Loops and the Geometry of Chance

Jens Jäger

forthcoming in *Noûs*
penultimate draft (January 8, 2025)

Abstract

Suppose your evil sibling travels back in time, intending to lethally poison your grandfather during his infancy. Determined to save grandpa, you grab two antidotes and follow your sibling through the wormhole. Under normal circumstances, each antidote has a 50% chance of curing a poisoning. Upon finding young grandpa, poisoned, you administer the first antidote. Alas, it has no effect. The second antidote is your last hope. You administer it—and success: the paleness vanishes from grandpa's face, he is healed. As you administered the first antidote, what was the chance that it would be effective? This essay offers a systematic account of this case, and others like it. The central question is this: Given a certain time travel structure, what are the chances? In particular, I'll develop a theory about the connection between these chances and the chances in ordinary, time-travel-free contexts. Central to the account is a Markov condition involving the boundaries of spacetime regions.

Contents

1 Introduction

You're a star pharmacist, and you've invented a universal antidote, able to cure any poisoning. Unfortunately, the antidote isn't perfectly reliable: normally, given any poisoning, there's a 50% chance that it'll cure it. One day, your evil sibling travels back in time, intending to lethally poison your grandfather, back when he was still an infant. Determined to save grandpa, you grab two antidotes and follow your sibling into the wormhole. ("Better to bring more than one!", you think.) Upon finding infant grandpa, poisoned, you administer the first antidote. Alas, it doesn't work. The second antidote is your last hope. You administer it—and success: the paleness vanishes from grandpa's face, he is cured.

As you administered the first antidote, what was the chance that it would be effective? Perhaps 0? After all, it already failed: its failure is what causes the second antidote's success, which causes grandpa's survival, which causes your being born... On the other hand, the antidote's failure is also still future—some time will pass until it occurs—and the present leaves it open which of the two antidotes is ultimately effective. So perhaps the chance is 0.5, because that's what it normally is? No. I'll argue that, on a salient interpretation of "as you administered the first antidote", the answer is 2/3.

The essay's broader question is this: *Given* a time travel structure, what are the chances? This is distinct from asking what the chances are of wormholes and other time-travel structures arising in the first place—a question reserved for physicists studying the mechanisms behind such structures.

Why care about *our* question? Where time travel involves *spacetime loops*—informally, trajectories which travel back in time to their origins—it has been of continued interest to physicists.[1] Part of a philosopher of science's responsibility is to interpret what physicists study.

But more importantly, our inquiry yields significant philosophical insights. Previously, philosophical investigations into chance have largely assumed standard spacetime backgrounds.[2] As we'll see, this practice misses important lessons. Assuming that spacetime loops are metaphysically possible, our account challenges two orthodoxies about chance. First, it'll show that chances aren't as intimately tied to time or causation as is usually thought; and second, that chances aren't necessarily constant across intrinsically duplicate trials. I'll replace these orthodoxies with a view on which chances are tied, not to temporal histories, but to *chance setups*, and on which chances on loops differ from the ordinary chances in systematic, scrutable ways. Our final account provides a complete theory for

[1] E.g., Gödel (1949), Carter (1968), Echeverria, Klinkhammer, and Thorne (1991), and Earman (1995).

[2] *Viz.*, globally hyperbolic Lorentzian backgrounds, or classical Newtonian or Galilean backgrounds.

chances on loops, for any setup.[3]

The account is founded on two principles about the relationship between chance and spacetime structure. One key principle concerns the familiar idea of "screening off". In assessing the probability of some proposition C, we say that A screens off B from C if A renders any information provided by C about B irrelevant. For example, suppose I'd like to know if I carry a certain genetic marker. Given complete information about my *parents'* genes (A), no information about my *grandparents'* genes (C) should affect my confidence of my carrying the gene (B): complete information about A *screens off* any information C would provide about B. With respect to *chance*, what screens off what is partially determined by spacetime structure. Specifically, I defend the idea that, given a local dynamics, what's happening at a spacetime region's *boundary* screens off what happens on the region's inside from what happens on its outside. This provides a systematic connection between chances on loops and the ordinary, time-travel-free chances.

The essay is structured as follows. To streamline the discussion, section 2 introduces a simple, stripped-down spacetime loop scenario—grandpa will reappear later. In Sections 3–5, I survey two accounts of this case. One is based on the orthodox "temporalist" framework of chance, promulgated by Lewis (1987); the other is based on the idea that chances are invariant across intrinsically duplicate trials. The first account often trivializes chances on spacetime loops, and the second account leads to inconsistency. Both should thus be rejected. This sets the stage for my positive proposal. In Sections 6 and 7, I explain the proposal's two core principles, *Acyclic Chance Invariance* and *Strong Boundary Markov*. Section 8 applies these two principles to the simple loop scenario. Section 9 revisits the stochastic grandfather paradox and develops a general recipe for deriving chances on loops. Section 10 corrects a misconception regarding chances on loops. It shows that, in stark contrast to the acyclic case, in *cyclic* spacetimes the dynamics alone manages to fix *unconditional* chances over the states of the universe, thus generating a "complete probability map of the world". Similarly, the dynamics generically fixes precise expectations about what will emerge from future wormholes. Section 11 concludes.

2 A Simple Case

The previous scenario involves a spacetime loop: your saving your infant grandfather leads to him growing up and having children, one of whom bears you and your sibling,

[3] Provided the background geometry is static, as e.g. in special relativity. A generalization to theories with dynamical spacetime structure (such as general relativity) is a topic for future work. Still, I hope the present proposal marks significant progress in that direction.

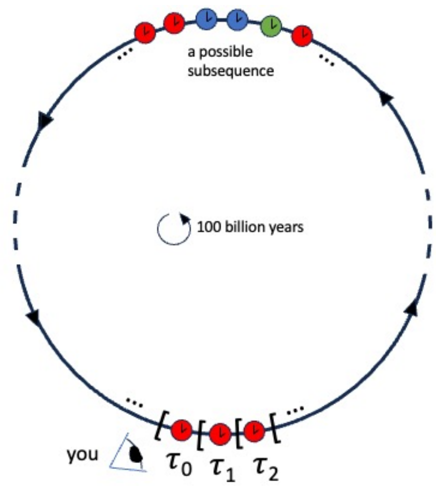


Figure 1: A sketch of CIRCLE.

which leads you to eventually enter the wormhole, ending up in front of infant grandfather. For the purpose of discovering general principles about cases like this, let's start with a particularly simple spacetime loop world, called CIRCLE.

Imagine a single stationary particle, occupying a single point of space, persisting in circular time. For concreteness, let the cycle have a period of 100 billion years—that is, persisting for 100 billion years from now gets the particle back to the present. Heuristically, you may picture the spacetime as an infinitely thin strip of paper, with an arrow drawn parallel to the strip, and whose ends you've glued together. This circular strip is a crude representation of a universe consisting of a single point of space, persisting in circular time, with the arrow indicating the time's direction. Now picture a small marble sitting at every point along the strip. The marbles represent the particle's different time slices, successively occupied by the particle as it persists through time. Together, this yields a crude representation of a single particle in a circular, one-dimensional spacetime.[4]

Let's stipulate a simple dynamics for the particle. Let the particle have two intrinsic magnitudes, *color* and *clock*. The particle's clock grows in proportion to the time passed, until it reaches 24 hours, at which point it restarts from 0. The particle sometimes changes colors, exactly at clock restart points. You only ever observe three possible colors and three possible transitions: *red to green*, *green to blue*, and *blue to red*. (See fig. 1.) Moreover, you observe a color change at about 2 out of 10 reset points.

[4]Mathematically, we can represent CIRCLE by a one-dimensional, oriented, closed Lorentzian (or, equivalently in the one-dimensional case, Riemannian) manifold—that is, a circle equipped with a metric. Note that the one-dimensional is illustrative not only because it's particularly simple, but also because one-dimensional oriented spaces commonly appear as base spaces in fibre bundle constructions of other spaces, e.g. of Galilean spacetime or of the configuration-space-in-time relevant to some interpretations of quantum mechanics.

Suppose the particle is currently red, and you'd like to calculate the chance that it'll still be red tomorrow. Two hypotheses may jump out, paralleling those about the stochastic grandfather paradox. One might think that the chance that the particle changes color at the next reset point is 1 if it actually changes color, and 0 if it doesn't. After all, whatever the particle does has *already* happened. But then again, whatever happens is *still* to happen, and the particle's current color leaves tomorrow's color open. So perhaps the chance of a color switch is just what it ordinarily is. The following sections survey two proposals, capturing the two hypotheses. I argue that both should be rejected.

3 Against Temporalism

Chances vary in time: Wilbur and Orville Wright flipped a coin to settle who would fly first. As the coin was flipped, both had a positive chance of being the first to fly. But because Orville lost the flip, *today* there's chance 0 that he flew first. Lewis's (1987) framework therefore relativizes chances to times:

Temporalism about chance:

1. Necessarily, chance is a function of two arguments, a proposition and a time.[5]
2. Necessarily, if chance is the function ch, then for any time t, $ch(\cdot, t)$ assigns chance 1 to t's temporal history, i.e. to the strongest truth "entirely about matters of particular fact" (Lewis, 1987) at times at or before t.[6]

The second clause causes trouble. In CIRCLE, every time precedes every other time. Given this, clause 2 trivializes all chances in CIRCLE: it implies that every chance function assigns

[5]Officially, Lewis adds a *third* argument, a world, where "$ch(A, t, w)$" (or "$P_{tw}(A)$" in his notation) refers to "the chance, at time t and world w, of A's holding" (87). By contrast, we're setting up chance as a *contingent* relation here (between a proposition, time, and real number). I find this latter setup more perspicuous, because it's neutral on the underlying account of contingency (e.g., if it's worlds-based, or something else). But nothing of significance will hang on this here.

[6]Actually, the original quote says "entirely about matters of particular fact at times *no later than t*" (Lewis, 1987). This, incidentally, doesn't fall prey to the objection below: since in CIRCLE every time is later than any other time, this account simply places no constraints at all on $ch(\cdot, t)$ in CIRCLE. Now, it's clear that Lewis himself doesn't use "no later than t" to distinguish it from "at or before t". (As evidenced a few paragraphs later, where he writes, of a proposition A about states of affairs at time t_A, that "[i]f t is later than t_A, then A is admissible at t" (Lewis, 1987). Temporalism entails that A is admissible, and hence receives chance 1, at t only if "t is later than t_A" entails "t_A is *no later* than t". But "t is later than t_A" is equivalent to "t_A is before t", and so Lewis must assume that "t_A is before t" entails "t_A is no later than t". Since Lewis also thinks that A is admissible at t if t_A is *simultaneous* with t, it follows that he must assume that "t_A is at or before t" entails "t_A is no later than t".) In any case, an account which places no constraints at all on $ch(\cdot, t)$ in CIRCLE is seriously incomplete. For example, surely $ch(\cdot, t)$ will at least assign chance 1 to the state of the world at t. Temporalism, as formulated in the main text, is thus a proposal for filling out the account.

chance 1 to the particle's actual complete color history, and 0 to every non-actual history. As a valid argument:

(i) For all times t and t', t' is at or before t.

(ii) For all times t and t', if t' is at or before t, then the temporal history of t entails what color the particle is at t'.

(iii) For any time t, $ch(\cdot, t)$ assigns chance 1 to t's temporal history.

(iv) For all propositions A and B and times t, if A entails B and $ch(A, t) = 1$, then $ch(B, t) = 1$.

∴ For all times t and t', $ch(\cdot, t)$ assigns chance 1 to what color the particle is at t'.

Premise (ii) follows from the definition of temporal history; (iii) follows from temporalism's clause 2; and (iv) follows from the fact that $ch(\cdot, t)$ is a probability function. Premise (i), meanwhile, is supported by two thoughts: (a) there is some small duration ε such that, for any t, all times no more than ε into t's past are before t, and (b) "before" is transitive. While it's logically possible to either deny (a) or deny (b), neither possibility seems attractive. Regarding (a): surely, if *any* times are before t in CIRCLE, it includes those in t's most recent past. Meanwhile, holding that *no* times are before t falsifies evident truths: for example, despite the particle's changing color from red to green, it wouldn't be the case that the particle has *previously* been red. Moreover, the reply would at best secure silence about CIRCLE. Yet our aim is a positive theory about chances on loops. Regarding (b): denying transitivity burdens us with arbitrary cutoffs—when is t' just far enough in t's past that it's no longer "before" t? I see no principled way to draw this distinction.

The trivialization of chances in CIRCLE is problematic for two connected reasons.[7] First, recall the regularity you observe: whenever the particle has a given color, in about 2 out of 10 cases it'll have a different color the next day. It would seem extremely natural—and useful, and informative—to try to describe this behavior in non-trivial chance-theoretic terms. Indeed, it seems just as natural to do so in CIRCLE as it does in any *linear* world. But that's incompatible with Clause 2.

[7] Lewis (1987) is aware that temporalism has issues with time travel. He notes that the existence of time travelers may make some past information inadmissible: "That is why I qualified my claim that historical information is admissible, saying only that it is so 'as a rule'." (ibid., 274) But Lewis mentions this problem only to discard it: he states that he merely wants to argue that "the Principal Principle captures our common opinions about chance" and those common opinions, he says, "may rest on a naive faith that past and future cannot possibly get mixed up". (ibid., 274) I find it doubtful that our common opinions include any clear judgment about the possibility of time travel. In any case, Lewis admits that "[a]ny serious physicist, if he remains at least open-minded both about the shape of the cosmos and about the existence of chance processes, ought to do better" (ibid., 274). Philosophers should, too.

Secondly, a universe with circular time is still compatible with our empirical evidence.[8] If this speculative scenario was true, would it falsify all scientific theories involving non-trivial chances? Would it mean, for example, that radioactive decay wasn't stochastic after all? The answer is clearly no. But this contradicts temporalism's clause 2. So we should reject temporalism.[9]

4 Urchance

The problem with temporalism is that its chance functions must be certain of complete temporal histories. To deliver non-trivial chances in cyclic spacetimes, the correct chance theory has to be more flexible.

As Hall (2004) observes, temporalism can be reformulated by stipulating the existence of an "urchance" function, given to us by the fundamental physical laws. The chance function at a time t, ch_t, is then the result of conditioning said urchance function on the temporal history up to t, H_t, i.e., $ch_t(\cdot) = \text{urchance}(\cdot|H_t)$. One advantage of the urchance approach is that it automatically ensures that chance functions at different times "cohere" with one another, *viz.* that chance functions at later times result from conditioning those at earlier times on intervening history. Coherence is mandated by canonical chance deference principles like the Principal Principle.[10]

Now, the temporalist considers *only* the results of conditioning the urchance function on complete temporal histories. But sometimes the result of conditioning the urchance function on *non-history* propositions is also well-defined. This much follows already from the probability laws alone: where urch is the urchance function and H_t the world's history up to t, whenever $\text{urch}(\cdot|H_t)$ assigns positive probability to some A, $\text{urch}(\cdot|H_tA) = \text{urch}(\cdot \wedge A|H_t)/\text{urch}(A|H_t)$ is a well-defined probability function even if H_tA isn't a

[8]For example, it is compatible with our empirical evidence that our universe is (representable by) a four-dimensional time-like closed Lorentzian manifold. Such a manifold is permitted by Einstein's field equations, and if spatially flat, accords with general astronomical observations about the shape of our universe. (The role of the Past Hypothesis in such a world would be played by the posit that there is a low-entropy macrostate *at some time*, with entropy increasing bidirectionally from there.)

[9]Cusbert (2018; 2022) has recently suggested replacing temporalism with a "causal history view" of chance: instead of having $ch(\cdot, t)$ assign chance 1 to t's temporal history, have it assign chance 1 to t's *causal* history instead. (Cusbert's formalism replaces "times" with "globally connected sets" of events—this difference doesn't matter here.) But this fares no better than temporalism in CIRCLE: the particle's color at each day is caused by the color the previous day. Moreover, *causal histories* are transitively closed (even if causation isn't). So Cusbert's view also trivializes chance in CIRCLE.

[10]To see this: the Principal Principle (in one of its canonical formulations, Lewis (1987)) says that, where Cr_0 is any rational initial credence function, t any time, H_t the world's actual history up to t, and T the true chance theory, $Cr_0(A|H_tT) = ch_t(A)$. Where $H_{[t,t^+]}$ is the intervening history from t to t^+, it follows that $ch_{t^+}(A) = Cr_0(A|H_{t^+}T) = Cr_0(A|H_{[t,t^+]}H_tT) = ch_t(A|H_{[t,t^+]})$, provided $H_{[t,t^+]}$ has positive chance at t.

complete temporal history. The fundamental dynamical laws go yet beyond this. Consider any bounded history segment. Since it is bounded, it doesn't entail any complete temporal history; yet, together with the fundamental dynamical laws, the segment's state generically entails a well-defined probability distribution over the segment's possible futures. (In the deterministic case, this probability distribution is trivial.) So, insofar as urchance encodes exactly the content of the fundamental dynamical laws, the result of conditioning it on non-history propositions is often well-defined too.

In my view, urchance *is* naturally thought of as encoding (exactly) the content of the fundamental dynamical laws. The results of conditioning it on non-history propositions are then objective chance functions: since they follow from the fundamental dynamical laws alone, they are objective; they are "single-case" probabilities—*viz.*, $\text{urch}(\cdot|B)$ exists and is generally non-trivial, even if an event described by B occurs only once; they respect dynamical symmetries; and they go with straightforward deference principles.[11]

Now, the fundamental dynamical laws aren't generally so powerful that $\text{urch}(A|B)$ is precisely defined for *all* physically specifiable A and B. Some B are too weak for some A: for example, a contingent proposition generally won't have a precise probability conditional on a logical tautology. In the following paragraphs, I outline a framework for capturing this predictive weakness using imprecise probability. While the framework underpins the remainder of the paper, it is designed so that the paper's main philosophical insights remain accessible without it. (Specifically, the formalism is designed so that all main text equations are interpretable under the simplifying, albeit false, assumption that urch is a precise probability function.) Readers who prefer to skip formalism may proceed to the final paragraph of this section.

To capture urchance's predictive weaknesses, we express it not in terms of a single probability function, but in terms of a *set* of probability functions. Intuitively, these functions are all the *precisifications* of the fundamental laws' probability judgments.[12] Accordingly, A has a precise chance x conditional on B iff all functions in the set assign A probability x conditional on B. Where the functions disagree, the best we can do is assign A a set of chance values conditional on B. Officially, we let these functions be two-place,

[11] One such principle, compatible with the Principal Principle: where Cr_0 is any rational prior credence function, A and E are any propositions, and $\langle \text{urchance}(A|E) = x \rangle$ is the proposition that the urchance of A, conditional on E, is x: $Cr_0(A|E \wedge \langle \text{urchance}(A|E) = x \rangle) = x$.

[12] Cf. van Fraassen (1984), who presents a set-based (or "representor"-based) formalism for credences. There are significant philosophical advantages to a set-based formalism over less expressive alternatives, such as one using partial urchance functions. For example, a set-based formalism can express probabilistic independencies even where no precise unconditional probabilities exist (cf. Joyce (2010))—this enables our definitions of Markov properties in Section 7. A set-based formalism also streamlines the proofs in Appendix C.

total, *primitively conditional* probability functions (cf. Hájek (2003)).[13]

Now, one way to proceed would be to identify urchance directly with the set of probability functions induced by the true fundamental dynamical laws. More conveniently, however, we define it as follows. Where $\mathbf{u}$ is the set of all precisifications of the dynamical laws' probability judgments, urchance $\text{urch}(\cdot|\cdot)$ is a function from pairs of propositions into functions on $\mathbf{u}$: specifically, for any pair (X, Y) of propositions, $\text{urch}(X|Y)$ is the function from $\mathbf{u}$ into $[0, 1]$ mapping $u \in \mathbf{u}$ to $u(X|Y)$—that is, $\text{urch}(X|Y)(u) = u(X|Y)$. Intuitively, $\text{urch}(X|Y)$ isn't merely a set of values, but (additionally) keeps track of which member of $\mathbf{u}$ assumes which value.

This way of defining urchance has the advantage that arithmetic operations on $\text{urch}(\cdot|\cdot)$ can be defined in the standard functional way, i.e., they apply point-wise. For example, $\text{urch}(A|B) \cdot \text{urch}(C|D)$ denotes the function which maps each $u \in \mathbf{u}$ to the value $u(A|B) \cdot u(C|D)$. Similarly, equality is point-wise: $\text{urch}(A|B) = \text{urch}(C|D)$ says that the both sides agree everywhere on their domain, i.e. for all $u \in \mathbf{u}$, $u(A|B) = u(C|D)$, while $\text{urch}(A|B) = x$ says that the left-hand side has constant value x, i.e. for all $u \in \mathbf{u}$, $u(A|B) = x$. Intuitively, whenever you see an expression like $\text{urch}(A|B)$ in an equation, read it as the equation's holding *determinately*, i.e. for all members of $\mathbf{u}$. (The same goes for inequalities, $>$, and approximate equalities, $\approx$.) Conditioning is equally straightforward: for any proposition B, $\text{urch}(\cdot| \cdot \wedge B)$ maps any pair (X, Y) of propositions to the function on $\mathbf{u}$ mapping each $u \in \mathbf{u}$ to $u(X|Y \wedge B)$.

