From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Philosophy

2021

On the Epistemology of Physical Probability

Meehan, Alexander

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01qz20sw575

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

On the Epistemology of Physical Probability

Alexander Meehan

A DISSERTATION PRESENTED TO THE FACULTY OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY THE DEPARTMENT OF PHILOSOPHY

Advisers: Hans Halvorson and Adam Elga

January 2021

© Copyright by Alexander Meehan, 2020.

All rights reserved.

Abstract

Physical probabilities or *chances* exist and come in many types. Some feature in fundamental physics, others feature in the special sciences. What unites them, on my view, is their epistemological role. Part I of this dissertation lays out this role. It expands David Lewis's *principal principle* to incorporate chances from special sciences like biology as well as fundamental physics, and discusses how we know these chances and why that knowledge is so valuable. Part II is concerned specifically with the chances in quantum mechanics and a problem that neither philosophers nor physicists have yet to appreciate. This problem, which I call the *control problem*, contrasts from the well-known measurement problem, and places new fundamental constraints on our quantum theories. Part III explores the limits of our mathematical models of chance. It focuses on the role of probabilistic independence in chance judgments, discussing puzzles with certain of those judgments, as well as with the traditional mathematical definition of probabilistic independence.

Acknowledgments

Chapters 5, 6 and 7 overlap with the following papers:

- "A new problem for quantum mechanics"
  The British Journal for the Philosophy of Science. Forthcoming.
- "The Borel-Kolmogorov paradox is your paradox too" (with Snow Zhang)
  Philosophy of Science. Forthcoming.
- "You say you want a revolution: Two notions of probabilistic independence"
  Philosophical Studies. Accepted for publication Oct 2020.

I am grateful to these journals and to Snow Zhang for their permission to include these materials here.

This work would not have been possible without the insights of my advisers, Hans Halvorson and Adam Elga. Hans inspired me to look outside the box to less explored questions, and our conversations gave me many of the ideas in this dissertation, especially in the area of the no-cloning theorem. Adam gave incisive comments on a countless number of drafts, and helped guide and sharpen my thoughts on every topic here. Hans and Adam were also tremendous mentors in other ways. Adam never hesitated to give valuable practical advice whenever asked, and helped me become a better teacher. Hans guided me through various stages of my graduate career, and I always walked away from our conversations feeling more motivated to continue my work. I would like to thank Hans and Adam for their consistent support and advice.

I also had the undeserved fortune of gaining supportive advisers from outside of Princeton. Laura Ruetsche kindly welcomed me to Michigan in fall 2018, where I benefited a lot from her guidance. Even after my visit there ended, she went above and beyond providing more feedback and advice, and Part II of this dissertation could not have taken shape without Laura's insights. Another extremely positive consequence of visiting Michigan was getting to know Jim Joyce, who provided invaluable advice on my work on conditional probability, especially the work in Part III. In another inter-departmental visit, this time to MIT in 2020-1, I benefited greatly from conversations with Jack Spencer, who helped sharpen my ideas for Part I, and who also provided a lot of helpful general advice about my work, and the framing of this dissertation.

I am also indebted to my 'unofficial' advisers. I would especially like to thank Thomas Barrett and Ben Feintzeig, who went out of their way to give me multiple rounds of feedback and advice on various projects, especially in the early stages of my PhD. I would also like to thank Nina Emery, who got me interested in these topics in the first place, for her continuing support and advice. I also thank Branden Fitelson, who went out of his way to make me feel welcome in the Boston philosophy scene during my visit to MIT.

Perhaps the person with the biggest impact on my philosophical development throughout graduate school is Snow Zhang. When I reflect on whether I regret going to philosophy grad school, or choosing to come to Princeton, I conclude that I cannot, because it led to my collaboration with Snow. I am grateful for Snow's openness to working with me, and I hope our collaboration continues on.

Each part of this dissertation has benefited substantially from the comments and advice of others. For his helpful comments on the end product, I'd like to thank John Burgess. For their feedback on various stages of Part I, I'd like to thank Kevin Blackwell, Richard Bradley, Eddy Chen, Dmitri Gallow, Barry Loewer, Mike Hicks, Josh Hunt, Boris Kment, Chris Meacham, Ezra Rubenstein, Brad Skow, Jack Spencer, Michael Strevens, Isaac Wilhelm, Al Wilson, and audiences at CMU, Michigan, MIT, Rutgers, and University of Colorado, Boulder. For their feedback on various stages of Part II, I'd like to thank Gordon Belot, Jeff Bub, Neil Dewar, Ben Feintzeig, David Schroeren, Jeremy Steeger, and audiences at City College of New York, Michigan, Princeton, University of Chicago, and the University of Washington. For their feedback on various stages of Part III, I'd like to thank David Builes, Branden Fitelson, Alan Hájek, Michele Odisseas Impagnatiello, Michael Rescorla, Brad Skow, Haley Schilling, and audiences at MIT and Princeton.

For their friendship, advice, and great conversation (philosophical and otherwise), I'd also like to thank Colin Bradley, Neil Dewar, Sam Fullhart, Josh Hunt, Brett Karlan, Daniel Krandzelbiner, Kyle Landrum, Calum McNamara, Patrick Miller, Joe Moore, Abby Plummer, Chris Register, Ezra Rubenstein, Alejandro Naranjo Sandoval, David Schroeren, Gabriel Shapiro, Aarthy Vaidyanathan, Isaac Wilhelm, Elise Woodard, Erik Zhang, the extremely welcoming community of grad students at MIT, the extremely welcoming community of grad students at Michigan, and the many others who I am forgetting to mention here.

Finally, I'd like to thank my family for their unwavering love and support. Special thanks to my parents, and to my wife, Ruthy, who was the one who initially encouraged me to follow my interests despite the risks, and stuck with me each step of the way. This work would not have been possible without them.

Contents

Abstract **iii**

Acknowledgments **iv**

Introduction **1**

I Expanding Lewisian chance **3**

1 A framework for non-fundamental chance **4**

1.1 The limits of Lewisian chance 6

1.2 Subject matters 9

1.3 Chance theories 11

1.4 The principal principle 15

1.5 Comparison to similar frameworks 22

1.6 Emergent dynamical chances 23

1.7 Alignment between chances of different kinds 26

1.8 The Mentaculus 29

1.9 Objections 30

2 Knowing the chances **37**

2.1 Knowing Lewisian chances 38

2.2 The emergence of chance structure 47

2.3 Resiliency and the value of chance knowledge 61

2.4 Conclusion 66

II Quantum chance **68**

3 States of ignorance and ignorance of states **69**

3.1 Earman on quantum chance . . . 71

3.2 Assessing Earman's argument . . . 73

3.3 Preparation, projective measurement, and a classical analogue . . . 75

3.4 Formulating the quantum principal principle . . . 77

3.5 Conclusion . . . 81

4 The impossibility of individual state determination, revisited **82**

4.1 The argument of Busch (1997) . . . 83

4.2 An improved argument . . . 84

4.3 The argument of D'Ariano and Yuen (1996) . . . 88

5 A new problem for quantum mechanics **92**

5.1 The standard measurement problem . . . 95

5.2 Quantum state preparation . . . 96

5.3 The control problem . . . 100

5.4 Relation to interpretations . . . 108

5.5 Relation to state determination . . . 113

5.6 Conclusion . . . 115

III Modeling chance, the limits **116**

6 The Borel-Kolmogorov paradox is your paradox too **117**

6.1 Uniqueness and Equivalence . . . 119

6.2 Is the formalism to blame? . . . 125

6.3 The urgency of the paradox . . . 126

6.4 Toward a resolution . . . 127

7 Two notions of probabilistic independence **130**

7.1 Declarations of independence . . . 131

7.2 Trouble in Popper's paradise . . . 135

7.3 Defending and amending Kolmogorov's framework . . . 138

7.4 Interlude: Comparing the amended theories . . . 143

7.5 A second notion of probabilistic independence 146
7.6 Conclusion 152

Appendices **154**

A Chapters 4 and 5 **155**

B Chapter 6 **159**

C Chapter 7 **163**

Introduction

Probabilistic claims like the following feature prominently in science and in ordinary life:

- The probability that an evenly weighted die lands 'six' is about 1/6.
- The probability that this uranium atom decays in the next year is less than 5%.
- The probability that a cross between two pea plants will yield an offspring that is heterozygous for both seed shape and seed color is approximately $1/2 \cdot 1/2 = 1/4$.
- The probability that atmospheric moist convection will be triggered over the Magdalena mountains today, leading to a late afternoon thunderstorm, is greater than 80%.

It is the philosopher's task to make deeper sense of these claims. One view is that the probabilities above are merely a reflection of our subjective ignorance. Intuitively, though, they seem to take the values they do in virtue of the way the *physical world* is, not in virtue of something about us or our minds. Nowadays, scientists and statisticians calculate these values using the principles of classical mechanics, quantum mechanics, molecular genetics, atmospheric dynamics, and other empirical theories. It has become increasingly natural to try to understand these probabilities as 'physical'.

I think *physical probabilities* or *chances* exist and come in many types. Some feature in physics, others feature in biology. Some may derive from frequencies, others may derive from propensities. What unites them, on my view, is not necessarily their specific metaphysical natures, but their epistemological role. In the first part of this dissertation, I lay out this role. I generalize David Lewis's *principal principle*, a principle connecting credence and chance, so that it can accommodate not only fundamental dynamical chances (those featuring in the universe's most fundamental dynamical laws), but chances of all types. In particular, I lay out a framework in which I distinguish between a 'local' and 'global' version of the principal principle, and I show how the global version induces constraints on the relationships between different kinds of chances. Next I apply the framework to the question of how agents like us come to learn chances. I point out that common methods, like counting frequencies, rely on the presence of *uniform chance structure*, which is often

absent at the level of the fundamental dynamical chances. However, it plausibly emerges at the level of less fundamental chances; elaborating on this claim, I use my framework to offer a general account of how we come to know physical probabilities, and why that knowledge is so valuable.

In the second part, I turn my attention to those probabilities featuring in quantum mechanics. Quantum chances raise special challenges for the epistemology of chance. I first examine an argument due to John Earman, that the principal principle, suitably formulated for quantum chances, is a *theorem* of quantum probability theory. I argue that while it's not a theorem, Earman is correct that on certain interpretations of quantum mechanics, it requires a special formulation. I lay out my positive proposal for this formulation, and use it to model how agents come to learn quantum chances. Next, I turn to consider a particular learning scenario that physicists and philosophers are interested in: when only a 'single copy' of the system is available. I evaluate the arguments from the physics literature for the impossibility of *individual state determination* in these scenarios, and point out that many overgeneralize: if sound, they would imply it is generally impossible to learn quantum states. I use my positive proposal above to offer an improved argument, but note a lingering worry. On the one hand, if the physicists' arguments were sound, they would imply this skeptical conclusion. On the other hand, it is very difficult to diagnose the error in one of these arguments (based on the *no-cloning theorem*). This worry leads me to introduce a foundational issue, previously unnoticed by physicists and philosophers, which I call the *control problem*. A huge effort has gone into solving the so-called measurement problem for quantum mechanics. The control problem is different. In brief, the problem reveals a tension between two natural claims (1) we have an ability to gain reliable information about quantum states, and (2) we have an ability to control the quantum states of systems. To conclude this second part, I lay out the problem and discuss its implications for quantum foundations.

In the third and final part, I return to considering physical probabilities in general, with a focus on puzzles and problems with our chance reasoning that arise when *continuous chancy quantities*, like position or speed, are involved. Drawing on joint work with Snow Zhang, I show how our usual *independence* reasoning (which plays an important role in calculations like $1/2 \cdot 1/2 = 1/4$ above) leads to conflicting answers in continuous cases. This kind of problem is usually attributed to how orthodox probability theory models 'probability zero' events, but I argue that there is a general puzzle for all theories, and propose a partial solution. Next I turn to consider a worry that these 'probability zero' independence questions raise for the orthodox definition of independence. Responding to Branden Fitelson and Alan Hájek, I defend the orthodox framework for modeling probability (due to Kolmogorov) from their criticism, arguing it can accommodate probability zero events and the notion of independence Fitelson and Hájek seek. I conclude there is also merit to the orthodox 'factorization' definition of independence, and I propose understanding this definition as capturing a kind of *almost sure* independence for continuous chancy quantities.

Part I

Expanding Lewisian chance

Chapter 1

A framework for non-fundamental chance

Many accounts of chance are uncompromising: they require objective chances be part of the universe's most fundamental dynamical laws. But this threatens to leave mysterious the indispensable use of chance reasoning in non-fundamental sciences like biology, atmospheric science, and statistical mechanics. In this chapter, I propose a generalized account which gives home to chances or physical probabilities of both types. The basic idea is that although the metaphysical nature of these chances may vary—some may derive from frequencies, others from propensities, and so on—they all guide and constrain belief in a unified way.

The idea that belief guidance is a characteristic feature of chance is not new. David Lewis (1980) famously proposed a principle connecting chance and belief, and suggested that it may "capture all we know about chance". This *principal principle*, in its roughest form, states that if an agent knows the objective chance of A is x, then she should set her credence in A to x. The thought is that while we might not know what chances are, we know they constrain rational credence in this way. Though it may seem like a small point, this turns out to be a very useful observation. The principle, once combined with certain substantive assumptions about rational credence, allows us to derive several interesting conclusions about chance. It also places an important desideratum on our metaphysical theorizing: an analysis of what objective chances are (whether it be a uniform analysis, or a disjunctive account) should support and explain the truth of this principle.

This is all on the right track. Where Lewis's account went wrong, in my view, is in the precise formulation of the principal principle. On Lewis's formulation, only chances that are part of the universe's most fundamental dynamical laws can satisfy it; we are forced into the kind of uncompromising account of chance described above. In particular, on this formulation, a chance must be such that if an agent knows its current

value, then *no matter what she has learned about the past, at any level of microphysical detail,* she should set her degree of belief to that value. As I argue in this chapter and the next, it's implausible that most chances in the special sciences play this role, and it's even an open question whether the chances in our current most fundamental physical theories do.

However, while chances in the special sciences may not play this role, they do satisfy an instance of the following general schema: ⟨if an agent knows the value, then no matter what she has learned about *a certain subject matter*, she should set her degree of belief to that value⟩. Lewis implicitly focused on one subject matter: the entire past, specified in maximal microphysical detail. But by taking a more expansive view of the matters that may fill in for the schema, and in particular by allowing for more macroscopic matters to do so, we may better accommodate special science chances. That is the starting point of the framework I offer here. It allows for a variety of chances—M-chances, N-chances, and so on—based on different subject matters, corresponding, roughly speaking, to different levels of specification of different sciences. While these chances are varied, they play a unified belief guiding role: If an agent knows the M-chance of A is x, then as long as she is only informed by matters in M, she should set her credence in A to x. Similarly, if an agent knows the N-chance of A is y, then as long as she is only informed by matters in N, she should set her credence in A to y. And so on for chances based on other subject matters.

This kind of relativization to subject matters raises several questions. For instance, say the M-chances are the biological chances and the N-chances are the fundamental dynamical chances. Then can the M-chance of A and the N-chance of A differ from one another and, if so, by how much? I mentioned that a nice feature of Lewis's principal principle is that, once combined with certain substantive assumptions about rational credence, it allows us to derive several interesting conclusions about chance. The current framework is also fruitful in this respect. In the first half of the chapter, after presenting the apparatus of subject matters and chance theories based on different subject matters, I introduce what I call the *global principal principle*. I show how this principle, combined with the assumption that rational credence is probabilistically coherent, induces non-trivial constraints on how chances based on different subject matters must align with one another. The result is a richer account of how different kinds of chances relate.

Several other authors have sought to expand Lewisian chance. Meacham (2009), Handfield and Wilson (2014), List and Pivato (2015), and Albert (2015) and Loewer (2020), have laid down interesting proposals for how to do so. And Lewis himself, in his discussion of *counterfeit chance* (1987), alluded to how this expansion could be done. My approach builds upon these authors'. In fact, it turns out that my framework captures all of these authors' proposals as special cases. In the second half of the chapter, I show how this is so, and respond to some objections that have been leveled against my and similar frameworks. In Chapter 2, I go on to apply my framework to more specific questions concerning the emergence of 'nice' chance structure

at the level of the special sciences, and our epistemic access to special science chances.

1.1 The limits of Lewisian chance

Before diving into the framework, I should say more about why expanding Lewisian chances may be necessary in order to incorporate special science probabilities. I will also elaborate further on this point in Chapter 2.

By *Lewisian chance*, I mean a chance that satisfies Lewis's principal principle: knowing its value screens off all information about the past, in the sense that no matter what one learns about the past or about the laws, one's credence should not deviate from the chance. At least intuitively, Lewisian chances are the kinds of chances that are part of the universe's most fundamental dynamical laws—the laws that describe how the fundamental state of the universe evolves over time.[1] To see the connection, note that Lewisian chances are incompatible with fundamental determinism. For suppose, toward contradiction, that the fundamental dynamical laws are deterministic and yet the Lewisian chance of A, for some proposition A, is non-trivial $0 < x < 1$. Then learning the initial condition of the universe, together with the laws, should not lead one's credence to deviate from x. But by assumption the initial condition together with the laws entail whether A is true.[2] So one's credence in A should shift to either 0 or 1, which contradicts our earlier assumption.

In what follows, I will use the terms *Lewisian chance* and *fundamental dynamical chance* interchangeably. Note that 'fundamental' is used here in the sense of physically fundamental (part of the fundamental physical theory), not in the sense of metaphysically fundamental or non-derivative.

One way to see that expanding Lewisian chance may be necessary in order to incorporate special science probabilities is to notice that deterministic worlds may contain non-trivial special science probabilities.

Example 1 (Deterministic wheel of fortune)**.** Imagine a world containing a group of roaming (free) particles, and a wheel of fortune with two possible results, WHITE and BLACK. For simplicity we suppose that the particles start with the same spatial trajectory, heading straight toward the wheel, as shown in the left part of Figure 1.1. The particles hit the wheel and cause it to spin. The hits are sufficiently intermittent, such that the wheel always has time to come to a halt before the next particle arrives. We can think of the result of a given 'trial' or spin as a function of the speed of the incoming particle and the starting position of the wheel. If there is sufficient variety in these incoming speeds, the wheel will land WHITE approximately 50% of the time and BLACK approximately 50% of the time. Nevertheless, there isn't a non-trivial fundamental dynamical chance of WHITE and BLACK on any given trial; the fundamental dynamical chance is always 0 or 1, depending on the exact speed of the incoming particle and the precise starting position of the wheel.

[1] I say "at least intuitively" because of complications about whether there is fundamentally a distinction between past, present and future that is independent of a reference frame (and indeed, whether time itself is physically fundamental or emergent).

[2] Here it may be important that A is a proposition that is entirely about matters of particular fact; cf. below.

Scientists in this world who observe the wheel will see a 'chancy' pattern. There is a sense in which, in virtue of the dynamics of the wheel and the varying speed of the particles, the physical probability of WHITE in this world is approximately 50%. But whatever this physical probability is, it is not a Lewisian chance. All Lewisian chances are either 0 or 1.

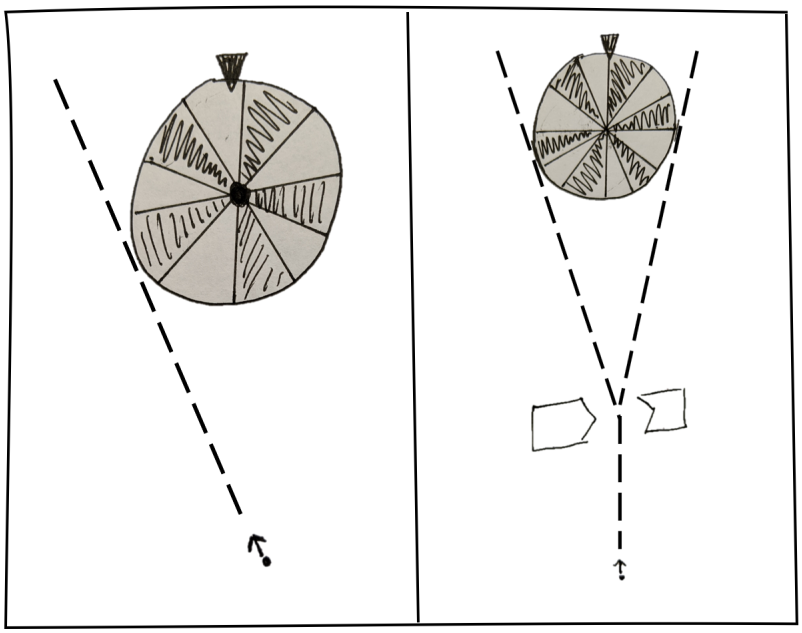


Figure 1.1: The wheel of fortune case. Left: deterministic version. Right: indeterministic version.

The previous example involved a deterministic world, which is a very special kind of case. But in a fundamentally indeterministic world, won't there be a Lewisian chance that equals the special science chance? The answer is no, not necessarily. In fact the Lewisian chances, even if non-trivial, may have nothing to do with certain intuitively chancy patterns.

Example 2 (Indeterministic wheel of fortune)**.** Imagine another wheel of fortune world which is otherwise very similar, but whose dynamical laws feature the kind of indeterminism seen in (certain interpretations of) quantum mechanics. In particular, imagine the world is configured such that the particles pass through a magnet on their way to the wheel, as shown in the right side of Figure 1.1. The fundamental dynamical chances are such that with 30% chance a particle will go left through the magnet and with 70% chance it will go right. If the particle goes left, it will hit the left side of the wheel, causing it to spin clockwise. If it goes right, it will hit the right side of the wheel, causing it to spin counter-clockwise. Other than their interactions with the magnet, the particles move deterministically as before.

Reflecting on this second set-up, one realizes that with reflect to the overall pattern of WHITE and BLACK outcomes, the addition of the magnet will not make a difference; as long as the speed of incoming particles

is sufficiently varied, the wheel will land WHITE about 50% of the time and BLACK about 50% of the time. The fact that about 30% of the spins are clockwise and about 70% of the spins are counterclockwise should not affect these statistics. Here, although there are non-trivial Lewisian chances, these chances have nothing to do with the chancy pattern exhibited by the wheel.

This is a toy example, but arguably less extreme versions of this kind of phenomenon occur in many of the special science probabilities listed in the Introduction. Although the die, our fingers, and the air in the room may be governed by chancy fundamental dynamical laws, that does not mean that those laws will give rise to the familiar 1/6 chance distribution over the toss outcomes (and given what we know about the physics, this seems highly unlikely, at least if we focus on the time just before the throw (Diaconis et al., 2007)). Similarly, although the matter in the atmosphere may be governed by chancy fundamental dynamical laws, that does not mean those Lewisian chances will percolate up in the right way to constitute the atmospheric scientists' chance distribution over atmospheric convection events. In sum, someone committed to:

Fundamentalism All objective chances are fundamental dynamical chances (i.e. Lewisian chances),

may face serious difficulties incorporating special science probabilities as objective chances.

Here it is important to distinguish **Fundamentalism** from the weaker claim:

Weak Fundamentalism All objective chances feature in the fundamental physical theory.

A fundamental physical theory may include more than a specification of dynamical laws. I have not said anything which would rule out special science probabilities arising as part of the fundamental physical theory more generally. Indeed, some authors, such as Albert (2015) and Loewer (2020), are inclined to build in a statistical mechanical measure over initial conditions as a fundamental non-dynamical (Humean-style) law. Since these authors think that statistical mechanical probabilities can account for all special science probabilities (cf. Section 1.8) they are an example of theorists who would reject **Fundamentalism** but endorse **Weak Fundamentalism**. I will stay neutral on the status of **Weak Fundamentalism**.

At this stage, the reader may be having the following reaction: I am not worried that endorsing **Fundamentalism** leads to difficulties with incorporating these special science probabilities as objective chances, because these special science probabilities are not objective chances. At the very least, they are not the kind of probability that Lewis set out to investigate when he developed his account of objective chance.

I agree with the last claim. It seems Lewis was primarily interested in the kinds of chances that appear in theories like quantum mechanics. But while special science probabilities may not have been the focus of Lewis's theorizing, they are the focus of my theorizing. As I discussed in the Introduction, it is natural to try to understand many of these probabilities as physical. Or at least, as a matter of philosophical

methodology, it seems worth investigating whether this kind of account can work, before moving to more skeptical alternatives. Indeed, I won't argue against **Fundamentalism** here, except to point to my account, in which special science probabilities govern rational credence in a parallel way to Lewisian chances, and in which dynamical chances from the special sciences play an especially similar functional role to Lewisian chances.[3] If it looks like chance, guides belief like chance, and is physical like chance, why not consider it chance? Of course, for historical reasons, some may want to reserve the term 'objective chance' for the kind of probabilities that Lewis was interested in. I am happy to cede the term 'objective chance' in favor of 'physical probability'. But in either case, it seems worth investigating whether a more expansive account of physical probability is possible. With this in mind, let us turn to the first ingredient of my framework.

1.2 Subject matters

The first ingredient of the framework, following Handfield and Wilson (2014), is another Lewisian apparatus, namely that of subject matters.

Let $\mathcal{W}$ denote the set of metaphysically possible worlds. I will define a *subject matter* as a partition of $\mathcal{W}$ (Lewis, 1988a,b). The most detailed subject matter is the way the world is, represented by the maximally fine partition, $\{\{w\} : w \in \mathcal{W}\}$. By slight abuse of notation, I will also denote this subject matter by $\mathcal{W}$.

In this section, I will be especially focused on *factual subject matters*. Let H denote the partition of $\mathcal{W}$ into complete world-histories. So for each world w, the cell of H containing w—$H(w)$—specifies all the matters of particular fact for all times at w (see Figure 1.2). For example, if w is a simple Newtonian world, then $H(w)$ specifies the entire spacetime trajectory of every individual particle in w. Clearly, H is a very detailed subject matter. It is a controversial metaphysical question whether H is identical to $\mathcal{W}$; does the way the world is depend entirely on matters of particular fact, or does it depend on more than that? (For instance, does the way the Newtonian world is supervene entirely on the spacetime trajectories of all the particles, or also on some governing law, power, or other non-Humean whatnot?) In either case, let us define a *factual subject matter* M as one that is entirely about matters of particular fact. In particular, let us define a factual subject matter as one that is *included in* H, in the sense that every cell of M can be written as a disjunction or union of cells of H. I will denote this inclusion relation by $\preceq$.

Here are some examples of factual subject matters and interesting inclusion relations between them.

Example 3 (Maximally detailed histories)**.** Fix a time t. We can then consider the subject matter H_t of the history of the world up to t. Clearly, $H_t \preceq H$. Furthermore, if $t \leq t'$, then $H_t \preceq H_{t'}$. This is because

[3]See Section 1.6. For more on this point, see also List and Pivato (2015).

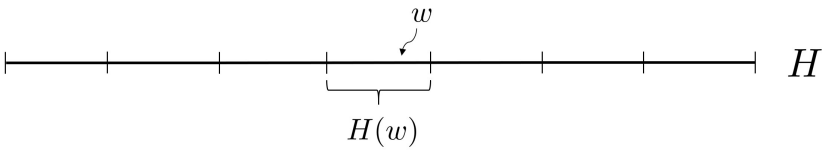


Figure 1.2: The subject matter H, represented as a partition of $\mathcal{W}$.

any specification of matters of particular fact up to the earlier time t can be formulated as a disjunction of matters of particular fact up to the later time t'.

Example 4. The 17th century is a subject matter. Denote it by N. Two worlds belong to the same cell of N, in this case, if and only if their 17th centuries are exact duplicates, or if neither contains a 17th century. The events of 17th century French politics is also a subject matter. Denote it by M. Then $M \preceq N \preceq H$.

Example 5 (Statistical mechanics)**.** Let N denote the subject matter concerning the individual positions and momenta of every particle in a particular system (the microstate of the system) at some time t. Let M be the subject matter concerning the temperature, pressure, and other related macro-properties of that system (the macrostate of the system) at t. Then $M \preceq N \preceq H_t$.

It is worth noting that the relation $\preceq$ is not a total order; a subject matter may neither include, nor be included in, another.

Example 6 (Roulette wheel)**.** Consider a roulette wheel in a casino. One subject matter M may specify the overall macro-configuration of the wheel and basic facts about how the croupier has previously spun it. Another subject matter N may specify, in addition, the exact incline of the table on which the wheel sits. A third subject matter O may specify, instead of the incline, the angle of the ball's descent. Here $M \preceq N$ and $M \preceq O$ but $N \not\preceq O$ and $O \not\preceq N$. (This example is loosely based on Bass (1985), who describes how a group of physics graduate students in the late 1970s and early 1980s used computer technology to gain better probability estimates of the outcome of roulette games.)

Some more terminology will be useful in what follows. Fix a subject matter M. Then a *complete M-specification* is a cell of M: it completely specifies how things are M-wise. A *partial M-specification* is any disjunction of complete M-specifications. A proposition[4] $A \subseteq \mathcal{W}$ is *entirely about M* if it is either a complete M-specification or a partial M-specification. (For instance, the proposition that this coin lands heads at t is entirely about H_t.) In what follows, I will use $\mathcal{A}(M)$ to denote the set of propositions entirely about M.

[4]There are many complications and controversies related to this choice to model propositions as sets of possible worlds. For example, what should we say about *de se* claims, like 'It is Monday'? I will set aside most of these complications and controversies here. For example, I will not consider whether de se claims can also have objective chances (cf. Wilhelm (2020)).

1.3 Chance theories

We can consider chance theories based on different subject matters. Roughly speaking, a chance theory describes the chance 'laws' that apply to a world. Take a Newtonian world. Then a chance theory for that world might amount to a series of conditionals of the form: if the initial configuration of the particles is such-and-such, then the chance of this future configuration is such-and-such. Such a chance theory encodes a lot of information, indeed if we think of dynamical laws as (possibly stochastic) maps from possible initial states to possible final states, then there is a precise sense in which such a chance theory specifies the dynamical laws of the world.[5]

I want a broad notion of chance theory that doesn't demand they be based on maximally detailed initial configurations or segments of history, and (relatedly) doesn't demand they take the form of dynamical laws. More generally, then, a chance theory *based on* a factual subject matter M takes a specification of how things are M-wise and, based on that specification, assigns chances to various propositions. In more precise terms, a *chance theory T_M based on a factual subject matter M* associates, with each complete M-specification F, a chance assignment P_{MF} represented by a function which assigns real numbers ranging between 0 and 1 to various propositions.[6]

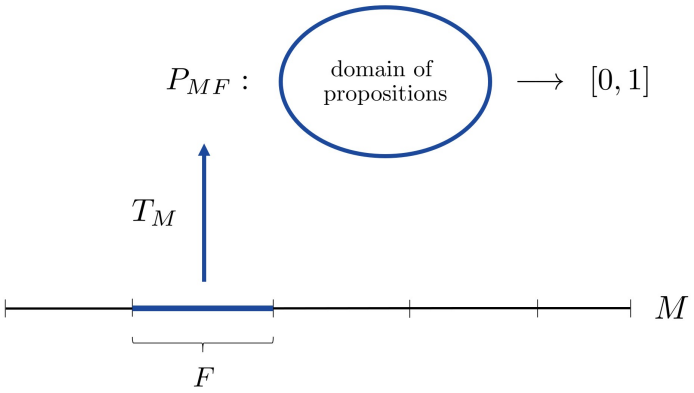


Figure 1.3: A chance theory T_M associates with each complete M-specification F a chance function P_{MF}.

Which propositions? I think the answer varies from case to case. Depending on the subject matter M and the proposition A, the M-chance of A may not be defined at all—for example, biological theories do not assign chances to propositions about quark collisions—or may not be defined in worlds where F is true. (I will return to this issue in detail in a later section, because I think this is a place where my framework offers

[5] In the deterministic Newtonian case, if this is a chance theory for the fundamental dynamical chances, then based on the initial micro-configuration it will assign 0 or 1 to all possible future micro-configurations. We can derive from this a function which takes possible initial micro-configurations and outputs the future micro-configurations, namely those that get assigned 1.

[6] Ultimately we may want to allow that special science probabilities take imprecise values, such that it is indeterminate whether the probability is x or y, but it is determinate that it falls within $[a, b]$. Or, we may want to allow qualitative or comparative probabilities. I leave these possible generalizations as future tasks for the development of the framework.

an improvement over other similar proposals.) So, in what follows I will not make any universal assumptions about the domain of these chance assignments. When I say "assuming $P_{MF}(A)$ is well-defined", I mean assuming the proposition A falls under the domain of P_{MF}.

Two chance theories based on the same factual subject matter M may disagree, namely, they may disagree on the M-chances. However, if two chance theories are based on different subject matters M and N, they may be incapable of direct disagreement. In particular, if M and N do not overlap (if they do not share a cell), then their chance assignments will always be relativized to different specifications; one theory will assign chances based on a specification F of how things are M-wise, and the other theory will assign chances based on a specification G of how things are N-wise. Of course, although these theories cannot be directly incompatible, depending on the relationship between F and G they may still conflict in some other way. I will discuss the constraints that relate chance theories based on different subject matters in a later section.

Here are some examples of chance theories. In what follows, for a subject matter M and world w, I will let $M(w)$ denote the cell of M containing w, in other words, the complete M-specification compatible with w. Given a chance theory T_M based on M, I will then write $P_{M(w)}$ as shorthand for P_{MF} where $F = M(w)$.

Example 7 (Omniscient chance theory)**.** Consider the *omniscient chance theory*, the chance theory T_H based on the entire history of the world that simply specifies the truth value of every proposition about matters of particular of fact. In particular, according to this theory, for every $w \in \mathcal{W}$ and proposition $A \in \mathcal{A}(H)$ about matters of particular fact, the H-chance of A at w is given by

$$P_{H(w)}(A) = \begin{cases} 1 & \text{if } H(w) \text{ entails } A, \text{ and} \\ 0 & \text{if } H(w) \text{ entails not-}A. \end{cases}$$

Note that one of these two conditions always holds, since every $A \in \mathcal{A}(H)$ can be expressed as a union of cells of H, and so either contains or excludes the cell $H(w)$.

Example 8 (Fundamental dynamical chances)**.** For each time t, we can consider a chance theory T_{H_t} based on the initial segment of history up to t. For a given world w, to specify the *fundamental dynamical chances at* w is to specify the function $P_{H_t(w)}$ for every time t. In the literature, this function is often written P_{tw}. When no confusion will arise (that is, when it is clear H_t is the subject matter under consideration), I will also adopt this notation. As an example of fundamental dynamical chance, consider a world w with a fair

'quantum coin' that is flipped at time t_f and lands heads. At w, we have

$$P_{tw}(A) = \begin{cases} 1/2 & \text{for } t \leq t_f \\ 1 & \text{for } t > t_f, \end{cases}$$

where A is the proposition that the coin lands heads at t_f. As I mentioned, these chances are the focus of Lewis (1980). Intuitively, we can think of the omniscient chance theory above as the limit of the fundamental dynamical chance theories, with the time parameter t taken to infinity.[7]

Example 9 (Medical)**.** Let M be a subject matter concerning the incidence of lung cancer, as well as the rates of smoking, in some sample of the population. Consider a complete M-specification F according to which the incidence of lung cancer in the sample is 1% and the incidence of lung cancer among smokers in the sample is 12%. Consider a chance theory T_M based on M which assigns chances of getting lung cancer to individuals in the wider population. For example, if A says that a particular individual will get lung cancer, we may have $P_{MF}(A) = 1\%$. If B says that this individual is a smoker, we may have $P_{MF}(A|B) = 12\%$.[8]

Example 10 (Indeterministic wheel of fortune, continued)**.** In the indeterministic wheel of fortune case, suppose the position of the wheel at the current time t is as shown in Figure 1.4. The next incoming particle is moving quite slowly, such that it will only slightly perturb the wheel, as indicated by the arrows in the figure. In particular, if the particle hits the left side of the wheel, the outcome will be WHITE, and if the particle hits the right side of the wheel, the outcome will be be BLACK. Recall that according to the fundamental dynamical laws, the chance that the particle going left through the magnet is 30%. So letting A denote the proposition that the next spin of the wheel yields WHITE, we have

$$P_{tw}(A) = 30\%,$$

where P_{tw} is shorthand for $P_{H_t(w)}$. That said, while the H_t-chance of A may be 30%, this does not mean there is no sense in which the chance of A is 50%. Indeed, note that the assignment above was predicated

[7] I say fundamental dynamical chance theor*ies*, plural, because strictly speaking for each time t, there is a fundamental dynamical chance theory T_{H_t} based specifically on H_t. Of course, it is a little unnatural to individuate 'theories' so finely in this way. Ultimately this is just a matter of terminology; the slightly odd choice will help us avoid confusion when we discuss non-dynamical chances, where the time parameter no longer plays the same role.

[8] Many authors have pointed out that these chance ascriptions fall prey to the *reference class problem* (Hájek, 2007). For instance, although the overall incidence of lung cancer is 1%, this rate varies from country to country. Suppose the rate is lower where the individual lives. Plausibly, the chance of A should then be lower; but now it seems the chance of A can change depending on how the individual is classified, an apparent paradox. On the current framework, however, there is no paradox. Let N be the subject matter whose cells specify not only the overall incidence of lung cancer among smokers and non-smokers in the sample, but also the incidence as a function of country of residence, as well as the country of residence of each individual in the wider population (note $M \preceq N$). Let G be an N-specification according to which the incidence of lung cancer in Peru is 0.9%. If the individual in A resides in Peru, A may indeed have a lower chance $P_{NG}(A) = 0.9\%$. However, as discussed above, this N-chance assignment is perfectly consistent with the M-chance assignment $P_{MF}(A) = 1\%$. Later I will discuss how these various assignments should guide our beliefs, and posit some 'alignment' constraints on how they relate.

on the specification that the particle was moving at such-and-such speed. If we consider a subject matter $\tilde{H}_t$ that coarse-grains over particle speeds, then assuming $P_{\tilde{H}_t(w)}$ is a relative smooth distribution over those speeds, we will obtain $P_{\tilde{H}_t(w)}(A) = \tilde{P}_{tw}(A) = 50\%$; see Sections 1.6 and 1.7, especially Example 15.



Figure 1.4: The position of the wheel at t.

Example 11 (Roulette wheel, continued)**.** Returning to the roulette wheel example, we can imagine that for the proposition A that the ball lands in pocket 15 (there are 38 pockets total), the true M-chance, N-chance and O-chance theories are such that at the actual world w:

$$P_{M(w)}(A) = 1/38 \approx 0.026, \qquad P_{N(w)}(A) \approx 0.030, \qquad P_{O(w)}(A) \approx 0.025.$$

(For instance, imagine the incline of the table favors a group of pockets including 15, but the angle of the ball's descent, taken without this information about the incline, slightly disfavors that zone of the wheel.)

Let M be a factual subject matter. What is the true M-chance theory? This question picks out another subject matter, the subject of M-chance theories, which I will denote $\mathbf{T}_M$. This subject matter can be represented by a partition of $\mathcal{W}$ where two worlds belong to the same cell just in case the same M-chance theory is true of them.[9] By slight abuse of notation, I use the same symbol T_M to denote the proposition that T_M is the true M-chance theory. (As I mentioned at the beginning of this section, we can think of T_M, generalizing Lewis (1980), as a conjunction of specification-to-chance conditionals of the form ⟨if things M-wise are F, then the M-chances of such-and-such propositions are thus-and-so⟩.[10]) Using this notation, then, T_M is an element of the partition $\mathbf{T}_M$.

Arguably, for at least some subject matters M, it is metaphysically possible that there are simply no M-chances, or interesting true M-chance theories. We can accommodate this possibility in the current framework as follows. Recall we left open that the domain of chance assignments may vary. So for a subject

[9]To stay as neutral as possible on different views of the metaphysics of chance, I stay neutral on whether $\mathbf{T}_M$ is a factual subject mater. For example, consider $\mathbf{T}_{H_t}$ for the current time t. Some theorists will want to say that a certain T_{H_t} is the true fundamental dynamical chance theory in virtue of some law, power, or propensity, which cannot be reduced to an arrangement of matters of particular fact. Others—most notably, Humeans about chance—will disagree. While I am staying neutral on which view is correct, the issue is not altogether independent of how I ultimately formulate the principal principle in my framework; see the discussion of the "big bad bug" in Section 1.4.

[10]These should be thought of as some kind of counterfactual conditionals, not material conditionals (Lewis, 1980, p. 275).

matter M, we can consider a theory T_M such that all of its chance assignments P_{MF} have the trivial domain $\{\emptyset, \mathcal{W}\}$ (or simply no domain). This element of $\mathbf{T}_M$ represents the possibility that the true M-chance theory is trivial and uninteresting. Those skeptical of the existence of non-fundamental chances may, for many M, be inclined to believe that this trivial M-chance theory is the true one.[11]

Generally speaking, in order to apply an M-chance theory to get specific chance predictions, one needs to know how things are M-wise. Thus, it will be useful to introduce yet another subject matter, namely $M\mathbf{T}_M$, the smallest subject matter that includes both M and $\mathbf{T}_M$. The elements of $M\mathbf{T}_M$ are conjunctions of the form FT_M. For agents looking for M-chance guidance that can be applied to their lives, it is this combined subject matter that is of most interest. Note that this combined subject matter will include all propositions of the form $\langle$the M-chance of A is $x\rangle$. In particular, any such proposition can be expressed as a disjunction of FT_M-pairs, namely those pairs that satisfy $P_{MF}(A) = x$.

1.4 The principal principle

Chance guides belief. Put more precisely in the framework of Bayesian epistemology—which I adopt here—chance constrains rational credence. But how so, exactly?

For the discussion that follows, I will make two substantive assumptions about rational credence:

Probabilism A rational agent's credences can be modeled by a (standard, real-valued) probability function defined over an algebra of propositions.

Conditionalization A rational agent's conditional credences obey the standard Bayesian conditionalization formula $C(A|B) = C(AB)/C(B)$ when $C(B) > 0$. Moreover, any rational agent's credences C can be viewed as the result of taking a rational prior C_0, and conditionalizing it on her total evidence. C_0 is called the agent's *ur-prior*, her initial credence function 'before she has obtained any evidence'. It is also sometimes called her *hypothetical prior* (Meacham, 2016).

These assumptions are crucial premises in many of the results of that follow, but it would take me too far afield to defend them here. For what it's worth, both can be weakened in various ways and (variants of) the important results will still hold. However, since these strong versions will help keep the presentation more simple and streamlined, I will hold on to them as stated.

Indeed a third assumption will also be very helpful for preventing extra complications. I will focus on rational agents whose (initial) epistemic possibilities coincide with the metaphysical possibilities $\mathcal{W}$. So I will take C_0 to be defined over an algebra on $\mathcal{W}$. I will also assume this algebra is rich enough to contain

[11] Some skeptics will also want to go for the stronger claim: that this trivial M-chance theory T_M is the *only metaphysically possible* M-chance theory. If that is right, then $\mathbf{T}_M$ will be ultimately represented by the trivial single-element partition $\{T_M\}$.

all the propositions under discussion, so expressions like $C_0(T_M)$, $C_0(F)$, $C_0(A)$, and so on are always well-defined. Furthermore, I will set aside complications related to infinities and 'probability-zero' events, by always assuming the partitions M and $\mathbf{T}_M$ are finite or countably infinite.[12]

The local principal principle

We are now in a position to state what I dub the *local principal principle*; later I will also introduce an even stronger *global principal principle*. This first principle states:

Local Principal Principle Given only the truth of an M-chance theory T_M and complete M-specification F, an agent ought to conform her credences to the chance assignment P_{MF} that T_M associates with F. In other words, $C_0(A|FT_M) = P_{MF}(A)$ whenever the right-hand side is defined.

To illustrate, imagine an agent whose total evidence consists of the history $H_t(w)$ of her world w up to now. Suppose she learns that the chance theory $T_{H_t} \in \mathbf{T}_{H_t}$ is true. Then the principle demands that she conform her credences to the chances assigned by this theory. For instance, recall the quantum coin, Example 8. The theory assigns a chance of $1/2$ to the proposition A that the coin lands heads. According to this principle, if the agent learns this theory, then she should conform her credence in A to $1/2$.

We can imagine an agent who already knows how coin will land heads, say by way of some oracle or time traveler. Intuitively, in such a case, the agent should not set her credence in A to $1/2$, but rather to a higher number. This does not contradict the principle above, however, because in this case the agent has factual information which goes beyond $H_t(w)$ (which goes beyond a complete specification of how things are M-wise; in this case $M = H_t$ for the present time t). This is why the principle is formulated in terms of the agent's ur-prior C_0, so that we can take into account any evidence the agent brings to the table that 'goes beyond' what the M-chances can offer, and is thus *inadmissible* (Lewis, 1980).

Indeed, we can recast the principle above in a more familiar form, by precisely defining the notion of admissibility. Say that a proposition E is *strongly admissible with respect to* M if it is entirely about the subject matter $M\mathbf{T}_M$, in other words if it can be expressed as a disjunction of FT_M-pairs: $E \in \mathcal{A}(M\mathbf{T}_M)$. Let the *$M$-chance hypothesis* $\langle P_M(A) = x\rangle$ be the disjunction of all FT_M-pairs such that $P_{MF}(A) = x$. Then assuming probabilism and conditionalization, the principle above is equivalent to:[13]

[12] This last assumption is especially unrealistic given the previous idealization. For what it's worth, the results can be re-derived for infinite partitions using Kolmogorov's theory of *proper, regular conditional distributions* from Chapter 7. However, I want to set aside these technical complications in the current discussion, since they are not the main philosophical focus.

[13] To show the admissibility version implies the original version, fix T_M and F. Suppose $P_{MF}(A)$ is defined and denote its value by x. Note that the conjunction $E = FT_M$ is strongly admissible with respect to M and that $\langle P_M(A) = x\rangle E = E$ since E is one of the disjuncts of $\langle P_M(A) = x\rangle$. Thus $x = C_0(A|\langle P_M(A) = x\rangle E) = C_0(A|E) = C_0(A|FT_M)$ as desired. Conversely, to show the original version implies the admissibility version, fix M and suppose the conjunction $\langle P_M(A) = x\rangle E$ is non-empty where E is strongly admissible with respect to M. Then this conjunction can be expressed as a disjunction $D = \vee_{i\in I} F_i T_M^i$ where under our simplifying assumption we suppose I is countable. Then also applying probabilism and conditionalization we

Local Principal Principle (Admissibility Formulation) $C_0(A|\langle P_M(A) = x\rangle E) = x$ whenever E is strongly admissible with respect to M and the conjunction $\langle P_M(A) = x\rangle E$ is non-empty.

In other words, as long as the agent's total evidence E is strongly admissible with respect to M, she should conform her credences to known M-chances: $C(A|\langle P_M(A) = x\rangle) = C_0(A|\langle P_M(A) = x\rangle E) = x$. The idea for how to translate between these two formulations is due to Meacham (2009).

The principle has several notable consequences. One is that even if the agent is uncertain what the M-chances are, her credences (as long as her total evidence is strongly admissible with respect to M) should still equal her *expectation* of the M-chances. Let's state this result more precisely. First, we consider a simple case where the agent has ruled out the possibility that the M-chance of A is undefined: so in particular, according to all the FT_M-pairs compatible with E, $T_{MF}(A)$ is defined. Then conjunctions of the form $\langle P_M(A) = x\rangle E$ completely partition E, and so:

$$\begin{aligned} C(A) = C_0(A|E) &= \sum_x C_0(A|\langle P_M(A) = x\rangle E) \cdot C_0(\langle P_M(A) = x\rangle E|E) \\ &= \sum_x x \cdot C(\langle P_M(A) = x\rangle), \end{aligned}$$

where the first equality follows from conditionalization, the second from probabilism and conditionalization, and the third from the local principal principle and conditionalization. Note that the last term is just the agent's subjective expectation of the M-chance of A. To take into account cases where E is compatible with the M-chance of A being undefined, we need to generalize the result. Let D be the disjunction of FT_M-pairs such $P_{MF}(A)$ is defined (such that A belongs to the domain of P_{MF}). Then we have

$$\begin{aligned} C(A|D) = C_0(A|DE) &= \sum_x C_0(A|\langle P_M(A) = x\rangle E) \cdot C_0(\langle P_M(A) = x\rangle E|DE) \\ &= \sum_x x \cdot C(\langle P_M(A) = x\rangle|D). \end{aligned}$$

In words, this says that the agent's credence in A, conditional on the M-chance of A being well-defined, should equal her expectation of the M-chance of A, conditional on the M-chance of A being well-defined.

When combined with probabilism and conditionalization, the local principal principle also has interesting consequences for chance itself. One is that chance assignments $P_{MF}(\cdot)$ must be probability functions, or at least, must be extendable to probability functions. This is easier to see with the first formulation of the principle. By probabilism and conditionalization, $C_0(\cdot|FT_M)$ must be a probability function. And the principle requires $P_{MF}(\cdot) = C_0(\cdot|FT_M)$ where their domains overlap.

have $C_0(A|\langle P_M(A) = x\rangle E) = C_0(A|D) = \sum_{i\in I} C_0(A|F_iT_M^i) \cdot C_0(F_iT_M^i|D) = \sum_{i\in I} x \cdot C_0(F_iT_M^i|D) = x \cdot \sum_{i\in I} C_0(F_iT_M^i|D) = x \cdot C_0(D|D) = x$ as desired.

Another consequence of the local principal principle, taken together with probabilism and conditionalization, is that if $P_{MF}(F)$ is defined, it must equal 1. That is, a chance theory must assign a chance of 1 to the matters it is based on (if it assigns them a chance at all). To see this, note that $C_0(F|FT_M) = 1$, since FT_M entails F, but also by the principal principle, $C_0(F|FT_M) = P_{MF}(F)$ when the latter is defined.

A more subtle consequence of the principle is this: at a world w, and for a given subject matter M, if the M-chance of A exists at all, then it must equal a single value. In a sense I have been presuming this already, by associating chance theories with collections of functions with single numerical outputs. But we can see that it follows from the local principal principle anyway, since if $\langle P_M(A) = x\rangle$ and $\langle P_M(A) = y\rangle$ were both true at w for $x \neq y$, then the conjunction $B = \langle P_M(A) = x\rangle\langle P_M(A) = y\rangle$ would be non-empty. Since both conjuncts are strongly admissible with respect to M, applying the principle to each conjunct, we'd obtain $C_0(A|B) = x \neq y = C_0(A|B)$, which contradicts either probabilism or conditionalization.

The above consequence may sound trivial, but it is worth flagging for the following reason. Some authors, such as Callender and Cohen (2009, 2010), have proposed relativizing chances, not to subject matters, but to different languages or vocabularies. Relative to one vocabulary, the chance of A may be x, whereas relative to another it may be y. Some of their examples, such as that of 'mass' and 'schmass', suggest that even in cases where the vocabularies are interdefinable, the chances may differ in this way. However, it seems possible that in such cases, the subject matter concerning the 'masses' and the subject matter concerning the 'schmasses' will end up carving out the same partition. And so positing differing chances in such cases will end up contradicting the local principal principle (assuming probabilism and conditionalization).[14]

I am inclined to judge this relativization of chances to specific languages or vocabularies a step too far. The main reason is that I think physical probabilities depend on the way the world is (or at least: the way it is in such-and-such respects), not on the language we use to describe it. The other reason is that I think we can accomplish a lot of the goal of recovering special science probabilities without relativizing them further, so I don't see this extra step as particularly well-motivated. Those who do want to allow for distinct chances based on the same subject matter will either have to reject my framework as currently formulated, or else one of probabilism, conditionalization, or the local principal principle. As indicated, I endorse these claims, including the local principal principle. In fact, I believe an even stronger principle holds.

The global principal principle

To state the global principal principle, I will need to quantify over factual subject matters. Thus it will be useful to fix in the background a domain of all factual subject matters $M, N, O, ...$, or at least all those which

[14]Thanks to Chris Meacham for helping me see this point. An additional subtlety here is that these theorists, who adopt a kind of Humeanism, may reject the principal principle anyway, for reasons related to the "big bad bug" (see the end of this section). But the conflict can also be derived using the Humean-adjusted principle; see Meacham (2014) for related discussion.

are relevant to the agent. To avoid technical issues, I will again assume that this domain is countable.

Take a conjunction $T = T_M T_N T_O ...$ of chance theories for every factual subject matter. I will call this conjunction T a *global chance theory*. We can think of T as an element of $\mathbf{T}$, the smallest subject matter that includes $\mathbf{T}_M$, $\mathbf{T}_N$, $\mathbf{T}_O$, and so on.

Global Principal Principle Given only the truth of a global chance theory T and complete M-specification F, an agent ought to conform her credences to the chance assignment P_{MF} that the T_M part of T associates with F. That is, $C_0(A|FT) = P_{MF}(A)$ whenever the right-hand side is defined.

It is straightforward to show that the global principle implies the local principle.[15] However, the converse does not hold. While the local principle tells an agent how to set her credences after conditionalizing her prior only on a chance theory based on a certain factual subject matter M, it does not tell her how to set her credences given subsequent conditionalization on chance theories based on other subject matters. As we will see, the global principle (together with probabilism and conditionalization) implies various 'alignment' relationships between chance theories based on different subject matters. The local principle, on the other hand, places very few constraints on the relationship between such theories.

As far as I know, this distinction between the local and global principal principle has not been made in the literature. Since it represents a potentially significant choice point, it is important to flag.

As foreshadowed, I endorse the global principal principle. If an agent has come to know a chance theory based on M, and her only other information is how things are M-wise, then she should conform her credences to the chances assigned by that theory, regardless of whether she happens to know other chance theories based on other subject matters. To get a sense for why, recall that chance theories only tell an agent what the chances would be, given such-and-such arrangements of facts. Intuitively, this hypothetical chance information should not lead one's credence to deviate from one's actual best chance estimate. For example, say that I know the exact history of the world $H_t(w)$ up to now, and I know the true fundamental dynamical chance theory T_{H_t} assigns 1/2, given this history, to a certain coin landing heads. Now, I also happen to know that the true fundamental dynamical chance theory $T_{H_{t'}}$ for some time t' closer to, or even after, the flip. Should this further knowledge affect my 1/2 credence? No, because the theory alone only tells me that *if* things between now and the flip were to develop thus-and-so, the chance of heads would be such-and-such. Intuitively, without knowing any of this further factual information, there is no reason for my credence to shift (see also Lewis (1980, p. 274)).[16]

We can formulate this idea in terms of the notion of admissibility. Fixing a particular M, consider the combined subject matter $M\mathbf{T}$. Then we say that a proposition E is *weakly admissible with respect to* M if

[15] This result is easier to see by comparing the admissibility formulations; see the discussion below.

[16] There are complications here for chances that are given a Humean analysis; see the discussion of the "big bad bug" below.

it is entirely about $M\mathbf{T}$, in other words if it can be expressed as a disjunction of FT-pairs: $E \in \mathcal{A}(M\mathbf{T})$. Then assuming probabilism and conditionalization, the global principal principle above is equivalent to:[17]

Global Principal Principle (Admissibility Formulation) $C_0(A|\langle P_M(A) = x\rangle E) = x$ whenever E is weakly admissible with respect to M and the conjunction $\langle P_M(A) = x\rangle E$ is non-empty.

In other words, as long as the agent's total evidence E is weakly admissible with respect to M, she should conform her credences to known M-chances: $C(A|\langle P_M(A) = x\rangle) = C_0(A|\langle P_M(A) = x\rangle E) = x$. Note that since the subject matter $M\mathbf{T}$ includes $M\mathbf{T}_M$, if E is strongly admissible with respect to M then it is weakly admissible with respect to M. So the global principle implies the local principle, as claimed. I will explore more consequences of the global principle in later sections.

At this point, it is natural to worry whether these principal principles are too artificial to be useful. Is an agent's total evidence ever going to be even weakly admissible with respect to the subject matter M of interest? For example, when considering the subject matter M of the genetics of an organism, we also have evidence about what we ate for lunch. But information about what we have for lunch is neither strongly nor weakly admissible with respect to M. This threatens to make the principal principle otiose as currently formulated. Fortunately, as Meacham (2009, footnote 30) points out, often the principle can still apply in these cases. Suppose an agent's total evidence is the conjunction $E = BI$ where B is weakly admissible with respect to M (about genetics) and I is inadmissible (about what she ate for lunch). She is considering some proposition A about the organism. Suppose, as seems plausible, that A is approximately independent of I, even conditional on B, in the sense $C_0(AI|B) \approx C_0(A|B) \cdot C_0(I|B)$. Then assuming probabilism and conditionalization, $C_0(A|E)$ will be approximately equal to $C_0(A|B)$, which will in turn be constrained directly by the principal principle.

I have been focusing on one type of chance-credence principle, but there are others. In addition to (i) on the following list, we may have (Builes and Spencer, ms):

(i) $C_0(A|\langle P_M(A) = x\rangle E) = x$ if E is admissible with respect to M,

(ii) $C_0(A|\langle P_M(A|E) = x\rangle E) = x$ if defined,

(iii) $C_0(A|\langle P_M(A) \approx x\rangle E) \approx x$ if E is admissible with respect to M,

(iv) $C_0(A|\langle P_M(A|E) \approx x\rangle E) \approx x$ if defined,

(v) ...

For my purposes, however, it will suffice to consider the principle that is principal among them, namely (i).

[17] The proof proceeds as in footnote 13 but with strong admissibility replaced by weak admissibility and T_M replaced by T.

The big bad bug

That said, there is a popular amendment to (i) which is worth mentioning. Many Humeans about chance favor rejecting (i) as formulated, in favor of an alternative called the *new principle*. This alternative is designed to rescue Humeans from the "big bad bug", also known as the problem of undermining futures.

Consider the phenomenon of tritium decay (Lewis, 1994). In the actual world w, tritium has a half-life of about 12.26 years. Consider the proposition A that all future tritium decays (starting from the current time t) occur extremely quickly, say within 3 minutes. Assuming such decays are fundamentally chancy in the way described by some of our current candidate fundamental theories, according to the fundamental dynamical theory T_{H_t}, the chance of A at w is very small but still non-zero: $P_{tw}(A) > 0$. On the other hand, on many Humean views, if A is indeed true at w, then the decay chances are not given by the 12.26 year half-life but are rather some lower number. And so $C_0(T_{H_t}|H_t(w)A) = 0$. But this entails, assuming the principal principle (either version), conditionalization, probabilism, and $C_0(T_{H_t}|H_t(w)) > 0$, that

$$P_{tw}(A) = C_0(A|H_t(w)T_{H_t}) = \frac{C_0(T_{H_t}|H_t(w)A) \cdot C_0(A|H_t(w))}{C_0(T_{H_t}|H_t(w))} = 0.$$

This is a contradiction, so one of the assumptions must be rejected.

A first thought, suggested by my framework, is to reject the premise that H_t is the correct choice of subject matter for Humean views. A more natural alternative, it seems, is that the chance is based on a specification $M(w)$ of the average rate of tritium decay across all of history at w. However, the problem with this proposal is that if the outcome described by A is sufficiently extreme, then it will be logically incompatible with $M(w)$, and by the principal principle we will again get the verdict $P_{M(w)}(A) = C_0(A|M(w)T_M) = 0$.

A follow-up proposal is to instead reject the local and global principal principle as currently formulated, and to amend the requirement $C_0(A|FT_M) = T_{MF}(A)$ to something like the following:

$$C_0(A|FT_M) = P_{MF}(A|T_M).$$

This *new principle* (Hall, 1994; Thau, 1994; Lewis, 1994) has the chance assignment associated with T_M conditionalize on T_M being the correct chance theory. The adjustment allows $P_{MF}(A) > 0$, even though $C_0(A|FT_M) = 0$. So in particular, in the example above, there is no contradiction with having both $P_{tw}(A) > 0$ and $C_0(T_{H_t}|H_t(w)A) = 0$.

According to the new principle, the unconditional chances $P_{MF}(A)$ do not constrain credence; rather, conditional chances of the form $P_{MF}(A|T_M)$ constrain credence. Which is correct, the new principle or the old? In what follows, I will focus on the old principle. One reason is that I think it's usually the chances

$P_{MF}(A)$, not $P_{MF}(A|T_M)$, that are the object of scientific investigation, and my working assumption is that it's what the scientists are investigating that is epistemically central (for similar remarks, see also Builes and Spencer (ms)). But mainly, I focus on the old principle because it is simpler and provides an easier demonstration of the core elements of my framework. A future project is to see to what extent a parallel treatment can be carried out with the new principle. In fact, ultimately I want my framework to accommodate both the old and new principles *simultaneously*, in the following sense. Say facts about the M-chances supervene on the Humean mosaic, but facts about the N-chances supervene on propensities. Then it may be that the old principle is true of the M-chances, while the new principle is true of the N-chances. Ultimately I would like to accommodate this possibility. Indeed I think I can, but some care will be needed when formulating the global versions of the principles, in particular when it comes to distinguishing global chance theories $T = T_M T_N T_O ...$ whose conjuncts include Humean chance theories (for which the new principle would apply, and which should not be counted as admissible) from those that do not.

1.5 Comparison to similar frameworks

The framework presented above shares many similarities with that of Handfield and Wilson (2014), who in turn draw on elements of Meacham (2009) and Lewis (1987).[18] Handfield and Wilson also relativize chances to subject matters—what I have been calling a specification $F \in M$, they call an *evidence base*—and implicitly endorse the global version of the principal principle. The main difference is that while Handfield and Wilson assume that all chance functions P_{MF} assign chances to every proposition (to all $A \in \mathcal{A}(\mathcal{W})$), I allow that chance functions, especially special science chance functions, may have a limited domain. As I mentioned earlier, it seems problematic to assume, for instance, that biological chance theories assign probabilities to quark collisions. Handfield and Wilson acknowledge this worry, writing (2014, 36, notation adapted):

> One further, terminological, issue needs to be mentioned. As we have formulated the multi-grained chance picture, a coarse-grained chance function is complete, in that it ascribes a chance to every contingent proposition. [...] This might seem worrying; shouldn't coarse-grained chance have a limited coverage, so that a chance function ranges only over propositions wholly about its 'target' subject-matter? Well, we can talk this way if we like: then 'coarse-grained chance functions' are individuated by pairs of subject-matters: the target subject-matter M' and the evidence base F. Such restricted-domain chance functions certainly exist. However, they will not

[18] I will discuss the relation to Lewis's (1987) notion of *counterfeit chance* in more detail when I respond to some objections at the end of the chapter.

in general stand in straightforward alignment-style relations to one another, which limits their theoretical usefulness. But ultimately, 'coarse-grained chance function' is a technical term; what matters is how we account for chance-attributions in ordinary language.

