

Michael Janas
Michael E. Cuffaro
Michel Janssen

Understanding Quantum Raffles

Quantum Mechanics on an
Informational Approach:
Structure and Interpretation

With a Foreword by Jeffrey Bub



Springer

Boston Studies in the Philosophy and History of Science

Volume 340

Editors

Alisa Bokulich, Boston University

Jürgen Renn, Max Planck Institute for the History of Science

Managing Editor

Lindy Divarci, Max Planck Institute for the History of Science

Editorial Board

Theodore Arabatzis, University of Athens

Heather E. Douglas, University of Waterloo

Jean Gayon, Université Paris 1

Thomas F. Glick, Boston University

Hubert Goenner, University of Goettingen

John Heilbron, University of California, Berkeley

Diana Kormos-Buchwald, California Institute of Technology

Christoph Lehner, Max Planck Institute for the History of Science

Peter McLaughlin, Universität Heidelberg

Agustí Nieto-Galan, Universitat Autònoma de Barcelona

Nuccio Ordine, Università della Calabria

Sylvan S. Schweber, Harvard University

Ana Simões, Universidade de Lisboa

John J. Stachel, Boston University

Baichun Zhang, Chinese Academy of Science

Kostas Gavroglu, National and Kapodistrian University of Athens

The series *Boston Studies in the Philosophy and History of Science* was conceived in the broadest framework of interdisciplinary and international concerns. Natural scientists, mathematicians, social scientists and philosophers have contributed to the series, as have historians and sociologists of science, linguists, psychologists, physicians, and literary critics.

The series has been able to include works by authors from many other countries around the world.

The editors believe that the history and philosophy of science should itself be scientific, self-consciously critical, humane as well as rational, sceptical and undogmatic while also receptive to discussion of first principles. One of the aims of Boston Studies, therefore, is to develop collaboration among scientists, historians and philosophers.

Boston Studies in the Philosophy and History of Science looks into and reflects on interactions between epistemological and historical dimensions in an effort to understand the scientific enterprise from every viewpoint.

More information about this series at <https://link.springer.com/bookseries/5710>

Michael Janas • Michael E. Cuffaro
Michel Janssen

Understanding Quantum Raffles

Quantum Mechanics on an Informational
Approach: Structure and Interpretation

Foreword by Jeffrey Bub

 Springer

Michael Janas
School of Physics and Astronomy
University of Minnesota
Minneapolis, MN, USA

Michael E. Cuffaro 
Fakultät für Philosophie
Ludwig Maximilian
University of Munich
München, Bayern, Germany

Michel Janssen
School of Physics and Astronomy
University of Minnesota
Minneapolis, MN, USA

ISSN 0068-0346 ISSN 2214-7942 (electronic)
Boston Studies in the Philosophy and History of Science
ISBN 978-3-030-85938-1 ISBN 978-3-030-85939-8 (eBook)
<https://doi.org/10.1007/978-3-030-85939-8>

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2022

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

We dedicate this volume to the memory of Itamar Pitowsky (1950–2010) and William Demopoulos (1943–2017).

In addition, Cuffaro dedicates his efforts to the memory of Giuseppe, his father (1932–2019) and Janssen dedicates his efforts to the memory of Heinrich, his father (1929–2015).

Foreword

Understanding Quantum Raffles was inspired by *Bananaworld*, as the authors say, but it is very much more than that. My initial aim in writing *Bananaworld* was to de-mystify quantum entanglement for non-physicists—as Schrödinger remarked, ‘*the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought.*’ I wanted to show that entanglement is essentially a new sort of nonlocal correlation, explain why it is puzzling, and point out how it can be used as a resource. The device I used to exhibit entanglement was the Popescu-Rohrlich nonlocal box, or PR-box, which I dramatized as a pair of bananas that each acquires one of two possible tastes when peeled in one of two allowable ways, from the stem end or the top end. The PR-box correlation is a superquantum correlation but can be expressed quite simply, without the mathematical machinery of quantum mechanics. It has all the puzzling features of quantum entanglement and, with a little poetic license, can even be exploited to show how entanglement works to enable feats like quantum teleportation, unconditional security in quantum cryptography, and apparently exponential speed-up in quantum computation.

In spite of the bananas, the book did not turn out to be the sort of thing you could pick up and enjoy over a beer. So I wrote *Totally Random: Why Nobody Understands Quantum Mechanics* with my daughter, Tanya Bub. *Totally Random* deals with some of the topics discussed in *Bananaworld*, but in a way that’s much more accessible and, we hoped, fun to read. We presented the book as ‘a serious comic on entanglement’—serious because we felt that the general reader could come away with a real understanding of entanglement: what it is, what the patriarchs of quantum mechanics have said about it, and what you can do with it. The authors of *Understanding Quantum Raffles*—the three Mikes—have evidently also given a great deal of thought to pedagogical issues. While some of the discussion, notably Chapter 4, tackles advanced material, a major part of the book, especially Chapters 2 and 3, is clearly intended for the general reader, so if you want to understand what is really new and interesting about quantum mechanics, this is the book to read.

In *Bananaworld*, I brought out the difference between classical and quantum mechanics by considering to what extent it is possible to simulate a PR-box correlation with various resources, classical or quantum. Bell's nonlocality proof amounts to a demonstration that two separated agents, Alice and Bob, restricted to classical, and so local resources (effectively what computer scientists call 'shared randomness'), can achieve an optimal success rate of no more than 75%. If Alice and Bob are allowed to use quantum resources, entangled pairs of photons or electrons, they can do better, about 85%. Equipped with PR-boxes, they can, of course, achieve a 100% success rate. Another way to put this is in terms of the Clauser-Horne-Shimony-Holt (CHSH) inequality for two bivalent Alice-observables and two bivalent Bob-observables. The CHSH correlation for the four pairs of observables is constrained to values between -2 and 2 for local classical correlations, between $-2\sqrt{2}$ and $2\sqrt{2}$ for quantum correlations, and between -4 and 4 for PR-box correlations, which are maximal for correlations that do not allow instantaneous signaling. Geometrically, as Pitowsky showed,¹ the classical or local correlations for this case can be represented by the points in an 8-dimensional polytope with facets characterized by the CHSH inequality and similar inequalities, the quantum correlations by the points in a convex set that includes the polytope, and the no-signaling correlations by a polytope that includes the quantum convex set.

The three Mikes do something brilliantly different. Instead of the CHSH inequality, they consider the Mermin inequality for three bivalent observables for each agent. In terms of bananas, Alice and Bob peel their bananas in one of three possible ways associated with three directions in which they are required to hold their bananas while peeling. This complication, which I blush to admit I first thought was pointless, results in a tetrahedron for the classical or local correlations, an ellipsope for the quantum convex set (a 'fat' tetrahedron that includes the classical tetrahedron), and a cube for the no-signaling correlations—easily visualizable in three dimensions. The three Mikes produce two derivations for the non-linear inequality characterizing the ellipsope: a derivation 'from within' quantum mechanics, which uses the Born rule for probabilities, and a derivation 'from without,' which follows work by Yule in the late 19th century on Pearson correlation coefficients. In Yule's derivation, the inequality is a general constraint on correlations between three random variables. In the 'proof from without,' the random variables are the eigenvalues of Hilbert space operators representing observables and the

¹ I. Pitowsky, 'On the geometry of quantum correlations,' *Physical Review A* 77, 062109 (2008).

constraint follows quite generally, without assuming the Born rule for quantum probabilities.

The Mermin inequality refers to spin-1/2 particles in the singlet state. Remarkably, it turns out that singlet state quantum correlations are confined to the ellipsope even for higher spin values, while the tetrahedron for local classical correlations is replaced by a succession of polyhedra with more and more facets for higher spins, approaching the ellipsope in the limit of infinite spin. All this is beautifully illustrated in 3-dimensional visualizations. The analysis is particularly impressive because it shows clearly and precisely how classical and quantum correlations are related in this particular case.

This is certainly the first book in which the word ‘Bubism’ appears. The three Mikes use the term to refer to ‘an interpretation of quantum mechanics along the lines of *Bananaworld*, belonging to the same lineage, or so we will argue, as the much-maligned Copenhagen interpretation.’ *Bananaworld* began as a discussion of entanglement, but as I wrote the book it evolved into a way of thinking about the transition from classical to quantum mechanics. The three Mikes have taken this perspective and articulated and developed it into an interpretation that I fully endorse but which owes as much to their careful analysis of the conceptual issues as my own thinking.

I added the last chapter to *Bananaworld*, ‘Making Sense of it All,’ because I thought I should say something about the measurement problem of quantum mechanics as it is usually understood, and how various interpretations propose to solve the problem. But the chapter doesn’t fit well with the rest of the book, which, taken as a whole, was already an attempt to make sense of it all. The revised version in the paperback edition is an improvement, but not entirely satisfactory. Chapter 6 of *Understanding Quantum Raffles*, on interpreting quantum mechanics, nails it.

Here, following the account by the three Mikes, is how I now see the view they call Bubism. Quantum mechanics began with Heisenberg’s unprecedented move to ‘reinterpret’ classical quantities like position and momentum as non-commutative. In a commutative algebra, the 2-valued quantities, representing propositions that can be true or false, form a Boolean algebra. A Boolean algebra is isomorphic to a set of subsets of a set, with the Boolean operations corresponding to the union, intersection, and complement of sets. The conceptual significance of Heisenberg’s proposal lies in replacing the Boolean algebra of subsets of classical phase space, where the points represent classical states and subsets represent ranges of values of dynamical variables, with a non-Boolean algebra. Later, following the Born-Heisenberg-Jordan *Dreimännerarbeit* and further developments by Dirac, Jordan, and von Neumann, this

non-Boolean algebra was formalized as the algebra of closed subspaces of Hilbert space, a vector space over the complex numbers, or equivalently a projective geometry. So the transition from classical to quantum mechanics is, formally, the transition from a Boolean algebra of subsets of a set to a non-Boolean algebra of subspaces of a vector space.

In his 1862 work ‘On the Theory of Probabilities,’ George Boole characterized a Boolean algebra as capturing ‘the conditions of possible experience.’ Classical theories are Boolean theories. The non-Boolean algebra of quantum mechanics (for Hilbert spaces of more than two dimensions) can be pictured as a family of Boolean algebras that are ‘intertwined,’ to use Gleason’s term,² or ‘pasted together,’ in such a way that the whole family can’t be embedded into a single Boolean algebra.³ So in a quantum theory, the single Boolean algebra of a classical theory is replaced by a family of Boolean algebras, in effect, a family of Boolean perspectives or Boolean frames associated with different incompatible measurement experiences. The upshot, as von Neumann pointed out, is that quantum probabilities are ‘perfectly new and *sui generis* aspects of physical reality’⁴ and ‘uniquely given from the start.’

The sense in which quantum probabilities are ‘uniquely given from the start’ is explained in an address by von Neumann on ‘unsolved problems in mathematics’ to an international congress of mathematicians in Amsterdam, September 2–9, 1954.⁵ Here is the relevant passage:

² A. N. Gleason, ‘Measures on the closed subspaces of Hilbert space,’ *Journal of Mathematics and Mechanics* 6, 885–893 (1957). The term is used to refer to intertwined orthonormal sets, which are Boolean algebras, on p. 886.

³ Kochen and Specker proved non-embeddability for the ‘partial Boolean algebra’ of subspaces of a Hilbert space of more than two dimensions in S. Kochen and E.P. Specker, ‘On the problem of hidden variables in quantum mechanics,’ *Journal of Mathematics and Mechanics* 17, 59–87 (1967). Bell proved a related result as a corollary to Gleason’s theorem in J.S. Bell, ‘On the problem of hidden variables in quantum mechanics,’ *Reviews of Modern Physics* 38, 447–452 (1966), reprinted in J.S. Bell, *Speakable and Unsayable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987).

⁴ From an unpublished manuscript ‘Quantum logics (strict- and probability-logics),’ reviewed in A.H. Taub in *John von Neumann: Collected Works* (Macmillan, New York, 1962), volume 4, pp. 195–197.

⁵ In Miklós Rédei and Michael Stöltzner (eds.), *John von Neumann and the Foundations of Quantum Mechanics*, pp. 231–246 (Kluwer Academic Publishers, Dordrecht, 2001). The quoted passage is on pp. 244–245. Also quoted (without the last sentence) in M. Rédei, ‘“Unsolved Problems in Mathematics” J. von Neumann’s address to the International Congress of Mathematicians Amsterdam, September 2–9, 1954,’ *The Mathematical Intelligencer* 21, 7–12 (1999).

Essentially if a state of a system is given by one vector, the transition probability in another state is the inner product of the two which is the square of the cosine of the angle between them [sic].⁶ In other words, probability corresponds precisely to introducing the angles geometrically. Furthermore, there is only one way to introduce it. The more so because in the quantum mechanical machinery the negation of a statement, so the negation of a statement which is represented by a linear set of vectors, corresponds to the orthogonal complement of this linear space. And therefore, as soon as you have introduced into the projective geometry the ordinary machinery of logics, you must have introduced the concept of orthogonality. This actually is rigorously true and any axiomatic elaboration of the subject bears it out. So in order to have logics you need in this set a projective geometry with a concept of orthogonality in it.

In order to have probability all you need is a concept of all angles, I mean angles other than 90° . Now it is perfectly quite true that in geometry, as soon as you can define the right angle, you can define all angles. Another way to put it is that if you take the case of an orthogonal space, those mappings of this space on itself, which leave orthogonality intact, leave all angles intact, in other words, in those systems which can be used as models of the logical background for quantum theory, it is true that as soon as all the ordinary concepts of logic are fixed under some isomorphic transformation, all of probability theory is already fixed.

What I now say is not more profound than saying that the concept of a priori probability in quantum mechanics is uniquely given from the start.

In *Bananaworld*, I defended what I called an ‘information-theoretic’ interpretation of quantum mechanics. The term is perhaps unfortunate. In the first place, it invites objections like those by Bell: ‘*Whose* information? Information about *what*?’⁷ In the second place, the emphasis should be on probability, as the three Mikes make clear, with the understanding that information theory is a branch of probability theory specifically concerned with probabilistic correlations.

If relativity is about space and time, quantum mechanics is *about probability*, in the sense that quantum probabilities are ‘*sui generis*’ and ‘uniquely given from the start’ as an aspect of the kinematic structure of the theory and are not imposed from outside as a measure of ignorance, as in classical theories, where probability is a measure over phase space. In this new framework, new sorts of nonlocal probabilistic correlations associated with entanglement are possible, which makes quantum information fundamentally different from

⁶ Von Neumann evidently meant to say that the transition probability is the square of the (absolute value of) the inner product, which is the square of the cosine of the angle between them.

⁷ J.S. Bell, ‘Against measurement,’ in *Physics World* 8, 33–40 (1990). The comment is on p. 34.

classical information. In a Boolean theory such correlations are impossible without introducing what Einstein called ‘spooky’ action at a distance.

Quantum probabilities are revealed in measurement, and a measurement is associated with the selection of a particular Boolean frame in the family of Boolean algebras that ‘captures the conditions of possible experience.’ In terms of observables, a measurement involves the selection of a basis of commuting observables in Hilbert space. As a consequence, the observer is no longer ‘detached,’ unlike the observer in classical mechanics, as Pauli observed.⁸ The measurement outcome is a random assignment of truth values to the elements in the Boolean frame, or a random assignment of values to the observables in the corresponding basis. What’s puzzling, from a Boolean perspective, is that measurement in a non-Boolean theory is not passive—not just ‘looking’ and registering what’s there in a passive sense. Measurement must produce a change in the description, and that’s not how we are used to thinking of measurement in a Boolean theory. Here’s how Schrödinger puts it:⁹

(1) The discontinuity of the expectation-catalog [the quantum pure state] due to measurement is *unavoidable*, for if measurement is to retain any meaning at all then the *measured value*, from a good measurement, *must* obtain. (2) The discontinuous change is certainly *not* governed by the otherwise valid causal law, since it depends on the measured value, which is not predetermined. (3) The change also definitely includes (because of ‘maximality’ [the ‘completeness’ of the quantum pure state]) some *loss* of knowledge, but knowledge cannot be lost, and so the object *must* change—*both* along with the discontinuous changes and *also*, during these changes, in an unforeseen, *different* way.

Quantum probabilities don’t simply represent ignorance about what is the case. Rather, they represent a new sort of ignorance about something that doesn’t yet have a truth value, something that simply isn’t one way or the other before we measure, something that requires us to act and do something that we call a measurement before nature supplies a truth value—and removes the truth values of incompatible propositions that don’t belong to the same Boolean frame, associated with observables that don’t commute with the measured observable.

⁸ M. Born, *The Born-Einstein Correspondence* (Walker and Co., London, 1971). Pauli talks about the classical ideal of the ‘detached observer’ in a letter to Born dated March 30, 1954 on p. 218.

⁹ ‘Die gegenwärtige Situation in der Quantenmechanik,’ *Die Naturwissenschaften* 48, 807–812; 49, 823–828, 844–849 (1935). The quotation is from p. 826. The translation is by John Trimmer, *Proceedings of the American Philosophical Society* 124, 323–338 (1980).

Schrödinger calls the measurement problem ‘the most difficult and most interesting point of the theory.’¹⁰ As the three Mikes aptly put it, the measurement problem is *a feature of quantum mechanics* as a non-Boolean theory, *not a bug*.

Interpretations of quantum mechanics that oppose the Copenhagen interpretation begin with Schrödinger’s wave theory as conceptually fundamental, rather than Heisenberg’s algebraic formulation of quantum mechanics, and propose dynamical solutions to what then seems to be a problem: how does what we do when we perform a measurement by manipulating some hardware in a laboratory select a Boolean frame in Hilbert space, a basis of observables that have definite values, and what explains the particular assignment of truth values to the elements in the Boolean frame, or the particular assignment of values to observables.

Bohm’s theory tells a one-world Boolean story: position in configuration space is always definite, associated with a Boolean algebra, and other quantities become definite through correlation with position via the measurement dynamics. The problem here, as Bell showed, is that Bohm’s theory is nonlocal in configuration space, allowing instantaneous action at a distance, which Einstein regarded as ‘spooky’¹¹ and so non-physical (although averaging over the Born distribution hides the nonlocality). I suspect that it was for this reason that Einstein dismissed Bohm’s theory as ‘too cheap for me’ in a letter to Born.¹²

The Everett interpretation tells a multi-world Boolean story in which everything that can happen does happen in some Boolean world. This avoids having to explain why *this* measurement outcome rather than *that* measurement outcome, since every possible outcome actually occurs in some world. The trick is to show how this fits Schrödinger’s wave theory of quantum mechanics. There is no spooky action at a distance in the Everettian interpretation, but the measurement problem appears as the basis problem: how to explain the selection of a particular basis with respect to which the multiplicity associated with ‘splitting into many worlds’ occurs in a measurement process. Everettians solve the basis problem by appealing to the dynamics of environmental decoherence: as the environment becomes increasingly entangled with the measuring ap-

¹⁰ *ibid.*, p. 826.

¹¹ M. Born, *op. cit.*. The term is used in a letter from Einstein to Born dated March 3, 1947 on p. 158.

¹² M. Born, *op. cit.* The comment is on p. 192 in a letter from Einstein to Born dated May 12, 1952.

paratus, it becomes more and more difficult, but not in principle impossible, to distinguish an entangled state from the corresponding mixture with respect to a particular coarse-grained basis. Quantum probabilities with respect to the elements of this basis are explained in terms of the decision theory of an agent-in-a-world about to make a measurement. Even granting decoherence as an effective solution to the basis problem, it seems contrived to interpret the ‘perfectly new and *sui generis* aspects of physical reality,’ the Hilbert space probabilities that are ‘uniquely given from the start,’ in this way.

Understanding Quantum Raffles is likely to be a classic in the foundational literature on quantum mechanics. The three Mikes have produced an exceptionally lucid book on quantum foundations that is also suitable for readers, with some tolerance for basic algebra and geometry, who are looking for answers to conceptual questions that are typically glossed over in standard courses on quantum mechanics.

Jeffrey Bub
Philosophy Department
Institute for Physical Science and Technology
Joint Center for Quantum Information and Computer Science
University of Maryland, College Park

Preface

The volume you just got yourself entangled with was inspired by Jeffrey Bub's (2016) *Bananaworld: Quantum Mechanics for Primates*. Our original plan had been to contribute an article to a special issue of the journal *Studies in History and Philosophy of Modern Physics* devoted to Jeff's book. That article eventually grew and morphed into this monograph, which we feel can now stand on its own feet. We are proud to present it as a volume in the series *Boston Studies in the Philosophy and History of Science*. In this volume, on the basis of some novel technical results (Chapters 2–5), we present and defend an informational interpretation of the basic framework of quantum mechanics (Chapters 1, 6–7). Our primary target audience for this book is physicists, philosophers of physics and students in these areas interested in the foundations of quantum mechanics. However, in the spirit of *Bananaworld* and its sequel, the graphic novel *Totally Random: Why Nobody Understands Quantum Mechanics* written by Jeff and his daughter Tanya (Bub & Bub, 2018), we wrote parts of our book (especially Chapter 2 and Sections 3.1–3.2) with the idea that they could be used as the basis for courses introducing non-physics majors to quantum mechanics, or for self-study by those outside of a university setting with an interest in quantum mechanics. Such readers, however, should be prepared to brush up on some high-school mathematics along the way (basic algebra and geometry; sines and cosines; vectors, matrices and determinants—but absolutely no calculus).¹³ We hope that all readers, even those who disagree with us on the basic issue of how their entanglement with our book results in them forming a definite view of its contents, will find something of value between its covers. This preface serves two purposes. First, we will briefly describe the contents of this volume. Second, we will give a brief history of how we came to write it, which will also give us an opportunity to thank the many people who helped us along the way.

Let us begin then by laying out the overall argumentative strategy of our book (which is in broad outline the same as it was in our original plan for a paper). We use correlation arrays, the workhorse of *Bananaworld*, to analyze the correlations found in an experimental setup due to David Mermin (1981) for measurements on pairs of spin- $\frac{1}{2}$ particles in the singlet state. Adopting an approach pioneered by Itamar Pitowsky (1989b) and promoted in *Banana-*

¹³ See Section 2.6.2, note 28, for some recommendations for non-expert readers looking for introductions to the basic formalism of quantum mechanics.

world, we geometrically represent the class of correlations allowed by quantum mechanics in this setup as an ellipsope in a non-signaling cube, which represents the broader class of all correlations that cannot be used for the purpose of sending signals traveling faster than the speed of light. To determine which of these quantum correlations are allowed by so-called local hidden-variable theories, we investigate which ones we can simulate using raffles with baskets of tickets that have the outcomes for all combinations of measurement settings printed on them. The class of correlations found this way can be represented geometrically by a tetrahedron contained within the ellipsope. We use the same Bub-Pitowsky framework to analyze a generalization of the Mermin setup for measurements on pairs of particles with higher spin in the singlet state. The class of correlations allowed by quantum mechanics in this case is still represented by the ellipsope; the subclass of those whose main features can be simulated with our raffles can be represented by polyhedra that, with increasing spin, have more and more vertices and facets and get closer and closer to the ellipsope.

We use these results to advocate for Bubism (not to be confused with QBism), an interpretation of quantum mechanics along the lines of *Banana-world*, belonging to the same lineage, or so we will argue, as the much-maligned Copenhagen interpretation. Probabilities and expectation values are primary in this interpretation. They are determined by inner products of state vectors in Hilbert space. State vectors do not themselves represent what is real in quantum mechanics. Instead the state vector gives a family of probability distributions over the values of subsets of observables, which do not add up to one overarching joint probability distribution over the values of all observables. As in classical theory, these values (along with the values of non-dynamical quantities such as charge or spin) represent what is real in the quantum world. Hilbert space puts constraints on possible combinations of such values, just as Minkowski space-time puts constraints on possible spatio-temporal constellations of events. To illustrate how generic such constraints are, we show that the one derived in this volume, the ellipsope inequality, is a general constraint on correlation coefficients, which can already be found in much older literature on statistics and probability theory. Udney Yule (1897) already stated the constraint. Bruno de Finetti (1937) already gave it a geometrical interpretation sharing important features with its interpretation in Hilbert space.

As this brief synopsis shows, polytopes and philosophy form two pillars of this volume. The third pillar is pedagogy. As noted above, we wrote parts of this volume as an introduction to quantum mechanics for non-specialists. For many years, one of us (Janssen) used a combination of the paper by Mermin (1981)

mentioned above and chapters from David Albert's *Quantum Mechanics and Experience* (Albert, 1992) to introduce quantum mechanics to non-physics majors in college and in high-school physics classes. Over the past few years, Janssen (assisted by Janas) has been developing a different approach, informed by and informing the material presented in this book. Like Albert (1992, Ch. 1, pp. 1–16, “Superposition”), we start, in Chapter 2, with certain stochastic experiments and show that classical theory (more precisely: local hidden-variable theories) cannot account for the statistics found in these experiments. Following Mermin rather than Albert, however, we choose (variations on) an experiment highlighting *entanglement* rather than *superposition* as the key feature that distinguishes quantum theory from classical theory (cf. Chapter 2, note 2 and Chapter 6, note 44). Albert (1992, Ch. 2, pp. 17–60) proceeds to give a concise and elementary exposition of the formalism of quantum mechanics (which we highly recommend to readers unfamiliar with it) and shows how it can account for the puzzling statistics presented in the opening chapter of his book. Yet it remains unclear how anybody would come up with this way of accounting for these puzzling statistics in the first place. Bub's *Bananaworld*, especially the notion of correlation arrays, allows us to do better. The correlation arrays for the puzzling statistics we start from can be parametrized by the sines and cosines of certain angles. In quantum mechanics such sines and cosines naturally emerge as components of vectors in various bases in what is called a Hilbert space. In Section 2.6, we introduce just enough formalism to get this basic idea across to non-specialists. More rigorous and more general versions of the arguments in Chapter 2 will be given in Chapter 4, which the reader can skip or skim (along with Chapter 5) without losing the thread of the overall argument (but we hope the reader will at least take a look at the pictures of correlation polyhedra in Figures 4.11, 4.13 and 4.17). The connection between quantum mechanics and general statistics and probability theory will be explored further in Chapter 3, also accessible to non-specialists with the exception of the later parts of Section 3.4. The upshot of Chapters 2–5 is summarized at the beginning of Chapter 6, making that chapter largely self-contained and thus suitable, all by itself, for courses on the foundations of quantum mechanics.

Polytopes, philosophy and pedagogy are the main interests of Janas, Cuffaro and Janssen, respectively. Accordingly, even though all three of us made substantial contributions to all seven chapters, Janssen had final responsibility for Chapters 1–2, Janas for Chapters 3–5 and Cuffaro for Chapters 6–7. The three of us came to this project from different directions. Janssen, a historian of science, is a recovering Everettian who has been defending Bub's

information-theoretic interpretation with the zeal of the converted. Cuffaro, a philosopher of science, was and is mainly interested in quantum computation and information, but began to think seriously again about the interpretation of quantum mechanics through conversations with Bill Demopoulos before meeting Janssen in 2017. Janas, a theoretical physicist, was and remains a Bohm sympathizer. Though we each have our own unique interests and histories, one thing the three of us share is a broadly Kantian outlook, something careful readers familiar with that outlook will not fail to notice as they go through the pages of this volume.

This project started in the Fall of 2016 when, at Janssen's suggestion, the Physics Interest Group (PIG) of the Minnesota Center for Philosophy of Science of the University of Minnesota, devoted most of its biweekly meetings that semester to *Bananaworld*. This book rekindled Janssen's interest in Bub and Pitowsky's heretical contribution to the Everett@50 conference in Oxford in 2007, "Two dogmas about quantum mechanics" (Bub & Pitowsky, 2010). In these PIG sessions, Janssen presented his reworking of Mermin's setup for testing a Bell inequality in terms of Bub's correlation arrays along with a (clumsy) derivation of the so-called Tsirelson bound for this setup. Janas attended these sessions. On a return visit to *Bananaworld* in the Fall of 2017, Janas began to explore the geometrical representation of correlation arrays by polyhedra and polytopes. He thereupon joined Janssen and Cuffaro, who, at the 2017 edition of the conference New Directions in the Foundations of Physics in Tarquinia, had decided to write a response to *Bananaworld* together. In the Fall of 2017, Janssen gave a physics colloquium at Minnesota State University Mankato on our joint project, and then a lunchtime talk at the Center for Philosophy of Science at the University of Pittsburgh in the Spring of 2018. By that time Laurent Taudin, illustrator extraordinaire for many projects of the *Max-Planck-Institut für Wissenschaftsgeschichte* in Berlin, had drawn the figures of the chimps and the bananas that we have been using in talks and lectures since (see Figures 2.1 and 2.2).