Our definition entails that, whenever all members of $\mathbf{u}$ satisfy an equation involving only arithmetic operations and conditioning, urch satisfies an equation of the exact same form. These equations include the probability axioms. For example, since all members of $\mathbf{u}$ satisfy the multiplicative axiom, urch satisfies an expression of the same form, $\text{urch}(AB|C) = \text{urch}(A|BC) \cdot \text{urch}(B|C)$. Defining urchance as we have allows us to use syntactically familiar equations.

Where urch is the urchance function, I'll call B the "background proposition" for the function $\text{urch}(\cdot| \cdot \wedge B)$.[14] The urchance formalism enables theories of non-trivial chances in

[13] Primitive conditionality is needed because sometimes A has precise chance conditional on B even if B has (precisely) zero prior chance. Popper ([1959] 1968, App. *IV and *V) offers a convenient axiomatization of total primitively conditional probability. His axioms—specifically, the version with mere finitely additivity—are what I'll mean throughout by "the laws of probability". The axioms include the standard multiplicative axiom, $p(A \wedge B, C) = p(A, B \wedge C) \cdot p(B, C)$, finite additivity, and the probabilistic analogue of explosion, $p(A, B \wedge \neg B) = 1$. We'll also assume that any member of the set induced by some fundamental dynamical laws is "maximally sure" exactly of anything metaphysically entailed by those laws. That is, for all members p of the set, $p(A|\neg A) = 1$ iff A is metaphysically entailed by L. In particular, no member assigns maximal prior chance to any nomically contingent matter of fact.

[14] The terminology of "background proposition" is also used in Nelson (2009) and Cusbert (2018). Other authors with flexible chance formalisms include Meacham (2005), Nelson (2009), Briggs (2010), Handfield

cyclic worlds: chance functions with background propositions weaker than any temporal history can assign non-trivial probabilities to contingent matters of fact. The second account attempts to build on this flexibility to construct a theory of non-trivial chances—unsuccessfully so.

5 Against Chance Invariance

Consider again an infinitely thin strip of paper, but this time leave its ends unjoined. Call the represented world LINE (or L for short).[15] This world is inhabited by the same sort of particle as CIRCLE, subject to the same laws. Now suppose you are told that, in LINE, upon clock reset there's (precisely) a *0.2 chance* that the particle changes color (from red to green, green to blue, or blue to red). What can we infer from this about the chances in CIRCLE? What would the transition chances be if the spacetime was cyclic rather than linear?

A natural idea is that the chances would be unchanged—that transition chances in circular spacetimes are just what they are in identical "linear" situations. This is suggested by the popular idea that objective chances are, as Schaffer (2003) puts it, "stable". Arntzenius and Hall (2003, p. 178) express the idea as follows: "if ... two processes going on in different regions of spacetime are exactly alike, your [theory should assign] to their outcomes the same single-case chances". Or, as Schaffer (2007, p. 125) puts it, "chance values should remain constant across intrinsically duplicate trials" within the same world. While these principles strictly speaking only concern chance assignments within the *same* world, they have obvious and natural generalizations that also cover chance assignments *across* worlds with the same laws. These generalizations dictate the same transition chances for CIRCLE as for LINE.[16]

Arntzenius and Hall don't provide a formally precise version of their principle, and Schaffer's (2007) presentation assumes temporalism.[17] So let's formulate a principle ourselves, for the case of CIRCLE and LINE, using urchance. (Instead of "stability" I choose the label "invariance", which I find more fitting.) Informally, the idea is that the chance of

and Wilson (2014), and Cusbert (2018).

[15]Mathematically, we can represent the world by a one-dimensional, oriented, open Lorentzian (or, equivalently in the 1D case, Riemannian) manifold—that is, a line equipped with a metric.

[16]We could also make CIRCLE and LINE world-mates, by connecting them (say) by a space-like line. Schaffer's and Arntzenius and Hall's principles then apply directly.

[17]Horacek (2005, p.428) and Effingham (2020, p.152) each propose similar principles which also presuppose temporalism. Cusbert (2022, p. 617, 625–6) meanwhile formulates a version of stability weak enough to count even the trivialized chances on CIRCLE as "stable" relative to the chances on LINE. This doesn't strike me as a promising formulation of stability, going against the intuitions the principle is meant to capture. Moreover, Cusbert's principle doesn't get us out of the trivialization trap.

an event conditional on the state of an interval in LINE equals the chance of a duplicate event, conditional on the state of a duplicate interval in CIRCLE, provided the temporal distances between the event and the interval are the same in both worlds.

More carefully, where t and s are times, let the *forward distance from t to s* be the smallest duration from t to s, i.e., the smallest amount of time yo have to persist to get from t to s. In CIRCLE, there's a forward distance from any time to any other time. In LINE, a forward distance from t to s exists iff t occurs earlier than s. Where I and J are *intervals*, let the *forward distance from I to J* be the forward distance from the starting point of I's closure to the starting point of J's closure.[18] For example, in CIRCLE (fig. 1), the forward distance from τ_0 to τ_2 is 2 days, and the forward distance from τ_2 to τ_0 is 100 billion years minus 2 days. Moreover, say that two pairs of intervals, (I, J) and (I^*, J^*), are *temporally congruous* iff I and I^* have equal duration, J and J^* have equal duration, and the forward distance from I to J equals the forward distance from I^* to J^*.

Since CIRCLE and LINE have the same fundamental dynamical laws, they share the same urchance function. Denote by $\text{urch}_C(\cdot|\cdot)$ and $\text{urch}_L(\cdot|\cdot)$ the results of conditioning that function on a complete description of CIRCLE's and LINE's spacetime geometry, respectively.[19]

> ***Hypothesis.*** **Chance Invariance:** Let (I, J) and (I^*, J^*) be temporally congruous pairs of intervals in CIRCLE and LINE, respectively. Then
>
> $$\text{urch}_C(P(J)|Q(I)) = \text{urch}_L(P(J^*)|Q(I^*)),$$
>
> for any qualitative intrinsic properties P and Q.

Chance Invariance captures the thought that locally duplicate situations—$Q(I)$ and $Q(I^*)$—generate the same chances for locally duplicate outcomes—$P(J)$ and $P(J^*)$—provided the relevant temporal distances are the same.

To illustrate the principle, let RED(τ) be the proposition that τ is a 24-hour interval and that a single particle exists throughout τ, is red throughout τ, and has a clock reading of 0 at the start of τ. Chance Invariance then requires that $\text{urch}_C(\text{RED}(\tau_1)|\text{RED}(\tau_0)) = 0.8$. For consider any two successive days d_0 and d_1 in LINE. The pairs (d_0, d_1) and (τ_0, τ_1) are temporally congruous. Moreover, from the dynamics in LINE, we have

[18] If one of the two closures doesn't have a starting point or if there is no forward distance from I's starting point to J's starting point, the forward distance from I to J is ill-defined.

[19] That is, where $\mathcal{C}$ is a complete description of CIRCLE's geometry and $\mathcal{L}$ is a complete description of LINE's geometry, $\text{urch}_C(\cdot|\cdot) :=\text{urch}(\cdot|\cdot \wedge \mathcal{C})$ and $\text{urch}_L(\cdot|\cdot) :=\text{urch}(\cdot|\cdot \wedge \mathcal{L})$. We'll always understand complete geometric descriptions to include a "that's all" clause—i.e., they say that they describe all geometrical relationships between spacetime regions.

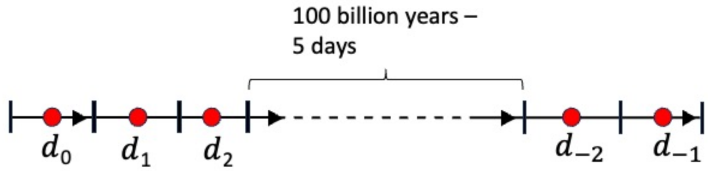


Figure 2: A sketch of LINE.

$\text{urch}_L(\text{RED}(d_1)|\text{RED}(d_0)) = 0.8$. So, by Chance Invariance,

$$\text{urch}_C(\text{RED}(\tau_1)|\text{RED}(\tau_0)) = \text{urch}_L(\text{RED}(d_1)|\text{RED}(d_0)) = 0.8.$$

Alas, Chance Invariance is inconsistent: it defines conditional chances in CIRCLE twice over, with conflicting results. Essentially, what goes wrong is that, except for antipodes, any two days in CIRCLE bear two distinct forward distances to each other: depending on which day comes first, the forward distance is either the short way or the long way around CIRCLE. For example, the pair (τ_0, τ_1) has a forward distance of one day, whereas the pair (τ_1, τ_0) has a forward distance of 100 billion years minus one day. As a result, the two pairs are temporally congruous to very different pairs in LINE. This makes Chance Invariance yield conflicting results.

For illustration, let's calculate $\text{urch}_C(\text{GREEN}(\tau_1)|\text{RED}(\tau_2) \wedge \text{RED}(\tau_0))$—the chance, in CIRCLE, of the particle's being green at τ_1, conditional on its being red on the two adjacent days. First consider the pairs (τ_0, τ_1) and (τ_1, τ_2) in CIRCLE and the temporally congruous days (d_0, d_1) and (d_1, d_2) in LINE—see figure 2. Chance Invariance yields the following:[20]

$$\text{urch}_C(\text{GREEN}(\tau_1)|\text{RED}(\tau_2) \wedge \text{RED}(\tau_0)) = \text{urch}_L(\text{GREEN}(d_1)|\text{RED}(d_2) \wedge \text{RED}(d_0)). \quad (5)$$

[20] *Proof:* By the multiplicative axiom,

$$\begin{aligned} \text{urch}_C(\text{GREEN}(\tau_1)|\text{RED}(\tau_2)\wedge\text{RED}(\tau_0)) \cdot \text{urch}_C(\text{RED}(\tau_2)|\text{RED}(\tau_0)) = \\ = \text{urch}_C(\text{GREEN}(\tau_1) \wedge \text{RED}(\tau_2)|\text{RED}(\tau_0)). \end{aligned} \quad (1)$$

From the congruity of (τ_0, τ_1) and (d_0, d_1) and of (τ_1, τ_2) and (d_1, d_2), we also have that (τ_0, τ_2) and (d_0, d_2) are congruous and that $(\tau_0, \tau_1 \cup \tau_2)$ and $(d_0, d_1 \cup d_2)$ are congruous. So, by Chance Invariance,

$$\text{urch}_C(\text{RED}(\tau_2)|\text{RED}(\tau_0)) = \text{urch}_L(\text{RED}(d_2)|\text{RED}(d_0)), \quad (2)$$

and

$$\text{urch}_C(\text{GREEN}(\tau_1) \wedge \text{RED}(\tau_2)|\text{RED}(\tau_0)) = \text{urch}_L(\text{GREEN}(d_1) \wedge \text{RED}(d_2)|\text{RED}(d_0)). \quad (3)$$

From eqs. 1–3,

$$\begin{aligned} \text{urch}_C(\text{GREEN}(\tau_1)|\text{RED}(\tau_2)\wedge\text{RED}(\tau_0)) \cdot \text{urch}_L(\text{RED}(d_2)|\text{RED}(d_0)) = \\ = \text{urch}_L(\text{GREEN}(d_1) \wedge \text{RED}(d_2)|\text{RED}(d_0)). \end{aligned} \quad (4)$$

But, from the given dynamics in LINE, $\text{urch}_L(\text{RED}(d_2)|\text{RED}(d_0)) = 0.8^2 > 0$, and so, from eq. 4, But the pair (τ_0, τ_1) is also temporally congruous with (d_{-2}, d_{-1}) while the "inverse" pair (τ_2, τ_1) is temporally congruous with (d_0, d_{-1}). This yields:[21]

$$\text{urch}_C(\text{GREEN}(\tau_1)|\text{RED}(\tau_2) \wedge \text{RED}(\tau_0)) = \text{urch}_L(\text{GREEN}(d_{-1})|\text{RED}(d_0) \wedge \text{RED}(d_{-2})). \tag{10}$$

But eqs. 5 and 10 conflict. Because the dynamics in LINE disallows RED-GREEN-RED transitions,

$$\text{urch}_L(\text{GREEN}(d_1)|\text{RED}(d_2) \wedge \text{RED}(d_0)) = 0.$$

$$\begin{aligned}\text{urch}_C(\text{GREEN}(\tau_1)|\text{RED}(\tau_2) \wedge \text{RED}(\tau_0)) &= \frac{\text{urch}_L(\text{GREEN}(d_1) \wedge \text{RED}(d_2)|\text{RED}(d_0))}{\text{urch}_L(\text{RED}(d_2)|\text{RED}(d_0))} \\ &= \text{urch}_L(\text{GREEN}(d_1)|\text{RED}(d_2) \wedge \text{RED}(d_0)),\end{aligned}$$

where the second line follows by the multiplicative axiom. ■

[21] *Proof:* By the multiplicative axiom,

$$\begin{aligned}\text{urch}_C(\text{GREEN}(\tau_1)|\text{RED}(\tau_2) \wedge \text{RED}(\tau_0)) \cdot \text{urch}_C(\text{RED}(\tau_0)|\text{RED}(\tau_2)) &= \\ = \text{urch}_C(\text{GREEN}(\tau_1) \wedge \text{RED}(\tau_0)|\text{RED}(\tau_2)).\end{aligned} \tag{6}$$

From the congruities of (τ_2, τ_1) and (d_0, d_{-1}) and of (τ_0, τ_1) and (d_{-2}, d_{-1}), we obtain that (τ_2, τ_0) is congruous with (d_0, d_{-2}) and $(\tau_2, \tau_0 \cup \tau_1)$ is congruous with $(d_0, d_{-2} \cup d_{-1})$. So, by Chance Invariance,

$$\text{urch}_C(\text{RED}(\tau_0)|\text{RED}(\tau_2)) = \text{urch}_L(\text{RED}(d_{-2})|\text{RED}(d_0)), \tag{7}$$

and

$$\text{urch}_C(\text{GREEN}(\tau_1) \wedge \text{RED}(\tau_0)|\text{RED}(\tau_2)) = \text{urch}_L(\text{GREEN}(d_{-1}) \wedge \text{RED}(d_{-2})|\text{RED}(d_0)). \tag{8}$$

Plugging eqs. 7 and 8 into eq. 6 yields

$$\begin{aligned}\text{urch}_C(\text{GREEN}(\tau_1)|\text{RED}(\tau_2) \wedge \text{RED}(\tau_0)) \cdot \text{urch}_L(\text{RED}(d_{-2})|\text{RED}(d_0)) &= \\ = \text{urch}_L(\text{GREEN}(d_{-1}) \wedge \text{RED}(d_{-2})|\text{RED}(d_0)).\end{aligned} \tag{9}$$

But $\text{urch}_L(\text{RED}(d_{-2})|\text{RED}(d_0)) \approx 1/3 > 0$, for the particle's current color provides essentially no evidence about its far-future color. (To derive this formally, divide the probability-weighted sum of color trajectories compatible with $\text{RED}(d_{-2}) \wedge \text{RED}(d_0)$ by the probability-weighted sum of color trajectories compatible with $\text{RED}(d_0)$, leading to a formula similar to the expression in fn. 44.) So, from eq. 9,

$$\begin{aligned}\text{urch}_C(\text{GREEN}(\tau_1)|\text{RED}(\tau_2) \wedge \text{RED}(\tau_0)) &= \frac{\text{urch}_L(\text{GREEN}(d_{-1}) \wedge \text{RED}(d_{-2})|\text{RED}(d_0))}{\text{urch}_L(\text{RED}(d_{-2})|\text{RED}(d_0))} \\ &= \text{urch}_L(\text{GREEN}(d_{-1})|\text{RED}(d_0) \wedge \text{RED}(d_{-2})),\end{aligned}$$

where the second line follows from the multiplicative axiom. ■

But the same dynamics also entails that[22]

$$\text{urch}_L(\text{GREEN}(d_{-1})|\text{RED}(d_0) \wedge \text{RED}(d_{-2})) = 0.2 > 0.$$

Contradiction. So Chance Invariance is inconsistent.

6 Two Vestiges of Invariance

Despite its inconsistency, Chance Invariance has intuitive appeal. Moreover, I think we can salvage its appealing parts in the form of weaker, and jointly consistent, principles.

First, its restriction to *loop-free* worlds remains consistent: two intrinsically duplicate situations in two nomically compatible *loop-free* worlds generate the same chance distributions. Let's call this weaker principle *Acyclic Chance Invariance*.

Below I'll spell out a precise version of this principle. To do this, a few concepts are needed, which will reappear throughout the rest of the paper (and thus shouldn't be skipped). Informally, a *curve* is any directed line (straight or curved) through spacetime.[23] A *causal curve* is any curve through spacetime which could be the trajectory of a material particle.[24,25] Causal curves can be either *future-directed* or *past-directed*. A *closed causal curve* (or "spacetime loop", as I called it earlier) is any causal curve which loops back in on itself, i.e. ends at its starting point. Informally, a region R's *causal past*, $J^-(R)$, is that part of spacetime from which a material particle can eventually reach R—we include in this the entirety of R itself. Meanwhile, R's *causal future*, $J^+(R)$, is that part of spacetime which a material particle can eventually reach *from* the region—again, including R itself.[26] R's

[22]*Proof:* Because d_{-2} is located in between d_2 and d_{-1}, $\text{RED}(d_{-2})$ "screens off" $\text{RED}(d_2)$ from $\text{GREEN}(d_{-1})$:

$$\text{urch}_L(\text{GREEN}(d_{-1})|\text{RED}(d_0) \wedge \text{RED}(d_{-2})) = \text{urch}_L(\text{GREEN}(d_{-1})|\text{RED}(d_{-2})). \quad (11)$$

(In section 7 I'll discuss "screening off" much more—the relevant principle, Parental Markov, is true in acyclic worlds and yields eq. 11.) From the dynamics for LINE we moreover have

$$\text{urch}_L(\text{GREEN}(d_{-1})|\text{RED}(d_{-2})) = 0.2. \blacksquare$$

[23]Formally, a *curve* is a function $c : I \to \mathcal{M}$ from an interval $I \subseteq \mathbb{R}$ into the spacetime $\mathcal{M}$. A *line* is the image of a curve. For bounded curves, we'll set $I = [0, 1]$ without loss of generality.

[24]Formally, we define a *causal curve* to be any differentiable curve with everywhere light-like (null) or time-like tangent vector.

[25]"Causal" thus has its technical meaning from physics, in terms of spacetime structure. I don't presuppose that this exactly—or even approximately—tracks the philosopher's various notions of causality. The fundamental dynamical laws speak the language of spacetime structure, not causation—so we're interested in the former, not the latter.

[26]More exactly, I define R's causal future [causal past], $J^+(R)$ [$J^-(R)$], as the union of R with all points p

proper causal future $K^+(R)$ is the difference between the causal future and R, i.e. $K^+(R) := J^+(R) \backslash R$; R*'s proper causal past* $K^-(R)$ is defined analogously, $K^-(R) := J^-(R) \backslash R$. Finally, throughout, where urch is a possible urchance function and $\mathcal{M}$ a possible spacetime, $\text{urch}_{\mathcal{M}}$ denotes the result of conditioning urch on a complete description of $\mathcal{M}$'s geometry.

Now for the precise version of Acyclic Chance Invariance (those who'd like to avoid formalism may skip this paragraph). We'll formulate the principle for pairs (R_1, R_2) of disjoint regions where R_2 is "strictly to the future" of R_1, i.e., $R_2 \subseteq K^+(R_1)$ and $R_1 \cap K^+(R_2) = \emptyset$, which I'll also write as $R_1 < R_2$. Intuitively, whenever $R_1 < R_2$, $J^+(R_1) \cap J^-(R_2)$ comprises the region "between" R_1 and R_2.

> ***Thesis.*** **Acyclic Chance Invariance:** Let $\mathcal{M}$ and $\mathcal{M}'$ be spacetimes with no closed causal curves. Let R_1 and R_2 be regions in $\mathcal{M}$ with $R_1 < R_2$. Finally, let R_1' and R_2' be regions in $\mathcal{M}'$ with $R_1' < R_2'$, such that there is an isometry $\Phi : J^+(R_1) \cap J^-(R_2) \to J^+(R_1') \cap J^-(R_2')$ with $\Phi(R_1) = R_1'$ and $\Phi(R_2) = R_2'$.[27] Then,
>
> $$\text{urch}_{\mathcal{M}}(P(R_2)|Q(R_1)) = \text{urch}_{\mathcal{M}'}(P(R_2')|Q(R_1')),$$
>
> for any qualitative intrinsic properties Q and P.

Acyclic Chance Invariance is our first vestige of Chance Invariance.

The second vestige of Chance Invariance concerns the *cyclic* chances. As we've learned in the previous section, they can't be strictly identical to the acyclic chances. Still, they shouldn't just be arbitrary either. Instead, they should be derivable from the acyclic chances in a principled way.

