Chances on Loops

forthcoming in *Noûs*

Jens Jäger
New York University

Note: This paper includes an extended technical appendix, for those interested. It may be skipped without losing the paper's main thread.

Contents

1 Introduction

A star pharmacist, 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: this failure is what causes the second antidote's success, which causes grandpa's survival, which causes your being born, which causes your administering the first antidote. So the first antidote's failure is already past. On the other hand, its failure is also still *future*: some time passes between the first antidote's administration and its failing to take effect, 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?

To the contrary, I'll argue that, on a salient interpretation of "as you administered the first antidote", the answer is 2/3. Note that this is the chance in an analogous time-travel-free situation, of the first antidote's working, conditional on *at least one* antidote's working. We'll arrive at this answer via a systematic account of chances on loops, using two general principles about chance.

The essay's central question is then this: *Given* a time travel structure, what are the chances? This question is distinct from the following (also interesting!) question: what could the chances of wormholes and other time-travel structures be, in the first place? This latter question, I suspect, falls primarily into the jurisdiction of those physicists who study actual or hypothetical physical mechanisms for such structures. As I hope to show here, however, philosophy *can* productively contribute to *my* question.[1]

Why care about what the chances are in time travel situations? The question is intrinsically interesting, I think. Moreover, where time travel gives rise to *causal loops*—situations where a causal influence travels back in time and interacts with its source, like in our grandfather case above—it has been of interest to physicists too (see Earman (1995)), who

[1] One limitation: my essay deals specifically with static spacetime structures (i.e., static metric fields), as found e.g. in special relativity. A generalization to theories with dynamical spacetime structure (such as general relativity) will have to wait for another day. Still, I hope that the current proposal constitutes a significant step in that direction.

have studied several causal loop worlds in detail (e.g. Stockum 1938; Gödel 1949; Carter 1968). Part of a philosopher of science's responsibility is to study what physicists take seriously. Moreover, the study pays off philosophically. Provided causal loops are at least metaphysically possible, it'll challenge the orthodox view which ties chance fundamentally to time, and the doctrine according to which chance is necessarily "stable", or invariant, across duplicate experiments. We'll replace these dogmas with a view on which chances are tied, not to temporal histories, but to chance setups, and on which chances on loops are unstable, in systematic, scrutable ways.

Additionally, it bears noting that our final framework isn't limited to particular setups, like the grandfather case. Instead, it can answer almost any query about chances on loop, for any setup. Among other things, it corrects a misconception I have encountered regarding wormholes: that there are no principled ways to assign precise prior chances to what emerges from future wormholes. I show that generically—namely in the presence of non-trivial chances—this is false. In other words, knowing the state of the world, and knowing the chance laws, generically you ought to have a precise expectation about what will emerge from a future wormhole. Wormholes are far from being the black boxes they are sometimes thought to be.

Core to my account is a proposal about what screens off what, with respect to chance. "Screening off" is a familiar idea: in assessing the probability of some proposition C, A screens off B with respect to C if A *overrides* any information B provides about C. Suppose I'd like to know if I'm a carrier of a certain genetic marker. Given complete information about my *parents'* genes, no information about my *grandparents'* genes can affect the probability of whether I carry the marker—complete information about the former *screens off* any information about the latter. With respect to chance, what screens off what is partially determined by spacetime structure. Specifically, I defend the idea that (given local laws) what's happening at a spacetime region's *boundary* screens off what's happening on the region's inside from what's happening on its outside. This connection will enable us to systematically derive chances on loops from the ordinary, time-travel-free chances.

The essay is structured as follows. Section 2 introduces a simple, stripped-down causal 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 "stability", or invariance, idea. The first account often trivializes chances on causal loops, and the second account is inconsistent. Both accounts should 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 explains how these two principles generate the chances in

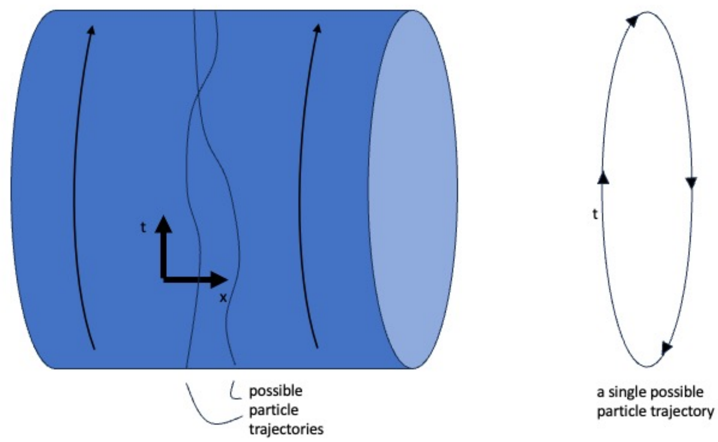


Figure 1: A pictorial representation of a two-dimensional world with circular time (left), and a one-dimensional world with circular time, CIRCLE (right)

the simple loop scenario. Section 9 generalizes the approach and revisits the stochastic grandfather paradox. Section 10 shows how dynamical laws alone can fix unconditional chances on cyclic spacetimes and moreover determine precise expectations about what emerges from future wormholes. Section 11 concludes.

2 A Simple Case

In our stripped-down case, time is a circle. Imagine a flat piece of paper, on which you've drawn two perpendicular arrows, one labeled "time", and the other labeled "space". This paper is a crude representation of a universe with one temporal and one spatial dimension. Any material body existing in that universe would trace out a line across the paper. Now roll up the paper, along the "time"-direction, into a cylinder. This provides a crude model of a universe with circular time.[2]

For our purposes, we'll go simpler still: instead of a sheet of paper, imagine a paper string cut along the direction labeled "time". This string represents a universe with one temporal and *zero* spatial dimensions. Its "spatial part" is a single point, existing at all times.[3] Glue the the string's ends together, and we'll have a universe with zero spatial dimensions and circular time. Call this world CIRCLE (or C for short). See fig. 1.[4]

[2]Mathematically, we can represent such a universe by a (two-dimensional, connected, oriented) flat and time-like closed Lorentzian manifold.

[3]Whether it *persists* or *endure*s won't matter here.

[4]Mathematically, we can represent CIRCLE by a one-dimensional, oriented, connected, closed Lorentzian (or, equivalently in the 1D case, Riemannian) manifold. Note that this manifold is illustrative not only because

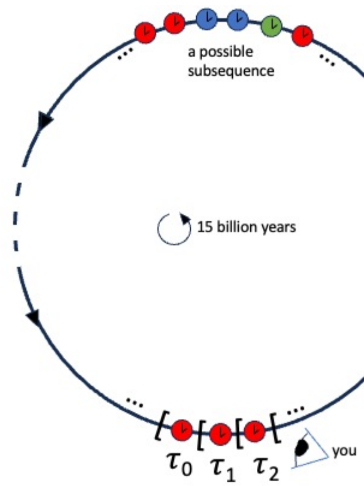


Figure 2: A sketch of the situation in CIRCLE.

Suppose CIRCLE is 15 billion years round-trip and populated by a single particle, existing at all times. The particle has two intrinsic magnitudes, *color* and *clock*. Both magnitudes take values in the real numbers, but for the color magnitude we'll choose familiar color names. The value of the clock magnitude is given in units of time (seconds, hours, etc). After learning about CIRCLE's geometry, you decide to closely follow the particle for a few years. (Imagine you are God, looking down on CIRCLE.) Throughout this period, the clock grows in proportion to the spacetime distance traversed by the particle, until it reaches 24 hours, at which point it restarts from 0. The particle occasionally changes colors, but only exactly at clock restart points. You only ever observe three transitions: *red to green*, *green to blue*, and *blue to red*. (See fig. 2.) Over your observation period, you observe a color change at about 2 out of 10 reset points.

When evaluating the chance of an impending color switch, two options may jump out, paralleling those about the stochastic grandfather paradox. One might think that the chance that it'll change color at the next restart point is 1 if it actually changes color, and 0 if it doesn't. After all, whatever the particle does, it has *already* happened—it is in our causal past. But then again, whatever happens is *still* to happen—it is in our causal future. So perhaps it is still open, and hence has a chance strictly between 0 and 1.

I'll now survey two proposals which capture these two ideas. We'll see that both should be rejected.

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 most interpretations of quantum mechanics.

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.

With this in mind, Lewis (1987) makes chance time-relative:

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.[6,7]

Clause 2 causes trouble. In CIRCLE, every time precedes every other time. Given this, clause 2 trivializes chance in CIRCLE: it implies that every chance function assigns chance 1 to the particle's actual complete color history, and 0 to every other 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.

[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. But nothing of significance will hang on this.

[6] To accommodate deterministic chance, weaken this to "the strongest truth entirely about *macroscopic* matters of particular fact at times at or before t" (cf. Loewer (2001)).

[7] 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" with the intention of distinguishing 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.

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

$\therefore$ 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 deny either (a) or (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 is extremely restrictive: having observed the particle change color from red to green, it'd be false to say that "the particle has *previously* been red". This reply would also 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 chance in CIRCLE is problematic for two connected reasons.[8] 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 in CIRCLE as it does in any *linear* world—say, in the one-dimensional spacetime we have before "gluing" the string's ends together. Clause 2 says that, in CIRCLE, this practice is always erroneous.

Secondly, a universe with circular time is still compatible with what we know.[9] If this speculative scenario was true, would that mean that all of our scientific theories involving chance are radically wrong? Would it mean, for example, that radioactive decay wasn't

[8]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 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.

[9]Specifically, it is compatible with what we know 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 is played by the posit that there is a low-entropy macrostate *at some time* (with entropy increasing bidirectionally from there).

stochastic after all? The answer is pretty clearly no. But this contradicts temporalism's clause 2. So we should reject temporalism.[10]

4 Urchance

Having rejected temporalism, let's consider what formalism should replace it.

A certain rendition of temporalism lends itself to a natural generalization. To avoid conflict with chance deference principles like the Principal Principle, temporalism's chance functions must "cohere" with each other: those at later times must be obtained from those at earlier times by conditioning on intervening history.[11] Hall (2004) observes that coherence follows naturally on the assumption that there is a single "urchance" function such that, for any time t, the chance function at t is the result of conditioning the urchance function on the history up to t.

Even for some non-history propositions the result will be well-defined: where H_t is the world's history up to t, and urch the urchance function, $\text{urch}(\cdot|H_t)$ will generically assign positive probability to some proposition A such that $H_t A$ is not a history—e.g., where A is an incomplete description of t's future. But then, by the ratio formula, $\text{urch}(\cdot|H_t A) = \text{urch}(\cdot \wedge A|H_t)/\text{urch}(A|H_t)$, where the quotient on the right-hand side is a well-defined probability function.

This is naturally extended further. Let us conceive of urchance as encoding exactly the content of the dynamical laws. But the dynamical laws go beyond what's definable from functions like $\text{urch}(\cdot|H_t)$ and the ratio formula alone. They can produce well-defined (and non-trivial) chance values from propositions that don't even entail any history proposition: consider a bounded temporal interval (i.e. a region between any two non-intersecting time-slices[12]): the conjunction of its state and the dynamical laws entails a well-defined chance distribution over possible futures.

Now, $\text{urch}(A|B)$ isn't well-defined for all A and B. Some B will simply be too weak: propositions about contingent matters of fact won't generally have a well-defined probability conditional on a logical tautology, for example. Officially, I'll therefore merely assume

[10] Cusbert (2018; 2022) suggests replacing temporalism with the "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 (2018) formalism replaces "times" with "globally connected sets" of events—this difference doesn't matter here.) This also trivializes chance in CIRCLE, and hence fares just as badly as temporalism.

[11] To see that the Principal Principle requires this: the Principle 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, we have $Cr_0(A|H_t T) = ch_t(A)$. Where $H_{[t,t^*]}$ is the intervening history between t and t^*, it follows that $ch_{t^*}(A) = Cr_0(A|H_{t^*}T) = Cr_0(A|H_{[t,t^*]}H_t T) = ch_t(A|H_{[t,t^*]})$.

[12] E.g. between two Cauchy surfaces from the same foliation.

that urch is a two-place, *partial*, and *primitively conditional* probability function (cf. Hájek (2003)).[13] Throughout I'll write urch$(\cdot|\cdot)$ instead of urch$(\cdot,\cdot)$. For any proposition B, I'll call B the "background proposition" for the function urch$(\cdot|\cdot \wedge B)$.[14]

It's easy to see the urchance formalism's promise: admitting propositions weaker than any history as background propositions opens a path out of the trivialization problem. In CIRCLE, the temporal history of any time fixes every contingent fact about the world. By contrast, plenty of weaker propositions leave many contingencies open. The second account uses this to develop a non-trivial chance theory of CIRCLE—unsuccessfully, as we'll see.

5 Against Invariance

Consider again the string of paper before its ends are glued together—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 a *0.2 chance* that the particle changes color to the next allowed one (i.e., 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" spacetimes. Suppose for a moment that CIRCLE and LINE were part of the same world (connected, if you wish, through a space-like line). In this case, the idea that the chances are unchanged follows from the thought that objective chances are, as Schaffer (2003) puts it, "stable". Arntzenius and Hall (2003, p. 178) express the same thought 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". Similarly, Schaffer (2007, p. 125) requires that "chance values should remain constant across intrinsically duplicate trials" within the same world. These principles are formulated to only cover chance assignments within

[13]Popper ([1959] 1968, App. *IV and *V) offers a convenient axiomatization of primitively conditional probability functions, which I generalize to the partial case in Chapter 2. This mild generalization is what I mean throughout by "the laws of probability". For our purposes, it's enough to note that those laws entail the following slight generalization of the multiplicative law: $p(AB|C) = p(A|BC) \cdot p(B|C)$, whenever any two of $p(AB|C)$, $p(A|BC)$, and $p(B|C)$ are defined.