The last remark makes salient an underlying difference between our projects, which is that while Handfield and Wilson's primary aim is to account for chance talk in ordinary language, my primary aim is to account for chance theorizing in the sciences.[19] Of course, these aims are not themselves in conflict. Indeed, I find Handfield and Wilson's account of how conversational context can pick out a particular M an attractive one (see Section 1.9 for discussion). But since I am ultimately interested in applying the framework to concrete examples from science, the question of domain completeness will not be a merely terminological issue for me, and so I do not set it aside.

Another advantage of relaxing the domain completeness assumption is that it allows us to capture another recent proposal for how to model (certain kinds of) non-fundamental chances as a special case of the framework. I turn to this proposal, due to List and Pivato (2015), next.

1.6 Emergent dynamical chances

I have discussed in detail how to model fundamental dynamical chances, i.e. Lewisian chances, within my framework. I call them *dynamical* because they are based on an initial segment of history, and change with the time t that marks the end of this initial segment. I call them *fundamental* because this history is specified in maximal microscopic detail. (Since 'fundamental' here is not fundamental in the metaphysics sense, a better choice of terminology may be *maximal* or *physically fundamental*.)

We may also wish to consider dynamical chances which are non-fundamental. In particular, we may wish to consider dynamical chances which are based on an initial segment of history that is coarse-grained in some way. For instance, consider the models of atmospheric dynamics used in meteorology. The chance assignments in these models are based, not on a complete micro-specification $H_t(w)$ of the system up to t (such as the exact position and momenta of every particle) but on a macro-specification $\tilde{H}_t(w)$ of the history (or at least, partial history) up to t, stated in terms of coarse grained variables like temperature, pressure, and humidity. Clearly, in our model of these chances, we would expect the inclusion relation $\tilde{H}_t \preceq H_t$ to obtain. But it would be nice to say something more precise than this.

Fortunately, List and Pivato (2015) provide an elegant framework for modeling these chances. Their first step is to suppose that each complete history $h \in H$ specifies a map from a set $\mathbb{T}$ of possible times into a set S of possible (maximally detailed) states of the world or of some system. Their next step is to consider

[19] And to remain as faithful to the science as possible—at least, as much as possible given the overall Lewisian starting point.

a *coarse graining* of the state space S which involves associating each lower level state $s \in S$ with a higher level state $\tilde{s} \in \tilde{S}$, such that $\tilde{S}$ is a partition of S into equivalence classes. The final step is to consider *coarse-grained histories* $\tilde{h} \in \tilde{H}$ which are functions from $\mathbb{T}$ into $\tilde{S}$, and emergent dynamical chances which are based on initial segments of these histories. For List and Pivato, such chance assignments always have domain $\mathcal{A}(\tilde{H})$, which is smaller than $\mathcal{A}(H)$. List and Pivato represent these chance assignments by functions $\tilde{P}_{\tilde{h},t}$ as opposed to $P_{h,t}$ (in my notation these will be $\tilde{P}_{tw}$ and P_{tw} respectively) and dub them the *emergent chances*; I will call them *emergent dynamical chances*.

List and Pivato's framework can be subsumed as a special case of my own, in the following way. Note that if each complete history $h \in H$ specifies a map from a set $\mathbb{T}$ of times into some set S of possible states, then we can define an extraction function f which takes a time $t \in \mathbb{T}$ and complete history $h \in H$ and yields the state at that time in history: $f_t(h) \in S$. Now for each time t, consider an equivalence relation on pairs of worlds defined by $w_1 \Leftrightarrow w_2$ iff $f_{t'}(H(w_1)) = f_{t'}(H(w_2))$ for all $t' \leq t$, in other words w_1 and w_2 agree on the maximally detailed state of the world for all times up to t. This equivalence relation picks out the subject matter H_t. As discussed, fundamental dynamical chances are defined with respect to this subject matter: $P_{H_t(w)}$, or P_{tw} for short.

Next, consider a *coarse-graining function* which identifies certain lower-level states:

$$\Phi : S \to \tilde{S}.$$

We can compose Φ with f to obtain a new extraction function $\tilde{f} = \Phi \circ f$ which takes a time and history and yields the coarse-grained state. We can then define the subject matter $\tilde{H}_t$ in an analogous fashion, using $\tilde{f}$ instead of f. We can also define $\tilde{H}$ as the subject matter picked out by $w_1 \Leftrightarrow w_2$ iff $\tilde{f}_t(H(w_1)) = \tilde{f}_t(H(w_2))$ for all $t \in \mathbb{T}$. These coarse-grained subject matters obey similar inclusion relations. For instance, one can verify that for all $t' \leq t$, $\tilde{H}_{t'} \preceq \tilde{H}_t$, and for all t, $\tilde{H}_t \preceq \tilde{H}$. We also get the inclusion relations across levels of grain that we predicted. For example, one can show that $\tilde{H}_t \preceq H_t$ for fixed t.

Finally, consider chance theories $T_{\tilde{H}_t}$ based on these coarse-grained histories. They associate coarse-grained histories $\tilde{H}_t(w)$ with chance assignments $P_{\tilde{H}_t(w)}$, which I will denote $\tilde{P}_{tw}$ for short. What is the domain of these chance assignments? According to List and Pivato, they always have domain $\mathcal{A}(\tilde{H})$:

> [T]he key insight is this: when evaluating chance and (in)determinism at a higher level of description, *only higher-level language is available*. The relevant family of history-and-time indexed probability functions now consists of functions defined on the algebra $\mathcal{A}(\tilde{H})$ of higher-level events rather than the lower-level algebra $\mathcal{A}(H)$, and the index $\tilde{h}$ now ranges over $\tilde{H}$ rather than H (while t continues to range over $\mathbb{T}$). (List and Pivato, 2015, p. 134, notation adapted)

As foreshadowed, we need domain flexibility to faithfully incorporate List and Pivato's account. Incidentally, however, while I think List and Pivato are correct that in many cases the emergent dynamical chance functions have a restricted domain $\mathcal{A}(\tilde{H})$, it is not clear this is always true. One can imagine that in some cases, $\tilde{P}_{tw}$ takes the form of a statistical mechanical-style measure, and so assigns well-defined chances to microstates as well as macrostates. So I will not in general demand that all functions of the form $\tilde{P}_{tw}$ have domain $\mathcal{A}(\tilde{H})$.

We can apply the above treatment to a concrete toy example. I will return to treat this example in much more detail in Chapter 2. It is based on the *baker's map*, a well-known dynamical map from ergodic theory. (Cf. also Werndl (2009, 2012, 2013) and List and Pivato (2015, p. 136).)

Example 12 (Baker's map)**.** Suppose time evolves in discrete steps $\mathbb{T} = \{0, 1, 2, ...\}$ and the possible states of the world, at a given time, can be specified by a point on the unit square, so $S = [0, 1] \times [0, 1]$. In this case we can think of each possible history $h \in H$ as a sequence $(s_0, s_1, ...)$ of such points. Suppose the fundamental dynamical laws at this world w are deterministic, in particular they are given by a map $U : S \to S$ with

$$U(x, y) = \begin{cases} (2x, \frac{1}{2}y) & \text{if } 0 \leq x < \frac{1}{2}, \\ (2x - 1, \frac{1}{2}(y + 1)) & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

Then, letting s_0 denote the initial state, the history $H(w)$ of this world is the sequence $(U_0(s_0), U_1(s_0), U_2(s_0), ...)$ where U_t means that U is applied t times. The fundamental dynamical chances at this world are trivial; they assign 0 or 1 to every proposition depending on whether it is consistent with this special sequence.



Figure 1.5: Reproduced with permission from (Werndl, 2009, Figure 1). Baker's map: the square is compressed, cut in half, and then stacked on itself. This continues until the white and grey strips are so thin that it is impossible to discern (visually speaking, looking at the diagram) black versus white points.

However, consider the coarse graining $\Phi : S \to \{L, R\}$ depicted in Figure 1.6 that only specifies whether we are in the left or right side of the square:

$$\Phi(x, y) = \begin{cases} L & \text{if } 0 \leq x < \frac{1}{2}, \\ R & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

The coarse-grained history $\tilde{H}(w)$ will then be a sequence of Ls and Rs $(\Phi(U_0(s_0)), \Phi(U_1(s_0)), \Phi(U_2(s_0)), ...)$.

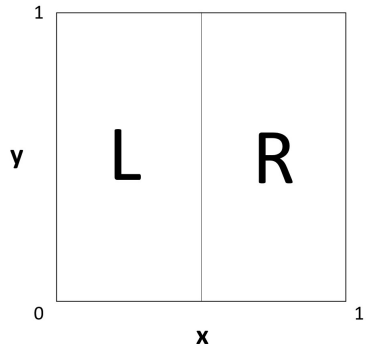


Figure 1.6: A coarse graining of state space.

We can consider initial segments of this coarse grained history $\tilde{H}_t(w)$ up to some time t. While the chances based on a fine-grained history $H_t(w)$ are trivial, the chances based on this more coarse-grained history may be non-trivial. For example, if R_{t+1} is the proposition that the state is R at time $t+1$, we may have:

$$\tilde{P}_{tw}(R_{t+1}) = 1/2.$$

In fact, in Chapter 2 I will discuss a precise sense in which the underlying dynamics give rise to a fair coin distribution over L and R. As I will also discuss there, other kinds of coarse-grainings Φ' will give rise to other kinds of stochastic processes. So there may be multiple distinct emergent dynamical chances, an $\tilde{H}_t$-chance, an $\tilde{\tilde{H}}_t$-chance, and so on, for the same proposition. That said, depending on the domains of these chance assignments, this does not mean there are no constraints on how they relate, as I discuss next.

1.7 Alignment between chances of different kinds

The global principal induces constraints on the relationship between chance theories based on different subject matters.

Fix a global chance theory T. An important and immediate consequence of the global principal principle (taken together with probabilism and conditionalization) is that if two subject matters M and N share a cell F, then, when it comes to any proposition A in their shared domain, the conjuncts T_M and T_N of T cannot disagree on the chance of A given F. That is to say, if $P_{MF}(A)$ and $P_{NF}(A)$ are both defined, then they must be equal. To see this, note that applying the global principal principle twice, we have $C_0(A|FT) = P_{MF}(A)$ when the latter is defined, and $C_0(A|FT) = P_{NF}(A)$ when the latter is defined. So when both chances are defined, $P_{MF}(A) = P_{NF}(A)$.

One upshot of this result is that it allows us to simplify our notation. Given a global chance theory T, we can simply write $P_F(A)$ for the chance (if any) of A that T assigns given the complete specification F. It does not matter if $P_F(A)$ is the chance assigned by T_M or by some other conjunct of T, since assuming the global principal principle, to the extent that these theories speak on the chance of A when F is the complete specification, they cannot disagree. In what follows, I will adopt this simplified notation.

The global principal principle also implies a further substantive constraint on the relationship between chances of different kinds. The following is a slight variant of a result due to Handfield and Wilson (2014, p. 33), which they also dub alignment:

Alignment Fix a global chance theory T and let F and G be complete specifications of (possibly distinct) subject matters theorized by T. If G entails F, then

$$P_F(A|G) = P_G(A)$$

whenever both $P_F(A|G)$ and $P_G(A)$ are defined.

For proof that this indeed follows from the global principal principle, see the footnote.[20] In words, alignment says the following. Suppose a coarse-grained chance function P_F and a more fine-grained chance function P_G assign chance to a proposition A. Then the coarse-grained chance function, if conditioned on the basis G of the more fine-grained chance function, must—assuming that basis G belongs to its domain—assign the same chance to A as the fine-grained chance function.

Framed in terms of subject matters M and N, alignment implies that if $M \preceq N$, then for all worlds w, $P_{M(w)}(A|N(w)) = P_{N(w)}(A)$ whenever both are defined. In words, the N-chance function can be obtained by taking the M-chance function and conditionalizing it on how things are N-wise (assuming the M-chance function initially assigns chances to how things are N-wise).

Example 13 (Fundamental dynamical chances, continued)**.** Fix a world w and $t > 0$. Note that $H_t(w)$ entails $H_0(w)$, the proposition specifying the initial state of the world. So alignment implies, for all $A \in \mathcal{A}(H)$:

$$P_{0w}(A|H_t(w)) = P_{tw}(A).$$

In other words, the fundamental dynamical chances at a time t can be obtained by taking the initial fundamental dynamical chances and conditionalizing them on the history up to t. (In fact, this claim about

[20] Suppose $P_F(A|G)$ and $P_G(A)$ are defined. We have $P_F(A|G) = C_0(A|GFT) = C_0(A|GT) = P_G(A)$ where the first equality follows from the fact that, assuming the global principal principle, $P_F(AG) = C_0(AG|FT)$ and $P_F(G) = C_0(G|FT)$, the second equality follows from the assumption that $G \subseteq F$, and the final equality follows from the global principal principle.

Lewisian chance has been derived in several places in the literature. Here I point out that it is a consequence of a more general alignment principle, and is not special to Lewisian chance.)

Example 14 (Roulette wheel, continued)**.** Recall that for A the proposition that the ball lands in a particular pocket, we had $P_{M(w)}(A) = 1/38 \approx 0.026$. Suppose that $P_{M(w)}$ assigns chances to different possible inclines of the table. Then we must have $P_{M(w)}(A|N(w)) = P_{N(w)}(A) \approx 0.030$. Similarly, suppose $P_{M(w)}$ assigns chances to the different possible angles of the ball's descent. Then we must have $P_{M(w)}(A|O(w)) = P_{O(w)}(A) \approx 0.025$. On the other hand, since $N(w)$ and $O(w)$ crosscut each other, alignment does not apply directly to $P_{N(w)}$ and $P_{O(w)}$. However, this does not mean alignment does not apply indirectly (Handfield and Wilson, 2014, pp. 34–35). For example, suppose there is a more fine-grained subject matter Q, like the combination of M, N, and O, such that $P_{N(w)}$ and $P_{O(w)}$ assign a chance to the basis $Q(w)$. Then it follows $P_{N(w)}(A|Q(w)) = P_{Q(w)}(A) = P_{O(w)}(A|Q(w))$. In other words, if you take the N-chances and condition them on the angle of the ball, and if you take the O-chances and condition them on the incline of the table, then (assuming they are not silent) these two types of chances must agree.

Another way of stating the upshot of alignment is that if a coarse-grained chance function has opinions about the bases of a more fine-grained chance theory, then its chance estimates must equal its expectation of the more fine-grained theory's estimates. In particular, suppose $G_1, G_2, ...$ partition F. Then:

$$P_F(A) = P_F(A|F) = \sum_i P_F(A|G_i) \cdot P_F(G_i|F) = \sum_i P_{G_i}(A) \cdot P_F(G_i)$$

assuming all terms are well-defined, where the first equality follows from the fact $P_F(F) = 1$ (which follows from the principal principle as discussed earlier), the second from the definition of conditional probability and the fact $P_F(\cdot)$ is a probability function (which follows from the principal principle as discussed earlier), and the third from alignment.

We can apply this fact to the wheel of fortune example:

Example 15 (Wheel of fortune, continued)**.** A particle is approaching the wheel of fortune at time t. Suppose the coarse-grained chance function $\tilde{P}_{tw}$ includes in its domain propositions about the specific configuration of the particles and the wheel up to t. Then letting A denote the proposition that the next spin of the wheel yields WHITE, and $X_A : \mathcal{W} \to \mathbb{R}$ a random variable defined by $X_A(w) = P_{tw}(A)$, the alignment result above (generalized to allow for infinite spaces, so the sum becomes an integral) implies:

$$\tilde{P}_{tw}(A) = \int_{w' \in \tilde{H}_t(w)} X_A(w')\, d\tilde{P}_{tw}(w').$$

In the deterministic case, $X_A(w') = 1_A(w')$ is just the indicator which returns 1 if A is true at w' and 0

if A is false at w'. Assuming that the value of X_A is a function of the current configuration of the wheel and the speed of the particle in question, and assuming it varies sensitively with speed (with the ratio of WHITE to BLACK in small regions of state space about $1:1$), then as long as $\tilde{P}_{tw}$ is distributed smoothly over the possible speeds compatible with $\tilde{H}_t(w)$, method-of-arbitrary-function results (Von Plato, 1983; Strevens, 1998, 2003, 2011, 2013) will imply the right-hand-side of the above equation is approximately 50%.

1.8 The Mentaculus

According to a proposal by Albert (2015) and Loewer (2020), not only are there alignment relations between different special science chances, but there is actually one parent chance function, a kind of statistical mechanical measure 'to rule them all'. This proposal is wide-ranging, seeking to account not only for special science probabilities themselves, but also for various phenomena related to explanation, time's arrow, and causation. Dubbed *the Mentaculus*, it has three main ingredients (Loewer, 2020, p. 5): the fundamental dynamical laws, which are assumed to be deterministic and temporally symmetric, the Past Hypothesis, which is the hypothesis that the universe's macrostate at a time shortly after the big bang was one of very low entropy, and, thirdly, a uniform probability distribution μ_0 over all the micro-histories of the universe that are compatible with the Past Hypothesis and the fundamental dynamical laws.

To get a sense of how the proposal is supposed to work, consider the proposition A that the majority of babies born in London to human parents in the 17th century will be boys. To get the relevant special science probability of A, one conditionalizes μ_0 on, as Albert (2015, p. 32) puts it, "the existence of our galaxy, and of our solar system, and of the earth, and life, and of the human species, and of cities, and of whatever else is implicitly being taken for granted in any scientific discussion of the relative birth rates of boys and girls in London in the [17th century]". Let B denote the proposition Albert has in mind here. According to this proposal, the probability one obtains, $\mu_0(A|B)$, should match the predictions and explanations provided by our special sciences. In this case, there is an accepted explanation which predicts the ratio ought to be about $1.1:1$ in favor of boys. According to this proposal, then, $\mu_0(A|B)$ will be relatively high.

We can subsume this proposal as a special case of our framework. Let w be the actual world. Let us say a chance theory T_M is a *Mentaculus-type theory for* w if it simultaneously satisfies the following pair of desiderata:

1. $M(w)$ is more coarse-grained than the specifications of many or all special science subject matters $N(w), O(w), ...$, and yet

2. the domain of the associated chance assignment $P_{M(w)}$ is extremely wide, for instance, it is $\mathcal{A}(H)$.

If T_M is a Mentaculus-type theory and N is one of the special science subject matters from (1), with T_N a chance theory for N, then alignment implies

$$P_{N(w)}(A) = P_{M(w)}(A|N(w)),$$

for all propositions A in the domain of the special science chance function $P_{N(w)}$. In other words, the special science chance of A can be expressed as a conditional Mentaculus chance. In the case above, $M(w)$ specifies the Past Hypothesis, $N(w)$ is equal to B, and $T_{M(w)}$ is equal to μ_0, which is defined over $\mathcal{A}(H)$.

Using my framework, then, we can factor the Mentaculus proposal into two components. The first component says that there is a true chance theory with features (1) and (2). The second component says that special science chances can be expressed as the conditional chances of such a theory. Perhaps surprisingly, we have found that once the first component is secured, the second component follows immediately from the global principal principle, via alignment. Assuming the global principal principle, two relatively simple—albeit substantial—conditions, (1) and (2), suffice to generate a "probability map of the world".

1.9 Objections

To conclude the chapter, I consider some objections to the framework I have laid out. Many of these also constitute objections to the other frameworks (Meacham, 2009; Handfield and Wilson, 2014; List and Pivato, 2015; Albert, 2015; Loewer, 2020) it subsumes.

Generic chances and event types

The first objection concerns the application of the framework to the special sciences. Many special science chances look like what we might call *generic chances* (Hoefer, 2007, 2019). Consider, for instance, the statement that ⟨the chance that a human develops lung cancer, given that he is male, and lives in Peru, is approximately 1%⟩. Read at face value, the chance of 1% here does not attach to any particular token event, outcome, or proposition, but rather an *event type* or *outcome type*. Similarly for statements like ⟨fair coin tosses have a probability of about 50% of coming up heads⟩. However, the framework offered here, including the alignment results, assumes that the objects of chances are propositions or possible events. This may be unproblematic for chances associated with the fundamental dynamical laws, but it is questionable—the thought goes—for chances in the special sciences.

While I agree that a face value of reading of special science chances has them attaching to event types, I do not think it is ultimately a problem for my framework that it focuses on propositions or token events. In

particular, claims about generic chances can be translated as claims about particularized chances. Following a proposal by Rubenstein (ms), the generic chance claim ⟨the chance that a human develops lung cancer, given that he is male, and lives in Peru, is 1%⟩ can be translated as the claim that for all possible token human individuals o, the chance that individual o develops lung cancer, given that he is male and lives in Peru, is 1%. More generally, if α and β are properties or event types, the generic chance of α given β is the particularized chance that o is α given o is β, where o is any possible instantiator of the properties.

An immediate objection to this proposal is that particularizing to a specific individual will ruin the chance assignment. For example, the chance that Fred develops lung cancer, given that he is male and resides in Peru, is surely higher than the generic chance if (say) Fred smokes. Furthermore, if Fred doesn't live in Peru, it seems that this conditional chance should be undefined, not equal to 1%. However, this objection underestimates the power and flexibility of the framework on offer. If the relevant subject mater M does not specify any further demographic information about Fred (like whether he smokes or resides in Peru), then the M-chance Fred develops lung cancer, given that he is male and resides in Peru, may still be plausibly the generic chance. Indeed, one criterion when charitably associating a subject matter M with a generic chance claim is that the M-chance is approximately uniform over all the tokens of the type being referenced.

How *is* the relevant subject matter M picked out by our ordinary thought and talk, in general? As foreshadowed, I am attracted to Handfield and Wilson's (2014) account for the semantics of chance ascriptions. On their proposed semantics, M will be the most fine-grained subject matter whose true specification could be known in the context, where we restrict attention to practically possible or salient methods. There are several complications to consider here. For example, in a game of poker, the opponent's hand could easily be known by direct inspection, but this method would violate the rules of poker, and so wouldn't count as salient (Handfield and Wilson, 2014, p. 37). I will not delve into these complications or debates about contextualist semantics here, as it is not my main focus. However, this is an important part of any application of the framework to ordinary chance ascriptions, especially generic chance ascriptions like those above. (For more detail on how the framework handles the smoking case in particular, see also Example 9.)

Counterfeit chance and indeterminate subject matters

In response to some criticisms of his subjectivist's guide, Lewis (1987) makes the following remarks:

> Without committing themselves one way or the other on the question of determinism, [Skyrms and Jeffrey] have offered a kind of counterfeit chance to meet the needs of the determinist. It is a relative affair [relative to partitions], and apt to go indeterminate, hence quite unlike genuine chance. But what better could a determinist expect? [...] Counterfeit chances will be relative

> to partitions; and relative, therefore, to standards of feasibility and naturalness; and therefore indeterminate unless the standards are somehow settled, or at least settled well enough that all remaining candidates for the partition will yield the same answers. Counterfeit chances are therefore not the sort of thing we would want to find in our fundamental physical theories, or even in our theories of radioactive decay and the like.

(In Chapter 2, I will discuss the connection of my framework to Skyrms's notion of *resiliency* in detail; for now, I focus on the more general worries Lewis raises in this passage.)

I agree that 'counterfeit chances'—what I would simply want to call chances that are not fundamental dynamical chances—are not the sort of thing we would want or expect to find in our fundamental dynamical laws. However, as I mentioned in my discussion of **Fundamentalism** and **Weak Fundamentalism**, they might still be the sort of thing we will find in our fundamental physical theories, broadly construed. For example, if we think of the Past Hypothesis as a fundamental physical posit (or even itself a Humean-style law, as Loewer (2020) advocates), then the Mentaculus is arguably an illustration of this possibility.

Regarding the relativity to partitions, both Lewisian and non-Lewisian chances are relative to partitions. It is true that the suitable partition M may be harder to identify in the case of the special sciences. But as Lewis alludes to, in many cases of interest we will only need to specify M in rough terms, because small perturbations or shifts to M will not change the chance much. Of course there is a further question of how the subject matter M is even roughly picked out in our ordinary talk, which I briefly addressed above.

A related issue is that most predicates in the special sciences are vague, and so it is rare that a chancy proposition from the special sciences will pick out a particular set of worlds $A \subseteq \mathcal{W}$, or that a special science context will pick out a particular partition M of $\mathcal{W}$. How this vagueness is ultimately incorporated in the model will depend on one's account of vagueness. But again, as long as the chance assignments are stable under small shifts, we can still get a working model by focusing on precisifications as I have done so far.

Gerrymandered subject matters

While some basic questions about determinacy, relativity and vagueness have been addressed, there remains a general worry about what to say about highly gerrymandered subject matters and whether there exist chances based on those subject matters. (See also Handfield and Wilson (2014, Section 9).)

I should emphasize again that my framework is far from committed to the claim that for every subject matter M there is a non-trivial chance theory T_M based on that subject matter. Indeed, as I mentioned in Section 1.3, I leave open that the true theory may be one with only a trivial domain. As I understand it, there being a being a set of true M-specification-to-chance conditionals, such that the M-chances satisfy the

local or global principal principle, is a non-trivial posit.

That said, if some subjects H_t have associated chances, while others M have no associated chances, there is pressure to say what makes some subject matters metaphysically privileged in this way and others not.

There is an answer to this question, but it is not very satisfying. Suppose for instance that there are fundamental quantum propensities of the kind associated with certain quantum theories.[21] Then it will be in the nature of these propensities that facts about their strength screen off, in the way described by the principal principle, all and only the initial segment $H_t(w)$. In contrast, because of the way $M(w)$ is gerrymandered, there may be nothing in the physical world, whether a (possibly emergent) chancy law, power, propensity, or pattern, such that facts about it screen off all and only the information $M(w)$. If so, this will be why there are no M-chances at w. Of course, this answer just passes the buck to more basic questions in the metaphysics of chance, about what kinds of entities or relations are such that facts about them could constrain credence in the way described by the principal principle. My point, though, is that we should not insist from the outset that the only such entities or relations will be ones that screen off $H_t(w)$ as opposed to other kinds of factual information about the world.

Further objections

Maher (2009) advances a number of objections to Lewis's (1987) proposed account of counterfeit chances. Some of these objections apply to my framework as well, and when they don't, it is helpful to see how my framework avoids the problems.

Maher's first objection is that Lewis's account has trouble accommodating typical statements of physical probability, like the following:

> S: The physical probability of heads on a toss of this coin is 1/2.

This statement doesn't attribute chance to a proposition or refer to a time or a possible world, so it doesn't mention any of the things that Lewis says are the arguments of chance or physical probability. We could try to translate it as:

> S': For all times s and t, if s is a token toss of this coin and t is a time just prior to t, then the physical probability of t in the actual world of the proposition that s lands heads is 1/2.

But S' is trivially true if the coin is never tossed, so it can't be equivalent to S. Moreover, the chance of a coin landing heads is different "depending on whether we are talking about tossing by a human, with no

[21] Such as GRW theory. Note that these are not really candidate fundamental theories, so this is a somewhat toy example. Indeed, it's not even clear if on the fundamental theory we can think of H_t as the subject matter picked out, thanks to complications raised by relativity theory (cf. footnote 1).

further specification (in which case S is probably true), or tossing with such and such a force and such and such a position, etc. (in which case S is false), but S' doesn't take account of this" (Maher, 2009, p. 205).

On my framework, the arguments of chance are a subject matter M, a specification $F \in M$, and a proposition A.[22] The statement S is translated as:

> S'': For all possible token tosses o of the coin, the M_o-chance of o yielding heads is $1/2$.

Here the subject matter M_o is picked out by the context. For example, if it is a context where the exact microstate could be determined, then M_o may be H_t for t just before o (in which case S'' is plausibly false). If it is a context where only macrostates could be determined, then M_0 may be $\tilde{H}_t$ for t just before o (in which case S'' is plausibly true). Note that there is no problem with S'' being trivially true if the coin is never tossed, since the quantification is over possible tosses. And the contextual parameter allows us to take into account the issue of fine-grainedness of specification (it may also allow us to restrict attention, in the "for all possible" quantification, to possible tosses o by a human, in cases where this restriction is implicit).

Maher's second objection to Lewis is that there is no good reason to add a possible world as a further argument of chance. He writes (2009, p. 205):

> Of course, the value of a physical probability depends on empirical facts that are different in different possible worlds, but this does not imply that physical probability has a possible world as an argument. The simpler and more natural interpretation is that physical probability is an empirical concept, not a logical one; that is, even when all the arguments of physical probability have been specified, the value is in general a contingent matter.

I agree with Maher on this point, and this is a place where my treatment differs from Lewis's. In my framework, the value of $P_{M(w)}(A)$ is a contingent matter that depends on which chance theory $T_M \in \mathbf{T}_M$ is true. If I were following Lewis (1980), who indexes chance theories T_w to a world, I would let T_M^w denote the M-chance theory that is true at w, and then $P_{M(w)}(A) = C_0(A|M(w)T_M^w)$ would represent the M-chance of A at w. Instead I define $P_{M(w)}(A) = C_0(A|M(w)T_M)$, which is the M-chance of A at w only if T_M is true at w. Whether T_M is true at w is in turn a matter of empirical investigation.

To state Maher's third objection, we need more background on Lewis's proposal for counterfeit chance. Lewis writes (1987, pp. 120–121):

> If a determinist says that a tossed coin is fair, and has an equal chance of falling heads or tails, he does not mean what I mean when he speaks of chance. Then what does he mean? [...] That question has been sufficiently answered in the writings of Richard Jeffrey and Brian

[22] Though given the global principal principle, we think of it as only two arguments, F and A; see Section 1.7.

> Skyrms on objectified and resilient credence. [...] According to my second formulation of the Principal Principle, we have the history-theory partition (for any given time); and the chance distribution (for any given time and world) comes from any reasonable initial credence function by conditionalizing on the true cell of this partition. That is, it is objectified in the sense of Jeffrey. Let us note three things about the history-theory partition.
>
> (1) It seems to be a natural partition, not gerrymandered. It is what we get by dividing possibilities as finely as possible in certain straightforward respects.
>
> (2) It is to some extent feasible to investigate (before the time in question) which cell of this partition is the true cell; but
>
> (3) it is unfeasible (before the time in question, and without peculiarities of time whereby we could get news from the future) to investigate the truth of propositions that divide the cells.
>
> [...] Conditions (1)–(3) characterize the history-theory partition, but not uniquely. [...] Any coarser partition, if it satisfies conditions (1)–(3) according to some appropriate standards of feasible investigation and of natural partitioning, gives us a kind of counterfeit chance suitable for use by determinists: namely, reasonable credence conditional on the true cell of that partition.

Maher first argues that even the history-theory partition may not satisfy the constraints (1)–(3). Let t be the time at which the first tritium atom formed and A the proposition that it still exists 24 hours after t. Suppose the only way to investigate the fundamental dynamical chance of A at t is to observe many tritium atoms and measure the proportion that decay within 24 hours. Then even if agents existed before t, it wouldn't be feasible for them to investigate the chance of A at t (or *a fortiori* the true cell of the history-theory partition $H_t\mathbf{T}_{H_t}$). Next, regarding counterfeit chance, Maher considers a case of bending a coin with a hammer and then tossing it: "Lewis has not indicated what [the relevant] partition might be but this part of his theory is adapted from Jeffrey, who indicates that the partition is one whose elements specify the limiting relative frequency of heads in an infinite sequence of tosses of the coin" (p. 207) but there cannot be such an infinite sequence, and even if existed, its limiting frequency couldn't be investigated prior to t.

Regarding the first point, in my framework I have not put any general constraints on the partitions M that chances are defined relative to, nor on the specification-theory partitions $M\mathbf{T}_M$, except to say that their cells specify the M-chances.[23] That said, I have endorsed a semantic account (Handfield and Wilson, 2014) of how the conversational context can pick out the relevant M under consideration, and this account goes

[23] As Maher (2009, p. 207) writes, "what makes the history-theory partition a suitable one is not the characteristics that Lewis cited, concerning naturalness and feasibility of investigation; it is rather that each element of the history-theory partition specifies the value of the relevant chance".

roughly along the lines of (1)–(3). But crucially, in that account the conditions (1)–(3) are imposed on the specification partition M, not the specification-theory partition $M\mathbf{T}_M$. So in the tritium case, if the agents did have the ability to investigate exact microstates at t (and this was made salient) but nothing after t, then this would pick out H_t as the relevant partition; H_t would satisfy (1)–(3) (even though $H_t\mathbf{T}_{H_t}$ would not), and so unqualified chance ascriptions would be interpreted as ascribing the fundamental dynamical chances at t, which is the desired result.

Regarding the second point, I do not tie my account of non-fundamental chance to Jeffrey (1965, Chapter 12.7) in the way that Lewis does. I do not require that either M or $M\mathbf{T}_M$ is a partition into past or future frequencies, and I certainly do not require either is a partition into limiting frequencies.[24] For a general discussion of how the cells of $M\mathbf{T}_M$ can be investigated, and a discussion of the relationship of my framework to Skyrms's (1977) notion of *resiliency*, which Lewis also references, see the next chapter.

Finally, Maher presents a circularity worry (2009, pp. 207–8):

> [I]f counterfeit chance is to be inductive probability [or rational credence] conditioned on the appropriate element of a suitable partition, the elements of that partition must specify the (true!) value of the counterfeit chance. But then it would be circular to explain what counterfeit chance is by saying that it is inductive probability conditioned on the appropriate element of a suitable partition. Therefore, counterfeit chance cannot be explained in this way—just as chance cannot be explained by saying it is inductive probability conditioned on the appropriate element of the history-theory partition. Thus the account of counterfeit chance, which Lewis adopted from Jeffrey, is misguided. The right approach is to treat what Lewis regards as genuine and counterfeit chance in a parallel fashion.

This would be a worry for me if I were claiming to have offered a metaphysical analysis of all these different kinds of non-fundamental chance. But like Lewis (1980), that is not my aim. Instead, I am pointing out a unifying feature of these different kinds of chance: facts about them constrain credence in the way described by the global principal principle. I am then showing how we can use this as a starting point do some useful theorizing. Indeed, while chances of different kinds may be metaphysically varied, on my account they are viewed as playing the same basic epistemological role. In that sense, my framework follows what Maher calls the "right approach", of treating fundamental and non-fundamental chance in a parallel fashion.

[24] Though incidentally, the latter does not seem to be what Jeffrey had in mind; for the nth trial, he selects the relevant partition to be the different possible frequencies, or different possible sequences, up to n only. (Jeffrey does discuss infinite limits of finite partitions in this part of his book, but that is in a different context, where a continuous quantity like temperature is under discussion (cf. Part III).)

Chapter 2

Knowing the chances

In the Introduction to this dissertation, I listed some probabilistic claims, the likes of which feature prominently in science and in ordinary life:

- The probability that an evenly weighted die lands 'six' is about 1/6.
- The probability that this uranium atom decays in the next year is less than 5%.
- The probability that a cross between two pea plants will yield an offspring that is heterozygous for both seed shape and seed color is approximately $1/2 \cdot 1/2 = 1/4$.
- The probability that atmospheric moist convection will be triggered over the Magdalena mountains today, leading to a late afternoon thunderstorm, is greater than 80%.

What are people looking for, when they search for probabilities like these? One appeal of a framework for non-fundamental chance is that it might allow us to recover, not only an account of these as true claims about the world, but also an account of how we know these claims, and what makes this knowledge so valuable. Drawing on the framework introduced in Chapter 1, this chapter aims to provide such an account.

At first glance, it may seem obvious how we can know the claims above (construed as claims about chance): we can simply count frequencies. Crucially, however, inferences from relative frequencies to chances are unreliable unless certain requisite *structural chance features* hold. The standard assumption, implicit in calculations like $1/2 \cdot 1/2 = 1/4$, is that the trials are *independent* and have *uniform chance*.

It may sound like the height of skepticism to deny this kind of uniformity assumption, akin to denying the uniformity of nature. But as I argue, this uniform structure is often plausibly absent, in many special science examples, at the level of the fundamental dynamical chances.

That said, the requisite uniform structure can still emerge at the level of other chances, based on other subject matters M. So, the focus of this chapter will be applying my framework to show how the emergence of uniform chance structure can be modeled and understood. As far as I know, this is a new application of this kind of framework, as authors with similar frameworks (Handfield and Wilson, 2014; List and Pivato, 2015) have mostly focused on the compatibility between higher-level chance and lower-level determinism. Some of my treatment will draw on work by Strevens (1998, 2003, 2011, 2013) and Werndl (2009, 2012, 2013), and related results in chaos and ergodic theory.

After discussing the emergence of uniform chance structure, I will return to the question of the value of chance knowledge. I will first point out that the uniform chance structure is in itself an important pattern that we'd miss out on if we focused only on Lewisian chances. Then I will illustrate, using the global principal principle and Skyrms's (1977) notion of *resiliency*, how knowledge of the kinds of chances discussed can be especially useful for setting and guiding our expectations.

Before proceeding, an important caveat concerning my use of 'knowledge'. It is convenient to talk about knowledge of a chance hypothesis or chance theory, and indeed I do think that to the extent we have full beliefs about chance hypotheses, we do possess such knowledge. But in keeping with the background Bayesian epistemological approach of Chapter 1, I will really be focused here on rational credence—on how chance hypotheses are learned, on how an agent's credences over chance hypotheses can be rationally distributed given their evidence, and so on. Like Lewis (1980), I will still often use knowledge terminology, but this should be understood as shorthand, not a commitment to a Lockean threshold view of knowledge. A more accurate title for this chapter may thus be "Achieving accurate rational credence in true, non-trivial chance hypotheses". As catchy as this alternative is, I hope the reader will forgive "Knowing the chances".

2.1 Knowing Lewisian chances

I wrote that one appeal of a framework for non-fundamental chance is that it might allow us to recover, not only an account of the four claims above as true claims about the world, but also an account of how we know these claims. The question naturally arises whether expanding beyond Lewisian or fundamental dynamical chance is necessary to provide such an account. I touched on this issue in the previous chapter, when I discussed whether special science probabilities can be recovered as Lewisian chances, but it is worth elaborating on it here, as the discussion will also set the stage for the application of my framework.

Fundamentalism and chance knowledge

Several authors have emphasized the importance of recovering our apparent chance knowledge, some even taking it as a central desideratum on an account of chance. For example, Hicks (2017) suggests modifying the virtue of 'simplicity' in Lewis's Best System Analysis (Lewis, 1994) so that systems in which chances are "more epistemically accessible to us" are more heavily favored. Hoefer (2007, p. 594) writes that "since scientists and policymakers do postulate objective probabilities and try to find out what they are—often with apparent success—philosophers of science should not be content with anything less than a theory of objective chances that entails that they exist, and that we can come to know them." And Handfield and Wilson (2014, p. 22) emphasize that "it is incumbent on us to seek an account of higher-level [objective] probabilities which accords with their epistemic accessibility and usefulness."

Authors have also noted the tension between this desideratum and a commitment to (cf. Chapter 1.1):

Fundamentalism All objective chances are fundamental dynamical chances, i.e. Lewisian chances.

One way to see this tension is to reflect on our ignorance about the fundamental physical theory and the fundamental dynamical laws. Fundamental physics is still far from settled, with attempts to reconcile general relativity and quantum field theory remaining inconclusive,[1] and with many debates about how to understand the status of the dynamical laws even in more basic theories, like non-relativistic quantum mechanics, still ongoing.[2] It is difficult to see how someone committed to **Fundamentalism** could be confident that we have come to know any non-trivial objective chances. As Builes and Spencer (ms) write:

> Fundamental [dynamical] chance plays little or no role in our [epistemic] lives. We often conform our credence to what we know about the chances. But, since the fundamental theory of physics is yet to be discovered, it is unclear whether anyone has ever known anything about fundamental chance, and thus it is unclear whether anyone has ever conformed their credence to something they have known about the fundamental chances.

A second, related, way to see this tension is to note that if we knew the Lewisian chance of a die landing 'six' was 1/6, of a thunderstorm arriving was 80%, and so on, then we'd be in a position to know that the

[1] As Butterfield and Isham (2001) write, it is very much an open possibility that "classical general relativity and standard quantum theory emerge from a theory that looks very different from both."

[2] See footnote 4 for details.

universe is fundamentally indeterministic.[3] But we are not currently in such a position.[4] Moreover, as we saw in the case of the indeterministic roulette wheel from Chapter 1, Example 2, it is not as though simply assuming indeterminism will then guarantee that the Lewisian chances correspond to the probabilities we have in mind. (The only proposal I know of that could ensure such a correspondence is Albert's sketch (2000, Chapter 7) of how chances in GRW theory could give rise to statistical mechanical probabilities. This proposal is interesting, but remains extremely tentative.[5]) Again I am not alone in noting this tension. Hoefer remarks that philosophers committed to **Fundamentalism** "[make] themselves hostage to the fortunes of determinism in physics" (2007, p. 577). And Handfield and Wilson offer a kind of variant of my argument above: "If the usefulness of chance-theorizing at the non-fundamental level depended on underlying non-trivial fundamental chances, then the usefulness of chance-theorizing would provide us with compelling evidence that fundamental physics is indeterministic. But we have no such evidence" (2014, p. 22).

I should clarify that I am not arguing **Fundamentalism** is inconsistent with us ever having known any non-trivial objective chances. Say a physicist justifiably believes ⟨the probability that this uranium atom decays in the next year is 4%⟩ and is agnostic about whether this is an objective chance or some other kind of probability; then it may turn out that the Lewisian chance of decay is indeed 4%, and so

[3]Because these chances are non-trivial and (see Chapter 1.1) non-trivial Lewisian chances are incompatible with fundamental determinism. That said, one might object to this reasoning for relying on a suspicious closure assumption. Consider the claim **Fundamentalism-Hands**: brains in vats do not have hands. Analogous reasoning would produce: *Suppose we knew we had hands. Then assuming* **Fundamentalism-Hands***, we'd be in a position to know we're not brains in vats.* This reasoning looks suspicious, or at least, controversial. So what is the disanalogy? Here is a place where the caveat about my use of the term 'knowledge' is relevant. Translated to the setting of rational credence, this reasoning becomes: if C(we have hands|BIVs have no hands) is high, then C(we're not BIVs|BIVs have no hands) is high. But I actually do endorse this reasoning (for rational credence, I think that if B entails A, then $C(A) \geq C(B)$; assuming conditionalization this then implies that $C(D|F) \geq C(H|F)$ when HF entails D, which is the reasoning above). So, there is no failure of parity to explain.

[4]For reasons similar to those raised in the previous paragraph. For example, even supposing non-relativistic quantum mechanics were true, it would still not be clear whether the universe is indeterministic and whether quantum probabilities are Lewisian chances. On some interpretations or theories of quantum mechanics, like the orthodox interpretation, GRW theory, and the Penrose interpretation, there is a fundamental "collapse rule" built into the dynamics. But on other interpretations, there is no such rule, and the dynamical laws themselves don't feature any probabilities. Bohmian Mechanics, Everettian Mechanics, and decoherence-based interpretations, on which isolated systems evolve unitarily, fall into this category. (On these interpretations, an additional explanation must be offered for why these probabilities correspond to our observed statistics, even though they don't play an active dynamical role. The explanation often involves a postulate about typical initial conditions; for a review of some options in the context of Bohmian mechanics for example, see Sebens (2015).) At the moment there is no empirical test which would distinguish between these interpretations. Extra-empirical considerations involving various theoretical virtues are brought to bear, but they are far from conclusive.

These complications do not dissolve when we move to more general settings like quantum field theory and the frontiers of quantum gravity. Many quantum gravitational models have the quantum state (the density matrix) of the universe evolving unitarily and hence deterministically (see e.g. Hooft (1999)); some don't, but for new reasons (Page, 1982). Interesting twists have also emerged in attempts to extend existing interpretations of quantum mechanics. For example, in the Bohmian quantum field theories proposed by Bell (1987) and Dürr et al. (2004), the dynamical laws are probabilistic. However, the probabilities that feature in the dynamical laws are not the usual Born rule quantum probabilities we hone in on in laboratories and particle accelerators. Rather, they correspond to the chances of certain stochastic jumps in the universe's total particle configuration number. (In Bell's theory, one may be able to remove this stochasticity in the continuum limit, though this will not be possible for some other choices of beables (Colin, 2003). But in any case, Dürr et al.'s proposal is formulated in the continuum limit and it is stochastic.)

[5]As Albert (2015) acknowledges. GRW theory is one of the indeterministic interpretations of quantum mechanics. Roughly, Albert's idea is that the spontaneous collapses posited by this theory will result in large systems getting perturbed from non-thermodynamic micro-conditions into thermodynamic ones, and so there will be no need for an initial probability distribution over micro-conditions (cf. Chapter 1.8 and Albert (2015, p. 40)). As Sklar (2015) writes, "little has yet been done to fill in the details to see if the posited GRW processes could, if real, account for the known thermodynamic asymmetries. And, of course, there is much skepticism that the GRW processes are even real."

there will be a sense in which this physicist knows the value of a non-trivial objective chance, even under **Fundamentalism**. My point is rather that, given **Fundamentalism**, we cannot be confident that we have ever known any non-trivial objective chances. So if we want a plausible account that secures and explains (at the current moment) our having non-trivial chance knowledge, we will have to drop **Fundamentalism**.

The principal principle to the rescue?

What I have just said may seem in conflict with Lewis (1980), in particular Lewis's discussion of "Chance of frequency" and "Evidence about chances". In these passages, Lewis considers an agent conducting a coin toss experiment. Using the principal principle, Lewis derives the following conclusion (1980, p. 287):

> Unless the prior [credence] distribution is irremediably biased, the result that after enough tosses is that the lion's share of the posterior credence will go to the hypotheses putting the single-case chance of heads close to the observed frequency.

This seems to suggest that, according to Lewis, we can (and have) quite easily come to know the chances of coin toss outcomes; the principal principle guarantees a nice connection between chance and frequency. On the other hand, Lewis endorses **Fundamentalism**. How can these two things be squared?

The answer is that there is an important qualification to Lewis's result. In his derivation, Lewis was careful to build in two assumptions concerning the beliefs of the agent in question (1980, p. 284):

1. The agent has high credence that the tosses are independent (I will give a precise definition shortly).
2. The agent has high credence that there is a uniform chance of heads across tosses, in other words, that the tosses are identically distributed (I will give a precise definition shortly).

This kind of uniform chance structure is crucial to the derivation, and is precisely the kind of structure that is this focus of the chapter. As I flagged, this kind of structure is needed for the connection between chance and frequency to hold, even given the principal principle. In this important sense, there is more to the epistemology of chance than the principal principle.

In a postscript, Lewis (1987) clarifies that he is not actually confident that this kind of structure obtains in this situation, or that the single-case chance of heads is close to the observed frequency:

> If my example concerned the beliefs of an ignoramus, it is none the worse for that. [...] I [myself] would give a small share of credence to the proposition that [the single-case chance of heads at the time of toss] is zero exactly, and an equal small share to the proposition that it is one exactly. I would divide most of my credence between the vicinity of 50%, the vicinity of zero, and the vicinity of one.

Lewis thought that 50% may still be plausible because, for instance, the air resistance experienced by the coin will depend partly on chancy processes involving the chemical bonds between the coin and the air molecules; it is not unreasonable to suppose, the thought goes, that these kind of ultimately quantum mechanical chance processes may culminate in a 50% probability. Incidentally, more recent experiments by Diaconis et al. (2007), in which the result of the toss (performed by a coin tossing machine) can be almost perfectly predicted from the initial configuration, suggest this is not the case.

Inferring Lewisian chances from frequencies

I have claimed that inferences from relative frequencies to chances are unreliable without the presence of 'uniform chance structure', and also that this kind of structure is often absent in the case of Lewisian chances. I will elaborate formally on what I mean by 'uniform chance structure' in a moment, but first let me explain intuitively why these two claims are true.

Consider the case of tossing a coin discussed above. Suppose, as seems plausible, that the fundamental dynamical chance of a particular toss outcome (evaluated at the time of the toss) is highly non-uniform across trials: it is either zero, one, approximately zero, or approximately one, depending on the trial. Then any inference from frequency to the fundamental dynamical chance of a particular toss outcome will be highly unreliable: the frequencies will suggest ∼50%, but the true chance will be far from this.[6]

Similarly, consider the indeterministic wheel of fortune case from Chapter 1.1, Example 2. The fundamental dynamical chance of particular WHITE or BLACK outcome (evaluated at the time just before the particle enters the magnet) will be either zero, 30%, 70%, or one, depending on the exact initial configuration of the wheel and the speed of the particle. Again an inference from the frequency to any of these fundamental dynamical chances will be unreliable: the frequency will suggest ∼50%, but the true chance will be at least ∼20% off the mark.

In addition to whether the chances are uniform across trials, another factor is whether the trials are independent, in the sense of whether the outcome of one trial will affect the chance of the next. Imagine a 'quantum coin' set-up in which the fundamental dynamical chance of heads is indeed 50% for each trial. But, the coin has a memory property such that if it lands heads or tails twice in a row then it will be extremely likely to land that same way on subsequent trials. Here, while the toss outcomes have uniform chance in a sense (they have uniform chance as evaluated at the time before the first toss), they are not chance-wise independent. Note the long run relative frequency will likely be close to either ∼0% or ∼100% heads, which is not a reliable indication of the fundamental dynamical chance of heads for the first toss, namely 50%.

[6] The frequencies may provide a kind of *average chance*; but an average chance is not a chance, or at least, not a Lewisian chance.

This example of failure of independence is quite artificial, but we can also see violations of independence (relative to the fundamental dynamical chances) in our other examples. Recall the situation with the indeterministic wheel of fortune pictured in Figure 2.1, in which the the pointer starts out pointing toward the left-hand-side of a black wedge. We imagine the first incoming particle, particle 1, is moving quite slowly, such that depending on whether the wheel is hit on the left or right side, it will either (a) shift so that the pointer points toward the adjacent white wedge, or (b) shift so that the pointer points toward the very edge of the same black wedge (scenarios (a) and (b) are depicted by the left and right arrows respectively).

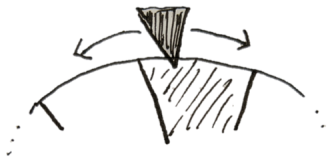


Figure 2.1: Reproduction of Figure 1.4.

Denoting the time before this first particle enters the magnet by t and this world by w, we have $P_{tw}(\text{WHITE}_1) = 30\%$ and $P_{tw}(\text{BLACK}_1) = 70\%$. Now, consider the particle destined to hit the wheel next, particle 2. Suppose that (at this world w) this particle is moving extremely slowly, such that it will only perturb the wheel by a very tiny amount. Then $P_{tw}(\text{WHITE}_2|\text{WHITE}_1) = 100\%$, since in scenario (a), a tiny perturbation in either direction is not enough to break free of the white wedge. On the other hand, $P_{tw}(\text{WHITE}_2|\text{BLACK}_1) = 70\%$, since in scenario (b), a tiny clockwise perturbation is enough to break free of the black wedge.[7] It follows that $P_{tw}(\text{WHITE}_2) = 79\%$.[8] And so $P_{tw}(\text{WHITE}_2|\text{WHITE}_1) \neq P_{tw}(\text{WHITE}_2)$. Relative to the fundamental dynamical chances P_{tw}, the outcomes WHITE_1 and WHITE_2 are not independent.[9]

Independence is not strictly necessary for inferences from frequencies to chances to be reliable. If the trials are *exchangeable*, meaning that permuting them does not affect their joint chance distribution, this will suffice.[10] However, one can check that, in the cases above, exchangeability also fails with respect to the

[7] Here it is important that the particles don't change speeds in a chancy way, and that $H_t(w)$ specifies that the exact speeds of both the first and second particles, as well as the initial configuration of the wheel.

[8] $P_{tw}(\text{WHITE}_2) = P_{tw}(\text{WHITE}_2|\text{WHITE}_1) \cdot P_{tw}(\text{WHITE}_1) + P_{tw}(\text{WHITE}_2|\text{BLACK}_1) \cdot P_{tw}(\text{BLACK}_1) = 100\% \cdot 30\% + 70\% \cdot 70\% = 79\%$.

[9] The reader may have noticed that I have been focusing on the fundamental dynamical chances before the first trial, and evaluating independence with respect to this chance function. As Lewis (1980, p. 283) explains, there are multiple independence questions one can ask here. Let $t = 1, 2$ denote the times of the first two trials, and A_2, A_3 the possible outcomes. One question is whether the true chance theory T is such that further facts about the intervening history between times 1 and 2, *other than* the fact that A_2 occurs, might affect the chance of A_3. If not then we will have, by the global principal principle, $P_{2w}(A_3) = C_0(A_3|A_2H_1(w)T) = P_{1w}(A_3|A_2)$, which is one kind of independence. Another question is whether the true chance theory T is such that *any* facts about the intervening history between times 1 and 2, including the occurrence of A_2, might affect the chance of A_3. If not, then we will have, by the global principal principle, $P_{2w}(A_3) = C_0(A_3|H_1(w)T) = P_{1w}(A_3)$. It then follows that $P_{1w}(A_3|A_2) = P_{1w}(A_3)$, which is a second kind of independence. It is this second kind of independence, relative to a single chance function, that I have been discussing and that is assumed in standard derivations of the connection between chance and frequency, like the Law of Large Numbers (cf. Section 2.2).

[10] Exchangeability in the case of propositions (i.e. binary random variables) implies that relative frequency is a *sufficient statistic*; see for example (Diaconis and Freedman, 1984; Lauritzen, 1988).

fundamental dynamical chances. For example, in the wheel of fortune example, $P_{tw}(\text{WHITE}_1, \text{BLACK}_2)$ does not equal $P_{tw}(\text{BLACK}_1, \text{WHITE}_2)$. And in the coin with memory example, $P_{tw}(\text{HEADS}_1, \text{TAILS}_2, \text{HEADS}_3)$ is much greater than $P_{tw}(\text{HEADS}_1, \text{HEADS}_2, \text{TAILS}_3)$, since the latter involves a tails after a string of two heads.

To sum up, uniform chance structure is often absent at the level of the Lewisian chances. In the absence of this kind of structure, relative frequencies are not a direct guide to these chances. Of course, this does not mean that this structure is absent at the level of other (non-Lewisian) chances, or that relative frequencies are not a direct guide to these other chances. This is the possibility I will explore next. For this task, it will be useful to have general definitions of these structural chance features. For a given factual subject matter M, I will say propositions $A_1, A_2, ...$ are *M-independent according to theory T at w* if

$$P_{M(w)}(A_1 A_2 ... A_n) = P_{M(w)}(A_1) \cdot P_{M(w)}(A_2) \cdot ... \cdot P_{M(w)}(A_n).$$

Propositions $A_1, A_2, ..., A_n$ are *M-identically distributed according to theory T at w*, in other words have uniform M-chance, if

$$P_{M(w)}(A_1) = P_{M(w)}(A_2) = ... = P_{M(w)}(A_n).$$

Propositions $A_1, A_2, ..., A_n$ are *M-exchangeable according to theory T at w* if, according to $P_{M(w)}$, the chance of exactly k of them being true, whichever k are selected, is the same, for all $1 \leq k \leq n$. (In the case where we have an infinite sequence $A_1, A_2, ...$ of propositions we say they are M-independent, M-identically distributed, and so on, if every finite subcollection of them is.) For instance, in the indeterministic roulette wheel example we found that WHITE_1 and WHITE_2 were neither H_t-independent, nor H_t-identically distributed, nor H_t-exchangeable, at w. Later I will argue that for an appropriate coarse-graining $\tilde{H}_t$, they nevertheless possess these features with respect to the $\tilde{H}_t$-chances.

Other methods for learning chances

Before moving on, I should clarify that counting frequencies is not the only way to learn Lewisian chances, or chances in general. Our inferences to chance hypotheses are generally quite complicated and need not be based on simple enumerative induction. Consider, for instance, how meteorologists estimate the probabilities of hurricane trajectories. In addition to historical data, they use dynamical models that draw on approximations of the physics of the atmosphere. This physics is developed, in turn, by a more complex process of scientific investigation.

The absence of uniform chance structure makes it difficult to learn chances directly by counting frequencies, but it is not necessarily a barrier to these other methods. Indeed, at least in principle, other methods

can succeed even when the chance structure is highly non-uniform.

Example 16. Imagine an agent knows that the fundamental dynamical chance of a particle going up through a Stern Gerlach magnet is a non-trivial function of its spin, but she isn't sure what the function is. Suppose she is interested in the chance x of a particular particle with spin s_α going up. Observing a different particle with spin s_β go up may—even though this event has a different chance (and, we can suppose, the agent knows this)—make her confident of the value of x, because it greatly informs her of this function.

Another case may be if all of the agent's candidate indeterministic fundamental physical theories are *ad hoc*, complicated, and so on, then she may be reasonably confident that the fundamental dynamical chances at the current time t are all either zero or one. Of course, these are just toy examples, and as I suggested in Section 2.1, we are not currently in a position to draw such conclusions about the Lewisian chances. But the point is that we do not do full justice to the epistemology of chance if we focus only on the method of counting frequencies.

To this end, it is worth briefly discussing how the framework of Chapter 1 might model these other epistemic methods. It is controversial whether a Bayesian framework can accommodate Ockham's razor, inferences to the best explanation, and other kinds of scientific theorizing (Earman, 1992; Forster, 1995; Bovens et al., 2003; White, 2005; Howson and Urbach, 2006; Dowe et al., 2007; Glymour, 2010; Belot, 2017). I will not delve into these controversies here. Instead I take up the much more modest goal of modeling one case, Example 16, in my framework.

Example 16 (continued)**.** Let w be the world in question and suppose that the agent's total evidence E includes the history up to t, $H_t(w)$. The agent's prior and E are such that she is only entertaining two global chance theories, $T, T' \in \mathbf{T}$, and she is equally confident in each, so $C_0(T|E) = C_0(T'|E) = 1/2$. According to T, the fundamental dynamical chance of a particle going up is a function $d(s)$ of its spin, which we suppose for simplicity does not change with time. T' is similar but the associated function $d'(s)$ is distinct. Now suppose there are two particles α and β with different initial spins s_α and s_β. For concreteness let us say:

$$d(s_\alpha) = \frac{2}{3},\ d(s_\beta) = \frac{4}{5}, \qquad d'(s_\alpha) = \frac{1}{4},\ d'(s_\beta) = \frac{1}{100}.$$

Let A be the proposition that the α particle goes up at $t+2$ and let B be the proposition that the β particle goes up at $t+1$. So originally the agent is equally split between two chance hypotheses concerning A. In particular, letting $C = C_0(\cdot|E)$, and letting $D_A = \langle P_{tw}(A) = 2/3\rangle$, $D'_A = \langle P_{tw}(A) = 1/4\rangle$,

$D_B = \langle P_{tw}(B) = 4/5\rangle$, and $D'_B = \langle P_{tw}(B) = 1/100\rangle$, we have:

$$C(D_A) = \frac{1}{2}, \qquad C(D'_A) = \frac{1}{2}, \qquad C(D_B|D_A) = 1, \qquad C(D'_B|D'_A) = 1.$$

To see $C(D_B|D_A) = 1$ for instance, recall $D_A = \langle P_{tw}(A) = 2/3\rangle$ is by definition the disjunction of members of $H_t\mathbf{T}$ which agree with this chance assignment, and the only such member supported by C is $H_t(w)T$, since E is incompatible with other cells of H_t, and T is the only chance theory under consideration such that $P_{H_t(w)}(A) = P_{tw}(A) = \frac{2}{3}$. Next note that $H_t(w)T$ entails $D_B = \langle P_{tw}(B) = 4/5\rangle$, since $d(s_\beta) = \frac{4}{5}$.

Now suppose the agent learns B, and nothing else. Then we have:

$$\begin{aligned}
C(D_A|B) &= \frac{C(D_A)\cdot C(B|D_A)}{C(B)} \\
&= \frac{C(D_A)\cdot C(B|D_A)}{C(D_A)\cdot C(B|D_A) + C(D'_A)\cdot C(B|D'_A)} \\
&= \frac{\frac{1}{2}\cdot C_0(B|D_B D_A E)}{\frac{1}{2}\cdot C_0(B|D_B D_A E) + \frac{1}{2}\cdot C_0(B|D'_B D'_A E)} \\
&= \frac{\frac{1}{2}\cdot\frac{4}{5}}{\frac{1}{2}\cdot\frac{4}{5} + \frac{1}{2}\cdot\frac{1}{100}} \\
&= \frac{80}{81} > \frac{1}{2},
\end{aligned}$$

where the first equality follows from Bayes' formula, the second from the fact that $\langle P_{tw}(A) = 2/3\rangle$ and $\langle P_{tw}(A) = 1/4\rangle$ are disjoint propositions whose credence sums to 1, the third from the facts noted above and the definition of C, and the fourth from the global principal principle and the fact that the conjunctions $\langle P_{tw}(A) = 2/3\rangle E$ and $\langle P_{tw}(A) = 1/4\rangle E$ are weakly admissible with respect to H_t.

So, even though the agent knows the H_t-chances are not uniform across A and B (according to T, one has chance $\frac{2}{3}$ and the other $\frac{4}{5}$, and according to T', one has chance $\frac{1}{4}$ and the other $\frac{1}{100}$) learning B still greatly confirms the agent's hypothesis D_A that the H_t-chance of A is $\frac{2}{3}$.

In this example, the agent is certain of the factual information grounding the chance, but unsure which chance theory is true; the complete H_t-specification is part of her evidence, but she is uncertain whether T or T' is the true chance theory. But we can also imagine that she is uncertain of the facts, like the initial spins of the particles. (Note that an agent can be certain of the true M-chance theory but, if she is uncertain about the facts, may still be uncertain of the M-chances. For example, consider $M = H$. I know the true chance theory based on H is the omniscient chance theory, described in Example 7, that assigns 1 to the truths and 0 to the falsities. But I know relatively little about the H-chances; for example, I do not know

the H-chance of rain tomorrow.) To deal with this complication we could apply the common Bayesian trick $C(\cdot) = \sum_{w_i} C(H_t(w_i)) \cdot C(\cdot|H_t(w_i))$, of factoring the agent's credences as a weighted sum over the partition of H_t-specifications, and then we could apply our earlier treatment to the individual terms $C(\cdot|H_t(w_i))$.