To sharpen this up, suppose you're given some possible fundamental dynamical laws, defining an urchance function urch. The laws' *acyclic chances* is simply the collection of all functions $\text{urch}_{\mathcal{K}}(\cdot|\cdot)$ such that $\mathcal{K}$ is a spacetime without closed causal curves. The laws' *cyclic chances* are defined in exactly the same way, except that $\mathcal{K}$ now ranges over all spacetimes *with* closed causal curves. Say that a theory of cyclic chances is *dynamically scrutable* iff the cyclic chances can be inferred from the acyclic chances in a principled way. A demand for dynamic scrutability is our second vestige of Chance Invariance. It and Acyclic Chance Invariance are, I say, what remains of Chance Invariance when stripped of its inconsistency.

Our work is now cut out for us: we must find a (1) *consistent* and (2) *dynamically scrutable* theory of cyclic chances that (3) *avoids trivialization*. Chance Invariance satisfies (2) and

---

such that there is a future-directed [past-directed] causal curve starting in R and ending in p.

[27] If Φ exists, you might call (R_1, R_2) and (R_1', R_2') *spatiotemporally congruous*, generalizing the concept of *temporally congruous* from Section 5.

(3), but not (1). A consistent but arbitrary way of assigning cyclic chances satisfies (1) and (3), but not (2). And Temporalism satisfies (1) and (2) but not (3)—it makes chances conditional on CIRCLE's geometry dynamically scrutable, but trivially so: the only allowed background proposition is the entire history. We are looking for a theory which checks all three boxes.

7 Spacetime Markov: What's Inside Doesn't Matter

As we saw in the introduction, when it comes to assessing the probability of a proposition, some information can override, or *screen off*, other information. When it comes to chance, what screens off what is partially determined by spacetime structure. For example, once we know the current state of a radioactive atom and its local environment, any additional information about the atom's causal past, e.g. how long it has been in its excited state, has no additional impact on the chance of its decaying within the next 10 seconds. Similarly, provided the dynamics are local[28] what happens at space-like separation also doesn't matter. In short, the atom's current state screens off everything outside of its own causal future.

Generalizing this yields the following proposal: a region's immediate proper causal past screens it off from everything outside of the region's causal future. This is the spacetime-theoretic analogue of the well-known "Causal Markov Condition". In its generic form, the Condition states that an event's immediate causes screen it off from any of its non-effects (cf. Hitchcock and Rédei 2021). The Causal Markov Condition and its predecessor, Reichenbach's "Common Cause Principle", occupy important roles in debates about the metaphysics of time (Reichenbach, 1956) and the nature of causation (Hitchcock and Rédei, 2021). Now, how might its spacetime-theoretic analogue bear on what the chances on loops are? The idea is this: for some spacetime regions we can screen off *whether or not they're part of a world with closed causal curves*. As we shall see, this allows us to derive the cyclic chances from the acyclic chances.

[28] Throughout this essay, I restrict myself to local dynamics. However, at least some non-local theories can be handled by my framework with mild modifications. For example, in the case of non-relativistic GRW—a stochastic theory of non-relativistic quantum mechanics, due to Ghirardi, Rimini, and Weber (1986)—one can maintain Parental Markov (see below) relative to Galilean spacetime structure (where any differentiable curve intersecting each time slice at most once counts as "causal"), provided we restrict urchance to the algebra generated by the maximal intrinsic states of time slices. Similarly, a theory involving superluminal particles might posit a non-Minkowskian spacetime structure with an alternative notion of "causal", relative to which Parental Markov (see below) could still hold.

7.1 Parental Markov

Let's start by stating the spacetime analogue of the Causal Markov Condition more precisely.

Informally, we say that a spacetime region S "screens off" region R from T iff, conditional on the complete state of S, urchance judges information about T as irrelevant to the state of R. More formally:

> ***Def.*** **Screening Off:** Where urch is the world's urchance function, $\mathcal{M}$ its spacetime, and R, S, T any spacetime regions: S *screens off* R *from* T iff
>
> $$\text{urch}_{\mathcal{M}}(Q_1(R)|Q_2(S) \wedge Q_3(T)) = \text{urch}_{\mathcal{M}}(Q_1(R)|Q_2(S)),$$
>
> for any maximal intrinsic property Q_2 and any intrinsic properties Q_1 and Q_3 such that $Q_2(S) \wedge Q_3(T)$ is possible according to $\text{urch}_{\mathcal{M}}$.[29]

Next we'll define what it means to contain a region's "immediate proper causal past". For any region A, denote A's complement by $A^{\perp}$. Let a *thick parent* of R be any (possibly empty) region P, disjoint from R, which is such that every future-directed causal curve which starts in $(P \cup R)^{\perp}$ and ends in R has a non-trivial subcurve[30] in P before ever intersecting R.[31] Intuitively, this says that approaching any region from the past, you'll have to spend at least some time in its thick parents. Any thick parent contains a region's "immediate proper causal past".

Why "thick" parent? One reason is that dynamical laws often require velocities or other time derivatives as inputs for generating chance distributions.[32] But many, including myself, are inclined to endorse reductive views about time derivatives, such as the "at-at"

[29] On "maximal" and "possible according to $\text{urch}_{\mathcal{M}}$": an intrinsic property of a spacetime region is *maximal* iff it includes a "that's all" clause. That is, where Q is a maximal intrinsic property, $Q(R)$ says that every particular matter of fact intrinsic to R is entailed by it. "Possible": in our Popperian formalism, where u is a primitively conditional probability function, a proposition A is *impossible according to* u iff $u(\neg A|A) = 1$. Where $\mathbf{u}$ is the set of primitively conditional probability functions induced by the fundamental dynamical laws, call any function from pairs of propositions into functions on $\mathbf{u}$ an *urchance candidate*. Then, where $\mathfrak{u}$ is an urchance candidate, say that A is *impossible according to* $\mathfrak{u}$ iff $\mathfrak{u}(\neg A|A) = 1$ (where "=" is point-wise). This explains the proviso that $Q_2(S) \wedge Q_3(T)$ be *possible* according to $\text{urch}_{\mathcal{M}}$. For example, where $S' \subseteq S$, we can generically choose a possible intrinsic property $Q_{2'}$ of S' such that $Q_{2'}(S') \wedge Q_2(S)$ is impossible according to $\text{urch}_{\mathcal{M}}$. This would have the consequence that S generically wouldn't even screen off R from any of S's own subsets—that's not a promising notion of "screening off".

[30] A *subcurve* of $c : I \to \mathcal{M}$ is any restriction of c to a subinterval of I. A continuous curve is *non-trivial* iff its image consists of at least two points (which, in a continuous spacetime, is equivalent to its consisting of continuum-many points).

[31] That is, where $c : I \to \mathcal{M}$, there are three disjoint subintervals, $I_1, I_2, I_3 \subseteq I$ with $I_1 < I_2 < I_3$, such that $c[I_1] \subseteq (R \cup P)^{\perp}$, $c[I_2] \subseteq P$ is non-trivial, and $c[I_3]$ intersects R.

[32] Specifically, this is the case if the dynamical law involves second- or higher-order differential equations.

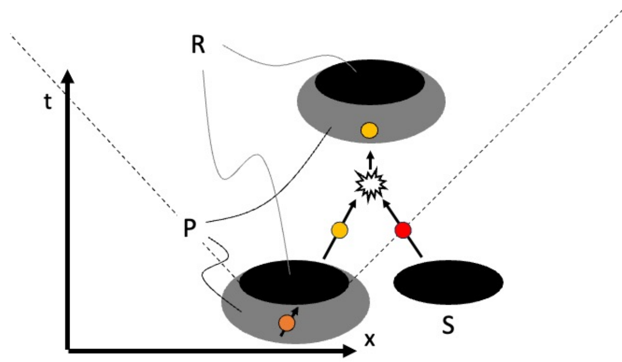


Figure 3: A counterexample to Unrestricted Parental Markov.

theory of velocities (Russell, 1903, Ch. 54). In this case, the maximal intrinsic state of a "thin" region doesn't specify particle velocities in that regions. A parent's *thickness* makes it so that at least one-sided derivatives are generally reductively definable.

In keeping with standard physics terminology, say that a spacetime region A is a *cause* of spacetime region B iff there is a future-directed causal curve starting in A and ending in B; *effect* is the converse of *cause*. One possible idea for the spacetime analogue of the Causal Markov Condition is this:

> ***Hypothesis.*** **Unrestricted Parental Markov:** Any thick parent of a spacetime region screens it off from any region not caused by it.

But Unrestricted Parental Markov is generically false. The problem is that, even in worlds with no closed causal curves, some regions generically cause their own thick parents: consider the disconnected region R (a union of two black ovals) in fig. 3. Conditioning on its thick parent P (the union of the two grey regions, disjoint from R) doesn't generally screen off R from S, despite S's being disjoint from R's causal future.[33]

This suggests a weaker principle. Let a *pure* thick parent of R be any thick parent of R not caused by R.

> ***Thesis.*** **Parental Markov:** Any *pure* thick parent of a spacetime region screens it off from any region not caused by it.

[33]For a concrete example, consider the world PLANE, inhabited by red and yellow particles, with the following dynamics: whenever two particles of the *same* color collide, they fuse into a red particle; and whenever two particles of *different* colors collide, they fuse into a yellow particle. Additionally, there are orange particles, which cannot collide with anything. The orange particles have a short lifespan, at whose end they decay, with equal chance, either into a yellow or a red particle. P's lower part contains such an orange particle (and nothing else), and P's upper part a yellow particle (and nothing else). Suppose that the spacetime distances are such that the orange particle is guaranteed to decay inside R. P's state together with S's containing a *red* particle entails that R's lower part contains a yellow particle. Yet P's state together with S's containing a *yellow* particle entails that R's lower part contains a red particle. So P's state doesn't screen off R from S.

This principle is plausibly true in any spacetime without closed causal curves and with a local dynamics.

Alas, we are interested in spacetimes *with* closed causal curves. Here Parental Markov is less useful: any region which intersects some closed causal curve without containing it whole causes all of its own parents, and so lacks *any* pure thick parents. In a world like CIRCLE, Parental Markov is therefore entirely vacuous. Even worse, in some loop worlds Parental Markov is outright *false*—Appendix A.1 provides an example. We can't use a principle that's at best vacuous and at worst false. So: back to the drawing board. Are there other properties about "screening off" we can exploit?

7.2 Boundary Markov

Yes. A cognate principle is that a region is screened off by its *thick boundaries.* In Minkowski spacetime, Parental Markov already entails that this is true for a fairly general class of regions (as I prove in Appendix C). But in contrast to Parental Markov, the generalized principle—which will be called *Boundary Markov*—remains non-trivial and plausibly true for *all* regions, in worlds with loops and in worlds without, provided that the dynamics is local. (In fact, satisfying Boundary Markov is plausibly part of what it *is* for a dynamics to be local.)

Before stating the principle, let's build an intuition for it. Consider again LINE (cf. section 5). Suppose that you know the world's urchance function. You then observe the particle on two days, d_9 and d_{11}, learning BLUE(d_9) and RED(d_{11}), respectively. Suppose you'd now like to calculate the chances of events outside of d_{10}. For this, is it worth examining d_{10}, to establish if BLUE(d_{10}) or RED(d_{10})? No: as far as the chances of events *outside* of d_{10} are concerned, d_{10}'s state doesn't matter given d_9's and d_{11}'s states. Any information d_{10} may carry about events outside of it is *screened off* by their union.

The union $d_9 \cup d_{10} \cup d_{11}$ contains what I'll call a *thick neighborhood* of d_{10}, and $d_9 \cup d_{11}$ is a *thick boundary* of d_{10}. More generally, a *thick neighborhood* of a region R is any open superset N of R such that every continuous curve starting in $N^\perp$ and ending in R has a non-trivial subcurve in $N \backslash R$ before ever intersecting R.[34] Intuitively, a thick neighborhood is like a city plus its suburbs: coming from the outside, you have to spend some time in the 'burbs to get to the city. Meanwhile, a *thick boundary* of R is any region B disjoint from R such that $R \cup B$ contains a thick neighborhood of R. In our analogy, the suburbs themselves are a thick boundary of the city, as is any region that *contains* the suburbs but no part of

[34]Formally: where $c : [0,1] \to \mathcal{M}$, the interval $[0,1]$ can be partitioned into subintervals I_0, I_1, I_2 with $I_0 < I_1 < I_2$, such that $c[I_0] \subseteq N^\perp$, $c[I_1] \subseteq N \backslash R$ is non-trivial, and $c[I_2]$ intersects R.

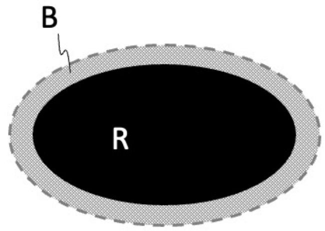


Figure 4: B is a thick boundary of R.

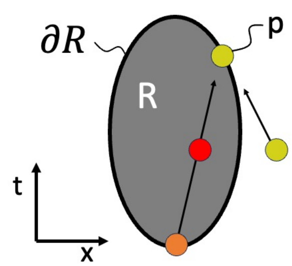
the city. See e.g. fig. 4, where B is a thick boundary of R.[35] As I prove in Appendix B, in any spacetime with the same topology as $\mathbb{R}^n$, N is a thick neighborhood of R iff N is an open superset of R's closure.

We are ready to state Boundary Markov:[36]

> ***Thesis.*** **Boundary Markov:** Any thick boundary of a spacetime region screens it off from any region disjoint from it.

[35]It's worth noting some edge cases: the entire spacetime has the empty set as a thick boundary ("the universe doesn't have suburbs"), and the empty set has every region as a thick boundary ("everything is a suburb of the empty set").

[36]Besides accommodating reductive accounts of derivatives, a boundary's thickness plays a second role. Even for non-reductivists, a *thin* boundary generally wouldn't screen off its inside from its outside. (This contrasts with parental thickness: to my knowledge, given instantaneous derivatives, thin parents do screen off a region from its non-effects.) To see this, consider again PLANE (cf. fn. 33), now with the following setup:



where ∂R is just the ordinary topological ("thin") boundary. Conditioning on the complete state of ∂R doesn't screen off R's inside from the outside. Let the distances be such that the orange particle on ∂R is guaranteed to decay inside R. Conditional on the state of ∂R, with its yellow particle at point p, the probability that R contains a red particle is then equal to the probability, conditional on the state of ∂R, that there is a yellow particle on course to collide with it at p. (Recall that the only way for a red particle to transform into a yellow one is to collide with another yellow particle.) Since ∂R contains little information about the universe's initial conditions, this probability is highly imprecise. (And if it was precise, it would be near zero—it would be more likely that the orange particle decayed into a yellow particle to begin with.) But conditional on the existence of a yellow particle on collision course with p, the chance rises to 1. So ∂R doesn't screen off R from the outside. Meanwhile, any *thick* boundary would automatically include either the information about the inside particle's color or the information about the outside particle's color (or both), and hence screen them off each other.

Appendix C proves that, in Minkowski spacetime, given plausible conglomerability and locality assumptions for urchance, Parental Markov already entails a certain restriction of Boundary Markov.

Every region has a thick boundary (indeed often infinitely many), even in worlds with closed causal curves. Moreover, there are plausibly no locality-respecting counterexample to Boundary Markov. (As I said, Boundary Markov may indeed be partially constitutive of locality.) Appendix A.2 explicitly demonstrates that in the cyclic case in which Parental Markov is false, Boundary Markov is still true. Henceforth I'll assume that, necessarily, if the dynamics is local, Boundary Markov is true.

7.3 Strong Boundary Markov

We are almost there. So far we've understood screening off a region as screening off its *matter content.* But plausibly not only matter content can be screened off, but also a region's *internal geometry.* This insight holds the key to our theory of chances on loops. For whenever a region intersects *all* spacetime loops, the existence of spacetime loops partially depends on the region's internal geometry. But then screening off the region's internal geometry also screens off whether there are any spacetime loops. This gives us the desired connection between cyclic and acyclic chances.

To first get a better feel for the strenghtened notion of *screening off*, consider LINE. Suppose you've studied the world's geometry everywhere outside of d_{10}, but haven't examined d_{10} at all; you're even unsure about aspects of d_{10}'s internal geometry—say its length, or whether it is connected, etc. When calculating the chance of events in d_{10}'s complement, does your ignorance matter? No, I say, because d_9 and d_{11} screen off d_{10}'s internal geometry. Where LINE* (or L^* for short) is a world just like LINE except that (say) d_{10} isn't connected, and where A is a proposition about the complement of d_{10}, we get the following identity:

$$\text{urch}_L(A|\text{BLUE}(d_9) \wedge \text{RED}(d_{11})) = \text{urch}_{L^*}(A|\text{BLUE}(d_9) \wedge \text{RED}(d_{11})).$$

(Where, as always, urch_{L^*} denotes the result of conditioning urch on a complete description of L^*'s geometry.)

The previous definition of "screening off" compares urchance functions on the same spacetime geometries. One way to define the stronger notion instead compares the original spacetime with the result of *deleting* the screened-off region. For any spacetime $\mathcal{M}$, let $\mathcal{M}\backslash X$ denote the result of deleting region X from $\mathcal{M}$.[37] Moreover, let $\text{urch}_{\mathcal{M}\backslash X}$ be the

[37] More precisely, we'll first define the deletion operation for manifolds. Let $M = (\Omega, \mathcal{A}, g)$ be a Lorentzian

result of conditioning urch on a complete geometrical description of $\mathcal{M}\backslash X$. (This description includes a "that's all" clause—cf. fn. 19—i.e. a proviso that $\mathcal{M}\backslash X$ is *all of spacetime*. Accordingly, $\text{urch}_{\mathcal{M}\backslash X}$ is not the result of conditioning urch on an *incomplete* description that's *true* at $\mathcal{M}$, but instead the result of conditioning urch on a *complete* description that's *false* at $\mathcal{M}$.)

> ***Def.* Strong Screening Off:** Where urch is the world's urchance function, $\mathcal{M}$ the spacetime, and R, S, T are any spacetime regions: *S strongly screens off R from T* iff
>
> $$\text{urch}_{\mathcal{M}}(Q_1(R)|Q_2(S) \wedge Q_3(T)) = \text{urch}_{\mathcal{M}\backslash T}(Q_1(R)|Q_2(S)),$$
>
> for any maximal intrinsic property Q_2 and any intrinsic properties Q_1 and Q_3 such that $Q_2(S) \wedge Q_3(T)$ is possible according to $\text{urch}_{\mathcal{M}}$.[38]

manifold with topological space Ω, atlas $\mathcal{A}$, and (pseudo-)metric field g. For $X \subseteq \Omega$, define $M\backslash X := (\Omega\backslash X, \mathcal{A}|_{\Omega\backslash X}, g|_{\Omega\backslash X})$, where $\mathcal{A}|_{\Omega\backslash X} := \{\phi|_{U\backslash X} | (\phi : U \to \mathbb{R}^n) \in \mathcal{A}\}$ is the set of all restrictions of coordinate charts to their domains minus X. (When X has non-differentiable boundary, $(\Omega\backslash X, \mathcal{A}|_{\Omega\backslash X})$ will generally be neither a manifold nor manifold-with-boundary. But no matter: $g|_{\Omega\backslash X}$ retains all the metric structure we need to make sense of the dynamics.) Let now $\mathcal{M}$ be a spacetime represented by M. Then we define $\mathcal{M}\backslash X$ to be the part of $\mathcal{M}$ represented (under the same representation) by $M\backslash X$.

[38]To see how Strong Boundary Markov entails that thick boundaries screen off a region's internal geometry, consider the operation of *adding* a region to a spacetime. I'll define this precisely at the end of the footnote; for now, simply note that if $\mathcal{M} + X$ is a result of adding region X to $\mathcal{M}$, the operation $\cdot\backslash X$ (fn. 37) reverses that addition: $(\mathcal{M} + X)\backslash X = \mathcal{M}$. Let now B be a thick boundary of R in $\mathcal{M}$, and let A be a proposition purely about $(R \cup B)^\perp$. Let $\mathcal{M}^* := (\mathcal{M}\backslash R) + R^*$ be the result of adding some region R^* to $\mathcal{M}\backslash R$ such that B is also a thick boundary of R^* in $\mathcal{M}^*$. (Think of $\mathcal{M}^*$ as resulting from "replacing" R by R^* in $\mathcal{M}$ in some way.) By the above, we have $\mathcal{M}^*\backslash R^* = ((\mathcal{M}\backslash R) + R^*)\backslash R^* = \mathcal{M}\backslash R$. Moreover, by Strong Boundary Markov,

$$\text{urch}_{\mathcal{M}^*}(A|Q_1(B) \wedge Q_2(R^*)) = \text{urch}_{\mathcal{M}^*\backslash R^*}(A|Q_1(B)), \text{ and}$$
$$\text{urch}_{\mathcal{M}}(A|Q_1(B) \wedge Q_3(R)) = \text{urch}_{\mathcal{M}\backslash R}(A|Q_1(B))$$

(where Q_1 is a maximal intrinsic property and Q_2 and Q_3 are intrinsic properties such that $Q_1(B) \wedge Q_2(R)$ is possible according to $\text{urch}_{\mathcal{M}}$ and $Q_1(B) \wedge Q_3(R^*)$ is possible according to $\text{urch}_{\mathcal{M}^*}$). So,

$$\text{urch}_{\mathcal{M}}(A|Q_1(B) \wedge Q_2(R)) = \text{urch}_{\mathcal{M}^*}(A|Q_1(B) \wedge Q_3(R^*)).$$

In other words, B screens off A from anything in B, including B's internal geometry.