[14]The terminology of "background proposition" is also used in Nelson (2009) and Cusbert (2018). Authors who employ similarly flexible chance formalisms to mine include Meacham (2005), Nelson (2009), Briggs (2010), Handfield and Wilson (2014), and Cusbert (2018).

[15]Mathematically, we can represent the world by a one-dimensional, oriented, simply-connected Lorentzian (or, equivalently in the 1D case, Riemannian) manifold.

the *same* world. But they have obvious and natural generalizations that also cover chance assignments across *nomically compatible* worlds. These generalizations then dictate the same transition chances for CIRCLE as for LINE, even when the two spacetimes aren't world mates (which is what we'll continue to assume in the following).

Arntzenius and Hall don't provide a formally precise version of the principle, and Schaffer's (2007) presentation assumes temporalism.[16] So let's formulate a principle ourselves, for the case of CIRCLE and LINE, using urchance. Informally, the idea is that the chances an interval in LINE generates over the states of an interval some distance into its future equal the chances that, in CIRCLE, an intrinsic duplicate of the former interval would generate over the states of an intrinsic duplicate of the latter located the same distance into the future.

More carefully, where t and s are times, define the *forward distance from t to s* as the smallest distance (i.e., duration) traversed by a future-directed curve starting in t and ending in s.[17] In CIRCLE, any pair of times has a forward distance. 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 fig. 2, the forward distance from τ_0 to τ_2 is 2 days, and the forward distance from τ_2 to τ_0 is 15 billion years minus 2 days. 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^*. So, for example, (τ_0, τ_2) is temporally congruous with (τ_5, τ_7).

Since CIRCLE and LINE have the same laws, they share the same urchance function. Denote by $\mathrm{urch}_C(\cdot|\cdot)$ and $\mathrm{urch}_L(\cdot|\cdot)$ the results of conditioning that function on a complete description of CIRCLE's and LINE's spacetime geometry, respectively.[19] I prefer the more perspicuous label "invariance" over "stability".

Hypothesis. **Chance Invariance:** Let (I, J) and (I^*, J^*) be temporally congruous

[16] Horacek (2005, p. 428) and Effingham (2020, p. 152) each propose similar principles which also presuppose temporalism. Cusbert's (2022) rendition presupposes temporalism's first clause (that chance functions are necessarily time-indexed), and it also weakens stability to such a degree that trivialized chances count as "stable" relative to the ordinary, non-trivial chances—not the alternative to trivialization we're looking for here.

[17] 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 range of a curve. Whenever we deal with bounded curves, we'll set $I = [0, 1]$ without loss of generality—the length of a curve's tangent vectors will never matter to us.

[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, let the forward distance from I to J be ill-defined.

[19] We'll always understand those descriptions to include a "that's all" clause: "the foregoing description entails all true geometric relationships"—this will be relevant later on.

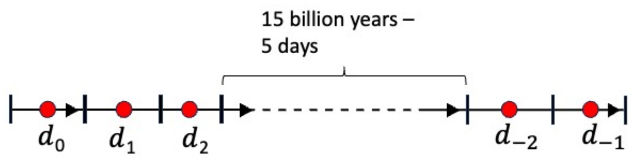


Figure 3: LINE

pairs of intervals in CIRCLE and LINE, respectively. Then, where Q and P are any qualitative intrinsic properties,

$$\text{urch}_C(P(J)|Q(I)) = \text{urch}_L(P(J^*)|Q(I^*)),$$

whenever at least one side is defined.

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 $P(J^*)$—irrespective of the global topology. To illustrate, 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.[20] From the dynamics in LINE, $\text{urch}_L(\text{RED}(d_1)|\text{RED}(d_0)) = 0.8$. 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, for all its initial attractiveness, Chance Invariance is inconsistent: it defines some probabilities multiple times over, with conflicting results. Essentially, what goes wrong is that each day in CIRCLE has two distinct forward distance relationships to any other given day: one because it comes before the latter, and another because the latter comes before it. Depending on what order we consider, the resulting pair is temporally congruous to very different pairs in LINE. We can exploit this fact to derive a contradiction from Chance Invariance.

For concreteness, 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 at both the previous day, τ_0, and the next day. Label the days in LINE as in figure 3. To calculate this chance with Chance Invariance, we can use the pairs (τ_0, τ_1) and (τ_1, τ_2), congruous with (d_0, d_1) and (d_1, d_2), respectively. This yields the following identity:[21]

[20] Note that the forward distances from τ_0 to τ_1 and from d_0 to d_1 (each 24 hours) are entailed by the description of CIRCLE's and LINE's geometries, respectively.

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

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

But we can *also* use the pair (τ_0, τ_1), congruous with (d_{-2}, d_{-1}),[22] together with the "inverted" pair (τ_2, τ_1), congruous with (d_0, d_{-1}). This yields the following identity:[23]

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

Since (τ_0, τ_1) and (τ_1, τ_2) are temporally congruous with (d_0, d_1) and (d_1, d_2), respectively, it follows that (τ_0, τ_2) and (d_0, d_2) are temporally congruous and that $(\tau_0, \tau_1 \cup \tau_2)$ and $(d_0, d_1 \cup d_2)$ are temporally 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

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

Plugging eqs. 2 and 3 into eq. 1 yields

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

But, from the dynamics in LINE, $\text{urch}_L(\text{RED}(d_2)|\text{RED}(d_0)) = 0.8 \cdot 0.8 = 0.64 > 0$, and so, from eq. 4,

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

where the second line follows by the multiplicative axiom. ■

[22] Note that (τ_0, τ_1) is congruous both with (d_0, d_1) and with (d_{-2}, d_{-1}).

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

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

Since (τ_2, τ_1) and (τ_0, τ_1) are temporally congruous with (d_0, d_{-1}) and (d_{-2}, d_{-1}), respectively, it follows both that (τ_2, τ_0) is temporally congruous with (d_0, d_{-2}) and that $(\tau_2, \tau_0 \cup \tau_1)$ is temporally congruous with $(d_0, d_{-2} \cup d_{-1})$. So, by Chance Invariance,

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

and

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

Plugging eqs. 7 and 8 into eq. 6 yields

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

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

Eqs. 5 and 10 jointly require

$$\mathrm{urch}_L(\mathrm{GREEN}(d_1)|\mathrm{RED}(d_2) \wedge \mathrm{RED}(d_0)) = \mathrm{urch}_L(\mathrm{GREEN}(d_{-1})|\mathrm{RED}(d_0) \wedge \mathrm{RED}(d_{-2})). \tag{11}$$

Because RED-GREEN-RED transitions are nomically disallowed in LINE, we have

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

But the dynamics in LINE also entail the following:[24]

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

So $\mathrm{urch}_L(\mathrm{GREEN}(d_1)|\mathrm{RED}(d_2) \wedge \mathrm{RED}(d_0)) \neq \mathrm{urch}_L(\mathrm{GREEN}(d_{-1})|\mathrm{RED}(d_0) \wedge \mathrm{RED}(d_{-2}))$, in contradiction with eq. 11.

6 Vestiges of Invariance

Chance Invariance fails: chances in loop worlds generally differ from those in loop-free worlds, even if they have the same chance laws. But note that invariance *among* loop-free

But $\mathrm{urch}_L(\mathrm{RED}(d_{-2})|\mathrm{RED}(d_0)) \approx 1/3 > 0$, 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 $\mathrm{RED}(d_{-2}) \wedge \mathrm{RED}(d_0)$ by the probability-weighted sum of color trajectories compatible with $\mathrm{RED}(d_0)$, leading to a formula similar to the expression in fn. 40.) So, from eq. 9,

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

where the second line follows from the multiplicative axiom. ■

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

$$\mathrm{urch}_L(\mathrm{GREEN}(d_{-1})|\mathrm{RED}(d_0) \wedge \mathrm{RED}(d_{-2})) = \mathrm{urch}_L(\mathrm{GREEN}(d_{-1})|\mathrm{RED}(d_{-2})). \tag{12}$$

In section 7 I'll discuss "screening off" much more—one principle for it (Parental Markov) is true in acyclic worlds and yields the first equality in eq. 12 as a special case. From the dynamics for LINE we moreover have

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

worlds remains unscathed. Plausibly, two intrinsically duplicate situations in two different nomically compatible *loop-free* worlds should generate the same chance distributions. Let's call this weaker principle *Acyclic Chance Invariance*. The following two paragraphs spell out a fairly general and precise version of the principle. Those who'd like to avoid formalism may skip the second of those paragraphs, continuing with the informal formulation just provided.

The following concepts are important as they'll reappear throughout the rest of the paper. A *causal curve* is any trajectory through spacetime which a material particle could take. (More exactly, a *causal curve* is any differentiable curve with only light-like (null) or time-like tangent vectors.) Curves have a direction, and causal curves can be either future-directed or past-directed. A *closed* causal curve is any causal curve which loops back in on itself—i.e., any non-trivial[25] causal curve which ends at its starting point. Next, a region R's *causal past*, $J^-(R)$, is that part of spacetime, including R, from which a material particle can eventually reach R. Its causal future, $J^+(R)$ is that part of spacetime, including R, a material particle can eventually reach *from* the region. More exactly, $J^+(R)$ is the union of R with all points p such that there is a future-directed causal curve starting in R and ending in p. $J^-(R)$ is defined analogously, with "future-directed causal curve" replaced by "past-directed causal curve". Note that $R \subseteq J^+(R)$ and $R \subseteq J^-(R)$.[26] Finally, R's *proper causal future* $K^+(R)$ is the difference between the causal future and R, i.e. $K^+(R) = J^+(R) \backslash R$. The *proper causal past* $K^-(R)$ is defined analogously, $K^-(R) = J^-(R) \backslash R$.

We'll formulate Acyclic Chance Invariance for pairs (R_1, R_2) of disjoint regions where $R_2 \subseteq K^+(R_1)$ and $R_1 \cap K^+(R_2) = \emptyset$—i.e., where R_2 is "strictly to the future" of R_1. If this is the case, I'll also write $R_1 < R_2$. For any R_1 and R_2 with $R_1 < R_2$, let the *region between R_1 and R_2* be the intersection of R_1's causal future and R_2's causal past, $J^+(R_1) \cap J^-(R_2)$.

> ***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 exists 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, where Q and P are any qualitative intrinsic properties,
>
> $$\mathrm{urch}_{\mathcal{M}}(P(R_2)|Q(R_1)) = \mathrm{urch}_{\mathcal{M}'}(P(R_2')|Q(R_1'))$$

[25] A curve $c : I \to \mathcal{M}$ is *non-trivial* iff its range, $c(I)$, consists of at least two (distinct) points. (For continuous curves in continuous manifolds, this is equivalent to $c(I)$ having continuum many points.)

[26] Not everyone defines causal past and future that way, but this choice will be convenient for our purposes.

[27] Here, of course, $J^+(R_1) \cap J^-(R_2)$ is the region between R_1 and R_2 *in* $\mathcal{M}$; and $J^+(R_1') \cap J^-(R_2')$ is the region between R_1' and R_2' *in* $\mathcal{M}'$. If the given isometry exists, you might call (R_1, R_2) and (R_1', R_2') *spatiotemporally congruous*, generalizing the concept of *temporal congruity* from Section 5.

if at least one side is defined.

This is the first vestige of chance invariance, covering acyclic chances.

As for the *cyclic* chances: they still shouldn't just be arbitrary, floating freely of the acyclic chances. Instead, they should be *derivable* from them in a principled way. This is the second vestige of chance invariance. To sharpen it up, suppose you're given a set of possible dynamical laws, defining an urchance function urch. The laws' *acyclic chances* are the set of all propositions entailed by the dynamical laws and of the form $\text{urch}_{\mathcal{K}}(Q(R)|P(S)) = x$, where $\mathcal{K}$ is any spacetime without closed causal curves, R and S are regions in $\mathcal{K}$, and Q and P are physically possible intrinsic properties of R and S, respectively, and x is a real number. (As always, $\text{urch}_{\mathcal{K}}$ denotes the result of conditioning urch on a complete description of $\mathcal{K}$'s geometry.) The laws' *cyclic chances* are defined in exactly the same way, except that $\mathcal{K}$ now ranges over all spacetimes *with* closed causal curves. The thought is now that the laws' *cyclic chances* can be inferred from their *acyclic chances* in a principled way. If so, say that the cyclic chances are "*dynamically scrutable*". Let *Dynamic Scrutability* be the claim that, necessarily, cyclic chances are dynamically scrutable. Acyclic Chance Invariance and Dynamic Scrutability are the core of 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* way of assigning chances on causal loops that (3) *avoids trivialization*. Chance Invariance satisfies (2) and (3), but not (1). A consistent but arbitrary way of assigning cyclic chances would satisfy (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 What's Inside Doesn't Matter

Dynamic Scrutability follows from a particular aspect about the relationship between the dynamical laws (urchance) and spacetime structure. In a slogan: dynamical laws are *superficial*—for the purpose of predicting what goes on outside of a region, only events at its boundary 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. Another example, from physics: once we know a radioactive atom's current state, anything else that happens further upstream, e.g. how long the atom's been in its current state, has no additional impact on the chance of its decaying within the next 10 seconds. Similarly, provided there's no action at distance,[28] what happens at space-like separation also doesn't matter. In short, the atom's current state screens off everything that's not one of its effects. Generalizing this yields the following general proposal: what's in a region's immediate proper causal past screens it off from anything outside of the region's causal future.

