise. In the choice distribution $\rho(x) \propto \mu(x)\exp(V(x)/\tau)$, the agent's behaviour is the product of two factors: the value-weighted preference, $\exp(V/\tau)$, and the doxic prior, μ. The first is what the agent wants; the second is what the agent expects, pre-reflectively, as a sensible thing to be doing. The score decomposition $\nabla \log \rho = \nabla V/\tau + \nabla \log \mu$ shows that real choice is the sum of value gradient and doxic gradient, with the relative weight set by τ. At high temperature, the doxic prior dominates: agents do what is expected because computing what is valuable is too costly. At low temperature, the value gradient dominates: agents compute and pursue what is valuable, with the doxic prior providing only the starting point. The empirical mix in any real community is somewhere between. For the ML reader, the closest analogue is the pretrained model's implicit distribution over outputs before any reward signal, but doxa is that same object at the scale of a community, carried not in weights but in the unspoken expectations its members share.

4.3 The stocks the structure can carry

The action $L = H + C + K$ decomposes the *costs* a system pays. A parallel question is what *stocks* a community accumulates to lower those costs and raise achievable value. The object answers it by reading the stocks off the three cost terms directly: each term is lowered by a different kind of capital, and the four standard capitals fall out of the three terms once one term is modelled as taking two distinct inputs.

The value term $H = -\int V\rho$, the cost of being wrong, is lowered by an accurate model of the value landscape, and accuracy has two independent ingredients that do not substitute. One is *material* capital, the data and the physical substrate from which the world is read and acted on; the other is *intelligence* capital, the patterns and models that turn data into prediction. Material capital without intelligence is a warehouse of uninterpreted facts; intelligence without material is ungrounded speculation. Both reduce H, and because neither substitutes for the other, the value term takes two inputs where the other two terms take one. This is why there are four capitals and not three.

The kinetic term K, the cost of changing, is lowered by *network* capital, the connective structure through which a community updates: the shared channels, relationships, and institutions that let a correction propagate without each agent paying the full cost of moving alone. The complexity term $C = \tau D_{\mathrm{KL}}(\rho\|\mu)$, the cost of fragility under a wrong or brittle reference, is lowered by *diversity* capital, the spread of strategies and references the community carries. A monoculture minimises C in the expected case and pays unboundedly when the unexpected arrives; diversity holds a portfolio that keeps C finite across states.

The four capitals are the three cost terms, read as stocks. The capitals that lower the action map onto its terms. Material and intelligence capital lower H, the cost of being wrong, as the two non-substituting inputs to an accurate model, data and patterns. Network capital lowers K, the cost of changing, as the connective structure through which updates propagate. Diversity capital lowers C, the cost of fragility, as the spread that keeps the complexity term finite when conditions shift. Achievable value requires all four: with no substrate there is nothing to read or act on, with no patterns nothing can be predicted, with no connection nothing can be coordinated, and with no diversity the system is fragile to the first shock outside its model, so any one missing is fatal regardless of the others. On this reading the multiplicative four-capital picture is the three cost terms counted as the stocks that lower them, with one modelling step, that accurate prediction takes data and patterns as distinct inputs, and with the complementarity that each at zero zeroes the whole following from the fact that each removes a different irreducible obstacle. The step is smaller than it looks: the object is already a product of two factors, the reference and the tilt, and both zero-laws are theorems above: a truth the reference excludes is unrecoverable at any evidence, and a constant V leaves the choice at μ exactly, no prediction beyond what was already carried. What the step supplies is the names: data for what feeds the tilt, patterns for what the reference carries. The value function V is not among them: V is what the system is trying to bring about, the end rather than a means, and one does not accumulate an objective, one pursues it with the capitals the structure provides.

These four are the measurable expression of three constraints the next subsection states as laws: material and intelligence capital serve the conservation and circulation of value, network capital serves openness, and diversity capital serves resilience. The stocks are what a community holds; the laws are the conditions under which holding them keeps the dynamics alive.

4.4 The conditions of viability: flow, openness, resilience

The capitals say what a community accumulates. A separate question is what a community must satisfy for its dynamics to persist at all, and the structure answers it with a single condition seen along three axes. The dynamics has a single attractor the whole population reaches, under the connected-support and regularity conditions of Link 4, only while the free energy remains finite and bounded below; it loses that attractor when the free energy runs away, and there are three principal modes in which it can. Each mode is the failure of one cost term to stay bounded, along a different axis: the value term unbounded in the population's stock of value, the complexity term divergent across the population's members, the complexity term divergent across the states of the world. These three are not moral premises imported from outside economics. They are the one viability condition, the free energy finite and bounded below, read three ways, and the corpus names them the laws of flow, openness, and resilience.

Flow. The value term must be conserved and circulated. A system that consumes its capital without accounting for it lets its model of the world diverge from reality: the error in H grows without bound, and the drift loses its rest point. A system that conserves value but does not circulate it stops generating the data that tests its predictions, and its model goes stale. Formally, flow is the condition that the value term stay bounded; its economic reading is that value keep moving through the population rather than pooling out of the dynamics, and it governs the material and intelligence capital that feed the value term.

Resilience. The complexity term must stay bounded across the states of the world, which requires diversity. A community whose references have collapsed to one is a monoculture: minimal C in the expected case, unbounded C when an unmodelled state arrives and the single shared prior is the wrong one. A population that instead holds a spread of references, a mixture, pays at worst the log of the weight on the component matched to the realised state, a finite price no matter how surprising that state, where the monoculture pays without bound. Resilience is the condition that C stay bounded across states, and it is what diversity capital secures.

Openness. The third axis is the one the dynamics states most sharply, because it is a condition on the connectedness of the distribution across the population, and it is where a result usually treated as a separate moral commitment turns out to be a viability requirement.

Openness is the connectedness of the distribution. For the gradient flow to carry the whole population to one attractor, ρ must remain a single connected object, and connectedness can fail in two ways, each a way of closing the system off. Exclusion is absorbing in the exact case: the support of the stationary target $\rho^* \propto \mu\, e^{V/\tau}$ is the support of μ, so where the community's reference assigns a subpopulation's configurations zero probability, the relaxation defined on that support cannot create mass there and the excluded cannot be drawn back. Once the doxa stops recognising a subpopulation, the exclusion is self-reinforcing; a merely small but positive reference does not close the system off, which is the formal content of the difference between neglect and exclusion. Detachment is divergent: where a group's realised distribution leaves bounded KL of the shared reference, its free energy diverges and it falls out of the community's attractor. A society that drives part of its population below viability, or lets part detach from the shared world, has not committed an injustice named from outside the framework; it has fragmented its distribution into disconnected components and lost the ergodicity that let one dynamics carry everyone. The Law of Openness, formally, is that the population remain one connected distribution relaxing to one attractor, and it is the same requirement the corpus reads as the condition that connection fights entropy: a closed-off subpopulation, whether excluded or detached, is a part of the system whose free energy is no longer bounded and whose distribution is no longer maintained.

The three laws are therefore one viability condition read in three modes. Under the doxa reading, they are what the corpus calls *fairness as an operational requirement*: the conditions a community must meet to keep its dynamics alive are among the conditions usually argued for on moral grounds, so the failures moral traditions condemn first, exclusion above all, are failures of viability as well. Flow secures *dignity*, the ability to meet one's needs without falling below viable support; openness secures *capability* in the sense of the welfare tradition, the freedom to move toward higher value that the kinetic term affords a member who remains connected; resilience secures *viability* proper, the system's survival of shocks it did not model. This is how the framework recovers the distributional concerns of welfare economics and the capability tradition as viability requirements, without importing them as axioms. Each is a face of the one condition that the dynamics reach a single attractor, and a society can be measured against it by asking whether its free energy remains finite and bounded below along all three axes. Here they are the one survival condition of the dynamics, read once per cost term.