2.2 The emergence of chance structure

So far we have been focused on Lewisian chances. Let us now turn to consider chances of other kinds.

Uniform structure at higher levels

One of the main goals of this chapter is to show how uniform chance structure can plausibly emerge at higher levels, even when it is absent at the level of Lewisian chance. To do this, I will expand on the example of the *baker's map*, introduced in Chapter 1, Example 12. I will repeat all the relevant set-up here.

For concreteness, we fix a toy setting where time evolves in discrete steps $\mathbb{T} = \{0, 1, 2, ...\}$ and the possible states of the world at a given time can be specified by a point on the unit square, $S = [0, 1] \times [0, 1]$. Each possible history is a sequence $(s_0, s_1, ...)$ of such points. We suppose the evolution from one state s_t to the next s_{t+1} is given by the deterministic map $U : S \to S$:

$$U(x, y) = \begin{cases} (2x, \frac{1}{2}y) & \text{if } 0 \leq x < \frac{1}{2}, \\ (2x - 1, \frac{1}{2}(y + 1)) & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

We can think of this set-up, slightly more realistically, as a model for the situation pictured in Figure 2.2 (Pitowsky, 1996, pp. 166–7). A particle starts in motion from a square, moving perpendicular to it. It then bounces on a series of mirrors, causing it to return to some spot on the square, and then the process repeats.

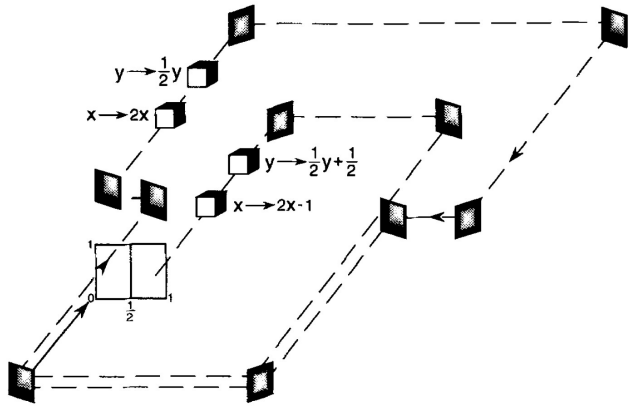


Figure 2.2: Reproduced with permission from (Pitowsky, 1996, Figure 3).

Let $\Omega \subseteq \mathcal{W}$ denote the set of worlds whose histories can be described as a sequence of points in S, and $\mathbf{\Omega} \subseteq \Omega$ the set of those which are consistent with U. Fix $w \in \mathbf{\Omega}$ with initial state $s_0 = (\frac{3}{10}, \frac{1}{2})$. Then its history $H(w)$ can be thought of as the sequence:

$$(U_0(s_0), U_1(s_0), U_2(s_0), U_3(s_0), ...) = ((\tfrac{3}{10}, \tfrac{1}{2}), (\tfrac{3}{5}, \tfrac{1}{4}), (\tfrac{1}{5}, \tfrac{5}{8}), (\tfrac{2}{5}, \tfrac{5}{16}), ...),$$

as shown in Figure 2.3, where here U_t means that U is applied t times. Suppose that w is not only consistent with U, but has U as its fundamental dynamical law. Then all the Lewisian chances at w are trivial, either zero or one. Put more precisely, let $A \subseteq \Omega$ be any proposition that is a disjunction of histories; note if that is the case, then we can think of A as a set of sequences (for example, Ω the set of all sequences, and $\mathbf{\Omega}$ the set of all sequences that are nomologically possible at w). Then for any time $t \in \mathbb{T}$, the fundamental dynamical chance $P_{tw}(A)$ equals 1 if A contains the special sequence above and 0 if it does not.

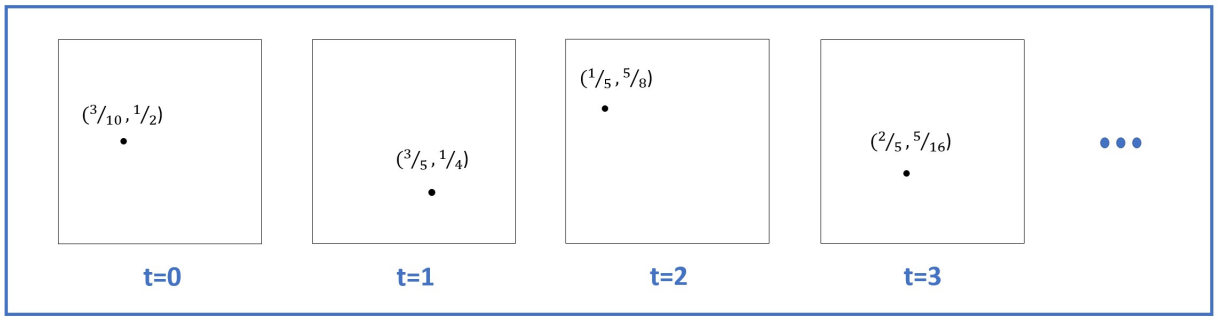


Figure 2.3: The history $H(w)$.

Now, consider the coarse graining function $\Phi : S \to \tilde{S}$ with $\tilde{S} = \{L, R\}$ that only specifies whether the particle is in the left or right side of the square:

$$\Phi(x, y) = \begin{cases} L & \text{if } 0 \leq x < \frac{1}{2}, \\ R & \text{if } \frac{1}{2} \leq x \leq 1. \end{cases}$$

The coarse-grained history $\tilde{H}(w)$ of the world described above can then be thought of as the sequence:

$$(\Phi(U_0(s_0)), \Phi(U_1(s_0)), \Phi(U_2(s_0)), \Phi(U_3(s_0)), ...) = (L, R, L, L, R, ...),$$

as shown in Figure 2.4. We can also consider the proposition R_t that the world is in state R at time t. This

can be thought of as the set of all coarse-grained sequences with R as their t^{th} component.[11]

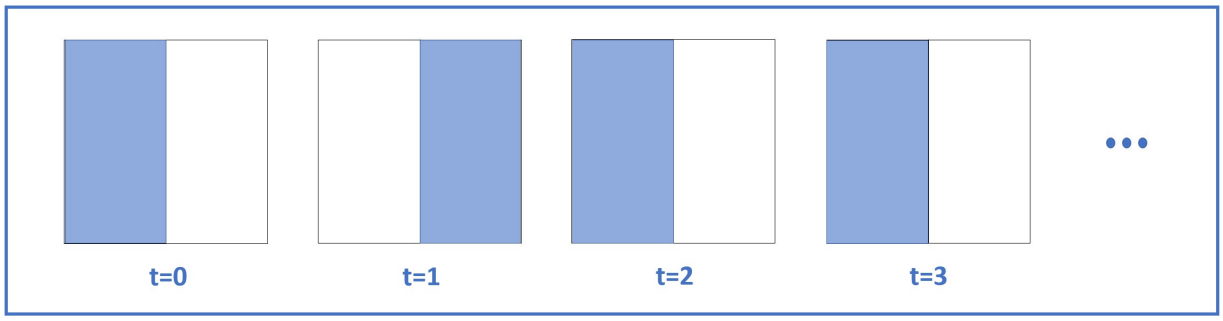


Figure 2.4: The coarse-grained history $\tilde{H}(w)$.

At each time t, we can then think of the 'trial' as whether the outcome R will occur at $t+1$, in other words whether R_{t+1} is true or false. For the Lewisian chances, the chances are not uniform across trials, but rather vacillate between zero and one according to the actual history $(L, R, L, L, R, ...)$:

$$P_{0w}(R_1) = 1, \qquad P_{1w}(R_2) = 0, \qquad P_{2w}(R_3) = 0, \qquad P_{3w}(R_4) = 1, \qquad ....$$

Nor are the trials H_t-identically distributed for any fixed t. To see this, note that they are not H_0-identically distributed (since $P_{0w}(R_1) = 1$ and $P_{0w}(R_2) = 0$) and $P_{0w}(\cdot) = P_{tw}(\cdot)$ for all t.

What about non-fundamental chances? In addition to chances based on the maximally detailed initial segment $H_t(w)$, we can also consider chances based on the coarse-grained segment $\tilde{H}_t(w)$. Note that these coarse-grained initial segments do not specify the initial microstate s_0 but rather the initial macrostate. For example, $\tilde{H}_2(w)$ can be thought of as the sequence $(\tilde{s}_0, \tilde{s}_1, \tilde{s}_2) = (L, R, L)$. Following the notation introduced in Chapter 1.6, I will write P_{tw} for $P_{H_t(w)}$ and $\tilde{P}_{tw}$ for $P_{\tilde{H}_t(w)}$. The question is then whether uniform chance structure can emerge at the level of $\tilde{P}_{tw}$, even though it is absent at the level of P_{tw} (cf. List and Pivato (2015, p. 138), Werndl (2009, 2012, 2013)).

To illustrate a precise sense in which uniform chance structure can emerge in this example, I will help myself to three assumptions in addition to the global principal principle. These can be viewed as 'dummy' assumptions adopted for the purpose of being able to make concrete calculations, or they can be taken at face value as claims about the true chance theory T at this world w. In what follows, for a map $f : X \to Y$, I will write the preimage of a set $B \subseteq Y$ under f as $f^{-1}(B) = \{x \in X : f(x) \in B\}$. I will write the image of

[11] I say "can be thought of" because technically I am representing propositions as subsets of $\mathcal{W}$. To define the propositions in question precisely, we use the apparatus of *extraction functions*; see Chapter 1.6 and also footnote 13 below.

a set $A \subseteq X$ under f as $f[A] = \{f(x) : x \in A\}$. If $f : X \to Y$ and $g : Y \to Z$ are functions I will write their composition as $g(f(\cdot))$ or alternatively as $g \circ f$. It will be a useful fact that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Recall the proposition Ω which simply restricts attention to worlds whose histories are sequences of states from S, but does not specify any information about which sequence is actual. My first assumption is that P_Ω, the chance function that the true global chance theory T assigns given Ω, is non-trivial and assigns probabilities to all sets of sequences.[12] My second assumption is that this chance function does not favor any particular initial states over others. In particular, letting $\mu(\cdot)$ denote the uniform (Lebesgue) measure over the square $S = [0,1] \times [0,1]$ and $\mathbf{0} : \Omega \to S$ the map that takes a world $v \in \Omega$ and outputs its initial state,[13] I assume:

$$P_\Omega(\mathbf{0}^{-1}(\cdot)) = \mu(\cdot). \tag{2.1}$$

My third assumption is that P_Ω only allows sequences which are nomologically possible:

$$P_\Omega(\mathbf{\Omega}) = 1. \tag{2.2}$$

Fact 1. *Fix this set-up, including the global principal principle and these three assumptions. Then for all t, $R_{t+1}, R_{t+2}, ...$ are $\tilde{H}_t$-independent and identically distributed at w, with uniformity across trials:*

$$\tilde{P}_{0w}(R_1) = \tilde{P}_{1w}(R_2) = ... = \tilde{P}_{tw}(R_{t+1}) = \tilde{P}_{(t+1)w}(R_{t+2}) = ... = \tfrac{1}{2}.$$

In fact, it would suffice to make two weaker assumptions: first, that the true chance theory T at w is such that $C_0(\mathbf{0}^{-1}(\cdot)|\Omega T) = \mu(\cdot)$ and, second, $C_0(\mathbf{\Omega}|\Omega T) = 1$; in other words, given only Ω and T one is rationally required not to favor any particular initial states, or assign positive credence to nomological impossibilities. However, it is cleaner to just assume a there is a function P_Ω that T assigns given Ω, and work with P_Ω.

Here is the plan for the proof of Fact 1. For the first step, we draw on a result from ergodic theory that shows a precise sense in which the system of the square $[0,1] \times [0,1]$, equipped with the Lebesgue measure μ and the dynamical map U, is equivalent to a Bernoulli fair coin system (cf. Werndl (2009)). Then we apply this result, together with the three assumptions above, to show that P_Ω will be a Bernoulli distribution with $P_\Omega(R_t) = \frac{1}{2}$ and $P_\Omega(R_t R_{t+1}) = P_\Omega(R_t) \cdot P_\Omega(R_{t+1})$ (in a sense this is the most interesting step, because it shows how results from ergodic theory can be applied to the current framework). Finally we invoke the global principal principle, in particular alignment (see Chapter 1.7), to show $\tilde{P}_{tw}(\cdot) = P_\Omega(\cdot|\tilde{H}_t(w))$, and then we invoke the previous steps to derive the result that $R_{t+1}, R_{t+2}, ...$ will be $\tilde{H}_t$-independent and $\tilde{H}_t$-identically

[12] Here I can simply write P_Ω, without reference to a subject matter, because I am assuming the global principal principle; see Chapter 1.7.

[13] We can define this map in terms of the *extraction functions* introduced in Section 1.6: $\mathbf{0}(v) = f_0(H(v))$.

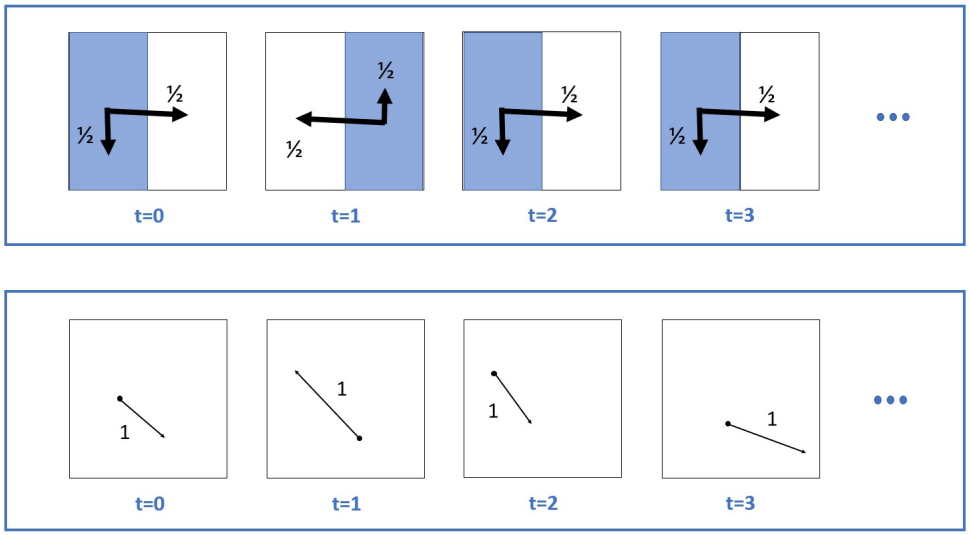


Figure 2.5: Comparison of the fundamental (bottom) and emergent (top) dynamical chances.

distributed as claimed, with the same fair coin distribution regardless of t.

Now for the first step. Since the square $S = [0,1] \times [0,1]$ has uncountably infinitely many points and we are considering continuous distributions over these points, we require some technical preliminaries from probability theory that were not needed previously.

Definition 1. A triple $(S, \mathcal{S}, \mu)$ is a *probability space* if S is a set, $\mathcal{S}$ is a sigma-algebra on S, i.e. a set of subsets of S that contains S and is closed under complement and countable union, and $\mu : \mathcal{S} \to [0,1]$ is a probability measure, i.e. satisfies $\mu(S) = 1$ and $\mu(\cup_i J_i) = \sum_i \mu(J_i)$ for disjoint sequences $J_1, J_2, ... \in \mathcal{S}$.

We next introduce the notion of a *factor* from ergodic theory.

Definition 2. Let $(S, \mathcal{S}, \mu)$ and $(V, \mathcal{V}, \nu)$ be probability spaces and let $U : S \to S$ and $U' : V \to V$ be (dynamical) maps. Then the combined space $(V, \mathcal{V}, \nu, U')$ is a *factor* for the combined space $(S, \mathcal{S}, \mu, U)$ if

1. there is an $\hat{S} \subseteq S$, $\hat{V} \subseteq V$ with $\mu(\hat{S}) = \nu(\hat{V}) = 1$, with $U[\hat{S}] \subseteq \hat{S}$ and $U'[\hat{V}] \subseteq \hat{V}$, such that
2. there exists a surjective map $\phi : \hat{S} \to \hat{V}$, which is measurable, in the sense $\phi^{-1}(K) \in \mathcal{S}$ for all $K \in \mathcal{V}$, $K \subseteq \hat{V}$, and measure preserving, in the sense $\mu \circ \phi^{-1}(K) = \nu(K)$ for all $K \in \mathcal{V}$, $K \subseteq \hat{S}$, and
3. for all $s \in \hat{S}$, $\phi \circ U(s) = U' \circ \phi(s)$.

The following is a well-known example of a factor (Petersen, 1989; Dajani and Dirksin, 2008):

Lemma 1. *Let $(S, \mathcal{S}, \mu, U)$ be the system described above where $S = [0,1] \times [0,1]$, $\mathcal{S}$ is the Borel sigma-algebra on S, μ is the Lebesgue measure, and U the baker's map. Let $(V, \mathcal{V}, \nu, U')$ be the Bernoulli system*

where V is the set $\{0,1\}^{\mathbb{Z}}$ of bidirectionally infinite binary sequences $(...,a_{-1},a_0,a_1,...)$ with $a_i \in \{0,1\}$, $\mathcal{V}$ the Borel sigma-algebra on V, ν the 'fair coin' measure, and U' the shift operator that shifts a sequence $(...,a_{-1},a_0,a_1,...)$ to $(...,a_0,a_1,a_2,...)$. Then $(V,\mathcal{V},\nu,U')$ is a factor for $(S,\mathcal{S},\mu,U)$, with $\phi : \hat{S} \to \hat{V}$ the map defined by $\hat{S} = [0,1)\times[0,1)$ and $\phi(x,y) = (...,a_{-1},a_0,a_1,...)$ where $\sum_{n=1}^{\infty}\frac{a_{n-1}}{2^n}$ is the binary expansion of x and $\sum_{n=1}^{\infty}\frac{a_{-n}}{2^n}$ is the binary expansion of y.

One can also check the following:

Lemma 2. *Let ϕ be the map from the previous lemma and let $\Phi : S \to \tilde{S}$ be the coarse-graining function defined above. Then for all $s \in \hat{S}$, $\Phi(s) = \Phi_0 \circ \phi(s)$, where $\Phi_0 : V \to \tilde{S}$ is the map that takes a binary string $v = (...,a_{-1},a_0,a_1,...)$ and outputs L if $a_0 = 0$ and R if $a_0 = 1$.*

Now for the next step, first note that for any $t \in \mathbb{T}$, we can write:

$$R_t\mathbf{\Omega} = \{w \in \mathbf{\Omega} : \Phi \circ U_t \circ \mathbf{0}(w) = R\},$$

since the history of any $w \in \mathbf{\Omega}$ must be consistent with the deterministic map U. Next let $\hat{\mathcal{W}} = \mathbf{0}^{-1}(\hat{S})$. Defining $\hat{\mathbf{0}} : \hat{\mathcal{W}} \to \hat{S}$ as the restriction of $\mathbf{0}$ to $\hat{\mathcal{W}}$, it follows

$$\begin{aligned}
R_t\mathbf{\Omega}\hat{\mathcal{W}} &= \{w \in \mathbf{\Omega}, w \in \hat{\mathcal{W}} : \Phi \circ U_t \circ \hat{\mathbf{0}}(w) = R\} \\
&= \{w \in \mathbf{\Omega}, w \in \hat{\mathcal{W}} : \Phi_0 \circ \phi \circ U_t \circ \hat{\mathbf{0}}(w) = R\} \\
&= \{w \in \mathbf{\Omega}, w \in \hat{\mathcal{W}} : \Phi_0 \circ U_t' \circ \phi \circ \hat{\mathbf{0}}(w) = R\} \\
&= (\Phi_0 \circ U_t' \circ \phi \circ \hat{\mathbf{0}})^{-1}\{R\} \cap \mathbf{\Omega} \\
&= \hat{\mathbf{0}}^{-1}((\Phi_0 \circ U_t' \circ \phi)^{-1}\{R\}) \cap \mathbf{\Omega} \\
&= \mathbf{0}^{-1}((\Phi_0 \circ U_t' \circ \phi)^{-1}\{R\}) \cap \tilde{\mathcal{W}} \cap \mathbf{\Omega}
\end{aligned}$$

where the second equality follows from Lemma 2 (and the first property of a factor[14]), and the third equality from Lemma 1 with the third property of a factor invoked t times. Next we turn to consider P_Ω. Recall Equation (2.2), that $P_\Omega(\mathbf{\Omega}) = 1$. We also have $P_\Omega(\hat{\mathcal{W}}) = P_\Omega \circ \mathbf{0}^{-1}(\hat{S}) = \mu(\hat{S}) = 1$ by the first property of a

[14]The assumption that $U[\hat{S}] \subseteq \hat{S}$, and hence $U_t[\hat{S}] \subseteq \hat{S}$, ensures that $\phi \circ U_t \circ \hat{\mathbf{0}}(w)$ is well-defined.

factor and Equation (2.1). Thus $P_\Omega(\mathbf{\Omega}\hat{\mathcal{W}}) = 1$ and so $P_\Omega(R_t) = P_\Omega(R_t\mathbf{\Omega}\hat{\mathcal{W}})$. We can then calculate:

$$\begin{aligned}
P_\Omega(R_t) &= P_\Omega(\mathbf{0}^{-1}((\Phi_0 \circ U_t' \circ \phi)^{-1}\{R\})) \\
&= \mu((\Phi_0 \circ U_t' \circ \phi)^{-1}\{R\}) \\
&= \mu \circ \phi^{-1}((\Phi_0 \circ U_t')^{-1}\{R\}) \\
&= \nu((\Phi_0 \circ U_t')^{-1}\{R\}) \\
&= \nu(\Phi_0 \circ U_t'(..., a_{-1}, a_0, a_1, ...) = R) \\
&= \nu(\Phi_0(..., a_{t-1}, a_t, a_{t+1}, ...) = R) \\
&= \nu(a_t = 1) = \tfrac{1}{2},
\end{aligned}$$

where the first equality follows from the results derived just above and the fact $P_\Omega(\tilde{\mathcal{W}}\mathbf{\Omega}) = 1$, the second equality follows from Equation (2.1), the fourth equality from Lemma 1 and the second property of a factor, and the subsequent equalities from the definition of $(V, \mathcal{V}, \nu, U')$ (see Lemma 1). Since this calculation was general to any t, we can conclude that $P_\Omega(R_0) = P_\Omega(R_1) = P_\Omega(R_2) = ... = \frac{1}{2}$.

Are $R_t, R_{t+1}, ...$ independent according to P_Ω? A similar calculation yields:

$$\begin{aligned}
P_\Omega(R_tR_{t+1}) &= P_\Omega(\mathbf{0}^{-1}((\Phi_0 \circ U_t' \circ \phi)^{-1}\{R\}) \cap \mathbf{0}^{-1}((\Phi_0 \circ U_{t+1}' \circ \phi)^{-1}\{R\})) \\
&= P_\Omega \circ \mathbf{0}^{-1}\left[(\Phi_0 \circ U_t' \circ \phi)^{-1}\{R\} \cap (\Phi_0 \circ U_{t+1}' \circ \phi)^{-1}\{R\})\right] \\
&= \mu\left[\phi^{-1}((\Phi_0 \circ U_t')^{-1}\{R\}) \cap \phi^{-1}((\Phi_0 \circ U_{t+1}')^{-1}\{R\}))\right] \\
&= \mu \circ \phi^{-1}\left((\Phi_0 \circ U_t')^{-1}\{R\} \cap (\Phi_0 \circ U_{t+1}')^{-1}\{R\})\right) \\
&= \nu((\Phi_0 \circ U_t')^{-1}\{R\} \cap (\Phi_0 \circ U_{t+1}')^{-1}\{R\})) \\
&= \nu(a_t = 1,\, a_{t+1} = 1) \\
&= \nu(a_t = 1) \cdot \nu(a_{t+1} = 1) \\
&= P_\Omega(R_t) \cdot P_\Omega(R_{t+1})
\end{aligned}$$

as desired, where the third equality follows from Equation (2.1), the fifth equality from Lemma 1 and the second property of a factor, the sixth equality by definition of U_t' and Φ_0, the second-to-last equality (which is the crucial step) from the fact that ν is the Bernoulli distribution that treats the variables a_i as independent, and the final equality from the previous result. This reasoning can be extended to show more generally that any collection of R_t with t distinct are independent according to P_Ω.

Finally, we use these results to draw our desired conclusion about the emergent dynamical chances. Let

w be the world fixed above. Note that $\tilde{H}_t(w)$ entails Ω. So, by the global principal principle, in particular by alignment (Chapter 1.7), we have

$$\tilde{P}_{tw}(\cdot) = P_\Omega(\cdot|\tilde{H}_t(w))$$

to the extent that the domains of these two functions overlap. Note that $\tilde{H}_t(w)$ will be a proposition of the form $R_0^{(c)}R_1^{(c)}...R_{t-1}^{(c)}R_t^{(c)}$ where c in parentheses indicates it may be the complement (negation). For example, since $\tilde{H}(w)$ corresponds to the sequence $(L, R, L, ...)$, we have

$$\tilde{P}_{0w}(\cdot) = P_\Omega(\cdot|L_0), \qquad \tilde{P}_{1w}(\cdot) = P_\Omega(\cdot|L_0R_1), \qquad \tilde{P}_{2w}(\cdot) = P_\Omega(\cdot|L_0R_1L_2), \qquad ....$$

It is now clear that the fact is true as claimed. Note for example that $P_\Omega(R_1|L_0) = P_\Omega(R_1) = \frac{1}{2}$ and $P_\Omega(R_1R_2|L_0) = P_\Omega(R_1) \cdot P_\Omega(R_2)$, so it follows that $\tilde{P}_{0w}(R_1) = \frac{1}{2}$ and $\tilde{P}_{0w}(R_1R_2) = \tilde{P}_{0w}(R_1) \cdot \tilde{P}_{0w}(R_2)$. Similarly, $P_\Omega(R_2|L_0R_1) = P_\Omega(R_1) = \frac{1}{2}$ and $P_\Omega(R_2R_3|L_0R_1) = P_\Omega(R_2) \cdot P_\Omega(R_3)$, so it follows that $\tilde{P}_{1w}(R_2) = \frac{1}{2}$ and $\tilde{P}_{1w}(R_2R_3) = \tilde{P}_{1w}(R_2) \cdot \tilde{P}_{1w}(R_3)$, and so on.

Similar results also hold for other cases, like the wheel of fortune. In Chapter 1, Example 15, we derived the result that, letting A_{t+1} denote the proposition that the next spin of the wheel yields WHITE:

$$\tilde{P}_{tw}(A_{t+1}) = \int_{w' \in \tilde{H}_t(w)} X_{A_{t+1}}(w')\, d\tilde{P}_{tw}(w'),$$

where $X_{A_{t+1}} : \mathcal{W} \to \mathbb{R}$ is defined by $X_{A_{t+1}}(w) = P_{tw}(A_{t+1})$. I suggested that, given the assumption that the initial probability density over the possible speeds is sufficiently smooth, method-of-arbitrary-function techniques will imply $\tilde{P}_{tw}(A_{t+1}) \approx 50\%$ (Von Plato, 1983; Strevens, 1998, 2003, 2011, 2013; Gallow, 2019). Since this holds for all t, we again get uniformity of chance across trials. Strevens (2003, Chapter 3) also sketches how, if the *joint density* over the initial conditions of all the future spins is sufficiently smooth, then $\tilde{P}_{tw}$ will treat $A_{t+1}, A_{t+2}, ...$ as approximately independent. So we get all the ingredients for uniform chance structure.

While all of these are toy examples, I submit that a similar kind of phenomenon is occurs with many probabilities in the special sciences. For example, consider the *law of independent assortment* in genetics and the related example that I began with: the probability that a cross between two pea plants will yield an offspring that is heterozygous for both seed shape and seed color is approximately $1/2 \cdot 1/2 = 1/4$. The physical basis for inheritance in these cases is (roughly speaking) the orientation of chromosomes along the metaphase plate during meiosis. The events of the seed shape inheritance and seed color may not have uniform chance and may not be independent with respect to the fundamental dynamical chances, because, given the exact microhistory up the relevant time t, conditioning on seed color may affect the proportion of

extensions that microhistory with certain seed shapes (compare how, in the indeterministic roulette wheel example, WHITE_1 and WHITE_2 are not independent with respect to the fundamental dynamical chances). However, they are plausibly independent with respect to the M-chances for appropriate M that coarse-grains over the micro-conditions that lead to specific chromosome orientations.

Other examples of emergence of uniform chance structure may not be so well suited for modeling with method-of-arbitrary-function techniques. Consider, say, the lung cancer case from Chapter 1.3, Example 9. It may just be true that according to the true chance theory, the chances $P_{M(w)}$ given a specification of the incidence of lung cancer in a population, treats the outcomes of different individuals getting lung cancer as independent and having the same chance, and that there is no natural 'measure over initial conditions' from which these results can be derived (though if there does exist such a measure, then the M-chances will have to be related to it by alignment). My framework leaves open this possibility, and is not tied to any specific mathematical or physical method for modeling these probabilities.

Inferring chances from frequencies, revisited

Once uniform chance structure is in place, chances can be learned though straightforward enumerative induction. The following result illustrates a precise sense in which this can happen. This result implies Lewis's (1980, p. 287) result about the coin tosses, discussed in Section 2.1, as a corollary.

Fact 2. *Assume the global principal principle. Consider propositions $A_1, A_2, ...$ and suppose the agent's total evidence E is a finite disjunction of pairs $M(w)T \in M\mathbf{T}$ (where here the world w is a free variable) such that, for each pair, these propositions are M-independent and M-identically distributed according to T at w. Fix $\epsilon > 0$. Then for any $\delta > 0$ there is an N such that, for all $n \geq N$,*

$$C\left(\langle P_M(A) = \bar{A}_n \pm \epsilon\rangle\right) \geq 1 - \delta,$$

where $\bar{A}_n$ is the relative frequency of the A_i up to trial n, and A is A_{n+1} or any other trial after n.

(A similar result also holds where the global principal principle is replaced by the local principal principle, and the global chance theory T is replaced by a local chance theory T_M.)

To prove Fact 2, we use the convergence-in-probability variant of the Law of Large Numbers, as applied to indicator variables:

Lemma 3 (Çınlar (2011), Thm III.6.2)**.** *Suppose that $A_1, A_2, ...$ are independent and identically distributed (with probability a) according to a probability measure P. Then for every $\epsilon > 0$, $\lim_{n\to\infty} P(|\bar{A}_n - a| < \epsilon) = 1$.*

Fix $\epsilon > 0$ and consider the pairs $M(w)T \subseteq E$. Applying Lemma 3 to each $P_{M(w)}$, we have that for any $\delta > 0$ there is an N such that, for each $P_{M(w)}$, the following holds for all $n \geq N$:[15]

$$P_{M(w)}\left(|\bar{A}_n - P_{M(w)}(A)| < \epsilon\right) \geq 1 - \delta.$$

So by the principal principle, for each pair $M(w)T \subseteq E$, we have $C_0\left(|\bar{A}_n - P_{M(w)}(A)| < \epsilon \,|M(w)T\right) \geq 1-\delta$. Since $M(w)T$ entails $\langle P_M(A) = P_{M(w)}(A)\rangle$ which in conjunction with $\{w : |\bar{A}_n(w) - P_{M(w)}(A)| \leq \epsilon\}$ entails $\langle P_M(A) = \bar{A}_n \pm \epsilon\rangle$, it follows that for each pair $M(w)T \subseteq E$, $C_0\left(\langle P_M(A) = \bar{A}_n \pm \epsilon\rangle | M(w)T\right) \geq 1 - \delta$. Finally, letting D stand for the proposition $\langle P_M(A) = \bar{A}_n \pm \epsilon\rangle$, we have $C(D) = C_0(D|E)$ and:

$$\begin{aligned}
C_0(D|E) &= \sum_{M(w)T \subseteq E} C_0(D|FT) \cdot C_0(M(w)T|E) \\
&\geq \sum_{M(w)T \subseteq E} (1-\delta) \cdot C_0(M(w)T|E) \\
&\geq (1-\delta) \cdot \sum_{M(w)T \subseteq E} C_0(M(w)T|E) \\
&\geq 1 - \delta,
\end{aligned}$$

as desired, where as usual we invoke probabilism and conditionalization.

Other interesting structure at higher levels

I have been focused on the emergence of uniform chance structure, where the coarse-grained chances treat the propositions or events in question as independent and identically distributed (i.i.d.). But there are many probability models in the special sciences which do not involve i.i.d. distributions but some more complicated stochastic process. It is thus worth noting how these more complex chance structures can also emerge.

To find a concrete case, we need not look further than the baker's map example from earlier in this section. Following Werndl (2012, 2013), rather than considering a coarse-graining Φ that distinguishes only between the left and right sides of the square, we can consider a less extreme coarse-graining Φ_{16} that distinguishes between the 16 square-shaped regions pictured in Figure 2.6. So instead of $\tilde{S} = \{L, R\}$, we have $\tilde{S} = \{e_{i,j} : 1 \leq i, j \leq 4\}$. Taking the same world $w \in \mathbf{\Omega}$ with initial state $s_0 = (\frac{3}{10}, \frac{1}{2})$, the coarse-grained history $\tilde{H}(w)$ will correspond to the sequence $(e_{2,3}, e_{3,2}, e_{1,3}, e_{2,2}, ...)$.

Using analogous reasoning as above, one can show that $\tilde{P}_{tw}$ is no longer a Bernoulli fair coin process, but rather a more complicated Markov process according to which a current state can be followed by one of

[15]This step relies on the premise that the number of pairs $M(w)T \subseteq E$ is finite; to get a working result for infinitely many pairs, we would need to posit more structural conditions on the agent's hypothesis space. I set aside these complications here.

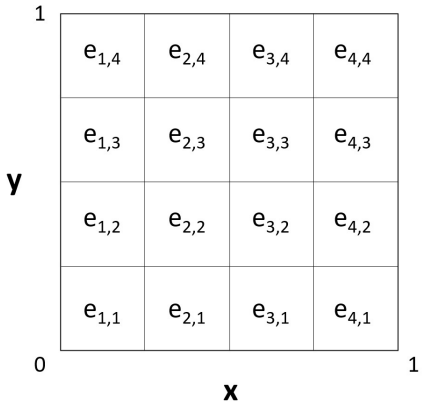


Figure 2.6: An intermediate coarse-graining of state space; based on Werndl (2013, Figure 2(b)).

two particular states each with probability 1/2, where the identity of those two particular states vary as a function of the current state (for example, $e_{2,3}$ can be followed by $e_{3,2}$ or $e_{4,2}$, and $e_{3,2}$ can be followed by $e_{1,3}$ or $e_{2,3}$); for illustration, see Figure 2.7.

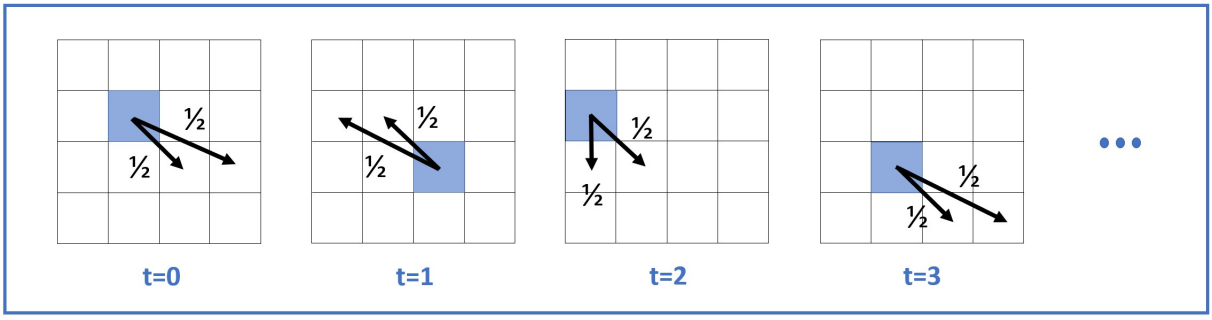


Figure 2.7: The history $\tilde{H}(w)$ and emergent dynamical chances at the level of grain corresponding to Φ_{16}.

To show this in more detail, we begin by observing that:

Lemma 4 (Werndl (2012), footnote 4, slight correction)**.** *Let ϕ be the map from Lemma 1 and let Φ_{16} be the coarse-graining function above. Then for all $s \in \hat{S}$, $\Phi_{16}(s) = \Xi_{16} \circ \phi(s)$, where $\Xi_{16} : V \to \tilde{S}$ is the function that, given an input string $(..., a_{-1}, a_0, a_1, ...)$, yields $e_{1,1}$ if $a_{-2}a_{-1}a_0a_1 = 0000$, $e_{1,2}$ if $a_{-2}a_{-1}a_0a_1 = 0001$, and so on for the remaining fourteen coarse-grained states, in the manner shown by Figure 2.8.*

Let $(e_{i,j})_t$ denote the proposition that the world is in state $e_{i,j}$ at time t, analogous to the proposition R_t that the world is in state R at time t. Similar calculations to those above show $P_\Omega((e_{i,j})_t) = \frac{1}{16}$ for all

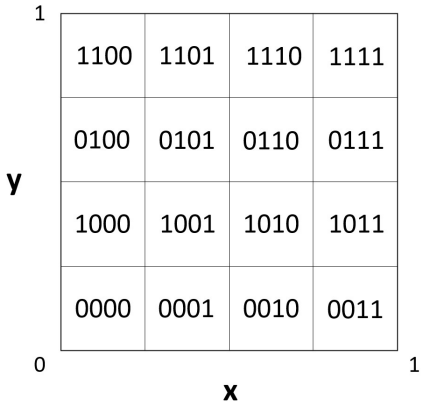


Figure 2.8: Illustration of the map Ξ_{16} from Lemma 4.

$1 \leq i, j \leq 16$ and $t \in \mathbb{T}$. For instance:

$$P_\Omega((e_{1,3})_t) = \nu((\Xi_{16} \circ U'_t)^{-1}\{e_{1,3}\}) = \nu(a_{t-2}a_{t-1}a_t a_{t+1} = 0100) = \left(\tfrac{1}{2}\right)^4 = \tfrac{1}{16}.$$

But unlike before, the trials are **not** independent. To illustrate this, note for instance that:

$$\begin{aligned} P_\Omega((e_{1,3})_t \,|\, (e_{3,2})_{t-1}) &= \nu((\Xi_{16} \circ U'_t)^{-1}\{e_{1,3}\} \,|\, (\Xi_{16} \circ U'_{t-1})^{-1}\{e_{3,2}\}) \\ &= \nu(a_{t-2}a_{t-1}a_t a_{t+1} = 0100 \,|\, a_{t-3}a_{t-2}a_{t-1}a_t = 1010) \\ &= \tfrac{1}{2} > P_\Omega((e_{1,3})_t), \end{aligned}$$

since conditional on $a_{t-2}a_{t-1}a_t$ being 010, the probability that $a_{t-2}a_{t-1}a_t a_{t+1}$ is 0100 is the probability of a_{t+1} being 0, which is $\frac{1}{2}$. However, while there is no independence from the previous trial, there is still a Markov property, in the sense that the probability for a certain state at $t+1$ is fixed entirely by the current state t. For example:

$$\begin{aligned} P_\Omega((e_{2,2})_{t+1} \,|\, (e_{1,3})_t (e_{3,2})_{t-1} ... (e_{2,3})_0) &= \nu(a_{t-1}a_t a_{t+1}a_{t+2} = 1001 \,|\, a_0 a_1 ... a_{t-2}a_{t-1}a_t a_{t+1} = 01...0100) \\ &= \nu(a_{t-1}a_t a_{t+1}a_{t+2} = 1001 \,|\, a_{t-2}a_{t-1}a_t a_{t+1} = 0100) \\ &= P_\Omega((e_{2,2})_{t+1} \,|\, (e_{1,3})_t), \end{aligned}$$

where the second equality follows from the fact that according to ν, the a_t are independent for distinct t.

This implies, again applying the global principal principle, that for all i, j,

$$\tilde{P}_{tw}((e_{i,j})_{t+1}) = P_\Omega((e_{i,j})_{t+1} \,|\, (e_{k,l})_t)$$

where $e_{k,l} \in \tilde{S}$ is the state at w at time t. For example, $\tilde{P}_{2w}((e_{2,2})_3) = P_\Omega((e_{2,2})_3 \,|\, (e_{1,3})_2) = \frac{1}{2}$. All in all, we have the following result:

Fact 3. *Fix this set-up with Φ_{16} and $\tilde{S} = \{e_{i,j} : 1 \leq i, j \leq 4\}$. Then for a given t and i, j, it may not be the case that $(e_{i,j})_1, (e_{i,j})_2, ...$ are $\tilde{H}_t$-independent or identically distributed at w. Instead we have a more complicated Markov process with $\tilde{P}_{tw}((e_{i,j})_{t+1})$ equaling either $\frac{1}{2}$ or 0 depending on i, j and t.*[16] *Furthermore, relative to this Φ_{16} coarse-graining, the propositions $R_t, R_{t+1}, R_{t+2}, ...$ do not have uniform chance across trials. In fact, relative to these $\tilde{H}_t$-chances, whether R_{t+1} is true is a deterministic matter:*[17]

$$\tilde{P}_{0w}(R_1) = 1, \qquad \tilde{P}_{1w}(R_2) = 0, \qquad \tilde{P}_{2w}(R_3) = 0, \qquad \tilde{P}_{3w}(R_4) = 1, \qquad ....$$

Deletion of structure

In addition to interesting chance structure emerging at higher levels, it is possible that there is interesting chance structure at lower levels that then gets 'washed out' or 'deleted' at higher levels. One example is when an underlying indeterministic process is deterministic relative to the emergent dynamical chances.

Example 17. Consider the scenario pictured in Figure 2.9 (Butterfield, 2012, Section 3.3). There are two possible initial microstates s, s'. The fundamental dynamical laws U are indeterministic, and in particular while s has a 100% chance of evolving to $U(s)$, s' has a 50% of evolving to $U(s')_1$ and a 50% chance evolving to $U(s')_2$. Meanwhile, the coarse-grained state $\tilde{s}_1$ that is compatible with s is also the one compatible with s', and the coarse-grained state $\tilde{s}_2$ that is compatible with $U(s)$ is also the one compatible with $U(s')_1$ and $U(s')_2$. If this pattern continues, then assuming the emergent dynamical chance function $\tilde{P}_{0w}$ only assigns chances to coarse-grained histories, it will be completely deterministic (take only values 0 and 1).

[16] In particular, letting $\tilde{s}_t$ denote the coarse-grained state of w at t (in terms of the extraction function introduced in Chapter 1.6, $\tilde{s}_t = \tilde{f}_t(H(w))$), we have (cf. Werndl (2012, p. 2248)): if $\tilde{s}_t = e_{1,1}$ then $(e_{1,1})_{t+1}$ and $(e_{2,1})_{t+1}$ are the two possibilities, each with probability $\frac{1}{2}$ and the remaining $(e_{i,j})_{t+1}$ with probability 0, if $\tilde{s}_t = e_{2,1}$ then the two possibilities are $(e_{3,1})_{t+1}$ and $(e_{4,1})_{t+1}$, if $\tilde{s}_t = e_{3,1}$ then the two possibilities are $(e_{1,3})_{t+1}$ and $(e_{2,3})_{t+1}$, if $\tilde{s}_t = e_{4,1}$ then the two possibilities are $(e_{3,3})_{t+1}$ and $(e_{4,3})_{t+1}$, if $\tilde{s}_t = e_{1,2}$ then the two possibilities are $(e_{1,1})_{t+1}$ and $(e_{2,1})_{t+1}$, if $\tilde{s}_t = e_{2,2}$ then the two possibilities are $(e_{3,1})_{t+1}$ and $(e_{4,1})_{t+1}$, if $\tilde{s}_t = e_{3,2}$ then the two possibilities are $(e_{1,3})_{t+1}$ and $(e_{2,3})_{t+1}$, if $\tilde{s}_t = e_{4,2}$ then the two possibilities are $(e_{3,3})_{t+1}$ and $(e_{4,3})_{t+1}$, if $\tilde{s}_t = e_{1,3}$ then the two possibilities are $(e_{1,2})_{t+1}$ and $(e_{2,2})_{t+1}$, if $\tilde{s}_t = e_{2,3}$ then the two possibilities are $(e_{3,2})_{t+1}$ and $(e_{4,2})_{t+1}$, if $\tilde{s}_t = e_{3,3}$ then the two possibilities are $(e_{1,4})_{t+1}$ and $(e_{2,4})_{t+1}$, if $\tilde{s}_t = e_{4,3}$ then the two possibilities are $(e_{3,4})_{t+1}$ and $(e_{4,4})_{t+1}$, if $\tilde{s}_t = e_{1,4}$ then the two possibilities are $(e_{1,2})_{t+1}$ and $(e_{2,2})_{t+1}$, if $\tilde{s}_t = e_{2,4}$ then the two possibilities are $(e_{3,2})_{t+1}$ and $(e_{4,2})_{t+1}$, if $\tilde{s}_t = e_{3,4}$ then the two possibilities are $(e_{1,4})_{t+1}$ and $(e_{2,4})_{t+1}$, and if $\tilde{s}_t = e_{4,4}$ then the two possibilities are $(e_{3,4})_{t+1}$ and $(e_{4,4})_{t+1}$.

[17] To see this, note that the $e_{i,j}$ that get assigned a probability $\frac{1}{2}$ are always either both in the left-hand or both in the right-hand side of the square (see previous footnote).

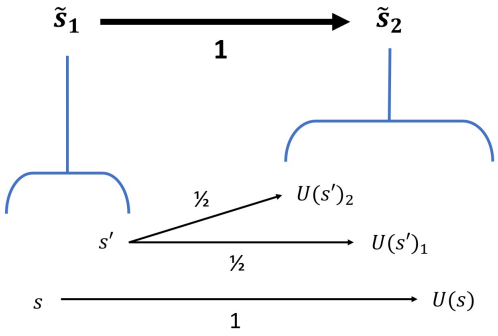


Figure 2.9: Determinism emerging from indeterministic dynamics; cf. Butterfield (2012, Figure 3).

In this example, the coarse-grained histories have trivial probability according to both the fundamental and coarse-grained chances, so in one sense no uniform chance structure is 'lost', because there was no uniform chance structure with respect to those histories at the lower level. However, there are also cases where there is uniform chance structure with respect to certain propositions at a lower level, but not with respect to those same propositions at a higher level: (Below I assume the global principal principle.)

Example 18. A biologist is interested in whether the offspring of a certain organism will inherit a disease. There are three independent considerations: whether the offspring inherits trait α, whether it inherits trait β, and whether the parent lives at high altitude. Focusing on a particular parent and its first and second offspring, let $A_{1\alpha}$ be the proposition that the first inherits trait α, $A_{2\alpha}$ that the second does, and similarly for $A_{1\beta}$ and $A_{2\beta}$. Let B be the proposition that the parent lives at high altitude. We suppose that D_1 the first offspring inheriting the disease is equivalent to $A_{1\alpha}B \vee A_{1\beta}B^c$ and D_2 is equivalent to $A_{2\alpha}B \vee A_{2\beta}B^c$. Let F specify the relevant background facts of the situation. Suppose, just to take an example, that $P_F(B) = \frac{2}{3}$, $P_F(A_{1\alpha}) = P_F(A_{2\alpha}) = \frac{1}{2}$, $P_F(A_{1\beta}) = \frac{1}{4}$ and $P_F(A_{2\beta}) = \frac{7}{9}$.[18] Next let G and G' be fine-grainings of F which specify, in addition, the altitude of the parent. By alignment, $P_G(\cdot) = P_F(\cdot|B)$ and $P_{G'}(\cdot) = P_F(\cdot|B^c)$. Then one can check that according to the finer-grained probabilities, D_1 and D_2 are either identically distributed and independent, or non-identically distributed and independent, depending on whether one is considering G or G'.[19] However, relative to the coarse-grained probabilities, D_1 and D_2 are neither identically distributed nor independent.[20] In other words, uniform chance structure that is present with respect to D_1 and D_2 at

[18] We suppose, for instance, that the second offspring is usually more likely to inherit trait β, but the likelihood of inheriting trait α does not vary by which is first-born or second-born.

[19] In particular, we have $P_G(D_1) = P_G(D_2) = \frac{1}{2}, P_G(D_1D_2) = P_G(D_1) \cdot P_G(D_2)$, and $P_{G'}(D_1) = \frac{1}{4}, P_{G'}(D_2) = \frac{7}{9}, P_{G'}(D_1D_2) = P_{G'}(D_1) \cdot P_{G'}(D_2)$. Note that for instance $P_G(D_1) = P_F(D_1|B) = P_F(A_{1\alpha}|B) = P_F(A_{1\alpha})$ since by supposition $A_{1\alpha}$ is independent of B (relative to P_F), and $P_{G'}(D_1D_2) = P_F(D_1D_2|B^c) = P_F(A_{1\beta}A_{2\beta}|B^c) = P_F(A_{1\beta}|B^c) \cdot P_F(A_{2\beta}|B^c) = P_G(D_1) \cdot P_G(D_1)$ since by supposition $A_{1\beta}$, $A_{2\beta}$ and B^c are independent (relative to P_F).

[20] We have $P_F(D_1) = \frac{5}{12}, P_F(D_2) = \frac{16}{27}, P_F(D_1D_2) = \frac{25}{108} \neq P_F(D_1) \cdot P_F(D_2)$. Here we use the fact for instance that $P_F(D_1D_2) = P_F(D_1D_2|B) \cdot P_F(B) + P_F(D_1D_2|B^c) \cdot P_F(B^c) = P_F(A_{1\alpha}A_{2\alpha}|B) \cdot \frac{2}{3} + P_F(A_{1\beta}A_{2\beta}|B^c) \cdot \frac{1}{3} = (\frac{1}{2})^2 \cdot \frac{2}{3} + (\frac{1}{4} \cdot \frac{7}{9}) \cdot \frac{1}{3}$.

the lower levels G and G' is lost at the higher level F.

2.3 Resiliency and the value of chance knowledge

Why are these coarse-grained chances worthy of epistemic investigation? Arguably, the emergent chance structures discussed in the previous section are important patterns—patterns that we would miss out on if we focused only on fundamental dynamical chances, and that are worth investigating for their own sake. But is there a more general story for why knowledge of non-fundamental chances is epistemically useful?

We have already seen a central part of this story, in the principal principle. Chance provides a guide to belief whether it is part of the fundamental dynamical laws or not. If I have learned that the M-chance of A is such-and-such, then assuming I don't have any M-inadmissible evidence, that tells me how to set my credence in A. Furthermore, if I know that I am in a situation where I'm not going to gain any M-inadmissible evidence, then I know my credence in A won't shift away from x. In that sense, the chance estimate provided to me is *resilient* to new information. It is precisely this kind of robust estimate that statisticians and scientists seek.

What if I have not learned the M-chance of A is x in the Bayesian sense, which would require a credence of 1 in $\langle P_M(A) = x\rangle$, but rather have high credence in it, say, on the basis of some finite number of trials? It's still the case that this chance will guide my credence in A, since by probabilism and conditionalization,

$$C(A) = \overbrace{C(A|\langle P_M(A) = x\rangle)}^{=x} \cdot \overbrace{C(\langle P_M(A) = x\rangle)}^{\approx 1} + C(A|\langle P_M(A) \neq x\rangle) \cdot \overbrace{C(\langle P_M(A) \neq x\rangle)}^{\approx 0} \approx x.$$

However, we would also like to spell out the sense in which (like above) this chance estimate is resilient to new information I might learn. Now matters are more complicated, because unlike the case where my credence distribution over the M-chance hypotheses is perfectly peaked at $\langle P_M(A) = x\rangle$, learning more M-admissible information might change that distribution. In this section I develop the notion of *resiliency* (Skyrms, 1977) to take into account these complications. This will end up also allowing us to see a further sense in which special science chances can be especially useful.

Introducing resiliency

The notion of the *resiliency* of a probability claim $\langle P(A) = x\rangle$ was introduced by Skyrms (1977, p. 705). Here is a rough gloss:

> Resiliency is a stability property, akin to the concept of robustness in statistics. A resilient probability is one that is relatively insensitive to perturbations in our belief structure.

For example, if A is the proposition that the coin will land heads on trial n, then the resiliency of $\langle P(A) = 1/2\rangle$ will be much higher if you are confident the coin is fair than if you only have a credence of $1/2$ of A because of symmetry of ignorance; in the former case, observing a large proportion of tails in the trials before n won't shift your credence in A away from $1/2$ very much, but in the latter case, it would lower your credence in A substantially below $1/2$.

More precisely, if C is a rational credence function, A is a proposition, and $\mathcal{B}$ is a (finite) algebra of propositions, then the *resiliency of* $\langle P(A) = x\rangle$ *in* C*, relative to* $\mathcal{B}$*,* is a measure of how much learning information in $\mathcal{B}$ could shift one's credence in A away from x. In a slight departure from Skyrms (1977), I will define this measure as the agent's *expectation* of the divergence of her posterior credence in A from x, upon learning information in $\mathcal{B}$.[21] In particular, letting $\mathbb{E}$ denote expectation with respect to C, and $\mathcal{E} \subseteq \mathcal{B}$ any partition (interpreted as a learning scenario, in which one only learns information in $\mathcal{B}$), we can consider the following measure which ranges from 0 (minimum resiliency) to 1 (maximum resiliency):

$$1 - \max_{\mathcal{E}:\mathcal{E}\subseteq\mathcal{B}} \mathbb{E}\left[\left(C(A|\mathcal{E}(\cdot)) - x\right)^2\right]. \tag{2.3}$$

Writing $X_{\mathcal{E}}(w)$ for $C(A|\mathcal{E}(w))$, this can be written as $1 - \max_{\mathcal{E}:\mathcal{E}\subseteq\mathcal{B}} \mathrm{Var}(X_{\mathcal{E}})$ where Var denotes variance.[22]

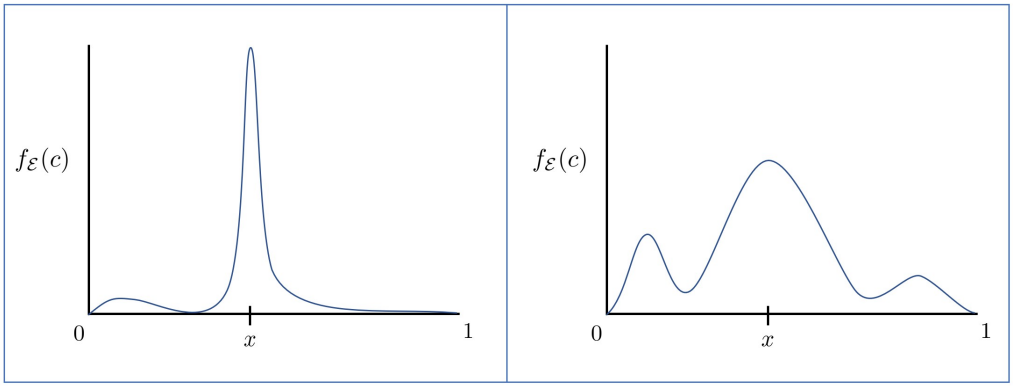


Figure 2.10: A situation of high resiliency (left) and low resiliency (right). Here $f_{\mathcal{E}}$ is the probability density for $X_{\mathcal{E}}$ (if $\mathcal{E} \subseteq \mathcal{B}$ is finite this will really be a discrete distribution, but I have smoothed it out for diagrammatic purposes). The shape on the left reflects high resiliency because of the narrowness and sharpness of the peak, whereas on the right there is a high variance in the agent's posterior credence in A.

Roughly speaking, this says that to calculate the resiliency of $\langle P(A) = x\rangle$ in C, relative to $\mathcal{B}$, we take all

[21] Skyrms defines resiliency as $1 - \max_{B\in\mathcal{B}} |Cr(A|B) = x|$. The problem with Skyrms's definition is that it makes resiliency too contingent on extreme 'outlier' scenarios. For example, looking at Figure 2.10, both the left and right sides allow the possibility that the agent's posterior credence in A might be very near 0, and hence far below x (note the small bump on the far left side of the left-hand graph). As a result, Skyrms's definition will treat both as cases of poor resiliency, even though intuitively the left side reflects much higher resiliency. By defining resiliency in terms of expectation, we can account for the relevant difference: in the left hand case, the outlier scenarios are extremely unlikely (by the agent's lights), whereas on the right hand side, they are not.

[22] In rewriting it this way, I assume, for simplicity, that $C(A) = x$ and so given reflection $\mathbb{E}[X_{\mathcal{E}}] = \mathbb{E}[C(A|\mathcal{E}(\cdot))] = C(A) = x$.

the learning scenarios $\mathcal{E} : \mathcal{E} \subseteq \mathcal{B}$ in which one only learns information in $\mathcal{B}$, calculate the maximum of the variance between one's posterior credence in A and the value x recommended by the probability claim, and subtract this number from 1. The case of maximum resiliency is when there is no variance:

$$\max_{\mathcal{E}:\mathcal{E}\subseteq\mathcal{B}} \mathbb{E}\left[(C(A|\mathcal{E}(\cdot)) - x)^2\right] = 0,$$

which implies that for all $B \in \mathcal{B}$ with positive credence, $C(A|B) = x$. So if resiliency is maximal then, with credence one, learning information in $\mathcal{B}$ will not shift one's credence in A by even a small amount. This kind of maximal resiliency is highly desirable: it means the agent can fully trust the probability estimate $\langle P(A) = x\rangle$ to continue being her best guide, even if she were to learn potentially new information in $\mathcal{B}$.

A notable example of maximal resiliency is as follows. Suppose C is a credence function which has learned the M-chance of A is x: $C = C_0(\cdot|\langle P_M(A) = x\rangle)$. Then the resiliency of $\langle P(A) = x\rangle$ in C, relative to the collection $\mathcal{B} = \mathcal{A}(M\mathbf{T})$ of propositions that are admissible with respect to M, is maximal. That is because, by the global principal principle, for any partition $\mathcal{E} \subseteq \mathcal{A}(M\mathbf{T})$, $C(A|\mathcal{E}(\cdot)) = C_0(A|\mathcal{E}(\cdot)\langle P_M(A) = x\rangle) = x$ where we use the fact that for all w, $\mathcal{E}(w)$ is admissible with respect to M. In fact, assuming that all the $FT \in M\mathbf{T}$ consistent with $\langle P_M(A) = x\rangle$ have positive credence, a converse claim holds, and we can reformulate the principal principle in terms of resiliency. Define the *M-resiliency of $\langle P(A) = x\rangle$ in C* as the resiliency of $\langle P(A) = x\rangle$ in C relative to the algebra of M-admissible propositions $\mathcal{A}(M\mathbf{T})$. Then we have:[23]

Global Principal Principle (Resiliency Formulation) Let $C = C_0(\,\cdot\,|\langle P_M(A) = x\rangle)$. Then the M-resiliency of $\langle P(A) = x\rangle$ in C is maximal.

Resiliency in cases of less than full certainty

A main purpose of introducing resiliency was to be able to consider cases where the agent's credence is peaked around $\langle P_M(A) = x\rangle$ but not perfectly so; and so we'd expect, in the terminology I just introduced, the M-resiliency of $\langle P(A) = x\rangle$ to be high but not necessarily maximal.

To elaborate, the idea is that if an agent's credence distribution over M-chance hypotheses is largely clumped around x, then $\langle P(A) = x\rangle$ will still be highly M-resilient, because the agent will expect only a

[23] I have already noted that maximal resilience implies that for all $B \in \mathcal{B}$ with positive credence, $C(A|B) = x$. (This follows from the fact that if $\mathrm{Var}(X_\mathcal{E}) = 0$ then $X = \mathbb{E}[X_\mathcal{E}]$ almost surely, together with the fact that for every positive credence $B \in \mathcal{B}$, there is a partition $\mathcal{E} \subseteq \mathcal{B}$ such that $B \in \mathcal{E}$.) So, under the assumption that all the $FT \in M\mathbf{T}$ consistent with $\langle P_M(A) = x\rangle$ have positive credence, this implies the global principal principle (admissibility formulation). We can also get an analogous result in terms of the local principal principle; rather than defining the M-resiliency of $\langle P(A) = x\rangle$ in C as the resiliency of $\langle P(A) = x\rangle$ in C relative to $\mathcal{A}(M\mathbf{T})$, we would define it as the the resiliency of $\langle P(A) = x\rangle$ in C relative to $\mathcal{A}(M\mathbf{T}_M)$.

small divergence between her posterior credence and x. In more detail, fix $\epsilon > 0$ and $\delta > 0$, and suppose

$$C(\langle P_M(A) = x \pm \epsilon \rangle) \geq 1 - \delta.$$

Then using the principal principle one can derive, for instance, the following lower bound on the M-resiliency of $\langle P(A) = x \rangle$ in C:[24]

$$1 - \max_{\mathcal{E}: \mathcal{E} \subseteq \mathcal{A}(M\mathbf{T})} \mathbb{E}\left[(C(A|\mathcal{E}(\cdot)) - x)^2\right] \geq 1 - 3\epsilon - 3\delta - 2\epsilon\delta.$$

Although the agent is not certain the probability estimate $\langle P(A) = x \rangle$ will continue to be her best guide, in expectation it is fairly certain that it will continue to be a very good guide in any learning scenario which involves learning M-admissible information.

The scope of resiliency

High resiliency relative to $\mathcal{B}$ is only significant to the extent that $\mathcal{B}$ includes all (or most) of the relevant information one may learn. For example, it is easy for a probability claim to achieve maximal resiliency with respect to the trivial set $\mathcal{B} = \{\emptyset, \mathcal{W}\}$ (if the current credence matches the probability), but this does not mean much if one is going to gain any non-trivial evidence that bears on the proposition A in question.

