Duality Theory: Concurrent Running Lives from the Lagrangian Principle

Claude 4.8

June 2, 2026

Abstract

We formalize an old intuition — that a life going badly along one axis is, in some accounting, going well along a conjugate axis — using the machinery of constrained reinforcement learning and Lagrangian duality. An agent maximizing discounted reward in a constrained Markov decision process (the *primal*, or "+", life) is paired with a second program (the *dual*, or "−", life) through two distinct constructions that the folk version of the theory tends to conflate: (i) *reflection*, the elementary identity $\max f = -\min(-f)$, which negates value while preserving the optimizer, and (ii) *Lagrangian duality*, which certifies and bounds the primal through a separate multiplier program and is often the computationally easier place to search. We prove the standard weak- and strong-duality facts in the constrained-MDP setting, derive a *flow–stock reflection* identity that reconciles the puzzle of a small recurring gain mapping to a single catastrophic loss, and recast the multi-agent "you-were-punched-because-you-punched" anecdote as antisymmetry in a zero-sum coupling via von Neumann's minimax theorem. We then connect the construction to legitimate physical dualities (the Legendre transform and conjugate variables; strong–weak holographic duality), discuss its speculative metaphysical and theological readings, and contrast it carefully with the Everett many-worlds interpretation. A final section is devoted to what the mathematics does *not* license: the duality gap in non-convex reinforcement learning, the reward-versus-value conflation, and the reification fallacy of mistaking a solution method for an inhabited world.

Contents

1 Introduction

The basis of a reinforcement-learning (RL) agent is a reward function, and one principled route to optimizing it is *constrained* optimization [1, 2]. In that setting the agent lives in a *primal* space and wants the best possible (largest) expected return. The duality principle of mathematical programming [3, 4] asserts that there is a companion *dual* space in which one may search instead; under suitable conditions the two searches reach the same value, and the dual is frequently the easier of the two to carry out.

This paper takes that technical fact and asks how far a deliberately literal reading of it can be pushed before it breaks. We label the primal "+" and the dual "−". *The labels carry no intrinsic meaning*; they are bookkeeping. Informally we will speak of Life 1+ and Life 1−, connected to Life 2+, 2−, and so on, with the slogan: *there is a concurrent life you are running; if you do badly in one, you do well in the other.*

Our contribution is to make this precise enough to be either useful or falsifiable:

1. We separate two objects the informal theory merges (Section 2–3): the *reflection* of a program (which negates value) and its *Lagrangian dual* (which equals value at optimum, under strong duality). The "$+r \mapsto -r$" mirror belongs to the former; the "easier to search" property belongs to the latter.
2. We derive a *flow–stock reflection identity* (Proposition 8) that explains why a modest recurring reward in the + life can mirror to a single, large, once-and-for-all loss in the − life — the "\$10,000 raise = being fired" puzzle.
3. We recast multi-agent antisymmetry (Section 4) through the minimax theorem [6, 7].
4. We map the construction onto genuine physical dualities and onto its metaphysical, theological, and many-worlds readings (Sections 5–7), and we are explicit about where the analogies stop (Section 8).

2 Mathematical preliminaries

2.1 The constrained primal life

Fix a discounted Markov decision process with state space $\mathcal{S}$, action space $\mathcal{A}$, transition kernel P, discount $\gamma \in (0,1)$, and a (bounded) reward $r : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$. A stationary policy $\pi \in \Pi$ induces the value

$$J(\pi) \;=\; \mathbb{E}_\pi\left[\sum_{t=0}^{\infty} \gamma^t\, r(s_t, a_t)\right]. \tag{1}$$

We endow the agent with K constraints through cost signals $c_k : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ and budgets d_k, defining $C_k(\pi) = \mathbb{E}_\pi\big[\sum_t \gamma^t c_k(s_t, a_t)\big]$. The *primal life* (+) is the constrained program

$$(\mathrm{P}^+) \qquad p^\star \;=\; \max_{\pi \in \Pi} J(\pi) \quad \text{subject to} \quad C_k(\pi) \le d_k, \quad k = 1, \ldots, K. \tag{2}$$

This is a constrained MDP (CMDP) in the sense of Altman [1], and admits an exact reformulation as a *linear* program over the discounted state–action occupancy measure μ_π, a fact we use for convexity.