The claim, the doxa it rests on, the capitals it accumulates, and the conditions under which it stays viable are what the framework *is*.

4.5 How doxa is built and maintained

Doxa is constructed, not given. It is built in layers running from raw possibility to particular cultural form, each layer the maximum-entropy distribution compatible with the constraints of all earlier ones, so each is the KL-minimal update of the last under new constraints. The layers run from the bare configuration space, through the symmetries physical law imposes and the salience perception supplies, to the institutions a community consciously builds and the history it accumulates. Doxa proper is the upper part of this stack, the institutional and historical content; the lower layers are the pre-social scaffolding the world supplies before any community acts. The framework's contribution is not the observation that doxa is layered, which is familiar from Bourdieu and others, but the formal claim that the layering is the iterated application of one maximum-entropy projection, each layer adding the minimum information needed to encode its constraints.

The same construction is visible in the engineering of foundation models. Pretraining installs the linguistic and cultural scaffolding, instruction tuning installs the institutional layer of how to respond, and reinforcement learning from human feedback installs the refined contextual layer. The anchoring is distributed as the doxic hierarchy distributes explicitness, implicit at depth and maintained at the surface: each deep layer is held by being what every later stage starts from, while the surface stage is anchored by an explicit KL-regularisation against the reference model, the very projection that defines the hierarchy. What AI laboratories have been refining computationally is hierarchical doxa engineering, named differently because it has not been recognised as such.

This gives a single ontology for the major economic objects: each is a maintained concentration of μ in some domain, with a mechanism that resists drift. Money is a concentration in the medium-of-exchange domain, property in the resource-ownership domain, a firm in the production-organisation domain, a market in the coordination-mechanism domain; each is the same operation applied to a different slice of the configuration space, which is the formal substrate institutional economics from Veblen through North [North, 1990] has wanted. The concentration requires maintenance because the same informational logic that drives physical entropy applies to the reference: unattended, μ relaxes back toward its scaffolding, so institutions decay unless work is done to sustain them, the work of enforcement, ritual, reproduction, and narrative, all of which pay the cost of holding a concentration of μ against drift. Flourishing in the Aristotelian sense is read, on this structure, as the state in which a community's activity sustains its doxic concentrations against that drift. A community whose institutions are maintained and whose shared expectations are reproduced is provisioning itself in the original sense of oikonomia; one whose μ is decaying faster than it is renewed is at structural risk whatever its conventional indicators show. This reading is a content claim, an interpretation of eudaimonia on the structure, not part of the grammar.

Two consequences of the structure carry implications sharp enough to box, both concerning alignment.

> **RLHF is doxic alignment.** The KL-regularised objective $\max_\rho \mathbb{E}_\rho[V] - \tau D_{\mathrm{KL}}(\rho \| \mu)$ that defines reinforcement learning from human feedback is the same variational form as the choice distribution. The pretrained reference policy is the doxic prior μ; the KL-regularisation term is the deviation cost of moving away from doxa; the temperature τ is the inverse rationality scale. RLHF is doing economic optimisation on the contractive form of the structure, with doxa as the reference; the only thing that has not been named is what the reference is.

The two-factor alignment requirement. A bounded-rational agent's behaviour is set by the score $\nabla \log \rho = \nabla V/\tau + \nabla \log \mu$, so behavioural alignment with a human community requires both terms to be compatible with that community's: value alignment for the first, doxic alignment for the second. The standard frame addresses only V, ensuring the AI wants what humans want. But an AI with aligned V and misaligned μ passes every conventional benchmark, its stated values matching, while its behaviour pulls systematically toward the directions the second term sets and away from the doxic concentrations the community maintains. Doxic misalignment shows up not in stated values but in unstated assumptions, what counts as obvious and what would never need to be said, exactly the failures hardest to detect because they live below the level of articulation. The framework names μ as a first-class alignment target alongside V, and makes the choice of which reference, by which procedure, at which level of the hierarchy, a question that can be precisely posed rather than handled implicitly.

5 Classical economics at the limits of this object

The claim that economics is choice under doxa on the structure has to answer for the economics that already exists. The answer: the classical results are not rivals to the forced object but special cases of it, each recovered at the limit its field assumed.

The recoveries are of two kinds, and the difference matters. Some reproduce a field's equation outright: the structure, specialised, returns the same formula the field derived by its own route, and the agreement is a check on both. Others identify a field's central object with one the structure already carries, claiming the two are the same thing seen twice. An identification of the second kind is easy to make and easy to abuse, because any sufficiently general formalism can rename a neighbouring field's primitives in its own vocabulary and call the renaming a unification. The test that separates a contentful identification from an empty one is *transfer*: an identification earns its keep when the structure's remaining machinery, the score, the temperature, the value-reference split, the cost terms, then carries over and yields a constraint the source field did not itself build in. A renaming transfers nothing and predicts nothing; a real identification imports the rest of the apparatus and says something new about the field it lands in. The identifications below are claimed at the level of transfer, and each is stated so that what it imports, and what new constraint that yields, is visible. That the object is an exponential tilt, a form ubiquitous across the sciences, is therefore not a weakness of the unification but its mechanism: bounded valued choice leaves the same signature wherever it occurs, and the content is in the slots the economic reading fills, not in the exponential form it shares with statistical physics.

The choice distribution $\rho \propto \mu\, e^{V/\tau}$ is the unique exponential tilt of a reference measure fixed by a constraint, and the classical equations of economics are what this one object becomes under the particular reference and limit each field assumes. With a linear value tilt $V(x) = -p \cdot x$, the mean choice is the gradient of the log-partition function, $x^*(p) = -\tau\, \nabla_p \log Z$, which is the demand system; its response to prices is the compensated price-response matrix (the Slutsky matrix of consumer theory), $\partial x^*/\partial p = -\tau\, \nabla_p^2 \log Z$. Both textbook properties then follow in one stroke: the matrix is symmetric because mixed partials of a scalar commute, and negative semidefinite because $\log Z$ is convex in p as a cumulant generating function, its Hessian is positive semidefinite, and the leading $-\tau$ carries it to negative. Slutsky symmetry and negativity, the compensated substitution structure of demand theory, are thus the single fact that $\log Z$ is a convex cumulant generating function, with no utility posited; income effects live in a richer V that encodes the budget, and are not claimed here.

The logit is the same object on a finite uniform support, with τ the information price rather than a fitting constant. The Euler equation of dynamic optimisation, the intertemporal first-order condition that governs saving and investment, is recovered by the same stroke: an interior mode of ρ^* satisfies $\nabla \log \rho^* = 0$, that is $\nabla V = -\tau \nabla \log \mu$, so when V is a discounted path objective carrying the budget through its price tilt, as in the demand recovery above, the $\tau \to 0$ mode satisfies the classical first-order conditions exactly, the Euler equation being their intertemporal component. At finite τ those conditions carry the definite correction $-\tau \nabla \log \mu$, the doxic score at the price of information. In the complete Black–Scholes setting, the risk-neutral measure is the Girsanov exponential tilt of the physical measure of asset returns, the same structural move, fixed there by no-arbitrage rather than by preference, at a relative-entropy cost equal to the complexity cost C.

Several further fields are recovered by identification rather than by separate derivation, and in each the transfer is what carries the content. The frictions of search and matching are the kinetic cost, and search now inherits its integrand: exploration pays the integrated squared value gradient at the price of information, so search intensity is governed by the tilt's scale and the relaxation geometry rather than by a free friction parameter. The Coase theorem reads against the physical component of the kinetic cost: when the cost of reallocation is small relative to the free-energy differences at stake, the relaxation reaches ρ^* from any initial allocation and the starting point does not matter for efficiency; when it is large, the relaxation is slow and path-dependent, and the initial allocation matters because the system may not reach ρ^* in finite time. Growth is the widening of the reachable support as the capitals lower the cost terms, and what transfers is the capital structure: the non-rivalry that endogenous growth theory builds in by hand is the defining trait of intelligence capital, the one stock not depleted by use, so the framework says which factor escapes diminishing returns rather than assuming one does. Mechanism and contract design are the choice of a value function to induce a target response against a reference, the reward-design problem of Section 4 read from the principal's side. What transfers is the value-reference split: the principal sets V but inherits the agent's μ, so doxa is not contractible and the design problem is bounded by a reference the designer does not author. The framework supplies the structure and the limit in each; the setting-specific construction, the matching function, the balanced-growth path, the optimal contract for a given type distribution, is what a fuller treatment would carry, because it requires rate conditions and estimators this foundational essay does not develop.