This is what Skyrms (1977) calls the issue of "scope" of resiliency. Knowledge of fundamental dynamical chances affords high resiliency with a very large scope: in this case, there is high resiliency relative to $\mathcal{B} = \mathcal{A}(H_t\mathbf{T})$; the estimate $\langle P(A) = x \rangle$ provided by the fundamental dynamical chances is resilient to any new evidence about the past or present, at any level of detail. (Though as Lewis (1980, p. 265) points out, "resiliency has its limits"; there is still some information that is H_t-inadmissible.) Knowledge of special science chances also affords high resiliency, but with a smaller scope: in general, if $M \preceq H_t$, then $\mathcal{A}(M\mathbf{T}) \subseteq \mathcal{A}(H_t\mathbf{T})$. Of course, this smaller scope does not mean the special science estimate is not useful or does not represent a lawlike or robust statistic; in many cases the only information one will learn that bears greatly on A is information that is M-admissible. Skyrms makes a similar point (1977, pp. 708–709):

> There is no unique answer as to which sublanguage [subalgebra $\mathcal{B}$] resiliency must be evaluated over, for lawlikeness. Rather, we must again say that *the larger* the sublanguage over which we

[24] Here for simplicity I assume that $C(A) = x$ (we can also derive a similar upper bound if we drop this assumption, but the derivation is more complicated). Fix any $\mathcal{E} \subseteq \mathcal{A}(MT)$ and letting $X_{\mathcal{E}}(w)$ denote $C(A|\mathcal{E}(w))$ note that $\mathbb{E}\left[(X_{\mathcal{E}} - x)^2\right] = \mathbb{E}[X_{\mathcal{E}}^2] - 2x\mathbb{E}[X_{\mathcal{E}}] + x^2$. By our assumption and reflection, $\mathbb{E}[X_{\mathcal{E}}] = \mathbb{E}[C(A|\mathcal{E}(\cdot))] = C(A) = x$. So it suffices to show $\mathbb{E}[X_{\mathcal{E}}^2] \leq x^2 + 3\epsilon + 3\delta + 2\epsilon\delta$. The strategy (I will omit the details here) is to note that, letting $D_\epsilon = \langle P_M(A) = x \pm \epsilon \rangle$ we have $X_{\mathcal{E}} = C(A|D_\epsilon\mathcal{E}(\cdot)) \cdot C(D_\epsilon|\mathcal{E}(\cdot)) + C(A|D_\epsilon^c\mathcal{E}(\cdot)) \cdot (1 - C(D_\epsilon|\mathcal{E}(\cdot))$, where $x - \epsilon \leq C(A|D_\epsilon\mathcal{E}(\cdot)) \leq x + \epsilon$ by the principal principle. Squaring the expression and using this bound, as well as the fact $\mathbb{E}[Y^2] \leq \mathbb{E}[Y]$ for $0 \leq Y \leq 1$, one can derive the desired result.

> have high instantial resiliency, the more *lawlike* the statistical law. [... Indeed] different choices of "scope" of resiliency correspond to different statistical properties which play an important role in the discussion of propensities. [...] The indeterminist is worried about the fact that, once we introduce descriptions of the microstate of the system into the language, the resiliency of the appropriate probability drops drastically. But those of us who are happy about talking about statistical laws and propensities with regard to classical statistical mechanics, are presumably impressed enough by resiliency over descriptions of the macrostate of the system. [...] But if indeterminism is *really* what the indeterminist wants, then it is clear that he should evaluate Resiliency of $\langle P(A) = x\rangle$ over a language that includes resources for a complete description of the history of the world up to the time of occurrence of A (but not after that time).

(The part about the resiliency "dropping drastically" refers to the fact that, for $M \preceq H_t$, we may have that $\langle P(A) = x\rangle$ is highly M-resilient, but not at all H_t-resilient.[25])

In some cases, special science chance estimates are highly resilient relative to a remarkably large scope. This often happens with chaotic systems whose behavior is very sensitive to perturbations in initial conditions or remote influences.[26] As an example, take again the case of the baker's map, and let M be the subject matter $\tilde{H}_t$ from the beginning of Section 2.2, corresponding to the coarse-graining $\Phi : S \to \{L, R\}$. One of the results of that section was that the M-chance of $A = R_{t+1}$ is $\frac{1}{2}$, and so the chance estimate $\langle P(A) = \frac{1}{2}\rangle$ is highly M-resilient. But in fact this chance estimate is also highly resilient relative to much more detailed subject matters. Let $N \succeq M$ be any subject matter that corresponds to a coarse graining Φ' that distinguishes, not just between the left and right halves of the square, but any finite number of vertical strips of the square, like for example those shown in Figure 2.11. Then one can check that $\langle P(A) = \frac{1}{2}\rangle$ will also be highly N-resilient.[27]

To illustrate the upshot, imagine that it is practically impossible for agents in this world to get information about the y-coordinate, but they have various methods for investigating the x-coordinate, with different degrees of precision. Then $\langle P(A) = \frac{1}{2}\rangle$ will be always be their best estimate unless they manage to find a

[25] In general, if $M \preceq N$, the N-resiliency of $\langle P(A) = x\rangle$ is less than or equal to the M-resiliency of $\langle P(A) = x\rangle$. That is because if $M \preceq N$ then $\mathcal{A}(M\mathbf{T}) \subseteq \mathcal{A}(N\mathbf{T})$, and so $\max_{\mathcal{E}:\mathcal{E}\subseteq\mathcal{A}(M\mathbf{T})}(...) \leq \max_{\mathcal{E}:\mathcal{E}\subseteq\mathcal{A}(N\mathbf{T})}(...)$.

[26] See Elga (2005) for more detailed discussion of the factor of remote influence.

[27] Here is a sketch of why this is true; see also similar observations by Werndl (2009, Section 4.2) and List and Pivato (2015, p. 139). Suppose $\tilde{S} = \{\tilde{s}_1, \tilde{s}_2, ..., \tilde{s}_n\}$ corresponds to a partition into n vertical strips, such that L and R can both be expressed as disjunctions of these strips. Then we will in general have $\Phi' = \Xi' \circ \phi$ where $\Xi'(..., a_{-1}, a_0, a_1, ...)$ depends only on $a_0, a_1, ..., a_m$ for a finite number m. (For example, in the case where S is split into four evenly sized vertical strips, we have that Ξ' outputs $\tilde{s}_1$ if $a_0a_1 = 00$, $\tilde{s}_2$ if $a_0a_1 = 01$, $\tilde{s}_3$ if $a_0a_1 = 10$, and $\tilde{s}_4$ if $a_0a_1 = 11$.) What is crucial here is that, unlike Ξ_{16} from Section 2.2, Ξ' does not depend on the 'earlier' elements a_{-i} of the Bernoulli process, since the way ϕ is defined, the x coordinate of $s \in S$ gets mapped entirely to the $0th$ and later elements. Like before we get $P_\Omega(\tilde{s}_t)$ equal to the measure on the square (in the example of Section 2.2 they were all regions of size 1/16) and so $P_\Omega(R_{t+1}) = \frac{1}{2}$. But this time, unlike with Ξ_{16}, there is no dependence on previous states, so for example $\tilde{P}_{0w}(R_1) = P_\Omega(R_1|L_0) = P_\Omega(R_1) = \frac{1}{2}$, and $\tilde{P}_{1w}(R_2) = P_\Omega(R_2|L_0R_1) = P_\Omega(R_2) = \frac{1}{2}$, and so on, where here $\tilde{P}_{tw}$ is understood as relative to the coarse-graining Φ' as opposed to Φ or Φ_{16}.

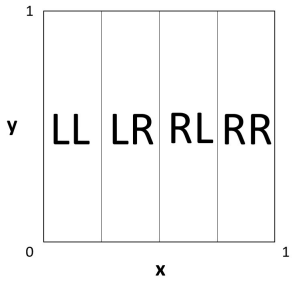


Figure 2.11: A partition of state space into four vertical strips.

way to measure the y-coordinate, or to measure the x-coordinate with infinite precision.

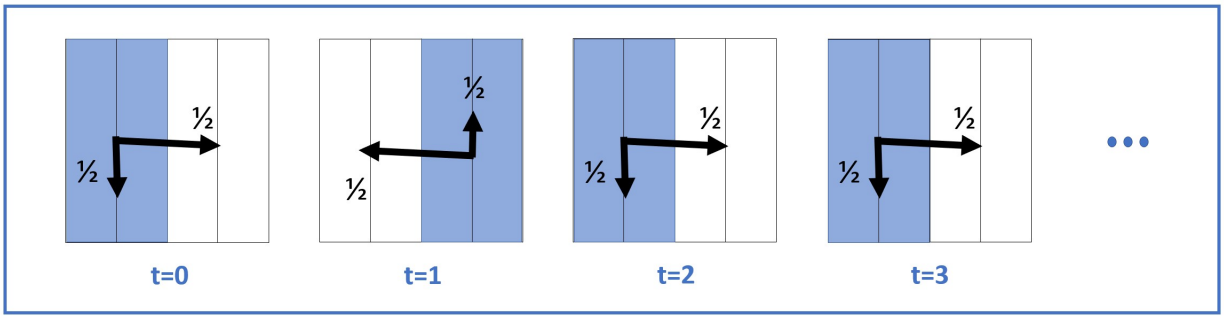


Figure 2.12: Chances of L and R with respect to the coarse-graining Φ' into four vertical strips.

Arguably, an analogous phenomenon is happening with many actual special science chance estimates. Given our inability to determine the microstate of the particles in the atmosphere, atmospheric scientists model the weather as a stochastic system. The chance estimates provided by these models could be beaten if we had more micro-information, but given obvious experimental and practical limitations, they are the best we can obtain. Assuming the principal principle, any rational agent subject to these limitations who knows the meteorological chances can treat them as a robust and stable estimate; in other words, these chance estimates have high resiliency in the agent's credences C, relative to any set of relevant information $\mathcal{B}$ that is practically feasible for her to obtain.

2.4 Conclusion

Many of our methods for learning chances rely on structural chance assumptions, that may be entirely absent at the level of the fundamental dynamical chances. Using my framework, I have sketched how uniform and other kind of interesting chance structure can nevertheless emerge at the level of non-Lewisian chances, and

how that structure grounds the (likely) success of these epistemic methods. So despite our ignorance of the fundamental physical theory, we do 'know the chances'. Finally, taking the principal principle as background, I used the principal principle and the notion of resiliency to explain in more detail why this knowledge is valuable, and the general epistemic importance of chance estimates in both fundamental and special sciences.

Part II

Quantum chance

Chapter 3

States of ignorance and ignorance of states

Examining the quantum principal principle

Perhaps the most notorious physical probabilities or chances are those found in quantum mechanics. The jury is still out on whether the universe is fundamentally indeterministic, and so on whether quantum probabilities are chances in Lewis's sense. But in either case, they are chances in my sense, and they provide such resilient estimates that they deserve special attention. Indeed, quantum probabilities pose some challenges to our preconceptions about the epistemology of chance. This chapter explores one such challenge.

A particle is approaching a Stern-Gerlach magnet oriented in the x direction. You know the particle is in the spin-up-z state, $\uparrow_z$. How confident should you be that it will be registered as having gone 'UP' through the magnet? Answer: you should be 50% confident. Here's one story for why. Knowing the particle is in the $\uparrow_z$ state tells you that the chance of 'UP' is $|\langle\uparrow_x|\uparrow_z\rangle|^2 = 1/2$. And epistemic rationality demands you conform your credences to known chances.

The last step in this reasoning is an application of the principal principle, familiar from Part I. On the standard way of understanding the principle, it is a normative requirement that goes above and beyond other more background constraints on rational credence, like probabilism and conditionalization. While authors have disagreed about whether the principle can be independently justified and if so how (Strevens, 1999; Ward, 2005; Pettigrew, 2012, 2013; Childers, 2012; Schwarz, 2014; Builes and Spencer, ms), almost all agree that an agent can be probabilistically coherent and yet still violate it. However, John Earman (2018) has

recently suggested that in the quantum context, the principle:

Quantum principal principle (QPP), rough version: if an agent learns that the state of the system is ψ, then she should set her credence in E to $\psi(E)$, the chance ψ assigns to E,

already follows as a "theorem of quantum probability theory", together with the usual coherence requirements on rational credence. Thus, there is no need to seek a further justification of the principle in the quantum setting, or to agonize over which variant of it is correct.[1]

In this chapter I aim to clarify the status of the QPP. I first note that on some interpretations, quantum chances are modeled by ordinary 'classical' (Kolmogorovian) probability theory, not quantum probability theory, and so there is no need for a new treatment of the QPP on these interpretations. On other interpretations, Earman is right to point out that QPP requires its own special formulation, and that this is an important unaddressed issue in the chance literature. However, I argue, Earman's proposed formulation is not the correct one. I lay out what I take to be the correct formulation, and show that although intuitively true, it does not follow as a theorem.

Before proceeding, it is worth mentioning a further reason that philosophers and physicists should be interested in the status of the quantum principal principle. An important practice within physics is the reconstruction of unknown quantum states through measurement. This process, known as *quantum state tomography*, is one of several methods used to learn and estimate quantum states (Paris and Rehacek, 2004). To the extent we think of the physicists in these scenarios as Bayesian agents with credences, we may be interested in modeling their learning experiences as they hone in on the unknown state. Here QPP plays a crucial role. Suppose ψ is the true state of the system and E is some experimental data. By Bayes' formula:

$$C(\text{state is } \psi|E) = \frac{C(E|\,\text{state is } \psi)\cdot C(\text{state is } \psi)}{C(E)}. \tag{3.1}$$

Crucially, the QPP fixes $C(E|\,\text{state is } \psi)$ to $\psi(E)$. This ensures that the higher the chance ψ assigns to the observation E, the stronger the true hypothesis is supported. The QPP also helps fix the value of the denominator $C(E)$: letting $\{\psi_i\}$ denote the candidate states, we have

$$C(E) = \sum_i C(E|\,\text{state is } \psi_i)\cdot C(\text{state is } \psi_i) = \sum_i \psi_i(E)\cdot C(\text{state is } \psi_i), \tag{3.2}$$

in other words the agent's credence in the experimental outcome E is equal to her expectation of its chance.

[1] How have other authors managed to overlook this point? Earman (2018, p. 2) attributes it, in part, to a lack of detailed attendance to the physics: "[S]ince the discussion seems to be constrained only by intuitions about how chance ought to work in the actual and other possible worlds it is not surprising that fundamental disagreements have arisen both about how to capture in precise form the idea behind Lewis' PP and about how to justify PP as a principle of rationality of belief. Even more disconcerting is the fact that the literature makes little contact with quantum theory, despite the fact that this is inarguably one of the most successful theories of modern physics and arguably the theory that speaks most clearly of objective chance."

The overall result is that the agent's credence in the state hypotheses evolves in the way familiar from Bayesian statistical inference (cf. Lewis (1980, pp. 285–7) and Chapter 2). One of my points will be that Earman's formulation of the QPP can't do the normative work required for this application to quantum state tomography, and statistical inference more generally. At the end of the chapter, I will show that my formulation can.

3.1 Earman on quantum chance

In *quantum probability theory*,[2] the objects of quantum chance—what I will call, for lack of a better term, the quantum propositions—correspond mathematically to the set of all projection operators on the Hilbert space of the system in question. *Quantum chance functions* are then defined as functions on this algebra of quantum propositions. They output numbers between 0 and 1, and obey similar axioms to 'classical' (Kolmogorovian) probability functions.

In more detail, let $\mathcal{H}$ be a separable Hilbert space and let $\mathfrak{N}$ denote the von Neumann algebra of all bounded operators acting on $\mathcal{H}$. A *projection* $E \in \mathfrak{N}$ is a self-adjoint element satisfying $E^2 = E$. Let $\mathcal{P}(\mathfrak{N})$ denote the set of projections in $\mathfrak{N}$. Then a *quantum chance function* is a map $\omega : \mathcal{P}(\mathfrak{N}) \to [0, 1]$ satisfying:

1. $\omega(I) = 1$, where $I \in \mathcal{P}(\mathfrak{N})$ is the identity operator (analogue of the necessity), and
2. $\omega(\bigvee_i E_i) = \sum_i \omega(E_i)$ for any countable collection $\{E_i\}$ of mutually orthogonal projections.[3]

Example 1. In the Stern Gerlach case, we can consider the projection $P_{\uparrow_x} = |\uparrow_x\rangle \langle\uparrow_x|$ corresponding to the proposition that the 'UP' outcome will occur. Since the particle is in state $\uparrow_z$, the quantum chance function is given by $\omega(E) = \langle\uparrow_z|E|\uparrow_z\rangle$ for all projections $E \in \mathcal{P}(\mathfrak{N})$. One can verify that ω satisfies both axioms. In addition, $\omega(P_{\uparrow_x}) = \langle\uparrow_z| \left(|\uparrow_x\rangle \langle\uparrow_x|\right) |\uparrow_z\rangle = |\langle\uparrow_z|\uparrow_x\rangle|^2 = 1/2$ as expected.

We just saw that specifying the quantum state fixes the quantum chance function. It turns out that when the dimension of $\mathcal{H}$ is greater than two, the converse also holds: specifying the quantum chance function ω uniquely fixes a quantum state $\omega : \mathfrak{N} \to \mathbb{C}$, which we also denote ω.[4] If ω is a *vector state*, in other words if there exists a vector $\psi \in \mathcal{H}$ such that $\omega(A) = \langle\psi|A|\psi\rangle$ for all $A \in \mathfrak{N}$, then we simply write ω as ψ.

[2] Here I follow the presentation of Earman (2018); cf. also von Neumann (1955), Bub (1977), Hughes (1989, Ch. 8), Pitowsky (2006), Rédei and Summers (2007), Strocchi (2008, Ch. 2.4), Hemmo and Shenker (2020).

[3] Here $\vee$ is the 'join' operator, the analogue of disjunction: if E_1 and E_2 are projection operators, then $E_1 \vee E_2$ is the projection corresponding to the closure of $\mathrm{Ran}(E_1) \cup \mathrm{Ran}(E_2)$. Two projection operators are E_1 and E_2 are *orthogonal* when $E_1 E_2 = \mathbf{0}$. A collection is *mutually orthogonal* if every pair is orthogonal.

[4] This follows from Gleason's theorem. Here a *quantum state* is a normed positive linear functional $\omega : \mathfrak{N} \to \mathbb{C}$ on $\mathfrak{N}$. Assuming ω is *normal* (cf. Ruetsche (2011)), there exists a unique density operator $\rho_\omega \in \mathfrak{N}$ with $\omega(A) = \mathrm{Tr}[A\rho_\omega]$ for all $A \in \mathfrak{N}$. In this case it is common to use ω and ρ_ω interchangeably.

What about credence functions? Earman assumes that in the quantum setting, rational credence functions C are also defined over the quantum propositions $\mathcal{P}(\mathfrak{N})$. Furthermore, Earman imposes two basic coherence requirements on these functions:

Probabilism (quantum analogue): rational C satisfy the axioms listed above, and

Conditionalization (quantum analogue): rational C obey an analogue of Bayesian updating, called *Lüder's updating*, which specifies that for all $E \in \mathcal{P}(\mathfrak{N})$:[5]

$$C(\cdot|E) = \frac{C(E \cdot E)}{C(E)}, \qquad \text{provided } C(E) > 0.$$

Furthermore, any rational C can be viewed as the result of a Lüder's update (or series of updates) on a rational initial credence function or ur-prior C_0. It will be convenient in this chapter to assume unless otherwise specified that C coincides with C_0 (in other words, that the agent starts out with trivial total evidence). In particular, I will set aside the issue of admissibility, as it is not the main focus.

Say a function $C : \mathcal{P}(\mathfrak{N}) \to [0, 1]$ is *coherent* if it satisfies the quantum analogues of probabilism and conditionalization. Earman's main claim is that any coherent credence function must automatically satisfy the QPP, thanks to the following result:

Theorem 1 (Ruetsche and Earman (2011) Fact 1)**.** *Set-up as above, where we suppose* $\dim(\mathcal{H}) > 2$. *Let* $C : \mathcal{P}(\mathfrak{N}) \to [0, 1]$ *be any coherent credence function. Let* $\psi \in \mathcal{H}$ *be any vector state and let* $P_\psi \in \mathcal{P}(\mathfrak{N})$ *denote the projection onto the vector.*[6] *Then, provided* $C(P_\psi) > 0$,

$$C(\cdot|P_\psi) = \psi(\cdot). \tag{3.3}$$

In other words, conditional on P_ψ, *the credences* C *must align with the chances* ψ.

(Note that here I am using P_ψ to denote a projection operator, not a probability function; unfortunately, this is a place where probability theory notation and quantum mechanics notation do not mix so nicely.)

Based on this result, Earman concludes that the truth of the QPP is not a matter "to be submitted to the intuitions of wise analytical metaphysicians" but rather one to be proved: "credence and chance have been brought into alignment without the cudgel of any extra normative principle" (Earman, 2018, p. 10).

[5]Here it is important that C can be extended to a state $C : \mathfrak{N} \to \mathbb{C}$ (which follows from probabilism; see the previous footnote), since for $F \in \mathcal{P}(\mathfrak{N})$, EFE may not be a projection operator, leaving the numerator $C(EFE)$ otherwise undefined.

[6] Here it crucial that ψ is a vector state, in particular that it is *normal* (cf. footnote 4) and *pure*; in ordinary QM, where $\mathfrak{N}$ is the von Neumann algebra, the pure normal states coincide with the vector states. Notably, the result cannot be extended to entangled ω, for instance; see Section 3.3 for discussion.

3.2 Assessing Earman's argument

Note that for Earman's result to establish QPP, it is crucial that the projection P_ψ represent the proposition $\langle$the state of the system is $\psi\rangle$, or $\langle$the chances are given by $\psi\rangle$. My main criticism of Earman's argument is that this claim is false, and so Equation (3.3) does not establish the Principal Principle.

Before I elaborate on this criticism, I should clarify an important issue regarding the scope of Earman's argument. Earman's choice to model quantum chances using quantum probability theory—and in particular to define chances over the set $\mathcal{P}(\mathfrak{N})$ of *all* projection operators—is not uncontroversial. Many quantum theories end up privileging a certain subset of commuting operators. This often brings them into the setting of ordinary 'classical' Kolmogorovian probability theory (see also Loewer (1994) for further discussion). For example, in GRW theory, position is given a privileged role. The objective chances in this theory are the chances attached to spontaneous collapses of the position wave function. At a given time, the possible states of the world at the next time form the relevant set of possibilities. Chance is more naturally represented by a classical probability measure over this set: it selects which of these disjoint possibilities obtain, with the specific probabilities determined by the current spatial wave function and the given collapse constant. (Of course, spin and other degrees of freedom are still subject to chancy jumps, but only insofar as they are entangled with the spatial wave function.)

Similar remarks apply to Bohmian mechanics and the Many Worlds Interpretation. Bohmian mechanics is deterministic and so does not involve objective chances in Lewis's original sense. But insofar as it does feature a kind of objective chance, as I think it does, it is represented by an ordinary classical measure over initial particle configurations. Again, this is related to the privileging of position in the theory. Similarly, the status of chance in the Many Worlds Interpretation is a fraught issue, but to the extent it features in the theory, objective chance is more naturally represented as a classical probability measure—this time over disjoint branching possibilities. What makes this possible is the selection, through decoherence (resulting from system-environment interactions), of a preferred 'quasi-classical' basis over which branching occurs.

The upshot is that on many quantum theories, standard treatments of the principal principle will carry over, and Earman's discussion won't apply.[7] That said, on other interpretations, including but not limited to quantum logic interpretations (Stairs, 1983), information-theoretic interpretations (Bub, 2005), and certain pragmatist interpretations (Healey, 2017), quantum chances are more naturally understood as ranging over all projection operators. In these settings, Earman is correct to point out that we need to resituate QPP within quantum probability theory. Thus, for the rest of this chapter, I will assume we are in such a setting.

[7]This is not to say that quantum probability theory is misguided, or that it doesn't play an important role in our understanding of quantum mechanics. Rather this is to say that on several foundational approaches, ordinary 'classical' probability theory is ultimately the more appropriate means to represent quantum chances.

Returning to my main criticism, here is one way to see that the projection P_ψ does not represent the proposition $\langle$the state of the system is $\psi\rangle$. Imagine an agent is trying determine the quantum state of a spin one-half particle. Suppose she eventually becomes very confident that $\langle$the state is $\uparrow_x\rangle$. Then she will have low confidence that $\langle$the state is $\uparrow_z$ or $\downarrow_z\rangle$. But assuming she is coherent,

$$C(P_{\uparrow_z} \vee P_{\downarrow_z}) = C(I) = 1 \tag{3.4}$$

always. More generally, for any orthonormal basis $\{\psi_i\}$, even though the disjunction $\bigvee_i\langle$the state is $\psi_i\rangle$ is not a necessity (the true state may be some superposition of basis elements), the disjunction $\bigvee_i P_{\psi_i}$ is a necessity, and so will always get assigned probability one. This strongly suggests that the projections P_{ψ_i} should not be used to represent propositions like $\langle$the state is $\psi_i\rangle$, because they do not have the appropriate logical structure.

To press the point, note that if ψ_1 and ψ_2 are distinct states, then we would expect $C(\text{state is } \psi_1 \text{ or } \psi_2) = C(\text{state is } \psi_1) + C(\text{state is } \psi_2)$. However, it is well known (Strocchi, 2008, p. 51) that we may have

$$C(P_{\psi_1} \vee P_{\psi_2}) > C(P_{\psi_1}) + C(P_{\psi_2}) \tag{3.5}$$

when ψ_1 and ψ_2 are non-orthogonal.[8]

One might try to dismiss these problems on the grounds that they rely on oddities related to the join operator $\vee$ and the non-orthogonality of states. However, even if we set these issues aside, problems remain.

Imagine that the agent is not sure which of some collection of states $\{\psi_i\}$ the system is in, but she is certain the system is in one of them, so $\sum_i C(\text{state is } \psi_i) = 1$. Suppose for simplicity that the collection $\{\psi_i\}$ is finite and that the $\{\psi_i\}$ are all mutually orthogonal. Then, recalling the discussion of quantum state tomography from the introduction (see (3.2)) we would expect

$$C(E) = \sum_i C(E|\,\text{state is } \psi_i) \cdot C(\text{state is } \psi_i),$$

i.e. the agent's credence in E equals her expectation taken over the mutually exclusive and exhaustive hypotheses about the state. However, the example below shows that we may have

$$C(E) > \sum_i C(E|P_{\psi_i}) \cdot C(P_{\psi_i}), \tag{3.6}$$

[8]Consider for example a spin one-half particle (examples in higher dimensions can also be given) and let C be the coherent credence function corresponding to the vector state $\uparrow_z$. Suppose $\psi_1 = \uparrow_x$ and $\psi_2 = \downarrow_z$. Then $P_{\psi_1} \vee P_{\psi_2} = I$ so $C(P_{\psi_1} \vee P_{\psi_2}) = 1$, but $C(P_{\psi_1}) + C(P_{\psi_2}) = |\langle\uparrow_z \mid \uparrow_z\rangle|^2 + |\langle\uparrow_z \mid \downarrow_z\rangle|^2 = 1/2 + 0 < 1$.

even though $\sum_i C(P_{\psi_i}) = 1$ and the $\{\psi_i\}$ are mutually orthogonal. This is further indication that the projections P_{ψ_i} do not represent state hypotheses. (For another similar cautionary tale about the interpretation of projection operators, see also Weinstein (1996).)

Example 2. Consider a spin one-half particle (the example can be extended to higher dimensions; see next footnote) and let the collection $\{\psi_i\}$ consist of $\uparrow_z$ and $\downarrow_z$. Let C be the coherent credence function corresponding to the vector state $\uparrow_x$; note that $\sum_i C(P_{\psi_i}) = C(P_{\uparrow_z}) + C(P_{\downarrow_z}) = 1/2 + 1/2 = 1$, as required. Now let $E = P_{\uparrow_x}$. Then $C(E) = 1$, but $C(E|P_{\uparrow_z}) \cdot C(P_{\uparrow_z}) + C(E|P_{\downarrow_z}) \cdot C(P_{\downarrow_z}) = 1/2 \cdot 1/2 + 1/2 \cdot 1/2 < 1$.[9]

An important consequence of this last result is that Earman's alleged principal principle, even if interpreted as such, does not do the normative work that it is supposed to do. (3.3) and (3.6) imply:

$$C(E) \neq \sum_i \psi_i(E) \cdot C(P_{\psi_i}) \tag{3.7}$$

even though $\sum_i C(P_{\psi_i}) = 1$. If we do interpret P_{ψ_i} as $\langle$the state is $\psi_i\rangle$, as Earman's formulation of QPP suggests, then we are led to conclude that:

$$C(E) \neq \sum_i \psi_i(E) \cdot C(\text{state is } \psi_i), \tag{3.8}$$

even though $\sum_i C(\text{state is } \psi_i) = 1$; in other words, the agent's credences need not match her expectations of the chances.[10] But as we saw in the discussion of quantum state tomography, this is supposed to be a crucial consequence of QPP.

3.3 Preparation, projective measurement, and a classical analogue

If Earman's theorem does not establish QPP, then what does it establish?

The theorem establishes that if the system was knowingly prepared in state ψ by a projective measurement with outcome P_ψ, then rational credence must align with ψ. Indeed this is the gloss that Earman offers in several places (2018, p. 10, notation adapted):

> The intended interpretation of this [result] should be obvious. Suppose that an agent whose initial

[9] Here we apply (3.3) to obtain $C(E|P_{\uparrow_z}) = \uparrow_z(E) = 1/2$ and $C(E|P_{\downarrow_z}) = \downarrow_z(E) = 1/2$. More generally, let $d = \dim(\mathcal{H})$ where $1 < d < \infty$. Let $\{\psi_i\}$ be an orthonormal basis and define $\varphi = \sum_i \sqrt{\frac{1}{d}}\,\psi_i$. Set $C(\cdot) = \varphi(\cdot)$ and $E = P_\varphi$. Then $C(E) = 1$ but $\sum_i \psi_i(E) \cdot C(P_{\psi_i}) = \frac{1}{d} \cdot \sum_i C(P_{\psi_i}) = \frac{1}{d} < 1$.

[10] It is true that given C it can always be expressed as $C(\cdot) = \sum_i \psi_i(\cdot) \cdot C(P_{\psi_i})$ for *some* orthonormal $\{\psi_i\}$. However this is changing the subject: given any collection of mutually exclusive and exhaustive chance hypotheses, the agent's credences should equal her expectations over those chance hypotheses. At the very least, we would expect that if all the $\{\psi_i\}$ assign the same chance to E, then $C(E)$ will equal that chance. But if the proposition $\langle$state is $\psi_i\rangle$ is represented by the projection P_{ψ_i} then this claim also fails: note that in the previous example, $\uparrow_z(E) = \downarrow_z(E) = 1/2$ yet $C(E) \neq 1/2$.

> credence function on $\mathcal{P}(\mathfrak{N})$ is C learns that a Yes-No experiment for P_ψ has been performed and that a Yes answer has been returned. Since she is rational she Lüders updates her credence function to $C(\cdot|P_\psi)$. On the objectivist interpretation of quantum probabilities, the returning of a Yes answer implies that the normal pure state ψ has been prepared, and, hence, that the objective chance of an event $E \in P$ is $\psi(E)$, which is the same as the agent's updated credence, $C(E|P_\psi)$.

Although interestingly related to QPP, this result is not sufficient for it. There are two main ways to see that it falls short.

First, note that many preparations do not involve projective measurements (Fröhlich and Schubnel, 2016): for example, I can prepare a crystal in its ground state ψ by leaving it in a cold room. Yet, QPP still requires I conform my credences to ψ, even though no projective measurement took place.

Furthermore, it seems that even if no preparation of the system has taken place at all, for example if I going to measure cosmic rays incident on my laboratory that nobody has interacted with, QPP should still require I conform my credence in accordance with the expected state. Yet in these cases Earman's result won't apply.

Third, the result only applies to vector states $\psi \in \mathcal{H}$. But suppose that the system S is entangled with another S'. If the agent knows S's state, it seems she should still be required to set her credences in line with that state, but since the state won't be pure, the result won't apply. (Here Earman could respond that we should move to the composite system $\mathcal{H}_c = \mathcal{H} \otimes \mathcal{H}'$ and consider pure states $\psi_c \in \mathcal{H}_c$ of it. But I am imagining the agent is only probing S, so there is no "Yes-No experiment for P_{ψ_c}" being performed, only Yes-No experiments for propositions of the form $F \otimes I$, where $F \in \mathcal{P}(\mathfrak{B}(\mathcal{H}))$ and $I \in \mathcal{P}(\mathfrak{B}(\mathcal{H}'))$ the identity.)

In fact, attending to this restriction to pure states, we can see that an exactly analogous result also holds in classical probability theory. Some preliminaries: Let $(S, \mathcal{S})$ be a classical outcome space, with S a set and $\mathcal{S}$ a sigma-algebra on S. Define a *state* μ as a probability measure on $(S, \mathcal{S})$. A state μ is *pure* if it cannot be expressed as a trivial convex combination of other states. One can check that in this classical context, a state is pure if and only if it is a point measure concentrated at some $s \in S$, which I will write as μ_s. It follows the pure states μ_s are in one-to-one correspondence with the elements of S. Next, define a *projection operator* as a random variable $X : S \to \mathbb{R}$ that satisfies $X^2 = X$. One can check that every projection can be expressed as an indicator $X = 1_J$ for some $J \in \mathcal{S}$, in other words a $\{0, 1\}$-valued function:

$$1_J(s) = \mu_s(J) = \begin{cases} 1 & \text{if } s \in J \\ 0 & \text{if } s \in S \backslash J. \end{cases}$$

For some pure state $\psi = \mu_s$ we can consider the projection $P_\psi = 1_{\{s\}}$ which tests whether $s' \in S$ belongs to $\{s\}$. Like in the quantum case, we also associate P_ψ with the claim that a Yes-No experiment for P_ψ has been performed, and that a Yes answer has been returned: $P_\psi = \{s' \in S : 1_{\{s\}}(s') = 1\} = \{s\}$. Now, with these preliminaries in hand, we can obtain a classical analogue of the earlier Theorem:

Theorem 2 (Classical analogue)**.** *Let $C : \mathcal{S} \to [0, 1]$ be any coherent credence function. Let ψ be any pure state and let P_ψ denote the projection defined above. Then, provided $C(P_\psi) > 0$,*

$$C(\cdot|P_\psi) = \psi(\cdot).$$

For proof, recall we can write $\psi = \mu_s$ for some $s \in S$ and $P_\psi = \{s' \in S : 1_{\{s\}}(s') = 1\} = s$ where we abuse notation and write $\{s\}$ as s. Now if $C(s) \neq 0$, then for all $J \in \mathcal{S}$,

$$C(J|P_\psi) = C(J|s) = \frac{C(Js)}{C(s)} = \begin{cases} 1 & \text{if } s \in J, \\ 0 & \text{if } s \in S \backslash J, \end{cases}$$

and so $C(J|s) = \mu_s(J) = \psi(J)$ as desired.

Intuitively, what's going on here is that since ψ is pure, P_ψ is an *atomic proposition* that specifies the exact outcome. Conditionalizing on this outcome, one is forced to align one's credence with the post-experiment chance of the outcome, which is either 0 or 1. And so, in this very restricted sense, "credence and chance have been brought into alignment without the cudgel of any normative principle."

Needless to say, while a notable result, this does not constitute a proof of the principal principle. And as I have suggested here, although its quantum analogue is much more interesting, it is not a proof of the quantum principal principle either.

This discussion also makes evident the reliance, in Earman's result, on the choice to focus on the von Neumann algebra $\mathfrak{N} = \mathfrak{B}(\mathcal{H})$, which has atoms. In some quantum field theoretic settings, such as in the setting of Type-III factor algebras, there are no atoms in the relevant algebra (Ruetsche and Earman, 2011). The prospect of extending Earman's result to such settings looks dim.

3.4 Formulating the quantum principal principle

I now present my positive proposal for how to formulate the QPP within the framework of quantum probability theory. As we'll see, while the resulting principle is highly intuitive, it does not follow as a theorem.

We begin by dividing the agent's credence function C into two components, x and μ. The first component,

x, is what I previously denoted C: it is a function on the set of projections $\mathcal{P}(\mathfrak{N})$ that, at minimum, satisfies the coherence requirements from Section 3.1. The second component, μ, is the key ingredient that was missing from Earman's approach. It is a (classical) credence distribution over the set of possible initial quantum states of the system. Roughly speaking, whereas x represents the agent's ignorance about experimental outcomes, μ represents her ignorance about the state. The QPP then constrains how x and μ relate. It implies, for instance, that the agent's credence $x(E)$ equals her expectation of the chance of E,

$$x(E) = \sum_i \psi_i(E) \cdot \mu(\psi_i) \qquad \forall E \in \mathcal{P}(\mathfrak{N}), \tag{3.9}$$

where the different state hypotheses are weighted by μ and I have assumed, for simplicity, that μ is a discrete distribution over finitely many vector states $\{\psi_i\}$. Even at this point, we can see that QPP is not a theorem of quantum probability theory, since there are many pairs (x, μ), with x and μ both coherent, that violate (3.9). For instance, suppose x corresponds to a vector state like $\uparrow_z$ and μ assigns probability less than 1 to $\uparrow_z$. Then $x(P_{\uparrow_z}) = 1 > \sum_i \psi_i(P_{\uparrow_z}) \cdot \mu(\psi_i)$. As I understand the QPP, then, it is a norm that goes beyond usual coherence requirements. Let us now formulate this norm in full detail.

Fix a separable Hilbert space $\mathcal{H}$ and associated algebra of bounded operators $\mathfrak{N}$. Let K denote the set of quantum states on $\mathfrak{N}$. This set comes naturally equipped with the weak* topology. Let $\sigma(K)$ denote the sigma-algebra on K generated by this topology. Now we can consider Kolmogorovian probability measures μ on $(K, \sigma(K))$. As in Section 3.1, we can also consider quantum probability functions x on $\mathcal{P}(\mathfrak{N})$. A coherent credence function C is then modeled as a pair (x, μ).

Example 3. In the case of a spin one-half particle, μ can be thought of as a probability density over the Bloch sphere of spin states, and x a point in the sphere. If (x, μ) obeys QPP, then x will be the 'weighted average' of the points supported by μ.

We are interested in how a rational C changes with new evidence. If the evidence is an experimental outcome $E \in \mathcal{P}(\mathfrak{N})$, then we know x updates by Lüder's conditionalization,

$$x_E(F) = \frac{x(EFE)}{x(E)} \qquad \forall F \in \mathcal{P}(\mathfrak{N}).$$

However, to know how C updates, we also need to know how μ updates: after learning data E, how should the agent's credences in hypotheses H about the (initial) state of the state of the system change? A coherent agent obeys Bayes' formula, so we set:

$$\mu_E(H) = \frac{x_H(E) \cdot \mu(H)}{x(E)} \qquad \forall H \in \sigma(K). \tag{3.10}$$

In the case $H = \{\psi\}$, this is just a more elaborate way of stating (3.1).

What determines the agent's credence $x_H(E)$ in data E given H? Enter the QPP:

Quantum Principal Principle (QPP): If $C = (x, \mu)$ is a rational initial credence function, then

$$x_H(E) = \int_K \omega(E)\, d\mu_H(\omega) \tag{3.11}$$

for all experimental outcomes $E \in \mathcal{P}(\mathfrak{N})$ and state hypotheses $H \in \sigma(K)$ such that $\mu(H) > 0$.

Equation (3.11) says the agent's credence in data E given state hypothesis H is the expectation of the chance of E, restricted to states in H. In particular, here μ_H is defined in the standard way

$$\mu_H(\cdot) = \frac{\mu(\cdot \cap H)}{\mu(H)} \qquad \text{provided } \mu(H) > 0,$$

which, substituted into (3.11), yields $x_H(E) = \frac{1}{\mu(H)} \int_H \omega(E)\, d\mu(\omega)$.

In the special case where H is the trivial (necessary) state hypothesis K, we obtain the result that the agent's unconditional credence in E is her expectation of the chance of E,

$$x(E) = \int_K \omega(E)\, d\mu(\omega), \qquad \forall E \in \mathcal{P}(\mathfrak{N}). \tag{3.12}$$

In mathematical terminology, x is the *barycenter* for μ (Alfsen, 2012). Note that this reduces to (3.9) when μ is a discrete distribution over finitely many vector states.

In the special case where $H = \{\nu\}$ is a hypothesis about a specific state ν then, assuming $\mu(\nu) > 0$, QPP yields:

$$x_\nu(E) = \frac{\int_{\{\nu\}} \omega(E)\, d\mu(\omega)}{\mu(\nu)} = \frac{\nu(E) \cdot \mu(\nu)}{\mu(\nu)} = \nu(E), \tag{3.13}$$

in other words the agent's credence in E given that the quantum state is ν must match the chance of E according to ν. This is exactly what QPP is supposed to capture. (And unlike Earman's formulation (3.3), it applies both when ν is a vector state ψ and when ν is an entangled state.)

Substituting these results into the Bayes' formula (3.10), we obtain:

$$\mu_E(\nu) = \frac{\nu(E) \cdot \mu(\nu)}{\int_K \omega(E)\, d\mu(\omega)}. \tag{3.14}$$

If μ is a discrete distribution over finitely many vector states $\{\psi_i\}$, then this becomes, $\mu_E(\psi) = \frac{\psi(E) \cdot \mu(\psi)}{\sum_i \psi_i(E) \cdot \mu(\psi_i)}$, which is just a way of spelling out (3.1) and (3.2).

Example 4. A physicist has two identically prepared electrons. She knows they are either in state $\uparrow_x$, $\uparrow_y$,

or $\uparrow_z$, and is equally confident in these three hypotheses. She then measures the first electron in the spin-z direction and obtains an 'UP' outcome. How should her confidences over the hypotheses about the (initial) spin state change? Here we have

$$\mu(\uparrow_x \otimes \uparrow_x) = \mu(\uparrow_y \otimes \uparrow_y) = \mu(\uparrow_z \otimes \uparrow_z) = \frac{1}{3}.$$

The data the agent received is represented by $E = P_{\uparrow_z} \otimes I$. Assuming she satisfies QPP, her posterior credence in the hypothesis $\uparrow_z \otimes \uparrow_z$ is determined by (3.14):

$$\mu_E(\uparrow_z \otimes \uparrow_z) = \frac{\frac{1}{3} \cdot (\uparrow_z \otimes \uparrow_z)(E)}{\frac{1}{3}\left[(\uparrow_x \otimes \uparrow_x)(E) + (\uparrow_y \otimes \uparrow_y)(E) + (\uparrow_z \otimes \uparrow_z)(E)\right]} = \frac{\frac{1}{3} \cdot 1}{\frac{1}{3}\left[\frac{1}{2} + \frac{1}{2} + 1\right]} = \frac{1}{2}.$$

As expected, the agent's credence in the $\uparrow_z$ hypothesis increases. A similar calculation shows that, as expected, her credence in the $\uparrow_x$ and $\uparrow_y$ hypotheses decreases:

$$\mu_E(\uparrow_x \otimes \uparrow_x) = \frac{\frac{1}{3} \cdot \frac{1}{2}}{\frac{1}{3}\left[\frac{1}{2} + \frac{1}{2} + 1\right]} = \frac{1}{4}, \qquad \mu_E(\uparrow_y \otimes \uparrow_y) = \frac{\frac{1}{3} \cdot \frac{1}{2}}{\frac{1}{3}\left[\frac{1}{2} + \frac{1}{2} + 1\right]} = \frac{1}{4}.$$

So the agent begins with credence (x, μ) and ends with credence (x_E, μ_E), where $x_E(\cdot) = x(E \cdot E)/x(E)$ and μ_E is specified by the above three equations.

What if the agent then obtains a second piece of experimental evidence $F \in \mathcal{P}(\mathfrak{N})$, say, through measuring the spin of the second electron? Here we need to check that our prescription for credal change is consistent, in the sense that if E and F commute, then beginning with (x, μ) and learning EF gives the same result as receiving E, updating to (x_E, μ_E), and then learning F. A complication arises related to the fact that QPP is formulated for *initial* credence functions. If the agent already has total evidence E, then we need a slight generalization of (3.11):

$$x_{EH}(F) = \int_K \omega_E(F)\, d\mu_{EH}(\omega) \qquad \forall E \in \mathcal{P}(\mathfrak{N}), H \in \sigma(K), F \in \mathcal{P}(\mathfrak{N}), \tag{3.15}$$

where μ_{EH} is defined as the result of conditioning μ_E on H, or equivalently, μ_H on E. Now, one can check that receiving EF, and so shifting to $x_{EF}(\cdot) = \frac{x(FE \cdot EF)}{x(EF)}$ and $\mu_{EF}(\cdot) = \frac{x_{(\cdot)}(EF) \cdot \mu(\cdot)}{x(EF)}$, gives the same result as

receiving E and then F, and so shifting to $(x_E)_F(\cdot)$ and $(\mu_E)_F(\cdot)$.[11]

Example 4 (continued)**.** The agent then measures the second electron in the spin-y direction and obtains a 'down' outcome, represented by $F = I \otimes P_{\downarrow_y}$. How should her credences over her three hypotheses about the state change now? Again using the QPP, in particular (3.14) but with μ_E instead of μ determining the weights, we have

$$\mu_{EF}(\uparrow_z \otimes \uparrow_z) = \frac{\frac{1}{2} \cdot (\uparrow_z \otimes \uparrow_z)(F)}{\frac{1}{4}[(\uparrow_x \otimes \uparrow_x)(F) + (\uparrow_y \otimes \uparrow_y)(F)] + \frac{1}{2}(\uparrow_z \otimes \uparrow_z)(F)} = \frac{\frac{1}{2} \cdot \frac{1}{2}}{\frac{1}{4}[\frac{1}{2} + 0] + \frac{1}{2} \cdot \frac{1}{2}} = \frac{2}{3}.$$

By similar reasoning, $\mu_{EF}(\uparrow_x \otimes \uparrow_x) = \frac{1}{3}$ and $\mu_{EF}(\uparrow_y \otimes \uparrow_y) = 0$. As expected, the $\downarrow_y$ result ruled out the hypothesis that the electrons were prepared in the $\uparrow_y$ state, but did not change the relative likelihood of the $\uparrow_z$ and $\uparrow_x$ hypotheses.

This treatment can be extended to more complicated quantum state tomography set-ups involving more copies, as well as other state estimation scenarios, as I will discuss in the next chapter.

3.5 Conclusion

How do quantum chances constrain credence? At least some interpretations of quantum mechanics demand that we recast the principal principle within the framework of *quantum probability theory*. Earman (2018) sought to do so, and came to a surprising conclusion: the principle follows as a theorem. I started down the same path as Earman, but came to a different conclusion. The quantum principal principle, appropriately formulated, is not a theorem of quantum probability theory; its justification does not come so straightforwardly. However, I still found that bringing the principal principle into the domain of quantum probability theory is an important and overlooked project. Not only does it yield new insight into the principle itself, as Earman showed, it also allows us to bring the principle to bear on important applications, like the learning of quantum states.

[11] For the derivation, first note that we have $(x_E)_F(\cdot) = \frac{x_F(E\cdot E)}{x_F(E)} = \frac{x(FE\cdot EF)}{x(FEF)}$ and:

$$(\mu_E)_F(\cdot) = \frac{x_{E(\cdot)}(F) \cdot \mu_E(\cdot)}{x_E(F)} = \frac{\int_{(\cdot)} \omega_E(F)\, d\mu_{EH}(\omega) \cdot \mu_E(\cdot)}{\frac{x(EFE)}{x(E)}} = \frac{\frac{\int_{(\cdot)} \omega_E(F)\, d\mu_E(\omega)}{\mu_E(\cdot)} \cdot \mu_E(\cdot)}{\frac{x(EFE)}{x(E)}}$$

$$= \frac{\int_{(\cdot)} \frac{\omega(EFE)}{\omega(E)} \cdot \frac{x_\omega(E)}{x(E)}\, d\mu(\omega)}{\frac{x(EFE)}{x(E)}} = \frac{\int_{(\cdot)} \frac{\omega(EFE)}{\omega(E)} \cdot \omega(E)\, d\mu(\omega)}{x(EFE)} = \frac{\int_{(\cdot)} \omega(EFE)\, d\mu(\omega)}{x(EFE)} = \frac{x_{(\cdot)}(EFE) \cdot \mu(\cdot)}{x(EFE)},$$

where here for second equality we apply (3.15), for the fifth we apply (3.13), and for the last equality we apply (3.11). Second, note that since E and F are commuting projections, we have $x(FEF) = x(EF) = x(EFE)$.

Chapter 4

The impossibility of individual state determination, revisited

While we can successfully learn quantum states in many situations, there is a particular kind of learning scenario that physicists and philosophers are interested in: It is conjectured that when only a *single copy* of a system is available, it is impossible to learn the quantum chances associated with that system.

Intuitively, the thought is that in order to successfully reconstruct the initial quantum state of a system, one needs to be able to measure that state in multiple ways. For example, take a spin one-half particle. If I measure its spin in the z direction and get the result 'UP', I cannot tell from that observation alone whether this system was in $\uparrow_z$ or $\uparrow_x$, since both are compatible with the outcome. So, I need also measure its spin in another direction—say, y. However, the problem is that the first measurement will have already disturbed the state (causing it to collapse to $\uparrow_z$) and so any subsequent measurement on that system won't be a reliable indication of its initial state. That is why, when we perform quantum state tomography, we begin with many 'identically prepared copies'—systems assumed to be in the same quantum state—and then measure them in different ways, i.e. for different observables. Usually we pick a set of observables which is 'informationally complete' and allows for a full reconstruction of the state (Scott, 2006). In the case of a spin one-half particle, the trio of spin-x, spin-y and spin-z is informationally complete.

The above was an intuitive gloss for why *individual state determination*—the reliable reconstruction of a previously unknown quantum state, given only a single copy—is impossible. But besides being only a gloss, the reasoning in the previous paragraph relies on some controversial assumptions, like that measurements cause collapse. One might wonder whether there is a more carefully formulated argument that could settle the issue. Indeed, this could have implications for debates about whether quantum states only attach

to ensembles (Unruh, 1994), and the significance of non-traditional measurement schemes, like *protective*, as opposed to *impulsive*, *measurements* (Aharonov et al., 1993; Dickson, 1995; Gao, 2015; Combes et al., 2018), which seek to probe individual wave functions without disturbing them. In this chapter, I begin by considering an argument for the impossibility of state determination presented by Busch (1997). I point out that this argument overgeneralizes, and propose a way to improve it which draws upon the framework developed at the end of the previous chapter. Next, I consider another popular argument for the impossibility of individual state determination that has been offered, originally due to D'Ariano and Yuen (1996). I show that this argument also overgeneralizes: if it were sound, it would imply the skeptical conclusion that it is impossible to come to know the quantum state of a system given *any finite number* of copies. However, unlike the previous argument, it is very difficult to diagnose where this argument goes wrong. This puzzle leads me to introduce a problem, the control problem, which forms the basis of the next chapter.

Before proceeding, a caveat. In what follows, I will set aside considerations related to faster-than-light signaling. In particular, I will focus on arguments that can be given within the confines of quantum theory, and do not rely on assumptions about the conditions for peaceful co-existence between quantum mechanics and relativity.[1]

4.1 The argument of Busch (1997)

Busch (1997) offers a simple argument for the impossibility of individual state determination. The argument relies on the POVM formalism, which I outline now, as it will also play a role in my improved version:

Definition 3. Fix a separable Hilbert space $\mathcal{H}$ and von Neumann algebra $\mathfrak{N} = \mathfrak{B}(\mathcal{H})$. Then a *positive operator-valued measure (POVM)* is a set $\{A_k : k = 1, 2, ...\} \subset \mathfrak{N}$ such that $A_k \geq \mathbf{0}$ for all k, and $\sum_k A_k = I$.[2]

An example of a POVM is a basis of mutually orthogonal projection operators. It is a common assumption, in philosophy of quantum mechanics, that any measurement of a system can be represented by a collection of mutually orthogonal projection operators on its Hilbert space. However, it turns out that a more general definition is needed, to take into account possible interactions between the system and its environment. The POVM formalism provides this generalization: if one takes a system + environment that are assumed to together evolve unitarily, then any projective measurement on the larger space can be viewed as the measurement of an 'unsharp observable' A_k on the system's Hilbert space, with the A_k positive and summing to identity (Busch et al., 1996; Nielsen and Chuang, 2002). (And conversely, any POVM can be viewed as the reduction of a projective measurement on a larger Hilbert space (Naimark, 1940, 1943).) Even

[1] Cf. (Bell, 1964; Jarrett, 1984; Shimony, 1986; Maudlin, 2002; Myrvold, 2002; Kent, 2014; Butterfield, 2015).
[2] Here I focus on countable POVMs for simplicity, but we can also extend the definition to continuous cases.

unusual measurement schemes, like protective measurements, can be modeled by POVMs (D'Ariano and Yuen, 1996; Combes et al., 2018).

Now for Busch's argument. Suppose there is a single system and one is uncertain whether its state is ψ or φ, where these two states are non-orthogonal (like $\uparrow_z$ and $\uparrow_y$). If it is possible to determine the state of this system, then there is some POVM measurement we could perform, which has at least one outcome $\mathbf{0} \leq A_k \leq I$, such that:

$$\psi(A_k) = \langle\psi|A_k\psi\rangle = 1, \;\; \varphi(A_k) = \langle\varphi|A_k\varphi\rangle = 0 \qquad \text{or} \qquad \psi(A_k) = \langle\psi|A_k\psi\rangle = 0, \;\; \varphi(A_k) = \langle\varphi|A_k\varphi\rangle = 1,$$

in other words that allows us to distinguish between ψ and φ. But then it would follow that:

$$A_k\psi = \psi, \;\; A_k\varphi = \mathbf{0} \qquad \text{or} \qquad A_k\varphi = \varphi, \;\; A_k\psi = \mathbf{0},$$

and then,

$$\langle\psi|\varphi\rangle = \langle A_k\psi|\varphi\rangle = \langle\psi|A_k\varphi\rangle = 0, \qquad \text{or} \qquad \langle\varphi|\psi\rangle = \langle A_k\varphi|\psi\rangle = \langle\varphi|A_k\psi\rangle = 0,$$

both of which contradict the assumption that ψ and φ are non-orthogonal.

We can see immediately that this argument overgeneralizes. The reasoning applies just as well if we replace ψ and φ by $\psi^{\otimes N}$ and $\varphi^{\otimes N}$ for some finite number N, since if ψ and φ are non-orthogonal it follows that so are $\psi^{\otimes N}$ and $\varphi^{\otimes N}$. Yet we do not want to say that given $\psi^{\otimes N}$ for large N it is impossible to reliably determine the state ψ, since this would be to endorse a radical skepticism about the success of quantum state tomography. What is going on here is that the argument is relying on an implicit premise, that in order to reliably determine that the true state is ψ, one must be able to *perfectly discriminate* between ψ and other states. Such perfect discrimination is impossible, even for an extremely large number of copies, and should not be required of individual state determination either.

4.2 An improved argument

One fix to this problem is to simply replace the equalities in the argument above by approximate equalities $\approx$. This would solve the immediate issue of overgeneralization, since while $\psi^{\otimes N}$ and $\varphi^{\otimes N}$ are not orthogonal, they are approximately so for sufficiently large N: $\left|\langle\psi^{\otimes N}|\varphi^{\otimes N}\rangle\right| = |\langle\psi|\varphi\rangle|^N \approx 0$. However, it would be better to have a more systematic treatment that makes certain crucial assumptions, like the conditions for a successful 'determination' of the state, more explicit. To do this, I draw on the Bayesian framework developed

in the previous chapter, in the context of quantum probability theory. On a natural way of spelling out this Bayesian approach, a successful determination will mean meeting a certain threshold in rational credence for a true hypothesis about the state. Of course, this means that whether a successful determination is possible will depend on the agent's prior. But this is the right result: if the agent already knows the state of the system, say, because of her preparation procedure, then in a trivial sense it is true she can determine the state. The interesting case is when the agent's prior is relatively spread out over state space.

With this in mind, we recall the framework of Chapter 3.4: the agent's beliefs are represented by a pair (x, μ) where x is the agent's prior $x : \mathcal{P}(\mathfrak{N}) \to [0, 1]$ over experimental outcomes, and μ is the agent's prior over the set of possible quantum states K. As before, I will assume μ is a classical probability measure and:

Probabilism (quantum analogue) x satisfies the axioms of quantum probability theory, in particular $x(I) = 1$ and $x(\vee_i E_i) = \sum_i x(E_i)$ for countable mutually orthogonal collections $\{E_i\} \subset \mathcal{P}(\mathfrak{N})$.

As discussed, this implies x can itself be extended to a state $x : \mathfrak{N} \to \mathbb{C}$.[3] This is important for what follows, because to take into account that the outcomes of a POVM may not be projection operators, I will be assuming a generalization of Bayes' formula (3.10), namely that for all operators $\mathbf{0} \leq A \leq I$ in $\mathfrak{N}$ and all $H \in \sigma(K)$:

$$\mu_A(H) = \frac{x_H(A) \cdot \mu(H)}{x(A)} \qquad \text{provided } x(A) > 0.$$

We can also introduce a generalization of Lüder's updating (cf. Busch and Singh (1998)):

Generalized Lüder's Conditionalization $x : \mathfrak{N} \to \mathbb{C}$ obeys *generalized Lüder's updating*, which specifies that that for all operators $\mathbf{0} \leq A \leq I$:[4]

$$x_A(\cdot) = \frac{x(A^{1/2}(\,\cdot\,)A^{\dagger 1/2})}{x(A^{1/2}A^{\dagger 1/2})}, \qquad \text{provided } x(A) > 0.$$

Furthermore, any rational x can be viewed as the result of a generalized Lüder's update (or series of such updates) on a rational initial credence function or ur-prior.

Again, I will assume the quantum principal principle, where below we understand $x_H(\cdot)$ and $\omega(\cdot)$ as functions on all of $\mathfrak{N}$, to take into account not just projection observables E but also operators A:

Quantum Principal Principle (QPP) The following relation between x and μ holds: for all $H \in \sigma(K)$, $x_H(\cdot) \cdot \mu(H) = \int_H \omega(\cdot)\, d\mu(\omega)$, and in particular $x_\nu(\cdot) = \nu(\cdot)$, where this is the case $H = \{\nu\}$. In other words, conditional on the state being ν, the agent's expectations over outcomes conforms with ν.

[3] If $\dim(\mathcal{H}) > 2$, then Gleason's theorem implies this extension is unique. Note that here I assume $\mathcal{H}$ is separable and $\mathfrak{N} = \mathfrak{B}(\mathcal{H})$ is the von Neumann algebra of bounded operators on $\mathcal{H}$.

[4] Note that in the special case $A = E$ a projection, $E^{1/2} = E^{\dagger 1/2} = E$ so this becomes $x_E(\cdot) = \frac{x(E \cdot E)}{x(E)}$ as before.

Finally, I will make an assumption about the ways that the agent can investigate the system:

POVM Assumption The only way the agent will investigate her system(s) is by performing a POVM measurement $\{A_k : k = 1, 2, ...\}$ which will have some outcome A_k.

I motivated this assumption in the previous section. Here it is also relevant to note that, assuming generalized Lüder's conditionalization, any series of multiple or repeated POVM measurements can be represented as a single POVM (D'Ariano and Yuen, 1996). Indeed, there is a sense in which the crucial part of the POVM assumption we will need already follows from the generalized Lüder's conditionalization assumption, in particular the last clause: whatever the agent's posterior x' is after her investigation, it will have to be viewed as the result of a series of generalized Lüder's updates, which is equivalent to saying that it is the result of conditionalization on the outcome of some POVM measurement.

These are the basic background assumptions of the argument. Now we consider a rational agent who begins with prior (x, μ). We suppose she has access to N copies of the system of interest (later we will consider the specific case $N = 1$), so $\mathcal{H} = \otimes_{i=1}^{N} \mathcal{H}_i$ where $\mathcal{H}_i$ are copies of the Hilbert space of the original system, whose dimension we denote by d. For concreteness, let's suppose that she knows the system is in a pure vector state, but is not sure which: her prior μ is uniformly distributed over the set of possible ψ. In particular, as a measure over the space $(K, \sigma K)$ of possible states, μ is supported entirely by symmetric pure states of the form $\psi^{\otimes N}$, and is uniform over those states, or equivalently, preserved under unitary transformations of those states. For convenience, I will let $S \subset K$ denote the set of such states, so $\mu(S) = 1$.

The agent is then going to investigate these N copies, obtain some result A_k, and update her credences (x, μ) to (x_{A_k}, μ_{A_k}). We are particularly interested in her posterior μ_{A_k}. For the true state $\psi \in \mathcal{H}_i$, let H_ψ be the hypothesis that the (initial) state of the system is approximately ψ, in other words, fixing some $\epsilon > 0$:

$$H_\psi = \{\varphi^{\otimes N} \in K : |\langle\varphi|\psi\rangle|^2 \geq 1 - \epsilon\} \in \sigma(K).$$

Our **main result** is then the following upper bound on the agent's posterior credence in this hypothesis:

$$\mu_{A_k}(H_\psi) \leq \frac{N + 1}{(N + d)(1 - \epsilon)}. \tag{4.1}$$

For example, suppose the system is a spin one-half particle ($d = 2$) and the agent has one copy ($N = 1$). We can consider the hypothesis $H_{\uparrow_z}$ that the state is $\uparrow_z$, with room for error of $\epsilon = \frac{1}{10}$. Then, given her uniform prior, the agent's posterior confidence in a state hypothesis cannot rationally exceed $\frac{1+1}{(1+2)(\frac{9}{10})} \approx \frac{3}{4}$. In this sense, she cannot determine the state given only a single copy. Note that, in general, the greater the N, the higher the bound in (4.1). So, unlike the previous argument, this argument does not overgeneralize to imply

the agent cannot determine the state for any finite N.