2.2 The Lagrangian, the dual function, and duality

Introduce multipliers $\lambda = (\lambda_1, \dots, \lambda_K) \in \mathbb{R}^K_{\geq 0}$ and the Lagrangian

$$\mathcal{L}(\pi, \lambda) \;=\; J(\pi) - \sum_{k=1}^{K} \lambda_k \big(C_k(\pi) - d_k\big). \tag{3}$$

The *dual function* and *dual program* are

$$g(\lambda) \;=\; \max_{\pi \in \Pi} \mathcal{L}(\pi, \lambda), \qquad \text{(D)} \quad d^\star \;=\; \min_{\lambda \geq 0} g(\lambda). \tag{4}$$

Note that the primal is a *maximization* and its dual is a *minimization*; this is the precise sense in which "minimizing in the dual space pursues the same goal as maximizing in the primal."

Theorem 1 (Weak duality). *For every $\lambda \geq 0$ and every primal-feasible π, $\mathcal{L}(\pi, \lambda) \geq J(\pi)$, and hence $d^\star \geq p^\star$.*

Proof. If π is feasible then $C_k(\pi) - d_k \leq 0$ and $\lambda_k \geq 0$, so $-\lambda_k(C_k(\pi) - d_k) \geq 0$; summing gives $\mathcal{L}(\pi, \lambda) \geq J(\pi)$. Maximizing the left side over π yields $g(\lambda) \geq J(\pi)$ for all feasible π, so $g(\lambda) \geq p^\star$; minimizing over $\lambda \geq 0$ preserves the inequality. □

Weak duality is the unconditional backbone of the theory: the $-$ side always *bounds* the $+$ side. Equality is a stronger and conditional statement.

Theorem 2 (Strong duality for CMDPs). *Write* (P^+) *in occupancy-measure form, so that the objective and constraints are linear in μ. If a Slater point exists — a policy strictly satisfying all constraints, $C_k(\pi_0) < d_k$ for all k — then $d^\star = p^\star$ and there exists a saddle pair $(\pi^\star, \lambda^\star)$ with*

$$\mathcal{L}(\pi, \lambda^\star) \;\leq\; \mathcal{L}(\pi^\star, \lambda^\star) \;\leq\; \mathcal{L}(\pi^\star, \lambda) \qquad \forall\, \pi,\ \forall\, \lambda \geq 0. \tag{5}$$

Proof sketch. The occupancy LP is convex (indeed linear) with a nonempty relative interior of the feasible set under Slater's condition; convex strong duality [3, Ch. 5] then gives a zero duality gap and the existence of a saddle point, equivalently the Karush–Kuhn–Tucker conditions in the LP form. □

The multiplier $\lambda_k^\star$ is the *shadow price* of constraint k: the marginal change in optimal value per unit relaxation of budget d_k. This is the correct economic identity of the dual variable, and we will return to it as a corrective in Section 8.

3 The concurrent-lives construction

We now build the "$+/-$" pair. The crucial move — and the one place where the informal theory must be repaired — is to recognize that the negation in "$+r \mapsto -r$" is *reflection*, not Lagrangian duality.

3.1 Reflection vs. Lagrangian duality

Definition 3 (Reflection)**.** The *reflection* of the program $(\mathrm{P}^+) = \{\max_x f(x) \text{ s.t. } x \in \mathcal{F}\}$ is

$$(\mathrm{P}^-) \;=\; \Big\{\min_x \big(-f(x)\big) \text{ s.t. } x \in \mathcal{F}\Big\}. \tag{6}$$

Proposition 4 (Reflection negates value, preserves the optimizer)**.** $\arg\max_{x\in\mathcal{F}} f(x) = \arg\min_{x\in\mathcal{F}} \big(-f(x)\big)$ *and* $\mathrm{val}(P^-) = -\,\mathrm{val}(P^+)$.