Now, to define *adding*, first define it for manifold-like structures. Specifically, let $M = (\Omega, \mathcal{A}, g)$ be the result of deleting a (possibly empty) region from a Lorentzian manifold. Then, for any X disjoint from Ω, $M + X = (\Omega \cup X, \mathcal{A}^{\Omega\cup X}, g^{\Omega\cup X})$ is *a result of adding X to M* iff

(i) $\mathcal{A}^{\Omega\cup X}$ is a set of charts whose domains form an open (relative to some chosen topology on $\Omega \cup X$, inducing Ω's topology) cover of $\Omega \cup X$,
(ii) $\mathcal{A}^{\Omega\cup X}|_\Omega = \mathcal{A}$ (see fn. 37 for the definition of $|.$ on sets of charts),
(iii) all charts in $\mathcal{A}^{\Omega\cup X}$ satisfy the usual smoothness condition (i.e., all transition maps are smooth),
(iv) $g^{\Omega\cup X}$ is an extension of g to $\Omega \cup X$ smooth relative to $\mathcal{A}^{\Omega\cup X}$.

The non-uniqueness of $M + X$ is reflected in (i), specifically in the choice of a topology on $\Omega \cup X$ and the

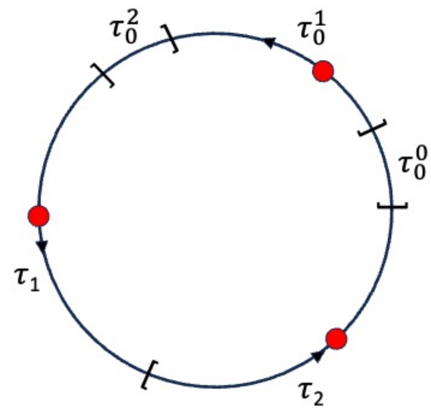


Figure 5: A sketch of SMALL CIRCLE. Each τ_i is a half-open interval (closed toward the past, open toward the future).

Replacing "screens off" in Boundary Markov by "strongly screens off" yields

> ***Thesis.* Strong Boundary Markov:** Any thick boundary of a spacetime region strongly screens it off from any region disjoint from it.

I think that, accepting Boundary Markov, you should also accept Strong Boundary Markov. Intrinsic geometrical information isn't privileged over information about matter content: both are screened off by thick boundaries.

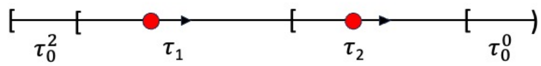
8 Cutting Loops

Our theory of chances on loops is now complete: it consists of Strong Boundary Markov and Acyclic Chance Invariance. Let us put them to work.

To start simply, consider SMALL CIRCLE (or SC for short), which is just like CIRCLE except only three days long. We'd like to calculate $\text{urch}_{SC}(\text{RED}(\tau_1)|\text{RED}(\tau_0))$—the chance of the particle's being red at τ_0 conditional on its being red the day prior. To do so, partition τ_0 into any three non-trivial intervals $\{\tau_0^0, \tau_0^1, \tau_0^2\}$ with $\tau_0^0 < \tau_0^1 < \tau_0^2$—see fig. 5. Note that $\tau_0^0 \cup \tau_0^2$ is a thick boundary of τ_0^1, and hence, given Strong Boundary Markov, strongly screens it off from τ_1.

charts (if any) whose domains overlap both Ω and X—intuitively, these two choices determine *how* X is added to M. Extending the deletion operation (fn. 37) to $M + X$ in the obvious way, we have, by conditions (ii) and (iv), $(M + X)\backslash X = ((\Omega \cup X)\backslash X, (\mathcal{A}^{\Omega\cup X})|_{(\Omega\cup X)\backslash X}, g^{\Omega\cup X}|_{(\Omega\cup X)\backslash X}) = (\Omega, \mathcal{A}, g) = M$, as desired. Finally, if $\mathcal{M}$ is a *spacetime* represented by M, *a result of adding X to $\mathcal{M}$* is any mereological fusion of $\mathcal{M}$ with a spacetime region such that the result is represented (under the same representation relation) by some $M + X$.

Now, whether the spacetime is a cycle partially depends on τ_0^1's internal geometry. In particular, if τ_0^1 has a "hole"—i.e., if it's empty or its ends otherwise don't connect—the spacetime isn't a cycle. For concreteness, consider the possibility where τ_0^1 is empty. The resulting spacetime is SMALL LINE (or SL for short), as follows:[39]



Strong Boundary Markov now lets us express the chances on SMALL CIRCLE in terms of the chances on SMALL LINE. But given Acyclic Chance Invariance, we already know the chances on SMALL LINE: they are based on the same transition chances as those on LINE (cf. Section 5). And now we are done: we've derived the cyclic chances from the acyclic chances.

Let's do this slowly. Let [$\text{RED}_i/\text{RED}_m/\text{RED}_f$] be the property of containing a red particle whose clock initially reads [$0:00/t_0^1/t_0^2$] and grows in proportion to the time passed—where t_0^1 is τ_0^1 's starting time and t_0^2 is τ_0^2's starting time. According to Acyclic Chance Invariance, the transition chances in SMALL LINE directly follow from those in LINE:[40]

$$\begin{aligned}
\text{urch}_{SL}(\text{RED}(\tau_1)|\text{RED}_f(\tau_0^2)) &= 0.8, \\
\text{urch}_{SL}(\text{GREEN}(\tau_1)|\text{RED}_f(\tau_0^2)) &= 0.2, \\
&\dots \\
\text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}(\tau_2)) &= 0.8, \\
\text{urch}_{SL}(\text{GREEN}_i(\tau_0^0)|\text{RED}(\tau_2)) &= 0.2, \\
&\text{etc.}
\end{aligned} \tag{12}$$

Strong Boundary Markov relates the urchance at SMALL CIRCLE to the urchance at SMALL LINE, as follows:

$$\begin{aligned}
\text{urch}_{SC}(\text{RED}(\tau_1)|\text{RED}(\tau_0)) &= \text{urch}_{SC}(\text{RED}(\tau_1)|\text{RED}_i(\tau_0^0) \wedge \text{RED}_f(\tau_0^2) \wedge \text{RED}_m(\tau_0^1)) \\
&= \text{urch}_{SL}(\text{RED}(\tau_1)|\text{RED}_i(\tau_0^0) \wedge \text{RED}_f(\tau_0^2)).
\end{aligned} \tag{13}$$

where the first line follows because, given a complete description of SMALL CIRCLE's

[39]Mathematically, this is a *manifold with boundary.*

[40]Using the official definition of Acyclic Chance Invariance, the relevant isometries are obvious: for example, for the first two lines in eqs. 12, any isometry from $J^+(\tau_0^2) \cap J^-(\tau_1) = \tau_0^2 \cup \tau_1$ into a subset of LINE will do; *mutatis mutandis* for the other lines.

geometry, $\text{RED}(\tau_0)$ is logically equivalent to $\text{RED}_i(\tau_0^0) \wedge \text{RED}_f(\tau_0^2) \wedge \text{RED}_m(\tau_0^1)$, and the second line follows by Strong Boundary Markov.

Eq. 13 gives the desired connection between cyclic and acyclic chances. We can easily derive its right-hand side, $\text{urch}_{SL}(A|\text{RED}_i(\tau_0^0) \wedge \text{RED}_f(\tau_0^2))$, from eqs. 12; the result is $64/65 \approx 0.98$.[41] So, by eq. 13,

[41] *Proof*: Using the multiplicative axiom twice,

$$\begin{aligned}&\text{urch}_{SL}(\text{RED}(\tau_1)|\text{RED}_i(\tau_0^0) \wedge \text{RED}_f(\tau_0^2)) \cdot \text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}_f(\tau_0^2)) = \\ &\quad = \text{urch}_{SL}(\text{RED}(\tau_1) \wedge \text{RED}_i(\tau_0^0)|\text{RED}_f(\tau_0^2)) \\ &\quad = \text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}(\tau_1) \wedge \text{RED}_f(\tau_0^2)) \cdot \text{urch}_{SL}(\text{RED}(\tau_1)|\text{RED}_f(\tau_0^2)).\end{aligned}$$

Provided that $\text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}_f(\tau_0^2)) > 0$—which we'll show below—we can rewrite this:

$$\begin{aligned}&\text{urch}_{SL}(\text{RED}(\tau_1)|\text{RED}_i(\tau_0^0) \wedge \text{RED}_f(\tau_0^2)) = \\ &\quad = \frac{\text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}(\tau_1) \wedge \text{RED}_f(\tau_0^2)) \cdot \text{urch}_{SL}(\text{RED}(\tau_1)|\text{RED}_f(\tau_0^2))}{\text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}_f(\tau_0^2))}.\end{aligned} \tag{14}$$

Let's evaluate each part of the quotient separately. Note that, since SMALL LINE is acyclic, Parental Markov is in good standing there. The first factor in the numerator:

$$\begin{aligned}&\text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}(\tau_1) \wedge \text{RED}_f(\tau_0^2)) = \\ &\quad = \text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}(\tau_1)) = \\ &\quad = \text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}(\tau_2) \wedge \text{RED}(\tau_1)) \cdot \text{urch}_{SL}(\text{RED}(\tau_2)|\text{RED}(\tau_1)) = \\ &\quad = \text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}(\tau_2)) \cdot \text{urch}_{SL}(\text{RED}(\tau_2)|\text{RED}(\tau_1)) = \\ &\quad = 0.8^2,\end{aligned}$$

where the first and third equalities follow from Parental Markov for SMALL LINE, the second equality follows from the probability laws plus the fact that only $\text{RED}(\tau_2)$ is nomically compatible with $\text{RED}(\tau_1) \wedge \text{RED}_i(\tau_0^0)$, and the final equality follows from the transition chances (eqs. 12) for SMALL LINE. The numerator's second factor follows immediately from the transition chances:

$$\text{urch}_{SL}(\text{RED}(\tau_1)|\text{RED}_f(\tau_0^2)) = 0.8.$$

Finally, to calculate the denominator, observe that

$$\begin{aligned}&\text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\text{RED}_f(\tau_0^2)) = \\ &\quad = \sum_{\pi \in \{\text{RED,GREEN}\}} \text{urch}_{SL}(\text{RED}_i(\tau_0^0)|\pi(\tau_1) \wedge \text{RED}_f(\tau_0^2)) \cdot \text{urch}_{SL}(\pi(\tau_1)|\text{RED}_f(\tau_0^2)),\end{aligned}$$

where the second line follows because LINE's dynamics disallow immediate RED-to-BLUE transitions. For $\pi = \text{RED}$, the right-hand side's summand is just the numerator, whose value we've just calculated: 0.8^3. (As promised, this also proves that the denominator is positive.) For $\pi = \text{GREEN}$, we perform exactly analogous calculations. Noting that $\text{BLUE}(\tau_2)$ is the only option for τ_2 compatible with $\text{GREEN}(\tau_1)$ and $\text{RED}_i(\tau_0^0)$, the right-hand side's summand then comes out to $0.2^2 \cdot 0.2 = 0.2^3$. Plugging everything into eq. 14,

$$\text{urch}_{SL}(\text{RED}(\tau_1)|\text{RED}_i(\tau_0^0) \wedge \text{RED}_f(\tau_0^2)) = \frac{0.8^3}{0.8^3 + 0.2^3} = \frac{64}{65} \approx 0.98. \ \blacksquare$$

$$\text{urch}_{SC}(\text{RED}(\tau_1)|\text{RED}(\tau_0)) = \frac{64}{65} \approx 0.98. \tag{15}$$

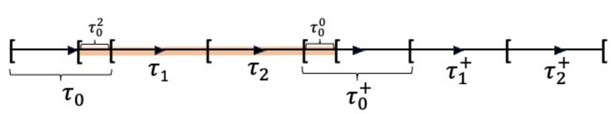
This is the cyclic transition chance we were looking for. *Contra* Chance Invariance, the chance for the particle to remain red is *higher* in SMALL CIRCLE than the usual 0.8.

We've derived this result from entirely general principles about chance. It's also worth noting that it makes sense pretheoretically. *Contra* Chance Invariance, we should really have expected that

$$\text{urch}_{SC}(\text{RED}(\tau_1)|\text{RED}(\tau_0)) > 0.8. \tag{16}$$

For, intuitively, $\text{RED}(\tau_0)$ *doubly* supports $\text{RED}(\tau_1)$: the latter is not only a likely *effect* of $\text{RED}(\tau_0)$, but also a likely *cause*. Going from $\text{RED}(\tau_1)$ to $\text{RED}(\tau_0)$ involves transitioning from $\text{RED}(\tau_1)$ to $\text{RED}(\tau_2)$ to $\text{RED}(\tau_0)$—two "likely" transitions. By contrast, going from $\text{GREEN}(\tau_1)$ to $\text{RED}(\tau_0)$ involves two "unlikely" transitions: $\text{GREEN}(\tau_1)$ to $\text{BLUE}(\tau_2)$ to $\text{RED}(\tau_0)$. So, pretheoretically, $\text{RED}(\tau_0)$ should favor $\text{RED}(\tau_1)$ both because it preferentially causes it *and* because it's preferentially caused *by* it. Eq. 13 captures this "double support" intuition. It says that the result of conditioning the urchance on today's particle being red equals (as far as events outside of τ_0 are concerned) the result of conditioning the acyclic urchance function on today's particle *and* the particle three days from now being red—thus *doubly supporting* tomorrow's redness.

We can also understand eq. 13 in terms of "unraveling" the cyclic spacetime. Consider SL^+, an extension of SMALL LINE obtained by adding[42] a copy of $\tau_0^0 \cup \tau_0^1$ to the beginning of τ_0^2 and a copy of $\tau_0^1 \cup \tau_0^2 \cup \tau_1 \cup \tau_2$ to the end of τ_0^0 in the obvious ways. The result is a concatenation of two copies of SMALL LINE. Denoting parts of the later copy with + superscripts, we can sketch SL^+ as follows (with the original SL highlighted):



Now, SL^+ is a result of "unraveling" the SMALL CIRCLE twice—more formally, it is a *double cover* of SMALL CIRCLE. Since $\tau_0^2 \cup \tau_0^0$ is a thick boundary of $\tau_1 \cup \tau_2$ in SL^+, by Strong Boundary Markov, it follows that[43]

$$\text{urch}_{SL}(\text{RED}(\tau_1)|\text{RED}_f(\tau_0^2) \wedge \text{RED}_i(\tau_0^0)) = \text{urch}_{SL^+}(\text{RED}(\tau_1)|\text{RED}(\tau_0) \wedge \text{RED}(\tau_0^+)).$$

[42]See fn. 38 for a rigorous definition of "adding".

[43]To see this rigorously: let RED^n be the property of being a concatenation of n 24h intervals the first of which satisfies RED. Given complete geometric descriptions of SL or of SL^+ (each entailing that τ_2 is a 24h interval following τ_1), $\text{RED}(\tau_1)$ is equivalent to $\text{RED}^2(\tau_1 \cup \tau_2)$. Similarly, given a complete geometric

So, by eq. 13,

$$\mathrm{urch}_{SC}(\mathrm{RED}(\tau_1)|\mathrm{RED}(\tau_0)) = \mathrm{urch}_{SL^+}(\mathrm{RED}(\tau_1)|\mathrm{RED}(\tau_0) \wedge \mathrm{RED}(\tau_0^+)).$$

Thus, our account implies that the transition chances for $\mathrm{RED}(\tau_0)$ on CIRCLE can also be obtained by unraveling the spacetime twice, into a double cover, and then conditioning on the proposition that both copies of τ_0 in the cover are red.

Our theory of chances on loops ticks all three boxes: since they're fully derived from the acyclic chances, the conditional chances in SMALL CIRCLE are dynamically scrutable. As eq. 15 shows, the approach also avoids trivialization. And if (as I claim) locality entails Strong Boundary Markov, the conjunction of Strong Boundary Markov and Acyclic Chance Invariance is (*ipso facto*) consistent with any local dynamics whose chance prescriptions are invariant across acyclic worlds—which arguably includes any *plausible* local dynamics.

There is a hidden fourth benefit to our account. Sometimes it salvages Chance Invariance in the short term, as it were, *asymptotically*. In the original CIRCLE case, with a period of 100 billion years, the short-term cyclic transition chances are essentially indistinguishable from the acyclic transition chances (the difference between $\mathrm{urch}_C(\mathrm{GREEN}(\tau_1)|\mathrm{RED}(\tau_0))$ and 0.2 is smaller than $0.8^{10^{12}}$). More generally, unless the chance to switch color in the *acyclic* case is *extremely* close to 0 or 1, the short-term cyclic transition chances converge extremely quickly to the acyclic transition chances as the loop length increases. This makes sense intuitively: knowing the particle's color in the far future provides little evidence about the near-term colors.[44] More generally, in a cyclic spacetime, our approach asymptot-

---

description of SL^+, $\mathrm{RED}(\tau_0^+)$ is equivalent to $\mathrm{RED}^3(\tau_0^+ \cup \tau_1^+ \cup \tau_2^+)$. We thus have

$$\begin{aligned}\mathrm{urch}_{SL^+}(\mathrm{RED}(\tau_1)|\mathrm{RED}(\tau_0) \wedge \mathrm{RED}(\tau_0^+)) &= \mathrm{urch}_{SL^+}(\mathrm{RED}^2(\tau_1 \cup \tau_2)|\mathrm{RED}(\tau_0) \wedge \mathrm{RED}^3(\tau_0^+ \cup \tau_1^+ \cup \tau_2^+)) \\ &\overset{SBM}{=} \mathrm{urch}_{SL}(\mathrm{RED}^2(\tau_1 \cup \tau_2)|\mathrm{RED}_f(\tau_0^2) \wedge \mathrm{RED}_i(\tau_0^0)) \\ &= \mathrm{urch}_{SL}(\mathrm{RED}\tau_1|\mathrm{RED}_f(\tau_0^2) \wedge \mathrm{RED}_i(\tau_0^0)),\end{aligned}$$

where the second line follows by Strong Boundary Markov because RED^2 and RED^3 are (non-maximal) intrinsic properties of $\tau_1 \cup \tau_2$ and $\tau_0^+ \cup \tau_1^+ \cup \tau_2^+$, respectively, $\mathrm{RED}(\tau_0)$ is equivalent to a conjunction $\phi(\tau_0 \backslash \tau_0^2) \wedge \mathrm{RED}_f(\tau_0^2)$ with ϕ intrinsic to $\tau_0 \backslash \tau_0^2$, and $\mathrm{RED}^3(\tau_0^+ \cup \tau_1^+ \cup \tau_2^+)$ is equivalent to a conjunction $\psi((\tau_0^+ \cup \tau_1^+ \cup \tau_2^+) \backslash \tau_0^0) \wedge \mathrm{RED}_i(\tau_0^0)$ with ψ intrinsic to $\tau_0^+ \cup \tau_1^+ \cup \tau_2^+) \backslash \tau_0^0$. ■

[44] More formally, where $q \in [0, 1)$ is the acyclic chance for a particle to switch color during the next transition and l is the loop length, we obtain (if $\lfloor l/3 \rfloor = 0$, the numerator sum is set to 0)

$$\mathrm{urch}_L(\mathrm{GREEN}(\tau_1)|\mathrm{RED}(\tau_0) \wedge \mathrm{RED}(\tau_{-1})) = \frac{\sum_{n=1}^{\lfloor l/3 \rfloor} (1-q)^{l-3n} q^{3n} \binom{l-1}{3n-1}}{\sum_{n=0}^{\lfloor l/3 \rfloor} (1-q)^{l-3n} q^{3n} \binom{l}{3n}} \tag{17}$$

For increasing l, this converges to q extremely quickly.

ically ensures short-term Chance Invariance whenever, according to the *acyclic* dynamics, far-future states are increasingly probabilistically independent of near-future states.[45]

9 Saving Grandpa

Let's apply our theory to the stochastic "grandfather paradox" from the introduction. For concreteness, let the spacetime $\mathcal{M}$ be Minkowskian except for one topological quirk, a *wormhole*. We can represent the wormhole with two duplicate (three-dimensional) bounded space-like surfaces, w_1 and w_2: every future-directed causal curve intersecting w_1 immediately exits at w_2 (without intersecting w_2), from where it continues future-ward, and every future-directed causal curve intersecting w_2 immediately exits at w_1 (without intersecting w_1), from where it continues future-ward.[46] See figure 6, which includes an example trajectory through the wormhole.[47]

Figure 6 also indicates six other spacetime regions: P is the region where the poisoning occurs; A_1 (A_2) is the region where you (would) administer the first (second) antidote, with A_1 additionally tri-partitioned into A_1^-, A_1°, and A_1^+; H_1 (H_2) is the region where the first (second) antidote takes (would take) effect; G is the region where grandpa has children, one of whom bears you and your sibling, who then travel through the wormhole at the end of G. (The regions aren't drawn to scale.)