After extensive preparatory work by Janas and Janssen in the Fall of 2018, we started writing what would eventually become this book during a visit by Cuffaro to Minnesota in January 2019. In March, Cuffaro presented a preliminary version of parts of Chapters 2, 3 and 6 at the Workshop on Interpreting Quantum Mechanics organized by Giovanni Valente at the *Politecnico di Milano* in Milan. In May, after a test run by Janas in a mathematics colloquium at the University of Minnesota, the three of us then presented parts of these same chapters at the 2019 edition of New Directions in the Foundations of Physics in Viterbo. A question for Janas by Wayne Myrvold in Q&A alerted

us to an important gap in one of our key results, which we have since managed to close (see Chapter 3, notes 10 and 11). In June 2019 the three of us met again in Minneapolis. Over the ensuing months we finalized (or so we thought) our manuscript and in October we posted it on the arXiv and on the PhilSci Archive preprint servers. By that time Janas had filled several whiteboards in Tate Hall, housing part of the School of Physics and Astronomy of the University of Minnesota, many times over to go over (preliminary versions of) the results presented in Chapters 2–5 with Janssen and, when in town, Cuffaro. Janas also did the computer programming needed for Section 4.2 and for Figures 2.8 and 2.16. Janssen is responsible for most other figures. Cuffaro handled whatever L^AT_EX issues we ran into.

Janssen gave two talks on parts of our preprint at the Second Chilean Conference on the Philosophy of Physics organized by Pablo Acuña in Santiago in December 2019, where he had the opportunity to discuss the material in person with Jeff Bub. A slightly revised version of our preprint was then pre-circulated among participants in a symposium on the foundations of quantum mechanics organized by Janssen, Jürgen Jost and Jürgen Renn at the *Max-Planck-Institut für Wissenschaftsgeschichte* in Berlin in January 2020. In this symposium, Cuffaro and Janssen presented parts of what was starting to get referred to as the “Three Mikes Manifesto,” a play on the famous *Dreimännerarbeit* (Three men paper) with which Max Born, Werner Heisenberg and Pascual Jordan (1926) consolidated matrix mechanics. Based on feedback from the participants in this symposium (especially Guido Bacciagaluppi, Jürgen Jost, Jürgen Renn and Matthias Schemmel) and from others who had read our preprint, we added further material to Chapter 5 and substantially rewrote Chapters 1 and 6 (especially Section 6.5 on measurement). We also changed the title. The title of our preprint, “Putting probabilities first: How Hilbert space generates and constrains them,” would have been fine for a journal article in a special issue devoted to *Bananaworld*. It would have been obvious, for instance, in that context that our topic is quantum mechanics even though the title does not explicitly mention this. Given the use of Hilbert space methods in general probability theory and statistics, however, this would not have been clear for a monograph with that same title. We settled on the new title *Understanding Quantum Raffles*. Raffles of various designs are ubiquitous in this volume. And while we are hardly the first to argue that the basic formalism of quantum mechanics is essentially a new framework for handling probabilities (cf. Chapter 1, notes 16 and 29), we are the first to do so on the basis of a sustained comparison between raffles serving as toy models of local hidden-variable theories and the statistical ensembles characterized by density operators in terms

of which John von Neumann (inspired by Richard von Mises) first formulated quantum mechanics (Von Neumann, 1927b). The “quantum raffles” in the title of our book refer to these statistical ensembles introduced by von Neumann.

In the Fall of 2020, after trying out some of the material in Chapter 2 in classes at the University of Minnesota and Washburn High School in Minneapolis, Janssen, assisted by Janas, taught a seminar in the Honors Program of the University of Minnesota under the title of Gilder’s (2008) *The Age of Entanglement*, covering—in addition to Gilder’s book and the graphic novel *Totally Random* by Tanya and Jeff Bub (2018)—Chapters 1–3 of the manuscript of *Understanding Quantum Raffles*. In response to student feedback, we reorganized some of the material in Chapters 2 and 3.

We are grateful for the questions from the audiences at the various workshops and talks mentioned above as well as for the feedback from students at the University of Minnesota and Washburn High School. In addition, we want to single out a number of individuals not explicitly mentioned so far and thank them for helpful comments and discussion: Jossi Berkovitz, Victor Boantza, Harvey Brown, Časlav Brukner, Adán Cabello, Joe Cain, Cindy Cattell, Radin Dardashti, Michael Dascal, Robert DiSalle, Tony Duncan, Lucas Dunlap, Laura Felling, Sam Fletcher, Mathias Frisch, Chris Fuchs, Louisa Gilder, Sona Ghosh, Peter Gilbertson, Peter Grul, Bill Harper, Stephan Hartmann, Geoffrey Hellman, Leah Henderson, Federico Holik, Luc Janssen, Christian Joas, Molly Kao, David Kaiser, Jim Kakalios, Alex Kamenev, Jed Kaniewski, Marius Krumm, Femke Kuiling, Samo Kutoš, Christoph Lehner, Charles Marcus, Tushar Menon, Eran Moore Rea, Markus Müller, Max Niedermaier, Sergio Pernice, Vincent Pikavet, Serge Rudaz, David Russell, Rob “Ryno” Rynasiewicz, Juha Saatsi, Ryan Samaroo, Chris Smeenk, Rob Spekkens, Jos Uffink, David Wallace and Brian Woodcock. We thank Lindy Divarci, Jürgen Renn and Matteo Valleriani of the *Max-Planck-Institut für Wissenschaftsgeschichte* for their help in turning our manuscript into a book. We thank an anonymous referee who reviewed our book for Springer both for the enthusiastic endorsement and for helpful comments. We thank Lucy Fleet, Prasad Gurunadham and Svetlana Kleiner at Springer for shepherding our manuscript through the production process.

We saved our most important intellectual debts for last. A heartfelt thanks to Jeff Bub for his enthusiastic support of our efforts and for his patience in explaining and discussing his views on the foundations of quantum mechanics with us, both in person and in email exchanges dating back to 2007. We are also grateful for all we learned from Itamar Pitowsky (1950–2010) and William

Demopoulos (1943–2017). Instead of dedicating this volume to them, we would have loved to discuss it with Bill and Itamar.

Finally, we want to express our thanks for generous institutional support. Janssen gratefully acknowledges support from the *Alexander von Humboldt Stiftung* and the *Max-Planck-Institut für Wissenschaftsgeschichte*. Cuffaro gratefully acknowledges support from the *Alexander von Humboldt Stiftung*, the Rotman Institute of Philosophy at Western University, the Foundational Questions Institute (FQXi), the Descartes Centre at Utrecht University, and the Institute for Quantum Optics and Quantum Information in Vienna. Janas thanks the University of Minnesota for travel support as well as the staff of Al's Breakfast in Dinkytown.

Lino Lakes, MN, USA
Montréal, Québec, Canada
Minneapolis, MN, USA

Michael Janas
Mike Cuffaro
Michel Janssen

March 2021

Contents

1	Introduction	1
2	Representing distant correlations by correlation arrays and polytopes	17
2.1	Taking Mermin to Bananaworld	17
2.2	Correlations found when peeling and tasting pairs of quantum bananas	20
2.3	Non-signaling correlation arrays	23
2.4	The non-signaling cube, the classical tetrahedron and the quantum ellipsope	27
2.5	Raffles meant to simulate the quantum correlations and the classical tetrahedron	32
2.6	The quantum correlations and the ellipsope	43
2.6.1	Getting beyond the classical tetrahedron: the ellipsope inequality	44
2.6.2	The quantum correlation array: the singlet state and the Born rule	50
3	The ellipsope and the geometry of correlations	65
3.1	The Pearson correlation coefficient and the ellipsope inequality	67
3.2	Why the quantum correlations saturate the ellipsope	71
3.3	Why our raffles do not saturate the ellipsope	83
3.4	The geometry of correlations: from Pearson and Yule to Fisher and De Finetti	87
4	Generalization to singlet state of two particles with higher spin	103
4.1	The quantum correlations	107
4.1.1	Quantum formalism for one spin- s particle	107
4.1.2	Quantum formalism for two spin- s particles in the singlet state	112
4.1.3	Wigner d -matrices and correlation arrays in the spin- s case	114
4.1.4	Non-signaling in the spin- s case	118
4.1.5	Anti-correlation coefficients in the spin- s case	119
4.1.6	Cell symmetries in the spin- s case	121
4.2	Designing raffles to simulate the quantum correlations	125

4.2.1	Spin- $\frac{1}{2}$	125
4.2.2	Spin-1	130
4.2.3	Spin- $\frac{3}{2}$	141
4.2.4	Spin- s ($s \geq 2$)	149
5	Correlation arrays, polytopes and the CHSH inequality	157
6	Interpreting quantum mechanics	171
6.1	The story so far	171
6.2	From within and from without	179
6.3	The new kinematics of quantum theory	183
6.4	Examples of problems solved by the new kinematics	187
6.5	Measurement	194
7	Conclusion	227



Chapter 1

Introduction

Genealogy of interpretations of quantum mechanics • Informational interpretations: objections (parochialism and anti-realism) and rejoinders • Bub and Pitowsky's "big" and "small" measurement problems: puzzles to be solved or lessons to be learned? • Bub's correlation arrays and Pitowsky's correlation polytopes for Mermin's setup to test a Bell inequality • Quantum mechanics as a general framework for handling probabilities.

This volume is a brief for a specific take on the general framework of quantum mechanics. In terms of the usual partisan labels, it is an *informational* interpretation in which the status of the state vector is *epistemic* rather than *ontic*. On the ontic view, state vectors represent what is ultimately real in the quantum world; on the epistemic view, they are auxiliary quantities for assigning probabilities to values of observables in a world in which it is impossible to do so simultaneously for all observables. Such labels, however, are of limited use for a classification of interpretations of quantum mechanics. A more promising approach might be to construct a genealogy.¹ As this is not a historical work, a rough characterization of the relevant phylogenetic tree must suffice for our purposes.²

¹ The contemporary literature on quantum foundations has muddied the waters in regards to the classification of interpretations of quantum mechanics, and it is partly for this reason that we prefer to give a taxonomy in terms of a genealogy. Ours is *not* an epistemic interpretation of quantum mechanics in the sense compatible with the ontological models framework of [Harrigan & Spekkens \(2010\)](#). In particular it is not among our assumptions that a quantum system has, at any time, a well-defined ontic state. Actually we take one of the lessons of quantum mechanics to be that this view is untenable (see Section 6.3 below). For more on the differences between a view such as ours and the kind of epistemic interpretation explicated in [Harrigan & Spekkens \(2010\)](#), and for more on why the no-go theorem proved by [Pusey, Barrett, & Rudolph \(2012\)](#) places restrictions on the latter kind of epistemic interpretation but is at most only indirectly relevant to ours, see [Ben-Menahem \(2017\)](#).

² One of us is working on a two-volume book on the genesis of quantum mechanics. The first volume has already been published ([Duncan & Janssen, 2019](#)).

The main thing to note then is that the mathematical equivalence of wave and matrix mechanics papers over a key difference in what its originators thought their big discoveries were.³ These big discoveries are certainly compatible with one another but there is at least a striking difference in emphasis. For Erwin Schrödinger the big discovery was that a wave phenomenon underlies the particle behavior of matter, just as physicists in the 19th century had discovered that a wave phenomenon underlies geometrical optics (Joas & Lehner, 2009). For Werner Heisenberg it was that the problems facing atomic physics in the 1920s call for a new framework to represent physical quantities just as electrodynamics had called for a new framework to represent their spatio-temporal relations two decades earlier (Duncan & Janssen, 2007, Janssen, 2019, pp. 134–142).⁴ What are now labeled ontic interpretations—e.g., Everett’s many-worlds interpretation, De Broglie-Bohm pilot-wave theory and the spontaneous-collapse theory of Ghirardi, Rimini and Weber (GRW)—can be seen as descendants of wave mechanics; what are now labeled epistemic interpretations—e.g., the much maligned Copenhagen interpretation⁵ and Quantum Bayesianism or QBism—as descendants of matrix mechanics.⁶

³ As Bacciagaluppi & Valentini (2009, pp. xv–xvi) explicitly acknowledge in the preface of their carefully annotated edition of the proceedings of the 1927 Solvay conference, their main objective was to show how the current debates over the foundations of quantum theory could benefit from being traced back to the discussions between the founding fathers attending this conference shortly after the theory in its various guises had first been formulated.

⁴ In *Bananaworld*, Bub makes a similar point. Heisenberg’s first paper on what would become matrix mechanics, he writes, “contained the germ of a radically new way of thinking about physical systems . . . while Schrödinger’s wave mechanics evoked a very different structural picture that has turned out to be misleading in many ways” (Bub, 2016, p. 4). The development of quantum mechanics—as Olivier Darrigol (2014, p. 237) puts it about as concisely as one can imagine—was driven by two analogies, “the analogy between classical and quantum theory [and] the analogy between matter and light. These analogies led to two different versions of the new mechanics: the Heisenberg-Born-Jordan matrix mechanics and the De Broglie-Schrödinger wave mechanics, which occurred in two different contexts: atomic constitution and the statistical properties of matter and light.”

⁵ In the preface of his book on quantum computing, Mermin (2007, p. xii) notes that his presentation “is suffused with a perspective on the quantum theory that is very close to the venerable but recently much reviled Copenhagen interpretation,” so much so, in fact, that he considered calling his book *Copenhagen Computation*.

⁶ David Wallace (2019) provides an example from the quantum foundations literature that shows that the “big discoveries” of matrix and wave mechanics are not mutually exclusive. He argues that the Everett interpretation should be seen as a general new framework for physics while endorsing the view that vectors in Hilbert space represent what is real in the quantum world. Wallace and other Oxford Everettians derive the Born rule for probabilities in quantum mechanics from decision-theoretic considerations instead of taking it to be given

The interpretation for which we will advocate in this volume can, more specifically, be traced to the (statistical) transformation theory of Pascual [Jordan \(1927a,b\)](#) and Paul [Dirac \(1927\)](#), amplified in his famous book, [Dirac, 1930](#)) and to the “probability-theoretic construction” (*Wahrscheinlichkeitstheoretischer Aufbau*) of quantum mechanics in the second installment of the trilogy of papers by John [Von Neumann \(1927a,b,c\)](#) that would form the backbone of *his* famous book ([Von Neumann, 1932](#)). While incorporating the wave functions of wave mechanics, both Jordan’s and Dirac’s version of transformation theory grew out of matrix mechanics. More strongly than Dirac, Jordan emphasized the statistical aspect. The “new foundation” (*Neue Begründung*) of quantum mechanics announced in the titles of Jordan’s 1927 papers consisted of a number of postulates about the probability of finding a value for one quantum variable given the value of another. Von Neumann belongs to that same lineage. Although he proved the mathematical equivalence of wave and matrix mechanics in the process (by showing that they correspond to two different instantiations of Hilbert space), he wrote his 1927 trilogy in direct response to Jordan’s version of transformation theory. His *Wahrscheinlichkeitstheoretischer Aufbau* grew out of his dissatisfaction with Jordan’s treatment of probabilities. Drawing on work in probability theory by Richard von Mises (soon to be published in book form; [Von Mises, 1928](#)), he introduced the now familiar density operators characterizing uniform (pure state) and non-uniform (mixed state) ensembles of quantum systems.⁷ He showed that what came to be known as the Born rule for probabilities in quantum mechanics can be derived from the Hilbert space formalism and some seemingly innocuous assumptions about properties of the function that gives expectation values ([Duncan & Janssen, 2013](#), sec. 6, pp. 246–251). This derivation was later re-purposed for the infamous von Neumann no-hidden variables proof, in which case the assumptions, entirely appropriate in the context of the Hilbert space formal-

by the Hilbert space formalism the way von Neumann showed one could (see below). For other Everettians, such as Christoph Lehner, state vectors are both ontic and epistemic. They help themselves to the Born rule *à la* von Neumann but also use state vectors to represent physical reality. One way to argue for such a position (suggested by a talk by Lehner during the workshop in Berlin in January 2020 mentioned in the preface) is to insist on a Lewisian modal realist interpretation of probabilities. One could then agree with members of the “epistemic camp” that quantum mechanics provides a new framework for handling probabilities but at the same time hold on to the notion of a multiverse.

⁷ For historical analysis of these developments, focusing on Jordan and von Neumann, see [Duncan & Janssen \(2013\)](#) and, for a summary aimed at a broader audience, [Janssen \(2019\)](#), pp. 142–161).

ism of quantum mechanics, become highly questionable (Bub, 2010a, Dieks, 2017).

A branch on the phylogenetic tree of interpretations of quantum mechanics close to our own is the one with Jeffrey Bub and Itamar Pitowsky's (2010) "Two dogmas about quantum mechanics," a play on W. V. O. Quine's (1951) celebrated "Two dogmas of empiricism." Bub and Pitowsky presented (an early version of) their paper in the Everettians' lion's den at the 2007 conference in Oxford marking the 50th anniversary of the Everett interpretation.⁸ It appears in the proceedings of that conference. Enlisting the help of his daughter Tanya, a graphic artist, Bub has since mounted an impressive PR campaign to bring his and Pitowsky's take on quantum mechanics to the masses. Despite its title and lavish illustrations, his first attempt, *Bananaworld: Quantum Mechanics for Primates* (Bub, 2016), is not really a popular book. Its sequel, however, the graphic novel *Totally Random* (Bub & Bub, 2018), triumphantly succeeds where *Bananaworld*, in that respect, came up short.⁹ The interpretation promoted overtly in *Bananaworld* and covertly in *Totally Random* has been dubbed *Bubism* by Robert Rynasiewicz (private communication).¹⁰ Like QBism, Bubism is an informational interpretation but for a Bubist quantum probabilities are objective chances whereas for a QBist they are subjective degrees of belief (Fuchs & Stacey, 2019). We will defend our own version of Bubism, building on the Bubs' two books and on "Two dogmas . . ." as well as on earlier work by (Jeff) Bub and Pitowsky, especially the latter's lecture notes *Quantum Probability—Quantum Logic* and his papers on George Boole's "conditions of possible experience" (Pitowsky, 1989a, 1994). We will rely heavily on tools developed by these two authors, Bub's *correlation arrays*

⁸ The video of their talk could, at the time we wrote this, still be watched at <users.ox.ac.uk/~everett/videobub.htm>. See Bub & Demopoulos (2010) for a moving obituary of Pitowsky.

⁹ See, e.g., the review in *Physics World* by Minnesota physicist Jim Kakalios (2018), well-known for his use of comic books to explain physics (Kakalios, 2009), the review in *Physics Today* by philosopher of quantum mechanics Richard Healey (2019), and the essay review co-authored by one of us (Cuffaro & Doyle, 2021).

¹⁰ In an essay review of Ball (2018), Becker (2018) and Freire (2015), Bub (2019a) gives a concise characterization of his views and places them in the lineage of Heisenberg sketched above (cf. note 4). In Chapter 6, we will quote various passages from Bohr (1928, 1935, 1937, 1948, 1958) that help convey the way that Bohr fits into this lineage (see Section 6.5, notes 23, 45, 47 and 49). For more on our lineage's connection with Bohr's views, see the preprint Bub (2017), the sections on Bohr were dropped in the published version of the paper), Bub (2019a, pp. 235–236) and an unpublished monograph by Demopoulos (2018, see below).

and Pitowsky's *correlation polytopes*. A third musketeer on whose insights we drew for our own work is William Demopoulos (see, e.g., Demopoulos, 2010, 2012, and, especially, 2018, a monograph he completed shortly before he died, which we fervently hope will be published soon).^{11, 12}

In the spirit of *Bananaworld*, *Totally Random* and Louisa Gilder's (2008) lovely *The Age of Entanglement*, we wrote Chapter 2 with a non-specialist audience in mind (cf. our comments in the preface).¹³ We will frame our argument in that chapter in terms of a variation of Bub's scheme for peeling and tasting quantum bananas (see Figures 2.1 and 2.2). This is not just a gimmick adopted for pedagogical purposes. It is also meant to remind the reader that, on a Bubist view, inspired by Heisenberg rather than Schrödinger, quantum mechanics provides a new framework for dealing with arbitrary physical systems, be they waves, particles, fields, or fictitious quantum bananas. The peeling and tasting of bananas also makes for an apt metaphor for the (projective) measurements we will be considering throughout (cf. Popescu, 2016). Our variation on Bub's peel-and-taste scheme makes it easy to pivot, in Section 4.1, from tasting bananas peeled in different ways to measuring the spin of particles sent through Du Bois magnets pointing in different directions in variations on the famous Stern-Gerlach experiment.¹⁴

Despite our phylogenetic proximity to Bub, we follow Jordan rather than Bub in arguing that quantum mechanics is essentially a new framework for

¹¹ Bub (2016, 2nd ed., p. 232, note 29) acknowledges the importance of discussions with Demopoulos for his thinking about quantum mechanics and cites Demopoulos (2010, 2012).

¹² We consider the views of Bub, Demopoulos and Pitowsky to be our neighbors on the phylogenetic tree of interpretations of quantum mechanics. Although we would not call QBism a neighboring view, we take it to be closer to us than many others. The same, we think, can also be said about the views of, among others, Scott Aaronson (2013), Časlav Brukner (2017), Lucien Hardy (2001), Richard Healey (2017) and Carlo Rovelli (1996).

¹³ We adopted the convention of using "Chapter" and "Section" when referring to other parts of this volume and "Ch." and "sec." when referring to chapters and sections in other sources we cite or quote. Equations are numbered by section (e.g., Eq. 2.3.1 refers to the first equation in Chapter 2, Section 2.3). Figures and tables are numbered by chapter (e.g., Figure 2.1 refers to the first figure in Chapter 2). Notes are also numbered by chapter. Unless a chapter number is specified, a reference to a note is to the note of that number in the same chapter.

¹⁴ See Section 6.5 (especially note 41 and Figure 6.1) for a description of such experiments, first performed by Walther Gerlach and Otto Stern (1922).

handling *probability* rather than *information*.^{15, 16} We are under no illusion that this substitution will help us steer clear of two knee-jerk objections to informational approaches to the foundations of quantum mechanics: *parochialism* and *instrumentalism* (or *anti-realism*).

What invites complaints of parochialism is the slogan “Quantum mechanics is all about information.” This slogan conjures up the unflattering image of a quantum-computing engineer, whose worldview is much like that of the proverbial carpenter who only has a hammer and sees everything as a nail. It famously led John Bell (1990, p. 34) to object: “*Whose* information? Information about *what*?” In *Bananaworld*, Bub (2016, p. 7) counters: “we don’t ask these questions about a USB flash drive. A 64 GB drive is an information storage device with a certain capacity, and whose information or information about what is irrelevant.” A computer analogy, however, is probably not the most effective way to combat the lingering impression of parochialism. We can think of two better responses.

The first is an analogy with meter rather than memory sticks. Consider the slogan “Special relativity is all about space-time” or “Special relativity is all about spatio-temporal relations.” These slogans, we suspect, would not provoke the hostile reactions routinely elicited by the slogan “Quantum mechanics is all about information.” Yet, one could ask, parroting Bell: “spatio-temporal relations of *what*?” The rejoinder in this case is simply that *what* could be any physical system allowed by the theory; and that, to qualify as such, it suffices that *what* can consistently be described in terms of mathematical quantities that transform as scalars, vectors, tensors or spinors under Lorentz transformations. When we say that a moving meter stick contracts by such-and-such a factor, we

¹⁵ This departure from Bub is not as big as it sounds. As Bub (2016, p. 6) makes clear in *Bananaworld*, “[i]n modern formulations, information theory is about random variables . . . and correlations between random variables. As such, information theory is a branch of the mathematical theory of probability.” We chose to call our interpretation “informational” (in part) to emphasize our proximity to Bub’s view, but unlike Bub we chose not to call our interpretation “information-theoretic.” The latter term would suggest an explicit connection to results from information theory in the Shannon or Schumacher senses, but in our case we are using information only in the general sense just described.

¹⁶ Others who have endorsed this view include Hardy (2001, p. 1), who has called quantum theory “simply a new type of probability theory” (quoted by Darrigol, 2014, p. 318), and Aaronson, who in the preface of *Quantum Computing since Democritus* relates the amusing story of how an Australian TV commercial made him realize that a passage from his book used in that commercial nicely captures the book’s central thesis: quantum mechanics is “about information, probabilities, and observables, and how they relate to each other” (Aaronson, 2013, p. xii–xiv, p. 110). See also note 29.

only have to specify its velocity with respect to the inertial frame of interest, not what it is made of. Special relativity imposes certain kinematical constraints on any physical system allowed by the theory. Those constraints are codified in the geometry of Minkowski space-time. There is no need to reify Minkowski space-time, no need for “fetishism of mathematics” as John Stachel (1994, p. 149) would put it. We can think of space-time in relational rather than substantival terms (Janssen, 2009, p. 28).

The slogan “Quantum mechanics is all about information/probability” can be unpacked in a similar way. Quantum mechanics imposes kinematical constraints on allowed values and combinations of values of observables. Which observables? Any observable that can be represented by a Hermitian operator on Hilbert space. The constraints quantum mechanics imposes on the values of such observables are codified in the geometry of Hilbert space. And as in the case of Minkowski space-time, there is no need to reify Hilbert space. So, yes, quantum mechanics is obviously about more than just information, just as special relativity is obviously about more than just space-time. Yet the slogans that special relativity is all about space-time and that quantum mechanics is all about information (or probability) do capture—the way slogans do—what is distinctive about these theories and what sets them apart from the theories they superseded.

In Chapter 6, we will further explore this parallel between quantum mechanics and special relativity.¹⁷ We should warn the reader upfront though that the kinematical take on special relativity underlying this comparison, while in line with the majority view among physicists, is not without its detractors. In fact, the defense of a kinematical interpretation of special relativity by one of us (Janssen, 2009) was mounted in response to an alternative dynamical interpretation of special relativity articulated and defended most forcefully by Harvey Brown (2005).¹⁸ Both Bub (2016, p. 228) in *Bananaworld* and Bub & Pitowsky (2010, p. 439) in “Two dogmas . . .” have invoked analogies with special relativity to defend their information-theoretic interpretation of quan-

¹⁷ Janssen (2009, p. 28) distinguishes between phenomena being kinematical in a narrow and in a broad sense, defined as being an example of default spatio-temporal behavior and being independent of the specifics of the dynamics, respectively. In this volume, we use the term kinematical in the latter sense. For a history of kinematics in the traditional sense of the science or geometry of motion, see Martínez (2009).

¹⁸ See Acuña (2014) for an enlightening discussion of the debate over whether special relativity is best understood kinematically or dynamically.

tum mechanics. [Brown & Timpson \(2006\)](#) have disputed the cogency of these analogies (see also [Timpson, 2010, 2013](#), sec. 8.3.3).¹⁹

Our second response to the parochialism charge is that the quantum formalism for dealing with intrinsic angular momentum, i.e., spin, laid out in Section 4.1 and used throughout in our quantum-mechanical analysis of an experimental setup to test the Bell inequalities, is key to spectroscopy and other areas of physics as well, such as, e.g., the theory of the electric susceptibility of diatomic gases. These two responses are not unrelated. In Section 6.4, drawing on papers on the history of quantum physics co-authored by one of us,²⁰ we will give a few examples of puzzles for the old quantum theory that physicists resolved not by altering the dynamical equations but by using key features of the kinematical core of the new quantum mechanics.

What about the other charge against informational interpretations of quantum mechanics, *instrumentalism* or *anti-realism*? What invites complaints on this score in the case of Bub and Pitowsky is their rejection of the second of the “two dogmas” they identified: “the quantum state is a representation of physical reality” ([Bub & Pitowsky, 2010](#), p. 433).²¹ This statement of the purported dogma is offered as shorthand for a more elaborate one: “[T]he quantum state has an ontological significance analogous to the significance of the classical state as the ‘truthmaker’ for propositions about the occurrence or non-occurrence of events” (*ibid.*). Of course, denying that state vectors in Hilbert space represent physical reality in and of itself does not make one an anti-realist. We can still be realists as long as we can point to other elements of the theory’s formalism that represent physical reality. The sentence we just quoted from “Two dogmas . . .” suggests that for Bub and Pitowsky “events” fit that bill. However, as John [Earman \(1989](#), p. 186) has argued regarding taking point coincidences to exhaust what is real in general relativity, such an event ontology smacks of “a crude verificationism and an impoverished conception of physical reality.” In Section 6.5 we argue that one can do considerably better without reinstating the dogma that Bub and Pitowsky want to reject.

¹⁹ What complicates matters here is that the distinction between kinematics and dynamics tends to get conflated with the distinction between constructive and principle theories ([Janssen, 2009](#), p. 38; see Section 6.1 below for further discussion).

²⁰ [Duncan & Janssen \(2008, 2014, 2015\)](#) and [Midwinter & Janssen \(2013\)](#).

²¹ In what he identifies as “orthodox quantum mechanics,” [Wallace \(2016](#), p. 4) detects “an inchoate attitude to the quantum state . . . where it is interpreted either as physically representational or as probabilistic, according to context.” This representational/probabilistic distinction is cited by [Bub \(2016](#), 2nd ed., p. 231, note 21). In Section 6.5, we will return to this point and characterize the “inchoate attitude” Wallace draws attention to more carefully.

By focusing on “events”, Bub and Pitowsky privilege dynamical observable quantities (e.g., energy, momentum, spin in the z -direction, . . .) over fixed parameters that may or may not be directly observable (e.g., rest mass and electric charge or bare masses and coupling constants, respectively) when it comes to representing physical reality. We do not. Neither in classical nor in quantum mechanics do we simply posit some grab bag of observable quantities to represent physical reality. Rather, constrained by factors such as the empirical and explanatory successes of prior theory, the existing mathematical toolkit, and the culturally specific reservoir of metaphors and analogies available for heuristic purposes (wave and particle imagery for instance),²² we posit systems with fixed properties such as mass, charge, and spin and introduce dynamical observable quantities as further properties of such systems. The objects of our everyday macroscopic world and their states are then constructed somehow out of catalogs of values, be they variable or constant, of these quantities, be they dynamical or fixed. How exactly this is done is a question physicists may want to leave for philosophers to ponder, especially since this is not what separates quantum from classical mechanics.