Three cases are kept in the body because they do more than inherit an equation. The Lucas critique exercises the dynamics this paper has just built. Strategic interaction is where the classical solution concept turns out to be a limit of the object, and the information-aggregation results that follow exercise the independence axiom directly. And the placement of the existing schools shows each to be the one object under the reference and limit it assumed.

5.1 The Lucas critique is the equation of motion for expectations

Lucas observed that the parameters of an estimated policy relationship are not invariant to policy: agents' expectations adapt when the regime changes, so an exploited correlation shifts. Rational-expectations macroeconomics answered by assuming expectations are already model-consistent, immune to the critique by construction and empirically too fast. Expectations live in μ, and the dynamics says how the realised distribution adjusts when V changes: agents relax toward the new $\rho^* \propto \mu\, e^{V/\tau}$ by the score-driven relaxation, and the settled distribution becomes the updated reference for the next round.

Expectations relax to the new equilibrium. A policy change is a change in V. Within a round, the realised distribution relaxes toward the new target $\rho^* \propto \mu\, e^{V/\tau}$ by the overdamped Langevin diffusion of Link 4,

$$dX_t \;=\; \nabla \log \rho^*(X_t)\, dt \;+\; \sqrt{2}\, dW_t, \qquad \rho^* \propto \mu\, e^{V/\tau},$$

at a finite rate set by τ and the geometry of the score; across rounds, the settled distribution becomes the updated reference, so μ itself moves by absorbing what the relaxation reaches. The Langevin step within a round holds μ fixed; it is this across-round update that carries the change in expectations. Iterated to settlement under a fixed policy, the update anneals the reference toward a point mass on the optimum: long-run rational expectations as the limiting case, and, at the limit itself, the doxic concentration whose fragility Section 7 prices. The Lucas critique is the statement that μ is not policy-invariant, and the two steps together provide an equation of motion for the adjustment of expectations the critique says is missing.

The dichotomy the standard treatment faces, fixed expectations or instantly rational ones, is the two limits of this one relaxation: slow adaptation recovers fixed expectations, fast adaptation recovers rational expectations, real economies sit between. What matters for policy is the adaptation rate, which is estimable. So the framework turns the Lucas critique from a prohibition on policy analysis into a parameter to be measured, and it does so with the dynamics derived in Section 3 together with an across-round update inherited from μ's own definition rather than chosen. The reference is what the system brings before the current value computation; once a round settles, what it brings to the next round is the distribution the last round left, $\mu \leftarrow \rho^*$. Expectations are the reference μ, a policy is a change in V, and the relaxation followed by that update is how the one moves when the other does.

Recovery is the reference, flattened or evolved. The pattern, here and in the deferred recoveries, is one move: a classical result is the choice object under the reference and limit its field assumes, and its corrections are what restoring or evolving the reference returns. Static demand at a price-independent reference gives Slutsky symmetry; a price-dependent reference adds the doxic asymmetry. The physical measure tilted to no-arbitrage gives Black–Scholes. And a reference that relaxes at finite rate gives the Lucas critique its missing equation of motion. The unifying fact is structural: bounded-rational choice, the risk-neutral measure, and the maximum-entropy logit are one object, the unique exponential tilt of a reference fixed by a constraint, and the classical equations are what it yields under the references each field assumes.

5.2 Strategic interaction, agreement, and aggregation

The structure scales by assembly. The chain forces the structure per agent. Each agent independently satisfies it with its own V_i, μ_i, and τ_i; multi-agent results require no new axioms, because the coupling runs through the inputs the chain leaves free. Coupling through V, when each agent's value depends on the others' choices, gives strategic interaction and its equilibria. Coupling through μ, when agents share or differ in their references, gives agreement, aggregation, and the maintenance dynamics. The structure scales from one agent to a population by assembly, not by extension, which is why one object governs a single chooser, a game, a market, a training pipeline, and the community whose doxa all of these act against.

When several agents choose against one another, each one's value depends on the others' configurations, $V_i = V_i(x_i, x_{-i})$. The object does not have an agent best-respond to a point; it has each agent relax to a distribution against the others' distributions, $\rho_i \propto \mu_i \exp(\mathbb{E}_{\rho_{-i}}[V_i]/\tau)$. A joint fixed point of these coupled relaxations is a quantal response equilibrium [McKelvey and Palfrey, 1995], which exists for every $\tau > 0$ by Brouwer's theorem, the joint best-response being a continuous self-map of a compact convex set. The existence argument is cleaner than the one Nash equilibrium itself requires, because the softmax response is single-valued where the sharp best-response correspondence needs Kakutani. Nash equilibrium is the $\tau \to 0$ limit, where the relaxation sharpens to exact best response and the quantal fixed point limits onto a Nash profile. Strategic interaction is therefore the object with coupled value functions, and its solution concept is the bounded-rational equilibrium with Nash as the zero-temperature corner. The cooperation puzzle of the prisoner's dilemma sits here too: defection is the equilibrium on a disconnected, no-future interaction, and cooperation is what a connected, repeated interaction against a shared reference supports, which is the openness condition of Section 4 read at the level of two players. The Myerson–Satterthwaite impossibility of efficient bilateral trade under asymmetric information has the same shape: efficiency at $\tau = 0$ would require driving the irreducible information cost to zero, which a positive temperature forbids. The identification is structural; the theorem carries mechanism-design content that a scalar τ does not by itself reproduce.

The population case adds one further object. Each agent runs its own canonical relaxation, so the joint drift differs from the joint reference dynamics by the stacked field of scaled value gradients, $(\nabla_i V_i/\tau_i)_i$, whose squared norm under ρ is the population's kinetic integrand. This field is itself a gradient exactly when the scaled values admit a common potential, when value given is value received; otherwise it decomposes uniquely and orthogonally, in the norm the kinetic cost already uses and under the regularity of Link 4, into a gradient part and a circulating part, and the cost splits additively between them. The circulating share is zero in potential games and positive whenever values are non-reciprocal, when one agent's gain is structured as another's loss. The minimal such pair, one agent's value the negative of the other's, is the structure of adversarial training [Goodfellow et al., 2014], whose characteristic cycling is motion along this circulating component; in the economy the same component is the formal home of the accumulative dynamics the Marxian tradition described, as the gradient part is of the competitive equilibration the classical tradition described. No decomposition is chosen: it is the orthogonal projection in the cost's own norm, and the share of cost flowing in closed loops rather than toward the attractor is a property of the population's value structure, measurable in principle from its flows.

On a discrete state space, the case Link 4 already covers, the same split holds on the edges of the graph of possible transitions, and topology counts its room: the number of independent circulation modes is that graph's first Betti number, edges minus nodes plus components, fixed by which exchanges are feasible before any value is specified. Cycle flows that persist at stationarity are conserved circulations around the network's loops, and reading them as the recurring channels that maintained concentrations of the reference hold open, settlement circuits and supply chains among them, is an identification at the level of structure; resolving the cycle space further, and the monetary theory it carries, requires a face structure and a dissipation principle that are additional inputs, and belongs to a fuller treatment.