To keep things simple, consider only finitely many possible maximal intrinsic states

---

[45] To see this: suppose we're given a cyclic spacetime $\mathcal{C}$. Let τ be a region in $\mathcal{C}$ bounded by two time-slices. As before, we partition τ into three parts τ^0, τ^1, τ^2. (In higher-dimensional cyclic spacetimes, $\tau, \tau^0, \tau^1, \tau^2$ are all hypercuboids.) Let $\mathcal{L} := \mathcal{C} \backslash \tau^1$. The longer the return time in $\mathcal{C}$, the greater the forward distance from τ^2 to τ^0 in $\mathcal{L}$. Let τ_+ be a time in τ^2's near-term future. Given a long return time and approximate probabilistic independence of far-future from near-future states, we have

$$\text{urch}_{\mathcal{L}}(Q(\tau_+)|P(\tau^2) \wedge P(\tau^0)) \approx \text{urch}_{\mathcal{L}}(Q(\tau_+)|P(\tau^2))$$

for any possible maximal intrinsic property P of τ—where "$P(\tau^i)$" denotes the strongest proposition entirely about τ^2 entailed by $P(\tau)$—and any intrinsic property Q. Hence, by Strong Boundary Markov,

$$\begin{aligned} \text{urch}_{\mathcal{C}}(Q(\tau_+)|P(\tau)) &= \text{urch}_{\mathcal{C}}(Q(\tau_+)|P(\tau^2) \wedge P(\tau^1) \wedge P(\tau^0)) \\ &\overset{SBM}{=} \text{urch}_{\mathcal{L}}(Q(\tau_+)|P(\tau^2) \wedge P(\tau^0)) \\ &\approx \text{urch}_{\mathcal{L}}(Q(\tau_+)|P(\tau^2)). \end{aligned}$$

But $\text{urch}_{\mathcal{L}}(Q(\tau_+)|P(\tau^2))$ is just the ordinary acyclic transition chance from $P(\tau)$ to $Q(\tau_+)$. ■

[46] To keep derivatives everywhere well-defined, w_1 and w_2's (two-dimensional) boundaries are deleted.

[47] The choice of w_1 and w_2 is non-unique: any other pair of duplicate bounded space-like surfaces with the same boundaries as w_1 and w_2 generates the same wormhole. Arguably, "wormhole" is therefore most naturally identified with the union of w_1's and w_2's domains of dependence. But to keep things simple, I'll keep framing things in terms of w_1 and w_2 specifically.

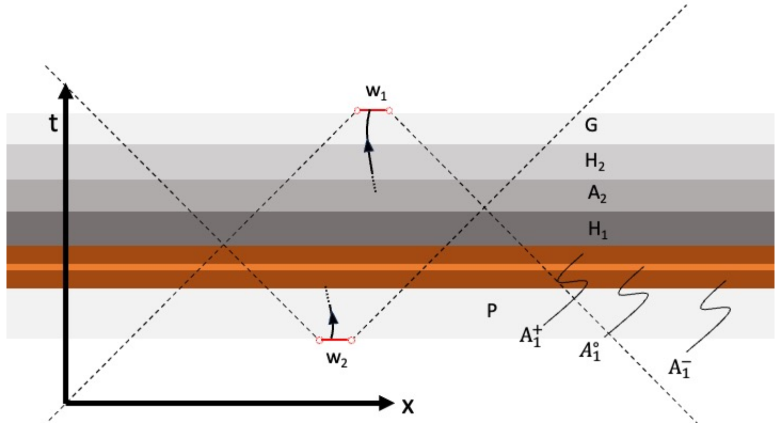


Figure 6: A sketch of $\mathcal{M}$, the spacetime in the grandfather case.

for each region.[48] Associating each state with a natural number, we can write property assignments to a region R conveniently as $R = n$. In the following all properties are to be read as intrinsic to the respective region.[49] Let a bracketed expression $(\phi)_i$ indicate that "ϕ" is to be added for the values i.

- $P = 0, 1$: infant $(\text{not})_0$ poisoned in P
- $A_1^- = 0, 1$: in A_1^-, infant $(\text{healthy})_0$ $(\text{sick})_1$ and antidote 1 $(\text{not})_0$ about to be administered
- $A_1^\circ = 0, 1$: in A_1°, infant $(\text{healthy})_0$ $(\text{sick})_1$and antidote 1 $(\text{not})_0$ being administered
- $A_1^+ = 0, 1$: in A_1^+, infant $(\text{healthy})_0$ $(\text{sick})_1$ and antidote 1 $(\text{not})_0$ just administered
- $A_1 = 0, 1 : (A_1^- = A_1^\circ = A_1^+ = 0)_0 \; (A_1^- = A_1^\circ = A_1^+ = 1)_1$
- $H_1 = 0, 1, 2$: antidote 1 $(\text{not})_{0,1}$ taking effect on $(\text{healthy})_0$ $(\text{sick})_{1,2}$ infant in H_1[50]
- $A_2 = 0, 1$: in A_2, infant $(\text{healthy})_0$ $(\text{sick})_1$ and antidote 2 $(\text{not})_0$ administered
- $H_2 = 0, 1, 2$: antidote 2 $(\text{not})_{0,1}$ taking effect on $(\text{healthy})_0$ $(\text{sick})_{1,2}$ infant in H_2

[48] We *could* consider infinitely many possible states for each region, and then condition on their disjunctions, using Regional Conglomerability (see Appendix C).

[49] So, for example: "infant" means something like *having the physiology typical of a neonate*, rather than *having been born some time ago*; "antidote about to be administered" means something like *your standing ready with the antidote, with the intention to administer it*, etc; and "antidote just administered" means *the bottle's being lifted from the infant's mouth, while the antidote is entering the infant's body*, etc. For readability I'll use the shorter expressions.

[50] An antidote can, of course, only take effect on a sick person.

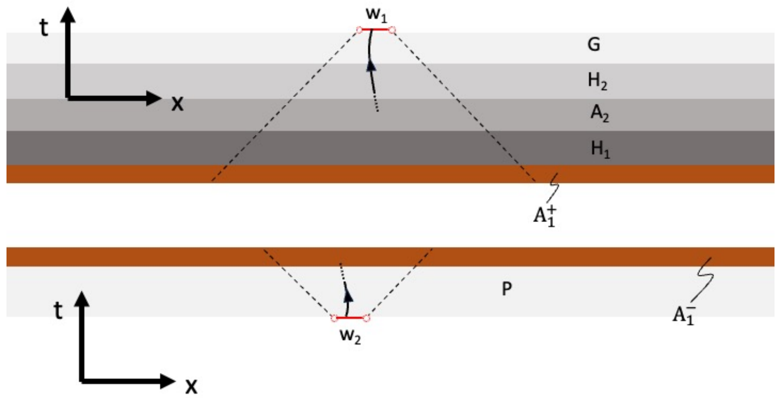


Figure 7: The result of deleting A_1° from $\mathcal{M}$.

- $G = 0, 1$: in G, infant (not)$_0$ alive, and (not)$_0$ eventually growing up to have two grandchildren

One salient interpretation of "chance, upon administration, of the first antidote's working" is $\text{urch}_\mathcal{M}(H_1 = 2|A_1 = 1)$—the urchance, in $\mathcal{M}$, of the first antidote's working conditional on administration in A_1. To calculate this, first note that $A_1^+ \cup A_1^-$ is a thick boundary of A_1° (cf. fig. 6). Thus consider $\mathcal{M}' := \mathcal{M} \backslash A_1^\circ$, the result of deleting A_1° from $\mathcal{M}$, sketched in figure 7. Since, given a complete geometric description of $\mathcal{M}$, $A_1 = 1$ is necessarily equivalent to $A_1^- = A_1^\circ = A_1^+ = 1$, we obtain, by Strong Boundary Markov:

$$\begin{aligned}\text{urch}_\mathcal{M}(H_1 = 2|A_1 = 1) &= \text{urch}_\mathcal{M}(H_1 = 2|A_1^- = A_1^\circ = A_1^+ = 1) \\ &\overset{SBM}{=} \text{urch}_{\mathcal{M}'}(H_1 = 2|A_1^- = A_1^+ = 1).\end{aligned} \tag{18}$$

Since A_1° intersects all closed causal curves in $\mathcal{M}$, $\mathcal{M}'$ contains no closed causal curves at all. So we can derive the right-hand side of eq. 18 from the acyclic chances.

To do this, let's assume for simplicity that the antidote's actions are the only indeterministic processes in $\mathcal{M}'$. In worlds without closed causal curves, there's a 50% chance that a given antidote works. Hence:

$$\text{urch}_{\mathcal{M}'}(H_1 = i|A_1^+ = 1) = \text{urch}_{\mathcal{M}'}(H_2 = i|A_2 = 1) = 0.5. \tag{19}$$

To pin down the rest of the acyclic dynamics, note that $\mathcal{M}'$ contains two disconnected infant spacetime worms: one starting in P (at birth) and ending at A_1^-'s future border, the other starting at A_1^+'s past border and ending somewhere in G (as the infant grows into an adolescent). These spacetime worms are disconnected because grandfather himself never travels through the wormhole. Call the spacetime worm from P to A_1^-'s future border *young infant*, and the spacetime worm from A_1^+'s past border to G *older infant*. For any propositions A, B, let $A \Rightarrow B$ denote that, deterministically, if A, then B.[51] (For convenience, I'll often suppress explicit mention of "deterministically" in the following, writing "...iff..." instead of "deterministically, ...iff...".) $A \Leftrightarrow B$ denotes $A \Rightarrow B \wedge B \Rightarrow A$.

(a) $H_1 = 0 \Leftrightarrow A_1^+ = 0$: older infant is healthy coming into H_1 iff he is healthy in A_1^+

(b) $A_2 = 1 \Leftrightarrow H_1 = 1$: an antidote is administered in A_2 iff older infant is still sick at the end of H_1

(c) $H_2 = 0 \Leftrightarrow A_2 = 0$: older infant is healthy coming into H_2 iff he is already healthy in A_2

(d) $G = 1 \Leftrightarrow H_2 = 0 \vee H_2 = 2$: older infant grows up to become a grandfather in G iff either he is healthy going into H_2, or he is sick going into H_2 but antidote 2 works

(e) $P = 1 \Leftrightarrow G = 1$: young infant is poisoned in P iff older infant in G grows up to be a grandfather[52]

(f) $A_1^- = 1 \Leftrightarrow P = 1$: antidote about to be administered in A_1^- iff young infant is poisoned in P

By eqs. (d), (e) and (f) we have $A_1^- = 1 \Leftrightarrow H_2 = 0 \vee H_2 = 2$—an antidote is about to be administered to the young infant in A_1^- iff the older infant survives in H_2 (either by being already healthy at the start of H_2 or by being healed in H_2). Meanwhile, from (a), (b), and (c) we have $A_1^+ = 1 \wedge H_2 = 0 \Leftrightarrow A_1^+ = 1 \wedge H_1 = 2$—if an antidote has entered the infant's stomach in A_1^+, then he is healthy at the start of H_2 iff the first antidote works in H_1. Those two equivalences jointly entail the following:

$$A_1^- = A_1^+ = 1 \Leftrightarrow (H_1 = 2 \vee H_2 = 2) \wedge A_1^+ = 1,$$

i.e., if an antidote has just been administered to the older infant in A_1^+, then [an antidote is *about* to be administered to the young infant in A_1^- iff at least one antidote works]. Plugging this into the right-hand side of eq. 18, we obtain the following:

$$\text{urch}_{\mathcal{M}}(H_1 = 2 | A_1 = 1) = \text{urch}_{\mathcal{M}'}(H_1 = 2 | A_1^+ = 1 \wedge (H_1 = 2 \vee H_2 = 2)). \qquad (20)$$

[51] In our urchance formalism, ⌜deterministically, A given B⌝ is equivalent to ⌜$\text{urch}(A|B \wedge Q) \equiv 1$ for all propositions Q⌝.

[52] This holds because one of the older infant's grandchildren poisons the young infant and, by assumption, nobody else possibly does.

That is, the probability, in $\mathcal{M}$, that the first antidote is effective conditional on its administration equals the probability, *in* $\mathcal{M}'$, that the first antidote is effective conditional on its administration *and* at least one of the two antidotes' working.

It's intuitively clear that the latter probability is greater than 1/2. That's because the guarantee that one of the two antidotes works raises the chance of each one's working—it excludes the possibility that both fail. More precisely, we find that[53]

$$\text{urch}_{\mathcal{M}'}(H_1 = 2 | A_1^+ = 1 \wedge (H_1 = 2 \vee H_2 = 2)) = \frac{1/2}{3/4} = \frac{2}{3}. \tag{22}$$

So, by eq. 20,

$$\text{urch}_{\mathcal{M}}(H_1 = 2 | A_1 = 1) = \frac{2}{3} > \frac{1}{2}.$$

Equally, the fact that one of the two antidotes works raises the chance of the second one's working: an analogous calculation yields $\text{urch}_{\mathcal{M}}(H_2 = 2 | A_1 = 1) = 1/3$, higher than the usual 1/4 (note that the second antidote is administered only if the first fails). We also obtain $\text{urch}_{\mathcal{M}}(H_2 = 2 | H_2 = 1) = 1$—given that antidote 1 fails, antidote 2 *must* work. In the presence of spacetime loops, chances differ in non-trivial but scrutable ways.

Our two scenarios provide us with a general recipe for deriving chances on loops.

General Recipe:

1. Given a spacetime $\mathcal{M}$ with closed causal curves, identify a region R intersecting all closed causal curves. Let Q_R a qualitative intrinsic state of R.
2. Identify a thick boundary B of R. Let Q_B be its complete intrinsic state.
3. Using Acyclic Chance Invariance, calculate, for any propositions X and Y

[53] *Proof:* By the multiplicative axiom,

$$\text{urch}_{\mathcal{M}'}(H_1 = 2 | A_1^+ = 1 \wedge (H_1 = 2 \vee H_2 = 2)) = \frac{\text{urch}_{\mathcal{M}'}(H_1 = 2 | A_1^+ = 1)}{\text{urch}_{\mathcal{M}'}(H_1 = 2 \vee H_2 = 2 | A_1^+ = 1)}, \tag{21}$$

provided the denominator isn't zero. Eq. 19 gives us $\text{urch}_{\mathcal{M}'}(H_1 = 2 | A_1^+ = 1) = 1/2$. Moreover, we obtain $\text{urch}_{\mathcal{M}'}(H_2 = 2 | A_1^+ = 1) = 1/4$ via the usual methods (the second antidote is administered only if the first one fails, and, if it is administered, has a 1/2 chance of healing the poisoning—all of this is encoded in eq. 19, (a), (b), and (c)). Finally, $H_1 = 2$ and $H_2 = 2$ are mutually exclusive (because, again, the second antidote is administered only if the first antidote fails). Plugging this all into eq. 21, we obtain

$$\text{urch}_{\mathcal{M}'}(H_1 = 2 | A_1^+ = 1 \wedge (H_1 = 2 \vee H_2 = 2)) = (1/2)/(3/4) = 2/3. \ \blacksquare$$

entirely about $\mathcal{M}\backslash R$,

$$\text{urch}_{\mathcal{M}\backslash R}(X|Y \wedge Q_B(B)).$$

4. By Strong Boundary Markov, this equals

$$\text{urch}_{\mathcal{M}}(X|Y \wedge Q_R(R) \wedge Q_B(B)).$$

This way, Strong Boundary Markov and Acyclic Chance Invariance determine certain *conditional* chances. A natural follow-up question is whether those conditional chances determine other chances too. The answer is a resounding Yes. In cyclic spacetimes, they determine precise *unconditional* chances for virtually all propositions: we generically obtain a full "probability map of the universe".[54] Meanwhile, where a wormhole is embedded into a larger spacetime, we obtain precise chance distributions over states of the loop region—including over what emerges from the wormhole—conditional on the state of the world *prior* to the loop region. The next section explains.

10 Marginal Chances

Several people have suggested to me in conversation that, generically, there are no well-defined chances of what comes out of a future wormhole, conditional only on the state of the world prior to it. In the same vein, one might think that there is no privileged way of assigning marginal (i.e., unconditional) chances over the states of a cyclic universe.

I once believed both things too. Certainly in *acyclic* worlds, like Minkowski spacetime, the transition chances alone generically *don't* fix marginal chance distributions over the states of the world; additionally, one requires a marginal chance distribution over the universe's possible initial conditions.[55]

But things are different with cyclic worlds. A specification of all transition chances generically fixes even a *marginal* chance distribution over the possible states of a cyclic world. Intuitively, the reason for this is that cyclic worlds have one "extra" transition compared to their linear counterparts: their "ends" also connect. This extra transition generically imposes additional constraints on the marginal distribution over the loop,

[54] Cf. Loewer (2019). Of course, in contrast to the Mentaculus, our probability maps for cyclic spacetimes don't require anything like a Past Hypothesis—everything is fixed by the *dynamics* alone.

[55] Some derive such a distribution from statistical mechanical considerations (e.g. Albert, 2000; Loewer, 2019). This abandons the idea that the urchance function is determined by the dynamical laws alone. In any case, the considerations cannot apply to closed causal curves, which lack unidirectional entropic arrows.

enough to fix it uniquely.[56]

To see this, consider a generalized version of CIRCLE, where the loop is n days round-trip and there are k possible particle colors, represented by natural numbers. (In the original CIRCLE case, $n \approx 5.5 \cdot 10^{12}$, and $k = 3$.) The probability axioms imply that, for every $i = 1, ..., n$ and $j = 1, ..., k$ (where we identify a 0 in an index with n):[57]

$$\text{urch}_C(\tau_i = j) = \sum_{l=1}^{k} \text{urch}_C(\tau_i = j | \tau_{i-1} = l) \cdot \text{urch}_C(\tau_{i-1} = l). \tag{23}$$

From the acyclic dynamics, our account obtains all cyclic transition probabilities, $\text{urch}_C(\tau_i = j|\tau_{i-1} = l)$. Hence eq. 23 yields $n \cdot k$ equations in $n \cdot k$ unknowns. Since we also know that, for every $i = 1, ..., n$,

$$\sum_{j=1}^{k} \text{urch}_C(\tau_i = j) = 1, \tag{24}$$

we can eliminate n equations and n unknowns from this system, leaving us with $n \cdot (k - 1)$ equations in $n \cdot (k - 1)$ unknowns. In Appendix D, we see that these equations generically are linearly independent and have a unique solution.

Solving the system (see eq. 34 in the Appendix—I'll skip the calculation here) for SMALL CIRCLE (i.e., $n = k = 3$, and every color i has only itself and $i + 1$ as a possible successor, with $1/65 \approx 0.015$ chance to transition to $i + 1$), we get the following result, for all $i = 1, 2, 3$:

$$\text{urch}_{SC}(\text{RED}(\tau_i)) = \text{urch}_{SC}(\text{GREEN}(\tau_i)) = \text{urch}_{SC}(\text{BLUE}(\tau_i)) = 1/3.$$

A sensible result, given the symmetry in transition probabilities between the colors: every color has, besides itself, a unique permissible successor, and each color has the same chance of switching to its respective successor. Breaking this symmetry in the transition probabilities also breaks the symmetry in the marginals. For example, if $\text{urch}_{SC}(\text{RED}(\tau_i)|\text{RED}(\tau_{i-1})) = 0.5$ for all $i = 1, 2, 3$ and the remaining transition chances are unchanged, we get $\text{urch}_{SC}(\text{RED}(\tau_i)) = 1/66 \approx 0.015$ and $\text{urch}_{SC}(\text{GREEN}(\tau_i)) = \text{urch}_{SC}(\text{BLUE}(\tau_i)) = 65/132 \approx 0.492$ for all $i = 1, 2, 3$. (The reader may verify this by plugging the given transition probabilities into eq. 34 in Appendix D.) We can of course also break the symmetry between the *times*, i.e., impose time-dependent transition probabilities,

[56]Mellor (1995, Sec. 17.3) once tried to leverage a mathematically similar fact into an argument *against* the possibility of spacetime loops, by claiming that it's impossible for transition chances to constrain marginal chances (or, in his framework, marginal *limiting frequencies*) in this way. See Berkovitz (2001, pp.14-5) for a cogent rebuttal.

[57]Marginal urchances are defined in the obvious way: $\text{urch}_C(A) := \text{urch}_C(A|\top)$, where $\top$ is any tautology.

which then also makes the marginal probabilities time-dependent.

Once we have marginal chance distributions over states of the loop, we can derive many other conditional chance distributions via the ratio formula. For example, consider a two-dimensional version of CIRCLE, i.e. a flat 2D spacetime rolled up along the time-like direction into a cylinder.[58] For any subset of the cylinder, our account determines the chance of any proposition conditional on any state of the subset, provided only that the latter has positive marginal chance.

I said that the linear equations are "generically" independent, because in special circumstances they aren't, allowing for multiple solutions (i.e., in our framework, a non-constant $\text{urch}_{\mathcal{M}}$). Roughly, this happens when there are *too many* deterministic transitions. For illustration, consider a variant of SMALL CIRCLE, where the particle is guaranteed to remain at its current color, i.e., for every $\text{COL} \in \{\text{RED}, \text{GREEN}, \text{BLUE}\}$,

$$\text{urch}_{SC}(\text{COL}(\tau_i)|\text{COL}(\tau_{i-1})) = 1.$$

Given these transition chances, the only constraint the system imposes on the marginals is $\text{urch}_{SC}(\text{COL}(\tau_{i-1})) = \text{urch}_{SC}(\text{COL}(\tau_i))$. Any probabilistically coherent[59] urchance function which satisfies this constraint is nomically allowed. (In Appendix D, I show this explicitly for the simplest non-trivial case, with $n = k = 2$. The case of SC, i.e. $n = k = 3$, is computationally more complex, but doesn't offer additional insight.)