The key difference between classical and quantum mechanics lies in *how values are assigned to dynamical quantities*. Bub and Pitowsky’s notion of a “truthmaker” provides a nice way to articulate this difference. In classical mechanics, dynamical observable quantities are represented by functions on the phase space of the system in question. Picking a point in phase space fixes the values of all these quantities. It is in this sense that points in phase space are “truthmakers”. In quantum mechanics, dynamical observable quantities are represented by Hermitian operators on Hilbert space. The possible values of these quantities are given by the eigenvalues of these operators. Picking a vector in Hilbert space, however, does not fix the value of these quantities in advance of a measurement. It fails to do so in two ways. First, the quantity or quantities being measured must be selected. Only those selected get to be assigned definite values. Quantum mechanics tells us that, once a selection is made, it is impossible for any quantity represented by an operator that does not commute with those representing the selected one(s) to be assigned a definite value as well. Second, even after a selection has been made, the state vector will in general only give a probability distribution over the various eigenvalues of the operators corresponding to the selected observables. Which of those values is found upon measurement of the observable is a matter of chance. Vectors in Hilbert space thus doubly fail to be “truthmakers”. *Pace* Bub and

²² For reflections by one of us on how new theories get introduced, see [Janssen \(2002, 2019\)](#).

Pitowsky, however, it does not follow that classical and quantum states have a different “ontological significance.” Neither vectors in Hilbert space *nor points in phase space* represent our physical world.²³ Both are mathematical auxiliaries for assigning values (albeit in radically different ways) to quantities that do.²⁴

What about the first dogma Bub & Pitowsky (2010, p. 433) want to reject: *Measurement outcomes should be fully explained in terms of the dynamical interaction between the system being measured and a measuring device*? Striking this dogma from the quantum catechism trivially solves the measurement problem in its traditional form of having two different dynamics side-by-side, unitary Schrödinger evolution as long as we do not make a measurement, state vector collapse as soon as we do. If we reject the demand for a dynamical account of how a measurement results in a particular definite outcome, this problem obviously evaporates. As we will show in Section 6.5, renouncing this dogma does not amount to black-boxing measurements.²⁵ On Bub and Pitowsky’s view, any measurement can be analyzed in as much detail as one can ask for. It does mean, however, that one accepts that there comes a point where no meaningful further analysis can be given of why a measurement gives one particular outcome rather than another. Instead it becomes a matter of irreducible randomness—the ultimate crapshoot.²⁶

Bub & Pitowsky (2010, p. 438) distinguish between what they, with thick irony, call a “big” and a “small” measurement problem. This distinction maps nicely onto the two ways distinguished above in which the quantum state vector fails to be a “truthmaker”. Those two ways correspond to two questions: (I) How does one set of observables rather than another get selected to be

²³ We will discuss this more carefully in Section 6.5.

²⁴ In *Wahrscheinlichkeitstheoretischer Aufbau*, von Neumann also resisted the idea that vectors in Hilbert space ultimately represent (our knowledge of) physical reality. He wrote: “our knowledge of a system \mathfrak{S}' , i.e., of the structure of a statistical ensemble $\{\mathfrak{S}'_1, \mathfrak{S}'_2, \dots\}$, is never described by the specification of a state—or even by the corresponding φ [i.e., the vector $|\varphi\rangle$]; but usually by the result of measurements performed on the system” (Von Neumann, 1927b, p. 260). He thus wanted to represent “our knowledge of a system” by the values of a set of observables corresponding to a complete set of commuting operators (Duncan & Janssen, 2013, pp. 251–252).

²⁵ Wallace raised this objection to Bub and Pitowsky’s presentation of “Two dogmas . . .” in Oxford in 2007 (Saunders, Barrett, Kent, & Wallace, 2010, p. 597).

²⁶ We realize that it is easier to swallow this “totally random” response for the observables we will be considering (where the spin of some particle can be up or down or a banana can taste yummy or nasty) than for others, such as, notably, position (where a particle can be here or on the other side of the universe).

assigned definite values? (II) Why does an observable, once selected, take on one value rather than another? These two questions can be seen as statements of the “small” and the “big” measurement problem, respectively.²⁷ Rejection of what Bub and Pitowsky identify as the first dogma of quantum mechanics then amounts to dismissing the “big” measurement problem as a pseudo-problem. They urge us to resist the call for a general dynamical account of how measurements result in definite outcomes and endorse their “totally random” response to question (II) instead. Though arguments from authority will not carry much weight in these matters, we note that a prominent member of the Copenhagen camp did endorse this very answer. In an essay originally published in 1954, Wolfgang Pauli wrote: “Like an ultimate fact without any cause, the individual outcome of a measurement is . . . in general not comprehended by laws” (Pauli, 1994, p. 32, quoted by Gilder, 2008, p. 169). Bub and Pitowsky thus do not see the “big” measurement problem as a problem but as a lesson quantum mechanics has taught us about how the world behaves. It is not a puzzle to be solved but a feature to be embraced.

The same is true for the “small” measurement problem. As we will explain in more detail in Chapter 6, our response to the question of which observables get assigned definite values is that this is decided by the experimenter. The experimenter chooses which question to put to nature. It is, we submit, not a task for physics to account for such decisions. It certainly never was for classical mechanics. One might object that this is only because, in classical mechanics, such decisions are irrelevant to the way in which observables acquire definite values. In classical mechanics, the decision to perform this rather than that measurement amounts to no more than choosing to ascertain the value of this rather than that observable, where all these values are taken to be pre-given. In quantum mechanics, by contrast, a decision to perform a measurement to obtain a value for one observable will typically preclude the assignment of a definite value to another observable that could have been measured instead. Such decisions thus select the catalog of values of observable quantities used in the construction of our conception of reality.

Yet the difference between classical and quantum mechanics on this score is not as big as it seems. Even though according to classical mechanics decisions about which measurement to perform do not influence the possible outcomes of that measurement (or any other measurement that could have been performed instead), these decisions clearly do affect the *actual* catalog

²⁷ The “small” measurement problem is familiar to Everettians as the preferred-basis problem.

of values of observables obtained and hence the reality constructed with the help of them. What door a contestant picks on Monty Hall’s “Let’s make a deal” does not affect what is behind that door but it obviously does make a difference whether that door has a check for a thousand dollars or a goat behind it (cf., e.g., [Janssen & Pernice, 2020](#)). As this simple example illustrates, the difference between quantum mechanics and classical mechanics is *not* whether decisions about what to measure affect what is real or what happens. What is different is that in quantum mechanics such decisions affect what is real or what happens at a deeper level than in classical mechanics, namely at the level of possibilities. Rather than insisting that quantum mechanics provide an account of the kind of decisions we did not expect classical physics to account for, however, we can take the fact that it does not as another lesson quantum mechanics taught us about the world—just as it taught us that, at rock bottom, the world is irreducibly stochastic.

The “small” measurement problem, we will argue in [Section 6.5](#), is only a problem if we insist that we, as observers, can be completely written out of our description of the physical world. This was certainly the goal of classical physics. But quantum mechanics tells us that we and our measuring devices are entangled with the rest of the world, that the probabilities quantum mechanics gives us are ultimately all marginal probabilities obtained once we trace out the degrees of freedom of us and our measuring devices. To use another slogan we will unpack in [Section 6.5](#), quantum mechanics is hard to square with the classical ideal of the “view from nowhere.” To bolster our case, we will briefly discuss the analysis of two historians of early-modern science of how this classical ideal took shape in the 17th century through developments in optics from Kepler to Descartes ([Gal & Chen-Morris, 2013](#), Ch. 1, “Science’s disappearing observer”). Their analysis suggests that while this classical ideal proved to be extremely useful for the development of optics and physics more generally from the 17th through the 19th century, it was never inevitable.

Whether or not we are able to convince the reader that either the “big” or the “small” or both measurement problems are pseudo-problems—not bugs but features—we hope to convince the reader of the more general thesis that quantum mechanics, at its core, is a new framework for handling probabilities (cf. notes [6](#), [16](#) and [29](#)). Our main argument for this thesis will come from our analysis—in terms of Bub’s correlation arrays and Pitowsky’s correlation polytopes—of correlations found in measurements on systems in a special but informative quantum state in a simple experimental setup due to David [Mermin \(1981, 1988\)](#) to test a Bell inequality (see [Figure 2.6](#) for the correlation array for Mermin’s example of correlations violating this particular inequality).

We introduce special raffles to determine which of these quantum correlations can be simulated by local hidden-variable theories (see Figure 2.11 for an example of tickets for such raffles and Figures 2.12 and 2.14 for examples of the correlation arrays that raffles with different mixes of these tickets give rise to). These raffles will serve as toy models of local hidden-variable theories. They are both easy to visualize and tolerably tractable mathematically (see Section 4.2). They also make for a natural classical counterpart to the statistical ensembles characterized by density operators introduced in von Neumann’s formulation of quantum mechanics in *Wahrscheinlichkeitstheoretischer Aufbau*, which were themselves inspired by von Mises’s classical statistical ensembles. The “quantum raffles” in the title of our book refer to these quantum statistical ensembles. In Section 6.5 we will rely heavily on a comparison between our raffles and these “quantum raffles” to bring out what is new about quantum theory compared to classical theory. It is with malice aforethought that we construct our raffles in such a way that they can serve as simple examples of local hidden-variable theories suffering from the “big” (albeit easily cured) measurement problem but not the “small” one (see note 9 in Section 3.2).

The quantum state we will focus on is that of two particles of spin s entangled in the so-called singlet state (with zero overall spin). For most of our argument it suffices to consider the $s = \frac{1}{2}$ case. In Chapter 2 we focus on this case. Our analysis, however, is informed (and justified) at several junctures by our analysis in Chapter 4 of cases with arbitrary integer or half-integer values of s . In Section 4.1, we analyze the quantum correlations for $s > \frac{1}{2}$; in Section 4.2 we analyze raffles designed to simulate as many features as possible of these quantum correlations.

In Chapter 5, returning to the special case that $s = \frac{1}{2}$, we show how our analysis in Chapters 2 and 4 can be adapted to the more common experimental setup used to test the CHSH inequality (Clauser, Horne, Shimony, & Holt, 1969). The advantage of the Mermin setup, as we will see in Chapter 2, is that in that case the classes of correlations allowed by quantum mechanics and by local hidden-variable theories can be pictured in ordinary three-dimensional space. The corresponding picture for the setup to test the CHSH inequality is four-dimensional.

The class of all correlations in the Mermin setup that cannot be used for sending signals faster than light can be represented by an ordinary three-dimensional cube, the so-called *non-signaling cube* for this setup; the class of correlations allowed by quantum mechanics by an *elliptope* contained within this cube; those allowed by classical mechanics by a *tetrahedron* contained

within this ellipsope (see Figures 2.15 and 2.16). This provides a concrete example of the way in which Pitowsky and others (see, e.g., Goh, Kaniewski, Wolfe, Vértesi, Wu, Cai, Liang, & Scarani, 2018) have used nested polytopes²⁸ to represent the convex sets formed by these classes and subclasses of correlations (compare the cross-section of the non-signaling cube, the tetrahedron and the ellipsope in Figure 2.8 to the familiar Vitruvian-man-like cartoon in Figure 2.5). Such polytopes completely characterize these classes of correlations whereas the familiar Bell inequalities (in the case of local hidden-variable theories) or Tsirelson bounds (in the case of quantum mechanics) only provide partial characterizations.

As Pitowsky pointed out in the preface of *Quantum Probability—Quantum Logic*:

The possible range of values of classical correlations is constrained by linear inequalities which can be represented as facets of polytopes, which I call “classical correlation polytopes.” These constraints have been the subject of investigation by probability theorists and statisticians at least since the 1930s, though the context of investigation was far removed from physics (Pitowsky, 1989a, p. IV).

The constraint expressed by the non-linear ellipsope inequality has likewise been investigated by probability theorists and statisticians before in contexts far removed from physics. As we will see in Chapter 3, it can be found in a paper by Udney Yule (1897) on what are now called Pearson correlation coefficients as well as in papers by Ronald A. Fisher (1924) and Bruno de Finetti (1937). Yule, like Pearson, was especially interested in applications in evolutionary biology (see Chapter 3, notes 2 and 3). We illustrate the results of these statisticians with a simple physics example, involving a balance beam with three pans containing different weights (see Figure 3.5). These antecedents in probability theory and statistics provide us with our strongest argument for the thesis that the Hilbert space formalism of quantum mechanics is best understood as a general framework for handling probabilities in a world in which only some observables can be assigned definite values.²⁹

²⁸ The term was coined by Alicia Boole Stott, George Boole’s mathematician daughter, who was introduced to the mathematics of higher dimensions by her brother-in-law, Charles Howard Hinton, also a mathematician (who was at the University of Minnesota in the late 1890s) and a notorious polygamist. We are grateful to Louisa Gilder for alerting us to the relevant Wikipedia entries.

²⁹ As the authors of a book on quantum information intended for a broad audience write: “the mathematical formalisms [sic] underlying quantum theory can be precisely and usefully viewed as an extension of probability theory” (Rieffel & Polak, 2011, p. 331, cf. note 16 above). The story of Pearson, Yule, Fisher and De Finetti in Chapter 3 illustrates their lament

In Section 2.6 we show that it follows directly from the geometry of Hilbert space that the correlations found in our simple quantum system are constrained by the elliptope inequality and do not saturate the non-signaling cube. This derivation of the elliptope inequality is thus a derivation *from within* quantum mechanics.

Popescu & Rohrlich (1994) and others have asked why quantum mechanics does not allow *all* non-signaling correlations. They introduced an imaginary device, now called a PR box, that exhibits non-signaling correlations stronger than those allowed by quantum mechanics.³⁰ Several authors have looked for information-theoretic principles that would reduce the class of all non-signaling correlations to those allowed by quantum mechanics and rule out devices such as PR boxes (Clifton, Bub, & Halvorson, 2003; Bub, 2016, Ch. 9; Cuffaro, 2020). Such principles would allow us to derive the elliptope inequality *from without*.³¹

What the result of Yule and others shows is that the elliptope inequality expresses a general constraint on the possible correlations between three arbitrary random variables. It has nothing to do with quantum mechanics per se. As such it provides an instructive example of a kinematical constraint encoded in the geometrical structure of Hilbert space, just as time dilation and length contraction provide instructive examples of kinematical constraints encoded in the geometrical structure of Minkowski space-time.

After summarizing the results presented in Chs. 2–5 in Section 6.1, we return to this and other analogies between quantum mechanics and special relativity in the remainder of Chapter 6. In Section 6.2, we take a closer look at the interplay between *from within* and *from without* approaches to understanding fundamental features of quantum mechanics. In Section 6.3 we present our take on the new kinematics of quantum mechanics. In Section 6.4, as mentioned above, we give some examples of puzzles of the old quantum theory solved with the help of the new kinematics of the new quantum mechanics. In Section 6.5, we address the thorny issue of measurement, arguing in support of Bub and Pitowsky’s claim that what they identified as the “big” and “small”

that “the close relationship between the formal structures underlying quantum mechanics and probability theory is surprisingly neglected” (ibid.).

³⁰ See Figure 2.3 for the correlation array for a PR box. Figures 2.9 and 2.10 show that it is impossible to design tickets for a raffle that could simulate the correlations generated by a PR box.

³¹ We took the within/without terminology from the chorus of Bob Dylan’s song “The Mighty Quinn”: “Come all without, come all within. You’ll not see nothing like the mighty Quinn.” Could “Mighty Quinn” be an oblique but prescient reference to a quantum computer?

measurement problems should be treated not as problems to be solved but as lessons to be learned from quantum mechanics. While both the “small” and the “big” problem point to striking new features of the quantum world, we will argue that the novelty revealed by the former is more profound. One way to describe this new feature is that a quantum state gives us a family of probability distributions that do not add up to one joint probability distribution. Put in the language preferred by Bub and Pitowsky, it gives us a collection of Boolean algebras that cannot be combined into one overarching Boolean algebra. We will unpack these slogans in Chapter 6. Finally, in Chapter 7, we briefly summarize the views defended in Chapter 6 on the basis of the results presented in Chapters 2–5. In line with Bub’s own views, we intend our defense of (our version of) Bubism to double as a way to make sense of the views of Niels Bohr (cf. note 10 above).³²

We want to make one more observation before we get down to business. As already noted above, it is not surprising that the correlations found in measurements on pairs of particles of (half-)integer spin s in the singlet state do not saturate the non-signaling cube. No such correlations between three random variables could. What is surprising is that these correlations saturate the ellipsope *even in the spin- $\frac{1}{2}$ case*. This is in striking contrast to the correlations that can be generated with the raffles designed to simulate the quantum correlations. In the spin- $\frac{1}{2}$ case, the correlations allowed by our raffles are all represented by points inside the tetrahedron inscribed in the ellipsope. As we will see in Chapter 3, this is because there are only two possible outcomes in the spin- $\frac{1}{2}$ case, $\pm 1/2$ (in units of \hbar , Planck’s constant divided by 2π). In the spin- s case, there are $2s + 1$ possible outcomes: $-s, -s + 1, \dots, s - 1, s$ (again, in units of \hbar). With considerable help from the computer (see the flowchart in Figure 4.7 and the discussion of its limitations in Section 4.2.4), we generated figures showing that, with increasing s , the correlations allowed by the raffles designed to simulate the quantum correlations are represented by polyhedra that get closer and closer to the ellipsope (see Figures 4.11, 4.13 and 4.17 for $s = 1, 3/2, 2, 5/2$). That the quantum correlations already fully saturate the ellipsope in the spin- $\frac{1}{2}$ case is due to a remarkable feature of quantum mechanics: it allows a sum to have a definite value even if the individual terms in this sum do not (see Section 3.2).

³² For the views of one of us on the basic framework of Bohr’s interpretation of quantum mechanics, see Cuffaro (2010).



Chapter 2

Representing distant correlations by correlation arrays and polytopes

Peeling and tasting quantum bananas in the Mermin-style setup • Trying to simulate the quantum correlations with classical raffles • Nested classes of correlations: non-signaling cube, quantum ellipsope, classical tetrahedron.

2.1 Taking Mermin to Bananaworld

In the preface and in Chapter 1, we promised that this chapter would be accessible to readers without any background in physics and would not presuppose any mathematics beyond basic high-school algebra and geometry (and would involve absolutely no calculus!). We intend to make good on that promise. But before we start, in Section 2.2, with our elementary introduction to some of the most puzzling facts about the physical world that quantum mechanics confronts us with, we need to say a few words about how our approach fits with earlier attempts to explain these matters to a broader audience.¹

In addition to *Bananaworld* (Bub, 2016), our approach owes much to the work of David Mermin (especially 1981, 1988). Like Mermin and Bub (and many other popularizers), we will focus on so-called *Bell inequalities*, named after John S. Bell (1964), who was the first to formulate one in the context of quantum mechanics. As we will explain in detail in the course of this chapter, quantum mechanics predicts that such inequalities will be violated by the results of certain measurements on various quantum systems. These violations have consistently been found in an ongoing series of remarkable experiments stretching back to the early 1970s. An excellent account of these developments, written for a general audience and based on interviews both with those who formulated the inequalities and those who did the experiments showing that they are violated can be found in *The Age of Entanglement* by Louisa Gilder (2008, Chs. 29–31, pp. 250–289), a book we highly recommend.²

¹ See also our comments in the preface, especially about the relation between our approach and the approach taken in Albert (1992).

² The term *entanglement* (*Verschränkung* in the original German) was introduced by Schrödinger (1935). In quantum mechanics, the state of a composite system $A + B$ can in general not be decomposed into separate states for the two subsystems, A and B . Instead,

The specific Bell inequality tested in most of these experiments was formulated by John Clauser, Michael Horne, Abner Shimony and Richard Holt (1969). Like the inequality originally proposed by Bell, the CHSH inequality—to use the acronym by which it has come to be known—is a bound on the strength of distant correlations allowed by so-called *local hidden-variable theories*.³ The reason these inequalities are violated in quantum mechanics is that the distant correlations found in measurements on pairs of systems in entangled quantum states (see note 2) are stronger than the bounds set by these inequalities. In Section 2.5, we will introduce special raffles that can serve as toy models of local hidden-variable theories. For now, it suffices to say that, in a local hidden-variable theory, outcomes of measurements are always predetermined by variables that (a) are not included in the quantum description of a system (hence: hidden) and (b) cannot be affected by superluminal signals (i.e., signals traveling faster than light; this is what “local” means in this context).

The setup used to test the CHSH inequality involves *two parties* (affectionately known as Alice and Bob in much of the literature on quantum foundations, quantum information and quantum computing), *two settings per party* of some measuring device,⁴ and *two outcomes per setting* (labeled ‘0’ and ‘1’, ‘+’ and ‘−’, or ‘up’ and ‘down’). Bell (1964, pp. 18–19) originally considered three rather than four settings, which we can label $\{\hat{a}, \hat{b}, \hat{c}\}$. In Bell’s setup, one party performs measurements using the pair $\{\hat{a}, \hat{b}\}$ while the other uses $\{\hat{b}, \hat{c}\}$. In the CHSH setup the two parties use two pairs that have no setting in common, $\{\hat{a}, \hat{b}\}$ and $\{\hat{a}', \hat{b}'\}$ in our notation. Mermin (1981, 1988) kept Bell’s three settings but in his setup both parties use all three settings rather than just two

the states of the components A and B are inextricably intertwined or “entangled” in the state of the compound system $A + B$. This linkage persists no matter how far the two subsystems are separated, famously leading Einstein to refer to the phenomenon as “spooky action at a distance” (*spukhafte Fernwirkung*; letter to Max Born of March 3, 1947; Born, 1971, p. 158). In an oft-quoted passage, Schrödinger (1935, p. 555) wrote that he “would not call [entanglement] *one* but *the* characteristic trait of quantum mechanics” (quoted, e.g., in Bub & Bub, 2018, p. 9). We will focus on a particular entangled state of two identical particles characterized by their intrinsic angular momentum or *spin*. Eq. (2.6.38) gives this state for the case of spin- $\frac{1}{2}$ particles. This is the only quantum state we need for our purposes in this chapter.

³ The Pearson correlation coefficient, which we will introduce in Chapter 3 (see Eq. (3.1.9)), provides a measure of the strength of a correlation.

⁴ E.g., a polarizer to measure the polarization of photons or a Du Bois magnet to measure the component of the spin of a particle in certain direction, as used in variations on the Stern-Gerlach experiment (cf. Section 6.5, especially note 41 and Figure 6.1).

of them. He derived a Bell inequality for this setup, so simple that even those without Mermin’s pedagogical skills can explain it to a general audience.⁵

We use a Mermin-style setup to illustrate the power of some of the tools in *Bananaworld*. We represent the correlations Mermin considered by *correlation arrays*, the workhorse of *Bananaworld*, and parametrize these arrays in such a way that they and the correlations they encode can, in turn, be represented as points in convex sets in so-called *non-signaling cubes* (a geometrical representation of all correlations one can imagine for a given setup that cannot be exploited to send superluminal signals). This approach was pioneered by Itamar Pitowsky (1989b) in *Quantum Probability—Quantum Logic* and is promoted in *Bananaworld*.⁶

The representation of classes of correlations in terms of (the geometry of) convex sets is well-established in the quantum foundations literature (see, e.g., Goh et al., 2018). Our efforts can be seen as another attempt to bring this approach to a broader audience by applying it to Mermin’s particularly simple and instructive example. Here is what makes it so. Since the CHSH setup uses four different settings, its non-signaling cube is a hypercube in four dimensions, which is hard if not impossible to visualize. The Mermin-style setup only uses three different settings and its non-signaling cube is an ordinary cube in three dimensions, which is easy to visualize. The convex set representing the non-signaling correlations allowed classically in this case is a tetrahedron spanned by four of the eight vertices of the three-dimensional non-signaling cube (see Figure 2.15); the convex set representing those allowed quantum-mechanically is an elliptope enclosing this tetrahedron (see Figure 2.16).

In *Bananaworld*, *settings* become *peelings*, *outcomes* become *tastes*, and *parties* become characters from *Alice in Wonderland* (Alice stars as Alice, the White Rabbit as Bob). Fictitious quantum bananas can be peeled “from the stem end (*S*)” or “from the top end (*T*)” and can only taste “ordinary (“o” or 0)” or “intense, incredible, indescribably delicious (“i” or 1)” (Bub, 2016, pp. 8–9, see also p. viii).⁷

⁵ For an excellent textbook treatment of the Bell inequality for this setup and the quantum-mechanical results maximally violating it, see McIntyre (2012, sec. 4.1, pp. 97–102).

⁶ In addition to *Quantum Probability—Quantum Logic*, Bub (2016, p. 120) cites Pitowsky (2006), his contribution to a *Festschrift* for Bub, as well as Pitowsky (1986, 1989a, 1991, 2008).

⁷ Betraying his information-theoretic leanings, Bub (2016) occasionally refers to *inputs* and *outputs* (both taking on the values 0 and 1) rather than peelings and tastes (see, e.g., p. 51, Figure 3.1).

Bub’s banana-peeling scheme suffices for the discussion of the CHSH inequality as well as for the analysis of so-called PR boxes, at least those of the original design of their inventors, [Popescu & Rohrlich \(1994\)](#). We will use the term “PR box” for any hypothetical system one can imagine that would allow *superquantum correlations* ([Bub, 2016](#), p. 106), i.e., correlations that are even stronger than those allowed by quantum mechanics but are still non-signaling (and would thus still be compatible with relativity theory). Like the CHSH setup, the original design of a PR box involves two parties, two settings per party, and two outcomes per setting. Bub’s scheme also works for the analysis of correlations that arise in measurements on so-called GHZ states ([Greenberger, Horne, & Zeilinger, 1989](#)). While these measurements involve three rather than two parties,⁸ they still fit the mold of two settings per party and two outcomes per setting. The Mermin-style setup breaks this mold by using (the same) *three* settings for two parties.

To recreate the Mermin-style setup in *Bananaworld* we introduce a new banana-peeling scheme (see Section 2.2). Our scheme allows infinitely many different settings, readily translates into the actual physics of measuring spin with Du Bois magnets (see note 4 above and Chapters 4–6 below), and, as such, highlights elements of spherical symmetry in the Mermin-style setup that turn out to be key to their quantum-mechanical analysis (see Sections 2.6 and 4.1).

2.2 Correlations found when peeling and tasting pairs of quantum bananas

Imagine a species of banana that grows in pairs on special banana trees in Bananaworld. These bananas can only taste yummy or nasty. Yet we cannot say that they come in two flavors, as they only acquire a definite flavor once they are peeled and tasted. We use these bananas in a long series of peel-and-taste experiments following a protocol that closely matches the one followed in Mermin’s (1981, 1988) setup for testing a Bell inequality (cf. Section 2.1). Our Mermin-style setup in Bananaworld is illustrated in Figure 2.1.⁹

We pick a pair of bananas, still joined at the stem, from the banana tree. We separate them and give one each to two chimps, Alice and Bob. Once they

⁸ Bub’s illustrator, his daughter Tanya, has the Cheshire Cat (starring as Clio) peel the third GHZ banana ([Bub, 2016](#), pp. 122–123, Clio and Charlie are introduced on p. 8).

⁹ This figure closely matches Figure 4.2 in the treatment of the Mermin-style setup in the quantum-mechanics textbook by [McIntyre \(2012\)](#), p. 99, cf. note 5 and note 28).

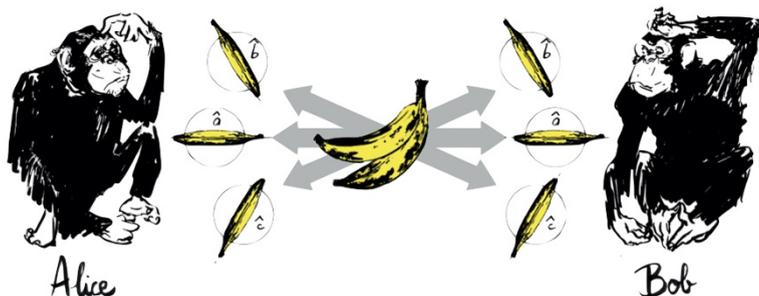


Fig. 2.1 Mermin-style setup in *Bananaworld* (I). *Two parties: the chimps Alice and Bob. Three settings per party: three peelings, $(\hat{a}, \hat{b}, \hat{c})$, given by three unit vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$, in the corresponding peeling directions (i.e., the direction of the line going from the top to the stem of the banana while it is being peeled). In Mermin’s specific example, the angles φ_{ab} between \mathbf{e}_a and \mathbf{e}_b , φ_{ac} between \mathbf{e}_a and \mathbf{e}_c , and φ_{bc} between \mathbf{e}_b and \mathbf{e}_c are all equal to 120° but we will also consider other values for these angles. Drawing: Laurent Taudin with a nod to Andy Warhol.*

have received their respective bananas, they randomly and independently of one another pick a particular *peeling*, defined by the *peeling direction*, i.e., the direction of the line going from the top to the stem of the banana while it is being peeled. Alice and Bob are instructed not to change the orientation of their bananas while peeling so that it is unambiguous which peeling they are using. In the Mermin-style setup, Alice and Bob get to choose between three peelings, labeled \hat{a} , \hat{b} and \hat{c} , represented by unit vectors, \mathbf{e}_a , \mathbf{e}_b and \mathbf{e}_c , with the angles φ_{ab} , φ_{ac} and φ_{bc} between them all equal to 120° (see Figure 2.1). Once they have randomly chosen one of these three peelings, they point the stem of their banana in the direction of the corresponding unit vector and peel their banana (it does not matter whether they peel from the top or from the stem). When done peeling, Alice and Bob reposition their bananas and take a bite to determine whether they taste yummy or nasty (see Figure 2.2). The whole procedure is then repeated with a fresh pair of bananas from the banana tree.