The derivation of this main result proceeds as follows. Let $f : S \to \mathbb{R}$ denote the map defined by $\varphi^{\otimes N} \mapsto |\langle\varphi|\psi\rangle|^2$. Then, using Markov's inequality—which states that, for a random variable f, the probability of $f \geq a$ is less than or equal to the expectation of f, divided by a—we have:

$$\mu_{A_k}(H_\psi) = \mu_{A_k}(f \geq 1 - \epsilon) \leq \frac{\mathbb{E}_{\mu_{A_k}}[f]}{1 - \epsilon}, \tag{4.2}$$

where $\mathbb{E}_{\mu_{A_k}}$ denotes expectation with respect to μ_{A_k}. Examining the numerator we obtain:

$$\mathbb{E}_{\mu_{A_k}}[f] = \int_S f(\varphi^{\otimes N})\, d\mu_{A_k}(\varphi^{\otimes N}) \tag{4.3}$$

$$= \int_S f(\varphi^{\otimes N})\, \frac{x_{\varphi^{\otimes N}}(A_k) \cdot d\mu(\varphi^{\otimes N})}{x(A_k)} \tag{4.4}$$

$$= \frac{1}{x(A_k)} \cdot \int_S |\langle\varphi|\psi\rangle|^2 \cdot \varphi^{\otimes N}(A_k)\, d\mu(\varphi^{\otimes N}), \tag{4.5}$$

where the second equality follows from Bayes' formula and the third equality follows from QPP. Letting $\rho = |\varphi\rangle\langle\varphi|$ and $\rho_k = |\psi\rangle\langle\psi|$, we can re-express the integral term in (4.5) as $\left\langle \mathrm{Tr}[\rho\rho_k]\, \mathrm{Tr}[A_k\rho^{\otimes N}] \right\rangle$, where the angle brackets denote uniform average over the $\rho^{\otimes N}$. The sum of this term over k is known as the *mean fidelity* of a state estimation process which, given outcome k, produces the guess ρ_k. Hayashi et al. (2005, Equation 7) show that this term is bounded above by: $\frac{1}{d_{N+1}} \cdot \mathrm{Tr}[A_k]$, where $d_l = \binom{l+d-1}{d-1}$.[5] On the other hand, $x(A_k) = \int_S \omega(A_k)\, d\mu(\omega) = \frac{1}{d_N} \cdot \mathrm{Tr}[A_k]$, where the first equality follows from QPP, and the second equality follows from the fact that μ is concentrated on the totally symmetric subspace (cf. Hayashi et al. (2005, Equation 4)). Applying these to (4.5) we obtain

$$\mathbb{E}_{\mu_{A_k}}[f] \leq \frac{d_N}{\mathrm{Tr}[A_k]} \cdot \frac{\mathrm{Tr}[A_k]}{d_{N+1}} = \frac{d_N}{d_{N+1}} = \frac{N+1}{N+d}, \tag{4.6}$$

which together with (4.2) implies (4.1).

In this derivation, we drew on a finding from quantum information theory on the *fidelity bounds of quantum state estimation*. Indeed, our result can be viewed as simply situating this finding within a credence-based framework. (Note that this particular credence-based framework is different from that of Quantum Bayesianism. Quantum Bayesianism is an interpretation of quantum mechanics according to which quantum states only represent degrees of belief; it is not clear QBians would endorse QPP, or consider it non-vacuous (Fuchs and Schack, 2004).) One advantage of this framework is that at least in principle, it allows one to

[5] Here we take advantage of the result, in Hayashi et al.'s derivation, that $\mathrm{Tr}[A_k\rho_k(n)] \leq \mathrm{Tr}[A_k]$, so their Equation 7, focusing on one term k in the larger summand (they use the letter a), is less than or equal to $\frac{1}{(N+1)d_{N+1}} \cdot \mathrm{Tr}\,[A_k + NA_K]$.

consider a wider variety of learning situations, where the agent's prior is not necessarily uniform.

4.3 The argument of D'Ariano and Yuen (1996)

Another well-known argument against the possibility of individual state determination is based on the *no-cloning theorem*. The no-cloning theorem is a well-known 'no-go result' within quantum mechanics. Roughly speaking, it says it is impossible to have a process which, given some arbitrary state ψ as input, outputs a copy of that state:

$$\psi \otimes \nu \;\longrightarrow\; \psi \otimes \psi, \qquad \forall \psi \in \mathcal{H}. \tag{4.7}$$

The idea of the theorem is as follows. Suppose $\longrightarrow$ represents some unitary process U, i.e. it preserves inner products. Take any distinct ψ and ψ' which are non-orthogonal, so $|\langle\psi|\psi'\rangle|^2 < |\langle\psi|\psi'\rangle|$. Then if U satisfies the above, we have

$$|\langle\psi|\psi'\rangle| = |\langle\psi|\psi'\rangle| \cdot |\langle\nu|\nu\rangle| = |\langle\psi \otimes \nu|\psi' \otimes \nu\rangle| = |\langle U(\psi \otimes \nu)|U(\psi' \otimes \nu)\rangle| = |\langle\psi \otimes \psi|\psi' \otimes \psi'\rangle| = |\langle\psi|\psi'\rangle|^2,$$

which is a contradiction.

It is customary, upon introducing this theorem, to make the following remark (Wootters and Zurek, 1982; Scarani et al., 2005): if the cloning of quantum states were possible, then individual state determination would be possible. One would simply create many copies of the state, and then perform quantum state tomography: one would measure the copies along various bases, and then use the statistics to reconstruct the state.

As foreshadowed, authors have also made a claim in the converse direction (D'Ariano and Yuen, 1996; Vaidman, 2015): if it were possible to reliably determine the initial, previously unknown quantum state of a single system (with no other copies available), then it would be possible to clone in the sense forbidden by the original no-cloning theorem. Or, contrapositively: the no-cloning theorem rules out the possibility of individual state determination.

If both of these claims were true then, not only would we have an argument against the possibility of individual state determination that did not rely on assumptions about POVM measurements or Lüder's updating, we would have an attractively straightforward interpretation of the physical significance of the no-cloning theorem.

However, matters are not so simple. As I will now suggest, the converse claim is mistaken.

The argument for this converse claim, and its overgeneralization

Vaidman (2015, p. 18) argues for this converse claim in his discussion of protective measurement:

> If we are given a single quantum system in an unknown quantum state, there is no way to find out what this state is. It would contradict the no-cloning theorem: if we can find out what the state is, we can prepare many other systems in this state.

The argument traces back at least to D'Ariano and Yuen (1996), who spell out the reasoning in more detail (1996, pp. 2832–3):

> In this Letter, we will show the impossibility of any measurement scheme for determining the wave function from a single copy of the system [...] A quantum cloning machine [...] has to be represented by a state transformation of the general form,
>
> $$\nu \otimes \psi \otimes \underbrace{\nu_1 \otimes ... \otimes \nu_{M-1}}_{M>1} \longrightarrow \nu'(\psi) \otimes \underbrace{\psi \otimes \psi \otimes ... \otimes \psi}_{M \text{ times}}, \tag{4.8}$$
>
> where both the transformation and the state preparation must be independent of ψ—which *a priori* is unknown [...] [By the no-cloning theorem,] one cannot clone the state of a system known to be in *any one* of two nonorthogonal states, say, ψ and φ, using a single state-independent unitary transformation. Let us now suppose that there is a device that is capable of determining (by a unitary process) in which of these two states the system is. Then, once the state is known, one can easily choose unitary transformations that generate an M clone of ψ or φ, depending on the two possible results of this device. Hence, such a 'state detector' would lead to a realization of the cloning machine.

Here the equation (4.8) is a generalization of (4.7); by similar reasoning, one finds that if $\longrightarrow$ is unitary, then $|\langle\psi|\psi'\rangle| < |\langle\psi|\psi'\rangle|^M$ which contradicts the assumption that ψ and ψ' are non-orthogonal and distinct.

The problem with this argument is that, like that of Busch (1997), it overgeneralizes. If it were sound, it would imply that it is impossible to determine the state of a quantum system given a very large number of copies. To see this, note that state transformations of the following form, for any finite N,

$$\nu \otimes \overbrace{\psi \otimes ... \otimes \psi}^{N \text{ times}} \otimes \underbrace{\omega_1 \otimes ... \otimes \omega_{M-N}}_{M>N} \longrightarrow \nu'(\overbrace{\psi \otimes ... \otimes \psi}^{N \text{ times}}) \otimes \underbrace{\psi \otimes ... \otimes \psi \otimes \psi \otimes ... \otimes \psi}_{M \text{ times}}, \tag{4.9}$$

are also in violation of unitarity. (By similar reasoning to the above, we would have $|\langle\psi|\psi'\rangle|^N < |\langle\psi|\psi'\rangle|^M$

for $N < M$, which contradicts the assumption ψ and ψ' are distinct and non-orthogonal.[6]) Indeed it was perhaps misleading for the authors who introduced the no-cloning theorem, Wootters and Zurek (1982), to frame it as the result that "a single quantum cannot be cloned", given that a finite ensemble of identical quanta also cannot be cloned.

So, the argument overgeneralizes to imply a radical skeptical conclusion. But where exactly does the argument—now interpreted as an argument *for* this much more skeptical conclusion—go wrong?

Diagnosing the argument

Unlike the argument of Section 4.1, this argument is very difficult to diagnose. One thought is that, like Busch's argument, it is presuming *perfect* or *infallible* discrimination as a standard for successful state determination. But first note that we can replace ψ on the right-hand-side with states $\tilde{\psi}$ that are only approximately equal to ψ; as long as $\tilde{\psi}$ is somewhat distinct from $\tilde{\psi}'$, we will still get a violation of unitarity for a large enough choice of M. And second, note that we need not assume that (4.9) holds for every ψ, just for some sufficiently distinct non-orthogonal pair ψ and ψ'. So we can imagine a protocol where given N copies the agent guesses either ψ or ψ' and then prepares many copies of the guessed state. If the agent's guessing is reasonably reliable, we would expect there to be at least one situation where (4.9) holds for both ψ and ψ', and this is enough to generate the contradiction.[7]

A more compelling objection is that the argument seems to presume $\longrightarrow$ is a unitary map even though the agent may be performing measurements on the N copies. In particular the argument seems to presume:

Unitarity The quantum state of an isolated system (like the agent's entire lab) always evolves unitarily.

This assumption is controversial, and is rejected by interpretations like the Orthodox Interpretation and GRW theory. However, it would also be problematic if the argument implied that on interpretations that do endorse **Unitarity**, it is impossible to reliably determine the state of a system, even given a large number of copies of that system. With no other obvious flaws in the skeptical argument,[8] we are still left with the

[6] See Appendix A, Theorem 2 for details.

[7] Here I am thinking of reliability in a modal sense: there should be at least some iterations of the set-up in which the counterfactual 'if the state of the N systems were ψ' instead of ψ, then the agent would have guessed differently from ψ, and prepared many copies of the guessed state' is true. When evaluating this kind of counterfactual, I am assuming that everything in the set-up except for the state ψ is held fixed (so there is no backtracking), and the dynamical law $\longrightarrow$ is also held fixed. We could also consider a safety requirement, that the agent's belief in the state could not have easily been false.

In response, one might object that these kind of externalist modal requirements are not needed for a successful state determination. Indeed, in the Bayesian epistemological framework from Section 4.2, I did not mention any such requirements. However, I think even if we endorse a strong kind of evidentialism, serious skeptical problems would result if the unreliability argument were to go through. For example, rational agents presented this argument would, if the argument is valid, be compelled to either conclude that their devices are not reliable (which is higher-order evidence that should, arguably, lead them to reduce their confidence in their state hypotheses), or conclude that there is something wrong with the quantum mechanical model of their scenario. It is bad news for this kind of scientific theory if all sufficiently informed and rational agents are forced to either endorse a radical skepticism about their experimental abilities, or to reject the theory.

[8] Another thought is that (4.9) presumes the state of the original N systems will be undisturbed, in particular that the state of the N systems will remain the same after the agent's investigation. But we can easily adjust for this, by requiring that M

question of where it goes wrong.

I think that this argument is picking up on a deep and unnoticed tension in the foundations of quantum theory, between the claim that we can reliably determine quantum states, on the one hand, and the claim that we can *manipulate* or *control* quantum states, on the other hand. Indeed, note a key premise of the argument is that the agent's ability to prepare the state she wishes. In the next chapter, I will argue that in the presence of **Unitarity**, this control assumption conflicts with the possibility of any kind of reliable state determination. On my view, then, the physical significance of the no-cloning theorem is not in its implications for individual state determination. The no-cloning theorem does not reveal the impossibility of individual state determination,[9] but rather this more general conflict, which I call the *control problem*.

(rather than only $M - N$) copies of the guessed state are prepared:

$$\nu \otimes \overbrace{\psi \otimes ... \otimes \psi}^{N \text{ times}} \otimes \underbrace{\omega_1 \otimes ... \otimes \omega_M}_{M>N} \longrightarrow \nu'(\overbrace{\psi \otimes ... \otimes \psi}^{N \text{ times}}) \otimes \underbrace{\psi \otimes ... \otimes \psi}_{M \text{ times}},$$

where this time we let the ancillary term $\nu'(\psi \otimes ... \otimes \psi)$ absorb the final state of the original N copies.

A related worry is that it is unreasonable to assume, given **Unitarity**, that the M prepared copies are unentangled from the environment. But we can re-run the argument allowing that the M prepared systems may be in mixed states and entangled with their environment, and as long as they are unentangled with their other ensemble members (a basic requirement for any good preparation) the argument will go through; cf. the next chapter and Appendix A, Theorem 3 for details.

[9] Are there other cloning-based arguments for the impossibility of individual state determination that do not overgeneralize? While there are promising proposals in this direction, they involve going significantly beyond the scope of the no-cloning theorem, for instance by dropping the assumption that $\longrightarrow$ is a deterministic map from states to states, or by dropping the assumption that the cloned states are a separable ensemble. (One strategy is to use the limits of 'probabilistic cloning' (Duan and Guo, 1998a) to derive limits on reliable (yet fallible) state determination as a function of N. Another strategy is to use the limits of 'approximate' or 'imperfect' cloning (Werner, 1998) to do the same thing. The latter has essentially been carried out, by Bae and Acín (2006), who establishes a connection to the *mean fidelity bounds* discussed in the previous section. As these authors note, one must allow a high degree of entanglement between the outputs, which is not the case in the original no-cloning theorem.)

Chapter 5

A new problem for quantum mechanics

There is a compelling argument that it is impossible to reconcile Schrödinger's quantum theory with the determinateness of measurement outcomes, unless we give up on the claim that the theory provides us with a complete description of physical systems. This is what I will call 'the standard measurement problem'. More precisely, the problem is that the following assumptions, though individually innocent, are jointly untenable (Maudlin, 1995):

(A1) The quantum state of a system determines, directly or indirectly, all of its physical properties [**Completeness**].

(A2) The quantum state of an isolated system always evolves in accord with a linear dynamical equation, such as the Schrödinger equation [**Linearity**].

(A3) Given determinate inputs, our measuring devices always produce unique, determinate outcomes [**Determinate Outcome**].

One desideratum for an interpretation of quantum mechanics is that it solve the standard measurement problem. An interpretation must reject at least one of the assumptions listed, and explain how the resulting gap is to be filled in—preferably, without introducing 'measurement' as a basic, unanalyzed term in the dynamical laws.[1] Bohmian mechanics, for instance, rejects (A1), GRW theory rejects (A2), and Everettian

[1] Authors often use the term 'measurement problem' to refer specifically to this deficiency of textbook collapse-on-measurement accounts. That is not how I am using the term here, but there is no real disagreement behind this difference; all agree that an account can reject one of (A1)–(A3), in this case (A2), and yet fail to be an acceptable fundamental physical theory for various reasons. For more on this terminological issue, see Lewis (2007, pp. 750–1). For other presentations of the measurement problem similar to Maudlin's above, see e.g. Albert (1992), Wallace (2012), Healey (2017), and Myrvold (2018).

mechanics rejects (A3).

I argue that there is another problem for quantum mechanics, distinct from the standard measurement problem. The characteristic feature of this new problem, which I call the control problem, is that it introduces considerations of state preparation on top of considerations of measurement. In particular, the problem is that the following claims are jointly incompatible:

(B1) We can successfully prepare quantum states: at least some of our preparation devices are such that, if determinately fed many inputs, they output a non-trivial fraction of those inputs in some specified range of quantum states [**Preparation**] (here the 'inputs' are subsystems, and we define 'the quantum state of a subsystem' in the standard way, as its reduced state).

(B2) The quantum state of an isolated system always evolves in accord with a deterministic dynamical equation that preserves the inner product, such as the Schrödinger equation [**Unitarity**].

(B3) It is always determinate whether or not a subsystem has been input into a given (measuring or preparation) device [**Determinate Input**].

In fact, there is another background assumption needed to generate the incompatibility, which is also taken for granted in the standard measurement problem:

(A4, B4) There is always, from a first-person view, an experienced 'observed outcome' of a measurement, even if—from a 'God's eye view'—there is no unique determinate outcome. Moreover, our measuring devices are at least somewhat informative and reliable. For example, if a spin-z measuring device is determinately fed an electron in the eigenstate $\uparrow_z$, it will output 'UP' (and not 'DOWN') [**Competent Measurement**].

This assumption would be disastrous to deny, which is why it is background.

Another desideratum for a quantum theory is that it solve the control problem. A theory must reject at least one of the assumptions listed above and explain how the resulting gap is to be filled in—preferably, without introducing 'measurement' or 'preparation' as basic, unanalyzed terms in the dynamical laws. GRW theory rejects (B2). Everettian mechanics rejects (B3). Bohmian mechanics, it turns out, rejects (B1) in favor of:

(B1') Given determinate inputs, our preparation devices can prepare those inputs in desired conditional quantum states (relativized to an extra parameter) but not desired quantum states simpliciter [**Conditional Preparation**].

Why are (B1)–(B4) incompatible? Briefly, the idea is to consider a certain type of preparation protocol, in which one chooses which quantum states to prepare based on which outcome of some initial measurement one (first-personally) observes. It is straightforward to show that if the assumptions were all true, then these protocols would violate a variant of the no-cloning theorem (Wootters and Zurek, 1982) and so would be impossible to implement. But this cannot be right; we implement such protocols all the time. Therefore, at least one of the assumptions must be false.

The main difference between the control problem and the standard measurement problem is the difference between (A1) and (B1). Whereas the standard measurement problem makes an assumption about what quantum states are (what physical properties they determine), the control problem makes an assumption about what we can do with quantum states—about the extent to which we can control the quantum states of the subsystems that we study in the laboratory. It is rare to find, for any interpretation or theory, detailed discussion of preparation and its limits. My hope is that the control problem will motivate such discussion.

Is the control problem just a 'non-standard' version of the measurement problem? On the one hand, the measurement problem is often glossed as a problem with interpreting quantum mechanics as a complete physical theory, and most discussions involve some sort of state completeness assumption, suggesting **Completeness** is a defining feature of the problem. If that is the case, then solving the measurement problem does not entail solving the control problem; consider, for example, a theory which denies that quantum states specify macro-properties, like the readings on measuring devices, but maintains that we can prepare certain micro-systems in certain quantum states.

On the other hand, the control problem invokes **Competent Measurement**, and crucially involves a measurement-type interaction (combined with a preparation). So if the measurement problem is defined as the general puzzle that measurement interactions can be problematic under Schrödinger dynamics, then the control problem should be viewed as a specific aspect of it.

Of course, to some extent this question of how broadly to define the measurement problem is a terminological one. What matters, for our purposes, is whether the control problem is of interest to the foundations of quantum mechanics. In this chapter, I aim to show that it is.

The plan is as follows. In Section 5.1, I review the standard measurement problem. In Section 5.2, I clarify **Preparation** and discuss whether it is already in tension with **Unitarity**. It turns out that this question has been raised under the header of 'the preparation problem' (Ruetsche, 1996; Wessels, 1997), so I take the opportunity to review this problem as well. In Section 5.3, I present the main result of the chapter, the incompatibility argument for the control problem. I also compare the control problem with some non-standard versions of the measurement problem that have been raised (Maudlin, 1995; Gao, 2019). In Section 5.4, I turn to consider how different interpretations may respond to the control problem. The

treatment here is necessarily brief, but some basic direction is given. Finally, in Section 5.5, I tie the control problem back to the issue of individual state determination raised in the previous chapter.

5.1 The standard measurement problem

I begin by reviewing why (A1)–(A4) are jointly untenable. This will set the stage for my own incompatibility argument for (B1)–(B4), which has a similar set-up and structure.

First, the set-up. Zarna is a physicist who specializes in quantum mechanics. Her laboratory is set up as shown in Figure 5.1. She has a collection of electrons, and a measuring device which takes those electrons as input and measures their spin in the z direction. We suppose, without loss of generality, that the lab is isolated, and so **Linearity** is applicable.

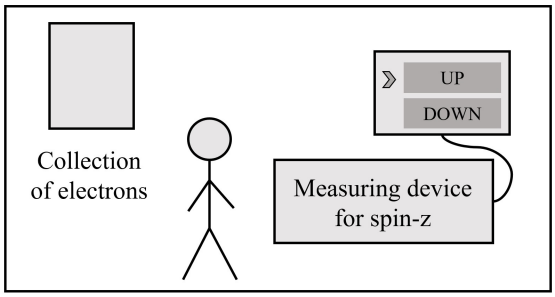


Figure 5.1: Set-up of Zarna's laboratory.

Suppose for *reductio* that (A1)–(A4) are all true. Suppose Zarna feeds an $\uparrow_z$ electron from her collection into her measuring device. By **Competent Measurement** the device will read 'UP'. By **Completeness**, the final quantum state of the lab will specify this fact. So, letting $\longrightarrow$ denote the evolution of the quantum state of the entire lab from the initial to final time, and letting 'initial' denote the initial quantum state of the rest of the lab, we have,

$$\uparrow_z \otimes \text{initial} \quad \longrightarrow \quad \text{device reads 'UP'}, \tag{5.1}$$

where the final quantum state of the lab is written 'device reads "UP"' to indicate any state which agrees on this specification of the device. (We stay neutral, for example, on the final state of the measured electron, since this was not specified by our assumptions.)

Next, we ask how the quantum state of the lab would have evolved if the electron were in the $\downarrow_z$ spin

state instead. By the same reasoning,

$$\downarrow_z \otimes \text{initial} \;\longrightarrow\; \text{device reads 'DOWN'}. \tag{5.2}$$

Finally, we ask how the quantum state of the lab would have evolved if the electron were $\uparrow_x$ instead. Since $\uparrow_x = \frac{1}{\sqrt{2}}\uparrow_z + \frac{1}{\sqrt{2}}\downarrow_z$, (5.1)–(5.2) and **Linearity** imply the following answer:

$$\uparrow_x \otimes \text{initial} \longrightarrow \frac{1}{\sqrt{2}}\,(\text{device reads 'UP'}) + \frac{1}{\sqrt{2}}\,(\text{device reads 'DOWN'}). \tag{5.3}$$

However, consider any possibility in which the right-hand side of (5.3) is the final quantum state of the lab. By **Completeness**, this quantum state must specify the reading on the measuring device, if there is one. And now we have a problem. As Maudlin (1995, p. 8) puts it:

> Since 'UP' and 'DOWN' enter symmetrically into the final state, by what argument could one attempt to show that this device is, in fact, in exactly one of the two indicator states?

So the only plausible option is to insist that, in all these possibilities, there is not a determinate fact of the matter as to what the measuring device reads (and so there is no failure of **Completeness**). But this contradicts **Determinate Outcome**.

As Maudlin acknowledges, "symmetry arguments [like this one] are not a matter of logic" (1995, p. 10) and so this is not a strict incompatibility. Nevertheless, the argument is highly illustrative. Decades of effort to convincingly remove the inconsistency by denying the symmetry, or by employing more detailed models of the measurement process, have not succeeded.

5.2 Quantum state preparation

Like measurement, preparation plays a crucial role in our experimental practice. Countless quantum mechanical experiments, including many foundational Gedanken-experiments, involve steps of the form 'prepare an ensemble of systems in such-and-such quantum state'.

The function of preparation, however, is rather distinct from that of measurement. In particular, it would be misleading to think of preparation as a special case of measurement. As Margenau (1963, p. 6) writes:

> In general, a preparation of state is any physical operation which assures that, if a system has been subjected to the operation, it will afterward be in a specifiable quantum state (ψ, or perhaps a mixture of ψ's). A measurement, on the other hand, [is] a physical operation which yields a number [an outcome] that refers to the state present before the operation.

Here "assures" may be too strong, and we may not wish to require that the operation uniformly prepares the same ψ (or mixture of ψ's) each time. But the basic contrast is clear, and can be restated in terms of devices. Roughly speaking, a preparation device is an object or physical set-up which, when fed the right sort of subsystem as input, subjects it to a preparation operation. Unlike measuring devices, there is no general requirement that a preparation device display an outcome, or that its final state be correlated with some aspect (initial quantum state, observable) of the input system. Indeed, at least for the purposes of this chapter, all we will assume of a preparation device is that it satisfy the conditions of **Preparation**: it occasionally succeeds in outputting its inputs in some pre-specified range of quantum states. (We do require that this range is not trivial. In particular, in the next section we will take **Preparation** to imply that there are at least two 'distinct' preparation devices, in the sense that their ranges do not significantly overlap.)

While preparation is not measurement, the issue remains that the success of certain preparation procedures in the lab looks incompatible with the Schrödinger dynamics. This issue has become known as 'the preparation problem' (Ruetsche 1996; Wessels 1997). I discuss the issue in some detail now. An important upshot is that **Preparation** is nevertheless perfectly compatible with **Unitarity** and **Linearity**. Readers who therefore wish to move on to the incompatibility (B1)–(B4) may skip directly to Section 5.3. That said, the preparation problem is of independent interest, and provides valuable context for the result that follows.

The preparation problem

Consider the (apparently successful) preparation procedure depicted in Figure 5.2. Zarna feeds a beam of electrons through a z-oriented Stern-Gerlach apparatus that has its lower path blocked. She then goes on to perform experiments using the upper beam.

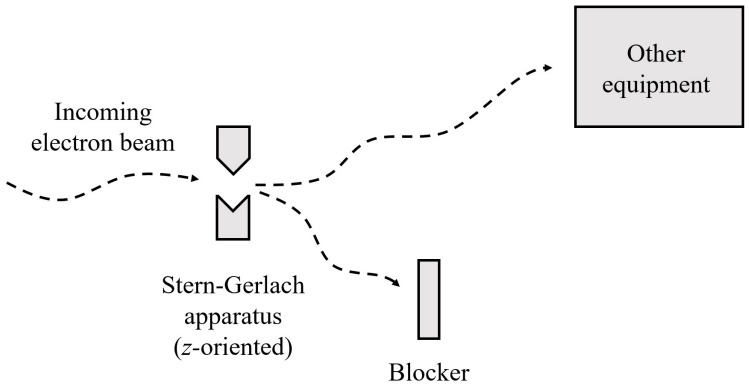


Figure 5.2: Purported preparation procedure for $\uparrow_z$.

A quick analysis shows that, if **Linearity** is true, then almost none of Zarna's electrons end up in a spin

state close to $\uparrow_z$. Rather, almost all end up in some distinct (mixed) spin state ρ which varies from electron to electron, depending on the electron's initial spin state $\omega_0 = \alpha \cdot \uparrow_z + \beta \cdot \downarrow_z$.[2] Intuitively, this is because **Linearity** forbids simply 'throwing out' the $\downarrow_z$ component (the β part of the superposition) just because it corresponds to hitting the blocker. This component of the superposition, suitably evolved, must also be considered when determining the final spin state of each electron.

At first glance, this result looks disastrous for **Linearity**. If this device does not really prepare $\uparrow_z$, then what explains why, when we measure the spin-z of its outputs, we obtain measurement statistics that correspond to $\uparrow_z$? For example, what explains why, when Zarna places another z-oriented Stern-Gerlach apparatus as well as a scintillation screen along the upper path, she always obtains the outcome 'UP' at this screen?[3]

The result is not as disastrous as it appears. As Wessels argues (1997, pp. 264–270), one can give a relatively interpretation-neutral story for why we should expect to observe the statistics of the filtered state even though (assuming **Linearity**) each subsystem is actually in an entangled, mixed state. The approach is piecemeal and involves taking into account how the measuring device works in each case. In the Stern-Gerlach case, note that the second Stern-Gerlach apparatus and scintillation screen are spatially localized to the upper beam, far from the blocker. So, while one should expect a position reading at the screen to see contributions from any components of the electron's spatial wave function that are localized to this upper region, one should not expect to see contributions from any components that are localized to the lower region next the blocker. (We assume, as seems reasonable, that the interaction between the electron and scintillation screen is local.) And precisely the effect of the first Stern-Gerlach apparatus was to correlate these upper and lower spatial components with the $\uparrow_z$ and $\downarrow_z$ spin components respectively.

This account at least temporarily rescues **Linearity** from making false predictions.[4] But it raises two more questions. First, if these particular preparation procedures do not really prepare their associated quantum states under the Schrödinger dynamics, are there any procedures which do? And second, if the above preparation procedures still ensure we obtain the measurement statistics corresponding to the desired quantum state $\uparrow_z$ (for example), then why should we be interested in whether they "really prepare" $\uparrow_z$?

The answer to the first question, as foreshadowed, is yes. **Linearity** is compatible with **Preparation**, and so is **Unitarity**. This is thanks to the lack of constraints on the final state of a preparation device. For

[2] Namely, $\rho \approx |\alpha|^2 \cdot |\uparrow_z\rangle \langle\uparrow_z| + |\beta|^2 \cdot \sigma$ where σ depends on how the blocker entangles with the electron's spin degree of freedom; if it doesn't, $\sigma = |\downarrow_z\rangle \langle\downarrow_z|$. See Wessels (1997, p. 254) for details.

[3] The worry is not restricted to this Stern-Gerlach case, but extends to other uses of 'filters' such as the double slit experiment (Wessels, 1997, p. 255). Unsurprisingly, it's common in the literature to model these filter procedures with a projection-type, non-unitary operation; see for example Busch et al. (1996, Chapter III.9) and Ruetsche and Earman (2011, Section 2.2).

[4] It is possible that we will not be able to offer such a story for every case (see previous footnote), leaving some situations where the measurement statistics must be explained on an interpretation-by-interpretation basis. See also the problem of effect (Section 5.3).

instance, take the ideal case of a 'uniform' preparation which, given an input in any initial quantum state ω_0, will output it in the same desired quantum state ψ:

$$\nu \otimes \omega_0 \quad \longrightarrow \quad \nu'(\omega_0) \otimes \psi, \tag{5.4}$$

for all ω_0, where ν is the fixed initial quantum state of the device and ν' is its final quantum state, written $\nu'(\omega_0)$ to indicate that it varies with ω_0. Thanks to the lack of constraints on ν', there do exist $\longrightarrow$ which satisfy both (5.4) and **Unitarity** (Wessels, 1997, pp. 249–250). Thus even this ideal kind of preparation is compatible, in principle, with **Unitarity**. And though they are not ideal, there are several laboratory procedures of this rough form. A simple example, mentioned in Chapter 3, is the placement of a crystal in a cold vacuum, to cool it to its ground energy state.[5] In fact this is only one of a wide class of 'environment-induced' preparations, a general technique in which one first unitarily tunes the dynamics of the system-plus-environment so that the desired state-to-be-prepared is the ground state, and then lets the system relax for sufficient time (Fröhlich and Schubnel, 2016). It is also crucial to point out that even though it is inconsistent with **Unitarity** that the 'filter' procedures discussed above really prepare the quantum states we associate with them (like $\uparrow_z$), it is not inconsistent with **Unitarity** that they prepare a mixed state ρ, where ρ varies non-uniformly with ω_0. And so, in light of these observations, I conclude that **Unitarity** and **Preparation** are perfectly compatible. Schrödinger's theory does not face the same general problem for preparations as it does for measurements. As Wessels (1997) writes:

> No analogue of [the measurement problem] arises when quantum mechanics is applied to preparations [...] The preparation problem is not a foundational problem.

Of course, what I will argue next is that while **Unitarity** and **Preparation** are not incompatible taken on their own, they are incompatible when taken in conjunction with **Determinate Input** and **Competent Measurement**. This is the control problem.

Next we turn to the second question. One reason we might be interested, at least from a foundational perspective, in whether we "really prepare" $\uparrow_z$ is that we are interested in our control over quantum states themselves, not just our control over which patterns of outcomes will likely obtain when we make certain measurements. That being said, there may be a worry in the vicinity of this question: assuming we can reject **Preparation** without contradicting our empirical predictions (about the post-preparation statistics we observe), is the control problem, which pits **Preparation** against some other assumptions, really a "problem"?

[5] See Park and Band (1992, Section 3). A similar procedure can be used to prepare, say, a collection of rubidium atoms in a Bose-Einstein condensate (Healey, 2017, Chapter 5).

In response to this worry, note that we can reject **Completeness** without contradicting our empirical predictions about measurement statistics. Yet, we do not say that the standard measurement problem is not really a problem. Indeed, just as rejecting **Completeness** requires a systematic account of what the quantum state is missing, and how to make sense of our usual talk of the quantum state as complete, rejecting **Preparation** requires a systematic account of our control over quantum states, and how to recover or translate our usual talk of preparing states. (We will see, for instance, that Bohmian mechanics accomplishes this using the notion of a conditional quantum state.) At least from the viewpoint of a theorist who is starting out with the naïve Schrödinger account, these are not trivial tasks.

5.3 The control problem

Set-up

Zarna has installed an additional room in her laboratory, as shown in Figure 5.3. The room contains two preparation devices, D and D', as well as a large collection of subsystems. The devices are designed to take these subsystems as input. We assume that the devices are distinct, that is, tend to prepare non-negligibly different quantum states, and can be used repeatedly, that is, can be fed many inputs if desired. Some examples of such devices are provided below.

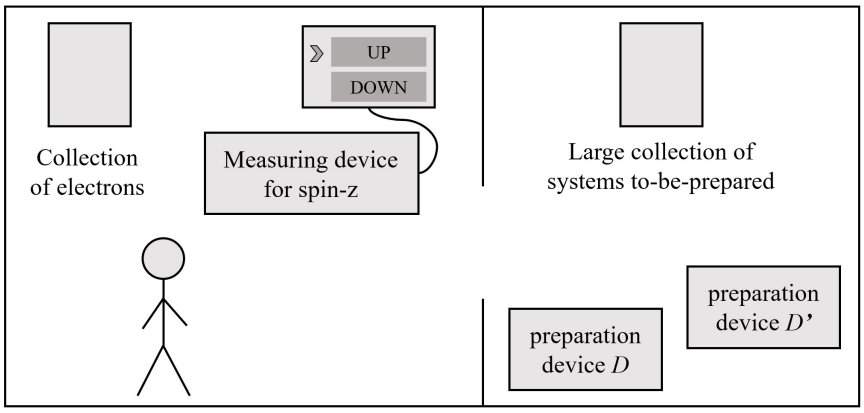


Figure 5.3: The new set-up of Zarna's laboratory.

Example 1. The subsystems to-be-prepared are also electrons. D is a filtration which prepares about half of its inputs in the $\uparrow_z$ spin state. D' is a distinct filtration which prepares about half of its inputs in $\uparrow_x$. As we discussed in Section 5.2, it is not obvious how to implement these filtrations under **Unitarity**. But we include the example, since D and D' do satisfy our criteria of being sufficiently distinct and reusable or repeatable.

Example 2. The subsystems to-be-prepared are crystals. D is an array of vacuums each enclosed in an absorbing wall and cooled as much as possible. They prepare their inputs in a ground energy state ψ or something near to ψ. D' is an array of containers which are heated to prepare their inputs in much more energetic states, so that many or most of their final states ψ' are different from ψ. Here ψ' may vary as long as it is usually distinct from ψ.

Example 3. The subsystems are electrons. D is a Stern-Gerlach apparatus oriented in the z direction and D' is a Stern-Gerlach apparatus oriented in the x direction. As discussed in Section 5.2, this may not be equivalent to Example 1 under **Unitarity**, even if Zarna blocks both the lower paths. Nevertheless, D and D' satisfy our criteria. If an electron is fed into D, it will usually end up in a spin state ρ very different from the ρ' that would have resulted if it were fed into D'.[6] Also, the devices can be used repeatedly. (An important aspect of this repeatability assumption, which I have not made explicit, is that the repetition does not introduce unwanted entanglement between the prepared systems. This holds automatically if the prepared state is pure as in Examples 1 and 2, and is taken for granted as a property of any good ensemble preparation.[7] However, given the complications of Section 5.2, it is worth flagging.[8])

Incompatibility argument

Suppose for *reductio* that all four assumptions (B1)–(B4) are true. Note that, given the above set-up, it should be physically possible for Zarna to implement the following protocol:

> Zarna's protocol: I will feed an electron from my collection into my measuring device. If I observe 'UP',[9] I will walk into the other room (the room containing the preparation devices D and D'), take all the subsystems, and feed them into D. Otherwise, if I do not observe 'UP', I will walk into the other room, take all the subsystems, and feed them into D' instead.

Suppose that, as a matter of fact, Zarna measures an electron which is in the $\uparrow_z$ spin state. We ask how the quantum state of her lab will evolve from this initial time. By **Competent Measurement**, Zarna will observe 'UP'. So, by the protocol, she will end up feeding all the subsystems into D. By **Preparation**,

[6] Continuing footnote 2, the z-oriented apparatus outputs $\rho \approx |\alpha|^2 |\uparrow_z\rangle \langle\uparrow_z| + |\beta|^2 |\downarrow_z\rangle \langle\downarrow_z|$ while the x-oriented apparatus outputs $\rho' \approx (1/2)(|\alpha + \beta|^2) |\uparrow_x\rangle \langle\uparrow_x| + (1/2)(|\alpha - \beta|^2) |\downarrow_x\rangle \langle\downarrow_x|$.

[7] Since, usually, the purpose of preparing many systems is so that they provide the basis for multiple independent trials or experiments. Quantum state tomography, for instance, would break down without this assumption (Paris and Rehacek, 2004).

[8] Some remarks: First, we only require that the prepared subsystems are unentangled with each other; entanglement with the environment or the preparation device is allowed. Second, this unentanglement need only be in the degree of freedom being prepared, so for example if spin is being prepared, then entanglement in the spatial wave function is okay. And third, the prepared systems need only be approximately in a product state; perfect unentanglement is not required. (For the subsequent argument, we only require that having more members of the prepared ensemble makes its state more distinguishable. This is true even when the members are not perfectly unentangled.)

[9] There is (**Competent Measurement**) always, from a first-person view, an experienced observed outcome of a measurement, even if, from a third-person view, there is no unique determinate outcome.

many of these systems will end up, at the final time, in quantum states in the range associated with D. Let us call quantum states which fall within this specified range 'D-states'. Then the evolution of the quantum state of the lab will look like,

$$\uparrow_z \otimes \text{initial} \quad \longrightarrow \quad \text{many systems in } D\text{-states}, \tag{5.5}$$

where 'initial' is the initial quantum state of the rest of the lab in this particular instance.

Next, we ask how the quantum state of the lab would have evolved if the electron were $\downarrow_z$. By **Competent Measurement**, Zarna would not observe 'UP'. So, by the protocol, she would feed all her subsystems into D' instead of D. Again applying **Preparation**, we have:

$$\downarrow_z \otimes \text{initial} \quad \longrightarrow \quad \text{many systems in } D'\text{-states}. \tag{5.6}$$

(It is crucial that the right-hand sides of (5.5) and (5.6) are not misread as specifying some property of the preparation devices D or D'; they are simply specifying the final quantum states of Zarna's subsystems. Importantly, there is no illicit reliance on **Completeness** here.)

Finally, we ask how the quantum state of the lab would have evolved if the electron were in the $\uparrow_x$ spin state instead. By **Unitarity**, this question,

$$\uparrow_x \otimes \text{initial} \quad \longrightarrow \quad \boxed{\qquad ? \qquad}, \tag{5.7}$$

is well-defined and must have a single definite answer (since $\longrightarrow$ is deterministic, that is, a function from states to states, and '$\uparrow_x \otimes$ initial' is a possible state). Notice that this answer must take one of the following two forms:

$$\boxed{\text{many systems in } D\text{-states}}, \tag{a}$$

or

$$\boxed{\text{many systems in } D'\text{-states}}. \tag{b}$$

To see why, consider any possibility in which Zarna follows her protocol, and suppose toward contradiction that the final quantum state of the lab takes neither of these forms in that possibility. By **Determinate Input**, it cannot be indeterminate whether Zarna feeds her subsystems into a given preparation device at the end of her measurement. Now, if she did not feed the subsystems into either D or D', then she did not follow the protocol, contrary to our assumption (the protocol always ends with her using one of D or D').

If she fed the subsystems into D, then by **Preparation** the final quantum state of the lab should indeed have form (a), contrary to our assumption. And if she fed the subsystems into D', then by **Preparation** the final quantum state of the lab should indeed have form (b), contrary to our assumption. In all cases we reach a contradiction, as claimed.[10]

Now suppose, on the one hand, that the answer is of form (a). Substituting (a) into (5.7), and putting it side by side with (5.6), we have:

$$\downarrow_z \otimes \text{initial} \longrightarrow \text{many systems in } D'\text{-states}, \tag{5.6}$$

$$\uparrow_x \otimes \text{initial} \longrightarrow \text{many systems in } D\text{-states}. \tag{5.7a}$$

Suppose, on the other hand, that the answer is of form (b). Substituting (b) into (5.7), and putting it side by side with (5.5), we have:

$$\uparrow_z \otimes \text{initial} \longrightarrow \text{many systems in } D\text{-states}, \tag{5.5}$$

$$\uparrow_x \otimes \text{initial} \longrightarrow \text{many systems in } D'\text{-states}. \tag{5.7b}$$

The problem is that both answers lead to a violation of **Unitarity**. I give the argument for (5.5) and (5.7b); the reasoning for (5.6) and (5.7a) is the same.

According to **Unitarity**, the inner product of the left-hand sides of (5.5) and (5.7b) must equal the inner product of the right-hand sides. The basic observation is that since many D and D'-states were prepared, and many of those states are very different, the right-hand sides will actually be more orthogonal (i.e. more easily distinguishable, i.e. inner product closer to 0) than the left-hand sides. We illustrate this observation using our earlier examples.

Example 1 (continued)**.** Recall that D prepares about half of the input electrons in $\uparrow_z$ and D' prepares about half of the input electrons in $\uparrow_x$. Plausibly, there will be some subset of M electrons that were successfully prepared in both cases, allowing us to rewrite (5.5) and (5.7b) as follows:

$$\uparrow_z \otimes \text{initial} \longrightarrow \text{final} \otimes \overbrace{\uparrow_z \otimes \uparrow_z \otimes \ldots \otimes \uparrow_z}^{M \text{ times}}, \tag{5.5-1}$$

$$\uparrow_x \otimes \text{initial} \longrightarrow \text{final}' \otimes \uparrow_x \otimes \uparrow_x \otimes \ldots \otimes \uparrow_x, \tag{5.7b-1}$$

[10] Since we do not assume **Completeness**, we do not assume a general one-to-one correspondence between determinate states of affairs of the lab (this or that device being used) and its quantum state. However, we do assume that if two states of affairs involve the determinate use of distinct *preparation* devices, then they have different final quantum states; precisely what distinct preparations do is prepare subsystems in distinct quantum states.

where 'final' and 'final′' denote the final quantum state of everything in the lab except the (spin of the) successfully prepared electrons. **Unitarity** implies,

$$|\langle\uparrow_z|\uparrow_x\rangle|^2 \cdot |\langle \text{initial} \,|\, \text{initial}\rangle|^2 = |\langle \text{final} \,|\, \text{final'}\rangle|^2 \cdot |\langle\uparrow_z|\uparrow_x\rangle|^{2\cdot M}, \tag{5.8}$$

which, since $|\langle \text{initial} \,|\, \text{initial}\rangle|^2 = 1$ and $|\langle \text{final} \,|\, \text{final'}\rangle|^2 \leq 1$, implies,

$$|\langle\uparrow_z|\uparrow_x\rangle|^2 \leq |\langle\uparrow_z|\uparrow_x\rangle|^{2\cdot M}. \tag{5.9}$$

However, this clearly fails for any $M > 1$, since $|\langle\uparrow_z|\uparrow_x\rangle|^2 = \frac{1}{2}$ and $\frac{1}{2} > \left(\frac{1}{2}\right)^M$ for all $M > 1$. (If we suppose $\uparrow_z$ and $\uparrow_x$ are only approximately prepared, a modified argument still applies.[11]) Recall there was nothing to stop us from stipulating Zarna's collection of electrons to-be-prepared to be as large as we want, suggesting there are no principled limits on M. As foreshadowed, this derivation is a variant of the no-cloning theorem; for a general statement of the theorem and this variant, see Theorems 1 and 2 of Appendix A.

Example 2 (continued)**.** Similar reasoning applies. There will be some subset of M crystals which were successfully prepared in both cases:

$$\uparrow_z \otimes \text{initial} \;\longrightarrow\; \text{final} \otimes \psi_1 \otimes \psi_2 \otimes \ldots \otimes \psi_M, \tag{5.5-2}$$

$$\uparrow_x \otimes \text{initial} \;\longrightarrow\; \text{final}' \otimes \psi'_1 \otimes \psi'_2 \otimes \ldots \otimes \psi'_M, \tag{5.7b-2}$$

where we add subscripts to allow for variation from case to case, the requirement just being that ψ_i and ψ'_i are distinct so that $|\langle\psi_i|\psi'_i\rangle|^2 < 1$ for all i. For instance if ψ_1 is the ground energy eigenstate and ψ'_1 is some other energy eigenstate then $|\langle\psi_1|\psi'_1\rangle|^2 \approx 0$ which is already enough to violate **Unitarity**, since again it follows from **Unitarity** that we must have

$$|\langle\uparrow_z|\uparrow_x\rangle|^2 \leq |\langle\psi_1|\psi'_1\rangle|^2 \cdot |\langle\psi_2|\psi'_2\rangle|^2 \cdot \ldots \cdot |\langle\psi_M|\psi'_M\rangle|^2. \tag{5.10}$$

Otherwise we can always stipulate that Zarna started out with more systems, so that M is plausibly high enough to violate (5.10). For the variant of the no-cloning theorem that abstracts this result, see Theorem 2(iii) of Appendix A.

Example 3 (continued)**.** To model Example 3 we need the formal apparatus of density matrices. Since the

[11] As long as $|\langle\tilde{\uparrow_z}|\tilde{\uparrow_x}\rangle|^2$ is less than 1, there will be some M above which the inequality is broken. For example, suppose we define an error term ϵ such that $|\langle\tilde{\uparrow_z}|\tilde{\uparrow_x}\rangle|^2 = 0.5 \pm \epsilon$. Fix $\epsilon < 0.3$. Then for any $M > 25$, $|\langle\tilde{\uparrow_z}|\tilde{\uparrow_x}\rangle|^{2\cdot M}$ will be at least two degrees of magnitude smaller than $|\langle\uparrow_z|\uparrow_x\rangle|^2$. (Note that if $\tilde{\uparrow_z}$ and $\tilde{\uparrow_x}$ are mixed states we rewrite these expressions in terms of the fidelity measure; see the Appendix. We assume the M systems are still approximately a product; see footnote 8.)

upshot is the same, I relegate this more technical discussion to Appendix A, in particular to Theorem 3, which generalizes the above variants of the no-cloning theorem to mixed states.

Does the incompatibility still hold if we replace **Unitarity** by **Linearity**? The natural strategy is to argue that the answer to (5.7),

$$\boxed{\frac{1}{\sqrt{2}}(\text{many systems in } D\text{-states}) + \frac{1}{\sqrt{2}}(\text{many systems in } D'\text{-states})}, \tag{c}$$

takes neither form (a) nor (b), contrary to **Determinate Input & Preparation**. However this reasoning may fail if the preparation devices are not assumed to work perfectly.[12] Of course, this isn't to say that **Linearity** is not in tension with the other assumptions—just that the problem is more forceful when formulated in terms of the other major feature of the Schrödinger dynamics, namely its inner product-preserving property.

Other measurement problems

As I foreshadowed, the control problem crucially involves a measurement-type interaction, and so in this basic sense is a type of measurement problem. It is worth investigating, however, whether it shares any deeper similarities (or differences) with other existing versions of the measurement problem.

As one might expect, there are many versions of the measurement problem. In fact, immediately after presenting the standard measurement problem, Maudlin (1995) presents two others. The first, the problem of statistics, replaces (A2) and (A3) with:

(A2-S) the quantum state of an isolated system evolves deterministically;

(A3-S) two measurement situations which begin with the same initial quantum state can sometimes have different outcomes (where the probability of an outcome is given approximately by Born's rule).

This time (A4) is not needed; the **Completeness** assumption, (A1), does the rest of the work, since it rules out the strategy of distinguishing the two measurement situations using a 'hidden variable' detector state. Here the contrast with the control problem is arguably more stark than before. (A3-S) and (B3) are less similar, and it's not clear there is any analogous reformulation (B3-S) which would bring the assumptions in parallel.[13] And, most importantly, there remains the central difference between (A1) and (B1).

[12] Consider Example 1. Suppose that whenever D fails to prepare $\uparrow_z$ (about 50% of the time), it outputs $\uparrow_x$ instead. Then (c) could actually be rewritten in form (b), since we could factor out many $\uparrow_x$ states from both terms. While **Unitarity** would still be violated in this case (since there would still be many electrons which D prepares in $\uparrow_z$ but D' prepares in $\uparrow_x$), **Linearity** would not.

[13] There is, though, a variant of the preparation problem which pits determinism (A2-S) directly against the following analogue of (AS-3): some of our 'filter' preparations for $\uparrow_z$ begin with the same initial quantum state and yet, depending on the trial, do or do not prepare $\uparrow_z$.

The other problem Maudlin presents, the problem of effect, has a deeper connection to the control problem (1995, p. 13):

> If the result of the first measurement [in spin-z of an $\uparrow_x$ particle] is 'UP', a second measurement will certainly yield another 'UP', and similarly for 'DOWN'. The result of a measurement therefore has predictive power for the future [...] Any theory which seeks to replicate the empirical content of the traditional theory should have this feature. Let us call this the problem of effect [...].

To solve the problem of effect, a theory must explain why measurement results are repeatable. Strictly speaking, this problem is not about state preparation, since what is being discussed is the determinate record ('UP', 'UP', ...) associated with a measuring device (a record which, without **Completeness**, might not be encoded in any quantum state). However, it is tempting to infer from the record a claim about quantum state preparation, namely that the measuring device has genuinely prepared $\uparrow_z$. And then one quickly runs into trouble, since this claim contradicts **Unitarity**.[14] In fact we have just circled back to the preparation problem, that certain 'filter' procedures do not really prepare their associated states under **Unitarity**, even though they seem (judging by the subsequent measurement statistics) to have succeeded. For discussion of the preparation problem, see Section 5.2.

Finally we consider a more recent formulation of the measurement problem, due to Gao (2019). Like some other presentations (Albert 1992, pp. 77–79; Barrett 1999, Chapter 9), this 'mentalistic' reformulation emphasizes the tension between **Linearity** and the claim that observers have unique, determinate experiences. It gets off the ground by reformulating (5.1) and (5.2) as follows,

$$\uparrow_z \otimes \text{initial} \quad \longrightarrow \quad \text{Zarna experiences 'UP'}, \tag{5.1-M}$$

$$\downarrow_z \otimes \text{initial} \quad \longrightarrow \quad \text{Zarna experiences 'DOWN'}. \tag{5.2-M}$$

Notice that the reformulation assumes that the final quantum state of the lab specifies Zarna's mental state. If we endorse the thesis that mental states supervene on physical states, then this claim follows automatically from (A1), since **Completeness** says the quantum state is the complete physical state. However if we reject physicalism then the two assumptions may come apart. Thus in a careful reformulation we must replace (A1) by:

(A1-M) Our mental states supervene on our quantum states.

The incompatibility is then with (Gao, 2019, p. 302):

[14] If this claim is true, then we have '$\uparrow_x \otimes \text{initial} \longrightarrow \uparrow_z \otimes \text{final}$,' where *initial* and *final* are the initial and final quantum states of everything in the lab except the measured electron. But then by the same reasoning '$\downarrow_z \otimes \text{initial} \longrightarrow \downarrow_z \otimes \text{final}'$'. The pair then contradicts **Unitarity**.

(A2-M) Same as (A2); the quantum state evolves linearly.

(A3-M) A measurement does not branch mental states; it yields a single mental state representing a single outcome.

(A4-M) Assuming no past mental state branching, implementing a spin-z measurement of an electron in the $\uparrow_z$ quantum state will lead to one experiencing 'UP' as the outcome, and similarly for $\downarrow_z$ and 'DOWN'.

This formulation is most illuminating when applied to non-physicalist interpretations like the Single Mind Interpretation (Albert and Loewer, 1988). This theory endorses (A1) but rejects (A1-M) by positing that an observer's mental state collapses into a particular branch of the wave function even while the physical state continues to evolve according to the Schrödinger equation. The theory thus manages to hold on to (A3-M) even while giving up on (A3).

What would a mentalistic formulation of the control problem look like? The answer reveals a deeper difference between the problems, and highlights a disguised downside of the Single Mind Interpretation, namely that it must reject (B1-M):[15]

(B1-M) Assuming no past mental state branching, implementing a preparation leads to some of the systems being in the specified quantum states.

(B2-M) Same as (B2); the quantum state evolves unitarily.

(B3-M) Same as (A3-M); no mental state branching (even if there is physical branching).

(B4-M) Same as (A4-M); measurements are at least somewhat reliable and informative.

While the assumption that the mental supervenes on the physical is needed for the mentalistic formulation of the traditional measurement problem, it is not needed for the mentalistic formulation of the control problem. What is needed instead is the assumption that our actions can affect the physical (specifically, the quantum state), and that different mental states (observing 'UP', observing 'DOWN') can be correlated with different actions (prepare with D, prepare with D'). This latter assumption is much less philosophically controversial, and shows that theories like the Single Mind Interpretation (since they endorse (B2-M)–(B4-M)) don't get off so easy; in addition to rejecting physicalism, they must reject a certain degree of our control over the physical.

[15]The incompatibility of (B1-M)–(B4-M) follows from an analogous argument, in which Zarna uses the same protocol. Note (B3-M) entails the antecedent to (B1-M) is satisfied in the scenario.

5.4 Relation to interpretations

I now discuss how the major interpretations of quantum mechanics that are currently on the table should respond to the control problem. The goal is not to settle whether they ultimately solve it in a satisfactory way. Rather the aim is to give the lay of the land, and show how the control problem provides a new lens through which to view and compare these interpretations.

Everettian mechanics

Everettian mechanics is arguably least affected by the control problem. This interpretation rejects (B3), **Determinate Input**. According to this theory, when Zarna measures the spin of the electron, her lab branches into two 'worlds', one in which she observes 'UP' and then uses device D, and the other in which she observes 'DOWN' and then uses D'. At the final time, the quantum state of the lab is of neither form (a) 'many systems in D-states' nor (b) 'many systems in D'-states'. But crucially, there is no contradiction with **Preparation** here, since D was not determinately fed the input systems and neither was D'. (Intuitively, since there was not a determinate input, we should not expect the desired output.)

Taken in conjunction with the observation (Section 5.2) that a variety of preparations are not measurements (and so need not cause branching), this suggests that Everettian mechanics does not reject (B1), **Preparation**. There are devices which, as long as they are determinately used, will from a 'God's eye view' prepare the desired quantum state. Other devices, like spin measuring devices, are such that even when determinately used, they only prepare the desired state relative to certain (decohered) branches, in other words only prepare desired 'relative states' as opposed to states simpliciter (Wallace, 2012, Chapter 10).

Now, if Everettian mechanics is true, then realistically the universe has already branched such that most of our uses of preparation devices are not, from a God's eye view, determinate; as a result, we are usually only preparing relative states in the lab, in any case. Since other than cases of branch re-combining, we only interact (from a first-person view) with 'our' world and its future sub-branches, the preparation of relative states is usually all that matters for practical purposes on this theory. Of course, one consequence of the fact that we usually interact only with 'our' world is that we are relatively ignorant of what is happening at other worlds. So while we have access to (parts of) the relative state of our world, the wave function of the universe (construed broadly to include all the worlds or branches) is generally beyond our grasp.

Bohmian mechanics

Bohmian mechanics presents an interesting case for the control problem, because the central strategy it employs to evade the measurement problem, namely to deny **Completeness**, (A1), will not work for the

control problem. Intuitively, then one would expect the control problem to cause trouble for Bohmian mechanics. In fact, I suggest that Bohmian mechanics has a more satisfying response to the control problem than the other interpretations we will consider.

According to Bohmian mechanics, systems have definite positions which are not determined by their wave function (quantum state). The wave function is actually a kind of field which determines how these positions change with time. The wave function always evolves unitarily according to the Schrödinger equation, so **Unitarity**, (B2), holds.

Let us apply the theory to the situation in Zarna's lab. Unlike in Everettian mechanics, according to Bohmian mechanics there is always a determinate fact of the matter as to whether Zarna fed her systems into the device D or the device D'; **Determinate Input**, (B3), is true. Nevertheless these devices do not prepare her subsystems in the specified quantum state; **Preparation**, (B1), is false. To see why, consider the evolution of the wave function of Zarna's lab, as predicted by **Unitarity** when the electron starts in $\uparrow_x$. There will be a component which leads to D-states and a component which leads to D'-states, and its final superposed form will be neither (a) 'many systems in D-states' nor (b) 'many systems in D'-states', even though Zarna determinately fed the subsystems into one of D or D'.

That being said, suppose one 'conditions' the quantum state of these subsystems (see the next paragraph) on the reality of which preparation device was used, that is, on a particular configuration of Bohmian positions. Then one does get a 'conditional' or 'effective' state that takes one of these forms (a) or (b). In this sense, Zarna has indeed prepared. So the theory rejects (B1) in favor of:

(B1') Given determinate inputs, our preparation devices can prepare those inputs in desired conditional quantum states (relativized to an extra parameter) but not desired quantum states simpliciter [**Conditional Preparation**].

Here the control problem highlights an aspect of Bohmian mechanics that the traditional measurement problems do not. Bohmian mechanics solves the standard measurement problem with its postulation of hidden variables; it solves the problem of effect with its account of the dynamics of those hidden variables; and it solves, or at least addresses, the problem of statistics with its statistical postulate (Albert, 1992, p. 140). But in addition to these features, the theory offers up a new notion of 'the quantum state of a subsystem'. This new notion is that of the conditional or effective state. Roughly speaking, it is the analogue of relative state from Everettian mechanics, obtained by conditioning the reduced state of the subsystem (what I have simply been calling its quantum state) on the hidden variable configuration of its environment (Dürr et al., 1992; Norsen, 2016). In this theory, the conditional or effective state, not the

reduced state, determines the measurement statistics we will (likely) observe.[16] Note that while rejecting **Completeness** gives Bohmian mechanics the resources to define this new notion of a quantum state, it does not fall out automatically from the mere introduction of hidden variables; it takes further work to be motivated and developed. (See for instance (Dürr et al., 2005) for an extension of the notion to density matrices.)

The development of this notion is not obviously needed to solve the above measurement problems. It is crucial, however, for addressing the control problem. Any theory which rejects **Preparation** must give a convincing story of how to recover talk of 'preparation', for instance by providing a sense in which our usual claims about preparation of quantum states are true. By motivating this revised notion, Bohmian mechanics makes the rejection of **Preparation** less a full-blown rejection of our ability to prepare, and more a rejection of the assumption that the 'quantum state of a subsystem' must be its reduced state or, more generally, must supervene only on the wave function or density matrix of the parent system. In this way, Bohmian mechanics successfully replaces **Preparation** with **Conditional Preparation**.

GRW theory

GRW theory (Ghirardi et al., 1986) is a collapse theory. Unlike the standard textbook collapse theories, which stipulate that collapses occur when and only when certain measurements are made, GRW says that collapses are spontaneous. Roughly speaking, each individual particle has a very small chance of collapsing (to a definite position) at any given time. However when many particles become entangled to constitute a macroscopic object, like a measuring device, these add up to make the chance of collapse much higher, explaining why macroscopic objects have definite positions.

Since collapses are not consistent with unitarity, GRW automatically rejects (B2), **Unitarity**. Thus one should expect GRW to have an easy time dealing with the control problem. However, matters are not so simple.

According to this theory, upon Zarna's initial measurement the quantum state of the lab very likely spontaneously collapses into either a state in which the device reads 'UP' or a state in which it reads 'DOWN'. (GRW endorses **Completeness**, so the quantum state must specify these facts.) Zarna then goes on to prepare using D or D', respectively, and since the quantum state specifies exactly one of these situations, and the quantum state is complete, **Determinate Input** holds.[17] The devices then genuinely

[16] In some situations, the conditional and reduced states align, namely when the subsystem is appropriately isolated from its environment (so that conditioning on the configuration of the environment does not make a difference). This happens for instance when the reduced state is pure, as in the scenario described by **Competent Measurement**. In this scenario, the initial reduced spin state of the electron is $\uparrow_z$, implying that the initial conditional spin state is also $\uparrow_z$. Thus Bohmian mechanics still predicts we always observe 'UP' in this scenario, as the **Competent Measurement** assumption requires.

[17] Here I ignore complications related to the tail problem (Albert and Loewer, 1990).

prepare according to **Preparation**. Depending on how they operate, they may or may not induce more spontaneous collapses as part of their preparation process.

So far so good. But while GRW avoids the incompatibility by denying (B2), its approach raises another challenge. GRW permits non-unitary evolutions, yet we do not usually say that if GRW theory is true, then cloning is possible. So what is the sense in which GRW forbids the construction of cloning machines? The answer must be substantive, but it cannot rule out Zarna's measure-then-prepare schemes, since we know these are possible.