This is a spacetime-theoretic version of the "(Parental) Causal Markov Condition". Generically, the condition states that an event's immediate causes screen it off from any of its non-effects (cf. Hitchcock and Rédei 2021). The (Parental) 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 (Halpern and Pearl 2005). But how might they—and more generally, a theory of screening off—bear on what the chances on loops are? The idea is this: for some regions we can screen off *whether they're part of a world with causal loops*. This allows us to relate cyclic chance functions to *acyclic* chance functions.

7.1 Parental Markov

Let's state the spacetime version of the parental Markov condition more precisely.

First, a region S "screens off" region T from R iff conditional on the complete state of S, the urchance function judges information about T as irrelevant to the state of R. As before, where $\mathcal{M}$ is a spacetime, $\text{urch}_{\mathcal{M}}$ denotes the result of conditioning the urchance function on a complete description of $\mathcal{M}$'s geometry.

> ***Def.* Screening Off:** For any spacetime $\mathcal{M}$, and any regions R, S, T in $\mathcal{M}$, region *S screens off R and T* (in $\mathcal{M}$) 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 qualitative intrinsic property Q_2 and qualitative intrinsic properties Q_1 and Q_3, such that at least one side of the equation is well-defined

[28] Throughout the essay I confine 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 relativistic GRW—a stochastic non-local theory of quantum mechanics—one can maintain Parental Markov (see below) relative to Minkowskian spacetime structure by limiting oneself to regions that are inextendible space-like surfaces. Other non-local accounts, e.g. ones involving superluminal particle, may altogether dispense with Parental Markov relative to *Minkowskian* structure. But such accounts should substitute an alternative non-Minkowskian spacetime structure, and thus a different notion of "causal". For example, on a theory of superluminal particles, some space-like curves may count as causal. My account would then apply to this alternative spacetime structure.

and $Q_2(S) \wedge Q_3(T)$ is possible according to $\text{urch}_{\mathcal{M}}$.[29]

It follows from the probability laws that "screening off" is symmetric in the last two arguments, R and T. Where the spacetime is clear from context, I'll also speak of regions' screening off *simpliciter*.

Second, 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 all of its thick parents. Any thick parent thus contains the region's "immediate proper causal past".

Why "thick" parent? One reason is that, in the following, we'll often talk about *maximal* intrinsic states of regions. In the desired sense of "maximal", this includes a specification of *velocities* (and potentially other time derivatives)—those are required as inputs to dynamical laws expressed by second- (or higher-) order differential equations. But many, including myself, are inclined to endorse a reductivist view about time derivatives, such as the "at-at" theory of velocities. 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.

Throughout the rest of the essay, when I say "cause" I mean the Lorentzian notion: a region A is a *cause* of 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 version of parental Markov is now this:

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

But Unrestricted Parental Markov is generically false—even in worlds with no closed causal curves. For some regions *still* cause their own parents (including their own immedi-

[29]On "maximal" and "possible": An intrinsic property of a spacetime region is *maximal* iff it comes with a "that's all" clause: where Q is a maximal intrinsic property, $Q(R)$ entails that every particular matter of fact intrinsic to R is entailed by $Q(R)$. In our Popperian formalism, where u is a partial primitively conditional probability function, a proposition A is *impossible according to* u iff, for all B, $\text{u}(B|A) = 1$. This explains the proviso that $Q_2(S) \wedge Q_3(T)$ be *possible* according to $\text{urch}_{\mathcal{M}}$: since for any subset $S' \subseteq S$ we can generically choose an intrinsic property $Q_{2'}$ such that $Q_{2'}(S') \wedge Q_2(S)$ is impossible according to u—simply let $Q_{2'}$ describe a nomic impossibility—this would have the non-sensical consequence that S generically wouldn't even screen off any of its subsets from R. That's not a useful notion of "screening off".

[30]A *subcurve* of c is a restriction of c to a subinterval of I.

[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 $c[I_1] \subseteq (R \cup P)^{\perp}$, and $c[I_2] \subseteq P$ is a non-trivial subcurve, and $c[I_3]$ intersects R.

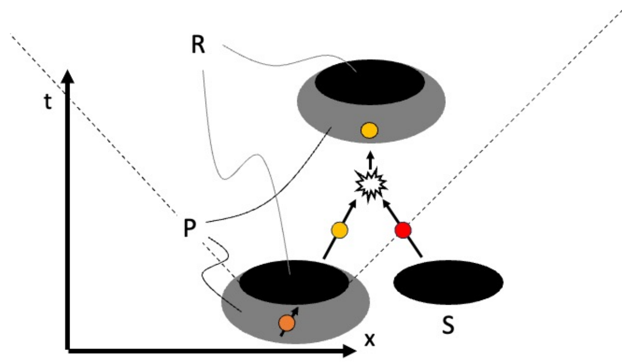


Figure 4:

ate proper causal past). Consider the discontiguous region R (the union of the two black ovals) in fig. 4. Conditioning on one of its parents, P (the union of the two grey regions, disjoint from R), doesn't generally screen off R from S, despite S being disjoint from R's causal future.[32]

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 region screens it off from any region not caused by it.

This claim is plausibly true in any spacetime without closed causal curves or action at a distance.

Alas, we are interested in worlds *with* closed causal curves—where 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 doesn't have any *pure* thick parent. In a world like CIRCLE, Parental Markov is therefore entirely vacuous. But it gets worse than that. In some worlds, Parental Markov is outright *false*; in Appendix 2A I provide 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?

[32] 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 annihilate each other and leave behind a red particle; and whenever two particles of *different* colors collide, they annihilate each other and leave behind 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.

7.2 Boundary Markov

Yes. In Minkowski spacetime, where Parental Markov *is* in good standing, it entails that regions are screened off by their *thick boundaries.* In contrast to Parental Markov, this latter principle has obvious application to, and very plausibly remains *true* in, worlds with closed causal curves.

To build an intuition for the new principle, consider again LINE (cf. section 5). Suppose that you, God, have studiously probed the world's urchance function. You then observe the particle in LINE on two days, d_9 and d_{11}; you learn that $\text{BLUE}(d_9)$ and $\text{RED}(d_{11})$. Suppose you'd like to calculate the chance distribution over times *other* than d_9, d_{10}, d_{11}.

To do this, is it worth examining d_{10}, to establish whether $\text{BLUE}(d_{10})$ or $\text{RED}(d_{10})$? No: as far as the chances over times *outside* of d_9, d_{10}, d_{11} are concerned, d_{10}'s state doesn't matter. Any information d_{10} may carry about the times outside of d_9, d_{10}, d_{11}—so long as it is at least compatible with the dynamical laws—is *screened off* by the states of d_9 and d_{10}. For example, where A is any proposition purely about times other than d_9, d_{10}, d_{11},

$$\begin{aligned} \text{urch}_L(A|\text{BLUE}(d_9) \wedge \text{RED}(d_{10}) \wedge \text{RED}(d_{11})) = \\ =& \text{urch}_L(A|\text{BLUE}(d_9) \wedge \text{BLUE}(d_{10}) \wedge \text{RED}(d_{11})) \\ =& \text{urch}_L(A|\text{BLUE}(d_9) \wedge \text{RED}(d_{11})). \end{aligned} \tag{13}$$

The interior of $d_9 \cup d_{10} \cup d_{11}$ is 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, let a *thick neighborhood* of region R be 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.[33] Intuitively, a thick neighborhood is like a city plus its suburbs: coming from the outside, you have to pass through the suburbs for some time to get to the city. A *thick boundary* of R is any region B disjoint from R such that $R \cup B$ contains a thick neighborhood of R. So a thick boundary contains the city's suburbs, but no part of the city itself.[34] See e.g. fig. 5, where B is a thick boundary of R. As I prove in Appendix 1, in Minkowski spacetime (indeed, any space 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. Accordingly, in Minkowski spacetime (and any space with the same topology as $\mathbb{R}^n$), B is a thick boundary of R iff B is disjoint from R and $R \cup B$ contains an open superset of R's closure.

[33] That is, 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 a non-trivial subcurve, and $c[I_2]$ intersects R.

[34] Note the edge cases: the entire spacetime has the empty set as a thick boundary ("the universe doesn't have suburbs"); the empty set has every region as a thick boundary ("everything is a suburb of the empty set").

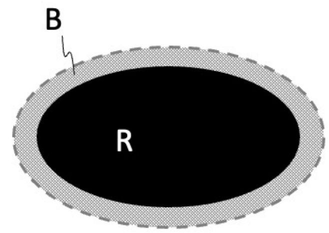


Figure 5:

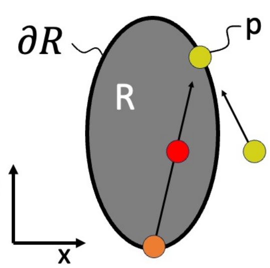
The claim about boundaries is now this:[35]

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

In Appendix 3, I prove that Parental Markov entails Boundary Markov in Minkowski spacetime, given plausible continuity and locality assumptions for urch.

Every region has a thick boundary—indeed most regions have infinitely many of them. This is true of any spacetime, with or without loops. So, in contrast to Parental Markov, Boundary Markov continues to make substantive predictions in worlds with

[35]Like parental thickness, boundary thickness accommodates reductive accounts of derivatives. But boundary thickness plays a vital second role. Even for non-reductivists, a *thin* boundary doesn't generally 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. 32), 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, which includes a future yellow particle at point p, the probability that R contains a red particle is then equal to the probability that there is a yellow particle on course to collide with it exactly at p. (Recall that the only way for a red particle to transform into a yellow one is to collide with another yellow particle.) This probability may be undefined, or alternatively some low value at or near 0. Conditioning additionally on the existence of such a yellow particle raises this probability to 1. So ∂R alone doesn't screen off R from the outside. Meanwhile, any thick boundary would already contain 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.

closed causal curves. Moreover, given local laws, those predictions are plausibly all *true*: besides cases involving action at a distance, I'm not aware of any metaphysically possible counterexample to Boundary Markov. (Indeed, I'm inclined to think that violations of Boundary Markov are *constitutive* of action at a distance: what it *means* for action at a distance to occur at a world is for Boundary Markov to be false at that world.) Appendix 2B demonstrates explicitly that in the cyclic case in which Parental Markov is false, Boundary Markov is still substantive and true.

7.3 Intrinsic Geometrical Information

We are almost there. So far, when we've spoken of a region's *intrinsic state* being screened off, we've implicitly understood this as being about the region's *matter content.* But what's screened off is plausibly not merely its matter content but also its *internal geometry.*

To illustrate, suppose you've studied LINE everywhere outside of d_{10}, such that you know the intrinsic geometry of $d_0 \cup ... \cup d_9 \cup d_{11} \cup d_{12} \cup ....$ Meanwhile, you haven't examined the interval d_{10} at all—you're even unsure about *how long* the interval is. For the purpose of calculating the chance of some proposition about times other than d_9, d_{10}, d_{11}, does your ignorance matter? No, because geometrical information is also screened off. Where LINE* (or L^* for short) is a world just like LINE except that d_{10} is two days long, we should have

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

where A is a proposition about the complement of d_{10} and urch_L and urch_{L^*} are, as usual, the results of conditioning urch on a complete description of LINE's geometry and LINE*'s geometry, respectively.

We thus obtain a strengthening of "screening off". One way to define it compares the original spacetime with the result of outright *deleting* the screened-off region. For any spacetime $\mathcal{M}$, let $\mathcal{M}\backslash X$ denote the result of deleting region X from $\mathcal{M}$.[36] Moreover, let $\text{urch}_{\mathcal{M}\backslash X}$ be the result of conditioning urch on a complete geometrical description of $\mathcal{M}\backslash X$. Note that these descriptions include a "that's all" clause (cf. fn. 19)—a proviso that $\mathcal{M}\backslash X$ is *all of spacetime*, i.e. that it captures all true geometric relationships. Thus $\text{urch}_{\mathcal{M}\backslash X}$ is *not*

[36] More precisely, we'll first define the deletion operation for manifolds. Let $M = (\Omega, \mathcal{A}, g)$ be a Lorentzian manifold with point set 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 members of $\mathcal{A}$ to their domains minus X. (When X has a non-differentiable boundary, $(\Omega\backslash X, \mathcal{A}|_{\Omega\backslash X})$ will generally be neither a manifold nor manifold-with-boundary. But no matter: it retains all the metrical 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 $M\backslash X$.

the result of conditioning urch on an incomplete description that's *true* at $\mathcal{M}$, but instead of conditioning urch on a complete description that's *false* at $\mathcal{M}$.

Def. **Strong Screening Off:** For any spacetime $\mathcal{M}$, and all regions R, S, T in $\mathcal{M}$: region S *strongly screens off* R *and* T (in $\mathcal{M}$) iff

$$\text{urch}_{\mathcal{M}}(Q_1(R)|Q_2(S) \wedge Q_3(T)) = \text{urch}_{\mathcal{M}\setminus T}(Q_1(R)|Q_2(S)),$$

for any maximal qualitative intrinsic property Q_2 and qualitative intrinsic properties Q_1 and Q_3 such that at least one side of the equation is well-defined and $Q_2(S) \wedge Q_3(T)$ is possible according to $\text{urch}_{\mathcal{M}}$.

Replacing "screens off" in Boundary Markov by "strongly screens off" yields[37]

Thesis. **Strong Boundary Markov:** Any thick boundary of a 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 in this regard: both are screened off by thick boundaries.

[37]To see how Strong Boundary Markov entails that thick boundaries screen off internal geometry, we need a notion of *adding* a region to spacetime. I'll define this precisely at the end of the footnote; for now, simply note that adding is an inverse of deleting (cf. fn. 36), "gluing" a region onto the existing spacetime. Let now B be a thick boundary of R in $\mathcal{M}$, and let A be a proposition purely about the complement of $R \cup B$. Let $\mathcal{M}^*$ be a result of adding to $\mathcal{M}\setminus R$ some region R^*, such that B is still a thick boundary of R^* in $\mathcal{M}^*$. Given Strong Boundary Markov, we then have both

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

(where Q_1 is a maximal qualitative intrinsic property and Q_2 and Q_3 are qualitative 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 $\text{urch}_{\mathcal{M}^*}$). But since adding is an inverse of deleting, $\mathcal{M}\setminus R = \mathcal{M}^*\setminus R^*$. So,

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

So B screens off A from whatever happens on B's "inside", including its internal geometry.