In each run of this peel-and-taste experiment, Alice and Bob record that run’s ordinal number, the peeling chosen (\hat{a} , \hat{b} or \hat{c}) and the taste of their banana, using “+” for “yummy” and “−” for “nasty”. Every precaution is taken to ensure that, as long as there are more bananas to be peeled and tasted, Alice and Bob cannot communicate. While they are peeling and tasting, the only contact between them is that the bananas they are given come from pairs originally joined at the stem on the banana tree.

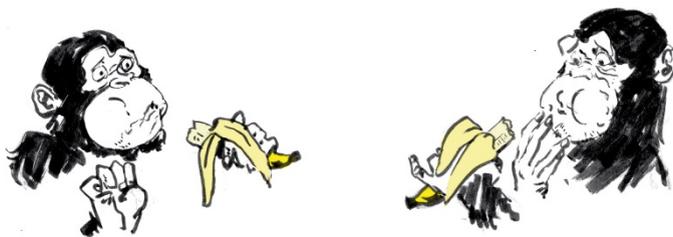


Fig. 2.2 Mermin-style setup in *Bananaworld* (II). *Two outcomes per setting*: the tastes “yummy” (+) or “nasty” (−) for different peeling directions. The peeling and tasting is done by the chimps Alice and Bob. Drawing: Laurent Taudin.

When all bananas are peeled and tasted, Alice and Bob are allowed to compare notes. Just looking at their own records, they see nothing out of the ordinary—just a sequence of pluses and minuses as random as if they had faked their results by tossing a fair coin for every run. Comparing their records, however, they note that, every time they happened to choose the same peeling (in roughly 33% of the total number of runs), their results are perfectly anti-correlated. Whenever one banana tasted yummy, its twin tasted nasty. In and of itself, this is not particularly puzzling. Maybe our bananas always grow in pairs in which one tastes yummy and the other tastes nasty and it is random which one goes to which chimp. This simple explanation, however, is ruled out by another striking correlation Alice and Bob discover while poring over their data. When they happened to peel differently (in roughly 66% of the runs), their results were positively correlated, albeit imperfectly. In 75% of the runs in which they used different peelings, their bananas tasted the same (Mermin, 1981, p. 86).¹⁰ The tastes of two bananas coming from one pair thus depend on the angle between the peeling directions used. This is certainly odd but one could still imagine that our bananas are somehow pre-programmed to respond differently to different peelings and that the set of pre-programmed responses is different for the two bananas in one pair. What Mermin’s Bell inequality

¹⁰ In Mermin’s (1981, p. 86) example, there is a perfect (positive) correlation in runs in which the two parties use the same setting and an imperfect anti-correlation in runs in which they use different settings (see also Mermin, 1988, pp. 135–136). To get Mermin’s original example, we should have used our pairs of bananas to represent entangled pairs of photons and let “peel and taste bananas using different peeling directions” stand for “measure the polarization of these photons along different axes”. We got our variation on Mermin’s example by having pairs of bananas represent pairs of spin- $\frac{1}{2}$ particles entangled in the singlet state and letting “peel and taste bananas using different peeling directions” stand for “measure spin components of these particles along different axes” (see Section 2.6).

shows, however, is that it is impossible to pre-program twin bananas in such a way that they would produce the specific correlations found in this case. Such correlations, however, can and have been produced with quantum twins (see Section 2.6). Given that they persist no matter how far we imagine Alice and Bob to be apart, another explanation of these curious correlations is also unavailing: it would take a superluminal signal for the taste of one banana peeled a certain way to either affect the way the other banana is peeled or affect its taste when peeled that way. In short, these correlations cannot be accounted for on the basis of any local hidden-variable theory.

2.3 Non-signaling correlation arrays

The correlations found in the Mermin-style setup can be represented in a correlation array consisting of nine cells, one for each of the nine possible combinations of peelings (see Figure 2.6 in Section 2.4). These cells form a grid with three rows for Alice's three peeling directions and three columns for Bob's. Each cell has four entries, giving the probabilities of the four possible pairs of tastes for that cell's combination of peelings (the entries in one cell thus always sum to 1).

Since *Bananaworld* focuses on setups with two settings per party, all correlation arrays in it have only four cells. These cells form a 2×2 grid with rows for Alice peeling from the stem and from the top and columns for Bob peeling from the stem and from the top. Before we turn to the 3×3 correlation arrays for the correlations found in our Mermin-style setup, we go over some properties of these simpler 2×2 correlation arrays.

An example of a correlation array of this simple form is the one in Figure 2.3 for a Popescu-Rohrlich or PR box (see Section 2.1; Bub, 2016, p. 89, Table 4.2, we switched Alice and Bob to match the usual convention that the index labeling the rows of a matrix comes before the index labeling its columns). This correlation array plays an important role in *Bananaworld* and is central to its sequel, Tanya and Jeffrey Bub's (2018) enchanting *Totally Random*. A version of it is prominently displayed on many pages of this graphic novel (Bub & Bub, 2018, pp. 15, 21, 33, 95, 115, 181, 200 and 227).¹¹

¹¹ The version in *Totally Random* differs from the version in Figure 2.3 (which follows *Bananaworld*) in that it is the cell in the upper-left corner rather than the one in the lower-right corner that is different from the other three cells. Instead of the four entries in each cell in Figure 2.3, the cells in *Totally Random* just have “=” for perfectly correlated and “≠” for perfectly anti-correlated. This follows the convention in Figs. 3.3 and 3.5 in *Bananaworld* (Bub, 2016, p. 57, p. 59).

		Bob				
		\hat{a}'		\hat{b}'		
Alice	\hat{a}	+	-	+	-	
	\hat{a}	+	$\frac{1}{2}$	0	$\frac{1}{2}$	0
\hat{a}	-	0	$\frac{1}{2}$	0	$\frac{1}{2}$	
\hat{b}	\hat{b}	+	$\frac{1}{2}$	0	0	$\frac{1}{2}$
	\hat{b}	-	0	$\frac{1}{2}$	$\frac{1}{2}$	0

Fig. 2.3 Correlation array for a Popescu-Rohrlich box.

In *Bananaworld*, the PR-box correlations in Figure 2.3 are realized with the help of PR bananas growing in pairs on PR banana trees. The settings $\{\hat{a}, \hat{b}\}$ and $\{\hat{a}', \hat{b}'\}$ now stand for Alice and Bob peeling their bananas from the stem (S) or from the top (T) (with $\hat{a} = \hat{a}' = S$ and $\hat{b} = \hat{b}' = T$). These peelings could be replaced by two of the peeling directions we introduced. In realizations of this PR box, we can (but do not have to) use the same pair of settings for Alice and Bob (in the case of the CHSH setup we definitely need different pairs of settings; see Chapter 5).

In *Totally Random*, the PR-box correlations in Figure 2.3 are realized with the help of an imaginary device, named for the inventors of the PR box, the “Superquantum Entangler PR01”. This gadget, which looks like a toaster, has slots for two US quarters. When we insert two ordinary coins, the PR01 turns them into a pair of entangled “quoins” (Bub & Bub, 2018, p. 7). The different settings now stand for Alice and Bob holding their quoins heads-up ($\hat{a} = \hat{a}'$) or tails-up ($\hat{b} = \hat{b}'$) when tossing them. The outcomes are the quoins landing heads or tails. What makes this a realization of a PR box is that the quoins invariably land with the same side facing up, except when both are tossed being held heads-up (\hat{a}, \hat{a}'), in which case they always land with opposite sides facing up (recall that in *Totally Random* it is the cell in the upper-left corner of the correlation array that is different from the other three).

The correlations between the outcomes found in a PR box—be it between the tastes of a pair of PR bananas or the landings of a pair of quoins—are preserved no matter how far its two parts are pulled apart.¹²

¹² Part of what makes it interesting to contemplate entangled quoins or bananas is that we are free to choose *when* to toss or peel them whereas with entangled photons or spin- $\frac{1}{2}$

An important feature of correlation arrays (no matter how many cells they have or how many entries each cell has) is that, in many cases, they allow us to see at a glance whether or not the correlations they represent can be used for the purposes of instant messaging or, more precisely, superluminal signaling. Suppose Alice wants to use the peeling of a pair of PR bananas to instant-message the answer to some “yes/no” question to Bob. They agree ahead of time that Alice will peel \hat{a} if the answer is “yes” and \hat{b} if it is “no”.¹³ This scheme will not work. No matter how Bob peels his banana, he cannot tell from its taste whether Alice peeled hers \hat{a} or \hat{b} . This is easy to prove with the help of the correlation array in Figure 2.3. Suppose Bob peels \hat{b}' (essentially the same argument works if Bob peels \hat{a}'). In that case, the correlation array in Figure 2.3 tells us that the *marginal probability* of Bob finding + if Alice were to peel \hat{a} (trying to transmit “yes”) is¹⁴

$$\Pr(+_B|\hat{a}\hat{b}') = \Pr(++|\hat{a}\hat{b}') + \Pr(-+|\hat{a}\hat{b}') = 1/2 + 0 = 1/2. \quad (2.3.1)$$

This is the same as the marginal probability of him finding + if Alice were to peel \hat{b} (trying to transmit “no”):

$$\Pr(+_B|\hat{b}\hat{b}') = \Pr(++|\hat{b}\hat{b}') + \Pr(-+|\hat{b}\hat{b}') = 0 + 1/2 = 1/2. \quad (2.3.2)$$

Hence Bob cannot tell on the basis of the outcomes of his measurements whether Alice is peeling \hat{a} (for “yes”) or \hat{b} (for “no”). Inspection of the correlation array in Figure 2.3 shows that *all* marginal probabilities like those in Eqs. (2.3.1)–(2.3.2) are equal to $1/2$ in this case. PR boxes—whether realized with the help of magic bananas, quoins, or other systems—cannot be used for instant messaging.

particles we have no choice but to measure their polarization or spin as soon as they arrive at our detectors.

¹³ It does not matter in what order Alice and Bob peel their bananas. The correlations in the correlation array in Figure 2.3 represent constraints on possible combinations of outcomes found by Alice and Bob, *not* some mechanism through which the outcome of one peeling would cause the outcome of the other.

¹⁴ Throughout this volume we will use the notation $\Pr(X|Y)$ for the *conditional* probability of X given Y . In Eqs. (2.3.1)–(2.3.2) and (2.3.3) below, Y stands for the pair of peelings used by Alice and Bob (in that order) and X stands either for the taste \pm of one of their bananas (indicated by a subscript A or B) or for the tastes of both bananas (in which case the first entry refers to the taste found by Alice and the second to the taste found by Bob and there is no need for the subscripts A or B). The conditional probabilities on the left-hand sides of Eqs. (2.3.1)–(2.3.2) are called *marginal* probabilities because we are interested in the probability of the taste found by Bob irrespective of the taste found by Alice.

Correlations that do not allow instant messaging are called *non-signaling*. It will be convenient to use this term for their correlation arrays as well. The correlations and correlation arrays for a PR box are always non-signaling. In fact, this is what makes these hypothetical devices intriguing. Even though they would give rise to correlations stronger than those allowed by quantum mechanics, they would not violate special relativity's injunction against superluminal signaling.

Generalizing the results in Eqs. (2.3.1)–(2.3.2), we can state the following *non-signaling condition*:

A correlation in a setup with two parties, any number of settings per party and two outcomes per setting is non-signaling if the probabilities in both rows and both columns of all cells in its correlation array add up to $1/2$.

This is a sufficient but not a necessary condition. A correlation array with entries

1	0	0	1
0	0	0	0
0	0	0	0
1	0	0	1

is non-signaling even though the entries in half the rows and columns of its cells add up to 1 while the entries in the other half add up to 0 (see Bub, 2016, p. 109, Table 5.2, for a similar example). The relevant marginal probabilities, however, are still equal to each other. For instance,

$$\begin{aligned} \Pr(+_{\text{B}}|\hat{a}\hat{b}') &= \Pr(+_{\text{B}}|\hat{b}\hat{b}') = 0, \\ \Pr(-_{\text{B}}|\hat{a}\hat{b}') &= \Pr(-_{\text{B}}|\hat{b}\hat{b}') = 1. \end{aligned} \tag{2.3.3}$$

In Chapter 4, we will encounter correlation arrays for setups with three outcomes per setting that are non-signaling even though not all rows and columns of its cells add up to the same number (see Figure 4.8 in Section 4.2.2).¹⁵

¹⁵ Consider the entries for the three correlation arrays, labeled (a), (b) and (c), below:

$$\begin{aligned} \text{(a)} \quad & \begin{array}{|c|c|c|c|} \hline 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ \hline \end{array}, & \text{(b)} \quad & \begin{array}{|c|c|c|c|} \hline 6/10 & 1/10 & 2/10 & 1/10 \\ \hline 1/10 & 2/10 & 1/10 & 6/10 \\ \hline 2/10 & 1/10 & 6/10 & 1/10 \\ \hline 1/10 & 6/10 & 1/10 & 2/10 \\ \hline \end{array}, & \text{(c)} \quad & \begin{array}{|c|c|c|c|} \hline 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ \hline \end{array}. \end{aligned}$$

2.4 The non-signaling cube, the classical tetrahedron and the quantum ellipsope

Any cell in a correlation array that satisfies the non-signaling condition we stated in Section 2.3 can be parametrized by a variable with values running from -1 to $+1$. Figure 2.4 shows such a cell for Alice using setting \hat{a} and Bob using setting \hat{b} . Let $-1 \leq \chi_{ab} \leq 1$ be the variable parametrizing this cell. If $\chi_{ab} = 0$, the results of Alice and Bob are uncorrelated; if $\chi_{ab} = -1$, they are perfectly correlated; if $\chi_{ab} = 1$, they are perfectly anti-correlated. On the basis of these three special values (0 , -1 and 1), we will call χ_{ab} and parameters like it *anti-correlation coefficients*. Once we have introduced the *Pearson correlation coefficient* in Chapter 3, we will see that it makes sense to call such parameters anti-correlation coefficients for *any* value between -1 and 1 (see Eqs. (3.2.5)–(3.2.6)).

		Bob	
		\hat{b}	
Alice		+	-
	+	$\frac{1}{4}(1 - \chi_{ab})$	$\frac{1}{4}(1 + \chi_{ab})$
-	$\frac{1}{4}(1 + \chi_{ab})$	$\frac{1}{4}(1 - \chi_{ab})$	

Fig. 2.4 Cell in a non-signaling correlation array parametrized by $-1 \leq \chi_{ab} \leq 1$.

A 2×2 non-signaling correlation array such as the one in Figure 2.3 for a PR box, with four cells of the form of Figure 2.4, can be parametrized by four anti-correlation coefficients

$$-1 \leq \chi_{aa'} \leq 1, \quad -1 \leq \chi_{ab'} \leq 1, \quad -1 \leq \chi_{ba'} \leq 1, \quad -1 \leq \chi_{bb'} \leq 1. \quad (2.4.1)$$

These correlation arrays not only violate the non-signaling condition but are actually signaling (see Bub, 2016, p. 109, Table 5.1, for an example equivalent to (a)). If there were a system producing the distant correlations in (a), be it pairs of bananas or pairs of coins, one pair would suffice for Alice and Bob to transmit one bit of information to the other party instantaneously. With a system producing the distant correlations in (b), several pairs would be needed to do so with some fidelity: (b) can be thought of as a noisy version of (a). With a system producing the distant correlations in (c), Bob could signal to Alice but Alice could not signal to Bob: Bob finds $+$ for setting \hat{a} and $-$ for setting \hat{b} regardless of Alice's setting.

Such a correlation array can thus be represented by a point in a hypercube in four dimensions with the anti-correlation coefficients serving as that point's Cartesian coordinates. The correlation array for a PR box is represented by one of the vertices of this hypercube:

$$(\chi_{aa'}, \chi_{ab'}, \chi_{ba'}, \chi_{bb'}) = (-1, -1, -1, 1). \quad (2.4.2)$$

The four-dimensional hypercube that represents the class of all non-signaling correlations in this setup (two parties, two settings per party, two outcomes per setting) is an example of a so-called *non-signaling polytope*, which can be defined (typically in some higher-dimensional space) for setups with two parties, any number of settings and any number of outcomes per setting.

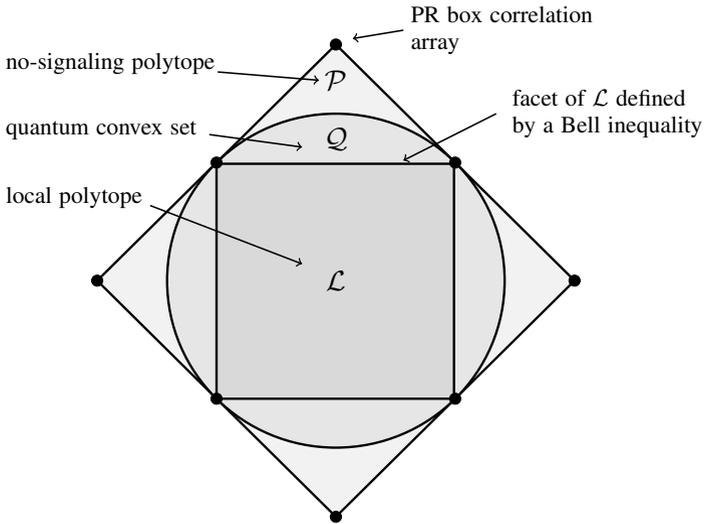


Fig. 2.5 Schematic representation, for some arbitrary experimental setup, of the set \mathcal{P} of all non-signaling correlations, the subset $\mathcal{Q} \subset \mathcal{P}$ of those allowed quantum-mechanically and the subset $\mathcal{L} \subset \mathcal{Q}$ of those allowed classically. One of the facets of \mathcal{L} represents a Bell inequality. The vertex of the non-signaling cube where this Bell inequality is maximally violated represents a PR box for the setup under consideration (Bub, 2016, p. 107, Figure 5.2).

Figure 2.5 gives a schematic representation of the non-signaling polytope for such a setup. The outer square and everything inside of it (the non-signaling polytope \mathcal{P}) represents the set of all non-signaling correlations. The inner

square and everything inside of it (the local polytope \mathcal{L}) represents the set of all non-signaling correlations allowed classically (i.e., by a local hidden-variable theory). The circle in between these two squares and everything inside of it (the quantum convex set \mathcal{Q}) represents the set of all correlations allowed quantum-mechanically. One of the facets of \mathcal{L} represents a Bell inequality, a bound on the strength of the correlations allowed classically. The vertex of the non-signaling cube where this bound is maximally violated represents a PR box for the setup under consideration.

		Bob					
		\hat{a}	\hat{b}			\hat{c}	
Alice	\hat{a}	+	-	+	-	+	-
	+	0	$\frac{1}{2}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
-	$\frac{1}{2}$	0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	
\hat{b}	+	$\frac{3}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$	$\frac{3}{8}$	$\frac{1}{8}$
-	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$	0	$\frac{1}{8}$	$\frac{3}{8}$	
\hat{c}	+	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$
-	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$	0	

Fig. 2.6 Mermin correlation array (cf. Figures 2.1 and 2.2 for our Mermin-style setup in Bananaworld and note 10 for how it relates to Mermin’s original setup).

Figure 2.6 shows the correlation array for our version of Mermin’s example of a quantum correlation violating a Bell inequality. We will refer to it as the *Mermin correlation array*. Its nine cells form a 3×3 grid. The cells along the diagonal of this grid, when Alice and Bob peel the same way, show a perfect anti-correlation. The six off-diagonal cells, when Alice and Bob peel differently, all show the same imperfect positive correlation. It is easy to see that this correlation array is non-signaling: the entries in both rows and both columns of all nine cells add up to $\frac{1}{2}$. Concisely put, this correlation (array) has *uniform marginals*.

		Bob					
		\hat{a}		\hat{b}		\hat{c}	
Alice		+	-	+	-	+	-
	\hat{a}	+	0	$\frac{1}{2}$	$\frac{1}{4}(1-\chi_{ab})$	$\frac{1}{4}(1+\chi_{ab})$	$\frac{1}{4}(1-\chi_{ac})$
-		$\frac{1}{2}$	0	$\frac{1}{4}(1+\chi_{ab})$	$\frac{1}{4}(1-\chi_{ab})$	$\frac{1}{4}(1+\chi_{ac})$	$\frac{1}{4}(1-\chi_{ac})$
\hat{b}	+	$\frac{1}{4}(1-\chi_{ab})$	$\frac{1}{4}(1+\chi_{ab})$	0	$\frac{1}{2}$	$\frac{1}{4}(1-\chi_{bc})$	$\frac{1}{4}(1+\chi_{bc})$
	-	$\frac{1}{4}(1+\chi_{ab})$	$\frac{1}{4}(1-\chi_{ab})$	$\frac{1}{2}$	0	$\frac{1}{4}(1+\chi_{bc})$	$\frac{1}{4}(1-\chi_{bc})$
\hat{c}	+	$\frac{1}{4}(1-\chi_{ac})$	$\frac{1}{4}(1+\chi_{ac})$	$\frac{1}{4}(1-\chi_{bc})$	$\frac{1}{4}(1+\chi_{bc})$	0	$\frac{1}{2}$
	-	$\frac{1}{4}(1+\chi_{ac})$	$\frac{1}{4}(1-\chi_{ac})$	$\frac{1}{4}(1+\chi_{bc})$	$\frac{1}{4}(1-\chi_{bc})$	$\frac{1}{2}$	0

Fig. 2.7 A non-signaling correlation array for three settings (peelings) and two outcomes (tastes) parametrized by the anti-correlation coefficients $-1 \leq \chi_{ab} \leq 1$ (for the $\hat{a}\hat{b}$ and $\hat{b}\hat{a}$ cells), $-1 \leq \chi_{ac} \leq 1$ (for the $\hat{a}\hat{c}$ and $\hat{c}\hat{a}$ cells) and $-1 \leq \chi_{bc} \leq 1$ (for the $\hat{b}\hat{c}$ and $\hat{c}\hat{b}$ cells).

The Mermin correlation array in Figure 2.6 is a special case of the more general correlation array in Figure 2.7. The three cells along the diagonal are the same, all showing a perfect anti-correlation (i.e., its diagonal elements are 0 and its off-diagonal elements are $1/2$). Moreover, cells on opposite sides of the diagonal are the same. This correlation array and the correlations they encode can thus be parametrized by three anti-correlation coefficients of the kind introduced in Figure 2.4. In the specific example of the Mermin-style setup in Figure 2.6, the three anti-correlation coefficients have the same value:

$$\chi_{ab} = \chi_{ac} = \chi_{bc} = -1/2. \quad (2.4.3)$$

The class of all non-signaling correlations in the Mermin-style setup can be visualized as a cube in ordinary three-dimensional space with the correlation coefficients, χ_{ab} , χ_{ac} and χ_{bc} , providing the three Cartesian coordinates of points in this cube. The non-signaling correlations allowed classically can be represented by a tetrahedron spanned by four of the eight vertices of this

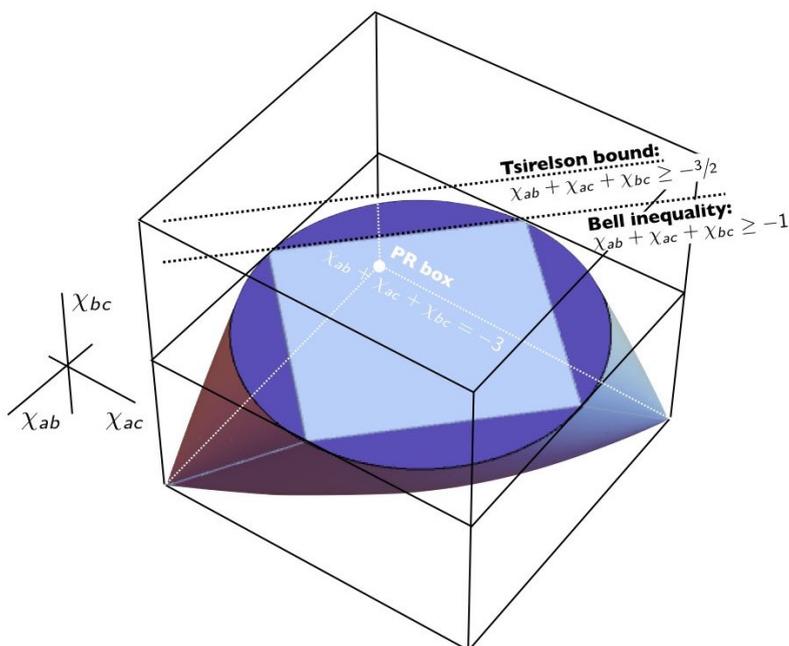


Fig. 2.8 Concrete version of the diagram in Figure 2.5 for correlations found in the Mermin-style setup. The figure shows the cross-section $\chi_{bc} = 0$ of the classical tetrahedron and the ellipsope in a non-signaling cube in ordinary three-dimensional space (cf. Figures 2.15 and 2.16 below). For the time being, ignore the dotted lines labeled “Bell inequality” and “Tsirelson bound”. These will be explained in Sections 2.5 and 2.6. The bound on the quantum correlations is maximally violated in the point $(-1, -1, -1)$, which thus represents the PR box for this setup.

non-signaling cube (see Figure 2.15 in Section 2.5); those allowed quantum-mechanically by an ellipsope enclosing this tetrahedron (see Figure 2.16 in Section 2.6). Figure 2.8 shows the cross-section $\chi_{bc} = 0$ of this non-signaling cube, the classical tetrahedron and the ellipsope. This cross-section has exactly the form of the cartoonish rendering in Figure 2.5 of the Vitruvian-man-like structure of the local polytope \mathcal{L} and the quantum convex set \mathcal{Q} inside the non-signaling polytope \mathcal{P} . In the next two sections, we will show in detail how one arrives at the classical tetrahedron and the quantum ellipsope in the Mermin-style setup.

2.5 Raffles meant to simulate the quantum correlations and the classical tetrahedron

To decide whether or not some correlation array is allowed classically (or quantum-mechanically), Bub, as he explains in the opening chapter of *Bananaworld* (Bub, 2016, p. 10), checks whether or not it can be simulated with classical (or quantum-mechanical) resources. Though we will use a more direct approach to find classes of correlations allowed quantum-mechanically (see Sections 2.6 and 4.1), we will adopt a variation on Bub's imitation game to find classes of correlations allowed classically (i.e., by some local hidden-variable theory).

We will use special raffles to simulate the correlations found in our quantum banana peel-and-taste experiments. These raffles involve baskets of tickets such as the ones in Figure 2.11. All tickets list the outcomes for both parties and for all settings in the setup under consideration. We randomly draw a ticket of the appropriate kind from a basket with many such tickets. We tear this ticket in half and randomly decide which side goes to Alice and which side goes to Bob. Alice and Bob then decide, randomly and independently of each other, which setting they will use. They record the outcome for that setting printed on their half of the ticket. We repeat this procedure a great many times.

Raffles of this kind provide a criterion for determining whether or not a certain correlation is allowed classically:^{16,17}

A correlation array is allowed by a local hidden-variable theory if and only if there is a raffle (i.e., a basket with the appropriate mix of tickets) with which we can simulate that correlation array following the protocol described above.

¹⁶ In Section 6.5 we will see that there is an extra bonus to discussing classical theory in terms of such raffles. It makes for a natural comparison between local hidden-variable theories and John von Neumann's (1927b) formulation of quantum theory in terms of statistical ensembles characterized by density operators on Hilbert space. *Single-ticket raffles*, i.e., raffles with baskets of tickets that are all the same, are the classical analogues of pure states in quantum mechanics; *mixed raffles*, i.e., raffles with baskets with different tickets, are the analogues of mixed states.

¹⁷ By using the imagery of baskets with different mixes of tickets, we admittedly sweep a mathematical subtlety under the rug: the fractions of different types of tickets in a basket will always be rational numbers. To simulate the quantum correlations we are interested in, however, we need to allow fractions that are real numbers. In Section 4.2.1 we will introduce a different mechanism for selecting tickets that gets around this problem (see Figure 4.5). From a practical point of view, the restriction to rationals is perfectly harmless, since the rationals are dense in the reals. This means that for any real number we can always find a rational number arbitrarily close to it.