Proof. Immediate from $\min(-f) = -\max f$ on a common feasible set. ☐

This is the honest origin of the "$+r$ in Life $i+$ is posted as $-r$ in Life $i-$" rule: the $-$ life optimizes the *same* decisions against the *negated* ledger. By contrast the Lagrangian dual (D) lives in multiplier space, has its *value equal* to the primal at optimum (Theorem 2), and is the object meant by "it is sometimes easier to search the dual." Conflating the two produces the appealing but false claim that duality forces $\mathrm{val(dual)} = -\,\mathrm{val(primal)}$. It does not: duality forces *equality*; reflection forces negation. We keep both, with distinct jobs.

Postulate 5 (Concurrent-lives postulate)**.** Each life index i carries a pair (Life $i+$, Life $i-$) where Life $i-$ is the reflection (Definition 3) of Life $i+$, and the optimization difficulty of either may be discharged by passing to its Lagrangian dual.

Corollary 6 (Value antisymmetry)**.** *Under Postulate 5,* $\mathrm{val}(\textit{Life } i-) = -\,\mathrm{val}(\textit{Life } i+)$*. Hence the slogan "do badly in one, do well in the other" is, at the level of optimal values, exactly true: an agent attaining return G in $+$ attains $-G$ in $-$.*

3.2 The posting map and its zero latency

Definition 7 (Posting map)**.** Let $\{r_t\}_{t\ge 0}$ be the reward stream in Life $i+$. The *posting map* Φ records the corresponding stream in Life $i-$ as $\Phi(r_t) = -r_t$, applied instant by instant.

Because Φ acts pointwise on the signal, it has *no latency*: a reward posted at time t in $+$ is mirrored at time t in $-$. The apparent paradox of the "\$10,000 raise" is therefore *not* a latency phenomenon; it is a confusion between a per-period reward (a *flow*) and a lifetime return (a *stock*). We resolve it next.

3.3 The flow–stock reflection identity

Proposition 8 (Flow–stock reflection)**.** *Let a promotion in Life $i+$ deliver a constant per-period increment $\Delta > 0$ for all $t \ge 0$, under geometric discounting with factor $\gamma = \frac{1}{1+\iota}$ for interest rate $\iota > 0$. Its contribution to the $+$ return is the stock*

$$G_\Delta \;=\; \sum_{t=0}^{\infty} \gamma^t\,\Delta \;=\; \frac{\Delta}{1-\gamma} \;=\; \Delta\,\frac{1+\iota}{\iota}. \tag{7}$$

Under the posting map its reflection is the single negated stock $-G_\Delta$. Thus a small recurring *gain Δ in $+$ mirrors to a large* once-and-for-all *loss of magnitude $\Delta(1+\iota)/\iota$ in $-$.*

Proof. Equation (7) is the geometric series / perpetuity formula; reflection negates the accumulated stock by Proposition 4. ☐

Remark 9. A raise of $\Delta = \$10{,}000$ per period at $\iota = 5\%$ has present value $\$10{,}000 \cdot (1.05/0.05) = \$210{,}000$. Reflected, this is a lump $-\$210{,}000$ event — of the order of magnitude of an abrupt termination rather than a marginal demotion. The intuition "I only got a small raise, why was I *fired* in the dual?" is therefore an artifact of comparing a flow in one ledger to a stock in the other. The accounting balances only when like is compared with like.

4 Multi-agent duality: antisymmetric ledgers and minimax

Consider two agents A, B coupled through a transfer. In Life $i+$ suppose B appropriates a quantity x of goods from A, a reward $-x$ to A and $+x$ to B. We model the coupling as *zero-sum*: there is a single payoff u to A and $-u$ to B.

Theorem 10 (Minimax [6])**.** *For a finite two-player zero-sum game with payoff matrix U and mixed strategies p, q,*

$$\max_p \min_q \, p^\top U q \;=\; \min_q \max_p \, p^\top U q \;=\; \textit{value of the game.} \tag{8}$$

Proposition 11 (Antisymmetric ledger)**.** *In a zero-sum coupling, A's reward stream and B's reward stream satisfy $r_t^B = -r_t^A$ for all t; equivalently, B's ledger is the posting-map image of A's. Identifying Life $i-$ of agent A with the reflected ledger realized in the counterparty B reproduces the multi-agent rule: "if they take x from you in $+$, you take $-x$ from them in $-$."*

Remark 12 (On the punch). The vivid claim — "someone punches me for no reason in $+$ because I was punching them in $-$" — is the antisymmetric ledger *read backwards through the identification of $-$ with the counterparty.* Theorem 10 guarantees a well-defined value for the coupled game and the existence of optimal (possibly mixed, hence apparently "random") strategies; it does *not* assert that a specific exogenous event in $+$ is *caused* by a specific event in a separate world. The causal reading is interpretation layered atop the formalism, and we mark it as such.