To define *adding*, first define it for a manifold or manifold-like structure. Specifically, let $M = (\Omega, \mathcal{A}, g)$ be the result of restricting a Lorentzian manifold by deleting a (possibly empty) region, as defined in fn. 36. Then $(\Omega \cup X, \mathcal{A}^{\Omega\cup X}, g^{\Omega\cup X})$ is *a result of adding set* X *to* M if $\mathcal{A}^{\Omega\cup X}$ is the union of $\mathcal{A}$ with a set of charts such that, for every $x \in X$ there is a neighborhood U_x of x in $\Omega \cup X$ such that U_x is a subset of the domain of at least one chart in $\mathcal{A}^{\Omega\cup X}$, all charts in $\mathcal{A}^{\Omega\cup X}$ satisfy the mutual smoothness condition (i.e. concatenations of a chart and another chart's inverse are smooth on the image, under the latter chart, of their domains' intersection), $\mathcal{A}^{\Omega\cup X}|_{\Omega} = \mathcal{A}$ (see fn. 36 for definition of $|.$ on atlases), and $g^{\Omega\cup X}$ is an extension of g to $\Omega \cup X$ that's smooth relative to $\mathcal{A}^{\Omega\cup X}$. Deleting is an inverse of adding provided $\mathcal{A}$ is a maximal atlas, containing all coordinate maps that satisfy the mutual smoothness condition—this is a standard assumption (e.g. Malament (2012)). If $\mathcal{M}$ is a spacetime represented by M, *a result of adding* X *to* $\mathcal{M}$ is then any fusion of a spacetime region and $\mathcal{M}$ represented by $M \cup X$.

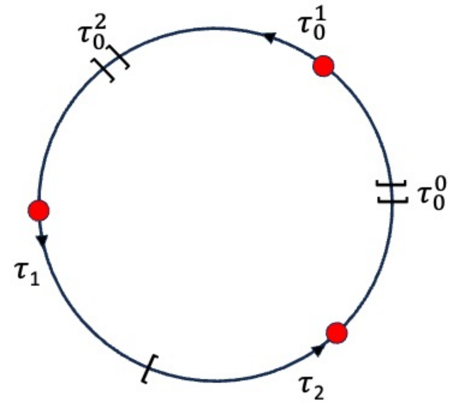


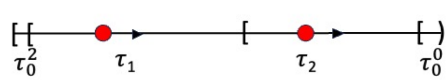
Figure 6: Each τ_i denotes a half-open region—closed toward the past, open toward the future)

8 Cutting Loops

Our theory of chances on loops is the conjunction of Strong Boundary Markov and Acyclic Chance Invariance. We are ready to apply it.

To start simply, consider SMALL CIRCLE (or *SC* for short), which is just like CIRCLE except only three days long. We'd like to determine the value of $\text{urch}_{SC}(A|\text{RED}(\tau_0))$, where A is any proposition about days other than τ_0. To do so, we partition τ_0 into three intervals. For concreteness, suppose they are the first minute, $\tau_0^0 = 0:00$ (up to, but not including, $0:01$), the last minute, $\tau_0^2 = 23:59$ (up to, but not including, $0:00$), and the rest, τ_0^1 (from $0:01$ up to, but not including, $23:59$). See fig. 6. The region $\tau_0^0 \cup \tau_0^2$ is a thick boundary of τ_0^1, and hence, given Strong Boundary Markov, strongly screens it off from A.

But crucially, the internal geometry of the world *outside* of τ_0^1 doesn't by itself determine whether the world forms a causal loop or not. Indeed, the result of deleting τ_0^1 from SMALL CIRCLE, SMALL LINE (or *SL* for short), is *loop-free*:[38]



But given Acyclic Chance Invariance, we already know the transition chances for SMALL LINE: they are the same as for LINE (cf. Section 5). Strong Boundary Markov then lets us express the transition chances on SMALL CIRCLE in terms of the transition chances on

[38]Mathematically, it's a loop-free *manifold with boundary.*

SMALL LINE. And now we are done: we've derived the cyclic transition chances from the acyclic transition chances.

Let's do this more slowly. Let RED_i be the property of containing a red particle whose clock reads $0:00$, RED_m be the property of containing a red particle whose clock goes from $0:01$ to $23:59$ in proportion to the time passed, and let RED_f be the property of containing a red particle whose clock reads $23:59$. According to Acyclic Chance Invariance, the transition chances in SMALL LINE are as follows:

$$\begin{aligned}
\mathrm{urch}_{SL}(\mathrm{RED}(\tau_1)|\mathrm{RED}_f(\tau_0^2)) &= 0.8,\\
\mathrm{urch}_{SL}(\mathrm{GREEN}(\tau_1)|\mathrm{RED}_f(\tau_0^2)) &= 0.2,\\
&\dots\\
\mathrm{urch}_{SL}(\mathrm{RED}_i(\tau_0^0)|\mathrm{RED}(\tau_2)) &= 0.8,\\
\mathrm{urch}_{SL}(\mathrm{GREEN}_i(\tau_0^0)|\mathrm{RED}(\tau_2)) &= 0.2,\\
&\text{etc.}
\end{aligned} \tag{14}$$

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

$$\begin{aligned}
\mathrm{urch}_{SC}(A|\mathrm{RED}(\tau_0)) &= \mathrm{urch}_{SC}(A|\mathrm{RED}_i(\tau_0^0) \wedge \mathrm{RED}_f(\tau_0^2) \wedge \mathrm{RED}_m(\tau_0^1))\\
&= \mathrm{urch}_{SL}(A|\mathrm{RED}_i(\tau_0^0) \wedge \mathrm{RED}_f(\tau_0^2)).
\end{aligned} \tag{15}$$

where the first line follows because, given a complete description of SMALL CIRCLE's geometry, $\mathrm{RED}(\tau_0)$ is logically equivalent to $\mathrm{RED}_i(\tau_0^0) \wedge \mathrm{RED}_f(\tau_0^2) \wedge \mathrm{RED}_m(\tau_0^1)$, and the second line follows by Strong Boundary Markov. We can derive the value of $\mathrm{urch}_{SL}(A|\mathrm{RED}_i(\tau_0^0) \wedge \mathrm{RED}_f(\tau_0^2))$ from eqs. 14 and Parental Markov for SMALL LINE. (Recall that Parental Markov is in good standing in worlds without closed causal curves.) The result is $64/65 \approx 0.98$.[39] Plugging this into eq. 15,

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

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

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

$$\begin{aligned}
&\mathrm{urch}_{SL}(\mathrm{RED}(\tau_1)|\mathrm{RED}_i(\tau_0^0) \wedge \mathrm{RED}_f(\tau_0^2)) =\\
&\quad = \frac{\mathrm{urch}_{SL}(\mathrm{RED}_i(\tau_0^0)|\mathrm{RED}(\tau_1) \wedge \mathrm{RED}_f(\tau_0^2)) \cdot \mathrm{urch}_{SL}(\mathrm{RED}(\tau_1)|\mathrm{RED}_f(\tau_0^2))}{\mathrm{urch}_{SL}(\mathrm{RED}_i(\tau_0^0)|\mathrm{RED}_f(\tau_0^2))}.
\end{aligned} \tag{16}$$

$$\mathrm{urch}_{SC}(\mathrm{RED}(\tau_1)|\mathrm{RED}(\tau_0)) = \frac{64}{65} \approx 0.98. \tag{17}$$

This is the cyclic transition chance we were looking for.

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

$$\mathrm{urch}_{SC}(\mathrm{RED}(\tau_1)|\mathrm{RED}(\tau_0)) > 0.8. \tag{18}$$

For it's quite natural to think that $\mathrm{RED}(\tau_0)$ *doubly* supports $\mathrm{RED}(\tau_1)$: $\mathrm{RED}(\tau_1)$ is not only

---

Let's evaluate each part of the quotient separately. The first factor in the numerator:

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

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 $\mathrm{RED}(\tau_2)$ is the only option for τ_2 nomically compatible with $\mathrm{RED}(\tau_1) \wedge \mathrm{RED}_i(\tau_0^0)$, and the final equality follows from the transition chances (eqs. 14) for SMALL LINE. The value of the numerator's second factor follows immediately from the transition chances:

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

Finally, to calculate the denominator, observe that

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

where the second line follows because LINE's dynamics disallow immediate RED-to-BLUE transitions. For $\pi = \mathrm{RED}$, the right-hand side's summand is just the numerator, whose value we've just calculated: 0.8^3. (Note that, as promised, this proves that the denominator is positive.) For $\pi = \mathrm{GREEN}$, we perform essentially the same calculations, noting that $\mathrm{BLUE}(\tau_2)$ is the only option nomically compatible with $\mathrm{GREEN}(\tau_1)$ and $\mathrm{RED}_i(\tau_0^0)$. We obtain:

$$\mathrm{urch}_{SL}(\mathrm{RED}_i(\tau_0^0)|\mathrm{GREEN}(\tau_1) \wedge \mathrm{RED}_f(\tau_0^2)) \cdot \mathrm{urch}_{SL}(\mathrm{GREEN}(\tau_1)|\mathrm{RED}_f(\tau_0^2)) = 0.2^2 \cdot 0.2.$$

Plugging everything into eq. 16,

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

a likely *effect* of RED(τ_0), but also a likely *cause*: going from RED(τ_1) to RED(τ_0) involves transitioning from RED(τ_1) to RED(τ_2) to RED(τ_0)—two "likely" transitions. By contrast, going from GREEN(τ_1) to RED(τ_0) involves two "unlikely" transitions: GREEN(τ_1) to BLUE(τ_2) to RED(τ_0). So RED(τ_0) should favor RED(τ_1) both because it preferentially causes it *and* because it's preferentially caused *by* it. Eq. 15 captures this "double support" intuition. It says that the result of conditioning the urchance on today's particle being red equals (as far as propositions about times other than τ_0 are concerned) the result of conditioning the acyclic urchance on today's particle *and* the particle three days from now being red—this *doubly supports* tomorrow's redness.

Our proposal ticks all three boxes: since they're fully derived from urch$_{SL}$, the conditional chances in SMALL CIRCLE are dynamically scrutable. As eq. 17 shows, the approach also avoids trivialization. And, since it just consists of the conjunction of Strong Boundary Markov and Acyclic Chance Invariance, the proposal is consistent with any local dynamics whose chance prescriptions are invariant across acyclic worlds—which includes any plausible local dynamics, including any of the dynamics we've encountered in this essay.

There is a hidden fourth benefit to our account. In CIRCLE, it salvages, as it were, *asymptotic* chance invariance: for long roundtrip times, the short-term cyclic transition chances quickly approach the acyclic transition chances. In CIRCLE, with its roundtrip time of 15 billion years, the value of (say) urch$_C$(GREEN(τ_1)|RED(τ_0)) is essentially indistinguishable from 0.2. (The error is smaller than $0.8^{10^{12}}$.) More generally, unless the chance, q, to switch color in the *acyclic* case is *extremely* close to 0 or 1, urch$_C$(GREEN(τ_1)|RED(τ_0)) converges quickly to q for increasing roundtrip time. This makes sense intuitively: knowing the particle's color in the far future provides little evidence about the near-term colors.[40] More generally, our approach ensures asymptotic invariance in a cyclic spacetime (of whatever dimension, provided we can partition it into times) whenever, according to the *acyclic* dynamics, far-future states are increasingly probabilistically independent of near-future states.[41]

[40] More formally, where $q \in [0, 1)$ is the acyclic chance for a particle to switch its current color during the next transition and l is the loop length, we obtain

$$\mathrm{urch}_L(\mathrm{GREEN}(\tau_1)|\mathrm{RED}(\tau_0) \wedge \mathrm{RED}(\tau_{-1})) = \frac{\sum_{n=1}^{\lfloor l/3 \rfloor} (1-q)^{l-3n} q^{3n} \binom{l-1}{3n-1}}{\sum_{n=0}^{\lfloor l/3 \rfloor} (1-q)^{l-3n} q^{3n} \binom{l}{3n}}. \tag{19}$$

(Where $\lfloor l/3 \rfloor = 0$, the numerator sum is set to 0.) For increasing l, this converges to q quickly, which can be proved by finding an analytic expression.

[41] To see this: suppose we're given a cyclic spacetime $\mathcal{C}$. Let τ be a day in $\mathcal{C}$, which we partition, as before, into τ^0, τ^1, τ^2. (In higher-dimensional spacetimes, $\tau, \tau^0, \tau^1, \tau^2$ are hypercuboids, bounded from above and below by time slices.) Let $\mathcal{L} := \mathcal{C} \backslash \tau^1$. The longer the return time in $\mathcal{C}$, the greater the forward distance from

9 Saving Grandpa

Let's apply our framework to the stochastic grandfather paradox from the introduction. For concreteness, let the spacetime $\mathcal{M}$ be flat (more specifically, Minkowskian) except for one topological quirk, the *wormhole*. We can represent the wormhole with two duplicate (three-dimensional) space-like surfaces, w_1 and w_2. Every future-directed causal curve intersecting w_1 exits at (without intersecting) w_2, from where it continues future-ward. Similarly, every future-directed causal curve intersecting w_2 exits at (without intersecting) w_1, from where it continues future-ward.[42] See figure 7, which includes an example trajectory through the wormhole.[43]

Figure 7 also indicates six additional spacetime regions: P is the region where the poisoning occurs; A_1 (A_2) is the region where the first (second) antidote is administered, with A_1 additionally partitioned into A_1^-, A_1°, and A_1^+; H_1 (H_2) is the region where the first (second) antidote would take effect if it did; G is the region where grandpa has children, one of whom bears you and your sibling.