Invoking this criterion, we can easily show that a PR box with the correlation array in Figure 2.3 is not allowed classically.¹⁸ These correlations place impossible demands on the design of the tickets for a raffle that would simulate them (see Figure 2.9).

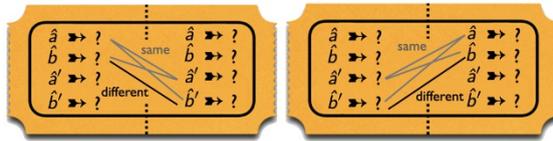


Fig. 2.9 When we try to design a raffle ticket for a PR box with the correlation array in Figure 2.3, with $\hat{a} \neq \hat{a}'$ and $\hat{b} \neq \hat{b}'$, we inevitably run into a contradiction no matter whether we start from the left or the right side of the ticket.

The perfect positive correlation between the outcomes for three of the four possible combinations of settings ($\hat{a}\hat{a}'$, $\hat{a}\hat{b}'$ and $\hat{b}\hat{a}'$) requires that the outcomes printed on the ticket for \hat{a} and \hat{b} on one side are the same as the outcomes for \hat{a}' and \hat{b}' on the other side. That makes it impossible for the outcomes for \hat{b} and \hat{b}' on opposite sides of the ticket to be different as required by the perfect anti-correlation for the remaining combination of settings ($\hat{b}\hat{b}'$).

This is even easier to see if we try to design tickets for a raffle to simulate the correlations found with PR boxes for which $\hat{a} = \hat{a}'$ and $\hat{b} = \hat{b}'$, like the ones used in *Bananaworld* and *Totally Random* (where $\hat{a} = \hat{a}'$ and $\hat{b} = \hat{b}'$ stand for “peeling from the top/bottom” and “flipping a quoin holding it heads-up/tails-up”, respectively).

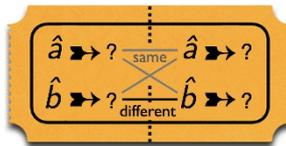


Fig. 2.10 When we try to design a raffle ticket for a PR box with the correlation array in Figure 2.3, with $\hat{a} = \hat{a}'$ and $\hat{b} = \hat{b}'$, we inevitably run into a contradiction.

Figure 2.10 shows that there can be no such tickets. Since the outcomes have to be the same for three of the four combinations of settings ($\hat{a}\hat{a}$, $\hat{a}\hat{b}$ and $\hat{b}\hat{a}$), all four outcomes listed on the ticket should be the same. But the correlation array for the PR box requires opposite outcomes for the remaining fourth combination of settings ($\hat{b}\hat{b}$).

¹⁸ Essentially the same argument can already be found in Rastall (1995, p. 970).

Figure 2.11 shows four different types of tickets, labeled (i) through (iv), for raffles meant to simulate correlations found in the Mermin-style setup in which Alice and Bob choose from the same three settings $(\hat{a}, \hat{b}, \hat{c})$ with two possible outcomes each $(+, -)$. Since in all setups that we will examine Alice and Bob find opposite results whenever they use the same setting, the outcomes on one side of the ticket dictate the outcomes on the other. That reduces the number of different ticket types to $2^3 = 8$. Given that it is decided randomly which side of a ticket goes to Alice and which side to Bob, two tickets that differ only in that the left and the right side are swapped are two equivalent versions of the same ticket type. This further reduces the number of different ticket types to four. As illustrated in Figure 2.11, we chose the ones that have $+$ for the first setting (\hat{a}) on the left side of the ticket.

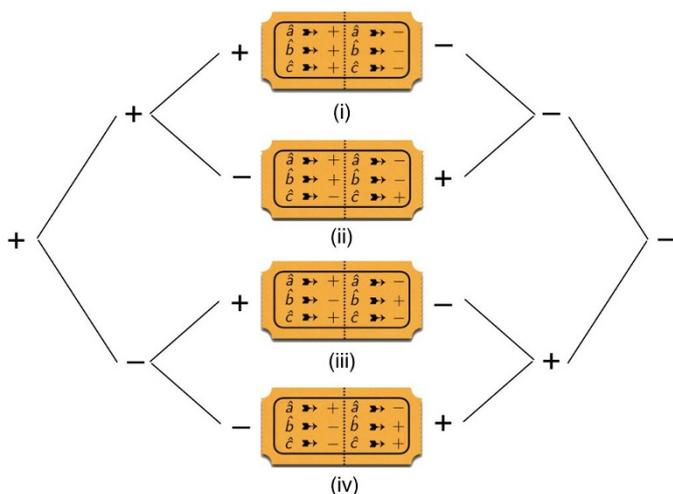


Fig. 2.11 The four different raffle tickets for three settings and two outcomes. Given the protocol of our raffles, two tickets that differ only in that their left and right sides are swapped are the same ticket.

Figure 2.12 shows the correlation arrays for raffles with baskets containing only one of the four types of tickets in Figure 2.11. We will call these raffles *single-ticket raffles* and those with baskets containing tickets of different types *mixed raffles* (cf. note 16). The design of our raffles—here and elsewhere in this volume—guarantees that the correlations between the outcomes found by Alice and Bob are non-signaling. The process of tearing a ticket in half, sending one side to Alice at one location and the other to Bob at another

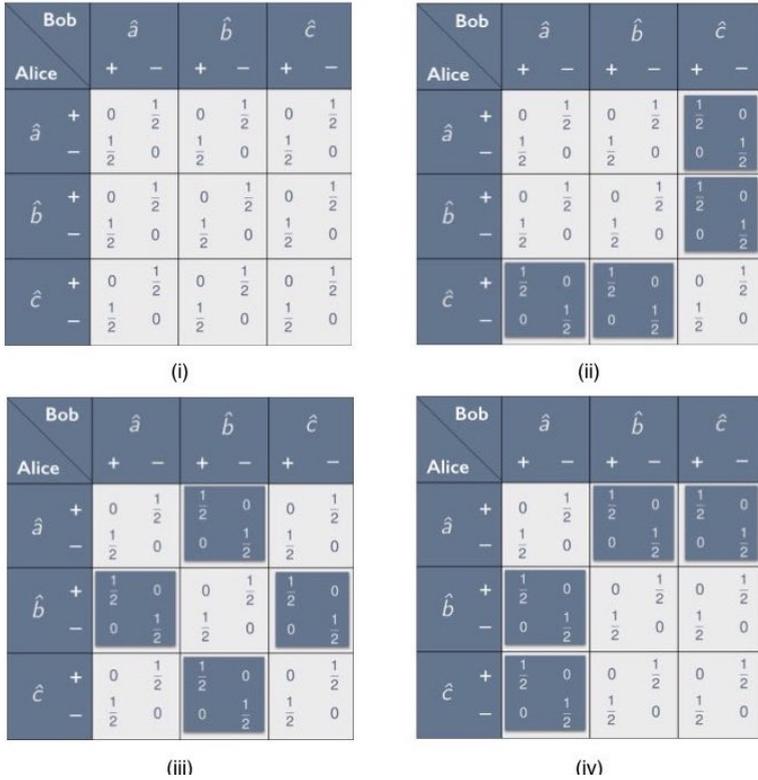


Fig. 2.12 Correlation arrays for the four different single-ticket raffles in the Mermin-style setup. In blue-on-white cells the outcomes are perfectly anti-correlated; in white-on-blue cells they are perfectly correlated.

location, and then having them inspect their respective ticket stubs at their respective locations does not allow Alice and Bob to instant-message each other through some manipulation of their ticket stubs agreed upon ahead of time. That the correlations found with these raffles are indeed non-signaling is borne out by the correlation arrays in Figure 2.12. The entries in both rows and both columns of all cells in these correlation arrays add up to $\frac{1}{2}$. In other words, these raffles all give uniform marginals.

The entries of correlation arrays like those in Figure 2.12 form 6×6 matrices. These matrices are symmetric. They stay the same when we switch rows and columns. This is true both for single-ticket and mixed raffles. All raffles

we will consider have this property. This too follows directly from the design of these raffles. It is simply because Alice and Bob are as likely to get the left or the right side of any ticket.

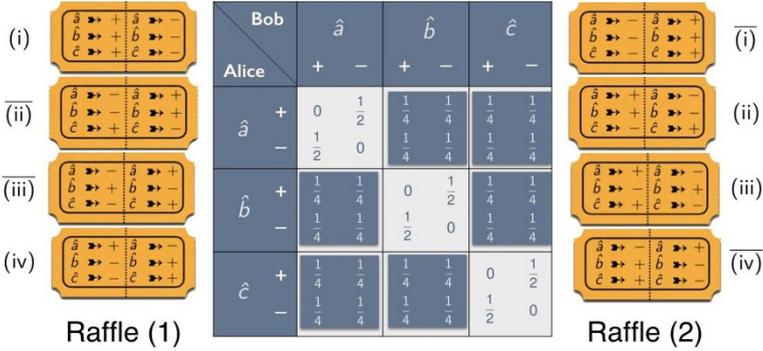


Fig. 2.13 Two raffles leading to the same correlation array (in blue-on-white cells the outcomes are perfectly anti-correlated; in white-on-blue cells they are completely uncorrelated). In both raffles, whenever a ticket is drawn, Alice gets the left and Bob gets the right side. In addition to tickets (i)–(iv) in Figure 2.11 we now have four more tickets, labeled (i)–(iv) and obtained by switching the left and the right side of the tickets (i)–(iv). Raffle (1) has equal numbers of tickets of type (i), (ii), (iii) and (iv). Raffle (2) has equal numbers of tickets of type (i), (ii), (iii) and (iv).

Before we continue our analysis, we show that changing the protocol of our raffles so that Alice is always given the left side and Bob is always given the right side of any ticket does not give rise to correlation arrays with symmetric associated matrices that cannot be simulated with our more economical protocol—more economical because it requires fewer ticket types. For the alternative protocol, we need four more tickets, labeled (i) through (iv), that differ from their counterparts (i) through (iv) in that the left and right sides of the ticket have been swapped. Figure 2.13 shows two raffles for this alternative protocol. Raffle (1) has equal numbers of tickets of type $\{(i), (ii), (iii), (iv)\}$. The matrix associated with the correlation array for this raffle is symmetric. That means that we get the same correlation array if we swap the left and the right sides of all tickets in raffle (1). This turns raffle (1) into raffle (2) with equal numbers of tickets of type $\{(i), (ii), (iii), (iv)\}$. Any raffle mixing raffles (1) and (2) will also give that same correlation array. Consider the special case of a raffle with equal numbers of all eight tickets. This raffle is equivalent to a basket with equal numbers of tickets $\{(i), (ii), (iii), (iv)\}$ with the understanding that it is decided at random which side of the ticket goes to Alice

and which side goes to Bob. This construction works for any correlation array with a symmetric associated matrix that we can produce using the protocol in which Alice always get the left side and Bob always get the right side of a ticket. This shows that we can indeed produce any such correlation array using our more economical protocol.

		Bob			
		\hat{a}	\hat{b}		
Alice		+	-	+	-
	\hat{a}	+	0	$\frac{1}{2}$	$\frac{3}{8}$
-		$\frac{1}{2}$	0	$\frac{1}{8}$	$\frac{3}{8}$
\hat{b}	+	$\frac{3}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$
	-	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$	0
\hat{c}	+	$\frac{3}{8}$	$\frac{1}{8}$	0	$\frac{1}{2}$
	-	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{2}$	0

Raffle (a)

		Bob			
		\hat{a}	\hat{b}		
Alice		+	-	+	-
	\hat{a}	+	0	$\frac{1}{2}$	$\frac{1}{3}$
-		$\frac{1}{2}$	0	$\frac{1}{6}$	$\frac{1}{3}$
\hat{b}	+	$\frac{1}{3}$	$\frac{1}{6}$	0	$\frac{1}{2}$
	-	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$	0
\hat{c}	+	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$
	-	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{3}$

Raffle (b)

Fig. 2.14 Correlation arrays for raffles with different mixes of the four tickets in Figure 2.11. Raffle (a) has 25% type-(i) tickets and 75% type-(iv) tickets. Raffle (b) has 33% each of type-(ii) through type-(iv) tickets. The entries in these correlation arrays are the weighted averages of the corresponding entries in the correlation arrays for the single-ticket raffles in Figure 2.12. Blue-on-white cells are the same as the corresponding cells in the Mermin correlation array in Figure 2.6, white-on-blue cells are different.

We are now ready to prove the most important claim we want to make about these raffles. *There is no raffle that can simulate the Mermin correlation array in Figure 2.6.* Figure 2.14 shows the results of two unsuccessful attempts to cook up such a raffle. In the first, we take a basket with 25% tickets of type (i) and 75% of type (iv). This results in correlation array (a) in Figure 2.14. Every entry in the correlation array for this mixed raffle is the weighted average of the corresponding entries in the correlation arrays for the single-ticket raffles (i) and (iv).¹⁹ This raffle correctly simulates all but two cells of the Mermin correlation array. We get the same result if we replace tickets (iv) by tickets (ii) or (iii), the only difference being that now two other cells will differ from the corresponding ones in the Mermin correlation array. The best we can do overall, it seems, is to take a basket with 33% each of tickets (ii) through (iv).

¹⁹ For a formal justification of this intuitively plausible rule, see Section 4.2.1, especially Eqs. (4.2.7)–(4.2.9).

This results in correlation array (b) in Figure 2.14. The entries in the correlation array for this mixed raffle are the straight averages of the corresponding entries in the correlation arrays for the single-ticket raffles (ii) through (iv). Like the Mermin correlation array we are trying to simulate, this one has the same positive correlation in all six off-diagonal cells but the correlation is weaker than in the Mermin case.²⁰

Showing that various mixed raffles fail to simulate the Mermin correlation array may suggest but obviously does not prove that *all of them* will fail. To prove this general claim, we consider the sum $\chi_{ab} + \chi_{ac} + \chi_{bc}$ of the anti-correlation coefficients for a raffle. From the tickets in Figure 2.11 we can read off the values of χ_{ab} , χ_{ac} and χ_{bc} for the four single-ticket raffles. Since all cells in the correlation arrays for the four single-ticket raffles in Figure 2.12 show either a perfect anti-correlation or a perfect positive correlation, the anti-correlation coefficients for these single-ticket raffles can only take on the values ± 1 . If the outcome for \hat{a} and the outcome for \hat{b} on opposite sides of the ticket are *opposite* (giving rise to perfectly anti-correlated $\hat{a}\hat{b}$ and $\hat{b}\hat{a}$ cells in the correlation array), $\chi_{ab} = 1$. If the outcome for \hat{a} and the outcome for \hat{b} on opposite sides of the ticket are the *same* (giving rise to perfectly correlated $\hat{a}\hat{b}$ and $\hat{b}\hat{a}$ cells), $\chi_{ab} = -1$. The values for χ_{ac} and χ_{bc} are found in the same way. The results are collected in Table 2.1.

ticket	χ_{ab}	χ_{ac}	χ_{bc}
(i)	+1	+1	+1
(ii)	+1	-1	-1
(iii)	-1	+1	-1
(iv)	-1	-1	+1

Table 2.1 Values of the anti-correlation coefficients parametrizing the off-diagonal cells of the correlation arrays (i) through (iv) in Figure 2.12 for single-ticket raffles with tickets (i) through (iv) in Figure 2.11.

Table 2.1 tells us that, in a single-ticket raffle, the sum $\chi_{ab} + \chi_{ac} + \chi_{bc}$ can only take on the value 3 (for ticket type (i)) or -1 (for ticket types (ii), (iii) or (iv)). For mixed raffles, $\chi_{ab} + \chi_{ac} + \chi_{bc}$ is the weighted average of the value of

²⁰ Interpreting the parameters $-\chi_{ab}$, $-\chi_{ac}$ and $-\chi_{bc}$ as Pearson correlation coefficients (see Chapter 3), we see that for the correlation array (b) in Figure 2.14 all three of them are $1/3$, whereas for the Mermin correlation array in Figure 2.6 they are $1/2$.

$\chi_{ab} + \chi_{ac} + \chi_{bc}$ for these four single-ticket raffles, with the weights given by the fractions of each of the four tickets in the raffle.²¹ Hence, for any mix of tickets, this sum must lie between -1 and 3 :

$$-1 \leq \chi_{ab} + \chi_{ac} + \chi_{bc} \leq 3. \quad (2.5.1)$$

The first of these inequalities, giving the lower bound on $\chi_{ab} + \chi_{ac} + \chi_{bc}$, is the analogue of the CHSH inequality for the Mermin-style setup. As we will see in Chapter 5, it is also the form in which Bell (1964) originally derived the Bell inequality. The CHSH-type Bell inequality is violated by the Mermin correlation array in Figure 2.6. In that case, $\chi_{ab} = \chi_{ac} = \chi_{bc} = -1/2$ (see Eq. (2.4.3)) and their sum equals $-3/2$. As we will see in Section 2.6, this is the maximum violation of this inequality allowed by quantum mechanics. Note that the absolute minimum value of $\chi_{ab} + \chi_{ac} + \chi_{bc}$ is -3 . This value is allowed neither classically nor quantum-mechanically. It is the value reached with the (hypothetical) PR box for this setup.

The values of χ_{ab} , χ_{ac} and χ_{bc} in Table 2.1 for tickets (i) through (iv) can be used as the Cartesian coordinates of four vertices in the non-signaling cube for the Mermin-style setup. These are the vertices labeled (i) through (iv) in Figure 2.15. The vertex $(-1, -1, -1)$ represents the PR box for this setup (see Figure 2.8). The vertices (i) through (iv) span a tetrahedron forming the convex set of all raffles that can be obtained by mixing the four types of tickets. The sum $\chi_{ab} + \chi_{ac} + \chi_{bc}$ takes on its maximum value of 3 at the vertex for tickets of type (i) and its minimum value of -1 for the facet spanned by the vertices for tickets of types (ii), (iii) and (iv). The pair of inequalities in Eq. (2.5.1) tell us that all correlations that can be simulated with raffles with various mixes of tickets must lie in the region of the non-signaling cube between the vertex (i) and the facet (ii)-(iii)-(iv).

This is a necessary but not a sufficient condition for a correlation to be allowed by a local hidden-variable theory. As Figure 2.15 shows, there are three forbidden sub-regions in the region between vertex (i) and facet (ii)-(iii)-(iv). A full characterization of the class of correlations allowed classically requires three additional pairs of inequalities like the pair given in Eq. (2.5.1), corresponding to the other three vertices and the other three facets of the tetrahedron. The following four pairs of inequalities do fully characterize the tetrahedron:

$$-1 \leq \chi_{ab} + \chi_{ac} + \chi_{bc} \leq 3, \quad (2.5.2)$$

²¹ For a formal proof of this intuitively plausible result, see Section 4.2.1.

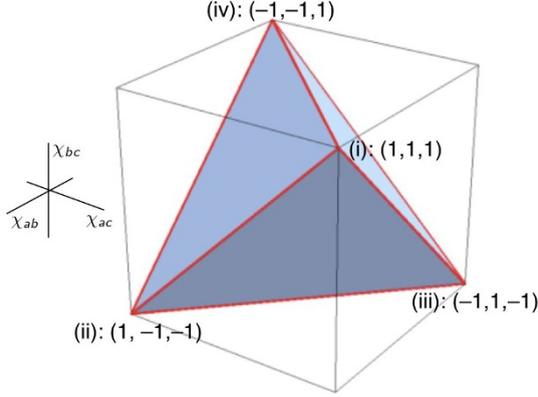


Fig. 2.15 Tetrahedron of triplets of anti-correlation coefficients $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ allowed by local hidden-variable theories in the Mermin-style setup.

a restriction to the part between facet (ii)-(iii)-(iv) and vertex (i);

$$-1 \leq \chi_{ab} - \chi_{ac} - \chi_{bc} \leq 3, \quad (2.5.3)$$

a restriction to the part between facet (i)-(iii)-(iv) and vertex (ii);

$$-1 \leq -\chi_{ab} + \chi_{ac} - \chi_{bc} \leq 3, \quad (2.5.4)$$

a restriction to the part between facet (i)-(ii)-(iv) and vertex (iii);

$$-1 \leq -\chi_{ab} - \chi_{ac} + \chi_{bc} \leq 3, \quad (2.5.5)$$

a restriction to the part between facet (i)-(ii)-(iii) and vertex (iv).

Using the symmetries of the tetrahedron we can easily get from any one of these pairs of inequalities to another. Another way to see this is to recall that the coordinates $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ are anti-correlation coefficients for different combinations of the measurement settings $(\hat{a}, \hat{b}, \hat{c})$ and to look at what happens when we flip the sign of the outcomes for one of these three settings. If we do this for \hat{a} , χ_{ab} and χ_{ac} pick up a minus sign and Eq. (2.5.2) turns into Eq. (2.5.5). If we do this for \hat{b} , χ_{ab} and χ_{bc} pick up a minus sign and Eq. (2.5.2) turns into Eq. (2.5.4). Finally, if we do this for \hat{c} , χ_{ac} and χ_{bc} pick up a minus sign and Eq. (2.5.2) turns into Eq. (2.5.3).

Mermin formulated a different inequality for this setup, one that implies the lower bound on the sum of anti-correlation coefficients in Eq. (2.5.1) but

requires an additional assumption. To derive Mermin’s inequality, we have to assume that *Alice and Bob randomly and independently of each other decide which setting to use in any run of the experiment* (whether with raffle tickets, spin- $\frac{1}{2}$ particles, or quantum bananas). This provision is part of the protocol we described in Section 2.2 but we had no need to invoke it so far. The CHSH-like inequality in Eq. (2.5.1) could be derived without it—and so, for that matter, can the CHSH inequality itself.

This is good news because it means that we can test these inequalities without having to change the settings in every run. We can make measurements for one pair of settings at a time, providing data for the correlation array one cell at a time. This is how Stuart Freedman and John Clauser (1972) originally tested the CHSH inequality. Changing the orientation of their polarizers was a cumbersome process.²² Because of this limitation of their equipment, the violation of the CHSH inequality they found could conceivably be blamed on the two photons generated as an entangled pair “knowing” ahead of time (i.e., the moment they separated) what the orientation of the polarizers would be with which they were going to be measured. To close this loophole, the settings should only be chosen once the photons are in flight. This was accomplished by Alain Aspect and his collaborators later in the 1970s and in the 1980s (Gilder, 2008, Ch. 31). We will not be concerned with the extensive experimental efforts to close this and other loopholes.²³

If we assume that Alice and Bob randomly and independently of each other decide which setting to use in each run,²⁴ the nine possible combinations of settings are equiprobable. Following Mermin (1981, pp. 86–87), we ask for the probability, $\Pr(\text{opp})$, that Alice and Bob find opposite results. Consider the Mermin correlation array in Figure 2.6. For the cells along the diagonal $\Pr(\text{opp}) = 1$ (the results are perfectly anti-correlated). For the off-diagonal cells $\Pr(\text{opp}) = 1/4$, the sum of the off-diagonal entries in those cells. Alice and Bob use the same setting in one out of three runs and different settings in two out of three. Hence, the probability of them finding opposite results is the weighted average of those two probabilities with weights $1/3$ and $2/3$,

²² For a photograph and a drawing of Freedman and Clauser’s apparatus see Gilder (2008, p. 231 and p. 262, respectively); for another photograph, see Kaiser (2011, p. 48); for a schematic drawing, see Freedman & Clauser (1972, p. 939, Figure 1).

²³ David Kaiser alerted us to a paper written by 20 authors (with Kaiser, Alan Guth and Anton Zeilinger listed in 17th, 18th and 20th place, respectively) about one of the latest initiatives in this ongoing effort (Handsteiner et al., 2017).

²⁴ We still do not need the stronger assumption that these decisions are made only *after* they receive their banana, their spin- $\frac{1}{2}$ particle, or their ticket stub.

respectively:

$$\Pr(\text{opp}) = 1/3 \cdot 1 + 2/3 \cdot 1/4 = 1/2. \quad (2.5.6)$$

Upon inspection of the four correlation arrays in Figure 2.12, however, we see that the minimum value for $\Pr(\text{opp})$ in a local hidden-variable theory is $5/9$. Here is how. In correlation array (i), the results in all nine cells are perfectly anti-correlated. In a single-ticket raffle with tickets of type (i), we thus have $\Pr(\text{opp}) = 1$. In each of the other three correlation arrays, there are five cells in which the results are perfectly anti-correlated and four in which they are perfectly correlated. In single-ticket raffles with tickets of type (ii), (iii), or (iv), we thus have $\Pr(\text{opp}) = 5/9$. For an arbitrary mix of tickets (i) through (iv), we therefore have the inequality

$$\Pr(\text{opp}) \geq 5/9. \quad (2.5.7)$$

This is the form in which Mermin states the Bell inequality for the setup we are considering.

It is easy to see that Eq. (2.5.7) implies the lower bound on the sum $\chi_{ab} + \chi_{ac} + \chi_{bc}$ in Eq. (2.5.1). Consider, once again, the general non-signaling correlation array in Figure 2.7 parametrized by χ_{ab} , χ_{ac} and χ_{bc} . Adding the off-diagonal elements in every cell and dividing by 9, as we are assuming that Alice and Bob use the settings of all nine cells with equal probability, we find

$$\begin{aligned} \Pr(\text{opp}) &= 3/9 + 2/9 \cdot 1/2 (1 + \chi_{ab}) \\ &\quad + 2/9 \cdot 1/2 (1 + \chi_{ac}) + 2/9 \cdot 1/2 (1 + \chi_{bc}) \\ &= 2/3 + 1/9 (\chi_{ab} + \chi_{ac} + \chi_{bc}). \end{aligned} \quad (2.5.8)$$

If $\Pr(\text{opp})$ must at least be $5/9$, then $\chi_{ab} + \chi_{ac} + \chi_{bc}$ cannot be smaller than -1 . Conversely, if $\chi_{ab} + \chi_{ac} + \chi_{bc} \geq -1$ and all nine combinations of the settings \hat{a} , \hat{b} and \hat{c} are equiprobable, then $\Pr(\text{opp}) \geq 5/9$.

Mermin's lower bound on the probability of finding opposite results may be easier to grasp for non-specialists than a lower bound on a sum of parameters such as χ_{ab} , χ_{ac} and χ_{bc} . The latter, however, does have its own advantages. First, as we just saw, it can be derived from weaker premises. Second, it immediately generates inequalities corresponding to other facets of the polyhedron of classically allowed correlations in the Mermin-style setup (see Eqs. (2.5.2)–(2.5.5)). Third, as we will show in detail in Chapter 5, it makes it easier to see the connection with the CHSH inequality.

2.6 The quantum correlations and the ellipsope

In this section, we investigate what correlations are allowed by quantum mechanics in the Mermin-style setup. In Section 2.4, we saw that, as long as a correlation array for this setup satisfies a non-signaling condition (i.e., the two entries in every row and every column of every cell add up to $1/2$), we can fully specify it using just three parameters, χ_{ab} , χ_{ac} and χ_{bc} , each having a value in the closed interval $[-1, +1]$ (see Figure 2.7). We noted that these parameters can be interpreted as anti-correlation coefficients (but postponed a formal proof till Chapter 3). In Section 2.5, we found that, in any single-ticket raffle, χ_{ab} , χ_{ac} and χ_{bc} must be either $+1$ or -1 (see Table 2.1 and the four correlation arrays in Figure 2.12). In quantum mechanics, it turns out, χ_{ab} , χ_{ac} and χ_{bc} can take on the values $+1$, -1 and *any value in between*. As we will show in Section 2.6.2, quantum mechanics sets these parameters equal to the cosines of the angles φ_{ab} , φ_{ac} and φ_{bc} between the three peeling directions \mathbf{e}_a , \mathbf{e}_b and \mathbf{e}_c :

$$\chi_{ab} = \cos \varphi_{ab}, \quad \chi_{ac} = \cos \varphi_{ac}, \quad \chi_{bc} = \cos \varphi_{bc}. \quad (2.6.1)$$

If φ_{ab} , φ_{ac} and φ_{bc} could be chosen independently of one another, the class of correlations allowed by quantum mechanics in the Mermin-style setup would saturate the non-signaling cube (in which case we could even realize the PR box represented by $\chi_{ab} = \chi_{ac} = \chi_{bc} = -1$). These angles, however, cannot be chosen independently of one another: elementary geometry tells us that picking two of them constrains our choice of the third. So quantum mechanics still puts restrictions on which triplets $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ are allowed. However, as we already noted in Section 2.4, these restrictions are not as stringent as the ones we found with our raffles (see Figure 2.8).

This is a remarkable result. Measurements on a collection (or ensemble) of copies of the same quantum system all prepared in the same state (in our example: pairs of bananas picked from a particular species of banana tree) produce correlations that we cannot reproduce in measurements on a collection of copies of the same classical system prepared in *any* mix of different states (such as a raffle with an arbitrary mix of different tickets). Moreover, the way in which quantum mechanics allows us to do this is, at least in this particular example, a natural generalization of what classical (i.e., local hidden-variable) theories allowed us to do. Again, instead of just allowing the values ± 1 for χ_{ab} , χ_{ac} and χ_{bc} , quantum mechanics allows the full range of values between -1

and $+1$, naturally given by the cosines of the angles φ_{ab} , φ_{ac} and φ_{bc} between peeling directions.