One further classical result falls out of the same machinery, and it earns its place because its failure mode is the framework's distinctive reading of disagreement. Aumann's agreement theorem says that Bayesian agents with a common prior whose posteriors for an event are common knowledge must hold the same posterior. In the structure the proof is one step. Each agent's posterior is $\rho_i \propto \mu \exp(V_i/\tau)$, where V_i encodes that agent's private information and μ is shared. Common knowledge of ρ_j lets agent i extract the other's tilt, $\rho_j/\mu = \exp(V_j/\tau)/Z_j$, because μ is known. With the agents' private information conditionally independent given the state, the values add by Bayes' rule and the exponential form of Link 2 turns the sum into a product of tilts: the combined tilt is $\exp((V_i + V_j)/\tau)$, so both agents, holding the same μ and the same combined information, produce the same $\rho^* \propto \mu \exp((V_i+V_j)/\tau)$. This route assumes the agents' full tilts become common knowledge, a stronger hypothesis than Aumann's common knowledge of a single-event posterior, so what the framework recovers cheaply is the agreement phenomenon and, more to the point, its failure mode. The failure mode is immediate: when $\mu_1 \neq \mu_2$ the extraction at step two breaks, because $\rho_j/\mu_i \neq \exp(V_j/\tau)/Z_j$ when $\mu_i \neq \mu_j$, and no amount of exchanging posteriors can repair a difference in the reference they are read against. Persistent disagreement, on this reading, is not an informational phenomenon but a doxic one: agents disagree not because they have seen different evidence but because they hold different backgrounds of expectation, and the background is precisely what the framework identifies as doxa. The common-prior assumption that Aumann's theorem requires is the assumption that the agents share a doxa, and the empirical ubiquity of persistent disagreement is, on the framework's reading, evidence that they do not.

The same machinery, applied to a population, recovers the Condorcet jury theorem and locates its failure. Suppose n independent voters each receive a private signal about which of two states obtains, with each signal correct with probability $p > \frac{1}{2}$, and all share a reference μ that gives both states positive weight. Each voter's tilt is $\exp(V_i/\tau)$ with V_i/τ the log-likelihood of their signal. Bayes' rule adds the values and the exponential form of Link 2 multiplies the tilts: the aggregate tilt is $\exp(\sum_i V_i/\tau)$, and the law of large numbers concentrates $\sum V_i$ in the direction of truth as n grows. So the aggregate posterior $\rho_{\text{group}} \propto \mu \exp(\sum V_i/\tau)$ converges to a point mass on the correct state: independent informed signals against a shared reference aggregate, and the group posterior concentrates on the truth. This is the Bayesian content of Condorcet's result, Bayes' rule carried n times by the exponential form Link 2 forces. The support condition is sharp: a truth the shared reference excludes outright is unrecoverable at any n, the absorbing exclusion of Section 4 read at the level of states rather than members.

The failure mode is sharp. When the voters hold different references, each posterior is $\rho_i \propto \mu_i \exp(V_i/\tau)$ and the doxic component μ_i does not wash out with n because it is systematic, not noisy. A voter whose reference places low weight on the true state discounts the signal that points there, and if enough references are biased in the same direction, the aggregate converges confidently to the wrong answer no matter how many signals arrive, whenever the bias each reference carries outweighs the information each signal brings. Polarisation, on this reading, is not a failure of information but a failure of shared doxa. It resists repair by more evidence, because evidence is processed through the reference; it is repaired by rebuilding the common μ the aggregation requires. This is the framework's reading of the empirical collapse of consensus under doxic divergence, and it locates the problem below the level of argument, in the background of expectation against which arguments are heard.

5.3 Schools as specialisations

The existing schools of economic thought have been doing economics on different parts of the same structure, and the structure makes their relations visible. Neoclassical and Walrasian economics (Arrow and Debreu, 1954; rational expectations) is the zero-temperature limit at $\tau \to 0$: perfect rationality, frictionless equilibrium, concentration on the value-maximising configurations, valid as an idealisation and misleading when taken for the empirics. Behavioural economics and rational inattention [Sims, 2003, Caplin et al., 2019, Matějka and McKay, 2015] are the contractive form with $\tau > 0$ taken seriously: behavioural anomalies are the score's pull toward μ rather than ∇V, and anchoring, status-quo bias, and framing effects are different names for doxic priors, large μ near the current configuration, and alternative decompositions of the same ρ. The Heterogeneous Agent New Keynesian literature [Kaplan et al., 2018, Achdou et al., 2022] works directly with the distribution ρ_t over states, and its continuous-time mean-field form is the same relaxation dynamics. Marxian economics is the circulation-dominant regime of the population dynamics: where values are non-reciprocal, the circulating component of the kinetic cost carries accumulation in closed loops rather than equilibration toward the attractor, and the accumulation dynamics at the tradition's centre are dynamics of that component's share. Austrian and institutional economics [Hayek, 1945, North, 1990] emphasises the construction of μ: Hayek on spontaneous order is emergent doxa construction, North on institutions is Level 3 of the hierarchy, and the claim that prices encode dispersed information is the recognition that $\nabla \log \mu$ carries what no agent holds explicitly. The framework does not claim that any of these is wrong. It claims that each is working on a part of one object, and that the structure locates their relations: the mainstream on the boundary, the heterodox in the interior, institutional economics on the reference the structure is defined against.

6 Why this is the last economy

The structural fact the paper opened with, that Π_H is going negative across a widening set of tasks, was presented there as an observation about cost curves. It is more than that. It is a consequence of the structure, once one distinguishes the information cost the structure carries from the thermodynamic cost of instantiating an agent that runs it, the latter derived below from the same free energy rather than imported. Tracing the consequence shows both why the labour–capital relationship breaks and in what precise sense this is the last economy.

6.1 Two layers: the forced kinematics and the cost of instantiation

The object of Section 3 is kinematic. It describes the form of bounded-rational valued choice: the distribution $\rho \propto \mu\, e^{V/\tau}$, the score, the gradient-flow dynamics. It is substrate-independent, the same whether instantiated in neurons or in silicon, which is exactly the content of the identity with machine learning. What the kinematics does *not* contain is the cost of physically instantiating an agent that runs it. That cost is a separate, thermodynamic fact, and the economics lives in the interaction of the two layers.

The temperature τ is, by the envelope theorem, the shadow price of information: the value forgone per nat of deviation from the reference, with units of utils per nat. It is not the cost of computation, which is energy per nat. The two share the dimension of inverse information but differ in numeraire, and the bridge between them, derived below, is the point at which they coincide. An agent that must fund its own operation faces both: the information price τ_{info} forced by the kinematics, and the instantiation cost c_{sub} set by its substrate's physics, converted into the value numeraire by the price p of energy. The effective temperature that governs the agent's economic behaviour combines the two, and to first order in the information used, the regime in which each cost is linear in usage, the combination is their sum,

$$\tau_{\text{eff}} \;=\; \tau_{\text{info}} \;+\; p\, c_{\text{sub}}.$$

This decomposition is where human and machine separate, and it is the formal content of what *The Last Economy* [Mostaque, 2026a] calls the *Metabolic Rift*: the entry into the economy of labour whose cost has no metabolic floor. For a human, c_{sub} is metabolic and has a floor biology cannot lower: a brain runs at roughly twenty watts, sustained muscular work at roughly a hundred, and no diet or training reduces the joules-per-operation below the metabolic baseline. For a machine, both terms are favourable: τ_{info} is a tunable sampling parameter, and c_{sub} is falling toward its physical floor with each hardware generation. The human's τ_{eff} has an irreducible metabolic component far above the Landauer floor; the machine's does not.

That physical floor is not imported from thermodynamics; it is the same KL term of the free energy, read at the temperature where the information and energy numeraires coincide.