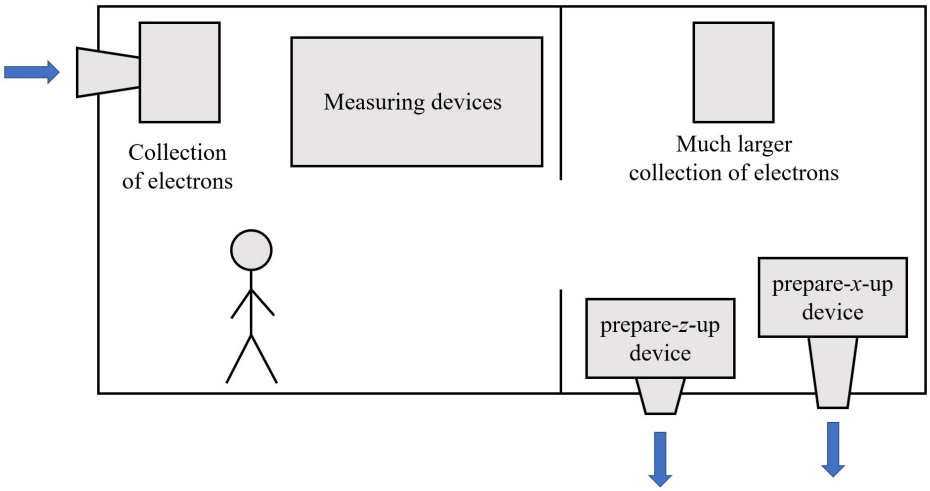


Figure 5.4: Cloning non-orthogonal spin states with Zarna's laboratory.

The needle is not trivial to thread. According to this theory (and other collapse theories), such schemes really do end with preparing many copies of the desired state, making them look like a kind of cloning process. To press the point, suppose we configure Zarna's set-up so that the lab itself has inputs and outputs like a machine, as shown in Figure 5.4. Then we can pick a scheme such that the machine behaves as follows:

> The user feeds the machine an electron in either one of the non-orthogonal states $\uparrow_z$ or $\uparrow_x$ as input (top left arrow of Figure 5.4). With about 71% probability, the machine outputs nothing. But with about 29% probability, the machine outputs many (can be set to any desired finite number) electrons from one of its exits (bottom right arrows). Assuming GRW, these outputs are guaranteed to be genuine copies of the input, that is, to have the same spin state.

In some sense, this machine clones. But it is perfectly possible to build.[18]

[18] Zarna carries out the unambiguous state discrimination in Nielsen and Chuang (2002, p. 92); if she obtains the POVM result corresponding to $\uparrow_z$, she feeds the much larger collection of electrons (we can set this as arbitrarily large as we want) into the prepare-z-up device shown, and similarly for $\uparrow_x$. If she obtains the 'inconclusive' result (probability $1/\sqrt{2}$), she does nothing.

A promising response to these cases is to suggest that GRW only forbids the construction of machines which 'unitarily clone' quantum states. The basic idea is that we should allow for cloning by means of probabilistic collapse, but forbid cloning by means of a unitary process. After all, it is really the latter type of cloning which the no-cloning theorem deems problematic.

This response is on the right track, but it raises a major complication: what to say about more intermediate cases, like machines which clone spin states by means of an 'approximately unitary' process. Since the spatial wave function evolves indeterministically in GRW (regardless of whether a measurement is being made), small entanglements between the spatial and spin wave functions can make it so that the spin state dynamics is also indeterministic and hence non-unitary, but still approximately unitary, in the sense that the chance of a spin state collapse is still very small. Intuitively, although there is no unitary process here, the cloning of spin states should still be forbidden to some extent. To address this type of complication, a starting place may be the relatively recent literature on 'approximate cloning' and its limits (Werner, 1998; Scarani et al., 2005), as well as the literature on 'probabilistic cloning' and its limits (Duan and Guo, 1998a,b; Pati, 2000; Chen et al., 2011).[19]

Other quantum theories

One of the interesting differences between the control problem and the standard measurement problem is that the control problem makes no assumption about what quantum states specify or represent. Thus, we should expect the problem to be particularly revealing when applied to quantum theories of a more subjectivist, pragmatist or instrumentalist flavor, which deny that quantum states directly represent physical properties, but maintain that we measure and prepare them.

There are several factors which complicate the discussion of these theories. One is whether they accept **Unitarity** as formulated. Another is whether they accept a premise presupposed in the arguments above:

(A5, B5) All physical systems—including, for instance, Zarna's lab—have a unique, determinate quantum state (wave function or density matrix) at a given time [**Determinate State**].

Quantum Bayesianism (Caves et al., 2002), for instance, rejects this assumption. Of course, in Quantum Bayesianism we can still attribute quantum states to systems as long as we also specify a third parameter: the epistemic perspective of a particular agent. Thus it is still possible to re-run the argument on this

[19]The approximate cloning literature considers imperfections in the cloning process, but a drawback (for current purposes) is that it still assumes the process is deterministic. The probabilistic cloning literature argues that it is possible to clone non-orthogonal quantum states as long as the cloning is probabilistic in nature, and then derives several limitations on this kind of process, including maximum probabilities for success and maximum cloning efficiencies. None of these results are formulated in the GRW framework, and so they do not take into account the subtleties of the spontaneous collapse theory. But they provide a possible starting point. (It should be noted that, incidentally, Ghirardi himself was one of the early discoverers of the no-cloning theorem (Peres, 2003).)

theory, by relativizing all the states to the perspective of a 'friend' outside the lab. (For hints of how a QBist solution to this revised problem might go, see the QBist analysis of how to translate talk of 'not knowing which quantum state was prepared' in tomography scenarios (Fuchs and Schack, 2004).) In any case, the point is that this is a complication which must be kept in mind.

Another theory of interest is Healey's (2017) pragmatist interpretation, according to which quantum states prescribe rational degrees of belief but need not specify physical properties (**Completeness** is not assumed). Two agents may rationally assign different quantum states to the same system (**Determinate State** is false), but only if they are in different physical situations. In a Wigner's friend scenario, the quantum state relative to the friend outside the lab satisfies **Unitarity** (Healey, 2017, Chapter 11). Thus there is a question of how to cash out **Preparation** so that, when applied to Zarna's lab, there is no inconsistency with unitarity, yet we still recover a sense in which (even according to her friend) Zarna successfully prepares.

Of course, there are many other theories of interest, including Bub's information-theoretic interpretation (Bub, 2005), Rovelli's relational interpretation (Rovelli, 1996), and Peres's operationalist view (Peres, 2006). I do not purport to adequately treat these frameworks here. I leave it as an open question how they solve, or dissolve, the control problem.

5.5 Relation to state determination

We are now in a position to diagnose the argument of D'Ariano and Yuen (1996) discussed in Chapter 4. As we already discussed in Chapter 4, it assumes **Unitarity**, (B2). And there is the assumption that we can determinately prepare distinct non-orthogonal states as a function of what state determination result we obtain, which amounts to the conjunction of (B1) & (B3). But these assumptions, though individually innocent, are together incompatible with:

(B4') if we have a sufficiently large (finite) ensemble of subsystems which we know to be in a product of the same initial wave function, then we can come to know, with reasonable reliability and to a good approximation, what that initial wave function is [**Competent State Determination**].

The derivation proceeds much the same as in Section 5.3, except for the final step, which invokes the reliability condition in (B4'). Consider the Zarna's set-up, except suppose she is given state determination equipment, and informed that her collection of electrons is an N-product of some 'source state'; some wave function. Her protocol: 'If I end up confident that the source state is approximately $\uparrow_z$, I will feed all my many ($>> N$) subsystems from the other room into D. Otherwise, I will feed them into D'.' Suppose the

source state is $\uparrow_z$. We have, by (B1, B2, B4'):

$$\uparrow_z{}^{\otimes N} \otimes \text{initial} \longrightarrow \text{many systems in } D\text{-states.}$$

But now we ask how the quantum state of the lab would have evolved if the source state were ψ where ψ is non-orthogonal to $\uparrow_z$ and non-negligibly different from it:

$$\psi^{\otimes N} \otimes \text{initial} \longrightarrow \boxed{\quad ? \quad}.$$

The answer 'many systems in D'-states' is forbidden by (B2) for the same no-cloning reasons as before. But (B1, B3) entail that the answer is not of this form only if Zarna did not feed her subsystems into D', which by the protocol means she is confident the source state was $\uparrow_z$. So if the source state were ψ, Zarna would still be confident it is $\uparrow_z$. And this violates the reliability clause in (B4'). Note that (B4'), like (B4), would be disastrous to deny. All interpretations, including Bohmian and Everettian mechanics, endorse it.[20] What these unitary theories may deny, though, is that the antecedent of (B4') is ever realistically satisfied. On these theories, we are typically only ever going to know that an ensemble is in a product of the same effective wave function or relative quantum state, respectively.[21]

Drawing on this observation, I suggest an alternate view of the connection between the no-cloning theorem and state determination. Note that the source of the incompatibility with (B4') is, as in our previous incompatibility result, the no-cloning theorem. In this way, the no-cloning theorem does have implications for state determination. On the one hand, if the kind of cloning described by the no-cloning theorem were possible, then in principle there would be no bounds on our state determination or preparation abilities. On the other hand, since the kind of cloning described by the no-cloning theorem is impossible, there is (assuming **Unitarity**) a tension between these abilities. This tension, put in more general terms below, is what I have been attempting to highlight throughout this chapter, in the form of the control problem.

[20] It is important, for the claim to be plausible, that the systems do not start out unduly entangled with their environments, especially with the relevant measuring devices. I have bypassed these complications by restricting attention to pure states (wave functions). A more general formulation with density matrices should make sure to rule out cases where the systems have the same reduced state, but (thanks to environment entanglement) different Bohmian conditional states or Everettian relative states. For another place in the chapter where this complication is relevant, see footnote 16.

[21] See Section 5.4. Regarding Bohmian mechanics, note that this claim—that we rarely know the quantum state simpliciter of a subsystem, but we do often know its effective wave function—is the counterpart to the claim that **Preparation** is false but **Conditional Preparation** is true.

5.6 Conclusion

In this chapter I have argued that there is another problem for quantum mechanics, namely the control problem. Fortunately, the control problem does not straightforwardly rule out any major existing interpretations of quantum mechanics. However, it brings to the foreground a number of issues requiring attention, the most central being quantum state preparation. As we have seen, the control problem puts pressure on our current theories to spell out more carefully the limits of quantum state preparation. Thanks to the incompatibility (B1)–(B4), any interpretive proposal which neglects state preparation, or simply subsumes it under measurement, cannot be complete.

The problem also places a constraint on our future theorizing. Whereas **Completeness** is a claim about what quantum states are or what quantum states represent, **Preparation** is a claim about what we can do with quantum states. As I have remarked, these two issues can come apart. Thus a candidate future theory might successfully avoid the traditional measurement problem, but end up being ruled out by the control problem.

I conclude with one more gloss of the control problem, which I find illuminating. Suppose we hold fixed **Unitarity** and **Determinate Input** as background assumptions. From this point of view, what the control problem shows is that there is a tension between two of our alleged abilities as agents. First, our ability to gain any sort of reliable information about quantum states. Second, our ability to systematically control quantum states. This tension (precisified as the incompatibility between **Competent Measurement** or **Competent State Determination**, on the one hand, and **Preparation**, on the other) exists regardless of whether we posit hidden variables in addition to quantum states (regardless of **Completeness**). Given our two background assumptions, it is a fundamental aspect of our epistemic relationship to quantum states; an aspect which is captured formally by the no-cloning theorem.

Part III

Modeling chance, the limits

Chapter 6

The Borel-Kolmogorov paradox is your paradox too[†]

A puzzle for conditional physical probability

Setting aside quantum mechanics and returning to chance in general, I now consider some difficulties and puzzles that arise with our chance reasoning when *continuous chancy quantities* are involved. In this chapter, Snow Zhang and I present a puzzle that calls into question the reliability of many of our *independence* judgments in these continuous cases. As we saw in Part I, the assumption of independence plays an important role in much of our chance reasoning.

Without further ado, here is the set-up. A projectile is about to be launched at an obstacle. Suppose the collision is a chancy process, such that X, the speed of the projectile right after impact, might range anywhere from zero to one meters per second, with each value equally likely. For instance, the probability of FAST, the proposition that X exceeds half a meter per second, is $1/2$. At the same time, a second projectile is to be launched toward another obstacle. The possible values of Y, the speed of this projectile right after impact, also range from zero to one meters per second, with each value equally likely. Suppose the quantities X and Y are causally and probabilistically independent. Given this set-up, what is $P(X > 1/2|X = Y)$? In other words, what is the probability of FAST conditional on SAME SPEED?[1]

[†]This chapter is based on joint work with Snow Zhang. The title is a play on another title "The reference class problem is your problem too" by Alan Hájek (2007).

[1] For these questions about *the* probability, I am implicitly assuming we've fixed a particular subject matter or level M; cf. Part I. For example, if the collision chances are fundamental dynamical chances, then $M = H_t(w)$ for the time t of the launch, and $P(X > 1/2|X = Y)$ is really $P_{tw}(X > 1/2|X = Y)$. Generally, in this chapter and the next, I will suppress the relativization to a subject matter (I do not think the puzzles can be solved by appealing to this particular relativity of chance).

A natural thought is that the probability of FAST should not be affected by SAME SPEED: $P(X > 1/2|X = Y) = P(X > 1/2) = 1/2$. Perhaps you don't share this thought, however, and are unready to commit to a specific answer until you have performed some calculations. In that case, you may still be willing to endorse the following weaker claim:

Uniqueness In this set-up, the physical probability of FAST given SAME SPEED is well-defined and determinable based on the information given above.[2]

In support of **Uniqueness**, note that the propositions FAST and SAME SPEED are consistent and refer to well-defined (possible) physical events. It would be odd to insist that there is no fact of the matter regarding the chance of one given the other. Moreover, it seems we have fully specified the relevant probabilistic and causal structure of the set-up. What further information is needed?[3]

Uniqueness is an attractive assumption. However, consider the following equally attractive claim:

Equivalence If B is necessarily equivalent to SAME SPEED, then the physical probability of FAST given B equals the probability of FAST given SAME SPEED (if well-defined).[4]

The thought is that if A, B are necessarily equivalent then one cannot happen without the other, and so (for example) if the occurrence of A raises the chance of C, then the occurrence of B must raise the chance of C, and vice-versa.

Unfortunately, **Uniqueness** and **Equivalence** are jointly untenable; they conflict. Moreover, the problem is not restricted to this one case. The projectile example was designed to be mathematically simple, but there are less artificial examples where the same issue arises. Here is a case from statistical mechanics (Kac and Slepian, 1959). Consider a particle whose trajectory is a Gaussian process $\{x(t)\}_{t\in\mathbb{R}}$. In particular, fixing a time t, the displacement $x(t)$ follows a Gaussian distribution with a mean of 0 meters and a standard deviation of, say, 1 meter from the origin. Let $v(t)$ denote the velocity of the particle at time t, which we suppose also follows a Gaussian distribution and is independent of $x(t)$. Given this set-up, what is $P(v(0) > 1/2|x(0) = a)$? In other words, what is the probability that the particle is fast at $t = 0$, given it is located at point a? As we will discuss, the relevant analogues of **Uniqueness** and **Equivalence** also conflict in this case.

Although this formulation in terms of physical probability and **Uniqueness** and **Equivalence** is new, the mathematical aspect of this puzzle is familiar. If we take physical probabilities to be measures that

[2]See also footnote 1. At least for now, we stay neutral here on whether determinability is understood in a metaphysical sense (e.g. supervenience) or an epistemic sense (e.g. derivability).

[3]To lend credence to the second clause, if $P(X = Y) > 0$ (a positive real), then the probability of FAST given SAME SPEED is indeed determinable based on the information specified above.

[4]Here we have in mind nomological (or physical) necessity, though similar puzzles will arise even if we strengthen to metaphysical or logical necessity.

satisfy the standard Kolmogorov axioms of probability, then the conditioning events in both examples are *probability zero*. Much emphasis has been placed, in the mathematical foundations of probability, on the difficulties associated with conditioning on such events. Most relevant here is the *Borel-Kolmogorov paradox*, which asks for the probability that a randomly chosen point on a sphere will fall on a certain arc of a great circle, given that it falls somewhere on that great circle. It is one of a wide class of *Borel-Kolmogorov-type paradoxes*, in which conditioning on the same probability-zero event B seems to lead to different values of $P(A|B)$ depending on how B is 'viewed' (Borel, 1909; Kolmogorov, 1933; DeFinetti, 1974; Rao, 1988; Jaynes, 2003; Builes, 2020). From a mathematical perspective, the projectile and particle examples are just particular instantiations of this well-known paradox from the probability formalism. Why, then, rehash the paradox in this physical probability setting? We believe that the paradox gains new force and significance when considered in the context of physical probabilities. The aim of this chapter is to bring out this lesson.

The plan is as follows. After presenting the conflict between **Uniqueness** and **Equivalence**, we point out that the paradox arises regardless of whether we use the standard Kolmogorov probability formalism or some other approach to modeling 'probability zero' events. Then we explain how the paradox gains new urgency in this physical probability setting. Finally, we propose a way to make headway on resolving the paradox. The proposal is inspired by the approach that scientists took to the statistical mechanical case (Kac and Slepian, 1959) outlined above.

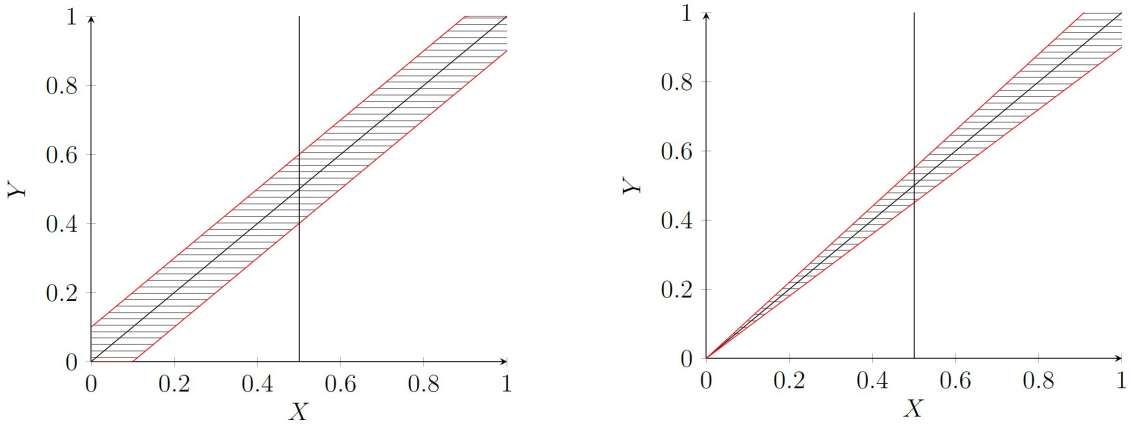
6.1 Uniqueness and Equivalence

To show that **Uniqueness** and **Equivalence** conflict, we begin by assuming **Equivalence** in the background. Then we show that all the standard mathematical and scientific methods for determining the conditional probability of FAST given SAME SPEED yield distinct, incompatible answers depending on how they are applied. Other methods we devise are shown to fail in similar ways. Reflecting on these attempts, we conclude that **Uniqueness** cannot plausibly be maintained under our background assumption.

Conditional densities and limits

In elementary probability theory, there is a standard method for computing conditional probabilities given probability zero events, known as the *conditional densities* method or the *limits* method. Roughly speaking, one calculates $P(A|B)$ as the limit of $P(A|B_n)$ where the B_n are an appropriately chosen sequence of positive-probability events that converge to B.

Recall the projectile set-up above. The idea is to consider the positive probability proposition that X is *close* to Y, $|X - Y| \leq \epsilon$, and then compute the limit of the conditional probability as ϵ shrinks. Under our



(a) $|X - Y| \leq \epsilon$.

(b) $|X/Y - 1| \leq \epsilon$.

Figure 6.1: Two measures of closeness.

assumption that X and Y are independent and both uniformly distributed, the calculation,

$$\begin{aligned} P(X > 1/2 | X = Y) &= \lim_{\epsilon \to 0} P(X > 1/2 \,|\, |X - Y| \leq \epsilon) \\ &= \lim_{\epsilon \to 0} \frac{P(X > 1/2, |X - Y| \leq \epsilon)}{P(|X - Y| \leq \epsilon)}, \end{aligned}$$

yields $1/2$, as (perhaps) expected. To see why, we can picture $|X - Y| \leq \epsilon$ as a band of thickness ϵ around the diagonal line $X = Y$, as shown in Figure 6.1a. The conditional probability of FAST given this event "CLOSE SPEED", one takes the proportion of this band that overlaps with the right-hand side of the square (the area corresponding to $X > 1/2$). By the symmetry of the band, this proportion is clearly $1/2$, and will stay at $1/2$ as the thickness of the band shrinks.

Crucially, however, different ways of quantifying the "closeness" between the two speeds leads to different results when the same method is applied.[5] We have considered whether the difference between the two recoil speeds is small, $|X - Y| \leq \epsilon$. But we might also consider whether their ratio is almost 1, $|X/Y - 1| \leq \epsilon$. And if we instead consider the limit as this ϵ shrinks, the calculation,

$$P(X > 1/2 | X = Y) = \lim_{\epsilon \to 0} P(X > 1/2 \,|\, |X/Y - 1| \leq \epsilon)$$

yields a value greater than $1/2$. Why? Note that smaller speeds are less likely to have smaller ratios (for instance, 0.01 and 0.001 have a ratio of 10, whereas 0.91 and 0.901 have a ratio near 1), and so X/Y being small raises the probability X is large. The situation is represented in Figure 6.1b. The lopsided band

[5]The mathematical backbone of this case is due to Lindley (1982) and Rescorla (2015).

corresponds to the proposition $|X/Y - 1| \leq \epsilon$. The proportion of the band that overlaps with the right-hand side is clearly greater than the proportion that overlaps with the left-hand side.

Let's make the tension between these two results more precise. On the one hand, defining $Z = X/Y$, the limits method implies that:

1. $P(X > 1/2 \,|\, |X - Y| = 0) = \lim_{\epsilon \to 0} P(X > 1/2 \,|\, |X - Y| \leq \epsilon) = 1/2$, and

2. $P(X > 1/2 \,|\, |Z - 1| = 0) = \lim_{\epsilon \to 0} P(X > 1/2 \,|\, |Z - 1| \leq \epsilon) > 1/2$.

On the other hand, since $X = Y$ is necessarily equivalent to $|X - Y| = 0$ and $|Z - 1| = 0$, our background assumption of **Equivalence** implies:

3. $P(X > 1/2 | X = Y) = P(X > 1/2 \,|\, |X - Y| = 0) = P(X > 1/2 \,|\, |Z - 1| = 0)$.

And these three claims are inconsistent. (We could maintain that $P(X > 1/2 | X = Y)$ simply takes more than one value, but this is already to deny **Uniqueness**.)

Now, one might insist that the difference $|X - Y| \leq \epsilon$ is the correct or more natural way to quantify closeness. First note, though, that we do not always take difference as the intuitive measure of closeness. Consider the population data of two counties C and D in years 1918 and 2018 shown in Table 6.1. The absolute difference between the populations is the same in both years, but there is an intuitive sense in which the population of C is "closer" to the population of D in 2018 than it was in 1918. Second, note that even if we grant that absolute difference is the correct measure of closeness, whether the response succeeds is sensitive to one's choice of units. Imagine the Logarians are in every aspect like their Earthian counterparts, except that they measure lengths in logameters rather than meters, and speeds in logameters per second rather than meters per second. The Logarians might agree that absolute difference is the natural measure of closeness. However, because of their choice of scale, they will judge two speeds as close in absolute difference whenever we judge them as close in ratio. For the Logarians, the closer the two speeds are stipulated to be, the more likely it is for them to be fast. And so in the limit, when two speeds are so close that they are equal, the current method will yield the verdict that the probability of $X > 1/2$ is greater than $1/2$.

	1918	2018
C	10	100
D	20	110

Table 6.1: Population of C and D by year.

In fact, by choosing a suitable rescaling in units, we can make the limiting conditional probability any value we want; proofs are in Appendix B (cf. also Arnold and Roberston (2002)):

Fact 1. *Fix the set-up with X, Y probabilistically independent and uniformly distributed on $(0, 1]$. For any real-valued r between 0 and 1, there is a strictly increasing function $h : \mathbb{R} \to \mathbb{R}$ such that $\lim_{\epsilon \to 0} P(X > 1/2 \,|\, |h(X) - h(Y)| \leq \epsilon) = r$.*

Using this result we can obtain a similar inconsistent trio:

1. $P(X > 1/2 \,|\, |X - Y| = 0) = \lim_{\epsilon \to 0} P(X > 1/2 \,|\, |X - Y| \leq \epsilon) = 1/2$,
2. $P(X > 1/2 \,|\, |h(X) - h(Y)| = 0) = \lim_{\epsilon \to 0} P(X > 1/2 \,|\, |h(X) - h(Y)| \leq \epsilon)$ and
3. $P(X > 1/2 | X = Y) = P(X > 1/2 \,|\, |X - Y| = 0) = P(X > 1/2 \,|\, |h(X) - h(Y)| = 0)$ for any strictly increasing h,

where the third claim follows from **Equivalence** together with the fact that, since h is strictly increasing, it is injective and so $X = Y$ if and only if $h(X) = h(Y)$.

Independence reasoning

Examining the limits method and reflecting on its failure, it seems it does not make full use of a key piece of information about the set-up, namely that X and Y are both probabilistically and causally independent. Perhaps these independence features can allow us to determine that the answer is $1/2$ without **Equivalence** raising trouble.

As a first attempt, we might propose the following principle:

Naive Independence If two physical quantities X and Y are causally and probabilistically independent of each other, then the chance that X takes a specific value or range of values is unaffected by whether or not Y happens to sync up with X. In particular, for any x, the probability of $X > x$ given $X = Y$ equals the unconditional chance of $X > x$.

Since X is causally and probabilistically independent of Y in the projectile scenario, Naive Independence yields $P(X > 1/2 | X = Y) = P(X > 1/2) = 1/2$ as desired.

As the name suggests, however, Naive Independence is false. It even gives the wrong verdict in simple finite cases. Imagine Jonny tosses a coin with bias a on the North Pole and Donny tosses a coin with bias b on the South Pole, where the two coin tosses are causally and probabilistically independent of one another. Let HEADS denote the proposition that Jonny's coin lands heads, and let SAME SIDE denote the proposition that Jonny and Donny's coins land on the same side. The unconditional chance of HEADS is a. What is the probability of HEADS given SAME SIDE? The Independence Principle entails that the answer is a. But this

is not true in general. Suppose $a = b = 0.6$. Then we obtain:

$$P(\text{HEADS}|\text{SAME SIDE}) = \frac{P(\text{HEADS \& SAME SIDE})}{P(\text{SAME SIDE})} = \frac{0.6^2}{0.6^2 + 0.4^2} > 0.6.$$

In general, one can check that $P(\text{HEADS}|\text{SAME SIDE}) = P(\text{HEADS})$ if and only if Donny's coin is unbiased ($b = 0.5$). Intuitively, this is because if Donny's coin is biased towards heads (tails), then SAME SIDE means it is more likely that Jonny's coin lands heads (tails). An upshot here is that causal and probabilistic independence of physical quantities is not always a direct guide to conditional chances involving those quantities.

Nevertheless, the observation that HEADS is independent of SAME SIDE when Donny's coin is unbiased also suggests a charitable amendment to Naive Independence. Crucially, in the projectile example, Y is uniformly distributed, not 'biased' toward low or high speeds. At least in that special case, the thought goes, the occurrence or truth of $X = Y$ should not affect the probability of $X \geq x$:

Independence Principle If two physical quantities X and Y are causally and probabilistically independent of each other, have the same (marginal) distribution and are uniformly distributed over their support,[6] then the distribution of X is unaffected by whether or not X happens to sync up with Y. In particular, for any x the chance of $X > x$ given $X = Y$ equals the unconditional chance of $X > x$.

This independence principle also implies the desired verdict in the projectile case. And it fares much better than Naive Independence. Indeed, one can prove (Appendix B) that, if chances are probabilities that satisfy the standard Kolmogorov axioms, then the Independence Principle is true whenever $P(X = Y) > 0$:

Fact 2. *Suppose X and Y are probabilistically independent with identical, uniform distributions. If $P(X = Y) > 0$, then $P(X > x|X = Y) = P(X > x)$ for any x.*

It would be very odd, one might think, if the Independence Principle were false only if $P(X = Y) = 0$. Intuitively, the truth of this principle should not be sensitive to the chance of $X = Y$.

Remarkably, though, trouble looms even for this revised principle. First notice that analogous reasoning supports the following *de*pendence principle.

Dependence Principle If two physical quantities W and V are causally and probabilistically independent of each other, have the same (marginal) distribution and are *not* uniformly distributed over their support, then there exists a w such that the probability of $W > w$ given $W = V$ does *not* equal the unconditional chance of $W > w$.

[6] The support of a real-valued random variable X is the set $\{x \in \mathbb{R} : P(X \in B(x, r)) > 0 \text{ for all } r > 0\}$ where $B(x, r)$ denotes the ball with radius r centered at x.

Indeed, one can prove (Appendix B) that this principle is also true in the case where $P(W = V) > 0$:

Fact 3. *Suppose W and V are probabilistically independent and have the same (marginal) distribution, and are not uniform over their support. If $P(W = V) > 0$, then there is an w such that $P(W > w|W = V) \neq P(W > w)$.*

However, now we can derive a conflict, as follows. Consider the quantities $E_X = \frac{1}{2}mX^2$ and $E_Y = \frac{1}{2}mY^2$ which represent the kinetic energy of the projectiles after impact, where we suppose the two projectiles have the same mass m. Note that E_X and E_Y are not uniformly distributed and, more generally, they satisfy the conditions of the Dependence Principle, with $W = E_X$ and $V = E_Y$. Thus, by the Dependence Principle,

1. There exists an a such that $P(E_X > a|E_X = E_Y) \neq P(E_X > a)$.

On the other hand, by the Independence Principle, we have:

2. For all x, $P(X > x|X = Y) = P(X > x)$.

But since $X > \sqrt{\frac{2a}{m}} \Leftrightarrow E_X > a$, and $E_X = E_Y \Leftrightarrow X = Y$, **Equivalence** implies:

3. $P(X > \sqrt{\frac{2a}{m}}|X = Y) = P(E_X > a|E_X = E_Y)$.

The inconsistency then follows from the fact that $P(X > \sqrt{\frac{2a}{m}}) = P(E_X > a)$, so if $P(E_X > a|E_X = E_Y) \neq P(E_X > a)$ then $P(X > x|X = Y) \neq P(X > x)$ for some x.

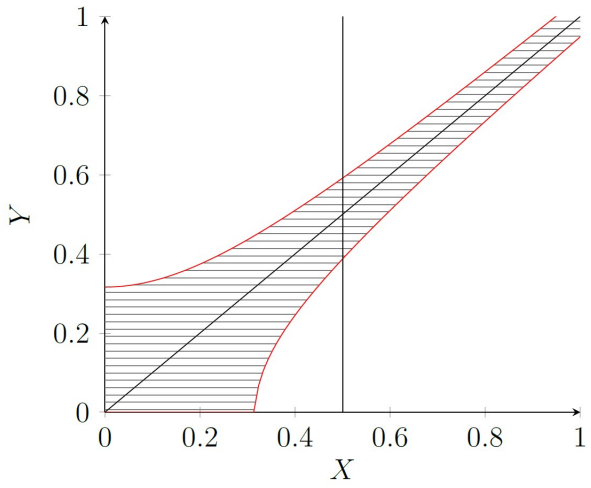


Figure 6.2: $|E_X - E_Y| \leq \epsilon$.

Other methods

There are other methods to derive $P(X > 1/2|X = Y) = 1/2$ that one might attempt. For instance, consider the *discretization* method. Unlike the limits method, which considers $|X - Y| \leq \epsilon$ as ϵ shrinks, the

discretization method begins by modeling the projectile speeds as discrete quantities $\tilde{X}_n$ and $\tilde{Y}_n$ with n distinct possible values, and then taking $\lim_{n\to\infty} P(\tilde{X}_n > 1/2|\tilde{X}_n = \tilde{Y}_n)$.

However, as one might predict, this method yields different results depending on how the quantities are discretized. For example, depending on whether we discretize state space into increments of equal speed or increments of equal kinetic energy, we get different results even in the limit.

We need not belabor the point. Reflecting on these cases, **Uniqueness** is not plausible under the background assumption of **Equivalence**. Crucially, this is not to say there is *no* good argument for the answer of 1/2. Perhaps there are good reasons to discretize X in a certain way, or to take the difference approximation $|X - Y| \leq \epsilon$ rather than the ratio. Our point is simply that such arguments would plausibly have to appeal to physical considerations beyond those supplied in the specification of the chance set-up. It seems implausible that the physical probability of FAST given SAME SPEED supervenes on, or can be determined from, only the facts about X, Y and the situation that were explicitly provided. And this is what **Uniqueness** requires.

6.2 Is the formalism to blame?

It is tempting to chalk up the puzzles and difficulties encountered in the last section to a limitation of our mathematical representation, and in particular a limitation of the standard Kolmogorov probability theory, according to which conditional probabilities are defined derivatively rather than being axiomatized directly. As Gyenis et al. (2017, p. 2595) write:

> According to [a certain popular view], the Borel–Kolmogorov Paradox poses a serious threat for the standard measure theoretic formalism of probability theory, in which conditional probability is a defined concept, and this is regarded as justification for attempts at axiomatizations of probability theory in which the conditional probability is taken as the primitive rather than a defined notion.

Proponents of this view include Kadane et al. (1999), who see the Kolmogorov-Borel paradox as a serious defect of Kolmogorov's theory (pp. 224–5, notation adapted):

> The seeming contradiction is often resolved by claiming that the transformation of variables only yields conditional probability given the sigma field of events determined by the random variable...not given individual events in the sigma field. This approach is unacceptable from the point of view of the statistician who, when given the information that $X = Y$ has occurred, must determine the conditional distribution of X.

We think this view is mistaken for two reasons. First, the tension between **Uniqueness** and **Equivalence** as discussed in Section 6.1 is formalism-neutral. It exists insofar as one grants that $P(A|B)$ equals the limit of $P(A|B_n)$, or that both the Independence Principle and the Dependence Principle are true. And these principles seem attractive even if one takes conditional probability to be primitive rather than derivative.

Second, it turns out that none of the existing alternative probability formalisms really solve the puzzle, conceptualized as a tension between **Uniqueness** and **Equivalence**. For instance, in the projectile example, if we model $P(\cdot|\cdot)$ as a primitive two-place function (as in Popper (1959) and Rényi (1955); for more details on these axiomatizations, see the next chapter), then there will be many $P(\cdot|\cdot)$'s that are compatible with the set-up, but differ on the value assigned to $P(X > 1/2|X = Y)$. In fact, one can prove that for any answer obtained from the limits method discussed in Section 6.1, there exists a Popper function $P(\cdot|\cdot)$ which yields that answer and is otherwise consistent with the set-up.[7] Similarly, if we allow P to range over infinitesimals in addition to positive reals between 0 and 1, then the value of $P(X > 1/2|X = Y)$ typically depends on the choice of ultrafilters that specify how to sum up infinitesimal probabilities.[8] In general, wheeling in these non-standard formalisms does not solve the mystery of whether $P(X > 1/2|X = Y) = 1/2$. All formalisms agree that this question can only be answered with respect to an appropriate Φ, where Φ is a primitive conditional probability function, an ultrafilter, a limiting procedure, a conditioning sigma-field, etc.—but they leave open how a particular choice of Φ can be justified.[9]

6.3 The urgency of the paradox

We have been arguing that the paradox is formalism-neutral, and in this sense, even if you don't ascribe to the formalism of Kolmogorov (1933), "the Borel-Kolmogorov paradox is your paradox too". But it is worth mentioning that if you doubt this conclusion, the paradox should still concern you for a further reason.

Scientists use the standard formalism. They use the formalism to represent continuous chancy quantities and, crucially, to infer claims about conditional probabilities involving those quantities. For example, Sauve et al. (1999) consider a gamma ray shot from position x toward an obstacle and allowed to scatter, as shown in the figure. They ask for the probability that ϕ (the ray's out-of-plane scatter angle) is greater than a, given that θ (the ray's Compton scatter angle) equals b, and reason from the premise that ϕ and θ are independent to the answer $P(\phi > a|\theta = b) = P(\phi > a)$, which they then use in further calculations.

However, our discussion in Section 6.1 casts into doubt the reliability of this independence-based reason-

[7] This follows from a result due to Dubins (1975); see also Chapter 7.4 for related details.

[8] See for example Benci et al. (2012).

[9] That being said, one's choice of formalism is not entirely irrelevant to the paradox. Different formalisms may fit more or less naturally with certain approaches to resolving the paradox. Kolmogorov's formalism, for example (see the next chapter for a presentation of this formalism), lends itself very naturally to approaches that reject **Equivalence**, whereas if one models probabilities using primitive conditional probability functions, for instance, **Equivalence** follows as a theorem of their axioms.

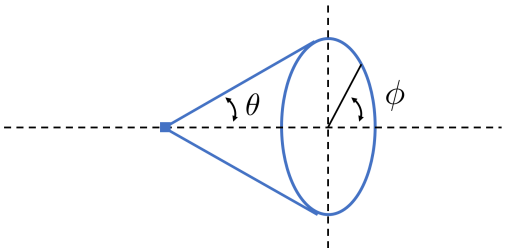


Figure 6.3: Particle scattering.

ing. Thus a question arises whether to try to recover the scientists' judgment, or to somehow explain away the role of these conditional probabilities in the model, not only in this case but also many other cases where theorists and practitioners apply the methods from Section 6.1. We think discarding the judgments is a last resort, so we should explore strategies for resolving the paradox that might vindicate them.

6.4 Toward a resolution

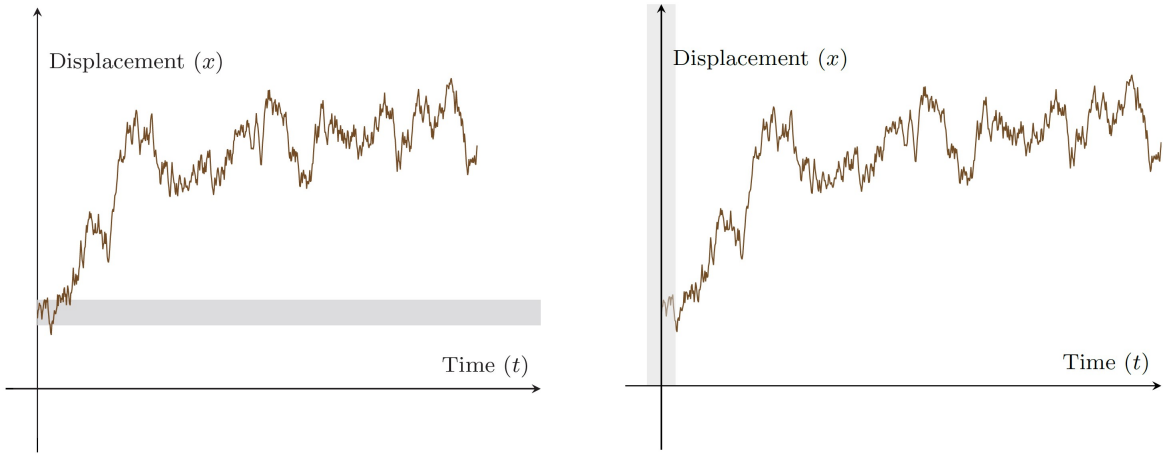
Since **Uniqueness** and **Equivalence** conflict, the first step in a proposed resolution to the paradox is to reject at least one of these premises.

We think that rejecting **Equivalence** is a last resort. In the setting of rational credence, $Cr(A|B)$ may plausibly differ from $Cr(A|B')$ even when B and B' are metaphysically necessarily equivalent. This may happen, for instance, if rational credence is sensitive to intensional or hyperintensional differences; for example, our credences conditional on Hesperus is Phosphorus may differ from our credences conditional on Phosphorus is Phosphorus.[10] However, we find this kind of failure of **Equivalence** much less plausible in the case of physical probability. Physical probability, it seems, should not be sensitive to intensional and hyperintensional differences, at least not in ordinary scenarios like the projectile example.

We propose instead rejecting **Uniqueness**, and in particular the determinability clause. In the projectile scenario, while there is a single true value of the physical probability $P(X > 1/2|X = Y)$ which does not change depending on how $X = Y$ is viewed or represented, this value is not determinable from the information provided. Either $P(X > 1/2|X = Y)$ taking the value it does is a primitive matter or, more plausibly, it takes the value it does in virtue of physical features of the chance set-up that were not explicitly specified.[11]

[10] Rescorla (2015) analyzes the Borel-Kolmogorov paradox in the subjective setting, and pursues this kind of idea.

[11] A related treatment of the paradox is the 'relativization' approach (Kolmogorov, 1933; Jaynes, 2003; Easwaran, 2008; Gyenis et al., 2017). The idea is that conditional probability is a three-place relation; the probability of A given B is always relativized to a sigma-field that represents some "hypothetical experiment" from which B is drawn. Our proposal goes beyond this approach in two respects. First, while we think conditional (physical) probabilities are sensitive to contextual information about the chance set-up, we are agnostic here about (i) whether this information is exhausted by information about the experiment, if there is one, and (ii) whether it can always be mathematically represented as a sigma-field of events (though cf. the next chapter). Second, as Rescorla (2015) points out, the relativization approach does not "explain why conditional probabilities are relativized" but merely "stipulates that they are". Our proposal gives a partial explanatory story: as one might expect, physical probabilities are sensitive to physical features of the chance set-up; what the paradox shows is that these features may



(a) Vertical window.

(b) Horizontal window.

Figure 6.4: The vertical and horizontal windows.

What kind of physical features could those be? On our treatment, this is the central question.

While we do not have a general answer to this question, we believe that the example from statistical mechanics (Kac and Slepian, 1959) described at the beginning of the chapter offers us an important clue. We conclude the chapter by examining this case.

In this case we are asked to compute $P(v(0) > 1/2|x(0) = a)$ where $v(t)$ is probabilistically independent of $x(t)$ for all t. Independence considerations yield the answer:

$$P(v(0) > 1/2|x(0) = a) = P(v(0) > 1/2) \approx 0.31.$$

This is equivalent to the answer we obtain by fixing the time $t = 0$, and the considering the limit as $x(0)$ gets closer to the point a:

$$\lim_{\epsilon \to 0} P(v(0) > 1/2| \, |x(0) - a| \leq \epsilon) = P(v(0) > 1/2) \approx 0.31.$$

This procedure is called taking the 'vertical window'.

However, suppose instead we consider the 'horizontal window' where we hold fixed the point a, and consider a shrinking interval of *time* toward $t = 0$:

$$\lim_{\epsilon \to 0} P(v(0) > 1/2|x(t) = a \text{ for some } t \in [-\epsilon, 0]) \approx 0.47.$$

This second procedure yields a different answer, similarly to the situation in Section 6.1.

not always be fully specified in textbook toy examples.

So far, these findings parallel our earlier findings in the projectile case. However, in this case there is a more obvious physical consideration that breaks the stalemate between these two answers—and surprisingly, not in favor of the independence-based verdict. The consideration is that the process governing the particle is *ergodic*. Roughly speaking, ergodicity means the probabilistic properties of the process $x(t)$ can be deduced from its frequency properties in a relevant time ensemble. It turns out that the horizontal window procedure is the only one that respects this equality between the probability and time ensemble limiting frequency. As Cramér and Leadbetter (2013, p. 221) write (slightly adapted to the current case):

> Further, one of the main properties which we shall require in many cases is that the conditional probability function should admit an appropriate interpretation in terms of a limit of relative frequencies. In the simple example given above, suppose that t_i denotes the time points in $(0, T)$ at which crossings of the level a [by the variable x] occur. Then, one may wish to show that the proportion of those t_i for which $v(t_i) > 1/2$ converges (with probability one) to $P(v(0) > 1/2|x(0) = a)$ as $T \to \infty$, under an appropriate definition of the conditional probability. The definition to be given will be such that "ergodic properties" of this nature may be demonstrated under appropriate conditions. Specifically, suppose we replace the conditioning event $a - \epsilon < x(0) \leq a$ in $P(v(0) > 1/2|a - \epsilon < x(0) \leq a)$ by the event "$x(t) = a$ for some t in the interval $[-\epsilon, 0]$." A conditional probability $P_2(v(0) > 1/2|x(0) = a)$ may then be defined by [the horizontal window definition above].

Granted, these specific considerations will not generalize to all chancy processes where the paradox arises. But it is a useful illustration of how, rather than rejecting **Equivalence**, we might appeal to further physical features of the chance set-up to determine the answer. The example also serves to further illustrate that the paradox discussed in this chapter, concerning the conflict between **Equivalence** and **Uniqueness**, is not just a contrived mathematical puzzle. It arises in routine scientific modeling, and presents a genuine challenge in the epistemology of physical probability.

Chapter 7

Two notions of probabilistic independence

A revolution is afoot in the foundations of probability. The call to arms has been issued: "The entire orthodox mathematical framework for thinking about probabilities should be reconsidered at a fundamental level." Why? It has to do with the notion of probabilistic independence. Independence is a central notion in probability theory, as we have seen throughout this dissertation. But the orthodox definition of independence,

FACTOR $P(F \cap G) = P(F) \cdot P(G)$,

has serious problems with how it treats *probability zero* events or propositions, like $X = Y$ and $x(0) = a$ from the previous chapter. Moreover, the orthodox mathematical framework due to Kolmogorov (1933) does not have the resources to overcome this problem, and so it must be replaced.

Or at least, so Fitelson and Hájek (2017, pp. 3993–4) argue:

> We conclude that it is time to bring to an end the hegemony of Kolmogorov's axiomatization, and with it, the Kolmogorovian account of independence. We seek independence, as it were, from that account of independence. Popper's axiomatization, and the conditioning construal of independence that it inspires, represent more promising alternatives. Long live the revolution!

This final chapter examines the basis for this radical revolt. I conclude that although some foundational revisions are needed, the orthodox framework has the resources to implement these revisions; it should be improved, not discarded. Similarly, in spite of its apparent issues, FACTOR should not be rejected. Instead, we should reconceptualize FACTOR as capturing a new type of probabilistic independence, which I call *almost sure independence*. Once we suitably interpret FACTOR as capturing this new type of independence, we see

that its verdicts are not so problematic after all, and that it plays an important theoretical role. An all-out revolution, then, would be premature.

The plan is as follows. After presenting Fitelson and Hájek's objections to orthodoxy, I point out that their positive proposal—to use Popper's axiomatization, and a conditioning construal of independence instead of FACTOR—leads to the same kind of bad verdicts they sought to avoid. To remedy the situation, I turn to my own positive proposal. I show how both the Kolmogorov and Popper approaches can be adjusted to avoid these verdicts. In the Kolmogorov case, the strategy involves introducing the conditioning construal of independence into the orthodox framework. This does not mean, however, that the conditioning construal should completely supplant the factorizing construal. Indeed, after comparing the amended Kolmogorov and Popper frameworks, I argue that FACTOR still picks out a coherent and important notion, almost sure independence, according to which it does make the right verdicts. So rather than abandoning the factorizing construal in favor of the conditioning construal, we should expand our conceptual repertoire to accommodate both as legitimate types of probabilistic independence. The result, I suggest, is a richer account of the notion.

7.1 Declarations of independence

The problem with FACTOR

The main problem with FACTOR is that it judges probability zero events to be independent of themselves. To see this, suppose $P(F) = 0$. Then we have:

$$P(F \cap F) = P(F) = 0, \qquad P(F) \cdot P(F) = 0 \cdot 0 = 0, \tag{7.1}$$

and so the equality in FACTOR is immediately satisfied.

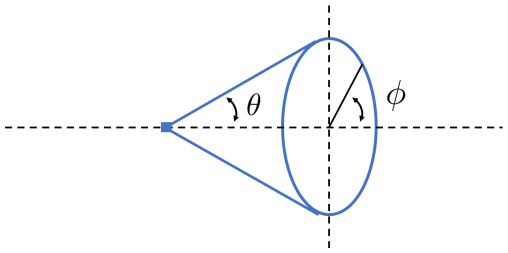


Figure 7.1: Particle scattering.

Example 1 (Scattering)**.** As in Chapter 6, imagine a particle is about to be shot toward an obstacle and allowed to scatter in a random direction. Let θ denote its *Compton scattering angle*, as shown in the figure.

Let F denote the proposition that θ equals 45 degrees. If θ is continuously distributed, then F has zero probability.[1] It then follows that according to FACTOR, F is independent of itself.

Is self-independence always a problem? Perhaps there is one exception which allows for probabilistic self-independence. If the probability of F is one, then there is a sense in which the probability of F is unmoved by the occurrence of F: if F occurs, its probability simply remains at one. But if the probability of F is less than one, and especially if the probability of F is zero, it should follow that F is *de*pendent on F. For Fitelson and Hájek (2017, p. 3987), this is a major sticking point: "We insist that there is *no* concept of independence that should regard probability zero events as independent of themselves. They could not be more hostage to their own occurrence, probabilistically speaking!"

One might object that it is illicit to ask whether an event F is probabilistically independent of itself, like it is illicit to ask whether F is causally independent of itself. However, even granting this point, there is a follow-up worry. Consider G distinct from F, such that G entails F. (For instance, let G be the event that $\theta = 45°$ and $\phi = 30°$, where ϕ is the other angle in Figure 7.1.) Intuitively, F should be dependent on G. But if F has zero probability, then FACTOR judges F independent of G. In fact, according to FACTOR, if F has zero probability then F is independent of *every* other event. To see this, note that if F has zero probability then for any G, $P(F \cap G) \leq P(F) = 0$ and $P(F) \cdot P(G) = 0 \cdot P(G) = 0$.

This last fact creates even more problems for FACTOR. Consider the following example due to Fitelson and Hájek (2017, p. 3989):

Example 2 (Coin machine)**.** Imagine a machine which constructs coins of random biases and then spits them out (in a tossing motion). The possible biases are uniformly distributed on the unit interval $[0, 1]$, with 1 maximally heads-biased and 0 maximally tails-biased. Let H denote the event that the next coin the machine spits out lands heads. Let G denote the event that it has bias 0.95. Is H probabilistically independent of G? Intuitively, the answer is no, since G occurring makes H much more likely. But according to FACTOR, since G has zero probability, it is independent of every event, including H.

In light of all these problems, Fitelson and Hájek (2017, p. 3989) conclude that FACTOR "is, in fact, an *incorrect* account of probabilistic independence".

[1] One might object that although $P(X = x)$ does not equal any positive real number, it is still not zero, it is *infinitesimal*. This could solve the worry with FACTOR, since plausibly an infinitesimal multiplied with itself $\delta \cdot \delta$ is smaller (higher-order) than δ itself. This is a natural route to take, but it is unavailable for the orthodox approach, which represents probabilities with real numbers. Moreover it is not clear this tactic ultimately solves the problem; see for instance Williamson (2007), Pruss (2013) and Easwaran (2014).

Moving from Kolmogorov's framework to Popper's

A natural response to these problems, and one which we implicitly took on board in the previous chapter, is to abandon FACTOR in favor of:

COND $P(F|G) = P(F)$.

In elementary probability texts, this condition is often presented as equivalent to FACTOR.[2] However, while COND and FACTOR are extensionally equivalent when $P(G) > 0$, they come apart when $P(G) = 0$.

At least intuitively, COND gives the right verdicts in the cases discussed above.[3] The issue, however, is that depending on one's characterization of conditional and unconditional probability, $P(F|G)$ may end up undefined when G has zero probability. In particular, according to Fitelson and Hájek, the "orthodox framework" characterizes probabilities by:

> (i) taking unconditional probabilities to satisfy Kolmogorov's axioms,
>
> (ii) taking these unconditional probabilities as primitive, and
>
> (iii) defining conditional probabilities in terms of unconditional probabilities in the standard way, using the so-called *ratio formula*:
>
> $$P(F|G) =_{def} \frac{P(F \cap G)}{P(G)}, \text{ whenever } P(G) > 0.$$

As Hájek (2003) emphasizes, a problem with the ratio analysis (iii) is that it leaves the conditional probabilities discussed above undefined.

If there were no other approaches which addressed this issue, perhaps we could rest on our laurels. But there are. The Popper framework (1955) characterizes probabilities by:

> (i') taking conditional probabilities to satisfy Popper's axioms,
>
> (ii') taking these conditional probabilities as primitive, and

[2] See for instance Dekking et al. (2005), Ross (2010) and Pishro-Nik (2014). Intermediate or advanced texts are usually careful to restrict the equivalence claim to the case $P(G) > 0$ (Feller, 1950; Billingsley, 1986; Grimmett et al., 2001; Bertsekas and Tsitsiklis, 2008).

[3] If G entails F then we should have $P(F|G) = 1$ which does not equal zero. And in the coin machine case (Example 2), we should have $P(H|G) = 0.95$ which does not equal $P(H)$.

(iii') understanding unconditional probabilities as a certain type of conditional probability, as described below.

In what follows we will let Ω denote a non-empty set of outcomes or possibilities, and $\mathcal{F}$ a non-empty set of events or propositions, which are represented by subsets of Ω. We will assume $\mathcal{F}$ is a *field*, i.e. it is closed under complement and finite union (negation and finite disjunction). In the context of Kolmogorov's theory, we will also assume $\mathcal{F}$ is a *sigma-field*, i.e. it is also closed under countably infinite union (countably infinite disjunction).

With these preliminaries in hand, a *Popper function* is a two-place function $P(\cdot|\cdot)$, defined over all pairs of propositions $\mathcal{F} \times \mathcal{F}$, that satisfies:[4]

A1 $0 \leq P(F|G) \leq 1$,

A2 $P(F|F) = 1$,

A3 if $P(H|G) \neq 1$ for some H, then $P(F|G) + P(F^c|G) = 1$ for all F,

A4 $P(F \cap G|H) = P(F|G \cap H) \cdot P(G|H)$, and

A5 there exist F and G such that $P(F|G) \neq 1$.

In this theory, $P(F)$ is reinterpreted as $P(F|\Omega)$, a conditional probability where the conditioning event is the trivial (necessary) event or proposition Ω.

We can precisely formulate the COND definition of independence within this framework. F is independent of G, relative to $P(\cdot|\cdot)$, if:

PCOND $P(F|G) = P(F|\Omega)$.

It is straightforward to show that Popper's axioms entail no probability zero event F is ever independent of itself, or any G which entails it. In particular, one can check **A2** and **A4** imply $P(F|F) = 1$, which is greater than zero, and $P(F|F \cap G) = P(F|G) = 1$ whenever $G \subseteq F$, which is greater than zero.

In light of these observations, Fitelson and Hájek conclude that it is time, not only to dispense with FACTOR, but to dispense with the orthodox framework altogether, and move to an alternative like Popper's.

Independence and intuition

As Fitelson and Hájek have framed the issue, a primary desideratum on a definition like FACTOR is that it satisfy our pre-theoretic intuitions about independence. The thought is that we have a concept of probabilistic

[4]Like Fitelson and Hájek, I focus on the set-theoretic formulation of Popper's theory.

independence, and our goal is to give a definition that is an attractive analysis of that concept. To the extent that a definition violates basic intuitions or platitudes about independence, it is a worse analysis and so a worse definition.

Of course, this framing is not uncontroversial. Some theorists may accept that intuitions should play a role, but take the main desideratum on a definition to be that it supports fruitful theory-building or theorem-proving. Others may reject that we have a pre-theoretic notion of "probabilistic independence" in the same way that we have pre-theoretic notions of truth and logical consequence, for instance (Fitelson and Hájek, 2017, p. 3981). While I am sympathetic to these alternative views, in this chapter I will take on board Fitelson and Hájek's framing. The idea is that, *even* on this framing, Kolmogorov's approach can be vindicated. That said, it is worth mentioning why I think even those who take an alternative view should not simply ignore the issues Fitelson and Hájek raise. It is not just intuition but scientific and statistical practice that suggests a COND-like alternative to FACTOR may be necessary, or at least useful.

As discussed in Chapter 6, independence judgments involving chance zero events play a role in scientific and statistical practice. For instance, recall that in the scattering case (Example 1) scientists reason from the claim that ϕ and θ are probabilistically independent to the claim $P(\phi \geq a|\theta = b) = P(\phi \geq a)$, which they then use in further calculations (Sauve et al., 1999, p. 2076). While this reasoning faces some skeptical worries, I argued in the previous chapter that it can ultimately be vindicated. The reasoning strongly suggests that a conditional construal of independence is operational here. Indeed, it is unclear how to make sense of their reasoning under FACTOR. According to FACTOR, events like $\phi \geq a$ and $\theta = b$ (or $v(0) > 1/2$ and $x(0) = a$) are trivially independent because $\theta = b$ has zero probability. Clearly, it would be a mistake to infer from *this* independence claim the substantive claim that $P(\phi \geq a|\theta = b)$ equals $P(\phi \geq a)$.

Such examples provide additional reason not to dismiss Fitelson and Hájek's independence worries. They are also why one may be increasingly concerned to discover that Fitelson and Hájek's positive proposal for resolving these worries, to use Popper's framework and the conditioning construal PCOND, has a serious difficulty. This difficulty, which I describe next, will take on further dialectical significance as I return to defend Kolmogorov in subsequent sections.

7.2 Trouble in Popper's paradise

It would be one thing if moving from Kolmogorov's framework to Popper's framework allowed us to recover all our intuitions about independence. But the situation is not so straightforward. We have been focused on whether an event can be independent of itself. However, there is another question in the vicinity: can an event [proposition] be independent of its non-occurrence [negation or complement]? That is, can F be

independent of F^c?

Perhaps in one case. If F has probability zero, then the probability of F is unmoved by its non-occurrence F^c; it simply remains at zero. But otherwise, if F has positive probability, then F should be dependent on F^c. This is most clear in the case where F has probability one. To borrow Fitelson and Hájek's turn of phrase, probability one events "could not be more hostage" to their own *non*-occurrence, probabilistically speaking.

However, we now run into trouble. While the Popper axioms entail that a probability zero event must be dependent on itself, they do not prevent a probability one event from being independent of its non-occurrence. Indeed one can show there are Popper functions such that $P(F|\Omega) = 1$ for a contingent event $F \subset \Omega$, and yet F is independent of F^c according to PCOND. The idea behind the proof is to note that there are Popper models in which a non-empty event G fails to satisfy the antecedent of the conditional in **A3** (Roeper and Leblanc, 1999, Chapter 3, Table 3). If G is indeed such that $P(H|G) = 1$ for all H, we say G is *abnormal*. It is the complements of these non-empty abnormal events, $F = G^c$, which meet the condition of the result. The axioms entail $P(F|\Omega) = 1$, but also by hypothesis, we have $P(F|F^c) = P(F|G) = 1$.

In fact, even more troublingly, there are situations in which *every* acceptable Popper function has this feature. Imagine a particle is released at the center of a sphere, and engages in Brownian motion until it hits the surface of the sphere (Pruss, 2015). We seek a probability function over the surface of the sphere which is invariant under rotations about the origin, so that if h is such a rotation, then

$$P(hF|hG) = P(F|G). \tag{7.2}$$

However, it follows from a result due to Pruss (2015, Theorem 2) that according to any Popper function satisfying this constraint, there will be infinitely many probability one events—such as 'the particle does not hit the north pole'—which are independent of their negations—'the particle hits the north pole'. Similar results also hold for the case of throwing a dart at a circular plane target where isometric invariance is required, and a bidirectionally infinite sequence of coin tosses where a natural sequence invariance is required (see Pruss (2015, Theorems 1 & 4); these results do not rely on the Axiom of Choice). Though Pruss does not discuss independence, these results follow as straightforward corollaries to his.

Fitelson and Hájek's proposal now looks to be in serious trouble. After all, the Popper framework allows—and even demands—the same type of counterintuitive verdicts they insisted we avoid. However, all is not lost; there is a fix available. We can add an axiom to the theory which rules out the possibility of these pathological probability one events:

A6 for all events $F \neq \emptyset$, there is an event G such that $P(G|F) \neq 1$.

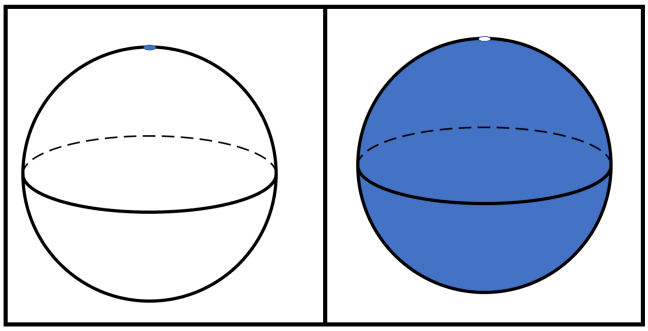


Figure 7.2: Intuitively, the event that the particle does not hit the north pole (right) should be probabilistically dependent on the event that the particle does hit the north pole (left).

It is straightforward to check that if **A6** is added, then no non-trivial probability one events are independent of their negations.[5] Some opposed to Popper functions might object that adding **A6** is *ad hoc*. But I do not share this objection; **A6** looks independently motivated. Indeed, it turns out that the system **A1**–**A6** with this axiom added already has a name: it corresponds to Dubins's theory of *full coherent conditional probabilities* (1975, Section 3). Letting $\mathcal{F}^\circ$ denote $\mathcal{F}$ but with the empty set removed, this theory models conditional probabilities using functions $P(\cdot|\cdot)$ over $\mathcal{F} \times \mathcal{F}^\circ$ that satisfy:[6]

D1 $P(\cdot|F)$ is finitely additive probability function,

D2 $P(F|F) = 1$, and

D3 $P(F \cap G|H) = P(F|G \cap H) \cdot P(G|H)$.

Of course, while the move to full coherent conditional probabilities prevents the bad verdicts, it does not solve Pruss's puzzles. According to this revised theory, in these cases (the sphere, dartboard, and bidirectional sequence of coin tosses) a probability function satisfying the relevant constraints simply does not exist. Thus, when it comes time to assign probabilities, the theory will go silent. This is a downside of the revised theory. But it is arguably better for a probability framework to be silent, than to say something that is false.

[5]Suppose $F \subset \Omega$ is such that $P(F|\Omega) = 1$. Let $H = F^c$. Applying **A6** to H, and invoking **A3**, we have $P(F|H)+Pr(F^c|H) = 1$, which in this case simplifies to $P(F|H) + P(H|H) = 1$ which, by **A2**, implies $P(F|H) = 0$. Thus $P(F|F^c) \neq P(F|\Omega)$ as desired.