We will proceed as follows. In Section 2.6.1, simply assuming for the time being that quantum mechanics does indeed give Eq. (2.6.1) for the Mermin-style setup, we will derive the constraint this result puts on triplets $(\chi_{ab}, \chi_{ac}, \chi_{bc})$. This will lead us to the *elliptope inequality* in Eq. (2.6.9), a non-linear constraint on χ_{ab} , χ_{ac} and χ_{bc} , which is the counterpart in quantum mechanics of the four linear CHSH-type inequalities in Eqs. (2.5.2)–(2.5.5). In Section 2.6.2, we then introduce just enough of the Hilbert-space formalism of quantum mechanics to derive expressions for the probabilities entering the correlation array for the Mermin-style setup, from which we can read off Eq. (2.6.1) for the anti-correlation coefficients (in Section 4.1, we will give more general derivations of these results).

2.6.1 Getting beyond the classical tetrahedron: the elliptope inequality

To derive the constraint that quantum mechanics puts on triplets $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ parametrizing correlations found in our banana taste-and-peel experiments in the Mermin-style setup, we introduce the matrix χ of anti-correlation coefficients or *anti-correlation matrix* for short:

$$\chi \equiv \begin{pmatrix} 1 & \chi_{ab} & \chi_{ac} \\ \chi_{ab} & 1 & \chi_{bc} \\ \chi_{ac} & \chi_{bc} & 1 \end{pmatrix}. \quad (2.6.2)$$

Using Eq. (2.6.1), which we will derive from quantum mechanics in Section 2.6.2, and replacing cosines of angles between peeling directions by inner products of units vectors in those directions, we can write this matrix as

$$\chi = \begin{pmatrix} \cos \varphi_{aa} & \cos \varphi_{ab} & \cos \varphi_{ac} \\ \cos \varphi_{ab} & \cos \varphi_{bb} & \cos \varphi_{bc} \\ \cos \varphi_{ac} & \cos \varphi_{bc} & \cos \varphi_{cc} \end{pmatrix} = \begin{pmatrix} \mathbf{e}_a \cdot \mathbf{e}_a & \mathbf{e}_a \cdot \mathbf{e}_b & \mathbf{e}_a \cdot \mathbf{e}_c \\ \mathbf{e}_b \cdot \mathbf{e}_a & \mathbf{e}_b \cdot \mathbf{e}_b & \mathbf{e}_b \cdot \mathbf{e}_c \\ \mathbf{e}_c \cdot \mathbf{e}_a & \mathbf{e}_c \cdot \mathbf{e}_b & \mathbf{e}_c \cdot \mathbf{e}_c \end{pmatrix}. \quad (2.6.3)$$

A matrix of this form is called a *Gram matrix*. Writing the components of the three unit vectors as

$$\mathbf{e}_a = (a_x, a_y, a_z), \quad \mathbf{e}_b = (b_x, b_y, b_z), \quad \mathbf{e}_c = (c_x, c_y, c_z), \quad (2.6.4)$$

we can write it as

$$\chi = \begin{pmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{pmatrix} \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix}. \quad (2.6.5)$$

Introducing

$$L \equiv \begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_x \end{pmatrix}, \quad (2.6.6)$$

we can write Eq. (2.6.5) more compactly as

$$\chi = L^\top L, \quad (2.6.7)$$

where L^\top is the transposed of L . It follows that the determinant of χ is non-negative:

$$\det \chi = \det(L^\top L) = (\det L^\top)(\det L) = (\det L)^2 \geq 0. \quad (2.6.8)$$

Using Eq. (2.6.2) to evaluate the determinant of χ , we can rewrite this condition as

$$1 - \chi_{ab}^2 - \chi_{ac}^2 - \chi_{bc}^2 + 2\chi_{ab}\chi_{ac}\chi_{bc} \geq 0. \quad (2.6.9)$$

This inequality is the constraint we have been looking for. It is both a necessary and a sufficient condition for a triplet $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ parametrizing correlations found in the Mermin-style setup to be allowed by quantum mechanics. The region of the non-signaling cube picked out by this inequality is the elliptope shown in Figure 2.16, an “inflated” version of the tetrahedron of classically allowed triplets of correlation coefficients in Figure 2.15. We will therefore call it the *elliptope inequality*.

It is easy to verify that the elliptope contains the classical tetrahedron. The vertices (i) through (iv) with the values

$$(1, 1, 1), \quad (1, -1, -1), \quad (-1, 1, -1), \quad (-1, -1, 1), \quad (2.6.10)$$

for $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ all satisfy Eq. (2.6.9) with an equality sign:

$$1 - \chi_{ab}^2 - \chi_{ac}^2 - \chi_{bc}^2 + 2\chi_{ab}\chi_{ac}\chi_{bc} = 0. \quad (2.6.11)$$

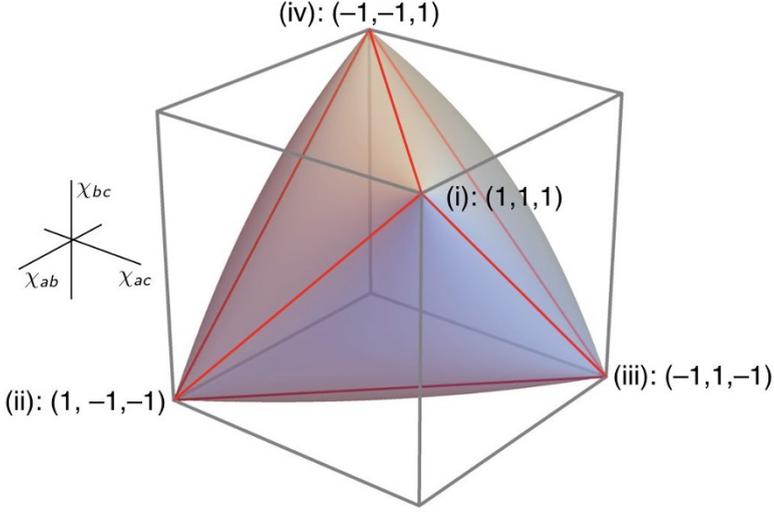


Fig. 2.16 Elliptope of triplets of anti-correlation coefficients $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ allowed by quantum mechanics in the Mermin-style setup.

So do the six lines connecting these four vertices. Consider the top and bottom face of the non-signaling cube, where $\chi_{bc} = 1$ and $\chi_{bc} = -1$, respectively. Eq. (2.6.11) then reduces to:

$$-\chi_{ab}^2 - \chi_{ac}^2 \pm 2\chi_{ab}\chi_{ac} = 0. \quad (2.6.12)$$

So the line $\chi_{ab} = \chi_{ac}$ on the $\chi_{bc} = 1$ face of the cube and the line $\chi_{ab} = -\chi_{ac}$ on the $\chi_{bc} = -1$ face of the cube satisfy Eq. (2.6.11). These are the lines connecting the vertices on these two faces of the cube. We similarly find that the other four lines connecting vertices of the tetrahedron satisfy Eq. (2.6.11): the $\chi_{ac} = \chi_{bc}$ line on the $\chi_{ab} = 1$ face, the $\chi_{ac} = -\chi_{bc}$ line on the $\chi_{ab} = -1$ face, the $\chi_{ab} = \chi_{bc}$ line on the $\chi_{ac} = 1$ face and the $\chi_{ab} = -\chi_{bc}$ line on the $\chi_{ac} = -1$ face.

Next, consider the planes $\chi_{ab} = 0$, $\chi_{ac} = 0$ and $\chi_{bc} = 0$. The cross-sections of those planes with the elliptope are the circles:

$$\chi_{ac}^2 + \chi_{bc}^2 = 1, \quad \chi_{ab}^2 + \chi_{bc}^2 = 1, \quad \chi_{ab}^2 + \chi_{ac}^2 = 1. \quad (2.6.13)$$

The last of these circles can be seen in Fig. 2.8, the cross-section $\chi_{bc} = 0$ of the non-signaling cube, the elliptope and the classical tetrahedron.

The volumes of the tetrahedron and the ellipptope are $8/3$ and $\pi^2/2$, respectively. Hence, the volume of the tetrahedron is only about 54% of that of the ellipptope. By this metric, the class of correlations allowed quantum-mechanically in this setup is thus almost twice the size of the class of correlations allowed by local hidden-variable theories.

Once again consider the special case that ellipptope inequality in Eq. (2.6.9) is satisfied with an equality sign. This has a simple geometrical interpretation. Note that $\det L = \pm V$, where V is the volume of the parallelepiped spanned by the three unit vectors \mathbf{e}_a , \mathbf{e}_b and \mathbf{e}_c . If this triplet is positively oriented, $\det L = V$; if it is negatively oriented, $\det L = -V$. Consider the case that $\det \chi = (\det L)^2 = 0$, which means that $V = 0$. This, in turn, means that \mathbf{e}_a , \mathbf{e}_b and \mathbf{e}_c are coplanar. It follows that the statement that $\det \chi = 0$ is equivalent to the statement that either the three angles $(\varphi_{ab}, \varphi_{ac}, \varphi_{bc})$ add up to 360° or one of them is the sum of the other two. Since the formal proof of this statement involves more trigonometry than we otherwise need, we placed it at the end of this section.

In the special case considered by Mermin, the three angles add up to 360° :

$$\varphi_{ab} = \varphi_{ac} = \varphi_{bc} = 120^\circ. \quad (2.6.14)$$

The corresponding values of the anti-correlation coefficients are (cf. Eq. (2.4.3)):

$$\chi_{ab} = \chi_{ac} = \chi_{bc} = \cos 120^\circ = -1/2. \quad (2.6.15)$$

This is the point at the center of the section of the surface of the ellipptope in Figure 2.16 behind the facet (ii)-(iii)-(iv) of the tetrahedron.

At this point, the sum of the anti-correlation coefficients has its minimum value. In quantum mechanics, this sum thus satisfies the inequality

$$\chi_{ab} + \chi_{ac} + \chi_{bc} \geq -3/2. \quad (2.6.16)$$

This is the quantum counterpart of the CHSH-type Bell inequality in Eq. (2.5.1), which says that $\chi_{ab} + \chi_{ac} + \chi_{bc} \geq -1$ in any local hidden-variable theory. The minimum value of $-3/2$ allowed by quantum mechanics is the so-called *Tsirelson bound* for this setup, named for Boris Cirel'son (1980). Both the Tsirelson bound and the CHSH-type Bell inequality for the Mermin-style setup are represented by dotted lines in Figure 2.8.²⁵

²⁵ The line marked "Tsirelson bound" in Figure 2.8 does not touch the circle representing the quantum convex set in the $\chi_{bc} = 0$ cross-section shown in the figure because the point

The point where the sum of the anti-correlation coefficients reaches the minimum value allowed by quantum mechanics in the Mermin-style setup is also the point where the probability of Alice and Bob finding opposite results reaches the minimum value allowed. Substituting the values for the anti-correlation coefficients in Eq. (2.6.15) into Eq. (2.5.8), we find

$$\Pr(\text{opp}) = 2/3 + 1/9 (\chi_{ab} + \chi_{ac} + \chi_{bc}) = 2/3 - 1/6 = 1/2. \quad (2.6.17)$$

As in the classical case (see Eqs. (2.5.2)–(2.5.5)), we can write down three pairs of inequalities like the one for which the lower bound is given in Eq. (2.6.16):

$$-3/2 \leq \chi_{ab} - \chi_{ac} - \chi_{bc} \leq 3, \quad (2.6.18)$$

$$-3/2 \leq -\chi_{ab} + \chi_{ac} - \chi_{bc} \leq 3, \quad (2.6.19)$$

$$-3/2 \leq -\chi_{ab} - \chi_{ac} + \chi_{bc} \leq 3. \quad (2.6.20)$$

However, while the classical counterparts of these four linear inequalities sufficed to fully characterize the classical tetrahedron, we need the non-linear inequality in Eq. (2.6.9) to characterize the elliptope with its curved surface.

To conclude this section, we give the proof alluded to above of the equivalence of the vanishing of the determinant of the anti-correlation matrix χ in Eq. (2.6.2) to the statement that the angles $(\varphi_{ab}, \varphi_{ac}, \varphi_{bc})$ between the different peeling directions either add up to 360° or that one of them is the sum of the other two.²⁶ Following Deza & Laurent (1997, p. 515), we do this with the help of the trigonometric identity

$$\begin{aligned} (2 \cos \alpha)(2 \cos \beta)(2 \cos \gamma) &= 2 \cos(\alpha + \beta + \gamma) \\ &\quad + 2 \cos(-\alpha + \beta + \gamma) \\ &\quad + 2 \cos(\alpha - \beta + \gamma) \\ &\quad + 2 \cos(\alpha + \beta - \gamma), \end{aligned} \quad (2.6.21)$$

which the reader can check for him- or herself with the help of the Euler formula, $2 \cos \alpha = e^{i\alpha} + e^{-i\alpha}$.

$(\chi_{ab}, \chi_{ac}, \chi_{bc}) = (-1/2, -1/2, -1/2)$ for which $\chi_{ab} + \chi_{ac} + \chi_{bc}$ reaches the minimum value quantum mechanics allows, lies below the $\chi_{bc} = 0$ plane (cf. Figure 2.16).

²⁶ The reader can skip this proof without loss of continuity though it is worth taking a look at Figure 2.17, which gives an alternative (tetrahedron) representation of the quantum elliptope.

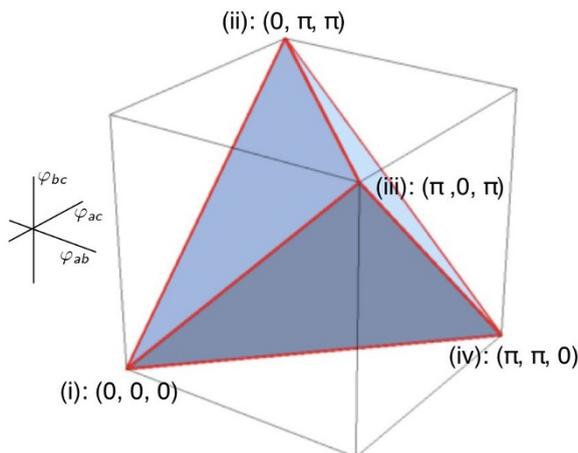


Fig. 2.17 Tetrahedron of triplets of angles $(\varphi_{ab}, \varphi_{ac}, \varphi_{bc})$ allowed by quantum mechanics in the Mermin-style setup.

In all cases in which one of the angles (α, β, γ) is either the sum of the other two ($\alpha = \beta + \gamma$, $\beta = \alpha + \gamma$, or $\gamma = \alpha + \beta$) or is 360° minus this sum ($\alpha + \beta + \gamma = 360^\circ$), the right-hand side of Eq. (2.6.21) reduces to

$$2 + 2 \cos 2\alpha + 2 \cos 2\beta + 2 \cos 2\gamma = 4 \cos^2 \alpha + 4 \cos^2 \beta + 4 \cos^2 \gamma - 4. \quad (2.6.22)$$

Eq. (2.6.21) can thus be rewritten as

$$2 \cos \alpha \cos \beta \cos \gamma = \cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma - 1. \quad (2.6.23)$$

Substituting $(\varphi_{ab}, \varphi_{ac}, \varphi_{bc})$ for (α, β, γ) and $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ for $(\cos \varphi_{ab}, \cos \varphi_{ac}, \cos \varphi_{bc})$, we see that Eq. (2.6.23) reduces to Eq. (2.6.9) with an '=' sign rather than a '≥' sign. This shows that the equality $\det \chi = 0$ is indeed equivalent to the statement that the angles $(\varphi_{ab}, \varphi_{ac}, \varphi_{bc})$ either add up to 360° or that one of them is the sum of the other two. More generally, the *inequality* $\det \chi \geq 0$ is equivalent to the statement that either their sum is no larger than 360° or the sum of any two of them is no smaller than the third, i.e.,

$$\varphi_{ab} + \varphi_{ac} + \varphi_{bc} \leq 360^\circ$$

or

$$(2.6.24)$$

$$\varphi_{ab} \leq \varphi_{ac} + \varphi_{bc} \quad \text{and} \quad \varphi_{ac} \leq \varphi_{ab} + \varphi_{bc} \quad \text{and} \quad \varphi_{bc} \leq \varphi_{ab} + \varphi_{ac}$$

(Deza & Laurent, 1997, p. 515; cf. the triangle on the right in Figure 3.2 in Chapter 3). These angle inequalities can be represented geometrically by the tetrahedron in Figure 2.17.²⁷

2.6.2 *The quantum correlation array: the singlet state and the Born rule*

In this section, we introduce the basic Hilbert-space formalism of quantum mechanics in Dirac notation and use it to derive expressions for the probabilities of Alice and Bob finding various combinations of ‘yummy’ and ‘nasty’ when peeling and tasting bananas in the Mermin-style setup.²⁸ These probabilities

²⁷ These inequalities can also be found in Accardi & Fedullo (1982, remark on p. 170). These authors also give Eq. (2.6.9) for the ellipsope in terms of *the cosines* of these angles, which are just our anti-correlation coefficients $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ (ibid., p. 166, Eq. 19; p. 168, Eq. 33; p. 169, Eq. 37):

$$2 \cos \varphi_{ab} \cos \varphi_{ac} \cos \varphi_{bc} \geq \cos^2 \varphi_{ab} + \cos^2 \varphi_{ac} + \cos^2 \varphi_{bc} - 1.$$

Citing Accardi & Fedullo (1982), Bas Van Fraassen (1991, pp. 120–121) states this same inequality in terms of probabilities rather than correlation coefficients, just as Mermin did with the Bell inequality (see Eqs. (2.5.6)–(2.5.8)).

²⁸ See Table 2.2 for a quick overview of the formalism. The non-expert reader looking for a more detailed introduction than we offer in this section has many options. Several books on the foundations of quantum mechanics written by philosophers of physics contain elementary expositions of this material. We particularly recommend Albert (1992, Ch. 2, pp. 17–60). Bub (2016, pp. 233–267) and Hughes (1989, Ch. 1, pp. 11–56) are also good though both cover far more ground than is needed for our purposes. The same is true for the coverage in books on quantum computation and information but these would nonetheless be the obvious choice for readers interested in these particular applications of quantum mechanics. We recommend Rieffel & Polak (2011, Chs. 2–4, pp. 9–70), specifically written for a non-expert audience, and Nielsen & Chuang (2016, Ch. 2, pp. 60–119), the introduction to quantum mechanics in the standard textbook on quantum computation and information for the past two decades. Other useful books in this category are Bernhardt (2019), pitched at a more elementary level than Rieffel and Polak, Mermin (2007), pitched at a level comparable to Nielsen and Chuang, and McMahon (2008). As Nielsen & Chuang (2016, pp. xxiv–xxv) observe in the preface of their book: “Conventional introductions to quantum mechanics rely heavily on the mathematical machinery of partial differential equations. We believe this often obscures the fundamental ideas. Quantum computation and quantum information offers [sic] an excellent conceptual laboratory for understanding the basic concepts and unique aspects of quantum mechanics, without the use of heavy mathematical machinery.” The mathematics needed instead is linear algebra. Nielsen and Chuang’s observation explains why most textbooks currently used to introduce physics students to quantum mechanics are not useful for our purposes. But this may be about to change. Several more recent introductory textbooks start

are the entries of the correlation arrays for the correlations found in these experiments. They involve the squares of sines and cosines of half the angles between peeling directions. With the help of two elementary trigonometry formulas they can be rewritten in terms of cosines of the full angle between peeling directions. In this way, we arrive at the one result from quantum mechanics we needed in Section 2.6.1: the anti-correlation coefficients χ_{ab} , χ_{ac} and χ_{bc} parametrizing these correlations are the cosines of the angles φ_{ab} , φ_{ac} and φ_{bc} between the unit vectors \mathbf{e}_a , \mathbf{e}_b and \mathbf{e}_c in the different peeling directions Alice and Bob can choose from.

We hate to break the spell but our *bananas*, *peelings* and *tastes* are ultimately window-dressing for *spin- $\frac{1}{2}$ particles*, *measurements* of components of their spin with Du Bois magnets in a Stern-Gerlach setup and *outcomes* of such measurements. Picking a pair of bananas from a quantum-banana tree corresponds to preparing a pair of spin- $\frac{1}{2}$ particles in the so-called *singlet state*, a term betraying the historical origins of quantum mechanics in spectroscopy and indicating that the sum of the components of the spin of these particles in any given direction is zero (just as the tastes of two bananas coming from the same pair and peeled in the same direction “cancel out”). The different peeling directions correspond to different orientations of the axes of the Du Bois magnets. ‘Yummy’ and ‘nasty’ correspond to ‘spin up’ and ‘spin down’.

To stay with the conceit of quantum bananas for now, the state of a pair of quantum bananas can be represented by a unit vector in the Hilbert space for this system, which we will denote as \mathcal{H}_{bb} . This is a four-dimensional vector space given by the *tensor product* $\mathcal{H}_b \otimes \mathcal{H}_b$ of two (identical) two-dimensional vector spaces, the one-banana Hilbert space \mathcal{H}_b . The dimension of a Hilbert space is equal to the number of basis vectors it takes to span it and the dimension of a tensor product of Hilbert spaces is the product of the dimensions of these Hilbert spaces. Before we deal with pairs of bananas, we introduce the Hilbert-space formalism in general and show how it works for one banana.

Table 2.2 shows the Hilbert space formalism at a glance. The state ψ of a physical system S is represented by a *state vector*, a unit vector in the Hilbert space \mathcal{H} for S . In Dirac notation this vector is written as $|\psi\rangle$. If S is a single quantum banana (mimicking the behavior of a single spin- $\frac{1}{2}$ particle), we can

with the same quantum systems we are considering here, i.e., spin- $\frac{1}{2}$ particles sent through Du Bois magnets in variations on the Stern-Gerlach experiment. We particularly recommend McIntyre (2012, Chs. 1–4, pp. 1–106) and the software it uses to simulate variations on the Stern-Gerlach experiment. See Section 6.5 for more detailed discussion of these experiments and note 41 in that section for a partial genealogy of textbooks taking this approach.

picture its state vectors in the one-banana Hilbert space \mathcal{H}_b as arrows in the xy -plane with their tails in the origin O and their tips on the unit circle $x^2 + y^2 = 1$ around O (cf. Figure 2.18). Following the usual “tip-to-tail” method, we can construct arbitrary linear combinations of vectors in \mathcal{H}_b .

Quantum system S	physical quantities	mathematical representation
States	ψ	Unit vector $ \psi\rangle$ in Hilbert space \mathcal{H}
Observables (possible values)	A (a_1, a_2, a_3, \dots)	Operator $\hat{A} : \mathcal{H} \rightarrow \mathcal{H}$ Eigenvectors/eigenvalues: $\hat{A} a_i\rangle = a_i a_i\rangle$
Measurement statistics (Born rule)	$\Pr(a_i \psi, A)$	Inner product squared: $ \langle a_i \psi \rangle ^2$

Table 2.2 The Hilbert-space formalism of quantum mechanics at a glance.

Observables A associated with the system S are represented by linear operators \hat{A} (the ‘hat’ distinguishes the operator from the associated observable) that map vectors $|\psi\rangle \in \mathcal{H}$ onto other vectors $\hat{A}|\psi\rangle \in \mathcal{H}$. That \hat{A} is *linear* means that, for any two vectors $|\psi\rangle, |\chi\rangle \in \mathcal{H}$ and arbitrary values of the coefficients α and β ,

$$\hat{A}(\alpha|\psi\rangle + \beta|\chi\rangle) = \alpha\hat{A}|\psi\rangle + \beta\hat{A}|\chi\rangle. \quad (2.6.25)$$

In general α and β will be complex numbers but, for our purposes in this chapter, we can take them to be real numbers (cf. note 36).

An *eigenvector* $|a\rangle$ of \hat{A} is a vector that \hat{A} maps onto a vector that is either in the same or in the exact opposite direction as $|a\rangle$. The associated *eigenvalue* a is the coefficient by which $|a\rangle$ needs to be multiplied to obtain $\hat{A}|a\rangle$. As we shall see shortly, for the class of operators representing observables in quantum mechanics, these eigenvalues are always (positive or negative) real numbers. The values that the observable A represented by the operator \hat{A} can take on are just the eigenvalues of \hat{A} . The set of eigenvalues of \hat{A} thus doubles as the set

of possible values of A and is called the *spectrum* of both \hat{A} and A . It can be discrete, continuous or a combination of both. We will focus on operators \hat{A} with a finite number of real, discrete and different eigenvalues. Let $\{a_i\}$ (with $i = 1, \dots, N$ and $a_i \neq a_j$ whenever $i \neq j$) be the set of eigenvalues of \hat{A} . And let $\{|a_i\rangle\}$ be the corresponding set of unit eigenvectors. So we have

$$\hat{A} |a_i\rangle = a_i |a_i\rangle. \quad (2.6.26)$$

Eigenvectors with different eigenvalues are orthogonal. To prove this, we consider the inner product of the vectors $|a_i\rangle$ and $\hat{A} |a_j\rangle$ where $i \neq j$. The Dirac notation for an inner product of two vectors $|\psi\rangle, |\chi\rangle \in \mathcal{H}$ is the ‘bracket’ $\langle\psi|\chi\rangle$. Vectors such as $|\psi\rangle$ and $|\chi\rangle$, entering inner products from the right, are therefore also called ‘kets’, while their counterparts, *dual vectors* such as $\langle\psi|$ and $\langle\chi|$, entering inner products from the left, are called ‘bras’.²⁹ The bra corresponding to the ket $\hat{A}|\psi\rangle$ is $\langle\psi|\hat{A}^\dagger$, where \hat{A}^\dagger is the so-called *adjoint* or *Hermitian conjugate* of \hat{A} . An operator that is its own adjoint ($\hat{A}^\dagger = \hat{A}$) is called *self-adjoint* or *Hermitian*. A key property of such operators is that all their eigenvalues are real numbers. This is why observables are always represented by *self-adjoint* linear operators in quantum mechanics. If \hat{A} is self-adjoint, the bracket $\langle a_i|\hat{A}|a_j\rangle$ can be parsed either as an inner product with the bra $\langle a_i|$ and the ket $\hat{A}|a_j\rangle = a_j|a_j\rangle$ or as one with the bra $\langle a_i|\hat{A} = \langle a_i|a_i$ and the ket $|a_j\rangle$. Hence

$$0 = \langle a_i|(\hat{A}|a_j\rangle) - (\langle a_i|\hat{A})|a_j\rangle = (a_j - a_i)\langle a_i|a_j\rangle. \quad (2.6.27)$$

If $i \neq j$, $a_j - a_i \neq 0$ and it follows that $\langle a_i|a_j\rangle = 0$. In other words, if $i \neq j$, $|a_i\rangle$ and $|a_j\rangle$ are orthogonal.

The set $\{|a_i\rangle\}$ of unit eigenvectors of \hat{A} form an *orthonormal basis* of the Hilbert space \mathcal{H} for the system S (i.e., a basis consisting of mutually *orthogonal* vectors with their lengths *normalized* to 1). In analogy with vectors in ordinary space, any vector $|\psi\rangle \in \mathcal{H}$ can be written as a linear combination of such basis vectors. As with vectors in ordinary space, the length of the component of $|\psi\rangle$ in the direction of $|a_i\rangle$ is given by the inner product $\langle a_i|\psi\rangle$. This then is the coefficient of $|a_i\rangle$ in the expansion of $|\psi\rangle$ in the $\{|a_i\rangle\}$ basis.³⁰

²⁹ We can think of kets as column vectors and of bras as row vectors. The inner product of two vectors can be written as the matrix product of the row version of the first and the column version of the second.

³⁰ Rewriting the right-hand side of Eq. (2.6.28) as $\sum_{i=1}^N |a_i\rangle\langle a_i|\psi\rangle$ (i.e., changing the order of $\langle a_i|\psi\rangle$ and $|a_i\rangle$ in every term of the sum), we see that the operator $\sum_{i=1}^N |a_i\rangle\langle a_i|$ maps $|\psi\rangle$

$$|\psi\rangle = \sum_{i=1}^N \langle a_i | \psi \rangle |a_i\rangle. \quad (2.6.28)$$

For $|\psi\rangle$ to be a state vector, it needs to be a unit vector, i.e., it needs to be normalized to 1. The normalization condition is that the sum of the absolute squares of the coefficients $\langle a_i | \psi \rangle$ (which will, in general, be complex numbers) must be equal to 1:^{31,32}

$$\sum_{i=1}^N |\langle a_i | \psi \rangle|^2 = 1. \quad (2.6.29)$$

For our purposes in this chapter, we can take the coefficients $\langle a_i | \psi \rangle$ to be real numbers (cf. note 36), in which case we can leave out the bars.

We now know how states and observables of a system S are represented in the Hilbert-space formalism in quantum mechanics but there is still one ingredient missing. Let \mathcal{H} be the Hilbert space for S . We saw that quantum mechanics says that if we measure an observable A of S with a discrete spectrum represented by an operator $\hat{A} : \mathcal{H} \rightarrow \mathcal{H}$ when S is in the state ψ represented by the unit vector $|\psi\rangle \in \mathcal{H}$, we will find some eigenvalue a_i of \hat{A} . What is still missing is a general rule for the probability of finding a particular eigenvalue. This is the *Born rule* (cf. Chapter 1), which says that the probability of finding a_i when measuring A if S is in the state ψ is given by the absolute square of

back onto itself, so it must be the unit operator $\hat{1}$. The equation $\sum_{i=1}^N |a_i\rangle\langle a_i| = \hat{1}$ is called the *completeness relation* or the *resolution of unity* (after the German *Zerlegung der Einheit*) for the basis $\{|a_i\rangle\}$. We will repeatedly make use of this relation, both in this section (see notes 32 and 35) and in Chapters 4 and 6.