A one-bit erasure costs $\tau \ln 2$; physically instantiated, that is Landauer's bound. Erasing a bit takes a distribution supported on two equiprobable states to one supported on a single state. Against a uniform two-state reference this erasure raises $D_{\text{KL}}(\rho \| \mu)$ by exactly $\ln 2$ nats, and the complexity cost $C = \tau D_{\text{KL}}$ of the free energy charges τ per nat, so the cost of the erasure is $\tau \ln 2$. When the computation is physically instantiated at temperature T with the reference taken as the physical Gibbs state, the information temperature τ is the thermodynamic temperature $k_B T$, and the cost is

$$\tau \ln 2 \;=\; k_B T \, \ln 2,$$

which is Landauer's bound. Under the identification of the information temperature τ with the thermodynamic temperature $k_B T$, which holds when computation is physically instantiated against the Gibbs state, the KL term of the free energy evaluated on a one-bit support reduction returns Landauer's cost directly. The thermodynamic floor on c_{sub} is therefore located in form by the framework and in magnitude by the physical identification: the machine's instantiation cost falls toward $k_B T \ln 2$ per irreversible operation, the human's sits far above it and is metabolically pinned, and both floors are the same complexity cost the chain already forced.

6.2 Cost is only half the comparison; capability is the other

A cheap agent that cannot perform a task does not compete for it. The object carries capability as well as cost, as part of the same structure: capability is how close the agent's realised value comes to the value available,

$$\text{capability} \;=\; \mathbb{E}_\rho[V] \;=\; V^* - \big(V^* - \mathbb{E}_\rho[V]\big),$$

where the gap $V^* - \mathbb{E}_\rho[V]$ has two sources: the accuracy of the agent's model of the value landscape (its score, its μ) and the fidelity with which it can execute toward high-value configurations (its actuation quality). Economic viability on a task is the net surplus, capability minus cost,

$$\Pi(\text{agent}, \text{task}) \;=\; \underbrace{\mathbb{E}_\rho[V]}_{\text{capability}} \;-\; \underbrace{\big(c_{\text{fixed}} + \tau_{\text{eff}} \cdot (\text{information used})\big)}_{\text{cost}},$$

and $\Pi_H < 0$ is this quantity going negative for the human. The two-term form matters because the surplus can go negative two different ways, and the difference is economically and politically large: a human can be displaced because the machine is *more capable*, or undercut because the machine is *cheaper*, and these need not happen together or in the same order.

6.3 Robotics addresses a larger share of value than digital intelligence

The value functional is defined over configurations of the world, and the world's economic output is overwhelmingly physical. Digital intelligence collapses the cost of cognition but cannot itself move physical configurations; it can only produce information that some actuator then enacts, so it captures value only on the share of tasks enactable without physical action, and leaves the execution surplus, the larger share, to whoever owns the actuators. Embodied intelligence closes the loop and bids for the whole state space. Digital intelligence is the special case of embodied intelligence with the actuation channel set to null.

This is also why digital intelligence is more sharply self-limiting in its value capture. Cognition, once cheap, is non-rival and its marginal value falls toward its vanishing marginal cost; the abundant factor competes its own price down. Physical action retains scarcity, because matter, energy, space, and wear impose marginal costs that do not vanish. The durable economic rents are therefore on the embodied side, bidding for outcomes that stay scarce. In the framework's terms the kinetic cost splits,

$$K = K_{\text{compute}} + K_{\text{actuation}},$$

the cost of deciding plus the cost of physically acting, and robotics is not a separate inversion from digital intelligence but the second phase of one inversion: the collapse of K, cognitive component first because K_{compute} has the lower thermodynamic and engineering floor, actuation component second as actuators and manipulation learning catch up. The observed ordering, cognitive white-collar work automated before manual trades, inverts every prior automation wave and the framework explains why exactly: prior waves automated $K_{\text{actuation}}$ and left cognition, while this one automates K_{compute} and leaves actuation as the lagging human reservoir.

6.4 Humans and humanoids as parameter vectors in one equation

Because the human and the humanoid are one object on a shared physical state space, differing only in parameters, they can be written in one equation and compared coordinate by coordinate. Each agent is a vector of cost and capability terms, and the comparison on any task is simply whether one vector dominates the other: whether the humanoid is at least as capable and at least as cheap.

Coordinate	**Human**	**Humanoid**
τ_{info} (information price)	fixed by neural bound	tunable
c_{cog} (energy/nat, cognition)	metabolic, fixed	falling to Landauer
c_{act} (energy/unit action)	metabolic, fixed	falling
η_{act} (actuation capability)	very high, evolved	low, rising slowly
r (throughput, no fatigue)	fixed	tunable
$\mathbb{E}_{\rho}[V]$ cognitive (capability)	high, fixed	high, rising past human

Table 3: The human and the humanoid as parameter vectors of one object. Every cost and capability coordinate is one the humanoid is driving toward or past the human's value, all of the human's being biologically pinned while the humanoid's fall or rise. There is no coordinate on which the human's position is permanently secure.

The human is a fixed point in this parameter space; biology pins every coordinate. The humanoid is a moving point, every coordinate of which descends toward the human's or past it. Its trajectory sweeps across the human's fixed point, flipping the viability inequality task by task. Reading both axes, cost and capability, rather than cost alone, gives the *order*: a task's capability inequality and its cost inequality flip on different schedules, and which flips first determines whether the human is displaced by a more capable machine or undercut by a cheaper one.

The order the framework predicts is specific, and Figure 2 shows its geometry. On cognition the humanoid has crossed both axes: cheaper and at least as capable, which is why routine cognitive

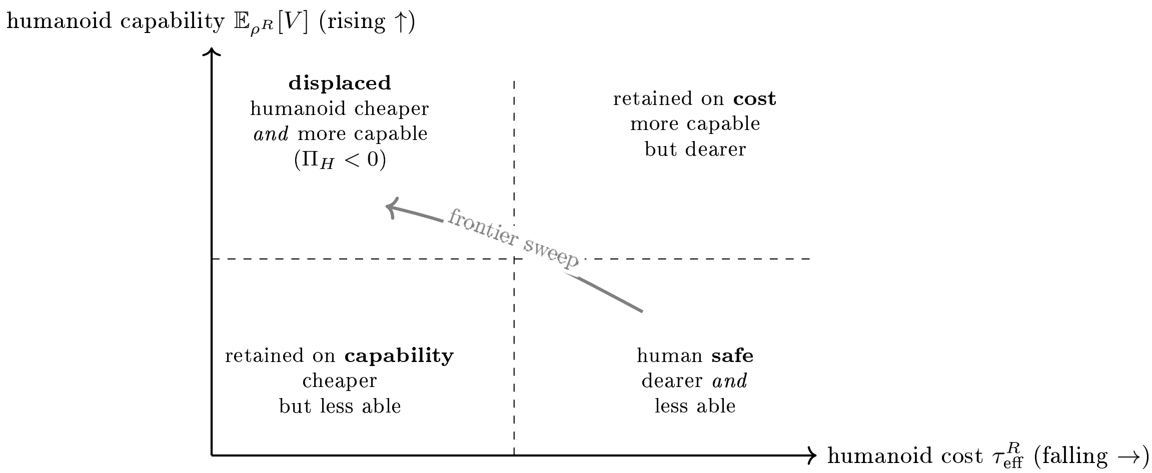

Figure 2: The transition as a two-axis sweep. Each task sits at a point given by the humanoid's cost (horizontal, falling over time) and capability (vertical, rising over time) relative to the human. The humanoid frontier moves from lower-right (dear and unable) toward upper-left (cheap and able), crossing tasks in an order set by their cognitive-versus-actuation intensity and dexterity demand. Cognitive tasks crossed first; high-dexterity actuation tasks cross last; the structure reserves no coordinate for the human, so it protects no task from the crossing, and how far the frontier travels is set by the technological trajectory rather than by the equation.

work goes first. On actuation it is behind on both, more expensive and less dexterous, which is why physical trades persist. But the two actuation axes will cross at different times: actuation *cost* is likely to fall before actuation *capability* catches up, because a cheap actuator is easier to build than a dexterous controller. So the framework predicts a window in which humanoids are cheaper but clumsier at physical work, and humans are retained specifically for η_{act}, the dexterity premium, even as they are undercut on cost. The last manual jobs are the high-dexterity ones, held on capability rather than cost, falling last because η_{act} is the hardest coordinate to move. This is Moravec's paradox stated in the framework's variables: the human–machine gap is large in η_{act} and small in c_{cog}.