To simplify things, we'll consider finitely many maximal intrinsic states for each region, represented by natural numbers. We'll then write property assignments to regions in the form of $A = n$, where A a region and n a natural number. Some care is needed to ensure that regions have their respective states *intrinsically*. For example, region P has "containing an infant" intrinsically,[44] but not "containing your grandfather".

τ^2 to τ^0 in $\mathcal{L}$. Let τ_+ be a time in τ^2's near-term future, with possible intrinsic property Q. Given a long roundtrip time and approximate probabilistic independence of far-future from near-future states, we have, in $\mathcal{L}$,

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

(where "$P(\tau^i)$" denotes the strongest proposition entirely about τ^2 entailed by $P(\tau)$). Hence, by Strong Boundary Markov,

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

But $\mathrm{urch}_{\mathcal{L}}(Q(\tau_+)|P(\tau^2))$ equals $\mathrm{urch}_{\mathcal{L}}(Q(\tau_+)|P(\tau))$, the ordinary acyclic transition chance from τ into τ_+. Hence

$$\mathrm{urch}_{\mathcal{C}}(Q(\tau_+)|P(\tau)) \approx \mathrm{urch}_{\mathcal{L}}(Q(\tau_+)|P(\tau)).$$

So the near-term transition chance in $\mathcal{C}$ is approximately invariant.

[42]To keep derivatives everywhere well-defined, w_1 and w_2's (two-dimensional) boundaries are deleted from spacetime (i.e., they end in two-dimensional singularities).

[43]The foregoing description isn't unique to (w_1, w_2) and instead true of any other pair of duplicate space-like bounded surfaces with the same boundaries (i.e., singularities) as w_1 and w_2. Any such pair represents the exact same wormhole. Arguably, "wormhole" therefore most naturally refers to the union of w_1's and w_2's domains of dependence—this region *is* distinguished. But to keep things as accessible as possible, I'll keep framing things in terms of w_1 and w_2 specifically.

[44]Where we understand "infant" as denoting someone with the physiology typical of a neonate.

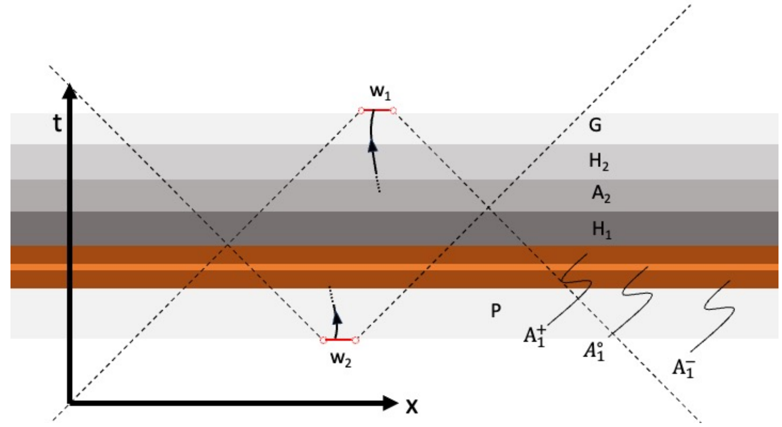


Figure 7:

In the following, let a bracketed expression $(\phi)_i$ indicate that "ϕ" is to be added for the values i:

- $P = 0, 1$: infant (not)$_0$ poisoned by adult in P
- $A_1^- = 0, 1$: in A_1^-, infant (healthy)$_0$ (sick)$_1$ and antidote 1 (not)$_0$ about to be administered
- $A_1^\circ = 0, 1$: in A_1°, infant (healthy)$_0$ (sick)$_1$and antidote 1 (not)$_0$ being administered
- $A_1^+ = 0, 1$: in A_1^+, infant (healthy)$_0$ (sick)$_1$ and antidote 1 (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 (not)$_{0,1}$ taking effect on (healthy)$_0$ (sick)$_{1,2}$ infant in H_1
- $A_2 = 0, 1$: in A_2, infant (healthy)$_0$ (sick)$_1$ and antidote 2 (not)$_0$ administered
- $H_2 = 0, 1, 2$: antidote 2 (not)$_{0,1}$ taking effect on (healthy)$_0$ (sick)$_{1,2}$ infant in H_2
- $G = 0, 1$: in G, infant (not)$_0$ alive, and (not)$_0$ eventually growing up to have two grandchildren

(In H_1 and H_2 we're assuming that the antidote only takes effect on sick people.)

One salient interpretation of "chance, upon administration, of the first antidote's working" is $\text{urch}_{\mathcal{M}}(H_1 = 2 | A_1 = 1)$—*viz.* the urchance, in $\mathcal{M}$, of the first antidote's working conditional on its being administered in A_1. To calculate this, first note that $A_1^+ \cup A_1^-$ is a thick boundary of A_1° (cf. fig. 7). Thus consider $\mathcal{M}' := \mathcal{M} \backslash A_1^\circ$,[45] sketched in

[45] For a definition of the $\backslash$-operation on spacetimes see fn. 36.

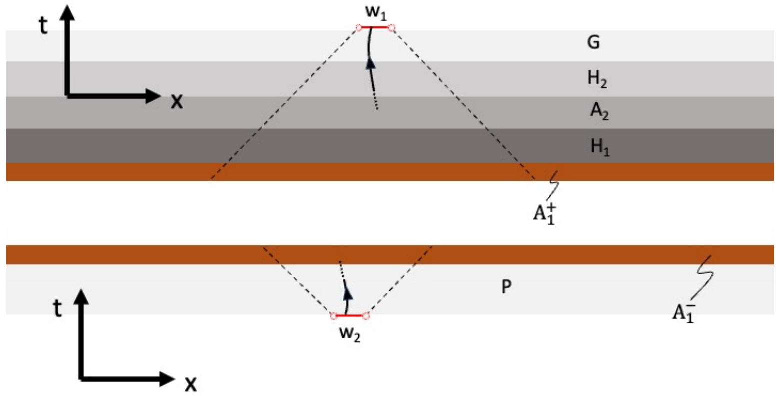


Figure 8:

figure 8.[46] Since $A_1 = 1$ is necessarily equivalent to $A_1^- = A_1^\circ = A_1^+ = 1$, we obtain, by Strong Boundary Markov:

$$\begin{aligned}\text{urch}_{\mathcal{M}}(H_1 = 2|A_1 = 1) &= \text{urch}_{\mathcal{M}}(H_1 = 2|A_1^- = A_1^\circ = A_1^+ = 1)\\ &= \text{urch}_{\mathcal{M}'}(H_1 = 2|A_1^- = A_1^+ = 1).\end{aligned} \tag{20}$$

But A_1° intersects all closed causal curves in $\mathcal{M}$, and so $\mathcal{M}'$ is entirely loop-free.

For simplicity, let's assume that the antidote's actions are the only indeterministic processes in $\mathcal{M}'$. This means that, among adjacent regions, only the transitions from A_1^+ into H_1 and from A_2 into H_2 are chancy. According to the introduction, *normally*, there's a 50% chance that an administered antidote will cure a given poisoning. These "normal circumstances" include ordinary Minkowski spacetime. By Acyclic Chance Invariance, the Minkowskian transition chances then also apply to $\mathcal{M}'$. That is:[47]

$$\text{urch}_{\mathcal{M}'}(H_1 = i|A_1^+ = 1) = \text{urch}_{\mathcal{M}'}(H_2 = i|A_2 = 1) = 0.5. \tag{21}$$

[46]The differential geometry for $\mathcal{M}'$ is slightly subtle: If A_1^- and A_1^+ are open, infinitely extended, non-intersecting hypercuboids, then $\mathcal{M}'$ is (representable by) a smooth manifold with boundary. If A_1^- and A_1^+ are bounded hypercuboids that do intersect, the result of removing A_1° *isn't* (representable by) a smooth manifold with boundary: the coordinate maps whose domains "wrap around" the edge of A_1° don't map into open subsets of the half-plane. However, while lacking (to my knowledge) a canonical name, the resulting structure is still well-defined and has all the needed differential and metrical structure: fields defined on $\mathcal{M}\backslash A_1^\circ$ are smooth on $\mathcal{M}\backslash A_1^\circ$ in $\mathcal{M}$ iff they are smooth on $\mathcal{M}'$.

[47]Explicitly, using the precise formulation of Acyclic Chance Invariance: let $(A_1^+)^M$ and H_1^M be regions in Minkowski spacetime, $\mathcal{M}^M$, with $(A_1^+)^M < H_1^M$, such that there's an isometry Φ from the region between A_1^+ and H_1 in $\mathcal{M}'$ into the region between $(A_1^+)^M$ and H_1^M in $\mathcal{M}^M$ with $\Phi(A_1^+) = (A_1^+)^M$ and $\Phi(H_1) = H_1^M$. Since there's no wormhole between A_1^+ and H_1 in $\mathcal{M}'$, the regions $(A_1^+)^M$ and H_1^M clearly exist. According to our stipulation, for $i = 0, 1$,

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

Note that $\mathcal{M}'$ contains two disconnected infant spacetime worms: one starting in P (at its birth) and ending at A_1^-'s future border, the other starting at A_1^+'s past border and ending in G (as it 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 the *younger infant*, and the spacetime worm from A_1^+'s past border to G the *older infant*. Let's pin down the rest of the dynamics. For any propositions A, B, let $A \Rightarrow B$ denote that, deterministically, if A, then B.[48] ($A \Leftrightarrow B$ 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) $P = 1 \Leftrightarrow G = 1$: younger infant is poisoned in P iff older infant in G grows up to be a grandfather[49]

(f) $A_1^- = 1 \Leftrightarrow P = 1$: the younger infant is sick in A_1^- iff he 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 brought up to the younger infant's mouth in A_1^- iff he 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 is in the older infant's body 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:

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

But, by Acyclic Chance Invariance,

$$\mathrm{urch}_{\mathcal{M}}(H_1 = i | A_1^+ = 1) = \mathrm{urch}_{\mathcal{M}^M}(H_1^M = i | (A_1^+)^M = 1),$$

and so $\mathrm{urch}_{\mathcal{M}}(H_1 = 1 | A_1^+ = 1) = 0.5$. The derivation for $\mathrm{urch}_{\mathcal{M}'}(H_2 = i | A_2 = 1)$ is analogous.

[48] In our Popperian formalism (fn. 13), ⌜deterministically, A given B⌝ is equivalent to: ⌜$\mathrm{urch}(A | B \wedge Q) = 1$ for all propositions Q⌝.

[49] This holds because one of the grandchildren of the older infant poisons the younger infant and, by assumption, nobody else possibly does.

i.e., if an antidote is in the older infant's body in A_1^+, then an antidote is brought up to the younger infant's mouth in A_1^- iff the older infant is healed by at least one antidote. Plugging this into the right-hand side of eq. 20, we obtain the following:

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

That is, the probability, in $\mathcal{M}$, that the first antidote is effective conditional on its administration in A_1 equals the probability, *in* $\mathcal{M}'$, that the first antidote is effective conditional on its just having entered the infant's body in A_1^+ *and* at least one of the two antidotes' working.

It's intuitively clear that the right-hand side of eq. 22 is greater than the usual 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[50]

$$\text{urch}_{\mathcal{M}'}(H_1 = 2|A_1^+ = 1 \wedge (H_1 = 2 \vee H_2 = 2)) = \frac{1/2}{3/4} = \frac{2}{3}. \tag{32}$$

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

$$\text{urch}_{\mathcal{M}'}(H_1 = 2|A_1^+ = 1 \wedge (H_1 = 2 \vee H_2 = 2)) = \frac{\text{urch}_{\mathcal{M}'}(H_1 = 2|A_1^+ = 1)}{\text{urch}_{\mathcal{M}'}(H_1 = 2 \vee H_2 = 2|A_1^+ = 1)}, \tag{23}$$

provided the denominator isn't zero. $H_1 = 2$ and $H_2 = 2$ are mutually exclusive—the infant is cured by the second antidote only if the first antidote has failed. (This follows from the combination of eq. 21, (a), (b) and (c).) Hence, the denominator becomes

$$\text{urch}_{\mathcal{M}'}(H_1 = 2 \vee H_2 = 2|A_1^+ = 1) = \text{urch}_{\mathcal{M}'}(H_1 = 2|A_1^+ = 1) + \text{urch}_{\mathcal{M}'}(H_2 = 2|A_1^+ = 1). \tag{24}$$

By eq. 21, we have

$$\text{urch}_{\mathcal{M}'}(H_1 = 2|A_1^+ = 1) = 1/2. \tag{25}$$

It remains to determine $\text{urch}_{\mathcal{M}'}(H_2 = 2|A_1^+ = 1)$. By Parental Markov in $\mathcal{M}'$,

$$\text{urch }_{\mathcal{M}'}(H_2 = 2|H_1 = i \wedge A_1^+ = 1) = \text{urch }_{\mathcal{M}'}(H_2 = 2|H_1 = i) \tag{26}$$

for $i = 0, 1, 2$. Moreover, by (a), (b), and (c),

$$\text{urch }_{\mathcal{M}'}(H_2 = 2|H_1 = 0) = 0 \tag{27}$$

(if the baby isn't sick to begin with, he is not cured by the second antidote). Likewise,

$$\text{urch }_{\mathcal{M}'}(H_2 = 2|H_1 = 2) = 0 \tag{28}$$

(if the baby is already cured by the first antidote, he is not cured by the second). By 21, (b), and (c),

$$\text{urch }_{\mathcal{M}'}(H_2 = 2|H_1 = 1) = 0.5. \tag{29}$$

So, by eqs. 22 and 32,

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

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 $1/4$ (recall 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 causal loops, chances differ in non-trivial but scrutable ways.