³¹ The absolute square $|z|^2$ of any complex number z is the square of its *modulus*. There are two standard ways for writing any complex number z in terms of two real numbers and the complex unit $i = \sqrt{-1}$, either as $\text{Re}(z) + i \text{Im}(z)$, where $\text{Re}(z)$ is the *real part* and $\text{Im}(z)$ is the *imaginary part* of z , or as $re^{i\varphi}$, where r is the *modulus* and φ is the *argument* of z . These two expressions are related via $\text{Re}(z) = r \cos \varphi$ and $\text{Im}(z) = r \sin \varphi$. With the help of the complex conjugate of z , defined as $\bar{z} \equiv \text{Re}(z) - i \text{Im}(z)$ or, equivalently, as $\bar{z} \equiv re^{-i\varphi}$, its absolute square can be written as $|z|^2 = z\bar{z}$.

³² To show that Eq. (2.6.29) ensures that $|\psi\rangle$ is a unit vector, we need the completeness relation (see note 30), a basic property of complex numbers (i.e., $|z|^2 = z\bar{z}$, see note 31) and a property of the inner product of vectors in \mathcal{H} (namely that $\langle \chi | \psi \rangle$ is the complex conjugate of $\langle \psi | \chi \rangle$):

$$\sum_{i=1}^N |\langle a_i | \psi \rangle|^2 = \sum_{i=1}^N \langle a_i | \psi \rangle \overline{\langle a_i | \psi \rangle} = \sum_{i=1}^N \langle a_i | \psi \rangle \langle \psi | a_i \rangle = \sum_{i=1}^N \langle \psi | a_i \rangle \langle a_i | \psi \rangle = \langle \psi | \psi \rangle.$$

Hence, Eq. (2.6.29) entails that $\langle \psi | \psi \rangle = 1$, i.e., that $|\psi\rangle$ is a unit vector.

the inner product of the state vector $|\psi\rangle$ and the eigenvector $|a_i\rangle$ of \hat{A} for that eigenvalue. In a formula:

$$\Pr(a_i|\psi, A) = |\langle a_i | \psi \rangle|^2. \quad (2.6.30)$$

As noted above, if the inner products $\langle a_i | \psi \rangle$ are real numbers, we can leave out the bars. Eq. (2.6.29) guarantees that

$$\sum_{i=1}^N \Pr(a_i|\psi, A) = 1. \quad (2.6.31)$$

If the system S is prepared in the same state ψ as before but we measure the observable B represented by an operator \hat{B} , the probability of finding one of its eigenvalues b_i is likewise given by $\Pr(b_i|\psi, B) = |\langle b_i | \psi \rangle|^2$ with $\sum_{i=1}^N \Pr(b_i|\psi, B) = 1$ on account of the normalization condition $\sum_{i=1}^N |\langle b_i | \psi \rangle|^2 = 1$.

The state vector $|\psi\rangle$ thus encodes a family of probability distributions, all subject to the Born rule. As in your average ordinary family, not all family members are compatible with one another. More specifically, if the operators \hat{A} and \hat{B} do not commute (i.e., if $\hat{A}\hat{B} \neq \hat{B}\hat{A}$),³³ they do not share a basis of eigenvectors, which makes it impossible to simultaneously assign probabilities to eigenvalues of \hat{A} and \hat{B} in accordance with the Born rule.

We will now illustrate this formalism for the simple case that the system S is a single quantum banana. The state ψ of such a banana is represented by a unit vector $|\psi\rangle$ in the two-dimensional one-banana Hilbert space \mathcal{H}_b . Its taste T_x when peeled in the x -direction is represented by the operator $\hat{T}_x: \mathcal{H}_b \rightarrow \mathcal{H}_b$ with only two eigenvalues, $+1/2$ and $-1/2$, both in units of \hbar , where \hbar (pronounced *b-bar*) is the *banana split*, defined as *Bub's constant* b divided by 2π . ‘Yummy’ is $+\hbar/2$, ‘nasty’ is $-\hbar/2$.³⁴ We will abbreviate these

³³ This means that letting \hat{A} and then \hat{B} act on some vector $|\psi\rangle \in \mathcal{H}$ will in general produce a different vector than letting \hat{B} and then \hat{A} act on $|\psi\rangle$. If operators are represented by matrices, this translates into the familiar property that a matrix product will in general depend on the order of the matrices multiplied (see note 35 for an example).

³⁴ We trust that it will be clear from context whether ‘ b ’ stands for Bub’s constant (as in \hbar), for a peeling direction (as in \mathbf{e}_b , φ_{ab} or χ_{ab}) or for ‘banana’ (as in \mathcal{H}_b). When ‘taste of a banana when peeled in a certain direction’ is replaced by ‘spin component of a spin- $\frac{1}{2}$ particle when sent through a Du Bois magnet pointing in a certain direction’, \hbar gets replaced by \hbar (pronounced *h-bar*), defined as Planck’s constant h divided by 2π (cf. Eq. (4.1.2)) and ‘yummy’ and ‘nasty’ get replaced by ‘spin up’ and ‘spin down’.

eigenvalues as ‘+’ and ‘-’, respectively. Let the unit vectors $|+\rangle_x$ and $|-\rangle_x$ be eigenvectors of \hat{T}_x for those eigenvalues:

$$\hat{T}_x |+\rangle_x = \frac{\hbar}{2} |+\rangle_x, \quad \hat{T}_x |-\rangle_x = -\frac{\hbar}{2} |-\rangle_x. \quad (2.6.32)$$

These two unit vectors form an orthonormal basis $\{|+\rangle_x, |-\rangle_x\}$ for \mathcal{H}_b . Figure 2.18 shows two such bases one for $x = a$ and one for $x = b$, associated with the observables ‘taste when peeled in the a -direction’ and ‘taste when peeled in the b -direction’. We will refer to these bases of \mathcal{H}_b as the a -basis and the b -basis. With malice aforethought, we have chosen the angle between the unit eigenvectors $|+\rangle_a$ and $|+\rangle_b$ in the one-banana Hilbert space \mathcal{H}_b to be half the angle φ_{ab} between the peeling directions given by the unit vectors \mathbf{e}_a and \mathbf{e}_b in ordinary three-dimensional space. The inner product ${}_a\langle + | + \rangle_b$ of these two vectors is the cosine of this half-angle. Upon inspection of Figure 2.18 and using that $\cos(90^\circ - \alpha) = \sin \alpha$, we similarly find that:

$$\begin{aligned} {}_a\langle + | + \rangle_b &= {}_b\langle + | + \rangle_a = \cos(\varphi_{ab}/2), \\ {}_a\langle + | - \rangle_b &= {}_b\langle - | + \rangle_a = -\sin(\varphi_{ab}/2), \\ {}_a\langle - | + \rangle_b &= {}_b\langle + | - \rangle_a = \sin(\varphi_{ab}/2), \\ {}_a\langle - | - \rangle_b &= {}_b\langle - | - \rangle_a = \cos(\varphi_{ab}/2). \end{aligned} \quad (2.6.33)$$

Any linear combination of these basis vectors with coefficients whose (absolute) squares add up to 1 represents a possible state of a quantum banana. Examples of such vectors with real coefficients are

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle_a + |-\rangle_a), \quad |\chi\rangle = \frac{1}{2}|+\rangle_b - \frac{1}{2}\sqrt{3}|-\rangle_b. \quad (2.6.34)$$

Suppose we peel a banana in the state $|\psi\rangle$ in the a -direction. We then expand $|\psi\rangle$ in the a -basis and appeal to the Born rule in Eq. (2.6.30) to set the probability of finding $T_a = \hbar/2$ (‘yummy’) equal to the square of the coefficient of $|+\rangle_a$ in this expansion and the probability of finding $T_a = -\hbar/2$ (‘nasty’) to the square of the coefficient of $|-\rangle_a$. In this case both probabilities are $1/2$:

$$\Pr(+|\psi, T_a) = {}_a\langle + | \psi \rangle^2 = \frac{1}{2}, \quad \Pr(-|\psi, T_a) = {}_a\langle - | \psi \rangle^2 = \frac{1}{2}. \quad (2.6.35)$$

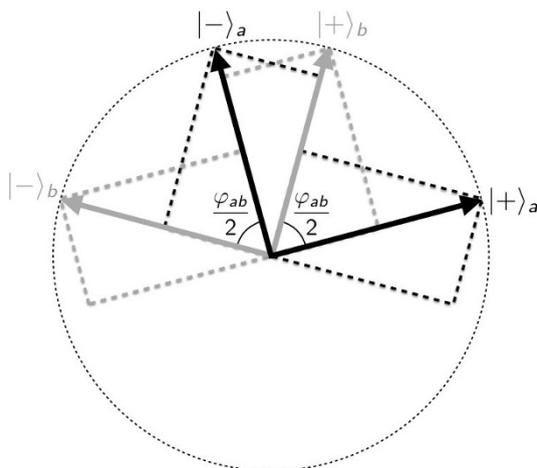


Fig. 2.18 Eigenvectors for the operators \hat{T}_a and \hat{T}_b acting in the one-banana Hilbert space \mathcal{H}_b representing the tastes of the banana when peeled in the a - and the b -direction, respectively.

Similarly, if we peel a banana in the state $|\chi\rangle$ in the b -direction, we expand $|\chi\rangle$ in the b -basis and appeal to the Born rule to set the probabilities of finding ‘yummy’ and ‘nasty’ equal to $1/4$ and $3/4$, respectively. However, if we decide to peel a banana in the state $|\psi\rangle$ in the b -direction or a banana in the state $|\chi\rangle$ in the a -direction, we need to expand the former in the b -basis and the latter in the a -basis. No matter what state it is in, the Born rule cannot give us probabilities for finding ‘yummy’ and ‘nasty’ for one and the same banana peeled in the a - and in the b -direction. Here the banana imagery is especially helpful. A banana can only be peeled once. As Sandu Popescu (2016, p. vi) put it in his foreword to *Bananaworld*: “once it’s peeled it’s peeled . . . once it’s eaten it’s eaten.” It is this impossibility of assigning a definite taste to one and the same banana for different peeling directions that lies behind quantum mechanics allowing a broader class of correlations in the Mermin-style setup than local hidden-variable theories.³⁵

³⁵ Representing the operators \hat{T}_a and \hat{T}_b as matrices, we can readily verify that they do not commute (cf. note 33). One obtains the matrix representing an operator in a given basis by “sandwiching” that operator between the bras and kets of that basis. Consider the operator \hat{T}_a in the a -basis. With the help of the completeness relation (see note 30), we can write the components of the vector $\hat{T}_a|\psi\rangle$ in the a -basis as ${}_a\langle i|\hat{T}_a|\psi\rangle = \sum_j {}_a\langle i|\hat{T}_a|j\rangle {}_a\langle j|\psi\rangle$, where i and j take on the values \pm . It follows that, in the a -basis, \hat{T}_a is represented by a diagonal matrix with its eigenvalues on the diagonal:

We are now finally ready to deal with the pairs of bananas peeled and tasted in the Mermin-style setup. We can use any orthonormal basis for the one-banana Hilbert space \mathcal{H}_b to construct an orthonormal basis for the two-banana Hilbert space \mathcal{H}_{bb} . For instance, we can use the two unit vectors of the a -basis

$$\left(\hat{T}_a\right)^{(a)} \equiv \begin{pmatrix} {}_a\langle +|\hat{T}_a|+\rangle_a & {}_a\langle +|\hat{T}_a|-\rangle_a \\ {}_a\langle -|\hat{T}_a|+\rangle_a & {}_a\langle -|\hat{T}_a|-\rangle_a \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The operator \hat{T}_b is represented by the same matrix in the b -basis. To find its matrix in the a -basis, we sandwich \hat{T}_b between the bras and kets of the a -basis and use the completeness relation in the b -basis: ${}_a\langle i|\hat{T}_b|j\rangle_a = \sum_{(k,l)} {}_a\langle i|k\rangle_b {}_b\langle k|\hat{T}_b|l\rangle_b {}_b\langle l|j\rangle_a$. These are the elements of a matrix that is the product of three matrices, the one for \hat{T}_b in the b -basis sandwiched between those governing transformations between the a - and the b -basis. The elements of these transformation matrices are inner products of the basis vectors in the two bases. Using Eq. (2.6.33), we can write them as:

$$\begin{pmatrix} {}_a\langle +|+\rangle_b & {}_a\langle +|-\rangle_b \\ {}_a\langle -|+\rangle_b & {}_a\langle -|-\rangle_b \end{pmatrix} = \begin{pmatrix} \cos(\varphi_{ab}/2) & -\sin(\varphi_{ab}/2) \\ \sin(\varphi_{ab}/2) & \cos(\varphi_{ab}/2) \end{pmatrix},$$

$$\begin{pmatrix} {}_b\langle +|+\rangle_a & {}_b\langle +|-\rangle_a \\ {}_b\langle -|+\rangle_a & {}_b\langle -|-\rangle_a \end{pmatrix} = \begin{pmatrix} \cos(\varphi_{ab}/2) & \sin(\varphi_{ab}/2) \\ -\sin(\varphi_{ab}/2) & \cos(\varphi_{ab}/2) \end{pmatrix}.$$

These two matrices are each other's inverse: multiplying one by the other gives the unit matrix. Sandwiching the matrix for \hat{T}_b in the b -basis between these two transformation matrices, we find the matrix for \hat{T}_b in the a -basis:

$$\begin{aligned} \left(\hat{T}_b\right)^{(a)} &= \begin{pmatrix} \cos(\varphi_{ab}/2) & -\sin(\varphi_{ab}/2) \\ \sin(\varphi_{ab}/2) & \cos(\varphi_{ab}/2) \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos(\varphi_{ab}/2) & \sin(\varphi_{ab}/2) \\ -\sin(\varphi_{ab}/2) & \cos(\varphi_{ab}/2) \end{pmatrix} \\ &= \begin{pmatrix} \cos \varphi_{ab} & \sin \varphi_{ab} \\ \sin \varphi_{ab} & -\cos \varphi_{ab} \end{pmatrix}, \end{aligned}$$

where in the last step we used the standard trigonometry formulas $\cos^2 \alpha - \sin^2 \alpha = \cos 2\alpha$ and $2 \sin \alpha \cos \alpha = \sin 2\alpha$. Now that we have the matrices for both \hat{T}_a and \hat{T}_b in the a -basis, we can calculate their commutator $[\hat{T}_a, \hat{T}_b]$, defined as $\hat{T}_a \hat{T}_b - \hat{T}_b \hat{T}_a$, in the a -basis:

$$\left([\hat{T}_a, \hat{T}_b]\right)^{(a)} = \left(\hat{T}_a\right)^{(a)} \left(\hat{T}_b\right)^{(a)} - \left(\hat{T}_b\right)^{(a)} \left(\hat{T}_a\right)^{(a)} = \begin{pmatrix} 0 & 2 \sin \varphi_{ab} \\ -2 \sin \varphi_{ab} & 0 \end{pmatrix}.$$

That not all elements of this matrix vanish tells us that \hat{T}_a and \hat{T}_b do not commute (unless $a = b$ in which case $\sin \varphi_{ab} = 0$).

for \mathcal{H}_b , to construct the four unit vectors of an aa -basis for \mathcal{H}_{bb} :

$$\left\{ |+\rangle_{1a}|+\rangle_{2a}, |+\rangle_{1a}|-\rangle_{2a}, |-\rangle_{1a}|+\rangle_{2a}, |-\rangle_{1a}|-\rangle_{2a} \right\} \quad (2.6.36)$$

(where the subscripts 1 and 2 refer to the bananas given to Alice and Bob, respectively). The inner product of vectors in \mathcal{H}_{bb} is defined in terms of the inner product of vectors in \mathcal{H}_b . The inner product between the first and the second vector in Eq. (2.6.36), for instance, is:

$$\left({}_{1a}\langle + | {}_{2a}\langle + | \right) \left(|+\rangle_{1a}|-\rangle_{2a} \right) \equiv {}_{1a}\langle + | + \rangle_{1a} {}_{2a}\langle + | - \rangle_{2a} = 0. \quad (2.6.37)$$

With this definition, the four vectors in Eq. (2.6.36) constitute an orthonormal basis of \mathcal{H}_{bb} .

The pairs of bananas in our peel-and-taste experiments are always prepared in the singlet state. The expansion of this state in the aa -basis in Eq. (2.6.36) is given by:

$$|0,0\rangle_{12} = \frac{1}{\sqrt{2}} \left(|+\rangle_{1a}|-\rangle_{2a} - |-\rangle_{1a}|+\rangle_{2a} \right). \quad (2.6.38)$$

In both terms, the taste of one banana ‘‘cancels’’ the taste of the other, hence the notation $|0,0\rangle_{12}$. It is clear upon inspection of Eq. (2.6.38) that $|0,0\rangle_{12}$ is an *entangled state* (cf. note 2 in Section 2.1): the state of the composite system (the pair of bananas) cannot be written as a product of states of its two components (the two bananas considered separately). The Born rule tells us (cf. Eqs. (2.6.34)–(2.6.35)) that, if Alice and Bob both peel their bananas in the a -direction, they will always find opposite results and they will each find ‘yummy’ and ‘nasty’ half the time.

As we will see in Section 4.1, the singlet state, not just for a pair of spin- $\frac{1}{2}$ particles but for a pair of particles of arbitrary spin, is invariant under rotation. This means that $|0,0\rangle_{12}$ has the same form regardless of whether we use the aa -basis or the bb -basis for the two-banana Hilbert space. Here we provide an intuitive proof of this property in the spin- $\frac{1}{2}$ case. In Section 4.1, we will prove this more rigorously for arbitrary spin.³⁶ From Figure 2.18, we read off that (cf. Eq. (2.6.33)):

³⁶ What makes the proof in this section intuitive and dubious at the same time is that we take the coefficients in Eqs. (2.6.39)–(2.6.40) to be real numbers. In Section 4.1 (see Eqs. (4.1.14)–(4.1.24)), we will show that we can always write the eigenvectors of one spin operator as linear combinations *with real coefficients* of the eigenvectors of another spin operator.

$$|+\rangle_a = \cos\left(\frac{\varphi_{ab}}{2}\right)|+\rangle_b - \sin\left(\frac{\varphi_{ab}}{2}\right)|-\rangle_b, \quad (2.6.39)$$

and that

$$|-\rangle_a = \sin\left(\frac{\varphi_{ab}}{2}\right)|+\rangle_b + \cos\left(\frac{\varphi_{ab}}{2}\right)|-\rangle_b \quad (2.6.40)$$

Inserting these expressions into Eq. (2.6.38), we find that

$$\begin{aligned} |0,0\rangle_{12} = \frac{1}{\sqrt{2}} \{ & \left(\cos\left(\frac{\varphi_{ab}}{2}\right)|+\rangle_{1b} - \sin\left(\frac{\varphi_{ab}}{2}\right)|-\rangle_{1b} \right) \\ & \times \left(\sin\left(\frac{\varphi_{ab}}{2}\right)|+\rangle_{2b} + \cos\left(\frac{\varphi_{ab}}{2}\right)|-\rangle_{2b} \right) \\ & - \left(\sin\left(\frac{\varphi_{ab}}{2}\right)|+\rangle_{1b} + \cos\left(\frac{\varphi_{ab}}{2}\right)|-\rangle_{1b} \right) \\ & \times \left(\cos\left(\frac{\varphi_{ab}}{2}\right)|+\rangle_{2b} - \sin\left(\frac{\varphi_{ab}}{2}\right)|-\rangle_{2b} \right) \}. \end{aligned}$$

Terms with $\sin\left(\frac{\varphi_{ab}}{2}\right)\cos\left(\frac{\varphi_{ab}}{2}\right)$ cancel; terms with $\cos^2\left(\frac{\varphi_{ab}}{2}\right)$ and $\sin^2\left(\frac{\varphi_{ab}}{2}\right)$ add up to:

$$|0,0\rangle_{12} = \frac{1}{\sqrt{2}} \left(|+\rangle_{1b}|-\rangle_{2b} - |-\rangle_{1b}|+\rangle_{2b} \right). \quad (2.6.41)$$

The state $|0,0\rangle_{12}$ thus has the exact same form in the bb -basis as in the aa -basis (see Eq. (2.6.38)). The Born rule thus tells us that as long as Alice and Bob peel their bananas in the same direction—be it the a -, the b - or any other direction—they will always find opposite results and they will each find ‘yummy’ and ‘nasty’ half the time, in accordance with what we found in our peel-and-taste experiments in the Mermin-style setup.

To find the probabilities for the four possible combinations of tastes found when Alice peels her banana in the a -direction and Bob peels his banana in the b -direction, we need to expand the singlet state $|0,0\rangle_{12}$ in a basis for the two-banana Hilbert space in which we use the a -basis of the one-banana Hilbert space for Alice’s banana and the b -basis of the one-banana Hilbert space for Bob’s banana. The unit vectors in this ab -basis are:

$$\left\{ |+\rangle_{1a}|+\rangle_{2b}, |+\rangle_{1a}|-\rangle_{2b}, |-\rangle_{1a}|+\rangle_{2b}, |-\rangle_{1a}|-\rangle_{2b} \right\}. \quad (2.6.42)$$

These four vectors correspond to the four combinations of tastes Alice and Bob can find in this particular peel-and-taste experiment. Using Eqs. (2.6.39)–(2.6.40) to express $|+\rangle_{2a}$ and $|-\rangle_{2a}$ in Eq. (2.6.38) in terms of $|+\rangle_{2b}$ and $|-\rangle_{2b}$,

we find the expansion of the singlet state in the ab -basis:

$$|0,0\rangle_{12} = \frac{1}{\sqrt{2}} \left(\sin\left(\frac{\varphi_{ab}}{2}\right) |+\rangle_{1a}|+\rangle_{2b} + \cos\left(\frac{\varphi_{ab}}{2}\right) |+\rangle_{1a}|-\rangle_{2b} - \cos\left(\frac{\varphi_{ab}}{2}\right) |-\rangle_{1a}|+\rangle_{2b} + \sin\left(\frac{\varphi_{ab}}{2}\right) |-\rangle_{1a}|-\rangle_{2b} \right). \quad (2.6.43)$$

The Born rule now tells us that the probabilities of Alice and Bob finding the combination of tastes ‘++’, ‘+-’, ‘-+’ and ‘--’, respectively, are the squares of the coefficients of the corresponding terms in the expansion of $|0,0\rangle_{12}$:

$$\begin{aligned} \Pr(++) &= \Pr(--)=\frac{1}{2} \sin^2\left(\frac{\varphi_{ab}}{2}\right), \\ \Pr(+-) &= \Pr(-+)=\frac{1}{2} \cos^2\left(\frac{\varphi_{ab}}{2}\right). \end{aligned} \quad (2.6.44)$$

To find the probabilities for combinations of peeling directions other than $(\mathbf{e}_a, \mathbf{e}_b)$, we can simply relabel a and b in Eq. (2.6.43). Note that quantum mechanics correctly predicts that it makes no difference whose banana is peeled first (cf. note 13). Since $\varphi_{ab} = \varphi_{ba}$, $\varphi_{ac} = \varphi_{ca}$ and $\varphi_{bc} = \varphi_{cb}$, it also correctly predicts the probabilities of the various combinations of tastes if we have Alice and Bob swap peeling directions.

Thinking in terms of spin- $\frac{1}{2}$ particles rather than bananas for a moment, we can now also present an intuitive argument as to why the angle between, say, $|+\rangle_a$ and $|+\rangle_b$ in Hilbert space is only half the angle between the directions of \mathbf{e}_a and \mathbf{e}_b in ordinary space. Imagine that we place two Du Bois magnets in a beam of spin- $\frac{1}{2}$ particles, one right after the other, with the second one rotated 180° with respect to the first. In this setup we would only find two possible outcomes $(+1, -2)$ and $(-1, +2)$ (where \pm refers to spin up/down and 1 and 2 refer to the two Du Bois magnets). The probability of finding $(+1, +2)$ and $(-1, -2)$, in other words, vanishes. For the Born rule to reproduce this result, the angle between the eigenvectors $|+\rangle_a$ and $|+\rangle_b$ should be 90° if the angle between the vectors \mathbf{e}_a and \mathbf{e}_b specifying two orientations of the Du Bois magnet is 180° . In Section 4.1, we will give a more general derivation of the relation between these two angles in the spin- $\frac{1}{2}$ case (see Eqs. (4.1.16)–(4.1.24)).³⁷

³⁷ Another way to understand the relation between angles and half-angles in the special case of spin- $\frac{1}{2}$ particles is to consider a stereographic projection that gets us from a point P representing the state of the particle on a unit sphere known as the *Bloch sphere* to the

This argument, incidentally, reveals the limitations of the banana metaphor in two ways. First, as we noted above, quoting Popescu (2016, p. vi), we can only peel and taste a banana once. Secondly, there is an intrinsic difference between the tastes ‘yummy’ and ‘nasty’ of a banana (or, for that matter, between ‘heads’ and ‘tails’ of a quoin) whereas there is no intrinsic difference between ‘spin up’ and ‘spin down’. ‘Spin up’ and ‘spin down’ are defined with respect to some preferred axis, given, for instance, by the orientation of a Du Bois magnet.

This second complication also provides a simple argument for why we should not expect to succeed, using only elementary quantum systems as our components, in building a device such as the Superquantum Entangler PR01 of Bub & Bub (2018) realizing the PR box with the correlation array in Figure 2.3. Recall that the diagonal cells of this correlation array are different. Now one can easily imagine that tossing two coins initially facing heads and tossing two coins initially facing tails would give different results. But this cannot be true for the spin-components of spin- $\frac{1}{2}$ particles in the singlet state. Suppose we measure the spin-in-the- z -direction of both particles with two Du Bois magnets. Spherical symmetry requires that the statistics for this experiment do not change if we rotate *both* Du Bois magnets by the same angle to measure, say, spin-in-the- x -direction. Measurements on the singlet state of two spin- $\frac{1}{2}$ particles can thus not be used to produce a PR box with the correlation array in Figure 2.3. As we saw in Section 2.6.1, they also cannot be used to produce a PR box for the Mermin-style setup with a correlation array represented by the point $\chi_{ab} = \chi_{ac} = \chi_{bc} = -1$ in Figure 2.8 since this combination of values for the triplet $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ violates the elliptope inequality in Eq. (2.6.9).

Returning now to Eqs. (2.6.43)–(2.6.44), relabeling a and b as needed, we fill out the correlation array for the Mermin-style setup. The result is shown in Figure 2.19. Note that the rows and columns of all cells add up to $\frac{1}{2}$. The correlation array thus has uniform marginals. As we saw in Section 2.3, this is a sufficient condition for it to be non-signaling.

In Mermin’s concrete example (see Figure 2.1), the peeling directions \mathbf{e}_a , \mathbf{e}_b and \mathbf{e}_c corresponding to the settings \hat{a} , \hat{b} and \hat{c} are such that $\varphi_{ab} = \varphi_{ac} = \varphi_{bc} = 120^\circ$. Inserting

components of the corresponding state vector in some orthonormal basis (see, e.g., Rieffel & Polak, 2011, pp. 23–25). Let N , S and C denote the north pole, the south pole and the center of the Bloch sphere, respectively. Elementary geometry tells us that the angle NSP that comes into play in the stereographic projection of P from S onto the tangent plane at N is half the polar angle NCP , one of the spherical coordinates of P on the Bloch sphere.

		Bob						
		\hat{a}	\hat{b}					
Alice	\hat{a}	+	-	+	-	+	-	
		+	0	$\frac{1}{2}$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{ab}}{2}\right)$	$\frac{1}{2} \cos^2\left(\frac{\varphi_{ab}}{2}\right)$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{ac}}{2}\right)$	$\frac{1}{2} \cos^2\left(\frac{\varphi_{ac}}{2}\right)$
	-	$\frac{1}{2}$	0	$\frac{1}{2} \cos^2\left(\frac{\varphi_{ab}}{2}\right)$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{ab}}{2}\right)$	$\frac{1}{2} \cos^2\left(\frac{\varphi_{ac}}{2}\right)$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{ac}}{2}\right)$	
	\hat{b}	+	$\frac{1}{2} \sin^2\left(\frac{\varphi_{ab}}{2}\right)$	$\frac{1}{2} \cos^2\left(\frac{\varphi_{ab}}{2}\right)$	0	$\frac{1}{2}$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{bc}}{2}\right)$	$\frac{1}{2} \cos^2\left(\frac{\varphi_{bc}}{2}\right)$
	-	$\frac{1}{2} \cos^2\left(\frac{\varphi_{ab}}{2}\right)$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{ab}}{2}\right)$	$\frac{1}{2}$	0	$\frac{1}{2} \cos^2\left(\frac{\varphi_{bc}}{2}\right)$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{bc}}{2}\right)$	
	\hat{c}	+	$\frac{1}{2} \sin^2\left(\frac{\varphi_{ac}}{2}\right)$	$\frac{1}{2} \cos^2\left(\frac{\varphi_{ac}}{2}\right)$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{bc}}{2}\right)$	$\frac{1}{2} \cos^2\left(\frac{\varphi_{bc}}{2}\right)$	0	$\frac{1}{2}$
	-	$\frac{1}{2} \cos^2\left(\frac{\varphi_{ac}}{2}\right)$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{ac}}{2}\right)$	$\frac{1}{2} \cos^2\left(\frac{\varphi_{bc}}{2}\right)$	$\frac{1}{2} \sin^2\left(\frac{\varphi_{bc}}{2}\right)$	$\frac{1}{2}$	0	

Fig. 2.19 Correlation array given by quantum mechanics for banana peel-and-taste experiments in the Mermin-style setup.

$$\begin{aligned} \frac{1}{2} \sin^2\left(\frac{\varphi_{ab}}{2}\right) &= \frac{1}{2} \sin^2 60^\circ = \frac{3}{8}, \\ \frac{1}{2} \cos^2\left(\frac{\varphi_{ab}}{2}\right) &= \frac{1}{2} \cos^2 60^\circ = \frac{1}{8}, \end{aligned} \quad (2.6.45)$$

etc., in the correlation array in Figure 2.19, we recover the Mermin correlation array in Figure 2.6.