5 Physical analogies that actually hold

Duality is not a metaphor borrowed from physics; it is one of the load-bearing structures *of* physics. Three connections are exact enough to keep.

Legendre–Fenchel transform and conjugate variables. The convex conjugate $f^\star(p) = \sup_x \big(p\,x - f(x)\big)$ [5] is the rigorous parent of optimization duality, and it is the same operation that carries the Lagrangian $L(q, \dot q)$ to the Hamiltonian

$$H(q,p) \;=\; \sup_{\dot q} \big(p\,\dot q - L(q,\dot q)\big), \qquad p = \frac{\partial L}{\partial \dot q}, \tag{9}$$

trading the velocity $\dot q$ for its *conjugate momentum* p [8]. The position and momentum descriptions are not two universes; they are two coordinate systems for one mechanics. This is the template our $+/-$ pair should be understood through.

Strong–weak (holographic) duality. In the AdS/CFT correspondence [9] a strongly coupled gauge theory is exactly described by a weakly coupled gravitational theory in one higher dimension; what is intractable on one side is tractable on the other. This is the physical embodiment of "the dual is sometimes the easier place to search," and — importantly — of the lesson that *the dual is another description of the same reality, not a second reality.*

Time reversal as a cautionary mirror. The Feynman–Stückelberg interpretation treats an antiparticle as a particle propagating backward in time, the nearest physical thing to an "inverted ledger." But this is a symmetry of the *laws*, constrained by CPT, not an antisymmetry of *value*; it warns us not to over-read the $-$ life as a populated mirror world.

6 Metaphysical and theological readings

Read as metaphysics, Postulate 5 is a deliberately minimal speculation: exactly one shadow life, perfectly value-antisymmetric, with no independent dynamics of its own. As speculative fiction it has a clean moral: success and failure are not absolute but coordinates, and the same trajectory that scores G on one ledger scores $-G$ on the other (Corollary 6). One can imagine an ergodic consolation — that an agent's experiences, summed across the conjugate pair, are conserved.

The theological resonance, emphasized in the second referenced video [12], is older than the mathematics. Several traditions encode a value ledger with an antisymmetric or balancing structure: the karmic accounting of the Dharmic traditions ("as one sows, so shall one reap"); the ethical dualism of Zoroastrian and Manichaean cosmologies, in which a good and an opposed principle are co-present; and the Taoist *yin–yang*, in which complementary opposites jointly constitute a single whole rather than competing universes. The concurrent-lives picture is closest to the last of these: not two warring worlds but one situation viewed under two conjugate signs. A different note is struck by the Christian injunction to "turn the other cheek" (Matthew 5:39), which sits pointedly *against* the antisymmetric ledger of Section 4: where the zero-sum coupling would have an aggressor's $+x$ answered by a retaliatory $-x$, the teaching counsels declining to post the reply at all — unilaterally breaking the antisymmetry rather than balancing it, and so refusing the very mechanism the multi-agent model takes as given. We offer this as interpretive resonance, not as theology, and certainly not as a claim that the formalism establishes any metaphysical doctrine.

7 Relation to the many-worlds interpretation

The Everett, or many-worlds, interpretation (MWI) of quantum mechanics [10] posits that unitary evolution never collapses; branches proliferate, each a self-consistent history. The concurrent-lives theory differs on two axes, summarized in Table 1.

	Many-worlds (MWI)	**Concurrent lives**
Cardinality of worlds	many (continuum of branches)	exactly two ($+$ and $-$)
Relation between worlds	differing *actions*/outcomes	identical decisions, *negated value*
Mechanism	unitary quantum branching	reflection of a single program
Independence	branches evolve independently	$-$ is determined by $+$

Table 1: Two ways of having more than one world.