So much for the case of a cyclic spacetime. The case of a spacetime with a wormhole is mathematically similar. Generically, our account yields well-defined transition chances on the region between the wormhole mouths, *conditional* on the state of the world prior to the wormhole. Once we have those transition chances, the remaining calculation is exactly the same. It follows that, conditional on the state of the world prior to the wormhole, there is (generically) a precise chance distribution over states of the loop region, including a precise chance distribution over what emerges from the wormhole. Only when enough transition probabilities are trivial—e.g. if the dynamics is deterministic—is there no such precise chance distribution.

[58]Mathematically, we can represent this by a two-dimensional, oriented, closed Lorentzian manifold.

[59]I.e., satisfying, for $i = 1, 2, 3$, $\sum_{\substack{\text{COL}\in \\ \{\text{RED,GREEN,BLUE}\}}} \text{urch}_{SC}(\text{COL}(\tau_i)) = 1$ (where addition is, as always, point-wise).

11 Conclusion

My theory of chances on loops consists of Strong Boundary Markov and Acyclic Chance Invariance. Strong Boundary Markov is an *a priori* plausible constraint on local laws, following from the idea that there shouldn't be a difference between *geometric* information about a region and other kinds of information about it—if the dynamics is local, both should be screened off by a thick boundary. Acyclic Chance Invariance, meanwhile, is a consistent weakening of an initially attractive, yet inconsistent, Chance Invariance principle. The weakening says that chances are invariant among *loop-free* worlds. Given the acyclic dynamics, these two general principles fix everything there is about chance in cyclic worlds.

Our theory satisfies all theoretical criteria we've set out. It avoids temporalism's triviality problem, and it avoids the inconsistency plaguing general Chance Invariance. Still, it manages to preserve the two next best things to Chance Invariance. The first is Acyclic Chance Invariance. The second is the idea that chances are "dynamically scrutable": while not identical to them, the cyclic chances should be *derivable from* the acyclic chances in a principled way. In our theory, Strong Boundary Markov and Acyclic Chance Invariance provide this principled connection. Finally, we also saw how under certain conditions—namely when, according to the acyclic chances, far-future events are increasingly probabilistically independent of near-future events—chances are *practically* invariant in the near-term.

While the essay's focus is objective chance, it naturally has implications for *rational credence*. Objective chance constrains credence via plausible deference principles.[60] For example, following the previous section, if you know the true (precise) transition chances governing a cyclic world, generically you should have precise prior credences over the world's possible states. Likewise, if you know the true chance laws, and are well-informed about the *current* state of the world, you should, in general, have precise expectations about what will emerge from a future wormhole.

Programmatically, this essay supports flexible chance formalisms, by demonstrating how they solve problems eluding other approaches. On our urchance formalism, *all* propositions—not just temporal or causal histories—are eligible background propositions. Of course, some propositions will be more informative than others. But we've seen that even regions much smaller than entire temporal or causal history regions can be highly informative.[61] Some of these regions cover exactly the local environments of coin flips,

[60] See fn. 11 for a deference principle for urchance.

[61] That is, assuming we additionally supply information about the world's background geometry—something which the temporalist also has to do.

of roulette wheels, or of decks of cards. The concept of a background proposition thus comes to subsume the concept of a *chance setup*. Indeed, conceptual economy suggests identifying the two: every background proposition is a chance setup, every chance setup a background proposition. With this identification, our framework then enshrines a view Popper (1959) embraced long ago: that chances are intimately tied, not to time or causation, but to chance setups ("*arrangements*", as he called them).

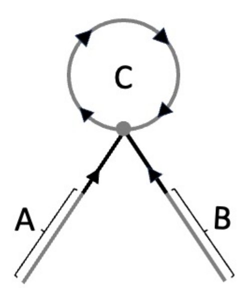
ACKNOWLEDGMENTS: The paper has benefited greatly from a great number of people. Special thanks to Cian Dorr, Dmitri Gallow, and Richard Roth for many hours of insightful conversation. I also thank David Albert, David Clyde, Sheldon Goldstein, Verónica Gómez Sánchez, Al Hajek, Moishe Kohan, Bar Luzon, Tim Maudlin, David Preber, Michael Strevens, Kun Zhang, Snow Zhang, and audiences at ANU Philosophy's lunch seminar and at the Society for the Metaphysics of Science 2024 Conference for helpful remarks, discussions, puzzles. Finally, I am indebted to an anonymous referee for their constructive and helpful feedback. Thank you all.

Appendix

A Parental Markov Failure

A.1 Where Parental Markov Is False...

Consider the following one-dimensional spacetime, FORK (or F for short):



Suppose the world is home to a scalar field, with deterministic dynamics. Specifically, suppose that the field everywhere takes the value 0 or 1, and that there are two kinds of spacetime points: *fork points*, where two or more lines converge (represented by the grey dot in the diagram above),[62] and *boring points*, all the others. In any interval consisting,

[62]Topologically, we can define a fork point as any point p such that there are two or more open lines

except for possibly its initial point, of boring points, the field remains constant. At a fork point, meanwhile, the field's value is determined by the values on the incoming lines. If it is 1 on exactly an even number of incoming lines, it is also 1 at the fork point and on all outgoing lines; otherwise it is 0 at the fork point and on all outgoing lines.

The above figure indicates three disjoint segments, A, B, and C. If we introduce for each a homonymous binary variable whose value represent the field value taken throughout the segment, the given dynamics entails the following (where $\overline{X} := (1 - X)$):

$$C = \overline{A}BC + A\overline{B}C + AB\overline{C} + \overline{A}\,\overline{B}\,\overline{C}. \tag{25}$$

If $A = B$, this becomes $C = \overline{C}$—contradiction. Hence eq. 25 entails that $A = \overline{B}$.

Now, the empty set is a pure thick parent of A, and A doesn't cause B. Hence, Parental Markov requires that the empty set screen off A from B:

$$\text{urch}_F(A = 1|B = 0) = \text{urch}_F(A = 1|B = 1).$$

But, since the dynamics requires $A = \overline{B}$,

$$\text{urch}_F(A = 1|B = 0) = 1 \neq 0 = \text{urch}_F(A = 1|B = 1).$$

So Parental Markov is false.[63]

A.2 ...Boundary Markov Is Still True

Yet Boundary Markov is still true. To see this, note that the fork point p partitions FORK into three disjoint segments: let $\mathcal{A}$ be the branch containing A, and $\mathcal{B}$ the branch containing B; the third segment is C itself, with $p \in C$. Since $\mathcal{A}$ and $\mathcal{B}$ are connected segments consisting of boring points only, the field value at one point in the segment nomically entails the field values everywhere else in the segment. The same is true for C, where all points can be connected to each other via segments consisting of boring points only. Let now R be any region, and let D be a thick boundary of R. Note that, whenever R and $R^{\perp}$ have non-empty intersection with $\mathcal{A}$, D also has non-empty intersection with $\mathcal{A}$; the same goes for $\mathcal{B}$ and C and arbitrary unions of $\{\mathcal{A}, \mathcal{B}, C\}$.[64] (Where X and Y are regions, I'll abbreviate "The field

containing p that do not share an open sub-line containing p.

[63]If you are worried that this case relies on trickeries with empty sets, it's straightforward to change it into one involving *non-empty* thick parents—just introduce additional intervals prior to A, generating a non-trivial chance distribution over the field values in A.

[64]In particular, note that if $R = \mathcal{A}$, then, due its thickness, D intersects both $\mathcal{B}$ and C. The same goes, *mutatis mutandis*, for $R = \mathcal{B}$ and $R = C$.

values throughout X entail the field values throughout Y" by "X entails Y".)

- Case 1: Suppose $D \cap \mathcal{A} \neq \emptyset$ or $D \cap \mathcal{B} \neq \emptyset$. It follows that D either entails $\mathcal{A}$ or entails $\mathcal{B}$. Since the laws (eq. 25) moreover entail $A = \overline{B}$, it follows that D entails $\mathcal{A} \cup \mathcal{B}$.
  - Case 1.1: Suppose $R \cap C \neq \emptyset$ and $R^{\perp} \cap C \neq \emptyset$. Then $D \cap C \neq \emptyset$ and so D entails C, and hence $\mathcal{A} \cup \mathcal{B} \cup C$. In particular, D entails R.
  - Case 1.2: Suppose $R^{\perp} \cap C = \emptyset$. Then $R^{\perp} \subseteq \mathcal{A} \cup \mathcal{B}$ and so D entails $R^{\perp}$.
  - Case 1.3: Suppose $R \cap C = \emptyset$. Then $R \subseteq \mathcal{A} \cup \mathcal{B}$ and so D entails R.
- Case 2: Suppose $D \cap \mathcal{A} = \emptyset$ and $D \cap \mathcal{B} = \emptyset$. Then $D \subseteq C$. It follows that either $\mathcal{A} \cup \mathcal{B} \subseteq R$ or $(\mathcal{A} \cup \mathcal{B}) \cap R = \emptyset$.[65]
  - Case 2.1: Suppose $\mathcal{A} \cup \mathcal{B} \subseteq R$. Then either $C \subseteq R$ or $D \cap C \neq \emptyset$. Since $R^{\perp} \subseteq C$, in either case D entails $R^{\perp}$.
  - Case 2.2: Suppose $(\mathcal{A} \cup \mathcal{B}) \cap R = \emptyset$. Then $R \subseteq C$ and either $R = \emptyset$ or $R \neq \emptyset$. If $R = \emptyset$, D triviallyentails R. If $R \neq \emptyset$, then either $C \subseteq R$ or $C \nsubseteq R$. If $C \subseteq R$, then $R = C$, and so $R^{\perp} = \mathcal{A} \cup \mathcal{B}$ and $D \cap (\mathcal{A} \cup \mathcal{B}) \neq \emptyset$. It follows that D entails $\mathcal{A} \cup \mathcal{B}$ and hence $R^{\perp}$. If $C \nsubseteq R$, then $R^{\perp} \cap C \neq \emptyset$ and so $D \cap C \neq \emptyset$. It follows that D entails R.

So, in every case, D either entails R or entails $R^{\perp}$. In particular,

$$\text{urch}_F(Q_1(R^{\perp})|Q_2(R) \cup Q_3(D)) = \text{urch}_F(Q_1(R^{\perp})|Q_3(D)),$$

for any intrinsic properties Q_1 and Q_2, and maximal intrinsic property Q_3 such that $Q_2(R) \cup Q_3(D)$ are nomically possible. ■

B Thick Neighborhoods = Neighborhoods of Closures

Let a *neighborhood of* R be any open superset of R. We've defined a *thick neighborhood* N as a neighborhood which satisfies the following additional condition: every continuous curve starting in $N^{\perp}$ and ending in R has a non-trivial subcurve in $N \backslash R$ before ever intersecting R. Here we prove that, in any space homeomorphic to $\mathbb{R}^n$, N is a thick neighborhood of R iff N is a neighborhood of R's closure, denoted $\overline{R}$. It follows that B is a thick boundary of

[65] For suppose otherwise, i.e. $(\mathcal{A} \cup \mathcal{B}) \cap R^{\perp} \neq \emptyset$ and $(\mathcal{A} \cup \mathcal{B}) \cap R \neq \emptyset$. Then $(\mathcal{A} \cup \mathcal{B}) \cap D \neq \emptyset$ and hence $D \nsubseteq C$.

R iff B is disjoint from R and $B \cup R$ contains a neighborhood of $\overline{R}$. We'll make ample use of this equivalence in Appendix C.

> ***Theorem 1.*** **Equivalence.** For any $R, N \subseteq \mathbb{R}^n$, N is a thick neighborhood of R iff N is a neighborhood of $\overline{R}$.

Proof: Right-to-left direction: Let N be a neighborhood of $\overline{R}$, and let c be a continuous curve which starts in $N^{\perp}$ and ends in R. Without loss of generality, assume $c : [0,1] \to \mathbb{R}^n$. Since c is continuous, $c^{-1}(\overline{R}) \subseteq [0,1]$ is closed and hence compact. Hence there is a first point $t^* \in [0,1]$ such that $q := c(t^*) \in \overline{R}$ and for all $t < t^*$, $c(t) \notin \overline{R}$. Since N is open and $q \in N$, there is an open ball $B(q) \subseteq N$ around q. Since c is continuous, $c^{-1}(B(q))$ is open, and so there is a $t^- < t^*$ such that $]t^-, t^*[$ is an open interval in $c^{-1}(B(q))$. Since $B(q) \subseteq N$ and for all $t < t^*$, $c(t) \notin \overline{R}$, the subcurve $c|_{]t^-,t^*[}$ is a non-trivial subcurve of c in $N \backslash \overline{R}$ prior to ever intersecting $\overline{R}$, and so in particular a non-trivial subcurve in $N \backslash R$ prior to ever intersecting R.

Left-to-right direction: Suppose, for contradiction, that N is a thick neighborhood of R but not a neighborhood of $\overline{R}$. Then $\overline{R} \nsubseteq N$, and so there is a $q \in \overline{R} \backslash N$. We now construct a continuous curve starting in q such that every non-trivial initial segment of it intersects R. Since $q \in \overline{R}$, for every $r > 0$ the open ball $B_r(q)$ has non-empty intersection with R. For every $n \in \mathbb{N}_{>0}$, choose a point q_n in $B_{1/n}(q) \cap R$. Let $c : [0,1] \to \mathbb{R}^n$ be such that $c(0) = q$, for all $n \in \mathbb{N}_{>0}$ $c(1/n) = q_n$, and c maps $]1/(n+1), 1/n[$ continuously to the straight line from q_{n+1} to q_n, excluding endpoints. We now prove that c is continuous. The subcurve $c|_{]0,1]}$ maps $]0,1]$ into a concatenation of straight lines, and hence is continuous. It remains to prove that c is continuous at 0. Let $\{a_k\}_{k \in \mathbb{N}}$ be any sequence in $[0,1]$ converging to 0. Since balls in $\mathbb{R}^n$ are convex, for every $r \in [0,1[$, $c(r) \in B_m(q)$ where m is the largest integer such that $r < 1/m$. Since $\{a_k\}_{k \in \mathbb{N}}$ converges to 0, for every $\delta \in]0,1[$ there is a $k \in \mathbb{N}$ such that for all $l > k$, $a_l \in [0, \delta[$, and hence (by the foregoing) $c(a_l) \in B_{1/m}(q)$ where m is the largest integer such that $\delta < 1/m$. Let now $\varepsilon \in]0,1]$. Then there is a smallest integer $m \geq 2$ such that $1/(m-1) < \varepsilon$. Since $1/m \in]0,1[$, it follows from the foregoing that there is a $k \in \mathbb{N}$ such that for every $l > k$, $a_l \in [0, 1/m[$ and $c(a_l) \in B_{1/(m-1)}(q)$ (since $m-1$ is the largest integer such that $1/m < 1/(m-1)$). Since $1/(m-1) < \varepsilon$, it follows that $c(a_l) \in B_\varepsilon(q)$. So, for every $\varepsilon \in]0,1]$, there is a $\delta \in]0,1]$ (namely $\delta = 1/m$) such that for all l with $|a_l| < \delta$, $c(a_l) \in B_\varepsilon(q)$. This proves that c is continuous at 0. So c is continuous over $[0,1]$. Hence c is a continuous curve which starts in $N^{\perp}$, ends in R, and every non-trivial initial segment intersects R; in particular, it has no non-trivial subcurve in $R^{\perp} \supseteq N \backslash R$ before ever intersecting R, in contradiction with the assumption that N is a thick neighborhood of R. So N is a neighborhood of $\overline{R}$. ■

C Parental Markov and Boundary Markov

Here we prove that, given plausible conglomerability and locality assumptions about urchance, Parental Markov already ensures, for sufficiently well-behaved regions, that they are screened off by their thick boundaries in Minkowski spacetime.

We'll first establish some auxiliary lemmas. As before, a *neighborhood* of A is any open superset of A. Where B is a thick boundary of R, let $B_R^+ := K^+(R) \cap B$, and $B_R^- := B \backslash B_R^+$ the rest of B. (To recall, $K^+(R)$ denotes R's proper causal future.) Note also the following elementary fact: if B is a thick boundary of R, then every continuous curve starting in $(R \cup B)^\perp$ and ending in R has a non-trivial subcurve in B before ever intersecting R.[66] Call a region *causally convex* iff it contains all causal curves starting and ending in it. ("Sufficiently well-behaved" will denote a mild strengthening of causal convexity.)

> ***Lemma 1*:** For any causally convex region R and any thick boundary B of R, B_R^- is a pure thick parent of R.

Proof of Lemma 1: Let c be a future-directed causal curve starting in $(R \cup B_R^-)^\perp$ and ending in R. Since c starts in $R^\perp$ and ends in R, c starts, specifically, in $K^-(R) \subseteq R^\perp$. Since R is causally convex, $K^+(R) \cap K^-(R) = \emptyset$. It follows that c starts in $(R \cup B)^\perp$ and ends in R. Since B is a thick boundary of R, c therefore has a non-trivial subcurve in B before ever intersecting R. Suppose, for contradiction, that there is a point q where c intersects B_R^+ and let $t \in [0,1]$ such that $c(t) = q$. Since q is in $K^+(R)$, there is a future-directed causal curve c' starting in R and ending in q. Concatenating c' and $c|_{[t,1]}$ thus yields a future-directed causal curve starting in R, intersecting B_R^+ and ending in R. Since $B_R^+ \subseteq R^\perp$, this contradicts R's causal convexity. So c doesn't intersect B_R^+. Hence c has a non-trivial subcurve in $B \backslash B_R^+ = B_R^-$ before ever intersecting R. So B_R^- is a thick parent of R. Finally, suppose for contradiction that R causes some point r in B_R^-. Then there is a future-directed causal curve c^* starting in R and ending in r. Since r is in $K^-(R)$, there is a future-directed causal curve c^{**} starting in r and ending in R. Concatenating c^* and c^{**} thus yields a future-directed causal curve starting in R, intersecting B_R^-, and ending in R. Since $B_R^- \subseteq R^\perp$, this contradicts R's causal convexity. Hence B_R^- is a pure thick parent of R. ■

[66] The reverse implication fails, however. In $\mathbb{R}$, consider $B = \bigcup_{n=1}^{\infty}]-\frac{1}{n}, -\frac{1}{n+1}[\cup \{0\}$ and $R =]0, +\infty[$. Every continuous curve which starts in $(R \cup B)^\perp$ and ends in R has a non-trivial subcurve in B before ever intersecting R. But no *open subset* of $(R \cup B)$ has that property—that is, no open set $N \subseteq R \cup B$ is such that every curve starting in $N^\perp$ and ending in R has a non-trivial subcurve in N before ever intersecting R. Hence B isn't a thick boundary of R. The Equivalence theorem (Appendix B) relies on this extra strength in the definition of "thick boundary".

***Lemma 2*:** If R is causally convex, $K^+(R)$ fully contains all future-directed causal curves starting in it; in particular, $K^+(R)$ is causally convex.

Proof of Lemma 2: Let R be causally convex and suppose for contradiction that there is a future-directed causal curve c starting in $K^+(R)$ and intersecting $(K^+(R))^\perp$. By the definition of $J^+(R)$, $J^+(R)$ contains all future-directed causal curves starting in it. Since $K^+(R) = J^+(R)\backslash R$, c thus intersects R in some point q; choose a $t \in [0,1]$ such that $c(t) = q$. Let p be c's starting point. Since $p \in J^+(R)$, there is a future-directed causal curve c^* starting in R and ending in p. Concatenating c^* and $c|_{[0,t]}$ yields a future-directed causal curve that starts in R, intersects $K^+(R) \subseteq R^\perp$, and ends in R, in contradiction with R's causal convexity. So $K^+(R)$ contains all future-directed causal curves starting in it. It follows that $K^+(R)$ is causally convex. ■

Let a *thick child* of R be any set C such that every future-directed causal curve starting in R and ending in $(R \cup C)^\perp$ has a non-trivial subcurve in C before ever intersecting $(R \cup C)^\perp$.

***Lemma 3*:** For any region R in Minkowski spacetime and any thick boundary B of R, if both R and $R \cup B$ are causally convex, then

(i) B_R^+ is a thick child of R, and
(ii) $R \cup B_R^-$ is a pure thick parent of B_R^+.

Proof of Lemma 3: (i): Let c be a future-directed causal curve starting in R and intersecting $(R \cup (B_R^+))^\perp$. Since R is causally convex, $K^+(R) \cap K^-(R) = \emptyset$, and so $K^+(R) \cap B_R^- = \emptyset$. Since also $R \cap B_R^- = \emptyset$, we have $J^+(R) \cap B_R^- = \emptyset$. But $J^+(R)$ contains c, so c doesn't intersect B_R^-. Since $B = B_R^+ \cup B_R^-$, c thus intersects $(R \cup B)^\perp$. Since B is a thick boundary of R, B is a thick boundary of $(R \cup B)^\perp$. Hence c has a non-trivial subcurve in B before ever intersecting $(R \cup B)^\perp$. Since c doesn't intersect B_R^-, it follows that c has a non-trivial subcurve in B_R^+ before ever intersecting $(R \cup B_R^+)^\perp$. So B_R^+ is a thick child of R.