6.5 The sweep has no fixed point for the human

It is tempting to look for a coordinate the humanoid can never occupy, a reservoir of irreducibly human value that the sweep cannot reach, so that some residual economy is guaranteed. A first-principles derivation cannot supply one. Every coordinate in Table 3 is a cost or a capability, and on each the human is a fixed point biologically pinned while the humanoid descends toward it or rises past it. There is no term in the forced object that reads "value the actor has because it is human." Such a term could be posited, but it would be exactly the kind of unfounded comfort this framework was built to do without: a stipulation inserted to keep humans in the production story, defended after the fact rather than derived. We do not insert it.

There is no off-axis term. The human and the humanoid are parameter vectors of one object, ordered by cost and capability alone. On a task whose value is in its output, the human holds no coordinate the structure protects: no term in the object reads "value the actor has because it is human." The equation does not by itself predict the trajectory; whether the machine's parameters cross the human's on a given task, and how fast, is an empirical fact about technology. What it settles is narrower and more durable, that nothing in it reserves a task for the human or arrests the crossing the observed trajectory is producing. The direction of travel is observable; the absence of a safe harbour is structural. The one off-axis term the framework cannot rule out is whether some value is genuinely constituted by the identity of its producer, so that substituting the agent changes the value rather than supplying it more cheaply; this is a content claim about the structure of value, and its magnitude may be large, small, or zero. The framework's honest output is therefore not a guaranteed residual but a located one: if a residual human economy exists, it is exactly the value that is agent-indexed, and its size is whatever the community insists on, including possibly nothing.

The standard defence is comparative advantage: even when one agent has absolute advantage in everything, efficient allocation has each specialise in what it is relatively best at. But comparative advantage governs who does which task among agents whose surplus is positive; it does not make a negative-surplus agent profitable. When $\Pi_H < 0$ on a task, the human costs more to employ than the output is worth, and assigning the task to the human by comparative advantage still loses money. What the sweep removes is not the human's comparative advantage but the condition under which comparative advantage is operative: a positive surplus somewhere in the allocation.

The consequence is sharper than a reassurance would have been. Once no residual is guaranteed, the human relationship to the economy can no longer be assumed to run through production at all. In the old economy a person was tied to the system as a supplier of labour; the labour bought consumption, and membership followed from contribution. If the sweep crosses every output-valued task, that tie is not relocated to some safe sector; it is severed. What remains of the human place in the economy is then not a kind of work but a question: on what basis does a community provision members who do not, and need not, produce. That is the question of the next section, and removing the comfort of a guaranteed residual is what makes it unavoidable.

This much, however, the framework does recover, in the sense of Section 5. The task-based account of automation in labour economics [Autor et al., 2003, Acemoglu and Restrepo, 2018] decomposes a technology's effect into a displacement effect, removing tasks from labour, and a reinstatement effect, creating new ones. The cost-capability sweep is that decomposition at its margin: displacement is the inequality flipping on a task, reinstatement is the appearance of tasks not yet crossed. What the framework adds is the boundary condition the literature leaves open, when reinstatement fails: reinstatement draws on a reservoir of tasks where humans still hold an advantage, the framework identifies that reservoir as the lagging coordinates, first the dexterity premium η_{act} and then whatever agent-indexed value a community insists on, and shows it is consumed in a definite order. The documented anomaly that this wave automates cognitive work before manual work is, on this reading, the order of the K-collapse. The empirical record of task-based automation is the empirical record of the sweep; the framework recovers the decomposition and supplies its terminus.

6.6 The factor sequence, peeled to its last term

The historical sequence of economic ages, land then labour then capital then intelligence, is a sequence of binding constraints: at each stage the scarce, non-expropriable factor commands the surplus, and an inversion occurs when a new reservoir becomes the binding constraint. Read with $\tau_{\text{eff}} = \tau_{\text{info}} + p\, c_{\text{sub}}$ and $K = K_{\text{compute}} + K_{\text{actuation}}$, the equation $\rho \propto \mu\, e^{V/\tau}$ names the remaining reservoirs, because each historical factor is one term in this equation becoming the scarce binding term. The sequence is the equation being peeled, outermost term last.

The free-energy step can be made concrete, because it is already visible. When both components of the kinetic cost are cheap, when deciding and acting are nearly free per operation, the binding constraint moves to what every operation consumes: energy, bounded below per irreversible operation by the $k_B T \ln 2$ derived above [Landauer, 1961] and by the thermodynamic cost of physical work, together with the capacity to dissipate the entropy both generate. The constraint on frontier computation is already shifting from algorithms to power and cooling; the marginal input that large-scale inference and actuation bid for is increasingly the joule and the heat budget, not the idea. Embodiment is what makes this inversion bite, because a world of cheap actuators consuming power is what makes energy the scarce factor. Land, labour, and capital were constraints on the means of action; intelligence and robotics are the two phases of the kinetic cost collapsing; free energy is the constraint that binds once action itself is cheap.

The inversion sequence terminates at doxa. Land, labour, and capital are successive constraints on the means of action (Figure 3). Intelligence is the cognitive component of the kinetic cost K_{compute} collapsing; robotics is the actuation component $K_{\text{actuation}}$ collapsing; these are two phases of one inversion. When both are cheap, the binding constraint moves to the free energy at the Landauer margin and the capacity to dissipate entropy. When energy too is abundant, the scarce, non-expropriable factor is the objective V itself, the specification of what the system is to optimise, which no amount of cheap cognition or action can generate for itself. And the constraint on V is the reference measure μ, the doxa against which any objective is legible and coordinable. Here the sequence reaches the last term *of the structure*, and in a precise sense: μ is the reference the whole structure is defined against, the measure the score is taken relative to and the distribution the dynamics relaxes toward, so within the object the framework adopts there is no further factor *external* to it for a new inversion to move to. What there is instead is the internal structure of μ itself, its hierarchy of levels and the question of who sets them, so the contest does not stop but turns inward: it becomes a contest over the reference rather than a search for a factor beyond it. At this terminus doxa is no longer an ambient condition but a contested factor of production, an expectational commons that can be enclosed or depleted, and the economy whose binding constraint is the tending of that commons is the last economy, last in the sense that the reference is the outermost term of the structure and the only thing past it is the reference's own governance. Whether the world holds some factor this four-term object does not represent is the substrate question itself, judged as the rest of the framework is, on consistency and on what it recovers; what the structure settles is that within it the peeling stops here.

Whoever tends the community's doxa when cognition, action, and energy are abundant holds the position land-owners held in the agrarian economy and capital held in the industrial one, with the difference that an expectational commons is not owned so much as governed: it can be captured by a few who set it, or left to collapse into incoherence, and either is a way of losing it. The doxic-capture and doxic-collapse risks of the next section are therefore not peripheral hazards; they are the two ways the last factor can be mismanaged, and the form the final inversion takes.