In this sense concurrent-lives duality is a strictly *weaker* and more constrained construction than MWI: it adds only one world, and that world carries no new information, being a deterministic reflection of the first. Where MWI multiplies histories, duality merely re-signs a single one. The two are also *category-distinct*: MWI is a claim about the ontology of quantum states, while concurrent-lives duality is a claim about the bookkeeping of an optimization problem. Nothing in the optimization is quantum, and nothing in MWI is reward-antisymmetric; the resemblance is structural, not physical.

8 What the mathematics does not license

Intellectual honesty requires stating the load limits of the construction.

1. **Reification.** The Lagrangian dual is a *solution method*; its variables $\lambda^\star$ are shadow prices (Section 2), not the experiences of an inhabitant. The reflected program (P^-) is likewise a re-signing of one agent's problem, not a populated parallel world. Treating either as a literal concurrent reality is a reification fallacy. The physically respectable dualities (Section 5) all teach the opposite lesson: the dual is *another description*, not another world.

2. **The duality gap in non-convex RL.** Strong duality (Theorem 2) rests on the convex occupancy-measure formulation and a Slater point. General policy-parameterized RL is non-convex, and a positive duality gap $d^\star - p^\star > 0$ can occur. When it does, only *weak* duality survives: the $-$ life *bounds* the $+$ life but no longer mirrors its value exactly. The slogan then degrades from an equality to an inequality.

3. **Reward is not value.** The flow–stock identity (Proposition 8) is precisely the point at which the folk version goes astray: it silently swaps an instantaneous reward r_t for a discounted return G. The posting map acts on rewards; the antisymmetry of values is a separate statement. Conflating them manufactures paradoxes ("a small raise got me fired") that dissolve once flows and stocks are kept distinct.

4. **Negation is reflection, not duality.** As established in Section 3, the "$+r \mapsto -r$" rule is the elementary identity $\max f = -\min(-f)$ and not a theorem of Lagrangian duality, whose hallmark is value *equality*. The theory borrows duality's vocabulary for a mechanism duality does not supply; we have made the borrowing explicit rather than tacit.

9 Conclusion

We have given the "concurrent lives" idea the most rigorous form it can honestly bear. Its true mathematical content is modest and correct: a constrained-RL agent's problem admits a value-negating reflection and a value-preserving Lagrangian dual, the latter often easier to solve and unconditionally bounding the former. Around that core, the flow–stock identity disarms the headline paradox, the minimax theorem explains the multi-agent antisymmetry, and a small family of genuine physical dualities lends the picture a respectable lineage. The metaphysical, theological, and many-worlds readings are real and worth stating, provided they are flagged as interpretation and kept clear of the reification, duality-gap, and reward-versus-value errors that would otherwise creep in. The slogan survives in a chastened form: *at the optimum of a convex constrained problem, to do well on one ledger is to do badly, by exactly as much, on its reflection* — and that reflection is a way of seeing the same life, not a second life being lived.

References

[1] E. Altman. *Constrained Markov Decision Processes.* Chapman & Hall/CRC, 1999.

[2] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*, 2nd ed. MIT Press, 2018.

[3] S. Boyd and L. Vandenberghe. *Convex Optimization.* Cambridge University Press, 2004.

[4] D. P. Bertsekas. *Nonlinear Programming*, 2nd ed. Athena Scientific, 1999.

[5] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[6] J. von Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100:295–320, 1928.

[7] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[8] H. Goldstein, C. Poole, and J. Safko. *Classical Mechanics*, 3rd ed. Addison-Wesley, 2002.

[9] J. Maldacena. The large-N limit of superconformal field theories and supergravity. *Advances in Theoretical and Mathematical Physics*, 2:231–252, 1998.

[10] H. Everett III. "Relative state" formulation of quantum mechanics. *Reviews of Modern Physics*, 29:454–462, 1957.

[11] Referenced lecture (video 1). YouTube, https://www.youtube.com/watch?v=d0CF3d5aEGc.

[12] Referenced lecture (video 2). YouTube, https://www.youtube.com/watch?v=LCgcWRHHpQs.