(ii): In Minkowski spacetime, the causal future of a neighborhood of A's closure is a neighborhood of the closure of A's causal future.[67] Since B is a thick boundary of R, $R \cup B$ contains a neighborhood N of $\overline{R}$ (cf. Equivalence, Appendix B). Thus $J^+(N)$ is a neighborhood of $\overline{J^+(R)}$. Since $J^+(N) \subseteq J^+(R \cup B)$, $J^+(R \cup B)$ thus contains a neighborhood of $\overline{J^+(R)}$, and since $\overline{B_R^+} \subseteq \overline{J^+(R)}$, $J^+(R \cup B)$ contains a neighborhood of $\overline{B_R^+}$. Therefore, all

[67] To see this, note the following three facts about Minkowski spacetime (the first is true for any spacetime):
1. If $A \subseteq B$, then $J^+(A) \subseteq J^+(B)$.
2. Closure and causal future "commute", i.e. $J^+(\overline{A}) = \overline{J^+(A)}$ for any region A.
3. If A is open, $J^+(A)$ is open.

Let then N be a neighborhood of $\bar{A}$. Since $\bar{A} \subseteq N$ we have, by the first fact, $J^+(\bar{A}) \subseteq J^+(N)$. By the second fact, $\overline{J^+(A)} \subseteq J^+(N)$. Finally, by the third fact, $J^+(N)$ is open, and hence a neighborhood of $\overline{J^+(A)}$.

continuous curves which start in $(J^+(R \cup B))^\perp$ and end in B_R^+ have a non-trivial subcurve in $J^+(R \cup B) \backslash B_R^+$ before ever intersecting B_R^+. Let c be a future-directed causal curve which starts in $(R \cup B)^\perp$ and ends in B_R^+. Since $(R \cup B)^\perp \subseteq (J^+(R \cup B))^\perp$, c has a non-trivial subcurve in $J^+(R \cup B) \backslash B_R^+$ before ever intersecting B_R^+. Suppose, for contradiction, that c doesn't have a non-trivial subcurve in $R \cup B_R^-$ before ever intersecting B_R^+. Then c must intersect $\left(J^+(R \cup B) \backslash B_R^+\right) \backslash (R \cup B_R^-) = J^+(R \cup B) \backslash (R \cup B) = K^+(R \cup B)$. Let q be a point in $K^+(R \cup B)$ which c intersects, and choose a $t \in [0, 1]$ with $c(t) = q$. Since $q \in K^+(R \cup B)$, there is a future-directed causal curve c^* starting in $R \cup B$ and ending in q. Concatenating c^* and $c|_{[t,1]}$ thus yields a future-directed causal curve that starts in $R \cup B$, intersects $K^+(R \cup B) \subseteq (R \cup B)^\perp$, and ends in $B_R^+ \subseteq (R \cup B)$, in contradiction with $R \cup B$'s causal convexity. So c has a non-trivial subcurve in $R \cup B_R^-$ before ever intersecting B_R^+. Hence $R \cup B_R^-$ is a thick parent of B_R^+. Finally, suppose for contradiction that B_R^+ causes $R \cup B_R^-$. Since $B_R^+ \subseteq K^+(R)$ and $R \cup B_R^- \subseteq K^+(R)^\perp$, it follows that $K^+(R)$ causes $K^+(R)^\perp$. Since $J^+(R) = K^+(R) \cup R$ contains all future-directed causal curves which start in it, it follows that $K^+(R)$ causes R. Since R causes every point in $K^+(R)$, there is thus a future-directed causal curve starting in R, intersecting $K^+(R) \subseteq R^\perp$, and ending in R, in contradiction with R's causal convexity. Hence $R \cup B_R^-$ is a pure thick parent of B_R^+. ■

For any spacetime region X, let $\mathcal{R}(X)$ denote the set of all possible maximally specific intrinsic properties of X. For any $x \in \mathcal{R}(X)$, I'll also write $X = x$ instead of $x(X)$. For any $\mathbf{x} \subseteq \mathcal{R}(X)$, let $X \in \mathbf{x}$ denote the proposition that $X = x$ for some $x \in \mathbf{x}$. Let an *urchance candidate* be any function $\mathfrak{u}$ mapping pairs of propositions to functions on total, primitively conditional probability functions, such that, where u is such a function, $\mathfrak{u}(X, Y)(u) = u(X, Y)$. (In particular, then, urchance functions are urchance candidates.)

> ***Def.* Regional Conglomerability:** An urchance candidate $\mathfrak{u}$ is *regionally conglomerable* iff, for any spacetime region X, any $\mathbf{x} \subseteq \mathcal{R}(X)$, any propositions A and B, and any $a, b \in [0, 1]$ with $a < b$: if $a \leq \mathfrak{u}(A|X = x \wedge B) \leq b$ for all $x \in \mathbf{x}$, then
>
> $$a \leq \mathfrak{u}(A|X \in \mathbf{x} \wedge B) \leq b.$$

Where $\mathfrak{u}$ is an urchance candidate, let $\mathcal{R}(Z)_\mathfrak{u}$ denote the set of properties which are possible maximally specific intrinsic properties of Z *according to* $\mathfrak{u}$.[68] Let $(\mathcal{R}(Y) \times \mathcal{R}(Z))_\mathfrak{u}$

[68] In our system, a proposition A is *possible according to* $\mathfrak{u}$ iff $\mathfrak{u}(\neg A|A) < 1$, and A is *necessary according to* $\mathfrak{u}$ iff $\mathfrak{u}(A|\neg A) = 1$. These are dual provided all functions in $\mathfrak{u}$'s range agree on what's impossible; that is, for all u, u' in $\mathfrak{u}$'s range and all A, $u(\neg A|A) = 1$ iff $u'(\neg A|A) = 1$. (To see this: $\neg A$ is impossible iff $\mathfrak{u}(A|\neg A) \not< 1$, i.e. iff for some u in $\mathfrak{u}$'s range, $u(A|\neg A) = 1$. But given that all u agree on what's impossible, this is the case iff, for *all* u in $\mathfrak{u}$'s range, $u(A|\neg A) = 1$, i.e., iff $\mathfrak{u}(A|\neg A) = 1$.) Recall that, by assumption, all functions in an *urchance function*'s range agree on what's impossible—cf. fn. 13. Hence "possible according to urch" and

denote the set of property *pairs* (y, z) such that: according to $\mathfrak{u}$, y and z are possible maximally specific intrinsic properties of Y and Z, respectively, *and* it's possible, according to $\mathfrak{u}$, that $Y = y \wedge Z = z$.

In addition to conglomerability, we need an additional locality assumption: information about a region is "nothing over and above" information about its parts. That is, necessarily, for any regions X, Y, maximally specifying X and Y, in a nomically compatible way, also *maximally* specifies their union $X \cup Y$ in a nomically possible way; moreover *every* nomically possible maximal property of the union can be so specified. Call this property *Separability*. Technically:

> ***Def.*** **Separability:** An urchance candidate $\mathfrak{u}$ is *separable* iff, for any spacetime regions X, Y, there is a one-to-one correspondence
>
> $$\phi : (\mathcal{R}(X) \times \mathcal{R}(Y))_{\mathfrak{u}} \to \mathcal{R}(X \cup Y)_{\mathfrak{u}},$$
>
> such that for any $(x, y) \in (\mathcal{R}(X) \times \mathcal{R}(Y))_{\mathfrak{u}}$, necessarily according to $\mathfrak{u}$, $(X = x \wedge Y = y) \leftrightarrow (X \cup Y = \phi(x, y))$.

Whenever such a ϕ exists I'll slightly abuse notation and write $(\mathcal{R}(X) \times \mathcal{R}(Y))_{\mathfrak{u}} = \mathcal{R}(X \cup Y)_{\mathfrak{u}}$, as well as $X \cup Y = (x, y)$ instead of $X \cup Y = \phi(x, y)$.

For any urchance candidate $\mathfrak{u}$ and for any regions X, Y, Z, let $(X \perp\!\!\!\perp Y|Z)_{\mathfrak{u}}$ denote that Z *screens off X from Y according to* $\mathfrak{u}$; that is:

$$(X \perp\!\!\!\perp Y|Z)_{\mathfrak{u}} \text{ iff: for all } \mathbf{x} \subseteq \mathcal{R}(X), \mathbf{y} \times \{z\} \subseteq (\mathcal{R}(Y) \times \mathcal{R}(Z))_{\mathfrak{u}},$$
$$\mathfrak{u}(X \in \mathbf{x}|Y \in \mathbf{y} \wedge Z = z) = \mathfrak{u}(X \in \mathbf{x}|Z = z).$$

We have the following lemma (the names of the conditions follow Pearl's (1985) nomenclature for graphoids):

> ***Lemma 4.*** **Spacetime Graphoid Theorems:** For any urchance candidate $\mathfrak{u}$, if $\mathfrak{u}$ is regionally conglomerable and separable, then for any regions X, Y, Z, and W:
>
> - *Contraction:* If $(X \perp\!\!\!\perp Y|Z)_{\mathfrak{u}}$ and $(X \perp\!\!\!\perp W|Z \cup Y)_{\mathfrak{u}}$, then $(X \perp\!\!\!\perp W \cup Y|Z)_{\mathfrak{u}}$.
> - *Weak Union:* If $(X \perp\!\!\!\perp Y \cup W|Z)_{\mathfrak{u}}$, then $(X \perp\!\!\!\perp Y|Z \cup W)_{\mathfrak{u}}$.[69]

"necessary according to urch" are dual if urch is an urchance function.

[69]The graphoid theorem *Decomposition*—if $(X \perp\!\!\!\perp W \cup Y|Z)_{\mathfrak{u}}$, then $(X \perp\!\!\!\perp W|Z)_{\mathfrak{u}} \wedge (X \perp\!\!\!\perp Y|Z)_{\mathfrak{u}}$—is also valid, and the proof is immediate. By contrast, only a restricted version of *Symmetry* is valid: *provided* that [for all $\mathbf{x} \subseteq \mathcal{R}(X)_{\mathfrak{u}}$ and $z \in \mathcal{R}(Z)_{\mathfrak{u}}$, $\mathfrak{u}(X \in \mathbf{x}|Z = z) > 0$], $(X \perp\!\!\!\perp Y|Z)_{\mathfrak{u}}$ implies $(Y \perp\!\!\!\perp X|Z)_{\mathfrak{u}}$. Likewise, only a restricted version of *Intersection* is valid: *provided* that [whenever $(w, z) \in (\mathcal{R}(W) \times \mathcal{R}(Z))_{\mathfrak{u}}$ and $(y, z) \in (\mathcal{R}(Y) \times \mathcal{R}(Z))_{\mathfrak{u}}$, $(w, y, z) \in (\mathcal{R}(W) \times \mathcal{R}(Y) \times \mathcal{R}(Z))_{\mathfrak{u}}$], $(X \perp\!\!\!\perp Y|Z \cup W)_{\mathfrak{u}}$ and $(X \perp\!\!\!\perp W|Z \cup Y)_{\mathfrak{u}}$ imply $(X \perp\!\!\!\perp W \cup Y|Z)_{\mathfrak{u}}$. As none of these three properties will play a role in the following, I'll omit the proofs.

Proof of Lemma 4:

Contraction: Suppose $(X \perp\!\!\!\perp Y | Z)_{\mathfrak{u}}$ and $(X \perp\!\!\!\perp W | Z \cup Y)_{\mathfrak{u}}$. That is: for any $\mathbf{y} \times \{z\} \subseteq (\mathcal{R}(Y) \times \mathcal{R}(Z))_{\mathfrak{u}}$,

$$\mathfrak{u}(X \in \mathbf{x} | Y \in \mathbf{y} \wedge Z = z) = \mathfrak{u}(X \in \mathbf{x} | Z = z), \tag{26}$$

and for any $\mathbf{w} \times \{(y,z)\} \subseteq (\mathcal{R}(W) \times \mathcal{R}(Y) \times \mathcal{R}(Z))_{\mathfrak{u}}$,

$$\mathfrak{u}(X \in \mathbf{x} | W \in \mathbf{w} \wedge Y = y \wedge Z = z) = \mathfrak{u}(X \in \mathbf{x} | Y = y \wedge Z = z). \tag{27}$$

By eqs. 26 and 27, for any $(w,y,z) \in (\mathcal{R}(W) \times \mathcal{R}(Y) \times \mathcal{R}(Z))_{\mathfrak{u}}$,

$$\begin{aligned} \mathfrak{u}(X \in \mathbf{x} | W = w \wedge Y = y \wedge Z = z) &= \mathfrak{u}(X \in \mathbf{x} | Y = y \wedge Z = z) \\ &= \mathfrak{u}(X \in \mathbf{x} | Z = z). \end{aligned} \tag{28}$$

But by Separability, $(\mathcal{R}(W) \times \mathcal{R}(Y) \times \mathcal{R}(Z))_{\mathfrak{u}} = (\mathcal{R}(W \cup Y) \times \mathcal{R}(Z))_{\mathfrak{u}}$. Thus, by eq. 28 and Regional Conglomerability, for any $\mathbf{v} \subseteq \mathcal{R}(W \cup Y)_{\mathfrak{u}}$ such that $\mathbf{v} \times \{z\} \subseteq (\mathcal{R}(W \cup Y) \times \mathcal{R}(Z))_{\mathfrak{u}}$,

$$\mathfrak{u}(X \in \mathbf{x} | W \cup Y \in \mathbf{v} \wedge Z = z) = \mathfrak{u}(X \in \mathbf{x} | Z = z).$$

But this is just $(X \perp\!\!\!\perp W \cup Y | Z)_{\mathfrak{u}}$. ■

Weak Union: Suppose $(X \perp\!\!\!\perp Y \cup W | Z)_{\mathfrak{u}}$, i.e. for all $\mathbf{x} \subseteq \mathcal{R}(X)_{\mathfrak{u}}$ and $\mathbf{v} \times \{z\} \subseteq (\mathcal{R}(Y \cup W) \times \mathcal{R}(Z))_{\mathfrak{u}}$,

$$\mathfrak{u}(X \in \mathbf{x} | Y \cup W \in \mathbf{v} \wedge Z = z) = \mathfrak{u}(X \in \mathbf{x} | Z = z). \tag{29}$$

For any $(w,z) \in (\mathcal{R}(W) \times \mathcal{R}(Z))_{\mathfrak{u}}$, let $\top_{w,z} \subseteq \mathcal{R}(Y)$ be such that, according to $\mathfrak{u}$, $W = w \wedge Z = z$ entails $Y \in \top_{w,z}$. Then

$$\begin{aligned} \mathfrak{u}(X \in \mathbf{x} | W = w \wedge Z = z) &= \mathfrak{u}(X \in \mathbf{x} | Y \in \top_{w,z} \wedge W = w \wedge Z = z) \\ &= \mathfrak{u}(X \in \mathbf{x} | Y \cup W \in \top_{w,z} \times \{w\} \wedge Z = z) \\ &\overset{\text{eq.29}}{=} \mathfrak{u}(X \in \mathbf{x} | Z = z). \end{aligned} \tag{30}$$

But now, for any $\mathbf{y} \times \{w\} \times \{z\} \subseteq (\mathcal{R}(Y) \times \mathcal{R}(W) \times \mathcal{R}(Z))_{\mathfrak{u}}$,

$$\begin{aligned} \mathfrak{u}(X \in \mathbf{x} | Y \in \mathbf{y} \wedge W = w \wedge Z = z) &= \mathfrak{u}(X \in \mathbf{x} | Y \cup W \in \mathbf{y} \times \{w\} \wedge Z = z) \\ &\overset{\textbf{eq.}29}{=} \mathfrak{u}(X \in \mathbf{x} | Z = z) \\ &\overset{\textbf{eq.}30}{=} \mathfrak{u}(X \in \mathbf{x} | W = w \wedge Z = z). \end{aligned}$$

But, by Separability, $(\mathcal{R}(W) \times \mathcal{R}(Z))_{\mathfrak{u}} = \mathcal{R}(W \cup Z)_{\mathfrak{u}}$. So this is just $(X \perp\!\!\!\perp Y | Z \cup W)_{\mathfrak{u}}$. ∎

The conglomerability and locality constraints on urchance are just this:

> ***Thesis.*** Necessarily, the urchance function is regionally conglomerable and separable.

This entails, by Lemma 4, that urchance validates Contraction and Weak Union.

As in the main text, when I speak of "Z **screens off** X **from** Y" (*simpliciter*)—denoted $X \perp\!\!\!\perp Y | Z$—I mean *screens off according to the urchance function, conditioned on a complete description of the world's geometry*. Finally, say that a region R is **tolerantly causally convex** iff it is causally convex and every thick boundary of R contains a thick boundary B' of R such that $R \cup B'$ is causally convex. (This is the aforementioned strengthening of causal convexity.) Many regions we typically consider are tolerantly causally convex—including any history segment (e.g., in Minkowski spacetime, the region between two Cauchy surfaces) and any light-cone segment (i.e., the intersection of a history segment and a past or future light-cone). We can now state the main result.

> ***Theorem.*** **Parental Markov & Boundary Markov.** In Minkowski spacetime, Parental Markov entails that, for all regions R, if R is tolerantly causally convex and B is a thick boundary of R, then B screens off R from $R^{\perp}$.

Proof of Theorem: Let R be causally convex and B^* a thick boundary of R. Since R is tolerantly causally convex, there is a $B \subseteq B^*$ such that B is a thick boundary of R and $R \cup B$ is causally convex. Once we show that $R \perp\!\!\!\perp R^{\perp} | B$, the desired result follows immediately by Weak Union: for, since $B^* \subseteq R^{\perp}$ and $B \subseteq B^*$, $R \perp\!\!\!\perp R^{\perp} | B$ entails $R \perp\!\!\!\perp R^{\perp} | B^*$.

Now, to prove $R \perp\!\!\!\perp R^{\perp} | B$, note the following three facts:

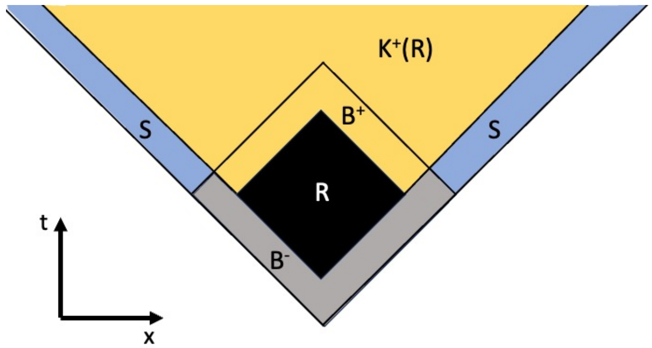
1. B_R^- is a pure thick parent of R.

2. There is a region S disjoint from $R \cup B$ such that $R \cup B_R^- \cup S$ is a pure thick parent of $K^+(R)$. ("S" stands for "spouse".)

3. B_R^+ is a thick child of R, and $R \cup B_R^-$ is a pure thick parent of B_R^+.

The first fact follows from Lemma 1 and R's causal convexity. The second fact follows like this: by Lemma 2, $K^+(R)$ is causally convex, and so, by Lemma 1 again, $K^+(R)$ has a pure thick parent. Since R is causally convex, $K^+(R)$ contains all causal curves which start in it, and so $K^+(R)$ doesn't cause $R \cup B_R^-$. Hence there is a pure thick parent P of $K^+(R)$ which contains $R \cup B_R^-$. Now simply choose $S := P \backslash (R \cup$

B_R^-). The third fact is just Lemma 3. The following sketch offers some orientation:



Below we show the following three facts:

$$R \perp\!\!\!\perp R^{\perp} \mid B \cup S \cup (K^{+}(R) \backslash B), \tag{31}$$
$$R \perp\!\!\!\perp (K^{+}(R) \backslash B) \mid B \cup S, \tag{32}$$
$$R \perp\!\!\!\perp S \mid B. \tag{33}$$

Once these are proven, $R \perp\!\!\!\perp R^{\perp} | B$ follows: facts 32 and 33 entail, by Contraction,

$$R \perp\!\!\!\perp (K^{+}(R) \backslash B) \cup S | B.$$

Contracting again with fact 31 gives

$$R \perp\!\!\!\perp R^{\perp} \cup (K^{+}(R) \backslash B) \cup S | B.$$

But $K^{+}(R) \backslash B \cup S \subseteq R^{\perp}$, hence

$$R \perp\!\!\!\perp R^{\perp} | B.$$

Proof of fact 31: Recall that $R \cup B_R^- \cup S$ is a pure thick parent of $K^{+}(R)$. By Lemma 1, B_R^- is a pure thick parent of R. Since S is disjoint from R, it follows that $B_R^- \cup S$ is a thick parent of $R \cup K^{+}(R) = J^{+}(R)$.[70] Moreover, since S and B_R^- are disjoint from R and subsets of a

[70]This follows from the following general lemma (instantiate V with $K^{+}(R)$; P with $R \cup B_R^- \cup S$; T with B_R^-; and P' with R):

> **Lemma:** Let V be any region and P be a thick parent of V. Then, if T is a thick parent of some subset $P' \subseteq P$, then $T \cup (P \backslash P')$ is a thick parent of $V \cup P'$.