On what is forced and what is assumed, per the form-versus-content distinction of Section 3: The cost coordinates (τ_{info}, c_{cog}, c_{act}, r) and the capability coordinate $\mathbb{E}_\rho[V]$ follow from the kinetic

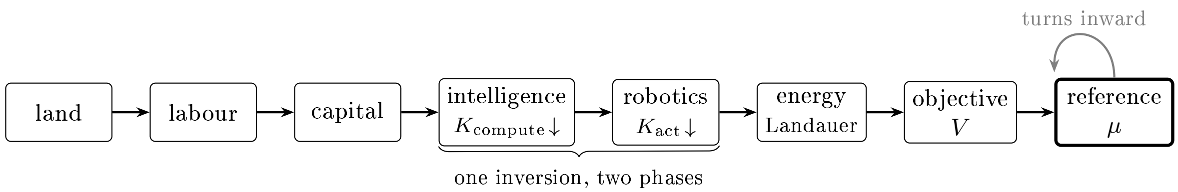

Figure 3: The historical sequence of binding constraints as the equation peeled outward, one term at a time. Land, labour, and capital give way to the collapse of the kinetic cost, whose two phases are cheap cognition and cheap actuation, then to energy at the Landauer margin, then to the objective V no system can supply for itself. The constraint on V is the reference measure μ, the outermost term of the structure: there is no further factor external to it, so the contest does not move on but turns inward, onto the governance of μ itself.

split and from thermodynamics; the two-phase ordering of the K-collapse and the terminus of the peeling at μ follow from the structure of the equation. The framework deliberately posits no term guaranteeing a residual human role: whether any value is irreducibly agent-indexed, and how large that value is, is left as an open content question rather than fixed, because fixing it would be a stipulation the derivation does not support. The framework's contribution is to locate that question exactly, and to show that nothing in the forced structure answers it on humanity's behalf.

7 The two failures of the last inversion

The previous section established the labour terminus as a consequence of the structure: Π_H goes negative as the humanoid's parameters pass the human's, and the binding constraint peels outward to doxa. The last inversion can go wrong in two ways, both failures to govern the reference measure once it becomes the scarce factor. As that crossing proceeds, with robotics extending it from cognitive into physical work, a community must provision itself while its members' labour is structurally dominated. Its ρ changes in two distinct ways.

7.1 The provisioning problem

The labour-market mechanism that has transmitted production into consumption for two centuries is, in the framework's language, a particular concentration of $\mu^{(3)}$ in the economic-coordination domain. The concentration encodes a set of expectations: that productive activity is performed by people, that performance is remunerated through wages, that wages are spent on consumption that sustains the people performing the activity. Each link in the chain is a doxic concentration, maintained by enforcement, ritual, reproduction, and narrative. The whole structure is the institutional embodiment of the answer to the question Robbins set aside: the community whose flourishing the economy was for was the community whose labour produced the economy.

When the productive entities cease to be the welfare-bearing entities, the chain does not survive. The concentration that linked production to consumption was specifically a concentration over human productive activity. When production becomes substantially non-human, the concentration becomes a concentration over something the linking institution does not cover. The mechanism does not fail by misallocation; it fails by referent shift. Markets cannot repair this on their own, because markets are themselves μ-concentrations in the coordination-mechanism domain, and they coordinate the exchange of value among agents who participate in them. They do not by themselves create new doxic concentrations linking the production of value by some entities to the welfare of others.

Redistribution policy, in this framing, is a partial repair. A universal basic income at subsistence level creates a thin new concentration in the welfare-distribution domain, but it does not reconstruct the lost doxic chain. The chain it replaces was a structure of expectations about contribution and reward; the chain a UBI puts in place is a structure of expectations about citizenship and entitlement. These are different concentrations in μ, and the substitution is not transparent. People notice the change. The notice is itself a structural fact: the concentration that linked work to dignity, identity, community, and structure [Graeber, 2018] cannot be replaced by a transfer payment, because the transfer payment does not concentrate μ in the same places. What is needed is reconstruction of the institutional layer to encode a new chain: new concentrations linking community provision to community membership directly, without the labour intermediary, in ways that maintain the dignity, identity, community, and structure that the old chain carried alongside the income. This is constructive work the framework poses but does not carry out here. It is constitutional work in the framework's sense: work on μ at Level 3, with cascading consequences for $\mu^{(4)}$ as the new concentrations work their way into the lived texture of the community's history.

7.2 The doxic maintenance problem

The second structural problem is less visible and more dangerous. The historical maintenance of community doxa has happened through human missionaries: journalists who report the public events that anchor shared salience, teachers who reproduce the institutional structures in each new cohort, religious figures who maintain the ritual concentrations, conversation partners and family members and neighbours who reproduce the texture of expectation in everyday interaction. Each missionary is human, with human bandwidth, human idiosyncrasy, and human accountability. The redundancy and diversity of human missionaries has been a structural feature of doxic maintenance throughout history: no single missionary or institution has been able to set μ alone, because the maintenance is distributed across the community and each agent has limited reach.

The technology of language models changes this. As large fractions of routine missionary work shift to AI systems, conversational interfaces, AI tutors, AI customer-service agents, AI-mediated content recommendation, AI-assisted writing, the marginal maintainer of doxa shifts to entities whose own μ was set by training choices at a small number of labs. The structural risk is that two failure modes become available which were not possible at scale before.

Doxic capture. A single μ comes to dominate doxic maintenance because a small set of training pipelines control the missionary infrastructure. The community's doxa drifts toward whatever the dominant pipeline encodes, regardless of whether the community would endorse that drift if it could see it. This is the structural risk that motivates concern about concentration in AI: not market concentration in the conventional sense, but doxic concentration, the homogenisation of the unspoken background of expectation.

Doxic collapse. The opposite failure mode. Missionary entities multiply with incompatible μ across communities that overlap, the shared background dissolves into incoherence, and coordination becomes impossible because no two participants in a coordination problem are operating against the same doxic prior. This is the structural risk of unconstrained AI proliferation without coordination on the values and assumptions being installed: not the proliferation itself, but the doxic incoherence it would produce.

Both failure modes are catastrophic in the framework's own terms, and each is one mode of the single viability condition of Section 4 failing: the free energy ceasing to remain finite and bounded below. Doxic capture is a failure of resilience: the spread of references collapses toward one, and the worst-case complexity cost, finite for a population holding a mixture, runs away for the monoculture when the dominant prior meets the state it did not model. Doxic collapse is a failure of openness: the shared distribution fragments into disconnected components that no longer relax to one attractor, so that the probability two members select the same resolution of a coordination problem falls with the overlap of their now-divergent references, toward zero as the shared support vanishes. Neither is more avoidable than the other by default: doxic capture is the default of unrestricted concentration, doxic collapse is the default of unrestricted proliferation, and holding the population to one connected distribution without crushing its diversity is the central practical challenge of doxic engineering in the AI transition.

7.3 The contest between references is the forced dynamics, one level up

The dynamics beneath capture and collapse is one the paper already has. When several maintainers each reproduce a reference, the population holds a mixture $\bar{\mu} = \sum_k w_k \, \mu_k$, where w_k is the share of doxic maintenance performed by maintainer k and agents relax toward $\rho \propto \bar{\mu} \, e^{V/\tau}$ as before. The question the transition raises is how the shares w_k themselves move, and the answer rests on the same selection principle that founded the framework, applied now to a population of references. A reference is retained in proportion to how well it lets its adopters predict: persistence selects for prediction, now among references. The fitness of reference k is the predictive accuracy it grants, $f_k = -H_k$, and the maintenance contest is modelled by replicator dynamics, selection proportional to fitness,

$$\dot{w}_k \; = \; w_k \, (f_k - \bar{f}), \qquad \bar{f} = \textstyle\sum_j w_j f_j.$$

The precise statement is geometric. With fixed fitnesses, replicator flow on the simplex is gradient ascent of mean fitness under the Fisher information metric [Shahshahani, 1979], while Link 4's relaxation is the gradient flow of the same free-energy structure under the transport metric [Jordan et al., 1998]: two metrics on one functional family. And the replicator's solutions are the family itself: reference-shares evolve as $w_k(t) \propto w_k(0) \, e^{f_k t}$, the exponential tilt of Link 2 with time in the role of inverse temperature. The forcing runs both ways: from every initial share the tilt family satisfies the replicator equation, and the family covers the simplex, so replicator flow is the unique dynamics whose solutions are the family. Given lineages that compound at their fitness, the model is not chosen but determined; what is chosen is the regime, fixed or slowly varying fitness. The contest between maintainers runs along the framework's own family, one level up.