The two scenarios provide us with a general blueprint for deriving chances on loops. In the first instance, Strong Boundary Markov and Acyclic Chance Invariance determine certain *conditional* chances: the chances, conditional on the state of a *spacetime region large enough to intersect all causal loops*, of events outside of that region. The general procedure is this:

General Recipe:

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

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

(if the baby is sick but the first antidote fails, there's a 0.5 chance that the second antidote will work). Hence,

$$\begin{aligned} \text{urch}_{\mathcal{M}'}(H_2 = 2|A_1^+ = 1) &= \sum_{i=0,1,2} \text{urch}_{\mathcal{M}'}(H_2 = 2|H_1 = i) \cdot \text{urch}_{\mathcal{M}'}(H_1 = i|A_1^+ = 1) \\ &= \text{urch}_{\mathcal{M}'}(H_2 = 2|H_1 = 1) \cdot \text{urch}_{\mathcal{M}'}(H_1 = 1|A_1^+ = 1) + \\ &\quad + \text{urch}_{\mathcal{M}'}(H_2 = 2|H_1 = 2) \cdot \text{urch}_{\mathcal{M}'}(H_1 = 2|A_1^+ = 1) \\ &= 1/2 \cdot 1/2 + 0 \cdot 1/2 = \frac{1}{4}, \end{aligned} \tag{30}$$

where the first equality follows from the multiplicative axiom and eq. 26, the second follows from 27, and the third follows from eqs. 28 and 29. Plugging eqs. 25 and 30 into eq. 24, the denominator becomes

$$\text{urch}_{\mathcal{M}'}(H_1 = 2 \vee H_2 = 2|A_1^+ = 1) = 3/4. \tag{31}$$

Plugging eqs. 25 and 31 into eq. 23,

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

4. By Strong Boundary Markov, this equals

$$\text{urch}_{\mathcal{M}}(X|Y \wedge Q_S(S)),$$

where $S := R \cup B$ and Q_S is the state of S entailed by $Q_R(R) \wedge Q_B(B)$.

Seeing this procedure, a natural follow-up question is what else can be derived from it about chances on loops. The answer is: *virtually everything*. The General Recipe generically fixes *marginal* (i.e., unconditional) chance distributions over states of a cyclic spacetime. These chances, in turn, fix chances conditional on arbitrary regions—including small regions, which *don't* intersect all causal loops—via the ratio formula. Moreover, where a loop is embedded into a larger, acyclic spacetime, we can derive chances over states of the loop conditional merely on the state of the world *prior* to the loop. A much larger swath of questions about chances on loops has definite answers than has previously been thought. I'll explain this in the next section.

10 Marginal Chances

Several people have suggested to me in conversation that, generically, there is no privileged way of assigning marginal (i.e., unconditional) chances over the states of a temporally circular universe; or, similarly, conditional only on the state of the world prior to a wormhole, there is no well-defined chance of what the wormhole will spit out.

I once believed this too. It's understandable why: certainly in *acyclic* worlds, like Minkowski spacetime, the transition chances alone generically *don't* fix a marginal chance distribution over the states of the world—one additionally needs a marginal chance distribution over the universe's possible initial conditions.[51] Similarly, one might have thought, closed causal curves need a "seed" marginal distribution to have well-defined marginal chances; e.g. in the case of CIRCLE, a marginal distribution over the possible states of some (arbitrary) day.

But this is not so. An assignment of transition chances on a cyclic world generically *also* fixes the marginal chance distribution over the possible states of that world. Similarly, conditional on the state of the world sometime prior to a wormhole, there is (generically) a well-defined chance for what will emerge from the wormhole.

Intuitively, the reason for this is that causal loops have one "extra" transition compared

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

to analogous linear situations: their "ends" also connect. This extra transition generically imposes an additional constraint on the marginal distribution over the loop, enough to fix it uniquely.[52]

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 $k = 3$.) As before, every color only has itself and a unique other color as a possible successor. The probability axioms imply that, for every $i = 1, ..., n$ and $j = 1, ..., k$ (where we identify k and 0 inside indices):

$$\mathrm{urch}_C(\tau_i = j) = \sum_{l=1}^{k} \mathrm{urch}_C(\tau_i = j | \tau_{i-1} = l) \cdot \mathrm{urch}_C(\tau_{i-1} = l). \tag{33}$$

From the acyclic dynamics, our account obtains all cyclic transition probabilities, $\mathrm{urch}_C(\tau_i = j|\tau_{i-1} = l)$. Hence eq. 33 yields $n \cdot k$ equations in $n \cdot k$ unknowns. Since we also know that, for every $i = 1, ..., n$,

$$\sum_{j=1}^{k} \mathrm{urch}_C(\tau_i = j) = 1, \tag{34}$$

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 4, I show that these equations generically are linearly independent and have a unique solution.

Solving the system (see eq. 43 in the Appendix—I'll skip the calculation here) for SMALL CIRCLE (i.e., $n = k = 3$, with $1/65 \approx 0.015$ transition chance to next allowed color), we get the following result for all $i = 1, 2, 3$:

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

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 $\mathrm{urch}_{SC}(\mathrm{RED}(\tau_i)|\mathrm{RED}(\tau_{i-1})) = 0.5$ for all $i = 1, 2, 3$ and the remaining transition chances are unchanged, we get $\mathrm{urch}_{SC}(\mathrm{RED}(\tau_i)) = 1/66 \approx 0.015$ and $\mathrm{urch}_{SC}(\mathrm{GREEN}(\tau_i)) = \mathrm{urch}_{SC}(\mathrm{BLUE}(\tau_i)) = 65/132 \approx 0.492$ for all $i = 1, 2, 3$. (The reader may verify this by plugging the given transition probabilities into eq. 43 in Appendix 4.) We can of course also

[52] Mellor (1995, Sec. 17.3) once tried to leverage a mathematically similar fact into an argument *against* the possibility of causal loops, by claiming that it's impossible for transition chances to constrain marginal limiting *frequencies* in this way. Berkovitz (2001, pp.14-5) cogently argues that Mellor's claim is unmotivated.

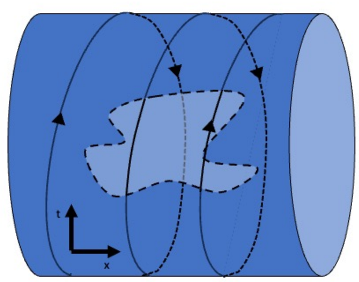
break the symmetry between the *times*, i.e., impose time-dependent transition probabilities, which then also makes the marginal probabilities time-dependent.

I said "generically" linearly independent, because in special circumstances the equations *are* dependent; they then allow for multiple solutions. 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 COL $\in$ {RED, GREEN, BLUE},

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

Given these transition chances, the only constraint the system imposes on the marginals is that $\mathrm{urch}_{SC}(\mathrm{COL}(\tau_{i-1})) = \mathrm{urch}_{SC}(\mathrm{COL}(\tau_i))$. Any probabilistically coherent[53] marginal probability function which satisfies this constraint is nomically allowed. (In Appendix 4, 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.)

Moreover, once we have marginal chance distributions over states of the loop, we can obviously derive all sorts of conditional chance distributions via the ratio formula. Consider a two-dimensional version of CIRCLE, i.e. a flat 2D spacetime rolled up along the time-like direction into a cylinder, like so (where the winding line indicates a possible particle trajectory):[54]



For any arbitrary region on the cylinder (an example indicated in light-blue), our account generically determines the chance, of any particular matter of fact, conditional on the state

[53] I.e., satisfying, for $i = 1, 2, 3$,

$$\sum_{\substack{\mathrm{COL}\in\\ \{\mathrm{RED,GREEN,BLUE}\}}} \mathrm{urch}_{SC}(\mathrm{COL}(\tau_i)) = 1.$$

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

of that region.

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

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 are local, both should be screened off by a thick boundary. Acyclic Chance Invariance, meanwhile, is a consistent weakening of an initially attractive, yet inconsistent, general invariance principle. The weakening says that chances are invariant among *loop-less* worlds.

These general principles fix nearly everything about chances on loops. Given the acyclic transition chances, they fix the cyclic transition chances, but also *unconditional* chance distributions and, as a result, chances conditional on any proposition with positive marginal chance. In this sense, chances on loops are more completely defined than chances off loops.

Accordingly, rational *credences* are also more sharply constrained in loop worlds. Following standard chance deference-principles for urchance,[55] provided you know the true chance laws governing a cyclic world, generically you should have precise credences about the state of that world. Similarly, provided you know the true chance laws, and are well-informed about the current state of the world, you should have precise expectations

[55] Where $< \text{urch} = u >$ is the proposition that the true urchance function is u and Cr_0 is any rational prior credence function, one chance deference principle is this:

$$Cr_0(A|E \wedge < \text{urch} = u >) = u(A|E),$$

for any propositions A and E for which $u(A|E)$ is well-defined.

about what will emerge from any future wormhole. Future wormholes are no black boxes.

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. Moreover, it preserves what are arguably the two next best things to general invariance. The first is Acyclic Chance Invariance. The second is the idea that urchance should be "dynamically scrutable": the cyclic chances should be derived from the acyclic chances in a principled way. The conjunction of Strong Boundary Markov and Acyclic Chance Invariance clearly is such a principled basis. 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 still *practically* invariant in the near-term.

In 2001, Berkovitz (2001, p.21) asked if there are "any non-arbitrary ... principles for relating [ordinary, acyclic] chances and long-run frequencies in causal loops?" His paper doesn't employ the concept of *cyclic* chances, instead following Mellor (1995) in cashing out questions about stochasticity on loops in terms of frequencies. But in a flexible chance formalism like urchance, the cyclic chance concept is well-formed and unmysterious. Indeed, principles linking it to frequencies are just instances of a general urchance-frequency principle.[56] Given such a principle, asking for non-arbitrary links between acyclic chances and cyclic long-run frequencies is just asking if there are non-arbitrary links between the acyclic chance and the *cyclic chances*. We've answered with Yes; and not only are the principles non-arbitrary, but they come with a whole host of other benefits, and go much further than one might initially think.

Programmatically, this essay supports flexible chance formalisms, by demonstrating how they solve problems eluding other approaches. On our rendering of the 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 temporal or causal history regions can be highly informative.[57] Some of these regions cover exactly the local environments of coin flips, of roulette wheels, or of decks of cards. The concept of a background proposition thereby subsumes the concept of a *chance setup*. Indeed, conceptual economy suggests

[56]Here's one formulation of such a principle:

$$\text{urch}(\text{lim-freq}(Q, P, G) = \text{urch}(Q(R)|P(S) \wedge G(R, S)) \text{ if the latter is defined}) = 1$$

where G is any qualitative relation, Q and P are intrinsic qualitative properties, and $\ulcorner\text{lim-freq}(Q, P, G) = x\urcorner$ is the proposition that, in the limit of infinitely many cases where two regions X and Y are such that $P(Y)$ and $G(X, Y)$, in a fraction x of those cases, $Q(X)$.

[57]That is, assuming we additionally supply a description of the world's background geometry—something which the temporalist also has to do.

identifying the two: every background proposition is a chance setup, every chance setup a background proposition. This identification enshrines a view Popper (1959) embraced long ago: that chances are intimately tied, not to time or causation, but to "*arrangements*", i.e., chance setups.

A Appendix

Appendix 1: Thick Neighborhoods = Neighborhoods of Closures

Let a *neighborhood* (simpliciter) *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 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 3.

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

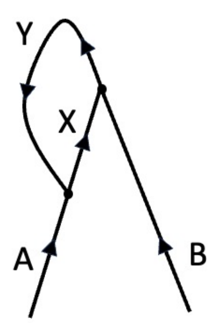
Proof: Right-to-left direction: Let N be a neighborhood of $\overline{R}$, and let c be a continuous curve which starts in $N^{\perp}$ and ends in R. Without loss of generality, assume $c : [0, 1] \to \mathbb{R}^n$. Since c is continuous, $c^{-1}(\overline{R}) \subseteq [0, 1]$ is closed and hence compact. 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 $c(0) = q$, and for all $n \in \mathbb{N}_{>0}$, $c(1/n) = q_n$, and $c|_{]1/(n+1),1/n)[}$ 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 (countable) 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 $\overline{R}$. ■

Appendix 2: Parental and Boundary Markov in Worlds with Loops

A: Where Parental Markov Is False...

Consider the following one-dimensional spacetime:



This structure is obtained from a one-dimensional oriented spacetime with an "inverse fork", by identifying the upper boundary (end point) of Y with the lower boundary (starting point) of X. Suppose the world is home to a scalar field, with completely deterministic dynamics. Specifically, suppose that the field takes the value 0 or 1 at each spacetime point, and that there are two kinds of spacetime points. *Fork points* are locations where two lines converge, represented by black dots in the figure above.[58] All other points are *boring points*. In any interval consisting entirely of boring points, the field remains constant. This also applies to any interval that starts at a fork point and is otherwise made up of boring points. At fork points, the field's value is determined by the values on the incoming lines: if the values on each line are the same, the value at the fork point is 1; otherwise, it is 0. Let's call this world FORK (or F for short). If we introduce, for each region, a binary variable whose value represent the fields value taken at that region, we can write the dynamics compactly as follows:

[58]Topologically, we can define a fork point as any point p such that there are two open lines containing p that do not share an open sub-line containing p.

$$X = A \cdot Y + (1 - A) \cdot (1 - Y),$$
$$Y = B \cdot X + (1 - B) \cdot (1 - X). \tag{35}$$

Jointly the two equations have exactly four distinct solutions:

A	B	X	Y
0	0	1	0
0	0	0	1
1	1	1	1
1	1	0	0

The empty set is a thick parent of A, and neither it nor B are caused by A. Hence, Parental Markov requires the following:

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

But, as the table shows, we have

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

So Parental Markov is false.[59]

B: ... Boundary Markov Is Still True

Yet Boundary Markov is still true. To show this, we examine each subregion.