Proof of lemma: Let P be a thick parent of V, and T be a thick parent of a subset $P' \subseteq P$. Let c be a future-directed causal curve starting in $(T \cup P \cup V)^{\perp}$ and ending in $V \cup P'$. We show that c has a non-trivial subcurve in $T \cup (P \backslash P')$ before ever intersecting $V \cup P'$. Then c either ends in P' or ends in V.

thick parent of $K^+(R)$, and thus also disjoint from $K^+(R)$, $B_R^- \cup S$ is a *pure* thick parent of $J^+(R)$. Since $J^+(R)$ contains every future-directed causal curve starting in it, $J^+(R)$ doesn't cause its complement $(J^+(R))^\perp$. Hence, by Parental Markov, $J^+(R) \perp\!\!\!\perp (J^+(R))^\perp | B_R^- \cup S$. By Weak Union, $R \perp\!\!\!\perp (J^+(R))^\perp | K^+(R) \cup B_R^- \cup S$, and thus $R \perp\!\!\!\perp (J^+(R))^\perp \cup K^+(R) | K^+(R) \cup B_R^- \cup S$. By definition, $(J^+(R))^\perp \cup K^+(R) = R^\perp$ and $K^+(R) \cup B_R^- = B \cup (K^+(R) \backslash B)$. So $R \perp\!\!\!\perp R^\perp | B \cup (K^+(R)\backslash B) \cup S$.

Proof of fact 32: Recall, again, that $R \cup B_R^- \cup S$ is a pure thick parent of $K^+(R)$. By Lemma 3(i), B_R^+ is a thick child of R. It follows that $B_R^+ \cup B_R^- \cup S = B \cup S$ is a thick parent of $K^+(R)\backslash B_R^+$.[71] We now show that $K^+(R)\backslash B_R^+$ doesn't cause $B \cup S$. Since R is causally convex, $K^+(R)$ contains all future-directed causal curves starting in $K^+(R)$. In particular, $K^+(R)$, and hence $K^+(R)\backslash B_R^+$, don't cause $B_R^- \cup S$. Now suppose, for contradiction, that $K^+(R)\backslash B_R^+$ causes B_R^+. Then there is a future-directed causal curve c starting in $K^+(R)\backslash B_R^+ = K^+(R)\backslash B \subseteq (R \cup B)^\perp$ and ending in $B_R^+ \subseteq R \cup B$. Since c starts in $K^+(R)$, there is a future-directed causal curve c^* from R to c's starting point. Concatenating c^* and c yields a future-directed causal curve that starts in $R \cup B$, intersects $(R \cup B)^\perp$, and ends in $R \cup B$, in contradiction with $R \cup B$'s causal convexity. So $K^+(R)\backslash B_R^+$ doesn't cause B_R^+. Hence $K^+(R)\backslash B_R^+$ doesn't cause $B_R^+ \cup B_R^- \cup S = B \cup S$, and so $B \cup S$ is a pure thick

Suppose c ends in P'. Since c starts in $(T \cup P')^\perp$ and T is a thick parent of P', there are $\tau_0, \tau_1 \in [0,1]$ with $\tau_0 < \tau_1$ such that $c(]\tau_0, \tau_1[) \subseteq T$ and $c([0,\tau_1[) \subseteq P'^\perp$. If $c([0,\tau_1[) \subseteq V^\perp$, then $c([0,\tau_1[) \subseteq (V \cup P')^\perp$, and so $c|_{[0,\tau_1[}$ is a non-trivial subcurve in T—and *a fortiori* in $T \cup (P\backslash P')$—which c has before ever intersecting $V \cup P'$. If instead $c([0,\tau_1[) \not\subseteq V^\perp$, $c^{-1}(V)$ has an infimum x in $[0,\tau_1[$. Since P is a thick parent of V, $c|_{[0,x[}$ contains a non-trivial subcurve in P before ever intersecting V. Since $c([0,x[) \subseteq P'^\perp$, it follows that $c|_{[0,x[}$ contains a non-trivial subcurve in $P\backslash P'$—and *a fortiori* in $T \cup (P\backslash P')$—before ever intersecting $V \cup P'$.

Suppose that c ends in V. Since c starts in $(V \cup P)^\perp$ and P is a thick parent of V, there are $s_0, s_1 \in [0,1]$ with $s_0 < s_1$ such that $c(]s_0, s_1[) \subseteq P$ and $c([0,s_1[) \subseteq V^\perp$. If $c([0,s_1[) \subseteq P'^\perp$, then $c|_{]s_0,s_1[}$ is a non-trivial subcurve in $P\backslash P'$—and *a fortiori* in $T \cup (P\backslash P')$—which c has before ever intersecting $V \cup P'$. If instead $c([0,s_1[) \not\subseteq P'^\perp$, $c^{-1}(P')$ has an infimum y in $[0,s_1[$. Since T is a thick parent of P', $c|_{[0,y[}$ contains a non-trivial subcurve in T before ever intersecting P'. Since $c([0,y[) \subseteq V^\perp$, it follows that $c|_{[0,x[}$ contains a non-trivial subcurve in T—and *a fortiori* in $T \cup (P\backslash P')$—before ever intersecting $V \cup P'$. ■

[71] This follows from the following general lemma (instantiate V with $K^+(R)$, P with $R \cup B_R^- \cup S$, C with B_R^+, and P' with R):

> **Lemma:** Let V be any region and P be a thick parent of V. Let C be a thick child of some $P' \subseteq P$. Then $(P\backslash P') \cup C$ is a thick parent of $V\backslash C$.

Proof of lemma: Let c be a future-directed causal curve starting in $((P\backslash P') \cup C \cup V)^\perp$ and ending in $V\backslash C$. Then c is, in particular, a future-directed causal curve starting in $V^\perp$ and ending in V. Since P is a thick parent of V, there are thus $a, b \in [0,1]$ with $b > a$ such that $c(]a,b[) \subseteq P$ and $c([0,b[) \subseteq V^\perp$. If $c(]a,b[) \subseteq P'^\perp$, then $c(]a,b[)$ is a non-trivial subcurve in $P\backslash P'$—*a fortiori*, in $P\backslash P' \cup C$—which c has before ever intersecting V, and hence before ever intersecting $V\backslash C$. If instead $c(]a,b[) \not\subseteq P'^\perp$, then the infimum x of $c^{-1}(P')$ is in $[0,b[$. Hence there is a $y \in]x,b[$ such that $c(y) \in P'$. But then $c|_{[y,1]}$ is a curve that starts in P' and ends in $V\backslash C \subseteq (P' \cup C)^\perp$. Since C is a thick child of P', $c|_{[y,1]}$ thus contains a non-trivial subcurve in C before ever intersecting $V\backslash C$. But $c([0,y]) \subseteq V^\perp \subseteq (V\backslash C)^\perp$, and so c itself contains a non-trivial subcurve in C before ever intersecting $V\backslash C$. In either case,$(P\backslash P') \cup C$ is a thick parent of $V\backslash C$. ■

parent of $K^+(R)\backslash B$. Finally, since $K^+(R)$ contains all future-directed causal curves starting in $K^+(R)$, $K^+(R)\backslash B$ doesn't cause R. So Parental Markov implies $R \perp\!\!\!\perp (K^+(R)\backslash B)|B \cup S$.

Proof of fact 33: Recall that $R \cup B_R^-$ is a thick parent of B_R^+. Moreover, by Lemma 1, B_R^- is a thick parent of R. It follows that B_R^- is a thick parent of $R \cup (B_R^+)$.[72] Since R is causally convex, $J^+(R)$ fully contains all future-directed causal curves starting in it. Hence $R \cup B_R^+ \subseteq J^+(R)$ doesn't cause $B_R^- \subseteq J^+(R)^\perp$. Hence B_R^- is a pure thick parent of $R \cup B_R^+$. Since S is not caused by $R \cup B_R^+$, Parental Markov entails that $R \cup B_R^+ \perp\!\!\!\perp S|B_R^-$. By Weak Union, $R \perp\!\!\!\perp S|B$. ■

D Marginal Chances over CIRCLE

Abbreviate $\text{urch}_C(\tau_i = j)$ as $[ij]$ and $\text{urch}_C(\tau_i = j|\tau_{i'} = j')$ as $[ij|i'j']$. For every $i = 1, ..., n$ and $j = 1, ..., k$, eqs. 23 and 24 are then more compactly written as follows (where 0 in an index is identified with n):

$$[ij] = \sum_{l=1}^{k} [ij|(i-1)l] \cdot [(i-1)l],$$
$$1 = \sum_{j=1}^{k} [ij].$$

This yields the following $n \cdot (k-1)$ equations, one for every $i = 1, ..., n$ and $j = 1, ..., k-1$:

[72] This follows from the following additional lemma (instantiate P with B_R^-, Q with R, and V with B_R^+):

> **Lemma:** Let Q be causally convex. Let P be a thick parent of Q and $Q \cup P$ be a thick parent of V. Then P is a thick parent of $Q \cup V$.

Proof: Let c be a future-directed causal curve starting in $(Q \cup P \cup V)^\perp$ and ending in $Q \cup V$. c either ends in Q or in V. We show that, in each case, c has a non-trivial subcurve in P before ever intersecting $Q \cup V$. First, suppose c ends in Q. Since P is a thick parent of Q and c starts in $Q^\perp$, c has a non-trivial subcurve in P before ever intersecting Q. Since Q is causally convex, $V \subseteq K^+(Q)$ doesn't cause Q, and so c intersects Q before ever intersecting V. Hence c has a non-trivial subcurve in P before ever intersecting $Q \cup V$; done. Second, suppose that c ends in V. Since $Q \cup P$ is a thick parent of V and c starts in $(Q \cup P \cup V)^\perp$ and ends in V, there are $a, b \in [0,1]$ such that $c(]a,b[) \subseteq Q \cup P$ and $c([0,b[) \subseteq V^\perp$. Suppose $c([0,b[) \subseteq Q^\perp$. Then $c|_{]a,b[}$ is a non-trivial subcurve in P which c has before ever intersecting $Q \cup V$; done. Suppose instead $c([0,b[) \nsubseteq Q^\perp$. Then there is a $q \in [0,b[\cap Q$. Since $c(0) \in Q^\perp$, $c|_{[0,q[}$ is a future-directed causal curve starting in $Q^\perp$ and ending in Q. Since P is a thick parent of Q, $c|_{[0,q[}$ thus has a non-trivial subcurve in P before ever intersecting Q. Since $c([0,q[) \subseteq c([0,b[) \subseteq V^\perp$, c thus has a non-trivial subcurve in P before ever intersecting $Q \cup V$. ■

$$[ij] = \left(\sum_{l=1}^{k-1}[ij|(i-1)l]\cdot[(i-1)l]\right) + [ij|(i-1)k]\cdot\left(1-\sum_{l=1}^{k-1}[(i-1)l]\right)$$
$$= \sum_{l=1}^{k-1}\left([ij|(i-1)l]-[ij|(i-1)k]\right)\cdot[(i-1)l] + [ij|(i-1)k],$$

which can be rearranged to

$$[ij] + \sum_{l=1}^{k-1}\left([ij|(i-1)k]-[ij|(i-1)l]\right)\cdot[(i-1)l] = [ij|(i-1)k].$$

Writing this linear system as a matrix equation yields the following:

$$\mathbf{M}\cdot\hat{\mathbf{p}} = \hat{\mathbf{v}}, \tag{34}$$

where

$$\hat{\mathbf{p}} = ([11], ..., [1(k-1)], [21], ..., [2(k-1], ..., [n1], ..., [n(k-1)])^{\mathrm{T}}$$
$$= ([ij])^{\mathrm{T}}_{i=1,...,n;j=1,...,(k-1)}$$

is a length $n(k-1)$ column vector of marginal probabilities ($\cdot^T$ denotes the transpose),

$$\hat{\mathbf{v}} = ([11|nk], ..., [1(k-1)|nk], [21|1k], ..., [2(k-1)|1k], ..., [n1|(n-1)k], ..., [n(k-1)|(n-1)k])^{\mathrm{T}}$$
$$= ([ij|(i-1)k])^{\mathrm{T}}_{i=1,...,n;j=1,...,(k-1)}$$

is a length $n(k-1)$ column vector, and

$$\mathbf{M} = \begin{pmatrix} \mathbb{I}_{k-1} & 0 & 0 & 0 & \dots & 0 & \mathbf{P}^1_{k-1} \\ \mathbf{P}^2_{k-1} & \mathbb{I}_{k-1} & 0 & 0 & \dots & 0 & 0 \\ 0 & \mathbf{P}^3_{k-1} & \mathbb{I}_{k-1} & 0 & \dots & 0 & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & ... & & \mathbf{P}^n_{k-1} & \mathbb{I}_{k-1} \end{pmatrix},$$

is a $n(k-1)\times n(k-1)$ matrix, where $\mathbb{I}_{k-1}$ is the $(k-1)\times(k-1)$ identity matrix and

$$\mathbf{P}^i_{k-1} = \mathbf{Q}^i_{k-1} - \mathbf{R}^i_{k-1}$$

is the $(k-1) \times (k-1)$ matrix such that

$$\mathbf{Q}^i_{k-1} = \begin{pmatrix} [i1|(i-1)k] & [i1|(i-1)k] & \dots & [i1|(i-1)k] \\ [i2|(i-1)k] & [i2|(i-1)k] & \dots & [i2|(i-1)k] \\ \vdots & & & \\ [i(k-1)|(i-1)k] & [i(k-1)|(i-1)k] & \dots & [i(k-1)|(i-1)k] \end{pmatrix},$$

and

$$\mathbf{R}^i_{k-1} = \begin{pmatrix} [i1|(i-1)1] & [i1|(i-1)2] & \dots & [i1|(i-1)(k-1)] \\ [i2|(i-1)1] & [i2|(i-1)2] & \dots & [i2|(i-1)(k-1)] \\ \vdots & & & \\ [i(k-1)|(i-1)1] & [i(k-1)|(i-1)2] & \dots & [i(k-1)|(i-1)(k-1)] \end{pmatrix}.$$

Both the matrix $\mathbf{M}$ and the enriched matrix $(\mathbf{M}|\hat{\mathbf{v}})$ generically have full rank $n(k-1)$, and so generically $\hat{\mathbf{p}}$ is unique.

To illustrate this further, consider CIRCLE. Because every color j only has itself and color $j+1$ as permissible successors, all entries of $\mathbf{Q}^i_{k-1}$ besides the first row are 0, and all entries of $\mathbf{R}^i_{k-1}$ besides the diagonal and the first lower diagonal are 0. In the simplest non-trivial case, $n = k = 2$, the loop is two days long, with two possible colors per day. In this case—call it SUPER SIMPLE CIRCLE, or SSC—we have

$$\mathbf{M} = \begin{pmatrix} 1 & [11|22] - [11|21] \\ [21|12] - [21|11] & 1 \end{pmatrix}$$

and

$$\hat{\mathbf{v}} = ([11|22], [21|12])^{\mathrm{T}}.$$

Note that $\mathbf{M}$ and $(\mathbf{M}|\hat{\mathbf{v}})$ both have rank 2 unless

$$[21|12] - [21|11] = [11|22] - [11|21] = \pm 1,$$

i.e. unless either[73]

$$\begin{aligned} \mathrm{urch}_{SSC}(\tau_2 = 1|\tau_1 = 2) &= \mathrm{urch}_{SSC}(\tau_1 = 1|\tau_2 = 2) = 1, \\ \mathrm{urch}_{SSC}(\tau_2 = 1|\tau_1 = 1) &= \mathrm{urch}_{SSC}(\tau_1 = 1|\tau_2 = 1) = 0, \end{aligned}$$

[73] We always assume that transition probabilities are precise.

or

$$\text{urch}_{SSC}(\tau_2 = 1|\tau_1 = 2) = \text{urch}_{SSC}(\tau_1 = 1|\tau_2 = 2) = 0,$$
$$\text{urch}_{SSC}(\tau_2 = 1|\tau_1 = 1) = \text{urch}_{SSC}(\tau_1 = 1|\tau_2 = 1) = 1.$$

In the first case, the particle is guaranteed to switch color every time. Any probabilistically coherent assignment of marginals respecting $\text{urch}_{SSC}(\tau_1 = 1) = \text{urch}_{SSC}(\tau_2 = 2)$ is a solution to the equations. In the second case, the particle is guaranteed to retain its color every time. Here, any probabilistically coherent assignment of marginals respecting $\text{urch}_{SSC}(\tau_1 = 1) = \text{urch}_{SSC}(\tau_2 = 1)$ is a solution to the resulting equations. These are the only two possible cases for SSC in which the dynamics fails to determine unique marginal chance distributions over the states of the loop.

References

Albert, David Z. (2000). *Time and Chance*. Harvard University Press.

Arntzenius, Frank and Ned Hall (2003). "On What We Know About Chance". In: *The British Journal for the Philosophy of Science* 54.2, pp. 171–179.

Berkovitz, Joseph (2001). "On Chance in Causal Loops". In: *Mind* 110.437, pp. 1–23.

Briggs, Rachael (2010). "The Metaphysics of Chance". In: *Philosophy Compass* 5.11, pp. 938–952.

Carter, Brandon (1968). "Global Structure of the Kerr Family of Gravitational Fields". In: *Physical Review* 174.5, pp. 1559–1571.

Cusbert, John (2018). "Backwards Causation and the Chancy Past". In: *Mind* 127.505, pp. 1–33.

— (2022). "How Stable Is Objective Chance?" In: *The British Journal for the Philosophy of Science* 73.3.

Earman, John (1995). "Recent Work on Time Travel". In: *Time's Arrows Today: Recent Physical and Philosophical Work on the Direction of Time*. Ed. by Steven F. Savitt. Cambridge: Cambridge University Press, pp. 268–310.

Echeverria, F., G. Klinkhammer, and K. S. Thorne (1991). "Billiard Balls in Wormhole Spacetimes with Closed Timelike Curves: Classical Theory". In: *Physical Review, D (Particles Fields); (United States)* 44:4.

Effingham, Nikk (2020). *Time Travel: Probability and Impossibility*. New York: Oxford University Press.

Ghirardi, GianCarlo, Alberto Rimini, and Tullio Weber (1986). "Unified Dynamics for Microscopic and Macroscopic Systems". In: *Physical Review D* 34.D, pp. 470–491.

Gödel, Kurt (1949). "An Example of a New Type of Cosmological Solutions of Einstein's Field Equations of Gravitation". In: *Reviews of Modern Physics* 21.3, pp. 447–450.

Hájek, Alan (2003). "What Conditional Probability Could Not Be". In: *Synthese* 137.3, pp. 273–323.

Hall, Ned (2004). "Two Mistakes About Credence and Chance". In: *Australasian Journal of Philosophy* 82.1, pp. 93–111.

Handfield, Toby and Alastair Wilson (2014). "Chance and Context". In: *Chance and Temporal Asymmetry*. Ed. by Alastair Wilson. Oxford University Press.

Hitchcock, Christopher and Miklós Rédei (2021). "Reichenbach's Common Cause Principle". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2021. Metaphysics Research Lab, Stanford University.

Horacek, David (2005). "Time Travel in Indeterministic Worlds". In: *The Monist* 88.3, pp. 423–436.

Joyce, James M. (2010). "A Defense of Imprecise Credences in Inference and Decision Making". In: *Philosophical Perspectives* 24.1, pp. 281–323.

Lewis, David (1987). "A Subjectivist's Guide to Objective Chance". In: *Philosophical Papers, Volume II*. Oxford University Press.

Loewer, Barry (Aug. 2019). "The Mentaculus Vision". In: *Statistical Mechanics and Scientific Explanation*. World Scientific Publishing Co.

Meacham, Christopher J. G. (2005). "Three Proposals Regarding a Theory of Chance". In: *Philosophical Perspectives* 19.1, pp. 281–307.

Mellor, D. H. (1995). *The Facts of Causation*. Routledge.

Nelson, Kevin (2009). "On Background: Using Two-Argument Chance". In: *Synthese* 166.1, pp. 165–186.

Pearl, Judea (1985). "Graphoids: A Graph-Based Logic For Reasoning About Relevance Relations".

Popper, Karl R. (1959). "The Propensity Interpretation of Probability". In: *The British Journal for the Philosophy of Science* 10.37, pp. 25–42.

Popper, Karl ([1959] 1968). *The Logic of Scientific Discovery*. Routledge.

Reichenbach, Hans (1956). *The Direction of Time*. Ed. by Maria Reichenbach. Mineola, N.Y.: Dover Publications.

Russell, Bertrand (1903). *Principles of Mathematics*. New York: Routledge.

Schaffer, Jonathan (2003). "Principled Chances". In: *British Journal for the Philosophy of Science* 54.1, pp. 27–41.

Schaffer, Jonathan (2007). “Deterministic Chance?” In: *British Journal for the Philosophy of Science* 58.2, pp. 113–140.
van Fraassen, Bas C. (1984). “Belief and the Will”. In: *Journal of Philosophy* 81.5, pp. 235–256.