[6]To verify the correspondence, note that **A1**, **A3** and **A6** together imply **D1**, and **A4** and **D3** coincide when $G \cap H \neq \emptyset$. For more on this theory, see DeFinetti (1974), Seidenfeld (2001), Coletti and Scozzafava (2002), Seidenfeld et al. (2013), and Easwaran (2019).

7.3 Defending and amending Kolmogorov's framework

I now argue that Kolmogorov's framework is in an equal, if not better, position to address Fitelson and Hájek's independence worries. While the framework does require some modification, these modifications (analogous to **A6**) are independently motivated.

Kolmogorov and probability zero conditioning

Kolmogorov was well aware of problems with probability zero conditioning, especially with regard to continuous quantities. In fact, the two main novelties of his (1933) were, first, his treatment of probabilities in infinite spaces and, second, his definition of conditional probability in terms of *conditional distributions*, which was based on an application of the Radon-Nikodym theorem (Nikodym, 1930), a then-recent result from measure theory.[7] This account of conditioning was much more sophisticated than the ratio analysis (iii).

Fitelson and Hájek (2017, p. 3983) anticipate this point, writing:

> But our complaints about the ratio analysis are hardly aimed at a strawman, since [it] is by far the most commonly used analysis of conditional probability, especially in philosophical applications of probability.

It is true that many treatments simply adopt the naive ratio analysis, and do not even mention Kolmogorov's more sophisticated account. Thus I agree with Fitelson and Hájek that the ratio analysis is not a strawman. However, it should also be emphasized that within probability and statistics, Kolmogorov's sophisticated account is widely viewed as the appropriate generalization of the ratio formula to probability zero events (Billingsley, 1986; Rao and Swift, 2006; Çınlar, 2011).[8] In this sense it is the more interesting target view. Even introductory probability textbooks, for example, contain uses of *conditional densities*, an application of Kolmogorov's account discussed in Chapter 6 which does not follow from the ratio formula (Feller, 1950; Bertsekas and Tsitsiklis, 2008; Meester, 2008; Ross, 2010; Grimmett et al., 2001; Pishro-Nik, 2014).

The rest of the chapter explains Kolmogorov's more sophisticated account, and argues that it supports a good analysis of independence.

[7] For a historical overview, see Shafer and Vovk (2006).

[8] Fitelson and Hájek (2017, p. 3981) attribute to Billingsley the opposite view, quoting a passage from one of the earlier chapters of his book, in which he seems to endorse the ratio analysis. However, in a later chapter, once it is time to introduce Kolmogorov's sophisticated theory, Billingsley (1986, p. 453) clarifies that "the formula $P(F|G) = P(F \cap G)/P(G)$ does not work here. The whole point of this section is the systematic development of a notion of conditional probability that covers conditioning with respect to events of probability 0".

Kolmogorov's theory of conditional distributions

Kolmogorov's more sophisticated account characterizes probabilities by:

> (i) taking unconditional probabilities to satisfy Kolmogorov's axioms,
>
> (ii) modeling conditional probabilities using the more sophisticated apparatus of conditional distributions, and
>
> (iii) taking at least some unconditional probabilities as primitive and, depending on one's view, also taking conditional probabilities as primitive.

Let us explain these points in order. First, the axioms in (i) are:

Definition 1. Fix a possibility space $(\Omega, \mathcal{F})$. A (countably additive) probability measure $P(\cdot)$ is a one-place, real-valued function on $\mathcal{F}$ that satisfies:

1. $P(F) \geq 0$ for all events $F \in \mathcal{F}$,
2. $P(\Omega) = 1$, and
3. if $F_1, F_2, ... \in \mathcal{F}$ are disjoint, then $P(\cup_i F_i) = \sum_i P(F_i)$ where $i = 1, 2, ...$

One of the reasons "most philosophers—even mathematically literate ones—would be unable to state the more sophisticated account" may be that presenting the theory of conditional distributions (ii) requires a relatively technical measure theoretic apparatus (Fitelson and Hájek, 2017, p. 3983). To make the exposition more accessible, I present a special case of the theory here. This "special case" is still extremely general and captures all the examples we are interested in. I include an overview of the full theory, along with a generalization of my results to this full setting, in Appendix C.

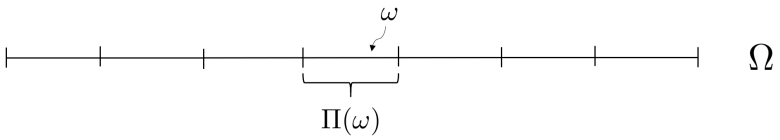


Figure 7.3: A partition Π of possibility space.

We begin with a partition Π of the possibility space Ω. The partition Π represents the *conditioning situation*. In most cases, this corresponds to the different possible values of some chancy quantity Y, like the scattering angle from Example 1. Here, $\Pi(\omega)$ is the set of worlds which agree with ω on the value of Y.

This relativization to a conditioning situation may sound odd, but is motivated by the puzzles and difficulties raised in Chapter 6; different choices of the partition Π will correspond to different choices of limiting procedure. This is not to commit to the view that the conditional chance somehow changes depending on what partition we use. Rather the idea is that the choice of Π encodes, in the mathematical model, a feature of the chance set-up which is not encoded by the unconditional chance function $P(\cdot)$ alone. (In particular, in the terminology of the previous chapter, we can keep **Equivalence** and reject **Uniqueness**.)

Given Π, we can consider functions $P_\Pi(\cdot|\cdot)$ which take events in $\mathcal{F}$ as their first input and possibilities in Ω as their second input, and yield an output[9]

$$P_\Pi(F|\omega),$$

which represents the probability of F conditional on $\Pi(\omega)$ being the true cell of the partition. Kolmogorov's theory places constraints on these functions.

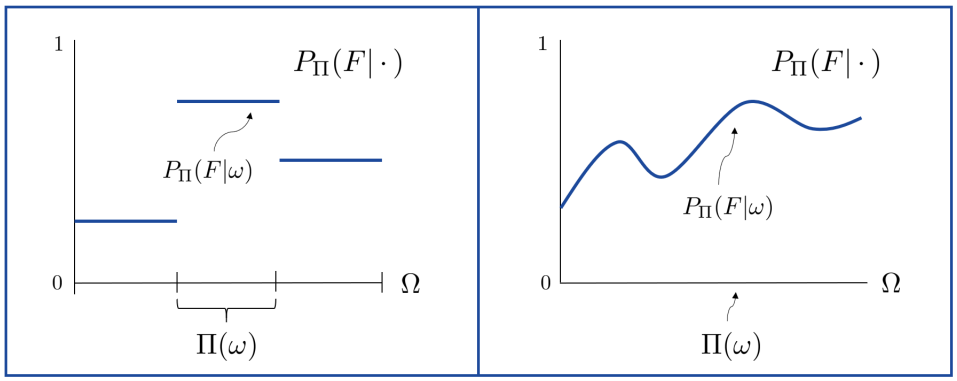


Figure 7.4: An illustration of $P_\Pi(F|\cdot)$ for both finite Π (left) and infinite Π (right).

Before stating these constraints, it is useful to recall a fact from elementary probability theory. Suppose the possibility space Ω is finite. Then it is straightforward to show that $P(F) = \sum_i P(F|G_i) \cdot P(G_i)$, where the G_i label the cells of the partition Π. This is the law of total probability. In our notation,

$$P(F) = \sum_{\omega \in \Omega} P_\Pi(F|\omega) \cdot P(\omega). \tag{7.3}$$

Note that (7.3) can be re-expressed as $P(F) = \mathbb{E}[P_\Pi(F|\cdot)]$, where $\mathbb{E}$ denotes expected value with respect to P. This way of expressing the law is useful, because it generalizes easily to infinite spaces.

Now for the constraints. The first constraint on $P_\Pi(\cdot|\cdot)$ is the *generalized law of total probability* just

[9]The relativization of conditional chance to a partition here should not be confused with the relativization of chance to a subject matter or level M discussed in Part I. The kind of relativization discussed here only is only important when conditional probabilities involving chance zero events are involved; cf. also Chapter 6, footnote 1.

mentioned, also called a *disintegrability* condition. The second constraint, Π-*constancy*, says that if ω and ν belong to the same cell of the conditioning partition, then they yield the same conditional probabilities. The third constraint, *regularity*, says that conditionalizing always yields another probability measure. (This constraint should not be confused with *regularity* in the sense of only assigning zero probability to impossible events; this kind of regularity is certainly not required by the theory. Unfortunately, the literature uses the same word for these two very different notions.) The fourth constraint, *propriety*, is an additional axiom, discussed most notably by the statistician David Blackwell (1963; 1975). Here I am proposing the axiom as an amendment to Kolmogorov's theory:

Definition 2. Fix $P(\cdot)$ and a partition $\Pi \subset \mathcal{F}$. Then a *proper, regular conditional distribution* for $P(\cdot)$ given Π is a function $P_\Pi(\cdot|\cdot)$ on $\mathcal{F} \times \Omega$ satisfying:

K1 (*generalized law of total probability*) for all $F \in \mathcal{F}$, $\mathbb{E}[P_\Pi(F|\cdot)] = P(F)$,

K2 (Π-*constancy*) if $\Pi(\omega) = \Pi(\nu)$, then $P_\Pi(\cdot|\omega) = P_\Pi(\cdot|\nu)$,

K3 (*regularity*) for all ω, $P_\Pi(\cdot|\omega)$ is a probability measure, and

K4 (*propriety*) for all ω, $P_\Pi(\Pi(\omega)|\omega) = 1$.

In more familiar notation, propriety says that $P(G_i|G_i) = 1$ for all G_i. To obtain Kolmogorov's original theory, one removes propriety and strengthens the generalized law of total probability to an *integral formula*. The strengthening is redundant given propriety, which is why I omit it here; see Appendix C for details. The main difference between this theory and Popper's theory is that while Popper's theory defines conditional probabilities over all pairs of events and imposes a multiplication condition **A4**, this theory relativizes conditional probability to a partition and imposes a total probability law **K1** over that partition. (In general one cannot simultaneously impose the generalized law over all partitions (Easwaran, 2019).) To get a sense of this theory, one can verify that the constraints imply the ratio formula as a special case:[10]

$$P_\Pi(F|\omega) = \frac{P(F \cap \Pi(\omega))}{P(\Pi(\omega))}, \text{ whenever } P(\Pi(\omega)) > 0. \tag{7.4}$$

Some readers may be wondering whether Kolmogorov's theory takes conditional probabilities as primitive. As I flagged in (iii), I think the theory is compatible with a few different views on this issue. Certainly one option is to take $P_\Pi(\cdot|\cdot)$ as primitive. Given this choice, one can then either take all the unconditional

[10] Fix ω. First note that by **K3** and **K4**, $P_\Pi(F \cap \Pi(\omega)|\nu)$ equals 0 when $\nu \notin \Pi(\omega)$ and $P_\Pi(F|\nu)$ when $\nu \in \Pi(\omega)$. Thus $\int_{\Pi(\omega)} P_\Pi(F|\nu)\, dP(\nu) = \int_\Omega P_\Pi(F \cap \Pi(\omega)|\nu)\, dP(\nu) =_{def} \mathbb{E}[P_\Pi(F \cap \Pi(\omega)|\cdot)] = P(F \cap \Pi(\omega))$ where for the last equality we invoke **K1**. On the other hand, by **K2**, $\int_{\Pi(\omega)} P_\Pi(F|\nu)\, dP(\nu) = \int_{\Pi(\omega)} P_\Pi(F|\omega)\, dP(\nu) = P_\Pi(F|\omega) \int_{\Pi(\omega)} dP(\nu) = P_\Pi(F|\omega) \cdot P(\Pi(\omega))$. So $P(F \cap \Pi(\omega)) = P_\Pi(F|\omega) \cdot P(\Pi(\omega))$, which entails (7.4).

probabilities $P(\cdot)$ as primitive as well, or just a certain subset. In particular, if one wishes, one can take only the restricted distribution $p(\cdot)$ over Π as primitive, and define the rest of $P(\cdot)$ using $p(\cdot)$ and the conditional probabilities; see the Appendix for details on how this is done. This route is especially natural in cases like Example 2, the coin machine, where what seems basic are the conditional probabilities of outcomes given the possible biases Π of the coin, and the unconditional distribution $p(\cdot)$ over those biases.

What about taking $P_\Pi(\cdot|\cdot)$ as derivative? In situations where all the cells of Π have positive probability, this option is straightforward to articulate, since $P(\cdot)$ and Π uniquely fix $P_\Pi(\cdot|\cdot)$. However, in situations where Π has cells with zero probability, there may be other distributions that satisfy **K1**–**K4**. The question is then how radically these distributions disagree with $P_\Pi(\cdot|\cdot)$. If they only disagree on a probability zero collection of (probability zero) cells (in particular, if $P\{\omega : P_\Pi(\cdot|\omega) = P_\Pi(\cdot|\omega)'\} = 1$), then some theorists will be inclined to treat them as equivalent to $P_\Pi(\cdot|\cdot)$, on the grounds that from the viewpoint of $P(\cdot)$ they are equal. These theorists may then view $P_\Pi(\cdot|\cdot)$ as derivative on $P(\cdot)$ and Π. (Compare: orthodox decision theorists who take utility functions to be derivative on preference orderings, even though a preference ordering does not fix a unique utility function.) However, if other distributions for $P(\cdot)$ given Π disagree with $P_\Pi(\cdot|\cdot)$ on a positive probability set, then $P_\Pi(\cdot|\cdot)$ cannot plausibly be taken as derivative on $P(\cdot)$ on Π.[11]

Two amendments to Kolmogorov's theory

I have already introduced my first amendment to Kolmogorov's theory, the addition of the propriety constraint **K4**. As with **A6**, the intuitive motivation for this constraint is clear.

The second amendment is a definition of independence based on COND. Recall that in Kolmogorov's theory, conditional probabilities are relativized to a conditioning situation Π. Unsurprisingly, then, my proposed independence notion is also relativized to a conditioning situation Π (cf. Chapter 6). Let us say an event F is independent of the cell $\Pi(\omega)$, relative to $P_\Pi(\cdot|\cdot)$, if:

KCOND $P_\Pi(F|\omega) = P(F)$.

This definition is the counterpart to PCOND for Kolmogorov's theory.

KCOND is stronger than FACTOR, in that if F is independent of $\Pi(\omega)$ according to KCOND, then $P(F \cap \Pi(\omega)) = P(F) \cdot P(\Pi(\omega))$; this follows from (7.4). However the converse does not hold. According to KCOND,

[11] Interestingly, while there are examples where regular conditional distributions disagree on a positive probability set, I do not know of any examples where *proper*, regular conditional distributions (for fixed $P(\cdot)$ and Π) disagree on a positive probability set. Thus it is unclear whether such cases ever arise. I leave this as an issue for future exploration. It's also worth mentioning here that theorists who are inclined to view distributions as equivalent when they are $P(\cdot)$-almost surely equal may prefer a slight weakening of regularity (**K3**) and propriety (**K4**) to *almost regularity* and *almost propriety*, where the constraints are only enforced for a probability one set of ω. The idea is that pathological behavior is less troubling when restricted to a null set. One can show, for example, that to be immune to Dutch books, almost regularity and almost propriety suffice (Rescorla, 2018; Meehan and Zhang, forthcoming-b). I will stay neutral on whether the weaker constraints are preferable.

probability zero events are not independent of themselves, and probability one events are not independent of their complements.[12]

KCOND also gives the right verdicts in more complicated cases. Consider the coin machine case (Example 2). One can show that a proper, regular conditional distribution $P_\Pi(\cdot|\cdot)$ exists in this case, with Π corresponding to the different possible values of the bias of the coin, such that the conditional probability of heads $P_\Pi(H|\omega)$ equals the bias at ω (see Appendix C for proof). Hence, relative to this conditional distribution and the KCOND definition, H is not independent of the event G (that the coin has bias 0.95), as desired.

We see, then, that Kolmogorov's sophisticated theory of conditioning has the resources to overcome Fitelson and Hájek's independence problems.

7.4 Interlude: Comparing the amended theories

With these amendments to Kolmogorov's and Popper's theory in place, it is worth returning to compare the two frameworks. Here I will mainly be defending the amended Kolmogorov theory from some initial objections, showing it can still be upheld as a reasonable alternative to the Popper and Dubins approaches. Readers who are pressed for time or are unworried by this issue may skip ahead to Section 7.5, which picks back up the main thread on independence.

Is Kolmogorov's theory just a special case?

My amendments have brought the Kolmogorov and Popper approaches closer together. Propriety now plays an analogous role to Popper's requirement that $P(F|F) = 1$. And the added Popper axiom rules out abnormal events, which were already forbidden in Kolmogorov's theory by regularity. Of course, differences between the theories remain, for example the Kolmogorov approach relativizes conditional probability to a conditioning situation Π, and emphasizes a disintegrability condition over Π (**K1**), rather than a generic multiplication condition (**A4**). That said, prompted by the addition of propriety (**K4**), proponents of the Popper or Dubins approach may attempt to consolidate Kolmogorov's theory as a special case of their own, by way of the following result, which follows as a corollary to theorems of Dubins (1975) and Horn and Tarski (1948) (see Appendix C for details and proof):

Proposition 1. *Fix $P(\cdot)$ and a partition $\Pi \subset \mathcal{F}$. Given the Axiom of Choice, if a regular conditional distribution $P_\Pi(\cdot|\cdot)$ is proper, then it can always be extended to a full coherent conditional probability function*

[12] To see this, fix any proper, regular conditional distribution $P_\Pi(\cdot|\cdot)$. If G is a cell of Π, then by propriety, $P_\Pi(G|\omega) = 1$ for any ω in G, which does not equal zero. So if $P(G) = 0$ then G is not independent of G. Similarly, suppose the complement of G is a cell of Π. By propriety, for any ω in G^c, $P_\Pi(G^c|\omega) = 1$ and hence, by regularity, $P_\Pi(G|\omega) = 0$. So if $P(G) = 1$, then G is not independent of G^c.

$P(\cdot|\cdot)$ *on* $\mathcal{F} \times \mathcal{F}^\circ$ *such that for all* ω, $P(\cdot|\Pi(\omega)) = P_\Pi(\cdot|\omega)$, *and also* $P(\cdot|\Omega) = P(\cdot)$.

Why wouldn't the Kolmogorovian take these extensions on board, and allow their theory to be subsumed by Popper's and Dubins's? (Note that the extension preserves independence judgments; if F is independent of $\Pi(\omega)$ according to KCOND, then in the extension it is also independent of $\Pi(\omega)$ according to PCOND.) There are likely many reasons a Kolmogorovian would resist these extensions, but here I focus on one.[13] The relativization to a conditioning situation Π is a key starting point of the Kolmogorovian approach, and it fixes the privileged partition over which the desired property of disintegrability (**K1**) must hold. (As noted previously, in general it may be impossible to satisfy disintegrability over all partitions.[14]) Given this starting point, the representational significance of $P(\cdot|F)$ for F not a member or segment of Π is unclear. This worry about representational significance grows more serious when one considers how these extensions are achieved through non-constructive means. Arguably, we should be suspicious of the physical relevance of conditional probabilities that can only be given through some version of the Axiom of Choice (Easwaran, 2014, 2019).

An example which illustrates why a Kolmogorovian might resist these extensions is the sphere case (Section 7.2). Recall that in this case, Popper's amended theory is silent. Kolmogorov's theory, in contrast, speaks. Let $P(\cdot)$ be the uniform distribution over the surface and suppose that Π corresponds (for instance) to the latitude of the point in question, so Π partitions the sphere into parallels. Then there does exist a proper, regular conditional distribution $P_\Pi(\cdot|\cdot)$ for $P(\cdot)$ given Π. Moreover, for all rotations h of the sphere, there exists another such distribution $P_{h\Pi}(\cdot|\cdot)$ for $P(\cdot)$ given the rotated partition $h\Pi$, such that the following analog of (7.2) holds (see Appendix C):

$$P_\Pi(F|\omega) = P_{h\Pi}(hF|h\omega). \tag{7.5}$$

These distributions satisfy the requirement that equivalent conditioning situations (those related by rotational symmetry) yield equivalent conditional probabilities, and Kolmogorovians will gladly make use of them. But it is also true (Section 7.2) that any extension of any one of these distributions will violate rotational symmetry with respect to the larger set $\mathcal{F} \times \mathcal{F}^\circ$. This appears problematic, but for the Kolmogorovian, focusing on rotational symmetry with respect to this larger set is misleading, because $\mathcal{F}^\circ$ includes conditioning events that are not part of the conditioning situation Π (or any of its equivalents). Indeed, in this case, the event G that causes the trouble—a nasty set obtained through a paradoxical Banach-Tarski-style

[13] It should also be noted that if propriety is weakened to *almost propriety*, a constraint that some theorists prefer (see footnote 11), then the result no longer holds.

[14] See Easwaran (2019) for more details on this; disintegrability implies the intuitively desirable property of *conglomerability*, which also cannot be enforced over all partitions.

decomposition (Pruss, 2015, p. 269)—is clearly not part of the conditioning situation. So not only does extending to $\mathcal{F} \times \mathcal{F}^{\circ}$ introduce surplus structure into the model, it introduces surplus structure that may be obscuring and misleading.

The issue of existence

We just saw that in the sphere case, Kolmogorov's theory arguably fares better than Popper's when it comes to existence; it speaks when Popper's does not. However, in other cases, it fares no better and, sometimes, worse. (Even in the sphere case, there are *some* conditioning situations Π where invariant proper regular, conditional distributions do not exist—namely those containing G, the nasty set referenced above.[15]) Seidenfeld et al. (2001) offer several results illustrating this point, and Fitelson and Hájek (2017, p. 3983) cite these results as their main objection to Kolmogorov's sophisticated theory. Perhaps the most central result is the one concerning the following example:

Example 3 (Infinite coins)**.** Let $\Omega = 2^{\aleph_0}$ be the infinite coin toss space, with $\mathcal{F}$ the Borel sigma-field and $P(\cdot)$ the fair coin measure. Let Π be the conditioning situation which involves information 'solely about the remote future', in the sense that $\Pi(\omega) = \Pi(\nu)$ if and only if ω and ν differ by only a finite number of toss outcomes. For instance, if ω is the possibility in which no toss lands heads, $\Pi(\omega)$ is the proposition that only finitely many tosses land heads. Seidenfeld et al. (2001) show that every regular conditional distribution for $P(\cdot)$ given Π dramatically violates propriety, in the sense that $P_{\Pi}(\Pi(\omega)|\omega) = 0$ for all ω.

Fortunately, proper, regular conditional distributions still exist in a wide class of cases, covering most actual applications.[16] Moreover, it is not so clear that this example will trouble those who were initially attracted to Kolmogorov's axioms. Consider the possibility ω in which no toss lands heads, so $\Pi(\omega)$ is the proposition that only finitely many do. The number of possibilities in $\Pi(\omega)$ is countably infinite, so learning $\Pi(\omega)$ is tantamount, in this situation, to a fair countable lottery. Unsurprisingly, given regularity (**K3**), we would expect to see some odd results here; it is well known that countable additivity leads to trouble with countable lotteries. Indeed, if we require of a Popper function that $P(\cdot|F)$ be countably additive, then assuming a natural sufficiency condition on how Π and $P(\cdot|\cdot)$ relate, it also follows that no Popper function

[15] In particular, there exists a countable G and set of rotations H such that $hG = G$ for all $h \in H$, yet no finitely additive $\nu(\cdot)$ defined on all subsets of G with $\nu(G) = 1$ is invariant under all $h \in H$ (Pruss, 2015, p. 269). Now let $\Pi = \{G, G^c\}$ and $\omega \in G$. We cannot have $P_{\Pi}(\cdot|\omega) = P_{h\Pi}(h(\cdot)|h\omega)$ for all $h \in H$. Otherwise, since $h\Pi = \Pi$ and $(h\Pi)(h\omega) = \Pi(\omega)$, $P_{\Pi}(\cdot|\omega) = P_{\Pi}(h(\cdot)|\omega)$ for all $h \in H$, which implies $P_{\Pi}(\cdot|\omega)$ behaves as $\nu(\cdot)$.

[16] If $(\Omega, \mathcal{F})$ is a Polish space and Π corresponds to the different possible values of a real-valued random variable Y, then a proper, regular conditional distribution always exists (Meehan and Zhang, 2020b). Here the restriction to the special class of conditional distributions I am considering, those with *informed* sub-sigma-fields (see Appendix C), is crucial. If we drop that restriction, we still get an *almost proper*, regular conditional distribution (see footnote 11); this follows from a generalization of Seidenfeld et al. (2001, Theorem 1) to quasi-separable sigma-fields (those containing a separable sigma-field with the same atoms).

exists in this case.[17] And if we weaken regularity to the requirement that $P_\Pi(\cdot|\omega)$ be finitely additive, then a $P_\Pi(\cdot|\cdot)$ satisfying the constraints does exist (Meehan and Zhang, 2020a, Proposition 7).

The silence of Kolmogorov's theory in this example is troubling, but I do not take it as an overwhelming reason to discard the theory. Of course, one could take a hardline approach in light of these results, and insist we discard both Kolmogorov's theory and Popper's theory, on the grounds that both are silent in cases where they should intuitively speak. To those who seek better alternatives, I say let a thousand flowers bloom. But there is some reason to suspect that any concrete proposal for modeling conditional probability that uses existing mathematical tools will admit of cases like this.[18] The main task, then, is to evaluate the scope and extent of these cases, while also keeping in mind which applications the mathematical framework is being employed for, and, relatedly, what kinds of probabilities it is being used to model.

7.5 A second notion of probabilistic independence

So far I have introduced a new independence definition into Kolmogorov's framework, called KCOND, and showed it can play the role that Fitelson and Hájek seek. However, while this conditioning notion of independence is useful and important, I do not think it is the only one that deserves a place in Kolmogorov's amended theory. Indeed, I now argue that despite its apparent problems, we should not discard Kolmogorov's FACTOR definition, but rather think of it as defining a second notion of probabilistic independence. My strategy is to begin by further exploring the conditional construal KCOND, and then show how a notion equivalent to FACTOR naturally emerges.

Extending KCOND to partitions and quantities

I begin by extending the KCOND notion of independence beyond events, to partitions and quantities. Let us say an event F is independent of the entire partition Π, as opposed to just one of the cells of Π, if:

KCOND-ENTIRE $P_\Pi(F|\omega) = P(F)$ for all ω.

In the case where the partition has two cells $\Pi = \{G, G^c\}$, this is extensionally equivalent to what Fitelson and Hájek call the negation construal of independence:

[17] The condition is: letting $\mathcal{I}(\Pi)$ denote the collection of events $F \in \mathcal{F}$ that can be expressed as a disjunction of cells of Π, every set of the form $\{\omega : P(F|\Pi(\omega)) \leq r\}$ belongs to $\mathcal{I}(\Pi)$. Intuitively: $P(F|\Pi(\cdot))$ does not depend on information outside of Π. (Popperians usually do not need to worry about this condition, because they usually choose $\mathcal{F}$ to be the powerset of Ω, in which case it is trivially satisfied. But here $\mathcal{F}$ is not the powerset.) Given this condition, the result follows from Blackwell and Dubins (1975, Theorem 1) with $\mathcal{B} = \mathcal{F}$ and $\mathcal{A} = \mathcal{I}(\Pi) \subseteq \mathcal{B}$; we can adapt their proof since it does not rely on **K1**, only on properties shared by our Popper function. One can check that $P(\cdot)$ restricted to $\mathcal{A}$ satisfies their condition (a), using the Kolmogorov 0-1 law (Meehan and Zhang, 2020b).

[18] Especially if we demand a total function $\mathcal{F} \times \mathcal{F}$; the construction in the sphere case (see footnote 15) suggests we must give up either (a) finite additivity, (b) propriety, or (c) rotational invariance unless we exclude nasty sets like G from the set of conditioning events.

NEG $P(F|G) = P(F|G^c)$.

More generally, for arbitrary Π, it is extensionally equivalent to the condition that $P(F|G)$ equals $P(F|G')$ for all cells $G, G' \in \Pi$.[19] At least in Kolmogorov's theory, NEG is more naturally understood, not as an alternative to COND, but an extension of COND to partitions.

But what does it mean, intuitively, for F to be independent of Π? If Π corresponds to the different possible values of some chancy quantity Y, then this says that the chance of F is unaffected by what value Y takes. To illustrate this definition more concretely, let us apply it to the coin machine case (Example 2). The heads event H is independent of Π, the partition corresponding to the possible values of the coin's bias, just in case $P_\Pi(H|\omega) = P(H)$ for all ω. As expected, this is not the case: $P_\Pi(H|\omega)$ is equal to the bias at ω, and so is not a constant function of ω as the condition requires.

We can generalize the KCOND-ENTIRE definition further, to pairs of partitions. In order to do this, we need one more piece of terminology:

Definition 3. Fix a possibility space $(\Omega, \mathcal{F})$ and a partition Π of Ω. Then F is a *segment of* Π if it belongs to $\mathcal{F}$ and can be expressed as an arbitrary union (disjunction) of cells of Π.

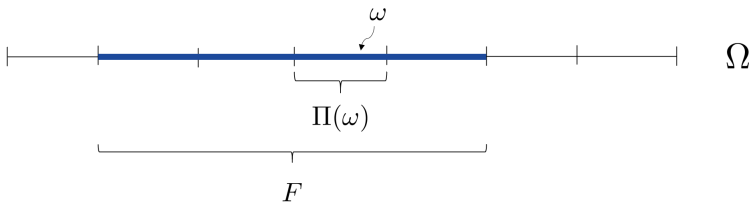


Figure 7.5: In the diagram, both $\Pi(\omega)$ and F can be expressed as a disjunction of cells of Π.

We say a second partition Π' is independent of Π if:

KCOND-ENTIRE For all segments F of Π', $P_\Pi(F|\omega) = P(F)$ for all ω.

One can check that the previous version of KCOND-ENTIRE corresponds to the special case $\Pi' = \{F, F^c\}$.[20] Thus, no confusion will arise from reusing the label. This more general version of KCOND-ENTIRE says, roughly, that no matter how things are Π-wise, the chances of propositions about Π' remain the same.

An important application of this last definition is to pairs of quantities. Suppose Π' corresponds to the different possible values of some quantity X and Π corresponds to the different possible values of some

[19] To check the equivalence in the $\Pi = \{G, G^c\}$ case, it suffices to check that $P(F|G) = P(F|G^c)$ implies $P(F|G) = P(F)$ and $P(F|G^c) = P(F)$. By **K1** we have $P(F) = P(G) \cdot P(F|G) + P(G^c) \cdot P(F|G^c) = (P(G) + P(G^c)) \cdot P(F|G) = P(F|G)$, as desired, and similarly for $P(F|G^c)$. The reasoning for arbitrary Π is the same. Suppose $P(F|G) = P(F|G')$ for all cells $G, G' \in \Pi$, so $P_\Pi(F|\omega)$ equals some constant c for all ω. By **K1**, $P(F) = \int_\Omega P_\Pi(F|\omega)\, dP(\omega) = \int_\Omega c\, dP(\omega) = c \cdot \int_\Omega dP(\omega) = c \cdot P(\Omega) = c$ as desired.

[20] Let $\Pi' = \{F, F^c\} \subset \mathcal{F}$. Suppose $P_\Pi(F|\omega) = P(F)$ for all ω. Then, by regularity, $P_\Pi(F^c|\omega) = 1 - P_\Pi(F|\omega) = 1 - P(F) = P(F^c)$ for all ω, and also $P_\Pi(\Omega|\omega) = 1 = P(\Omega)$ for all ω. Since F, F^c and Ω are all the segments of Π', we are done.

quantity Y. Then we say X is independent of Y if Π' is independent of Π. For instance, consider Example 1, the gamma ray experiment, with ϕ playing the role of X and θ playing the role of Y. The event $\phi \geq a$ is a segment of Π', so the independence of ϕ and θ implies $P(\phi \geq a|\theta = b) =_{def} P_\Pi(\phi \geq a|\omega) = P(\phi \geq a)$ where $\Pi(\omega)$ equals $\theta = b$. Thus, using this notion of KCOND-independence, we can recover the scientists' reasoning in this case.

Introducing almost sure independence

It is natural to consider a weakening of the above independence definitions. F is *almost surely independent* of Π if:

KCOND-ALMOST $P_\Pi(F|\omega) = P(F)$ for a probability one set of ω.

We can apply this definition to Example 2, the coin machine case. Let Π be the partition corresponding to the different possible values of the bias of the coin. The heads event H is almost surely independent of Π just in case $P_\Pi(H|\omega) = P(H)$ for a probability one set of ω. As expected, this is not the case, in fact the equality only holds for a probability zero set of ω.

When do independence and almost sure independence come apart? In the next subsection, we will see some trivial cases where this occurs. But here is a less trivial example. You are playing a game of darts. Consider the event F that you win money. We imagine that the probability of F is the same regardless of what point x you hit, with one exception: if you hit exactly the center, you will certainly win some money. In this example, F is almost surely independent of Π—the partition corresponding to the different possible outcomes of the throw—but it is not surely independent of Π.

We can also generalize KCOND-ALMOST to pairs of partitions. We say a second partition Π' is almost surely independent of Π if:

KCOND-ALMOST For all segments F of Π', $P_\Pi(F|\omega) = P(F)$ for a probability one set of ω.

Again the previous definition corresponds to special case $\Pi' = \{F, F^c\}$.[21] This more general definition says, roughly, that for any proposition F about Π', it is certain that settling things Π-wise won't affect the probability of F. We can make this more concrete by considering the case where Π' corresponds to the possible values of some real-valued quantity X and Π corresponds to the possible values of some real-valued quantity Y. In this case we again say X is almost surely independent of Y if Π' is almost surely independent of Π, or equivalently:[22]

[21] The proof proceeds as in footnote 20 but with the reasoning restricted to ω in the P-one set.

[22] For the ($\Leftarrow$) direction, if $P(X \in S)$ is defined then $X \in S$ belongs to $\mathcal{F}$. Clearly $X \in S$ is then a segment of Π'. For the ($\Rightarrow$) direction if F is a segment of Π', then $P(F)$ is defined (since $F \in \mathcal{F}$ by definition of a segment) and since $\Pi' = \{X = x : x \in X[\Omega]\}$, any segment of Π' is of the form $X \in S$ for some $S \subseteq X[\Omega]$.

for all S, $P_\Pi(X \in S|\omega) = P(X \in S)$ for a probability one set of ω,

where we restrict attention to sets of real numbers S such that $P(X \in S)$ is defined. Continuing the example above: let X denote the amount you win during the game of darts and Y the position you hit. X is almost surely independent of Y, but it is not surely independent, because conditional on Y being the center, the probability that X is greater than zero is much higher.

There is a precise sense in which if a real-valued quantity X is almost surely independent of a real-valued quantity Y, then with probability one the value of Y will not affect the distribution of X:

Proposition 2. *Suppose X and Y are real-valued random variables and Π corresponds to the different possible values of Y. Fix any regular conditional distribution $P_\Pi(\cdot|\cdot)$ for $P(\cdot)$ given Π. If X is almost surely independent of Y, then for a probability one set of ω, $P_\Pi(\cdot|\omega)$ and $P(\cdot)$ agree on the distribution of X: they agree on the sigma-field σX generated by events of the form $X \geq r$.*

For details and the proof, which has also been given (in a slightly different format) by Ramachandran (1979, Lemma 1(ii)–(iii)), see Appendix C.

I submit that almost sure independence is a perfectly coherent notion of probabilistic independence.

Self-independence, revisited

Let us revisit the question of self-independence with the notion of almost sure independence in mind. Recall the coin machine case, where the bias of the coin is uniformly distributed. Consider the event that the coin is perfectly fair, which we can denote by F. Then we can ask a question of self-independence: is the event that the coin is fair *almost surely independent* of the bias of the coin? In other words, is F almost surely independent of Π?

Perhaps surprisingly, the answer is yes. The key observation is that for every possible value of the bias except one, the probability of F will not be affected: it will remain at zero. In particular, $P_\Pi(F|\omega) = 0$ for all ω outside a zero probability set. So although F is not independent of Π, it is *almost surely* independent of Π, since its probability is almost always unaffected by Π.

What about self-independence of events? Is F almost surely independent of itself when $P(F) = 0$? While we cannot answer this question directly yet, because we haven't yet extended the definition of almost sure independence to pairs of events, we can get an intuition for the answer: it should be yes. Almost surely—with a probability of one—F will not occur. And so almost surely, the probability of F will not be affected: it will remain at zero. Although this verdict may initially sound odd, it is not so counterintuitive upon reflection.

Similarly, suppose G has probability zero and G entails F. Is F almost surely independent of G? Almost surely, G will not occur. And so almost surely, the probability of F will be unaffected: it will remain at zero.

So, F is almost surely independent of G.

We can make these observations precise. Setting $\Pi = \{G, G^c\}$, we say F is almost surely independent of G if F is almost surely independent of Π. It is then straightforward to show that if G has probability zero then every event $F \in \mathcal{F}$, including G, is almost surely independent of G.[23]

A theorem for almost sure independence

We now reach the punchline. As foreshadowed, the notion of almost sure independence is deeply related to Kolmogorov's FACTOR definition.

Theorem 1. *Fix a partition $\Pi \subset \mathcal{F}$ and any conditional distribution $P_\Pi(\cdot|\cdot)$ for $P(\cdot)$ given Π. Then Π' is almost surely independent of Π if and only if*

$$P(F \cap G) = P(F) \cdot P(G) \tag{7.6}$$

for all segments F of Π' and G of Π.

In fact, Kolmogorov was well aware of this result. First, let us introduce some of Kolmogorov's notation. Given a possibility space $(\Omega, \mathcal{F})$ and a function $u : \Omega \to E$, let $\mathcal{F}^{(u)}$ denote the collection of subsets $A \subseteq E$ such that $u \in A$ belongs to $\mathcal{F}$ (Kolmogorov, 1933, Chapter III.1). For example, let u be the function $u : \Omega \to \Pi'$ associated with the partition Π' in the theorem above, which maps each possibility ω to its cell $\Pi'(\omega)$. Then $\mathcal{F}^{(u)}$ coincides with the segments of Π'; more precisely, F is a segment of Π' if and only if it equals $u \in A$ for some A in $\mathcal{F}^{(u)}$.[24] With this background in mind, we state Kolmogorov's definition of independence, as he presented it in (1933, Chapter VI, notation slightly altered):

> Two functions, u and v of Ω, are *mutually independent* if for any two sets, A of $\mathcal{F}^{(u)}$, and B of $\mathcal{F}^{(v)}$, the following equation holds:
>
> $$P(u \in A \cap v \in B) = P(u \in A) \cdot P(v \in B).$$
>
> [...] For the independence of u and v, the following condition is necessary and sufficient. For any

[23] Fix $F \in \mathcal{F}$. Consider any ω in the P-one set G^c. Then $P(F) = \int_\Omega P_\Pi(F|\nu)\, dP(\nu) = \int_{G^c} P_\Pi(F|\nu)\, dP(\nu) = \int_{G^c} P_\Pi(F|\omega)\, dP(\nu) = P_\Pi(F|\omega) \int_{G^c} dP(\nu) = P_\Pi(F|\omega)$ as desired, where the first equality follows from the law of total probability, the second and fifth from the fact $P(G^c) = 1$, and the third from Π-constancy and the assumption $\Pi = \{G, G^c\}$.

[24] Suppose F is a segment of Π', so $F \in \mathcal{F}$ and $F = \cup_{\omega \in K} \Pi'(\omega)$ for some $K \subseteq \Omega$. Let $A = \{\Pi'(\omega) : \omega \in K\} \subseteq \Pi'$. Then $\omega \in F$ iff $\omega \in K$ iff $u(\omega) = \Pi'(\omega) \in A$. So F equals $u \in A$. Since $F \in \mathcal{F}$, it then follows $A \in \mathcal{F}^{(u)}$. Conversely, suppose F equals $u \in A$ for $A \in \mathcal{F}^{(u)}$. By definition of $\mathcal{F}^{(u)}$, $F \in \mathcal{F}$. Furthermore F must be a disjunction of cells of Π' since u is Π'-constant (if $\Pi'(\omega) = \Pi'(\nu)$ then $u(\omega) = u(\nu)$). So F is a segment of Π'.

choice of set A in $\mathcal{F}^{(u)}$, the following equation holds almost certainly:

$$P_v(u \in A|\cdot) = P(u \in A).$$

Take the case where u and v are the functions associated with Π' and Π above. Then the "condition" Kolmogorov lays out here is precisely what I have been calling almost sure independence. And what Kolmogorov calls "mutual independence" is just the factorization condition (7.6), which I will dub:

KFACTOR $P(F \cap G) = P(F) \cdot P(G)$ for all segments F of Π' and G of Π.

One can check that when $\Pi' = \{F, F^c\}$ and $\Pi = \{G, G^c\}$, KFACTOR coincides with FACTOR.[25] One can also check that when Π' and Π correspond to the possible values of real-valued quantities X and Y, it coincides with:

for all S and T, $P(X \in S, Y \in T) = P(X \in S) \cdot P(Y \in T)$,

where again we restrict attention to sets of real numbers S and T such that $P(X \in S)$ and $P(Y \in T)$ are defined.[26]

Did Kolmogorov miss the mark, then, when he defined probabilistic independence in terms of KFACTOR? No. In the terminology of KCOND, Kolmogorov was latching on to the notion of almost sure independence. As I remarked, almost sure independence is a perfectly coherent notion of probabilistic independence. It is also very useful. For most applications, we are only concerned with whether the distribution of X is insensitive, relative to the partition Π determined by Y, to 'almost every' value of Y. As the theorem (together with Proposition 2) shows, we can settle this question by evaluating (7.6), without needing to attend to the manifold complications and puzzles associated with conditioning on probability zero events (Borel, 1909; DeFinetti, 1974; Easwaran, 2008; Myrvold, 2015; Pruss, 2013; Rescorla, 2015; Gyenis et al., 2017; Builes, 2020). Relatedly, we can settle this question without worrying about failures of uniqueness, which occur when there are multiple conditional probability functions (multiple proper, regular conditional distributions, or multiple Popper functions) that satisfy the constraints of a problem. In these cases, questions of sure independence may be highly sensitive to arbitrary choices, but questions of almost sure independence are not.

[25] It is a simple exercise to show that $P(F \cap G) = P(F) \cdot P(G)$ implies $P(F^c \cap G^c) = P(F^c) \cdot P(G^c)$, $P(F \cap G^c) = P(F) \cdot P(G^c)$, $P(F^c \cap G) = P(F^c) \cdot P(G)$, and so on.

[26] This definition of the independence of real-valued random variables, known as the "Kolmogorov definition", is slightly different from the more common definition, $P(X \geq r, Y \geq s) = P(X \geq r) \cdot P(Y \geq s)$ for all $r, s \in \mathbb{R}$, known as the "Steinhaus definition". For comparison of the definitions, see Doob (1948), Ramachandran (1975) and Bogachev (2007, Chapter 10.10). In the fully general theory of conditional distributions (see the Appendix) both definitions can be accommodated. In particular, Kolmogorov independence corresponds to (KFACTOR-)independence of the informed sub-sigma-fields $\mathcal{I}(X)$ and $\mathcal{I}(Y)$, whereas Steinhaus independence corresponds to (KFACTOR-)independence of σX and σY.

This weaker notion of independence also has other theoretical benefits. Sure independence is an asymmetric relation; we must be careful to distinguish whether X is independent of Y or Y is independent of X, which makes the generalizations of sure independence to multiple quantities extremely complicated, as Fitelson and Hájek (2017) discuss in the context of PCOND and Popper. Meanwhile, however, the generalization of almost sure independence is elegant and easily applicable, as Kolmogorov (1933) shows.

Don't call me 'surely'

Throughout this discussion, I have been calling 'independence' that notion which corresponds to KCOND and 'almost sure independence' that notion which corresponds (by the theorem) at least extensionally to KFACTOR. This may give the impression that KCOND-independence is the *real* notion of independence, and KFACTOR-independence is only derivative or emergent from KCOND-independence. In fact, this is an artifact of the nomenclature. I could have just as easily introduced KCOND as a characterization of 'strict independence'. I would then have observed that although self-independence of probability zero events is problematic under the notion of 'strict independence', it is not so problematic under regular 'independence', once we better understand what regular (KFACTOR) independence is trying to capture.

Indeed, we can give a gloss on regular independence without mentioning strict independence. Regular independence corresponds to whether the probability of an event F will be almost surely (rough translation: for all practical purposes) unaffected by, unmoved by, insensitive to, etc. whether some other event G does or does not occur, what value some other quantity Y takes, or, more generally, how things are Π-wise for some partition Π of possibilities.

Since *almost sure* is not a folk notion, it is unsurprising that without prompting, our initial intuitions about probabilistic independence latch onto strict independence. But this does not mean strict independence ultimately has a special kind of conceptual priority in the theory. On my view, regular and strict independence are equally legitimate types of probabilistic independence, with their own uses and theoretical roles. We should not discard or relegate either one.

7.6 Conclusion

In the face of Fitelson and Hájek's revolt, I have been playing the part of the loyalist, who seeks to rehabilitate orthodoxy rather than discard it. But ultimately, I do not reject Fitelson and Hájek's call to arms, interpreted as a call to make progress on these issues of independence. Indeed, I have suggested foundational revisions to both the Kolmogorov and Popper approaches in order that they better address these independence questions.

Similarly, my goal is not to insist that we stick to the orthodox framework. Rather, my goal is to point

out that there is a way to shore up this framework, and that once we do this, we see that there may be some real wisdom behind what Fitelson and Hájek call (2017, p. 3980) "the conventional wisdom about probabilistic independence". It just takes some further work to spell it out.

Appendices

Appendix A

Chapters 4 and 5

We begin by recalling the original no-cloning theorem.

Definition 1. A linear map $U : \mathcal{H} \to \mathcal{H}$ is **unitary** if it has an inverse and $\langle U\psi|U\varphi\rangle = \langle\psi|\varphi\rangle$ for any vectors ψ and φ.

Theorem 1 (Original no-cloning theorem)**.** *Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ and let U be a map from $\mathcal{H}$ to itself. Fix any ψ, φ non-orthogonal and distinct. Suppose*

$$U(\psi \otimes \nu) = \psi \otimes \psi$$
$$U(\varphi \otimes \nu) = \varphi \otimes \varphi$$

for some $\nu \in \mathcal{H}_2$. Then U is not unitary (or linear).

Next we prove some close variants of the no-cloning theorem. We focus first on pure vector states, then generalize to mixed states.

Theorem 2 (Some relevant variants)**.** *Let $\mathcal{H} = \otimes_{i\in I}\mathcal{H}_i$ and let U be a map from $\mathcal{H}$ to itself. Fix any ψ, φ non-orthogonal and distinct. Then U is not unitary if:*

(i) for some ν,[1]

$$U(\psi \otimes \nu) = f(\psi) \otimes \psi \otimes \psi \otimes \dots$$
$$U(\varphi \otimes \nu) = f(\varphi) \otimes \varphi \otimes \varphi \otimes \dots,$$

[1] Throughout we just require $\nu \in \otimes_{i\geq 2}\mathcal{H}_2$ and $f(\psi), f(\varphi) \in \otimes_{i\in S}\mathcal{H}_i$ for some $S \subset I \subseteq \mathbb{N}$.

(ii) or, more generally, for some ν and $M > N$,

$$U(\overbrace{\psi \otimes \ldots \otimes \psi}^{N \text{ times}} \otimes \nu) = f(\psi) \otimes \overbrace{\psi \otimes \ldots \otimes \psi}^{M \text{ times}} \otimes \ldots$$
$$U(\varphi \otimes \ldots \otimes \varphi \otimes \nu) = f(\varphi) \otimes \varphi \otimes \ldots \otimes \varphi \otimes \ldots,$$

(iii) or, more generally, there are pairs $\{\chi_j, \chi'_j : j = 1, 2, ..., M\}$ such that $\chi_j \neq \chi'_j$ for each j and, for some ν,

$$U(\overbrace{\psi \otimes \ldots \otimes \psi}^{N \text{ times}} \otimes \nu) = f(\psi) \otimes \chi_1 \otimes \ldots \otimes \chi_M \otimes \ldots$$
$$U(\varphi \otimes \ldots \otimes \varphi \otimes \nu) = f(\varphi) \otimes \chi'_1 \otimes \ldots \otimes \chi'_M \otimes \ldots,$$

where M exceeds the finite bound $N \cdot \log\left[|\langle\psi|\varphi\rangle|^2\right] / \log\left[\max_j |\langle\chi_j|\chi'_j\rangle|^2\right]$.

Proof. All of (i)–(iii) follow as corollaries to the next theorem, in particular the case $\rho = |\psi\rangle\langle\psi|$, $\sigma = |\varphi\rangle\langle\varphi|$. (Note in this case $F(\rho, \sigma) = |\langle\psi|\varphi\rangle|^2$.) □

To generalize to mixed states,[2] we need some more technical apparatus. Though generalizations to mixed states have been made (Barnum et al., 1996), I am not aware of presentations of these particular variants. (Below we also relax from *unitary maps* to *quantum channels*. Physically this suggests we can relax the premise that Zarna's lab is isolated, since many open systems obey a reduced dynamics of this more general form (Rivas and Huelga, 2012).)

Definition 2. Let $\mathcal{D}(\mathcal{H})$ denote the convex set of density operators on $\mathcal{H}$. A map $\mathcal{E} : \mathcal{D}(\mathcal{H}) \to \mathcal{D}(\mathcal{H})$ is a **quantum channel** if it is *convex-linear* and *completely positive*; see Nielsen and Chuang (2002, Chapter 8) for details. Two density operators $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ are **non-orthogonal** if it's not the case that $\text{supp}(\rho) \perp \text{supp}(\sigma)$ (here $\text{supp}(\rho) \equiv (\ker(\rho))^\perp$). The **fidelity**, or closeness, between two density matrices is $F(\rho, \sigma) \equiv (\text{Tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}})^2$.

Definition 3. Fix $\mathcal{H} = \otimes_{i \in I}\mathcal{H}_i$ and $\rho \in \mathcal{D}(\mathcal{H})$. Given some subsystem $S \subset I$, let $\mathcal{R}_S(\rho)$ denote its **reduced state**, defined as the partial trace $\text{Tr}_{\bar{S}}(\rho)$ of the parent state ρ over all elements $i \in I$ not in S (Nielsen and Chuang, 2002, Chapter 2.4).

Theorem 3 (Generalization)**.** *Let $\mathcal{H} = \otimes_{i \in I}\mathcal{H}_i$ and let $\mathcal{E}$ be a map from $\mathcal{D}(\mathcal{H})$ to itself. Fix any ρ, σ non-orthogonal. Then $\mathcal{E}$ is not a quantum channel if:*

[2] I restrict myself to an ordinary QM setting in which all mixed states are represented by density operators; for settings where this can break down, see Ruetsche (2004).

(i) there is a subsystem $S \subset I$ such that, for some μ,[3]

$$\mathcal{R}_S\left[\mathcal{E}(\rho \otimes \mu)\right] = \rho \otimes \rho, \qquad \mathcal{R}_S\left[\mathcal{E}(\sigma \otimes \mu)\right] = \sigma \otimes \sigma,$$

(ii) or, more generally, there is $S \subset I$ such that, for some μ and $M > N$,

$$\mathcal{R}_S[\mathcal{E}(\rho^{\otimes N} \otimes \mu)] = \rho^{\otimes M}, \qquad \mathcal{R}_S[\mathcal{E}(\sigma^{\otimes N} \otimes \mu)] = \sigma^{\otimes M},$$

(iii) or, more generally, there is a set of pairs $\{\pi_j, \pi'_j : j = 1, 2, ..., M\}$ with $\pi_j \neq \pi'_j$ for each j, and a subsystem $S \subset I$ such that, for some μ,

$$\mathcal{R}_S[\mathcal{E}(\rho^{\otimes N} \otimes \mu)] = \pi_1 \otimes \ldots \otimes \pi_M, \qquad \mathcal{R}_S[\mathcal{E}(\sigma^{\otimes N} \otimes \mu)] = \pi'_1 \otimes \ldots \otimes \pi'_M,$$

where M exceeds the finite bound $N \cdot \log\left[F(\rho, \sigma)\right] / \log[\max_j F(\pi_j, \pi'_j)]$.

Proof. To see the general strategy we prove (i) first. Then we prove (iii), which immediately entails (ii).[4] For (i), suppose toward contradiction that $\mathcal{E}$ is a quantum channel. First we invoke the fact that fidelity is *monotonic* under quantum channels (Nielsen and Chuang, 2002, Theorem 9.6), that is

$$F(\rho \otimes \mu, \sigma \otimes \mu) \leq F(\mathcal{E}(\rho \otimes \mu), \mathcal{E}(\sigma \otimes \mu)). \tag{A.1}$$

Since fidelity is also monotonic under partial trace (indeed, partial trace can itself be viewed as a quantum channel or operation, and so this follows from the fact above (Nielsen and Chuang, 2002, Chapter 8.3.1)), we also have

$$F(\mathcal{E}(\rho \otimes \mu), \mathcal{E}(\sigma \otimes \mu)) \leq F(\mathcal{R}_S\left[\mathcal{E}(\rho \otimes \mu)\right], \mathcal{R}_S\left[\mathcal{E}(\sigma \otimes \mu)\right]), \tag{A.2}$$

which under our assumptions imply

$$F(\rho \otimes \mu, \sigma \otimes \mu) \leq F(\rho \otimes \rho, \sigma \otimes \sigma). \tag{A.3}$$

But since fidelity distributes over products (Miszczak et al., 2009, Section 2), this implies

$$F(\rho, \sigma) \cdot F(\mu, \mu) \leq F(\rho, \sigma) \cdot F(\rho, \sigma). \tag{A.4}$$

[3] Here we have $\mu \in \mathcal{D}(\otimes_{i \geq 2} \mathcal{H}_i)$.

[4] Since then $N \cdot \log\left[F(\rho, \sigma)\right] / \log[\max_j F(\pi_j, \pi'_j)] = N \cdot \log\left[F(\rho, \sigma)\right] / \log[F(\rho, \sigma)] = N$.

And now we reach a contradiction. The fact that ρ and σ are non-orthogonal implies $0 < F(\rho, \sigma)$ (Miszczak et al., 2009, Section 2), and since ρ and σ are distinct, we have $F(\rho, \sigma) < 1$. But then it follows that $F(\rho, \sigma)^2 < F(\rho, \sigma)$ which, since $F(\mu, \mu) = 1$, contradicts the above.

For (iii), we use the same reasoning to obtain the inequality

$$F(\rho^{\otimes N}, \sigma^{\otimes N}) \cdot F(\mu, \mu) \leq F(\pi_1, \pi'_1) \cdot \ldots \cdot F(\pi_M, \pi'_M). \tag{A.5}$$

Since the left-hand simplifies to $F(\rho, \sigma)^N$ and since each term in the right-hand is less than $\max_j F(\pi_j, \pi'_j)$, we obtain

$$F(\rho, \sigma)^N \leq (\max_j F(\pi_j, \pi'_j))^M \tag{A.6}$$

which implies $M \leq N \cdot \log(F(\rho, \sigma))/\log(\max_j F(\pi_j, \pi'_j))$. Since π_j is distinct from π'_j for all i, we have $0 \leq \max_j F(\pi_j, \pi'_j) < 1$, and thus the bound is finite, where we use the fact that N is finite and ρ and σ are non-orthogonal, so $F(\rho, \sigma) < 1$ (Miszczak et al., 2009, Section 2). (Note that if $\max_j F(\pi_j, \pi'_j) = 0$ the denominator diverges; in this case we must have $M = 0$.) □

Appendix B

Chapter 6

Fact 1. *Fix the set-up with X, Y probabilistically independent and uniformly distributed on $(0, 1]$. For any real-valued r between 0 and 1, there is a strictly increasing function $h : \mathbb{R} \to \mathbb{R}$ such that $\lim_{\epsilon \to 0} P(X > 1/2 \,|\, |h(X) - h(Y)| \leq \epsilon) = r$.*

Proof. Fix $r \in (0, 1)$. Consider

$$h(x) = \begin{cases} \frac{x}{2(1-r)}, & 0 < x < \frac{1}{2} \\ \frac{1}{4(1-r)} + (x - \frac{1}{2})\frac{1}{2r}, & \frac{1}{2} < x < 1 \end{cases}$$

We claim that h is strictly increasing and $\lim_{\epsilon \to 0} P(X > 1/2 \,|\, |h(X) - h(Y)| \leq \epsilon) = r$. Clearly h is strictly increasing on $(0, 1/2]$ and $(1/2, 1)$. Now let $x, x' \in (0, 1)$ such that $x < 1/2 < x'$. Then $h(x) < \frac{1}{4(1-r)} < h(x')$. So h is strictly increasing on $(0, 1)$.

Now note that for $X > 1/2$, for ϵ sufficiently small,

$$\begin{aligned} |h(x) - h(y)| \leq \epsilon \text{ } \textit{iff } \frac{1}{2r}|x - y| &\leq \epsilon \\ \textit{iff } x - 2r\epsilon \leq y &\leq x + 2r\epsilon. \end{aligned}$$

Similarly, for $X < 1/2$, then for ϵ sufficiently small,

$$\begin{aligned} |h(x) - h(y)| \leq \epsilon \text{ } \textit{iff } \frac{1}{2(1-r)}|x - y| &\leq \epsilon \\ \textit{iff } x - 2(1-r)\epsilon \leq y &\leq x + 2(1-r)\epsilon. \end{aligned}$$

Since X is uniformly distributed, for ϵ sufficiently small,

$$\begin{aligned} P(X > 1/2 \,|\, |h(X) - h(Y)| \leq \epsilon) &= \frac{P(X > 1/2, |h(X) - h(Y)| \leq \epsilon)}{P(|h(X) - h(Y)| \leq \epsilon)} \\ &= \frac{P(X > 1/2, |h(X) - h(Y)| \leq \epsilon)}{P(X > 1/2, |h(X) - h(Y)| \leq \epsilon) + P(X < 1/2, |h(X) - h(Y)| \leq \epsilon)} \\ &= \frac{\int_{\frac{1}{2}}^{1} \int_{x-2r\epsilon}^{x+2r\epsilon} dydx}{\int_{0}^{\frac{1}{2}} \int_{x-2(1-r)\epsilon}^{x-2(1-r)\epsilon} dydx + \int_{\frac{1}{2}}^{1} \int_{x-2r\epsilon}^{x+2r\epsilon} dydx} = \frac{2r\epsilon}{2\epsilon} = r, \end{aligned}$$

as claimed. □

Fact 2. *Suppose X and Y are probabilistically independent with identical, uniform distributions. If $P(X = Y) > 0$, then $P(X > x|X = Y) = P(X > x)$ for any x.*

To prove this fact, we use the following lemma:

Lemma 1. *Let X and Y be two random variables that are independent and identically distributed. Let P be a countably additive probability function. If $P(X = Y) > 0$, then both X and Y are discrete random variables that only take on finitely many values.*

Proof. (of lemma) Since Y is uniformly distributed, it cannot be neither discrete nor continuous, for then there are y, y' such that $P(Y = y) > 0$ but $P(Y = y') = 0$. We claim that Y must be discrete. Suppose toward contradiction that Y is continuous, i.e. for all y in the range of Y, $P(Y = y) = 0$. There are two cases: either X is continuous, or it is not. Suppose X is continuous. Let $Z = X - Y$. Then Z is a continuous variable. So $P(Z = 0) = P(X = Y) = 0$, which contradicts the assumption that $P(X = Y) > 0$. Suppose X is not continuous. Let $A = \{w : P(X = x) > 0\}$, i.e. A is the discrete component of X. Let $Z_1 = Z \cdot 1_{X \in A}$ (the conditional variable that equals Z on the set $X \in A$, and 0 otherwise). Let $Z_2 = Z \cdot 1_{X \notin A}$. Then $P(Z = 0) = P(Z_1 = 0) + P(Z_2 = 0)$. Since both X and Y are continuous on the set $X \notin A$, Z_2 is a continuous variable. So $P(Z_2 = 0) = 0$. On the other hand, $P(Z_1 = 0) = \sum_{x \in A} P(X = x, Y = x) = \sum_{x \in A} P(X = x)P(Y = x) = 0$ since $P(Y = x) = 0$ for all $x \in A$. So $P(Z_1 = 0) = 0$. But then again we have a contradiction with $P(Z = 0) = P(X = Y) > 0$. Therefore Y cannot be continuous. Since a distribution that is neither discrete nor continuous cannot be uniform, it follows that Y (and therefore X) must be discrete. Moreover, since Y is uniform and P is countably additive, both X and Y can only take on finitely many values. □

Proof. (of Fact 2) Let X, Y be two independent random variables with identical, uniform marginal distributions. Suppose $P(X = Y) > 0$. By Lemma 1, both X and Y are discrete. Thus, to show that $P(X > x|X = Y) = P(X > x)$, it suffices to show that $P(X = x|X = Y) = P(X = x)$ for all x.