The thick boundary of any (trivial or non-trivial) subinterval properly in the interior of a line segment obviously screens it off from the rest of the graph: it surrounds the subinterval with two non-trivial intervals from the same line segment, and since the field is stipulated to be constant along every line segment, this fixes the value of the field along the subinterval.

Slightly more interesting are the line segments themselves. Let's consider the segments A and X. Since the dynamics are symmetric, this also proves the case for B and Y. Every thick boundary of A includes the initial point x of X (the point where A, X, and Y meet). But to be a *thick* boundary, it must also contain *half-open intervals* at the start of X and the end of Y, containing x. It follows that every thick boundary of A fixes the field values on

[59] 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 an additional interval prior to A.

X and Z. But then, as per the eqs. 35, it also fixes the field values on A. In particular, then, it screens A off from the rest of spacetime.

The same goes for X: every thick boundary of it must contain half-open intervals at the end of A, at the start and end of Y, and at the end of B. It follows that it fixes the field values everywhere. In particular, then, it screens X off from the rest of spacetime. The same goes for the interiors and closures of line segments, as well as their end points.

From the probability laws one can then show that the same holds for any unions of these regions. So Boundary Markov is true in FORK.

Appendix 3: Parental Markov Entails Boundary Markov in Minkowski Spacetime

This appendix proves that Parental Markov entails Boundary Markov in Minkowski spacetime, given plausible continuity and locality assumptions about urchance. We'll first establish some auxiliary lemmas. As before, a *neighborhood* of A is an open superset of A. Where B is a thick boundary of R, let $B_R^+ := K^+(R) \cap B$ be the part of B to R's future, and $B_R^- := B \backslash B_R^+$ the rest of B. A region is *causally convex* if it contains all causal curves starting and ending in it. Note 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.[60]

> ***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^+. Choose a $t \in [0,1]$ 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 $B_R^+ \subseteq R^\perp$,

[60] 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 1) relies on the extra strength in the definition of thick boundary.

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^-, i.e. 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. ■

> ***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 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.[61] Since B is a thick boundary of R, $R \cup B$ contains a neighborhood N of $\overline{R}$ (by Equivalence, Appendix 1). Thus $J^+(N)$ is a neighborhood of $\overline{J^+(R)}$. Since $J^+(N) \subseteq J^+(R \cup B)$, $J^+(R \cup B)$ contains a neighborhood of $\overline{J^+(R)}$. Since $\overline{B_R^+} \subseteq \overline{J^+(R)}$, $J^+(R \cup B)$ contains a neighborhood of $\overline{B_R^+}$. Therefore, all 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^+. ■

Now the first main result:

> ***Theorem 2*:** Parental Markov entails that, for all regions R, if R is causally convex and B is a thick boundary of R such that $R \cup B$ is causally convex, then B screens off R from $R^\perp$.

Proof of Theorem 2: Let R be causally convex and B a thick boundary of R such that $R \cup B$ is causally convex. We have the following facts:

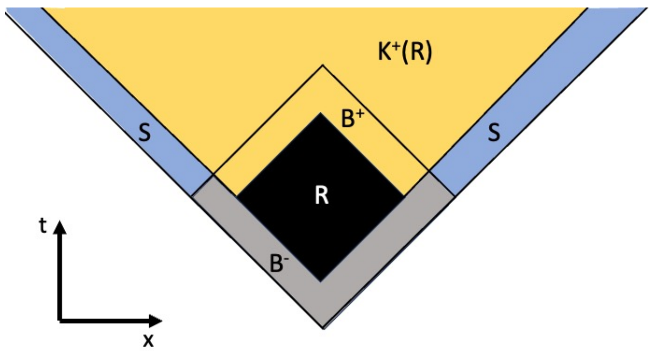
1. B_R^- is a pure thick parent of R.

[61] 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 $\overline{J^+(A)}$.

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:



For any X, Y, Z, let $X \perp\!\!\!\perp Y | Z$ denote that Z screens off X and Y (relative to $\text{urch}_{\mathcal{M}}$). The following are true by the probability laws (the terminology is due to Pearl (1985)):[62]

[62] *Proof*: For notational convenience, introduce for each region a homonymous variable, whose values represent the region's possible maximal intrinsic properties. Contraction is immediate: given $X \perp\!\!\!\perp Y | Z$ and $X \perp\!\!\!\perp W | Z \cup Y$, we have, for all values x, y, z, w:

$$\begin{aligned} \text{urch}_{\mathcal{M}}(X = x | W = w \wedge Y = y \wedge Z = z) &= \text{urch}_{\mathcal{M}}(X = x | Y = y \wedge Z = z) \\ &= \text{urch}_{\mathcal{M}}(X = x | Z = z). \end{aligned}$$

Weak Union follows from the multiplication axiom, plus the assumption that, if $X \perp\!\!\!\perp Y \cup W | Z$, then $\text{u}_{|wz}(y) := \text{urch}_{\mathcal{M}}(Y = y | W = w \wedge Z = z)$ is a well-defined probability measure on a sigma-algebra over Y's range, $\mathcal{R}(Y)$, for all for all z, w. For suppose that $\text{urch}_{\mathcal{M}}(X = x | Y = y \wedge W = w \wedge Z = z) = \text{urch}_{\mathcal{M}}(X = x | Z = z)$ for all x, y, z, w. Then

$$\begin{aligned} \text{urch}_{\mathcal{M}}(X = x | W = w \wedge Z = z) &= \int_{\mathcal{R}(Y)} \text{urch}_{\mathcal{M}}(X = x | W = w \wedge Y = y \wedge Z = z) \text{du}_{|wz}(y) \\ &= \text{urch}_{\mathcal{M}}(X = x | Z = z) \cdot \int_{\mathcal{R}(Y)} \text{du}_{|wz}(y) \\ &= \text{urch}_{\mathcal{M}}(X = x | Z = z) \\ &= \text{urch}_{\mathcal{M}}(X = x | Y = y \wedge W = w \wedge Z = z). \ \blacksquare \end{aligned}$$

- **Contraction:** If $X \perp\!\!\!\perp Y | Z$ and $X \perp\!\!\!\perp W | Z \cup Y$, then $X \perp\!\!\!\perp W \cup Y | Z$.
- **Weak Union:** If $X \perp\!\!\!\perp Y \cup W | Z$, then $X \perp\!\!\!\perp Y | Z \cup W$.

Below we show the following three facts:

$$R \perp\!\!\!\perp R^{\perp} \mid B \cup S \cup (K^{+}(R) \backslash B), \tag{36}$$

$$R \perp\!\!\!\perp (K^{+}(R) \backslash B) \mid B \cup S, \tag{37}$$

$$R \perp\!\!\!\perp S \mid B. \tag{38}$$

Once these are proven, the result follows: facts 37 and 38 entail, by Contraction,

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

Contracting again with fact 36 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 36: 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 $R \cup K^{+}(R) = J^{+}(R)$.[63] Moreover, since S and B_R^- are disjoint from R and subsets of a

[63]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.

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 $c([0, \tau_1[) \nsubseteq V^{\perp}$, $c^{-1}(V)$ 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 $c([0, s_1[) \nsubseteq P'^{\perp}$, $c^{-1}(P')$ 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 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 37: 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 $K^+(R) \backslash B_R^+$.[64] 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 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 38: 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 $R \cup (B_R^+)$.[65] Since R is causally

---

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 $V \cup P'$. ■

[64] 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 $V \backslash C$. ■

[65] 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 convex, $R \cup B_R^+ \subseteq J^+(R)$ doesn't cause $B_R^- \subseteq K^-(R)$. 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, $R \perp\!\!\!\perp S | B$. ■

One final lemma, which shows that the union of two causally convex regions is screened off by the union's thick boundary.

> **Lemma 4:** For any two regions R and R', and thick boundaries B of R and B' of R', $B \cup B' \backslash (R \cup R')$ is a thick boundary of $R \cup R'$. Moreover, if B screens off R from $R^\perp$ and B' screens off R' from $R'^\perp$, then $B \cup B' \backslash (R \cup R')$ screens off $R \cup R'$ from $(R \cup R')^\perp$.

Proof of Lemma 4: Let R and R' be two regions, and let B and B' be thick boundaries of R and R', respectively. Let $N \subseteq R \cup B$ and $N' \subseteq R' \cup B'$ be thick neighborhoods of R and R', respectively, and let c be a continuous curve starting in $(N \cup N')^\perp$ and ending in $R \cup R'$. Without loss of generality, suppose c ends in R. Since N is a thick neighborhood of R, there are $a, b \in [0,1]$ such that $c(]a, b[) \subseteq N \backslash R$ and $c([0, b[) \subseteq R^\perp$. If $c([0, b[) \subseteq R'^\perp$, then $c|_{]a,b[}$ is a non-trivial subcurve in $N \backslash R' \subseteq (N \cup N') \backslash (R \cup R')$ which c has before ever intersecting $R \cup R'$. If instead $c([0, b[) \nsubseteq R'^\perp$, $c^{-1}(R'^\perp)$ has an infimum $x \in]0, b[$. But since N' is a thick neighborhood of R', there's a $d \in [0, x[$ such that $c(]d, x[) \subseteq N' \backslash R$ and $c([0, x[) \subseteq R'^\perp$. Since also $c([0, x[) \subseteq R^\perp$, $c|_{]d,x[}$ is a non-trivial subcurve in $N' \backslash R \subseteq (N \cup N') \backslash (R \cup R')$ which c has before ever intersecting $R \cup R'$. In both cases, c has a non-trivial subcurve in $(N \cup N') \backslash (R \cup R')$ before ever intersecting $R \cup R'$. So $(N \cup N') \backslash (R \cup R')$ is a thick neighborhood of $R \cup R'$. But since $N \subseteq R \cup B$ and $N' \subseteq R' \cup B'$, $(N \cup N') \backslash (R \cup R') \subseteq (B \cup B') \backslash (R \cup R')$, and so $B \cup B' \backslash (R \cup R')$ is a thick boundary of $R \cup R'$.

For the second part, note that by assumption,

---

in Q or in V. We show that, in each case, c has a non-trivial subcurve in P before ever intersecting $Q \cup V$. Suppose c ends in Q. Since P is a thick parent of Q and c starts in $Q^\perp$, c then 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 hence $\inf(c^{-1}(Q)) < \inf(c^{-1}(V))$. Hence c has a non-trivial subcurve in P before ever intersecting $Q \cup V$. Suppose instead 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 $Q \cup V$. Suppose $c([0, b[) \nsubseteq Q^\perp$. Then there is a $q \in [0, b[\cap Q$. Thus $c|_{[0,q[}$ 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 $Q \cup V$. ■

$$R \perp\!\!\!\perp R^{\perp} | B, \tag{39}$$
$$R' \perp\!\!\!\perp R'^{\perp} | B'. \tag{40}$$

By Weak Union, facts 39 and 40 yield:

$$R \perp\!\!\!\perp R^{\perp} \mid B \cup (B' \backslash R) \cup (R' \backslash R) \cup (R \cap R'),$$
$$R' \perp\!\!\!\perp R'^{\perp} \mid B' \cup (B \backslash R') \cup (R \backslash R') \cup (R' \cap R).$$

Since $B \cap R = B' \cap R' = \emptyset$, we can rewrite this as follows:

$$R \perp\!\!\!\perp R^{\perp} | (B \cup B') \backslash (R \cup R') \cup R',$$
$$R' \perp\!\!\!\perp R'^{\perp} | (B \cup B') \backslash (R \cup R') \cup R.$$

Since $(R \cup R')^{\perp} \subseteq R^{\perp}$ and $(R \cup R')^{\perp} \subseteq R'^{\perp}$, these two facts imply the following:

$$R \perp\!\!\!\perp (R \cup R')^{\perp} | (B \cup B') \backslash (R \cup R') \cup R', \tag{41}$$
$$R' \perp\!\!\!\perp (R \cup R')^{\perp} | (B \cup B') \backslash (R \cup R') \cup R. \tag{42}$$

The following is also true in our system:[66]

- **Intersection:** If $X \perp\!\!\!\perp Y | Z \cup W$ and $X \perp\!\!\!\perp W | Z \cup Y$, then $X \perp\!\!\!\perp Y \cup W | Z$.