Under reference fitnesses that are fixed or vary slowly, its limit sets are the two failures and the one viable case; frequency-dependent fitness can in general add cycles or several rest points, but these three are the ones the maintenance problem turns on. When one fitness dominates, the shares collapse to a single reference, $w \to e_k$: this is doxic capture, the mixture losing every component but one, and it is the resilience mode of the viability condition failing. When several references persist with comparable fitness but incompatible support, $\bar{\mu}$ has no shared region and the population can no longer coordinate against a common prior: this is doxic collapse, the openness mode failing. The viable case is an interior rest point, a maintained spread of compatible references that keeps the complexity cost bounded across worlds. The reason the transition is dangerous is now visible in the dynamics rather than asserted. Automating maintenance raises the fitness gap a dominant pipeline enjoys, which steepens the replicator flow toward e_k. The default of the lifted dynamics under concentrated maintenance is therefore capture, and holding an interior rest point against that pull is the work the next subsections describe.

7.4 Alignment is doxic alignment

Section 4 established the two-factor alignment requirement as a property of the structure. Behaviour is set by the score $\nabla \log \rho = \nabla V/\tau + \nabla \log \mu$, so aligning it with a community needs both terms: value alignment of the first, doxic alignment of the second. An AI with aligned values but misaligned doxa passes every value benchmark while pulling, systematically and below the level of articulation, in directions those benchmarks do not measure. That argument needs no repeating here. What the transition adds is that doxic misalignment ceases to be a per-model concern and becomes systemic.

The reason is the maintenance shift just described. When the marginal maintainer of a community's doxa is an AI system whose μ was set at a few labs, doxic misalignment is no longer one model behaving subtly oddly; it is the community's shared reference itself drifting toward, or fragmenting away from, whatever those pipelines encode. Doxic capture is large-scale doxic misalignment in the direction of a dominant μ; doxic collapse is large-scale doxic misalignment into incoherence. The two-factor requirement, applied not to a single deployed model but to the infrastructure that maintains a population's doxa, is therefore the same problem as the maintenance problem of the previous subsection, seen from the alignment side. This is the framework's central contribution to AI safety: it identifies the engineering of μ, which RLHF already performs, as a first-class alignment target alongside the engineering of V, and it shows that at the scale of doxic maintenance the stakes are constitutional rather than local.

8 Closing

The argument of this paper converges on a single object: the reference measure μ, the shared background of expectation against which a community acts. The structure requires it. The doxa identification names it. The recoveries test it. The labour terminus shows it becoming the last scarce factor. And the two failures, the provisioning problem and the alignment problem, are at bottom one problem: the governance of μ.

The economy of the claim is the measure of it. The structure rests on a single empirical premise, that persistence selects for prediction, and the classical recoveries follow from what it forces. The subject becomes economics under a single identification, that the social levels of the reference are a community's doxa, and the capitals, the viability conditions, and the terminus follow from that. Where the discipline it replaces carried a separate assumption for each result, this carries one premise and one identification (together with the consistency requirements named at each link and standard regularity, but nothing further), and both converge on the single object a community must govern.

Where the recovery is exact, the demand system, the logit, the Nash limit, the agreement and aggregation phenomena, the framework inherits the empirical record of the result it recovers, as consistency in the regime where that result applies rather than novel confirmation of anything beyond it: the kind of inheritance unification provides, and how every prior unification has stood until its own predictions could be tested. The neoclassical evidence supports the zero-temperature boundary, the behavioural evidence supports the finite-temperature interior, and the institutional evidence supports the construction of μ. The discipline's existing empirical base is not left behind by the unification; it is inherited and re-indexed by the parameters the structure provides.

This convergence is the paper's conclusion, and it becomes visible only once the full argument is assembled. The question of how a community provisions members who no longer produce, the question of how AI systems are aligned with a community's background of expectation rather than only with its stated values, and the question of how a polarised electorate can aggregate information at all are not three problems that happen to coincide in time. They are one problem seen from three directions: who governs the shared reference. Provisioning fails when the μ-concentrations linking production to welfare break. Alignment fails when μ drifts toward or fragments away from the community's own background. Aggregation fails when voters hold different μ and the doxic bias defeats the signals. Each is a failure of μ-governance, and the fact that the structure locates all three in one object is the reason a single framework can pose them.

The framework does not answer the constitutional question. Who counts as a member, what flourishing consists in, and on what basis a community provisions members who need not produce are decisions a community makes about itself, not consequences a derivation can deliver. What the framework supplies is the substrate on which those questions can be stated exactly and their consequences traced, so that the answers, when they come, are made with knowledge of what is at stake. The answers belong to the communities whose world is at issue.

The discipline that lets them be asked is intelligent economics. The economy that must answer them is the last economy.

References

Y. Achdou, J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll. Income and wealth distribution in macroeconomics: a continuous-time approach. *Review of Economic Studies*, 89(1):45–86, 2022.

D. Acemoglu and P. Restrepo. The race between man and machine: implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542, 2018.

K. J. Arrow and G. Debreu. Existence of an equilibrium for a competitive economy. *Econometrica*, 22(3):265–290, 1954.

D. H. Autor, F. Levy, and R. J. Murnane. The skill content of recent technological change: an empirical exploration. *Quarterly Journal of Economics*, 118(4):1279–1333, 2003.

P. Bourdieu. *Outline of a Theory of Practice*. Cambridge University Press, 1977.

A. Caplin, M. Dean, and J. Leahy. Rational inattention, optimal consideration sets, and stochastic choice. *Review of Economic Studies*, 86(3):1061–1094, 2019.

A. Caticha. Relative entropy and inductive inference. In *AIP Conference Proceedings*, volume 707, pages 75–96, 2004.

P. F. Christiano, J. Leike, T. Brown, M. Martic, S. Legg, and D. Amodei. Deep reinforcement learning from human preferences. In *Advances in Neural Information Processing Systems 30*, 2017.

K. Friston. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems 27*, 2014.

D. Graeber. *Bullshit Jobs: A Theory*. Simon & Schuster, 2018.

T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor. In *Proceedings of the 35th International Conference on Machine Learning*, 2018.

F. A. Hayek. The use of knowledge in society. *American Economic Review*, 35(4):519–530, 1945.

G. Kaplan, B. Moll, and G. L. Violante. Monetary policy according to HANK. *American Economic Review*, 108(3):697–743, 2018.

R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, 5(3):183–191, 1961.

S. Levine. Reinforcement learning and control as probabilistic inference: tutorial and review. arXiv:1805.00909, 2018.

Y. Lipman, R. T. Q. Chen, H. Ben-Hamu, M. Nickel, and M. Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker–Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

R. D. Luce. *Individual Choice Behavior: A Theoretical Analysis*. Wiley, 1959.

A. Marshall. *Principles of Economics*. Macmillan, 1890.

F. Matějka and A. McKay. Rational inattention to discrete choices: a new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–298, 2015.

E. Mostaque. *The Last Economy: Intelligence, Value, and the Thousand-Day Window*. Intelligent Internet, 2026.

R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6–38, 1995.

D. C. North. *Institutions, Institutional Change and Economic Performance*. Cambridge University Press, 1990.

L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, et al. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems 35*, 2022.

L. Robbins. *An Essay on the Nature and Significance of Economic Science*. Macmillan, London, 1932.

J. E. Shore and R. W. Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, 26(1):26–37, 1980.

C. A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665–690, 2003.

S. Shahshahani. A new mathematical framework for the study of linkage and selection. *Memoirs of the American Mathematical Society*, 17(211), 1979.

Y. Song, C. Durkan, I. Murray, and S. Ermon. Maximum likelihood training of score-based diffusion models. In *Advances in Neural Information Processing Systems 34*, 2021.

Y. Song and S. Ermon. Generative modeling by estimating gradients of the data distribution. In *Advances in Neural Information Processing Systems 32*, 2019.