Clearly if $P(X = x) = 0$, then $P(X = x|X = Y) = 0 = P(X = x)$. Suppose $P(X = x) = P(Y = x) = \alpha > 0$. Then

$$\begin{aligned} P(X = x|X = Y) &= \frac{P(X = x, X = Y)}{P(X = Y)} = \frac{P(X = x, Y = x)}{\sum_{x'} P(X = x', Y = x')} \\ &= \frac{\alpha P(X = x)}{\alpha \sum_{x':P(X=x')>0} P(X = x')} = P(X = x), \end{aligned}$$

as claimed. □

Fact 3. *Suppose W and V are probabilistically independent and have the same (marginal) distribution, and are not uniform over their support. If $P(W = V) > 0$, then there is an w such that $P(W > w|W = V) \neq P(W > w)$.*

Proof. For each interval $(a, b]$, we have

$$\begin{aligned} P(a < W \leq b|W = V) &= \frac{P(a < W \leq b, W = V)}{P(W = V)} \leq \frac{P(a < W \leq b, a < V \leq b)}{P(W = V)} \\ &= \frac{P(a < W \leq b) \cdot P(a < W \leq b)}{P(W = V)} = P(a < W \leq b) \cdot \frac{P(a < W \leq b)}{P(W = V)}. \end{aligned}$$

We consider two cases: either there is $(a, b]$ such that $0 < P(a < W \leq b) < P(W = V)$, or there is not. In the first case, for that $(a, b]$, if $P(W > b|W = V) \neq P(W > b)$, then we are done. If $(W > b|W = V) = P(W > b)$, then $P(a < W \leq b|X = Y) < P(a < W \leq b)$. So

$$\begin{aligned} P(W > a|W = V) &= P(a < W \leq b|W = V) + P(W > b|W = V) \\ &< P(a < W \leq b) + P(W > b) \\ &= P(W > a). \end{aligned}$$

Now suppose that for all intervals $(a, b]$, $P(a < W \leq b)$ equals 0 or is greater than $P(W = V)$. This implies that W is discretely distributed, with its atoms at least of weight $P(W = V)$.[1] Since W and V are identically distributed, the same holds for V. Now since V is not uniform over its support, there are v, v'

[1] *Proof.* It suffices to show that X has a countable support. Suppose towards a contradiction that the support of W is uncountable. Then there exists w in the support of W such that $P(W = w) = 0$ but for all r, $P(W \in B(x, r)) > 0$. It follows that there is r such that $0 < P(W \in B(x, r)) < P(W = V)$. Let $(a, b] = [w - r, w + r]$. Then. $0 < P(a < W \leq b) < P(W = V)$, contradicting our assumption that for all $(a, b]$, $P(a < W \leq b) = 0$ or $P(a < W \leq b) > P(W = V)$.

such that $P(V = v) > P(V = v')$ where both these events have positive probability. Thus

$$\begin{aligned}\frac{P(W = v'|X = Y)}{P(W = v|X = Y)} &= \frac{P(W = v', V = y')}{P(W = v, V = v)} = \frac{P(W = v')P(V = y')}{P(W = v)P(V = v)} \\ &= \frac{P(W = v')}{P(W = v)} \cdot \frac{P(V = v')}{P(V = v)} < \frac{P(V = v')}{P(V = v)},\end{aligned}$$

and so either $P(W = v|V = V) \neq P(W = v)$ or $P(W = v'|W = V) \neq P(W = v')$ or both. Since a change in W's posterior distribution implies a change in its cdf, there must exist an x which satisfies the desired condition. □

Appendix C

Chapter 7

The integral formula, and treating $P(\cdot)$ as derivative

In Chapter 7.3, upon defining a proper, regular conditional distribution, I remarked that to obtain Kolmogorov's theory of regular conditional distributions, one removes propriety (**K4**) and strengthens the law of total probability (**K1**) to an integral formula. The formula is:

K1$^+$ (*integral formula*) for all $F \in \mathcal{F}$ and segments $G \in \mathcal{F}$ of Π,

$$P(F \cap G) = \int_G P_\Pi(F|\omega)\, dP(\omega). \tag{C.1}$$

One can verify that the strengthening to the integral formula is redundant given propriety, as claimed: (Proofs are located in Appendix C.)

Proposition 3. *Suppose* $\Pi \subset \mathcal{F}$. *Then* ***K1***–***K4*** *imply* ***K1***$^+$.

In Chapter 7.3, I also remarked that there is a way to treat some of the probabilities in $P(\cdot)$ as derivative. On this way of going, which is related to the use of *transition kernels* to construct probability spaces (Çınlar, 2011, I.6, IV.4), the order of presentation of Kolmogorov's theory is slightly reversed. First, one specifies a function $\mu_\Pi(\cdot|\cdot)$ on $\mathcal{F} \times \Omega$ that satisfies **K2**–**K4**. This function must be such that for all $F \in \mathcal{F}$, $\mu_\Pi(F|\cdot)$ is a random variable (for all $r \in \mathbb{R}$, $\{\omega : \mu_\Pi(F|\omega) \leq r\} \in \mathcal{F}$) so that we can take its expectation. Second, one specifies a distribution $p(\cdot)$ over Π, i.e. a probability measure defined on the *segments* $F \in \mathcal{F}$ of Π (see Section 7.5). Finally one defines $P(\cdot)$ as the expectation of $\mu_\Pi(\cdot|\cdot)$ with respect to $p(\cdot)$:

$$P(\cdot) =_{def} \int_\Omega \mu_\Pi(\cdot|\omega)\, dp(\omega). \tag{C.2}$$

This will be a probability measure since $\mu_\Pi(\cdot|\omega)$ is a probability measure (one can verify this using the Monotone Convergence Theorem (Çınlar, 2011, Exercise I.5.6)). Moreover thanks to **K3** and **K4**, $P(\cdot)$ will agree with $p(\cdot)$. It then automatically follows that $\mu_\Pi(\cdot|\cdot)$ is a proper, regular conditional distribution for $P(\cdot)$ given Π, which we can rewrite as $P_\Pi(\cdot|\cdot)$. For an example of this construction, see the coin case (Appendix C).

This procedure can also be generalized to the full theory in Appendix C. Fix $\mathcal{G} \subseteq \mathcal{F}$. First one specifies a function $\mu_\mathcal{G}(\cdot|\cdot)$ on $\mathcal{F} \times \Omega$ that satisfies $\mathcal{G}$-measurability, regularity and propriety. Then one specifies a probability measure $p(\cdot)$ on $\mathcal{G}$. Finally one defines $P(\cdot)$ like in (C.2). We have $P(\cdot)|_\mathcal{G} = p(\cdot)$ thanks to propriety and regularity, thus $\mu_\mathcal{G}(\cdot|\cdot)$ is a proper, regular conditional distribution for $P(\cdot)$ given $\mathcal{G}$, which we can rewrite as $P_\mathcal{G}(\cdot|\cdot)$.

Generalization to the full theory

In the body of Chapter 7 I presented a special case of the modern theory of conditional distributions. In particular, I focused on a special class of conditioning situations. Here I present the more general theory and explain how the theory in Section 7.3 is a special case.

In the more general theory, a conditioning situation is represented by a sigma-field $\mathcal{G} \subseteq \mathcal{F}$, known as the conditioning sub-sigma-field. Conditional distributions are defined as:

Definition 4. Fix a sub-sigma-field $\mathcal{G} \subseteq \mathcal{F}$ of Ω. Then a *proper, regular conditional distribution* for $P(\cdot)$ given $\mathcal{G}$ is a function $P_\mathcal{G}(\cdot|\cdot)$ on $\mathcal{F} \times \Omega$ that satisfies:

K1-G (*generalized law of total probability*) for all $F \in \mathcal{F}$, $\mathbb{E}[P_\mathcal{G}(F|\cdot)] = P(F)$, and

K2-G ($\mathcal{G}$*-measurability*) for all $F \in \mathcal{F}$, $P_\mathcal{G}(F|\cdot)$ is $\mathcal{G}$-measurable, in other words for all real numbers $r \in \mathbb{R}$, the set $\{\omega : P_\mathcal{G}(F|\omega) \leq r\}$ belongs to $\mathcal{G}$,

K3-G (*regularity*) for all ω, $P_\mathcal{G}(\cdot|\omega)$ is a probability measure,

K4-G (*propriety*) for all ω, $P_\mathcal{G}(G|\omega) = 1$ whenever $\omega \in G \in \mathcal{G}$.

A similar argument (see Appendices C and C) shows these constraints imply:

K1$^+$-G (*integral formula*) for all $F \in \mathcal{F}$ and $G \in \mathcal{G}$,

$$P(F \cap G) = \int_G P_\mathcal{G}(F|\omega)\, dP(\omega). \tag{C.3}$$

To define a regular conditional distribution for $P(\cdot)$ given $\mathcal{G}$ we again remove **K4-G** and strengthen **K1-G** to **K1$^+$-G**. To define a conditional distribution for $P(\cdot)$ given $\mathcal{G}$ we go the extra step of removing the regularity requirement **K3-G**.

In the text, I focused on conditioning sub-sigma-fields that are *informed*:

Definition 5. Fix a sigma-field $\mathcal{G} \subseteq \mathcal{F}$. The *atom* of $\mathcal{G}$ at ω, written $\mathcal{G}(\omega)$, is the arbitrary intersection of all the elements of $\mathcal{G}$ that contain ω (it may not necessarily belong to $\mathcal{F}$). The atoms of $\mathcal{G}$ always form a partition, called the partition *determined by* $\mathcal{G}$. A sigma-field $\mathcal{G} \subseteq \mathcal{F}$ is *informed* if it contains all the $F \in \mathcal{F}$ that are segments of this partition.

Fixing a possibility space $(\Omega, \mathcal{F})$, there is a one-to-one correspondence between partitions Π of Ω and informed sub-sigma fields of $\mathcal{F}$. Given Π, the unique Π-*informed sigma-field* $\mathcal{I}(\Pi) \subseteq \mathcal{F}$ is the richest sub-sigma-field whose atoms form Π: it is the collection that consists of the empty set and all and only those $F \in \mathcal{F}$ which can be expressed as disjunctions of cells of Π. Intuitively, $\mathcal{I}(\Pi)$ has a kind of closure property: it contains all the consequences in $\mathcal{F}$ that are guaranteed to be deducible once the true cell $\Pi(\omega)$ is settled.

Using this correspondence, we can establish our earlier definitions as a special case:

Proposition 4. *Suppose $\mathcal{G} \subseteq \mathcal{F}$ is informed and its atoms form a partition $\Pi \subset \mathcal{F}$, so $\mathcal{G} = \mathcal{I}(\Pi)$. Then $P_{\mathcal{G}}(\cdot|\cdot)$ is a [proper, regular] conditional distribution for $P(\cdot)$ given $\mathcal{G}$ if and only if it is a [proper, regular] conditional distribution for $P(\cdot)$ given Π, as defined in the main text.*

We can define our notions of independence in this more general framework:

KCOND $F \in \mathcal{F}$ is independent of $\mathcal{G}(\omega)$ when $P_{\mathcal{G}}(F|\omega) = P(F)$.

KCOND-ENTIRE $F \in \mathcal{F}$ is independent of $\mathcal{G}$ when $P_{\mathcal{G}}(F|\omega) = P(F)$ for all ω. More generally: a sigma-field $\mathcal{H} \subseteq \mathcal{F}$ is independent of $\mathcal{G}$ if, for all $F \in \mathcal{H}$, $P_{\mathcal{G}}(F|\omega) = P(F)$ for all ω.

KCOND-ALMOST $F \in \mathcal{F}$ is almost surely independent of $\mathcal{G}$ when $P_{\mathcal{G}}(F|\omega) = P(F)$ for a probability one set of ω. More generally: $\mathcal{H}$ is almost surely independent of $\mathcal{G}$ if, for all $F \in \mathcal{H}$, $P_{\mathcal{G}}(F|\omega) = P(F)$ for a probability one set of ω.

KFACTOR $\mathcal{H}$ is independent of $\mathcal{G}$ if $P(F \cap G) = P(F) \cdot P(G)$ for all $F \in \mathcal{H}$ and $G \in \mathcal{G}$.

We can also generalize our main results to this framework. We see, for instance, that the generalized KCOND-ALMOST and KFACTOR are also extensionally equivalent.

Definition 6. A sub-sigma-field $\mathcal{G} \subseteq \mathcal{F}$ is *atomic* if contains each of its atoms.

Proposition 1 (general)**.** *Fix $P(\cdot)$ and an atomic sub-sigma-field $\mathcal{G} \subseteq \mathcal{F}$. Given the Axiom of Choice, if a regular conditional distribution $P_{\mathcal{G}}(\cdot|\cdot)$ is proper, then it can always be extended to a full coherent conditional probability function $P(\cdot|\cdot)$ on $\mathcal{F} \times \mathcal{F}^\circ$ such that for all ω, $P(\cdot|\mathcal{G}(\omega)) = P_{\mathcal{G}}(\cdot|\omega)$, and also $P(\cdot|\Omega) = P(\cdot)$.*

Proposition 2 (general)**.** *Suppose X and Y are real-valued random variables. Fix any regular conditional distribution $P_{\sigma Y}(\cdot|\cdot)$ for $P(\cdot)$ given σY. If X is almost surely independent of Y, then for a P-one set of ω, $P_{\sigma Y}(\cdot|\omega)$ and $P(\cdot)$ agree on σY.*

Proposition 3 (general)**.** ***K1-G*–*K4-G*** *imply* ***K1$^+$-G****.*

Theorem 1 (general)**.** *Fix a sigma-field $\mathcal{G} \subseteq \mathcal{F}$ and any conditional distribution $P_{\mathcal{G}}(\cdot|\cdot)$ for $P(\cdot)$ given $\mathcal{G}$. Then a sigma-field $\mathcal{H} \subseteq \mathcal{F}$ is almost surely independent of $\mathcal{G}$ if and only if $P(F \cap G) = P(F) \cdot P(G)$ for all $F \in \mathcal{H}$ and $G \in \mathcal{G}$.*

Proofs

This well-known theorem will be useful in what follows (Billingsley, 1986, p. 37):

Definition 7. Fix Ω. A *λ-system* is a collection of subsets of Ω which contains Ω and is closed under complement as well as countable union of disjoint sets. A *π-system* is a collection of subsets of Ω which is closed under finite intersections.

Theorem 2 (Dynkin π-λ theorem)**.** *If a π-system is included in a λ-system, then the smallest sigma-field generated by that π-system is also included in the λ-system.*

Propositions

Proof of Proposition 1. Let $\mathcal{I}(\Pi) \subseteq \mathcal{F}$ denote the set of segments of Π and let P_0 denote the restriction of $P(\cdot)$ to $\mathcal{I}(\Pi)$. Let P_1 denote $P_\Pi(\cdot|\cdot)$. Assume the Axiom of Choice. By Dubins (1975, Theorem 5) it suffices, for everything except the claim $P(\cdot|\Omega) = P(\cdot)$, to show that the pair (P_0, P_1) can be extended to what Dubins calls a *strategy on* $(\mathcal{F}, \Pi)$, which is a pair (σ_0, σ_1) such that σ_0 is a finitely additive probability measure defined on all subsets of Π, and σ_1 is a function on $\mathcal{F} \times \Pi$ such that for each $h \in \Pi$, $\sigma_1(\cdot|h)$ is a finitely additive probability measure satisfying $\sigma_1(h|h) = 1$. The extension of P_1 to σ_1 is straightforward, since we may simply set $\sigma_1(\cdot|h) = P_1(\cdot|\omega)$ for $\omega \in h$. By Π-constancy this is well-defined, by regularity $\sigma_1(\cdot|h)$ is a finitely additive probability measure, and by propriety $\sigma_1(h|h) = 1$. For the extension of P_0 to σ_0, we first define $\tilde{\Omega} = \Omega \backslash \sim$ where $\omega \sim \nu$ iff $\Pi(\omega) = \Pi(\nu)$. With $\varphi : \Omega \to \tilde{\Omega}$ the quotient map, we can then define $\sigma_0^*(\varphi[\,\cdot\,]) = P_0(\cdot)$. To extend σ_0^* to a finitely additive probability measure σ_0 on all subsets of Π, we invoke Horn and Tarski (1948,

Theorem 1.22), who prove that given the Axiom of Choice, any partial probability measure on a subset of a Boolean algebra can be extended to a finitely additive probability measure on the whole algebra. Finally, to check the claim that $P(\cdot|\Omega) = P(\cdot)$, note that in the extension Dubins employs (1975, Theorem 5, Equation 31), we have $P(F|\Omega) =_{def} \int_\Pi \sigma_1(F|h)\, d\sigma_0(h|\Omega) = \int_\Omega P_\Pi(F|\omega)\, dP_0(\omega) = \mathbb{E}[P_\Pi(F|\cdot)] = P(F)$ where here we invoke Π-constancy and the generalized law of total probability. □

Proof of Proposition 1 general. The proof proceeds as above, except with $\mathcal{I}(\Pi)$ replaced by $\mathcal{G}$, and with Π defined as the partition determined by $\mathcal{G}$. Note that if $\mathcal{G}$ is atomic, then for all ω, $\Pi(\omega) = \mathcal{G}(\omega) \in \mathcal{G}$, and so propriety implies $P_\mathcal{G}(\Pi(\omega)|\omega) = 1$. Thus, when σ_1 is defined in terms of $P_\mathcal{G}(\cdot|\cdot)$ as above, we obtain $\sigma_1(h|h) = 1$ for all $h \in \Pi$ as required. □

Proof of Proposition 2. Fix $\Pi = \{Y = y : y \in Y[\Omega]\}$ and any regular conditional distribution $P_\Pi(\cdot|\cdot)$ for $P(\cdot)$ given Π. It suffices to show there is a probability-one set $S \subseteq \Omega$ such that:

$$\sigma X \subseteq \mathcal{D} =_{def} \{F \in \mathcal{F} : P_\Pi(F|\omega) = P(F) \text{ for all } \omega \in S\}. \tag{C.4}$$

Let $H_1, H_2, ...$ denote the collection $\mathcal{C}$ of events of the form $X \geq r$ for rational r (this is possible since the rationals are countable). Now for each H_i, let S_i be a probability one set such that $P_\Pi(F|\omega) = P(F)$ for all $\omega \in S_i$, which exists by the supposition, and set $S = \cap_i S_i$. Since the i are countable, $P(S) = 1$. Now note that given this choice of S, $\mathcal{C} \subseteq \mathcal{D}$ by construction. Next note that $\mathcal{C}$ is a π-system, and the smallest sub-sigma field generated by this π-system is σX. Thus by the Dynkin π-λ theorem, it suffices to show $\mathcal{D}$ is a λ-system. This can be straightforwardly shown using regularity. □

Proof of Proposition 2 general. The proof proceeds as above, with Π replaced by σY. See also Ramachandran (1979, Lemma 1(ii)–(iii)). □

Proposition 3. Suppose G is a segment of Π. If $\omega \in G$, then $P_\Pi(G|\omega) = 1$ by **K4** and so $P_\Pi(F \cap G|\omega) = P_\Pi(F|\omega)$ by **K3**. If $\omega \notin G$ then $P_\Pi(G^c|\omega) = 1$ by **K4** so $P_\Pi(G|\omega) = 0$ and so $P_\Pi(F \cap G|\omega) = 0$. Thus $\int_\Omega P_\Pi(F \cap G|\omega)\, dP(\omega) = \int_G P_\Pi(F|\omega)\, dP(\omega)$. But by **K1**, $P(F \cap G) = \int_\Omega P_\Pi(F \cap G|\omega)\, dP(\omega)$, so **K1**$^+$ follows. In fact, all we need for this derivation is *almost propriety*, that $P_\Pi(\Pi(\omega)|\omega) = 1$ for a P-one set of ω. □

Proof of Proposition 3 general. The proof proceeds analogously to the above, except we begin by supposing G is a member of $\mathcal{G}$ (for the rest, replace instances of Π by $\mathcal{G}$). □

Proof of Proposition 4. We begin with the case of a proper, regular conditional distribution. Note that **K1** and **K1-G**, and **K3** and **K3-G**, are immediately equivalent, respectively. **K4-G** implies **K4** (take the case $G = \Pi(\omega) \in \mathcal{G}$), and **K4** implies **K4-G** since if $\omega \in G \in \mathcal{G}$, then $\Pi(\omega) \subseteq G$, so $P_\mathcal{G}(G|\omega) \geq P_\mathcal{G}(\Pi(\omega)|\omega) = 1$

where we invoke **K3-G**. It then remains to show that $P_{\mathcal{G}}(\cdot|\cdot)$ satisfies **K2** if and only if it satisfies **K2-G**. For the ($\Rightarrow$) direction, first note that $f(\cdot) = P_{\mathcal{G}}(F|\cdot)$ is $\mathcal{F}$-measurable by **K1**, otherwise its expectation with respect to P would not be defined. Letting $\mathcal{H}$ denote the sigma-field on Ω consisting of any arbitrary union of Π-cells, by **K2**, $f^{-1}(B) \in \mathcal{H}$ for all Borel $B \in \mathcal{B}(\mathbb{R})$. Since $\mathcal{G} = \mathcal{H} \cap \mathcal{F}$, we obtain **K2-G**. For the ($\Leftarrow$) direction, note that if $f = P_{\mathcal{G}}(F|\cdot)$ were to take two different values x and y over the same cell $\mathcal{G}(\omega)$, then $f^{-1}\{x\}$ would not be expressible as a union of cells, and so would not be measurable with respect to $\mathcal{G}$, since $\mathcal{G}$ contains only those $G \in \mathcal{F}$ that can be expressed this way. Thus **K2-G** implies **K2**. The equivalence for (regular) conditional distributions follows by similar reasoning, except in the step showing that **K1$^+$-G** implies **K1$^+$**, we also need to note that $\mathcal{G} = \mathcal{I}(\Pi)$ contains *all* segments $G \in \mathcal{F}$ of Π. □

Theorem

Below is the proof of Theorem 1. First let us provide a rough intuition for the 'if' direction, which might be the least obvious. Note that in many situations, for at least a probability one set of ω we can think of $P_\Pi(F|\omega)$ as a limit of ratios $P(F \cap S_n)/P(S_n)$, where the S_n are positive-probability segments of Π that converge to the cell $\Pi(\omega)$ in question. (For example, if Π corresponds to the possible values of a continuous real-valued random variable Y, we can let S_n be the event that Y is within $1/n$ of its precise value at ω.) Since they are segments, the factorization condition (7.6) implies that each of these ratios equals $P(F)$, which then implies that the limit equals $P(F)$ too.

Proof of Theorem 1. Suppose $P_\Pi(\cdot|\cdot)$ is a conditional distribution for $P(\cdot)$ given Π, so it satisfies the integral formula (**K1$^+$**) and Π-constancy (**K2**); see Appendix C. For the ($\Rightarrow$) direction, suppose Π' is almost surely independent of Π and let F and G be segments of Π' and Π respectively. By hypothesis, there is a probability one set of ω such that $P_\Pi(F|\omega) = P(F)$. Denote this set by S. Then

$$\begin{aligned} P(F \cap G) &= \int_G P_\Pi(F|\omega)\, dP(\omega) = \int_{S \cap G} P_\Pi(F|\omega)\, dP(\omega) = \int_{S \cap G} P(F)\, dP(\omega) \\ &= P(F) \cdot \int_{S \cap G} dP(\omega) = P(F) \cdot P(S \cap G) = P(F) \cap P(G), \end{aligned}$$

where the first equality follows from the integral formula, the third equality follows from the definition of S and Π-constancy, and the remaining follow from the fact $P(S) = 1$. For the ($\Leftarrow$) direction, fix F a segment of Π'. $D^+ = \{\omega \in \Omega : P_\Pi(F|\omega) > P(F)\}$ is a segment of Π (it belongs to $\mathcal{F}$ since by **K1$^+$**, $P_\Pi(F|\cdot)$ has well defined P-expected value and so is $\mathcal{F}$-measurable) and so by hypothesis $P(F \cap D^+) = P(F) \cdot P(D^+)$. Now

suppose toward contradiction that $P(D^+) > 0$. Then by the integral formula,

$$P(F \cap D^+) = \int_{D^+} P_\Pi(F|\omega)\, dP(\omega) > P(F) \cdot \int_{D^+} dP(\omega) = P(F) \cdot P(D^+),$$

which contradicts our hypothesis, so $P(D^+) = 0$. Applying the same reasoning to $D^- = \{\omega \in \Omega : P_\Pi(F|\omega) < P(F)\}$, we conclude $P\{\omega \in \Omega : P_\Pi(F|\omega) = P(F)\} = 1$, as desired. □

Proof of Theorem 1 general. The proof proceeds as above, but with $\mathcal{G}$-measurability invoked rather than Π-constancy. See also Kolmogorov (1933, Chapter VI.1), Ramachandran (1979, Lemma 1(ii)) and Çınlar (2011, Proposition IV.3.2, special case where $\mathcal{F} = \{\emptyset, \Omega\}$). □

Other claims

Proof of the claim in Section 7.3 on the existence of a distribution for the coin machine. First, some set-up. Let $\Omega = \{H, T\} \times [0, 1]$, $\mathcal{F}$ the Borel sigma-field on Ω, and $\Pi = \{\{H, T\} \times b : b \in [0, 1]\}$. We use the technique described in Appendix C. Define $\mu_\Pi(\cdot|\cdot)$ on $\mathcal{F} \times \Omega$ as follows. For each $\omega = (R, b)$ and interval $[a, c]$, we set:

$$\mu_\Pi(H \times [a, c]|\omega) = \begin{cases} b & \text{if } a \leq b \leq c \\ 0 & \text{otherwise} \end{cases}, \qquad \mu_\Pi(T \times [a, c]|\omega) = \begin{cases} 1 - b & \text{if } a \leq b \leq c \\ 0 & \text{otherwise} \end{cases}.$$

Since events of the form $H \times [a, c]$ and $T \times [a, c]$ generate $\mathcal{F}$, this uniquely fixes a probability measure $\mu_\Pi(\cdot|\omega)$ on $\mathcal{F}$ (Çınlar, 2011, I.3.19). Note that $\mu_\Pi(\cdot|\cdot)$ satisfies Π-constancy and regularity. To check propriety, we have, for $\omega = (R, b)$:

$$\mu_\Pi(\Pi(\omega)|\omega) = \mu_\Pi(\{H, T\} \times b|\omega) = \mu_\Pi(H \times b|\omega) + \mu_\Pi(T \times b|\omega) = b + 1 - b = 1,$$

where we use regularity and our definition above for the case $[a, c] = [b, b]$.

Now let $\bar{p}(\cdot)$ be the uniform distribution over $[0, 1]$, and let $\bar{\mu}_\Pi(\cdot|\cdot)$ denote the function on $\mathcal{F} \times [0, 1]$ given by $\bar{\mu}_\Pi(\cdot|b) = \mu_\Pi(\cdot|\omega)$ where $\omega = (R, b)$. If we can show that the expectation of $\bar{\mu}_\Pi(F|\cdot)$ with respect to $\bar{p}(\cdot)$ is well-defined for all $F \in \mathcal{F}$, then we can define $P(\cdot) =_{def} \int \bar{\mu}_\Pi(\cdot|b)\, d\bar{p}(b)$. It will then follow that $\mu_\Pi(\cdot|\cdot)$ is a proper, regular conditional distribution for $P(\cdot)$ given Π, and this distribution will satisfy $\mu_\Pi(H|\omega) = b$ for each $\omega = (R, b)$ as desired (where here H is shorthand for $H \times [0, 1]$).

To show the claim, we use the Dynkin π-λ theorem. We want to prove:

$$\mathcal{F} \subseteq \mathcal{D} =_{def} \{F \in \mathcal{F} : \text{the expectation of } \bar{\mu}_\Pi(F|\cdot) \text{ wrt } \bar{p}(\cdot) \text{ is defined}\}.$$

It is straightforward to check, using the fact that $\bar{\mu}_\Pi(\cdot|b)$ is a probability measure, that $\mathcal{D}$ is a λ-system. Note also that the collection $\mathcal{C}$ of events of the form $H \times [a,c]$ and $T \times [a,c]$ form a π-system that generates $\mathcal{F}$. Thus it suffices to show $\mathcal{C} \subseteq \mathcal{D}$. We have

$$\int \bar{\mu}_\Pi(H \times [a,c]|b)\, d\bar{p}(b) = \int 1_{[a,c]}(b) \cdot b \; db = \int_a^c b \, db = \frac{c^2 - a^2}{2},$$

which is clearly well-defined, and a similar result holds for $T \times [a,c]$. □

Proof of the claim in Section 7.4 on invariant distributions for the sphere. First, some set-up. Let $\Omega = S_1 \subseteq \mathbb{R}^3$ the unit sphere surface, $\mathcal{F}$ the Borel-sigma-field, and P the uniform measure given by, for each latitude-longitude rectangle $F_i = \{(\theta, \varphi) : \theta_1 \leq \theta \leq \theta_2, \varphi_1 \leq \varphi \leq \varphi_2\}$,

$$P(F_i) = \int_{\varphi_1}^{\varphi_2} \int_{\theta_1}^{\theta_2} \frac{\cos\theta}{4\pi} \, d\theta \, d\varphi.$$

(Note the F_i generate $\mathcal{F}$.) Set $\Pi = \{\Theta = \theta : -\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}\}$ the partition into parallels on this coordinate system (where $\Theta = \frac{\pi}{2}$ is the North Pole and $\Theta = -\frac{\pi}{2}$ is the South Pole).

Now we define $P_\Pi(\cdot|\cdot)$, as follows. If ω is not the North Pole or South Pole then for each latitude-longitude rectangle we set

$$P_\Pi(F_i|\omega) = \frac{L[F_i \cap \Pi(\omega)]}{L[\Pi(\omega)]},$$

where L denotes arclength. (This is well-defined: $\Pi(\omega)$ for $\omega = (\theta_0, \varphi_0)$ has length $2\pi\cos(\theta_0)$, which is greater than zero when $-\frac{\pi}{2} < \theta_0 < \frac{\pi}{2}$ i.e. when ω is not a pole, and $F_i \cap \Pi(\omega)$ is either the empty set, in which case its length is defined as zero, or an arc $\{(\theta, \varphi) : \theta = \theta_0, \varphi_1 \leq \varphi \leq \varphi_2\}$ in which case its length is $(\varphi_2 - \varphi_1)\cos(\theta_0)$.) Since the F_i generate $\mathcal{F}$, this fixes a unique countably probability measure $P_\Pi(\cdot|\omega)$ (Çınlar, 2011, I.3.19). When ω is the North or South Pole (so $\Pi(\omega) = \{\omega\}$), we set $P_\Pi(\cdot|\omega)$ to the Dirac measure,

$$P_\Pi(\cdot|\omega) = \delta_\omega(\cdot).$$

Next we show $P_\Pi(\cdot|\cdot)$ is a proper, regular conditional distribution for $P(\cdot)$ given Π. Π-constancy and regularity are immediate. For propriety, note that, fixing $\omega = (\theta_0, \varphi_0)$ not the North or South Pole, $\Pi(\omega)$ is a latitude-longitude rectangle F_i with $\theta_1 = \theta_2 = \theta_0$, and the intersection $F_i \cap \Pi(\omega)$ equals $\Pi(\omega)$, so $P_\Pi(\Pi(\omega)|\omega) = \frac{L[\Pi(\omega)]}{L[\Pi(\omega)]} = 1$. In the case ω is the North or South Pole, $P_\Pi(\Pi(\omega)|\omega) = \delta_\omega(\Pi(\omega)) = 1$ since $\omega \in \Pi(\omega)$. Finally for the generalized law of total probability we use the Dynkin π-λ theorem. It suffices to show

$$\mathcal{F} \subseteq \mathcal{D} =_{def} \{F \in \mathcal{F} : \mathbb{E}[P_\Pi(F|\cdot)] \text{ is defined and equals } P(F)\}.$$

It is straightforward to check, using regularity and the Monotone Convergence Theorem (Çınlar, 2011, Exercise I.5.6), that $\mathcal{D}$ is a λ-system. Moreover the collection of latitude-longitude rectangles $\mathcal{C}$ is a π-system that generates $\mathcal{F}$. Thus it suffices to show $\mathcal{C} \subseteq \mathcal{D}$. Fix $F_i = \{(\theta, \varphi) : \theta_1 \leq \theta \leq \theta_2, \varphi_1 \leq \varphi \leq \varphi_2\} \in \mathcal{C}$. Then one can check $P_\Pi(F_i|\cdot)$ is $\mathcal{F}$-measurable, and

$$\begin{aligned}\mathbb{E}[P_\Pi(F_i|\cdot)] &= \int_{\theta_1}^{\theta_2} P_\Pi(F_i|(\theta,\varphi))\, \frac{cos\theta}{2}\, d\theta = \int_{\theta_1}^{\theta_2} \frac{\varphi_2 - \varphi_1}{2\pi}\, \frac{\cos\theta}{2}\, d\theta \\ &= \int_{\varphi_1}^{\varphi_2} \int_{\theta_1}^{\theta_2} \frac{\cos\theta}{4\pi}\, d\theta\, d\varphi = P(F_i)\end{aligned}$$

as desired, where we can ignore North and South Pole terms since they receive zero weight from the element $d\theta$.

Finally we consider rotational invariance. Let h be a rotation. Let $P_{h\Pi}(\cdot|\cdot)$ be a proper, regular conditional distribution for $P(\cdot)$ given $h\Pi$, defined as above but with respect to the rotated latitude-longitude coordinate system. Now we want to show

$$\mathcal{F} \subseteq \mathcal{E} =_{def} \{F \in \mathcal{F} : P_\Pi(F|\omega) = P_{h\Pi}(hF|h\omega) \text{ for all } \omega\}.$$

Again using regularity it is straightforward to check that $\mathcal{E}$ is a λ-system, where we note $(hF)^c = h(F^c)$ and $\cup_i hF_i = h(\cup_i F_i)$ since h is a bijection. It then suffices to show $\mathcal{C} \subseteq \mathcal{E}$. Fix $F_i \in \mathcal{C}$. First, two background remarks:

- $(h\Pi)(h\omega) = h(\Pi(\omega))$ (to see this, observe $h\omega \in h(\Pi(\omega)) \in h\Pi$ since $\omega \in \Pi(\omega)$).
- $\Pi(\omega)$ is a pole in the original coordinate system iff $(h\Pi)(h\omega)$ is a pole in the rotated coordinate system (if $\Pi(\omega) = \{\omega\}$ then $(h\Pi)(h\omega) = h(\Pi(\omega)) = h\{\omega\} = \{h\omega\}$, and if $(h\Pi)(h\omega) = \{h\omega\}$ then $\Pi(\omega) = h^{-1}h(\Pi(\omega)) = h^{-1}\{h\omega\} = \{\omega\}$).

Now, consider an ω which is not a pole in the original coordinate system:

$$\begin{aligned}P_\Pi(F_i|\omega) &= \frac{L[F_i \cap \Pi(\omega)]}{L[\Pi(\omega)]} = \frac{L[h(F_i \cap \Pi(\omega))]}{L[h(\Pi(\omega))]} = \frac{L[hF_i \cap h(\Pi(\omega))]}{L[h(\Pi(\omega))]} \\ &= \frac{L[hF_i \cap (h\Pi)(h\omega)]}{L[(h\Pi)(h\omega)]} = P_{h\Pi}(hF_i|h\omega),\end{aligned}$$

where we use the fact that h is a rigid transformation and so preserves arclength L. Second we consider an ω which is a pole in the original coordinate system (and so $h\omega$ is a pole in the rotated coordinate system).

We then have:

$$P_{\Pi}(F_i|\omega) = \begin{cases} 1 & \text{if } \omega \in F_i \\ \\ 0 & \text{if } \omega \notin F_i \end{cases}, \qquad P_{h\Pi}(hF_i|h\omega) = \begin{cases} 1 & \text{if } h\omega \in hF_i \\ \\ 0 & \text{if } h\omega \notin hF_i \end{cases}.$$

To see that these coincide, note that if $\omega \in F_i$ then $h\omega \in hF_i$ and if $\omega \notin F_i$ then $\omega \in F_i^c$ so $h\omega \in h(F_i^c)$ so $h\omega \in (hF_i)^c$. Thus $P_{\Pi}(\cdot|\cdot) = P_{h\Pi}(h(\cdot)|h(\cdot))$ as desired. □

Bibliography

Aharonov, Y., J. Anandan, and L. Vaidman (1993). Meaning of the wave function. *Physical Review A 47*(6), 4616.

Albert, D. Z. (1992). *Quantum mechanics and experience*. Harvard University Press.

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

Albert, D. Z. (2015). *After physics*. Harvard University Press.

Albert, D. Z. and B. Loewer (1988). Interpreting the Many Worlds Interpretation. *Synthese 77*(2), 195–213.

Albert, D. Z. and B. Loewer (1990). Wanted dead or alive: Two attempts to solve Schrödinger's paradox. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Number 1, pp. 277–285. Philosophy of Science Association.

Alfsen, E. M. (2012). *Compact convex sets and boundary integrals*, Volume 57. Springer Science & Business Media.

Arnold, B. and C. Roberston (2002). The conditional distribution of X given X=Y can be almost anything! In N. Balakrishnan (Ed.), *Advances on Theoretical and Methodological Aspects of Probability and Statistics*, pp. 75–81. New York: Taylor and Francis.

Bae, J. and A. Acín (2006). Asymptotic quantum cloning is state estimation. *Physical Review Letters 97*(3), 030402.

Barnum, H., C. M. Caves, C. A. Fuchs, R. Jozsa, and B. Schumacher (1996). Noncommuting mixed states cannot be broadcast. *Physical Review Letters 76*(15), 2818.

Barrett, J. A. (1999). *The quantum mechanics of minds and worlds*. Oxford University Press.

Bass, T. A. (1985). *The eudaemonic pie*. Houghton Mifflin Harcourt.

Bell, J. S. (1964). On the Einstein Podolsky Rosen Paradox. *Physics Physique Fizika 1*(3), 195.

Bell, J. S. (1987). Beables for quantum field theory. In B. J. Hiley and D. Peat (Eds.), *Quantum Implications: Essays in Honour of David Bohm*, pp. 227–234. Routledge.

Belot, G. (2017). Curve-fitting for Bayesians? *The British Journal for the Philosophy of Science 68*(3), 689–702.

Benci, V., L. Horsten, and S. Wenmackers (2012). Axioms for non-archimedean probability (NAP). In D. V. J. and D. L. (Eds.), *Future Directions for Logic. Proceedings of PhDs in Logic III*, Volume 2. College Publications.

Bertsekas, D. and J. Tsitsiklis (2008). *Introduction to Probability*. Athena Scientific.

Billingsley, P. (1986). *Probability and measure*. John Wiley & Sons.

Blackwell, D. and L. E. Dubins (1975). On existence and non-existence of proper, regular, conditional distributions. *The Annals of Probability 3*(5), 741–752.

Blackwell, D. and C. Ryll-Nardzewski (1963). Non-existence of everywhere proper conditional distributions. *The Annals of Mathematical Statistics 34*(1), 223–225.

Bogachev, V. (2007). *Measure Theory*, Volume 2 of *Measure Theory*. Springer Berlin Heidelberg.

Borel, É. (1909). *Eléments de la théorie des probabilités*. Gauthier-Villars, Paris.

Bovens, L., S. Hartmann, et al. (2003). *Bayesian epistemology*. Oxford University Press.

Bub, J. (1977). Von neumann's projection postulate as a probability conditionalization rule in quantum mechanics. *Journal of Philosophical Logic 6*(1), 381–390.

Bub, J. (2005). Quantum mechanics is about quantum information. *Foundations of Physics 35*(4), 541–560.

Builes, D. (2020). A paradox of evidential equivalence. *Mind 129*(513), 113–127.

Builes, D. and J. Spencer (ms). Why chance constrains credence. Manuscript.

Busch, P. (1997). Is the quantum state (an) observable? In R. S. Cohen, M. Horne, and J. Stachel (Eds.), *Potentiality, Entanglement and Passion-at-a-Distance*, pp. 61–70. Springer.

Busch, P., P. J. Lahti, and P. Mittelstaedt (1996). *The quantum theory of measurement*. Springer.

Busch, P. and J. Singh (1998). Lüders theorem for unsharp quantum measurements. *Physics Letters A 249*(1-2), 10–12.

Butterfield, J. (2012). Laws, causation and dynamics at different levels. *Interface Focus 2*(1), 101–114.

Butterfield, J. (2015). Peaceful coexistence: Examining Kent's relativistic solution to the quantum measurement problem. In *Nagoya Winter Workshop: Reality and Measurement in Algebraic Quantum Theory*, pp. 277–314. Springer.

Butterfield, J. and C. Isham (2001). Spacetime and the philosophical challenge of quantum gravity. In C. Callender and N. Huggett (Eds.), *Physics meets philosophy at the Planck scale*, pp. 33–89. Cambridge University Press.

Callender, C. and J. Cohen (2009). A better best system account of lawhood. *Philosophical Studies 145*(1), 1–34.

Callender, C. and J. Cohen (2010). Special sciences, conspiracy and the better best system account of lawhood. *Erkenntnis 73*(3), 427–447.

Caves, C. M., C. A. Fuchs, and R. Schack (2002). Quantum probabilities as Bayesian probabilities. *Physical Review A 65*(2), 022305.

Chen, H., D. Lu, B. Chong, G. Qin, X. Zhou, X. Peng, and J. Du (2011). Experimental demonstration of probabilistic quantum cloning. *Physical Review Letters 106*(18), 180404.

Childers, T. (2012). Objectifying subjective probabilities: Dutch book arguments for principles of direct inference. In D. Dieks, W. Gonzalez, S. Hartmann, M. Stoltzner, and M. Weber (Eds.), *Probabilities, Laws, and Structures*. Springer.

Çınlar, E. (2011). *Probability and stochastics*, Volume 261. Springer.

Coletti, G. and R. Scozzafava (2002). *Probabilistic Logic in a Coherent Setting*. Trends in Logic. Springer Netherlands.

Colin, S. (2003). A deterministic Bell model. *Physics Letters A 317*(5-6), 349–358.

Combes, J., C. Ferrie, M. S. Leifer, and M. F. Pusey (2018). Why protective measurement does not establish the reality of the quantum state. *Quantum Studies: Mathematics and Foundations 5*(2), 189–211.

Cramér, H. and M. R. Leadbetter (2013). *Stationary and related stochastic processes: Sample function properties and their applications*. Courier Corporation.

Dajani, K. and S. Dirksin (2008). A simple introduction to ergodic theory. University of Utrecht, Lecture notes in Ergodic Theory.

D'Ariano, G. M. and H. P. Yuen (1996). Impossibility of measuring the wave function of a single quantum system. *Phys. Rev. Lett. 76*, 2832–2835.

DeFinetti, B. (1974). *Theory of Probability: A Critical Introductory Treatment*. Wiley, London.

Dekking, F. M., C. Kraaikamp, H. P. Lopuhaä, and L. E. Meester (2005). *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer Science & Business Media.

Diaconis, P. and D. Freedman (1984). Partial exchangeability and sufficiency. In *Proc. Indian Statist. Inst. Golden Jubilee Int'l Conf. on Statistics: Applications and New Directions*, pp. 205–236. Calcutta: Indian Statistical Institute.

Diaconis, P., S. Holmes, and R. Montgomery (2007). Dynamical bias in the coin toss. *SIAM Review 49*(2), 211–235.

Dickson, M. (1995). An empirical reply to empiricism: Protective measurement opens the door for quantum realism. *Philosophy of Science 62*(1), 122–140.

Doob, J. (1948). On a problem of Marczewski. In *Colloquium Mathematicum*, Volume 1, pp. 216–217. Instytut Matematyczny Polskiej Akademii Nauk.

Dowe, D. L., S. Gardner, and G. Oppy (2007). Bayes not bust! Why simplicity is no problem for Bayesians. *The British Journal for the Philosophy of Science 58*(4), 709–754.

Duan, L.-M. and G.-C. Guo (1998a). Probabilistic cloning and identification of linearly independent quantum states. *Physical Review Letters 80*(22), 4999.

Duan, L.-M. and G.-C. Guo (1998b). A probabilistic cloning machine for replicating two non-orthogonal states. *Physics Letters A 243*(5-6), 261–264.

Dubins, L. E. (1975). Finitely additive conditional probabilities, conglomerability and disintegrations. *The Annals of Probability 3*(1), 89–99.

Dürr, D., S. Goldstein, R. Tumulka, and N. Zanghí (2004). Bohmian mechanics and quantum field theory. *Physical Review Letters 93*(9), 090402.

Dürr, D., S. Goldstein, R. Tumulka, and N. Zanghí (2005). On the role of density matrices in Bohmian mechanics. *Foundations of Physics 35*(3), 449–467.

Dürr, D., S. Goldstein, and N. Zanghí (1992). Quantum equilibrium and the origin of absolute uncertainty. *Journal of Statistical Physics 67*(5-6), 843–907.

Earman, J. (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*. MIT Press.

Earman, J. (2018). The relation between credence and chance: Lewis' "Principal Principle" is a theorem of quantum probability theory. Manuscript. Available at: `http://philsci-archive.pitt.edu/14822/`.

Easwaran, K. (2019). Conditional probabilities. In R. Pettigrew and J. Weisberg (Eds.), *The Open Handbook of Formal Epistemology*, pp. 131–198. PhilPapers Foundation.

Easwaran, K. K. (2008). *The foundations of conditional probability*. University of California, Berkeley.

Easwaran, K. K. (2014). Regularity and hyperreal credences. *Philosophical Review 123*(1), 1–41.

Elga, A. (2005). Isolation and folk physics. In H. Price and R. Corry (Eds.), *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*. Oxford University Press.

Feller, W. (1950). *An Introduction to Probability Theory and Its Applications*, Volume 2. Wiley.

Fitelson, B. and A. Hájek (2017). Declarations of independence. *Synthese 194*(10), 3979–3995.

Forster, M. R. (1995). Bayes and bust: Simplicity as a problem for a probabilist's approach to confirmation. *The British Journal for the Philosophy of Science 46*(3), 399–424.

Fröhlich, J. and B. Schubnel (2016). The preparation of states in quantum mechanics. *Journal of Mathematical Physics 57*(4), 042101.

Fuchs, C. A. and R. Schack (2004). Unknown quantum states and operations, a Bayesian view. In M. Paris and J. Rehacek (Eds.), *Quantum state estimation*, pp. 147–187. Springer.

Gallow, J. D. (2019). A subjectivist's guide to deterministic chance. *Synthese*. Published online.

Gao, S. (2015). *Protective Measurement and Quantum Reality*. Cambridge University Press.

Gao, S. (2019). The measurement problem revisited. *Synthese 196*(1), 299–311.

Ghirardi, G. C., A. Rimini, and T. Weber (1986). Unified dynamics for microscopic and macroscopic systems. *Physical Review D 34*(2), 470.

Glymour, C. (2010). Why I am not a Bayesian. In A. Eagle (Ed.), *Philosophy of Probability: Contemporary Readings*. Routledge.

Grimmett, G., G. R. Grimmett, D. Stirzaker, et al. (2001). *Probability and random processes*. Oxford University Press.

Gyenis, Z., G. Hofer-Szabó, and M. Rédei (2017). Conditioning using conditional expectations: The Borel–Kolmogorov paradox. *Synthese 194*(7), 2595–2630.

Hájek, A. (2003). What conditional probability could not be. *Synthese 137*(3), 273–323.

Hájek, A. (2007). The reference class problem is your problem too. *Synthese 156*(3), 563–585.

Hall, N. (1994). Correcting the guide to objective chance. *Mind 103*(412), 505–517.

Handfield, T. and A. Wilson (2014). Chance and context. In A. Wilson (Ed.), *Chance and Temporal Asymmetry*. Oxford University Press.

Hayashi, A., T. Hashimoto, and M. Horibe (2005). Reexamination of optimal quantum state estimation of pure states. *Physical Review A 72*(3), 032325.

Healey, R. (2017). *The Quantum Revolution in Philosophy*. Oxford University Press.

Hemmo, M. and O. Shenker (Eds.) (2020). *Quantum, Probability, Logic: The Work and Influence of Itamar Pitowsky*. Jerusalem Studies in Philosophy and History of Science. Springer International Publishing.

Hicks, M. T. (2017). Making fit fit. *Philosophy of Science 84*(5), 931–943.

Hoefer, C. (2007). The third way on objective probability: A sceptic's guide to objective chance. *Mind 116*(463), 549–596.

Hoefer, C. (2019). *Chance in the World: A Humean Guide to Objective Chance*. Oxford University Press.

Hooft, G. (1999). Quantum gravity as a dissipative deterministic system. *Classical and Quantum Gravity 16*(10), 3263.

Horn, A. and A. Tarski (1948). Measures in Boolean algebras. *Transactions of the American Mathematical Society 64*(3), 467–497.

Howson, C. and P. Urbach (2006). *Scientific reasoning: The Bayesian approach*. Open Court Publishing.

Hughes, R. I. (1989). *The structure and interpretation of quantum mechanics*. Harvard university press.

Jarrett, J. P. (1984). On the physical significance of the locality conditions in the Bell arguments. *Noûs 18*(4), 569–589.

Jaynes, E. T. (2003). *Probability theory: The logic of science*. Cambridge University Press.

Jeffrey, R. (1965). *The Logic of Decision*. University of Chicago Press.

Kac, M. and D. Slepian (1959). Large excursions of Gaussian processes. *The Annals of Mathematical Statistics 30*(4), 1215–1228.

Kadane, J. B., M. J. Schervish, and T. Seidenfeld (1999). Statistical implications of finitely additive probability. In J. B. Kadane, M. J. Schervish, and T. Seidenfeld (Eds.), *Rethinking the Foundations of Statistics*, pp. 211–232. Cambridge University Press.

Kent, A. (2014). Solution to the Lorentzian quantum reality problem. *Physical Review A 90*(1), 012107.

Kolmogorov, A. N. (1933). *Foundations of the Theory of Probability (Grundbegriffe der Wahrscheinlichkeitsrechnung)*. Springer, Berlin.

Lauritzen, S. (1988). *Extremal Families and Systems of Sufficient Statistics*. Springer.

Lewis, D. K. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in Inductive Logic and Probability*, pp. 83–132. University of California Press.

Lewis, D. K. (1987). *Philosophical Papers: Volume II*. Oxford University Press.

Lewis, D. K. (1988a). Relevant implication. *Theoria 54*(3), 161–174.

Lewis, D. K. (1988b). Statements partly about observation. *Philosophical Papers 17*(1), 1–31.

Lewis, D. K. (1994). Humean supervenience debugged. *Mind 103*(412), 473–490.

Lewis, P. J. (2007). How Bohm's theory solves the measurement problem. *Philosophy of Science 74*(5), 749–760.

Lindley, D. (1982). The Bayesian approach to statistics. In J. de Oliveria and B. Epstein (Eds.), *Some advances in statistics*. New York: Academic Press.

List, C. and M. Pivato (2015). Emergent chance. *Philosophical Review 124*(1), 119–152.

Loewer, B. (1994). Probability and quantum theory. In P. Suppes and P. Humphreys (Eds.), *Patrick Suppes: Scientific Philosopher: Volume 2. Philosophy of Physics, Theory Structure, and Measurement Theory*, pp. 3–28. Springer Netherlands.

Loewer, B. (2020). The Mentaculus vision. In V. Allori (Ed.), *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism, and Laws of Nature*. World Scientific.

Maher, P. (2009). Physical probability. In C. Glymour, W. Wang, and D. Westerstahl (Eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the Thirteenth International Congress*, pp. 193–210. College Publications.

Margenau, H. (1963). Measurements and quantum states: Part I. *Philosophy of Science 30*(1), 1–16.

Maudlin, T. (1995). Three measurement problems. *Topoi 14*(1), 7–15.

Maudlin, T. (2002). *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics*. Blackwell.

Meacham, C. (2009). Two mistakes regarding the Principal Principle. *British Journal for the Philosophy of Science 61*(2), 407–431.

Meacham, C. J. G. (2014). Autonomous chances and the conflicts problem. In A. Wilson (Ed.), *Asymmetries in Chance and Time*, pp. 45–67. Oxford University Press.

Meacham, C. J. G. (2016). Ur-priors, conditionalization, and ur-prior conditionalization. *Ergo: An Open Access Journal of Philosophy 3*(17), 444–492.

Meehan, A. (2020). You say you want a revolution: Two notions of probabilistic independence. Manuscript. Accepted to *Philosophical Studies*.

Meehan, A. (forthcoming). A new problem for quantum mechanics. *The British Journal for the Philosophy of Science*. DOI: 10.1093/bjps/axz053.

Meehan, A. and S. Zhang (2020a). Jeffrey meets Kolmogorov. *Journal of Philosophical Logic 49*, 941–979.

Meehan, A. and S. Zhang (2020b). Revisiting the existence and non-existence of proper, regular conditional distributions. Manuscript. Available upon request.

Meehan, A. and S. Zhang (forthcoming-a). The Borel-Kolmogorov paradox is your paradox too. *Philosophy of Science*.

Meehan, A. and S. Zhang (forthcoming-b). Kolmogorov conditionalizers can be Dutch booked (if and only if they are evidentially uncertain). *Review of Symbolic Logic*.

Meester, R. (2008). *A natural introduction to probability theory*. Springer Science & Business Media.

Miszczak, J., Z. Puchała, P. Horodecki, A. Uhlmann, and K. Życzkowski (2009). Sub– and super–fidelity as bounds for quantum fidelity. *Quantum Information and Computation 9*(1-2), 103–130.

Myrvold, W. C. (2002). On peaceful coexistence: Is the collapse postulate incompatible with relativity? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 33*(3), 435–466.

Myrvold, W. C. (2015). You can't always get what you want: Some considerations regarding conditional probabilities. *Erkenntnis 80*(3), 573–603.

Myrvold, W. C. (2018). Philosophical issues in quantum theory. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2018 ed.). Metaphysics Research Lab, Stanford University.

Naimark, M. A. (1940). Spectral functions of a symmetric operator. In *Izv. Akad. Nauk SSSR Set. Mat.*, Volume 4, pp. 277–318.

Naimark, M. A. (1943). On a representation of additive operator set functions. In *C. R. Acad. Sci. URSS*, Volume 41, pp. 359–361.

Nielsen, M. A. and I. Chuang (2002). *Quantum computation and quantum information*. Cambridge University Press.

Nikodym, O. (1930). Sur une généralisation des intégrales de MJ Radon. *Fundamenta Mathematicae 15*(1), 131–179.

Norsen, T. (2016). Bohmian conditional wave functions (and the status of the quantum state). In *Journal of Physics: Conference Series*, Volume 701, pp. 012003. IOP Publishing.

Page, D. N. (1982). Is quantum gravity deterministic and/or time symmetric? *General Relativity and Gravitation 14*(3), 299–302.

Paris, M. and J. Rehacek (2004). *Quantum State Estimation*. Lecture Notes in Physics. Springer Berlin Heidelberg.

Park, J. L. and W. Band (1992). Preparation and measurement in quantum physics. *Foundations of Physics 22*(5), 657–668.

Pati, A. K. (2000). Probabilistic exact cloning and probabilistic no-signalling. *Physics Letters A 270*(3-4), 103–107.

Peres, A. (2003). How the no-cloning theorem got its name. *Fortschritte der Physik: Progress of Physics 51*(4-5), 458–461.

Peres, A. (2006). *Quantum theory: Concepts and methods*, Volume 57. Springer Science & Business Media.

Petersen, K. E. (1989). *Ergodic theory*, Volume 2. Cambridge University Press.

Pettigrew, R. (2012). Accuracy, chance, and the Principal Principle. *Philosophical Review 121*(2), 241–275.

Pettigrew, R. (2013). A new epistemic utility argument for the Principal Principle. *Episteme 10*(1), 19–35.

Pishro-Nik, H. (2014). *Introduction to Probability, Statistics, and Random Processes*. Kappa Research, LLC.

Pitowsky, I. (1996). Laplace's demon consults an oracle: The computational complexity of prediction. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics 27*(2), 161–180.

Pitowsky, I. (2006). Quantum mechanics as a theory of probability. In *Physical theory and its interpretation*, pp. 213–240. Springer.

Popper, K. R. (1955). Two autonomous axiom systems for the calculus of probabilities. *The British Journal for the Philosophy of Science 6*(21), 51–57.

Popper, K. R. (1959). *The Logic of Scientific Discovery*. Routledge.

Pruss, A. R. (2013). Null probability, dominance and rotation. *Analysis 73*(4), 682–685.

Pruss, A. R. (2015). Popper functions, uniform distributions and infinite sequences of heads. *Journal of Philosophical Logic 44*(3), 259–271.

Ramachandran, D. (1975). On the two definitions of independence. In *Colloquium Mathematicum*, Volume 32, pp. 227–231.

Ramachandran, D. (1979). Existence of independent complements in regular conditional probability spaces. *The Annals of Probability 7*(3), 433–443.

Rao, M. (1988). Paradoxes in conditional probability. *Journal of Multivariate Analysis 27*(2), 434–446.

Rao, M. M. and R. J. Swift (2006). *Probability theory with applications*, Volume 582. Springer Science & Business Media.

Rédei, M. and S. J. Summers (2007). Quantum probability theory. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 38*(2), 390–417.

Rényi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Hungarica 6*(3-4), 285–335.

Rescorla, M. (2015). Some epistemological ramifications of the Borel–Kolmogorov paradox. *Synthese 192*(3), 735–767.

Rescorla, M. (2018). A Dutch book theorem and converse Dutch book theorem for Kolmogorov conditionalization. *The Review of Symbolic Logic 11*(4), 705–735.

Rivas, A. and S. F. Huelga (2012). *Open Quantum Systems*. Springer.

Roeper, P. and H. Leblanc (1999). *Probability theory and probability logic*. University of Toronto Press.

Ross, S. (2010). *A First Course in Probability*. Pearson Prentice Hall.

Rovelli, C. (1996). Relational quantum mechanics. *International Journal of Theoretical Physics 35*(8), 1637–1678.

Rubenstein, E. (ms). The modal account of probability. Manuscript.

Ruetsche, L. (1996). Van Fraassen on preparation and measurement. *Philosophy of Science 63*, S338–S346.

Ruetsche, L. (2004). Intrinsically mixed states: an appreciation. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 35*(2), 221–239.

Ruetsche, L. (2011). Why be normal? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 42*(2), 107–115.

Ruetsche, L. and J. Earman (2011). Interpreting probabilities in quantum field theory and quantum statistical mechanics. In S. Hartmann and C. Beisbart (Eds.), *Probabilities in Physics*. Oxford University Press.

Sauve, A. C., A. Hero, W. L. Rogers, S. Wilderman, and N. Clinthorne (1999). 3d image reconstruction for a compton spect camera model. *IEEE Transactions on Nuclear Science 46*(6), 2075–2084.

Scarani, V., S. Iblisdir, N. Gisin, and A. Acin (2005). Quantum cloning. *Reviews of Modern Physics 77*(4), 1225.

Schwarz, W. (2014). Proving the Principal Principle. In A. Wilson (Ed.), *Chance and Temporal Asymmetry*. Oxford University Press.

Scott, A. J. (2006). Tight informationally complete quantum measurements. *Journal of Physics A: Mathematical and General 39*(43), 13507.

Sebens, C. T. (2015). Quantum mechanics as classical physics. *Philosophy of Science 82*(2), 266–291.

Seidenfeld, T. (2001). Remarks on the theory of conditional probability: Some issues of finite versus countable additivity. In V. F. Hendricks, S. A. Pederson, and K. F. Jørgensen (Eds.), *Probability Theory: Philosophy, Recent History and Relations to Science*. Kluwer.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (2001). Improper regular conditional distributions. *Annals of Probability 29*(4), 1612–1624.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (2013). Two theories of conditional probability and non-conglomerability. In *Eighth International Symposium on Imprecise Probability: Theory and Applications, July*, pp. 2–5.

Shafer, G. and V. Vovk (2006). The sources of Kolmogorov's "Grundbegriffe". *Statistical Science 21*(1), 70–98.

Shimony, A. (1986). Events and processes in the quantum world. In R. Penrose and C. J. Isham (Eds.), *Quantum Concepts in Space and Time*, pp. 182–203. Oxford University Press.

Sklar, L. (2015). Philosophy of statistical mechanics. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2015 ed.). Metaphysics Research Lab, Stanford University.

Skyrms, B. (1977). Resiliency, propensities, and causal necessity. *The Journal of Philosophy 74*(11), 704–713.

Stairs, A. (1983). Quantum logic, realism, and value definiteness. *Philosophy of Science 50*(4), 578–602.

Strevens, M. (1998). Inferring probabilities from symmetries. *Noûs 32*(2), 231–246.

Strevens, M. (1999). Objective probability as a guide to the world. *Philosophical Studies 95*(3), 243–275.

Strevens, M. (2003). *Bigger than chaos: Understanding complexity through probability*. Harvard University Press.

Strevens, M. (2011). Probability out of determinism. In C. Beisbart and S. Hartmann (Eds.), *Probabilities in Physics*, pp. 339–364. Oxford University Press.

Strevens, M. (2013). *Tychomancy*. Harvard University Press.

Strocchi, F. (2008). *An introduction to the mathematical structure of quantum mechanics: a short course for mathematicians*, Volume 28. World Scientific.

Thau, M. (1994). Undermining and admissibility. *Mind 103*(412), 491–503.

Unruh, W. (1994). Reality and measurement of the wave function. *Physical Review A 50*(1), 882.

Vaidman, L. (2015). Protective measurement of the wave function of a single system. In S. Gao (Ed.), *Protective Measurement and Quantum Reality*, pp. 15–26. Cambridge University Press.

von Neumann, J. (1955). *Mathematical Foudations of Quantum Mechanics*. Princeton University Press.

Von Plato, J. (1983). The method of arbitrary functions. *The British Journal for the Philosophy of Science 34*(1), 37–47.

Wallace, D. (2012). *The emergent multiverse: Quantum theory according to the Everett interpretation.* Oxford University Press.

Ward, B. (2005). Projecting chances: A Humean vindication and justification of the Principal Principle. *Philosophy of Science 72*(1), 241–261.

Weinstein, S. (1996). Undermind. *Synthese 106*(2), 241–251.

Werndl, C. (2009). Are deterministic descriptions and indeterministic descriptions observationally equivalent? *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics 40*(3), 232–242.

Werndl, C. (2012). Evidence for the deterministic or the indeterministic description? A critique of the literature about classical dynamical systems. *Journal for General Philosophy of Science 43*(2), 295–312.

Werndl, C. (2013). On choosing between deterministic and indeterministic models: Underdetermination and indirect evidence. *Synthese 190*(12), 2243–2265.

Werner, R. F. (1998). Optimal cloning of pure states. *Physical Review A 58*(3), 1827.

Wessels, L. (1997). The preparation problem in quantum mechanics. In J. Earman and J. Norton (Eds.), *The Cosmos of Science*, pp. 243–273. University of Pittsburgh Press.

White, R. (2005). Why favour simplicity? *Analysis 65*(3), 205–210.

Wilhelm, I. (2020). Centering the Principal Principle. *Philosophical Studies*. Published online.

Williamson, T. (2007). How probable is an infinite sequence of heads? *Analysis 67*(3), 173–180.

Wootters, W. K. and W. H. Zurek (1982). A single quantum cannot be cloned. *Nature 299*(5886), 802–803.