Applying Intersection to facts 41 and 42 yields the desired result:

[66] *Proof:* Suppose $\text{urch}_{\mathcal{M}}(X = x | Y = y \wedge W = w \wedge Z = z) = \text{urch}_{\mathcal{M}}(X = x | Y = y \wedge Z = z)$ and $\text{urch}_{\mathcal{M}}(X = x | Y = y \wedge W = w \wedge Z = z) = \text{urch}_{\mathcal{M}}(X = x | W = w \wedge Z = z)$ for all x, y, z, w. It follows that $\text{urch}_{\mathcal{M}}(X = x | Y = y \wedge Z = z) = \text{urch}_{\mathcal{M}}(X = x | W = w \wedge Z = z)$ for all x, y, z, w. Let $\text{u}_{|z}(y) := \text{urch}_{\mathcal{M}}(Y = y | Z = z)$. Then, assuming $\text{u}_{|z}(y) := \text{urch}_{\mathcal{M}}(Y = y | Z = z)$ is a well-defined probability measure on a sigma-algebra over Y's range, $\mathcal{R}(Y)$,

$$\begin{aligned}
\text{urch}_{\mathcal{M}}(X = x | Z = z) &= \int_{y \in \mathcal{R}(Y)} \text{urch}_{\mathcal{M}}(X = x | Y = y \wedge Z = z) \text{du}_{|z}(y) \\
&= \int_{y \in \mathcal{R}(Y)} \text{urch}_{\mathcal{M}}(X = x | W = w \wedge Z = z) \text{du}_{|z}(y) \\
&= \text{urch}_{\mathcal{M}}(X = x | W = w \wedge Z = z) \\
&= \text{urch}_{\mathcal{M}}(X = x | Y = y \wedge W = w \wedge Z = z). \quad \blacksquare
\end{aligned}$$

$$R \cup R' \perp\!\!\!\perp (R \cup R')^{\perp} | (B \cup B') \backslash (R \cup R'). \ \blacksquare$$

Say that a region R is *tolerantly causally convex* iff it is causally convex and every thick boundary B of R contains a thick boundary B' of R such that $R \cup B'$ is causally convex. A corollary of Theorem 2 and Lemma 4 is a restriction of Boundary Markov to finite unions of tolerantly causally convex regions:

> ***Corollary* 2:** Parental Markov entails that, for any finite union U of tolerantly causally convex regions, and any thick boundary B of U, B screens off U from $U^{\perp}$.

Proof of Corollary 2: Suppose Parental Markov, and let $\{R_1, ..., R_n\}$ be a set of tolerantly causally convex regions. Let B be a thick boundary of $U := \bigcup_{i=1}^{n} R_i$. We can choose thick boundaries $B'_1, ..., B'_n$ for $R_1, ..., R_n$ such that $B = \bigcup_{i=1}^{n} B'_i \backslash \bigcup_{i=1}^{n} R_i$. Since $R_1, ..., R_n$ are all tolerantly causally convex, $R_1, ..., R_n$ are all causally convex and there are thick boundaries $B_1, ..., B_n$ for $R_1, ..., R_n$ such that, for all $i = 1, ..., n$, $B_i \subseteq B'_i$ and $B_i \cup R_i$ is causally convex. By Theorem 2, then, for all $i = 1, ..., n$, B_i screens off R_i from $R_i^{\perp}$. Applying Lemma 4 n times, we get that $\bigcup_{i=1}^{n} B_i \backslash \bigcup_{i=1}^{n} R_i =: B^*$ is a thick boundary of U and screens off U from $U^{\perp}$. Since $B^* \subseteq B \subseteq U^{\perp}$, we have $U \perp\!\!\!\perp B \cup U^{\perp} | B^*$, and hence by Weak Union $U \perp\!\!\!\perp U^{\perp} | B$. $\blacksquare$

The continuity assumption on urch comes in at the next step. It is this:

> ***Thesis.* Weak Continuity:** Where X and Y are any regions, $(R_1, R_2, ...)$ a countable set of regions, and $\bigwedge_{i=1}^{\infty} R_i = r_i$ a nomically possible maximal intrinsic state of $\bigcup_{i=1}^{\infty} R_i$,
>
> $$\lim_{n \to \infty} \mathrm{urch}_{\mathcal{M}} \left(X = x | \bigwedge_{i=1}^{n} R_i = r_i \wedge Y = y \right) = \\ = \mathrm{urch}_{\mathcal{M}} \left(X = x | \bigwedge_{i=1}^{\infty} R_i = r_i \wedge Y = y \right),$$
>
> provided there is a $k \in \mathbb{N}$ such that, for all $l > k$, $\mathrm{urch}_{\mathcal{M}} \left(X = x | \bigwedge_{i=1}^{l} R_i = r_i \wedge Y = y \right)$ is well-defined.

Weak Continuity says that the urchance, conditional on any nomically possible maximal state of a region, is equal to the limit of conditioning the urchance on the maximal intrinsic states of increasingly larger parts of the region (and that, except in cases where infinitely many members of the sequence are ill-defined, the limit exists). If the dynamical laws are local and the maximal properties of a region are *separable*, Weak Continuity is extremely plausible. Separability of maximal properties is another locality assumption we shall adopt throughout.

Given Weak Continuity, Corollary 2 implies this:

> ***Corollary 3*:** Parental Markov entails that, for any *countable* union $\mathbf{R}$ of tolerantly causally convex regions, and any thick boundary $\mathbf{B}$ of $\mathbf{R}$, $\mathbf{B}$ screens off $\mathbf{R}$ from $\mathbf{R}^{\perp}$.

Proof of Corollary 3: Suppose Parental Markov is true. Let $\{R_1, R_2, ...\}$ be a countable set of tolerantly causally convex regions, where for any $i \in \mathbb{N}$, R_i has thick boundary B_i. Denote $\mathbf{R}_n := \bigcup_{i=1}^n R_i$ and $\mathbf{B}_n = \bigcup_{i=1}^n B_i \backslash \bigcup_{j=1}^n R_j$. Abbreviate $\mathbf{R}_\infty =: \mathbf{R}$ and $\mathbf{B}_\infty =: \mathbf{B}$. For any $n \in \mathbb{N}$, $\mathbf{B}_n$ is a thick boundary of $\mathbf{R}_n$, and $\mathbf{B}$ is a thick boundary of $\mathbf{R}$. In the following, let $\mathbf{r}^{\perp}$ and $\mathbf{r}$ be intrinsic states of $\mathbf{R}^c$ and $\mathbf{R}$, respectively. Let $\mathbf{b}$ be a nomically possible maximal intrinsic state of $\mathbf{B}$. I'll denote by $\mathbf{b} \cup \mathbf{r}$ the state of $\bigcup_{i=1}^\infty (B_i \cup R_i)$ entailed by $\mathbf{B} = \mathbf{b} \wedge \mathbf{R} = \mathbf{r}$, and by $\mathbf{b}_n$ and $\mathbf{r}_n$ the states of $\bigcup_{i=1}^n B_i \backslash \bigcup_{j=1}^n R_j$ and $\bigcup_{i=1}^n R_i$, respectively, entailed by $\mathbf{B} = \mathbf{b} \wedge \mathbf{R} = \mathbf{r}$. Finally, I'll denote by $\mathbf{b_n} \cup \mathbf{r}_n$ the state of $\bigcup_{i=1}^n (B_i \cup R_i)$ entailed by $\mathbf{B} = \mathbf{b}_n \wedge \mathbf{R} = \mathbf{r}_n$. Note that, by separability, $\bigwedge_{n=1}^\infty \mathbf{R}_n = \mathbf{r}_n$ is *equivalent* to $\mathbf{R} = \mathbf{r}$, and similarly for other conjunctions. We then have the following:

$$
\begin{aligned}
\text{urch}_{\mathcal{M}}(\mathbf{R}^{\perp} = \mathbf{r}^{\perp} | \mathbf{B} = \mathbf{b} \wedge \mathbf{R} = \mathbf{r}) &= \text{urch}_{\mathcal{M}} \left(\mathbf{R}^{\perp} = \mathbf{r}^{\perp} | \left(\bigcup_{i=1}^{\infty} B_i \backslash \bigcup_{j=1}^{\infty} R_i \right) = \mathbf{b} \wedge \bigcup_{k=1}^{\infty} R_k = \mathbf{r} \right) \\
&= \text{urch}_{\mathcal{M}} \left(\mathbf{R}^{\perp} = \mathbf{r}^{\perp} | \bigcup_{i=1}^{\infty} (B_i \cup R_i) = \mathbf{b} \cup \mathbf{r} \right) \\
&\overset{\text{WC}}{=} \lim_{n \to \infty} \text{urch}_{\mathcal{M}} \left(\mathbf{R}^{\perp} = \mathbf{r}^{\perp} | \bigcup_{i=1}^{n} (B_i \cup R_i) = \mathbf{b}_n \cup \mathbf{r}_n \right) \\
&= \lim_{n \to \infty} \text{urch}_{\mathcal{M}} \left(\mathbf{R}^{\perp} = \mathbf{r}^{\perp} | \left(\bigcup_{i=1}^{n} B_i \backslash \bigcup_{j=1}^{n} R_j \right) = \mathbf{b}_n \wedge \bigcup_{k=1}^{n} R_k = \mathbf{r}_n \right) \\
&\overset{\text{Cor. 1}}{=} \lim_{n \to \infty} \text{urch}_{\mathcal{M}} \left(\mathbf{R}^{\perp} = \mathbf{r}^{\perp} | \bigcup_{i=1}^{n} B_i \backslash \bigcup_{j=1}^{n} R_j = \mathbf{b}_n \right) \\
&\overset{\text{WC}}{=} \text{urch}_{\mathcal{M}} \left(\mathbf{R}^{\perp} = \mathbf{r}^{\perp} | \mathbf{B} = \mathbf{b} \right),
\end{aligned}
$$

where the fifth step follows because $\mathbf{R}^{\perp} \subseteq (\mathbf{R}_n)^{\perp}$. ■

> ***Theorem 3*:** In Minkowski spacetime, Parental Markov entails Boundary Markov.

Proof of Theorem 3: The proof relies on three facts. First, Minkowski spacetime is *normal*: every two disjoint closed sets have disjoint neighborhoods. Second, Minkowski spacetime is *Lindelöf*: every open cover has a countable subcover. Third, "arbitrarily small" tolerantly causally convex regions are ubiquitous: for every point p in Minkowski spacetime and every neighborhood A of p, there is a tolerantly causally convex open region C with $p \in C \subseteq A$.[67]

Let now B be a thick boundary of R. Then $R \cup B$ contains a neighborhood N of R's closure. We now show the following:

> *Claim*: There's a countable union U of tolerantly causally convex regions such that $R \subseteq U \subseteq N$ and $N \backslash U$ is a thick boundary of R.

This Claim, together with Parental Markov, entails Boundary Markov. For by Corollary 3, Parental Markov implies that any countable union of tolerantly causally convex regions is screened off by its thick boundary. So, together with Claim, $U \perp\!\!\!\perp U^{\perp} | N \backslash U$. But since $R \subseteq U$, Weak Union implies $R \perp\!\!\!\perp U^{\perp} | N \backslash R$. Since $U^{\perp} \cup (N \backslash R) = R^{\perp}$ we have $R \perp\!\!\!\perp R^{\perp} | N \backslash R$. Done.

To prove the Claim, note that $N^{\perp}$ is closed and disjoint from $\overline{R}$. Hence, by the first fact (normality), there are disjoint neighborhoods N' of $N^{\perp}$ and N'' of R. (Pictorially, $N' \backslash N^{\perp}$ is an open "inner lining" of N, and $N'' \backslash R$ is an open "outer lining" of R.) By the third fact, N'' contains an open cover of R consisting of tolerantly causally convex regions. Hence, by the second fact (Lindelöf), N'' contains a *countable* open cover U of N'' consisting of tolerantly causally convex regions. Moreover, since $N' \backslash N^{\perp}$ is a thick boundary of $N^{\perp}$, and N' is disjoint from N'', $N \backslash N'$ is a thick boundary of N'. Since $U \subseteq N'$, $N \backslash U$ is thus a thick boundary of U. So U is a countable union of tolerantly causally convex regions such that $R \subseteq U \subseteq N$ and $N \backslash U$ is a thick boundary of R, which proves the Claim. ■

Appendix 4: 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. 33 and 34 are then more compactly written (addition in i is modulo n)

[67] To see this, note that any initial segment of a point's causal future, future-ward bounded by a space-like surface, is causally convex. Given a point p and a neighborhood C of p, there's always a point q in $\langle J^{-}(p) \rangle \cap C$ such that the interior of some initial segment of $J^{+}(q)$ is in C and contains p. This interior is obviously causally convex, and moreover tolerantly so, since we can always find initial segments of past points' causal futures that envelop it arbitrarily closely.

as follows:

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

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

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

which can be rearranged to

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

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

$$
\mathbf{M} \cdot \hat{\mathbf{p}} = \hat{\mathbf{v}}, \tag{43}
$$

where

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

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

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

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

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

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

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

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

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

and

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

In CIRCLE, because every color j only has itself and color $j+1$ as permissible successors, all entries of $\mathbf{Q}_{k-1}^{i}$ besides the first row are 0, and all entries of $\mathbf{R}_{k-1}^{i}$ besides the diagonal and the first lower diagonal are 0. 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 latter point, consider the simplest non-trivial case, $n = k = 2$, i.e. a loop of two days, with two possible colors per day. Call this world SUPER SIMPLE CIRCLE, or SSC. In this case,

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

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

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

or

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

In the first case, the particle is guaranteed to switch color every time. Any probabilistically coherent assignment of marginals respecting $\mathrm{urch}_{SSC}(\tau_1 = 1) = \mathrm{urch}_{SSC}(\tau_2 = 2)$ 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 $\mathrm{urch}_{SSC}(\tau_1 = 1) = \mathrm{urch}_{SSC}(\tau_2 = 1)$ 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, J. (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.

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

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.

Halpern, Joseph and Judea Pearl (2005). “Causes and Explanations: A Structural-Model Approach. Part I: Causes”. In: *The British Journal for the Philosophy of Science* 56.4, pp. 843–887.

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.

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

Loewer, Barry (2001). “Determinism and Chance”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32.4, pp. 609–620.

Malament, David (2012). *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory*. Chicago Lectures in Physics. Chicago, IL: University of Chicago Press.

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.

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

— (2007). "Deterministic Chance?" In: *British Journal for the Philosophy of Science* 58.2, pp. 113–140.

Stockum, W. J. van (1938). "IX.—The Gravitational Field of a Distribution of Particles Rotating about an Axis of Symmetry". In: *Proceedings of the Royal Society of Edinburgh* 57, pp. 135–154.