Using the trigonometric identities

$$\begin{aligned} \cos \alpha &= \cos^2\left(\frac{\alpha}{2}\right) - \sin^2\left(\frac{\alpha}{2}\right) \\ &= 2 \cos^2\left(\frac{\alpha}{2}\right) - 1 = 1 - 2 \sin^2\left(\frac{\alpha}{2}\right), \end{aligned} \quad (2.6.46)$$

we can replace the squares of sines and cosines of half the angles between peeling directions by cosines of the full angle. For instance,

$$\begin{aligned}\frac{1}{2} \sin^2\left(\frac{\varphi_{ab}}{2}\right) &= \frac{1}{4} (1 - \cos \varphi_{ab}), \\ \frac{1}{2} \cos^2\left(\frac{\varphi_{ab}}{2}\right) &= \frac{1}{4} (1 + \cos \varphi_{ab}).\end{aligned}\tag{2.6.47}$$

Comparison with the way we wrote the entries of a cell in a non-signaling array in Figure 2.4 then leads us to identify anti-correlation coefficients with these cosines:

$$\chi_{ab} = \cos \varphi_{ab}, \quad \chi_{ac} = \cos \varphi_{ac}, \quad \chi_{bc} = \cos \varphi_{bc}.\tag{2.6.48}$$

This is Eq. (2.6.1), the only result of our quantum-mechanical analysis of our peel-and-taste experiments in the Mermin-style setup that we needed in Section 2.6.1 to derive Eq. (2.6.9), the ellipsope inequality.

In closing, we want to highlight two aspects of the quantum-mechanical analysis of these experiments. First, the expressions for the probabilities entering the correlation arrays for these experiment are the squares of sines and cosines of angles that are directly related to the angles between the peeling directions. These expressions emerged naturally from the Born rule, the fundamental law governing probabilities in quantum mechanics, which sets these probabilities equal to the squares of cosines given by inner products of unit vectors in the Hilbert space for the system under consideration. In the case of our quantum bananas, the angles between the relevant unit vectors in Hilbert space, eigenvectors of operators representing tastes of bananas peeled in different directions, are half the angles between the corresponding peeling directions in ordinary space. The second aspect we want to highlight is that the Born rule is such that it precludes assigning probabilities to the taste of any one banana for more than one peeling direction. As we noted, this is what allows the class of correlations in our Mermin-style setup to be broader in quantum mechanics than in local hidden-variable theories. In other words, this is what allows these correlations to get beyond the tetrahedron in Figure 2.15 and saturate the ellipsope in Figure 2.16.



Chapter 3

The elliptope and the geometry of correlations

The Pearson correlation coefficient and the elliptope inequality • Why the quantum correlations saturate the elliptope • Why our raffles do not saturate the elliptope • The geometry of correlations: from Pearson and Yule to Fisher and De Finetti.

In Section 2.6, we used elements of quantum mechanics to derive the elliptope inequality (see Eq. (2.6.9)). We did so in two steps. In Section 2.6.2, we used the Born rule to show that the anti-correlation coefficients $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ parametrizing the correlations found in our banana peel-and-taste experiments in the Mermin-style setup are equal to the cosines of the angles $(\varphi_{ab}, \varphi_{ac}, \varphi_{bc})$ between the peeling directions given by the unit vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$. This means, as we noted in Section 2.6.1, that the elements of the anti-correlation matrix χ can be written as inner products of these unit vectors:

$$\chi \equiv \begin{pmatrix} 1 & \chi_{ab} & \chi_{ac} \\ \chi_{ab} & 1 & \chi_{bc} \\ \chi_{ac} & \chi_{bc} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{e}_a \cdot \mathbf{e}_a & \mathbf{e}_a \cdot \mathbf{e}_b & \mathbf{e}_a \cdot \mathbf{e}_c \\ \mathbf{e}_b \cdot \mathbf{e}_a & \mathbf{e}_b \cdot \mathbf{e}_b & \mathbf{e}_b \cdot \mathbf{e}_c \\ \mathbf{e}_c \cdot \mathbf{e}_a & \mathbf{e}_c \cdot \mathbf{e}_b & \mathbf{e}_c \cdot \mathbf{e}_c \end{pmatrix}.$$

The determinant of a matrix of this form (a so-called Gram matrix) cannot be negative, i.e., $\det \chi \geq 0$. Evaluation of this determinant then gave us the elliptope inequality.

This derivation of the elliptope inequality was thus a derivation *from within* quantum mechanics. In this chapter, we will show that the elliptope inequality can also be derived *from without*.¹ The constraint on (anti-)correlation coefficients it expresses thus has nothing to do with quantum mechanics per se. It is a general constraint on correlations between three random variables. From a paper on the early history of least-squares estimates (Aldrich, 1998), we learned that this general constraint can already be found in a paper by Udny Yule (1897, p. 487) on what today are called Pearson correlation coefficients.²

¹ For more on this from-within/from-without distinction, see Chapter 1, note 31, and Section 6.2.

² Yule was an associate of Karl Pearson and is remembered by historians of biology today for his role in bridging the divide between Mendelians and Darwinian biometrists, which

We will proceed as follows. In Section 3.1, we review some basic concepts in probability theory and statistics and introduce the Pearson correlation coefficient for a pair of random variables (see Eq. (3.1.8)). We then show that the three Pearson correlation coefficients for two out of three *balanced* random variables (see the definition labeled 3.1.1) must satisfy the ellipsope inequality. In Section 3.2, we apply this general result to the special case that the three balanced random variables are the tastes of Alice’s banana peeled in three different directions in the Mermin-style setup. We show that the three Pearson correlation coefficients for this case are equal to the quantities χ_{ab} , χ_{ac} and χ_{bc} parametrizing the correlations found in these experiments. This formally justifies our interpretation of these quantities as *anti*-correlation coefficients (the Pearson correlation coefficient for the taste of Alice’s banana peeled in the *a*-direction and the taste of Bob’s banana peeled in the *b*-direction is equal to *minus* χ_{ab}). Contrary to what our derivation in Section 2.6 suggested, χ_{ab} , χ_{ac} and χ_{bc} satisfy the ellipsope inequality, not because of something specific to quantum mechanics, but on much more general grounds, provided by probability theory and statistics.

What is remarkable, though, is that these quantum correlations fully saturate the ellipsope. The correlations produced by the raffles meant to simulate them only fill out the tetrahedron inscribed in the ellipsope (see Figure 2.16). Like all correlations between three random variables, they are subject to the ellipsope inequality but they are also subject to additional constraints, a CHSH-type Bell inequality associated with one of the facets of the tetrahedron and three inequalities like it associated with the other three facets (see Eqs. (2.5.2)–(2.5.5) and Eq. (3.2.31)). As we noted at the end of Chapter 1 and as we will discuss in detail in Section 3.2, what allows these additional inequalities to be violated in quantum mechanics is that in quantum mechanics a sum can have a definite value even when the individual terms in that sum do not.

would eventually result in the modern synthesis (Bowler, 2003, p. 329). Pearson and other biometrists, such as Francis Galton and Raphael Weldon are, of course, also remembered for their role in the eugenics movement (see, e.g., Kevles, 1985, Ch. II). Yule, in the paper of interest to us, which was received in December 1896 and presented in February 1897, refers to work by Auguste Bravais (1846) half a century earlier. In a historical note in Part III of his seminal paper on the mathematics of the biometrist version of Darwinian evolutionary theory, Pearson (1896) likewise writes that the correlation coefficient “appears in Bravais’ work, but a single symbol is not used for it” (quoted in Hald, 1998, p. 622). Pearson later “bitterly regretted this unbalanced evaluation” of Bravais’s contribution and tried to set the record straight in another historical note, “equally unbalanced, but in the other direction” (Hald, 1998, p. 623; see Pearson, 1921, p. 191).

As we will show in detail in Chapter 4, the situation is similar when we replace bananas behaving like spin- $\frac{1}{2}$ particles by more exotic bananas behaving like particles of higher spin s (with $s = 1, \frac{3}{2}, 2, \dots$) in our peel-and-taste experiments in the Mermin-style setup (still with two parties, Alice and Bob, and three settings/peelings per party but now with $2s + 1$ outcomes/tastes per setting). The quantum correlations continue to fill out the entire ellipptope, while the correlations produced by the raffles meant to simulate them only fill out inscribed polyhedra that, as the number of possible outcomes increases, have more and more facets and vertices and get closer and closer to the ellipptope (see Figures 4.11, 4.13 and 4.17). In Section 3.3, we will derive CHSH-like inequalities for these raffles designed to simulate the quantum correlations found in these higher-spin cases (see Eqs. (3.3.8)–(3.3.9)). These inequalities are associated either with one facet of the relevant polyhedron (in the case of half-integer spin values) or with one of the vertices where the polyhedron touches the ellipptope (in the case of integer spin values).

In Section 3.4, finally, we analyze a simple experiment involving what we will call a *3m balance* (see Figure 3.5) to illustrate the considerations in regression theory that originally led Yule to the ellipptope inequality. Following some remarkable papers by two famous statisticians a generation after Pearson and Yule, Ronald A. Fisher (1915, 1924) and Bruno De Finetti (1937),³ we will then sketch a geometrical perspective on Yule’s result, highlighting some features, in particular the notion of angles between random variables, highly suggestive of the Hilbert space formalism of quantum mechanics.

3.1 The Pearson correlation coefficient and the ellipptope inequality

Consider a random variable X that can take on the values x , where x can be an element of a discrete set $\{x_i\}$ or a continuous interval $[a, b]$ of real numbers or an element of the union of such sets and intervals. In other words, X can be discrete, continuous, or partly-discrete/partly-continuous. The random variables associated with the quantum observables we are interested in are all discrete with a finite number of possible integer or rational values x_i (with $i = 1, \dots, n$). An example of such a random variable would be the taste T_a^A of a

³ Fisher is remembered among many other things for his claim that Mendel faked his data (Franklin, Edwards, Fairbanks, Hartl, & Seidenfeld, 2008); De Finetti mainly for his advocacy of Bayesian personalism (McGrayne, 2011, Berkovitz, 2012, 2019). His first paper, however, was on population genetics and De Finetti diagrams are still used in that field.

quantum banana peeled in the a -direction by Alice represented by the operator \hat{T}_a^A in the one-banana Hilbert space \mathcal{H}_b (see Section 2.6.2). In that case $n = 2$ and the possible values x_1 and x_2 of this random variable are the eigenvalues $\pm \frac{1}{2}$ of this operator in units of b (see Eq. (2.6.32)). This is an example of a *balanced* random variable, which we define as follows. A random variable X is balanced if and only if

- (1) if x is a possible value, then $-x$ is a possible value as well; (3.1.1)
 (2) the value x is as likely to occur as $-x$, i.e., $\Pr(x) = \Pr(-x)$.

All quantum observables we are interested in are balanced discrete random variables. Focusing on this class of random variables will simplify the algebra we need for our argument in this chapter. This is because the *expectation value* $\langle X \rangle$ of a balanced variable vanishes.⁴ In the case of a discrete balanced random variable X , we have:⁵

$$\langle X \rangle = \sum_{i=1}^N x_i \Pr(x_i) = 0. \quad (3.1.2)$$

For a balanced random variable, the *variance*, defined as

$$\text{Var}(X) \equiv \langle (X - \langle X \rangle)^2 \rangle, \quad (3.1.3)$$

reduces to $\langle X^2 \rangle$ and the *standard deviation*, defined as its square root, becomes

$$\sigma_X \equiv \sqrt{\text{Var}(X)} = \sqrt{\langle X^2 \rangle}. \quad (3.1.4)$$

If X is balanced *and discrete*, the variance is given by

⁴ These angle brackets are different from the angle brackets used in the Dirac notation (cf. Section 3.2).

⁵ Consider a large random sample $\mathcal{S} = \{x_{i_1}, x_{i_2}, \dots, x_{i_N}\}$ of outcomes $x_{i_k} \in \{x_1, \dots, x_n\}$ of measurements of X , where $k = 1, \dots, N$ and $N \gg 1$ is the size of the sample \mathcal{S} . With the help of the *relative frequency* $f(x_i) \equiv \frac{\#x_i}{N}$ of the outcome x_i in \mathcal{S} (where $\#x_i$ is the number of occurrences of x_i in \mathcal{S}), the *average value* of X in \mathcal{S} can be written as

$$\bar{X} \equiv \frac{x_{i_1} + \dots + x_{i_N}}{N} = \sum_{i=1}^n x_i f(x_i),$$

where N is the size of the sample \mathcal{S} and n is the number of possible values of X . The *relative frequency* $f(x_i)$ in \mathcal{S} can be taken as a measure of the *probability* $\Pr(x_i)$ and the *average value* \bar{X} in \mathcal{S} can be taken as a measure of the *expectation value* $\langle X \rangle$.

$$\text{Var}(X) = \langle X^2 \rangle = \sum_{i=1}^n x_i^2 \text{Pr}(x_i). \quad (3.1.5)$$

We also need the *covariance* of X and some other random variable Y . This is defined as

$$\text{cov}(X, Y) \equiv \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle. \quad (3.1.6)$$

If X and Y are both balanced, the right-hand side is just the expectation value of their product. If they are also discrete and if (y_1, \dots, y_m) are the possible values of Y , we have, in analogy with Eq. (3.1.5):

$$\text{cov}(X, Y) = \langle XY \rangle = \sum_{i,j=1}^{n,m} x_i y_j \text{Pr}(x_i \text{ and } y_j). \quad (3.1.7)$$

The *Pearson correlation coefficient* is defined as the covariance of X and Y divided by the standard deviations for these two random variables:

$$\rho_{XY} \equiv \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}. \quad (3.1.8)$$

If X and Y are both balanced, this reduces to

$$\rho_{XY} = \frac{\langle XY \rangle}{\sigma_X \sigma_Y}. \quad (3.1.9)$$

Note that $\rho_{XY} = 1$ if $Y = X$, which makes sense as X is perfectly correlated with itself. Also note that ρ_{XY} is *symmetric*: $\rho_{YX} = \rho_{XY}$. The Pearson correlation coefficient then is the measure of the strength of a correlation that we alluded to in Section 2.1 (see note 3 in that section).

We will now show that the triplet $(\rho_{XY}, \rho_{XZ}, \rho_{YZ})$ of Pearson correlation coefficients for three arbitrary balanced (discrete or continuous) random variables X , Y and Z must satisfy the ellipptope inequality (see Eq. (2.6.9)). For any triplet of real numbers (v_1, v_2, v_3) , which can be thought of as the components of some vector \mathbf{v} , we have the inequality:

$$\left\langle \left(v_1 \frac{X}{\sigma_X} + v_2 \frac{Y}{\sigma_Y} + v_3 \frac{Z}{\sigma_Z} \right)^2 \right\rangle \geq 0. \quad (3.1.10)$$

This is trivially true: the expectation value of the square of some quantity cannot be negative. Expanding the expression before the \geq sign, we can rewrite this inequality as

$$\begin{aligned}
& v_1 \left(v_1 \frac{\langle X^2 \rangle}{\sigma_X^2} + v_2 \frac{\langle XY \rangle}{\sigma_X \sigma_Y} + v_3 \frac{\langle XZ \rangle}{\sigma_X \sigma_Z} \right) \\
& + v_2 \left(v_1 \frac{\langle YX \rangle}{\sigma_Y \sigma_X} + v_2 \frac{\langle Y^2 \rangle}{\sigma_Y^2} + v_3 \frac{\langle YZ \rangle}{\sigma_Y \sigma_Z} \right) \\
& + v_3 \left(v_1 \frac{\langle ZX \rangle}{\sigma_Z \sigma_X} + v_2 \frac{\langle ZY \rangle}{\sigma_Z \sigma_Y} + v_3 \frac{\langle Z^2 \rangle}{\sigma_Z^2} \right) \geq 0.
\end{aligned} \tag{3.1.11}$$

Comparing this expression with Eq. (3.1.9), we recognize a Pearson correlation coefficient in all nine terms before the \geq sign. The inequality thus turns into

$$\begin{aligned}
& v_1 \left(v_1 \rho_{XX} + v_2 \rho_{XY} + v_3 \rho_{XZ} \right) \\
& + v_2 \left(v_1 \rho_{YX} + v_2 \rho_{YY} + v_3 \rho_{YZ} \right) \\
& + v_3 \left(v_1 \rho_{ZX} + v_2 \rho_{ZY} + v_3 \rho_{ZZ} \right) \geq 0.
\end{aligned} \tag{3.1.12}$$

The expression before the \geq sign has the form of a matrix multiplied by \mathbf{v} both from the left and from the right:⁶

$$\left(v_1, v_2, v_3 \right) \begin{pmatrix} \rho_{XX} & \rho_{XY} & \rho_{XZ} \\ \rho_{YX} & \rho_{YY} & \rho_{YZ} \\ \rho_{ZX} & \rho_{ZY} & \rho_{ZZ} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \geq 0. \tag{3.1.13}$$

The correlation matrix ρ is defined in analogy with the anti-correlation matrix χ introduced in Eq. (2.6.2):

$$\rho \equiv \begin{pmatrix} \rho_{XX} & \rho_{XY} & \rho_{XZ} \\ \rho_{YX} & \rho_{YY} & \rho_{YZ} \\ \rho_{ZX} & \rho_{ZY} & \rho_{ZZ} \end{pmatrix} = \begin{pmatrix} 1 & \rho_{XY} & \rho_{XZ} \\ \rho_{XY} & 1 & \rho_{YZ} \\ \rho_{XZ} & \rho_{YZ} & 1 \end{pmatrix}, \tag{3.1.14}$$

where in the last step we used that, for any random variables X and Y , $\rho_{XX} = \rho_{YY} = 1$ and $\rho_{XY} = \rho_{YX}$.

A matrix satisfying the inequality in Eq. (3.1.12) for arbitrary \mathbf{v} is called *positive semi-definite*. The determinant of such a matrix cannot be negative:

⁶ Eq. (3.1.13) can be written more compactly as $\mathbf{v}^\top \rho \mathbf{v} \geq 0$, where \mathbf{v}^\top is the row vector that is the *transposed* of the column vector \mathbf{v} .

$\det \rho \geq 0$.⁷ Evaluation of this determinant gives

$$1 - \rho_{XY}^2 - \rho_{XZ}^2 - \rho_{YZ}^2 + 2\rho_{XY}\rho_{XZ}\rho_{YZ} \geq 0 \quad (3.1.15)$$

and Bub's your uncle: this is just the ellipptope inequality for $(\rho_{XY}, \rho_{XZ}, \rho_{YZ})$ instead of $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ as in Eq. (2.6.9).

As we did with $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ and $(\varphi_{ab}, \varphi_{ac}, \varphi_{bc})$ in Section 2.6, we can write the Pearson correlation coefficients $(\rho_{XY}, \rho_{XZ}, \rho_{YZ})$ as the cosines of the angles $(\vartheta_{XY}, \vartheta_{XZ}, \vartheta_{YZ})$. These angles will satisfy the angle inequality in Eq. (2.6.24) but, in general, will be unrelated to angles in either ordinary or Hilbert space. In Section 3.4, we return to the geometrical interpretation of random variables and angles between them. But first we apply the general results in probability and statistics presented in this section to the random variables in our quantum banana peel-and-taste experiment (Section 3.2) and to the random variables in the raffles meant to simulate those experiments (Section 3.3)

3.2 Why the quantum correlations saturate the ellipptope

Let the three random variables we considered in Section 3.1 be the taste of Alice's banana peeled in different directions, labeled a , b and c , in the Mermin-style setup:

$$X = T_a^A, \quad Y = T_b^A, \quad Z = T_c^A. \quad (3.2.1)$$

An obvious problem for this choice of random variables is that one and the same banana can only be peeled once (cf. Popescu, 2016, p. vi). Fortunately, as we will see, there are ways around this problem.

Random variables like T_a^A are balanced (cf. the definition labeled (3.1.1)). They have only two possible outcomes, $\pm 1/2$ in units of \hbar , and these two outcomes are equiprobable: $\Pr(1/2) = \Pr(-1/2) = 1/2$. This means that the *expectation value* $\langle T_a^A \rangle$ vanishes. It follows that the *variance* is given by

$$\langle (T_a^A)^2 \rangle = (1/2)^2 \Pr(1/2) + (-1/2)^2 \Pr(-1/2) = 1/4; \quad (3.2.2)$$

and the *standard deviation* by:

$$\sigma_a^A = \sqrt{\langle (T_a^A)^2 \rangle} = 1/2. \quad (3.2.3)$$

⁷ Let \mathbf{u} be an eigenvector of ρ with some (real) eigenvalue λ , i.e., $\rho \mathbf{u} = \lambda \mathbf{u}$. Since \mathbf{u} will satisfy the inequality in Eq. (3.1.13), we have (cf. note 6) $\mathbf{u}^\top \rho \mathbf{u} = \lambda \mathbf{u}^\top \mathbf{u} \geq 0$. Since $\mathbf{u}^\top \mathbf{u}$ is the inner product of \mathbf{u} with itself, which is the square of its length, it must be the case that $\lambda \geq 0$. A square matrix with non-negative eigenvalues has a non-negative determinant.

The variances and standard deviations for T_b^A , T_c^A and the corresponding variables for Bob, T_a^B , T_b^B and T_c^B , all have these same values, $1/4$ and $1/2$, respectively.

To find the *covariance* between, say, T_a^A and T_b^A , we run into the first of three hurdles that we will need to clear if we want to apply the results of Section 3.1 to these particular random variables. If Alice peels her banana in the a -direction to determine the value of T_a^A , she cannot peel that same banana in the b -direction to ascertain the value of T_b^A . However, we can have Bob taste the other banana in the same pair in the b -direction and use the opposite of the taste he finds as a proxy for what Alice would have found had she peeled her banana in the b -direction.⁸ Replacing T_a^A by $-T_b^B$ we can compute the covariance of T_a^A and T_b^A , using the correlation array in Figure 3.1 (cf. Figure 2.4) to find the probabilities of the four possible combinations of tastes.

		Bob		
		\hat{b}		
Alice			$+1/2$	$-1/2$
	\hat{a}		$+1/2$	$\frac{1}{4}(1 - \chi_{ab})$
		$-1/2$	$\frac{1}{4}(1 + \chi_{ab})$	$\frac{1}{4}(1 - \chi_{ab})$

Fig. 3.1 Cell in correlation array parametrized by $-1 \leq \chi_{ab} \leq 1$ (cf. Figure 2.4).

Proceeding in this way, we find:

$$\langle T_a^A T_b^A \rangle = -\langle T_a^A T_b^B \rangle = -\frac{1}{4} \cdot \frac{1}{2}(1 - \chi_{ab}) + \frac{1}{4} \cdot \frac{1}{2}(1 + \chi_{ab}) = \frac{1}{4}\chi_{ab}. \quad (3.2.4)$$

We use this result to calculate the *Pearson correlation coefficient* for T_a^A and T_b^A :

$$\rho_{T_a^A T_b^A} = \frac{\langle T_a^A T_b^A \rangle}{\sigma_a^A \sigma_b^A} = \frac{1/4 \chi_{ab}}{1/2 \cdot 1/2} = \chi_{ab}. \quad (3.2.5)$$

⁸ Counterfactual reasoning about the values that would have been obtained in a measurement, unlike counterfactual reasoning about the properties or ‘elements of reality’ possessed by systems (cf. Einstein, Podolsky, & Rosen, 1935), is innocuous; but if one prefers, for this particular situation, the derivation of the elliptope inequality in Section 3.1 can easily be modified in such a way that such counterfactual reasoning is avoided (see Eqs. (3.2.24)–(3.2.28) below).

This then is our formal justification for calling χ_{ab} an anti-correlation coefficient. This coefficient, introduced in Section 2.4 to parametrize the correlation between T_a^A and T_b^B found in our banana peel-and-taste experiment in the Mermin-style setup, is *minus* the Pearson correlation coefficient for this pair of variables:

$$\chi_{ab} = -\rho_{T_a^A T_b^B}. \quad (3.2.6)$$

Inserting the triplet of Pearson correlation coefficients,

$$(\rho_{XY}, \rho_{XZ}, \rho_{YZ}) = (\rho_{T_a^A T_b^A}, \rho_{T_a^A T_c^A}, \rho_{T_b^A T_c^A}) = (\chi_{ab}, \chi_{ac}, \chi_{bc}), \quad (3.2.7)$$

into the ellipptope inequality that we derived in Section 3.1 (see Eq. (3.1.15)), we recover the ellipptope inequality that we derived in Section 2.6 *from within* quantum mechanics (see in Eq. (2.6.9)):

$$1 - \chi_{ab}^2 - \chi_{ac}^2 - \chi_{bc}^2 + 2\chi_{ab}\chi_{ac}\chi_{bc} \geq 0. \quad (3.2.8)$$

So it looks as if we now derived this inequality *from without*, from general results in probability and statistics. In fact, we still need to clear two more hurdles before we can draw that conclusion. Like the first hurdle—how to make sense of covariances such as $\langle T_a^A T_b^A \rangle$ —these additional hurdles are created by the impossibility of peeling one and the same banana more than once. And like the first one, the other two hurdles are not insurmountable, even though clearing the third and final one turns on a result from quantum mechanics, which shows that this derivation of the ellipptope inequality is not entirely a derivation ‘from without’ (we will return to this point in Section 6.2).

The second hurdle is that, in forming linear combinations of χ_{ab} , χ_{ac} and χ_{bc} , we are combining data from different runs of the experiment. When Alice peels in the a -direction and Bob peels in the b -direction, giving us a value for $T_a^A T_b^A = -T_a^A T_b^B$, neither of them can, in the same run, peel in the c -direction to give us values for $T_a^A T_c^A = -T_a^A T_c^B = -T_a^B T_c^A$ or $T_b^A T_c^A = -T_b^A T_c^B = -T_b^B T_c^A$ as well. The way around this problem is to note that every pair of bananas that we pick from our quantum banana tree starts out in the same (singlet) state (see Eq. (2.6.38)). This is what makes it meaningful to consider expressions that combine, say, χ_{ab} and χ_{ac} , the former based on data obtained in runs in which one banana is peeled in the a -direction and the other one is peeled in the b -direction, the latter on data obtained in runs in which one is peeled in the a -direction and the other one is peeled in the c -direction.

In deriving the CHSH-type inequality in Eq. (2.5.1) and the Tsirelson bound in Eq. (2.6.16), we tacitly made the assumption that data from different runs of the experiment (randomly drawing a ticket from a basket in the case of these raffles) can be combined in this way. In the kind of local hidden-variable theories for which our raffles provide a model, this assumption can easily be avoided. We can simply change the protocol for our raffles and have Alice and Bob record the tastes for all three peelings on their halves of each ticket rather than pick just one. We trust that a moment's reflection will convince the reader that such a change of protocol does not change the correlation arrays for any of our raffles. In the quantum case, the assumption is unavoidable but equally innocuous.⁹

The third and final hurdle, however, looks more serious than the other two and to clear it we need to take an advance on our coverage of the quantum-mechanical formalism for spin in Section 4.1.1.¹⁰ The reason the expression in Eq. (3.1.15) is greater than or equal to zero is that it is simply a rewritten version of the expectation value of the square of a linear combination of X , Y and Z (see Eq. (3.1.10)). If $(X, Y, Z) = (T_a^A, T_b^A, T_c^A)$, there is no way we can assign values to all three of these variables in one run. Fortunately, as we noted at the end of Chapter 1 and in the introduction to this chapter, quantum mechanics routinely allows something inconceivable in classical theory, namely to assign a value to a linear combination of T_a^A , T_b^A and T_c^A without assigning values to those variables individually. It is this feature of quantum mechanics that gets us over the third hurdle.¹¹

⁹ Incidentally, this demonstrates the point we made in Chapter 1 that our raffles provide a (toy) model of a theory that suffers from the “big” but not the “small” measurement problem. Whenever a ticket is drawn from the basket, it is totally random who gets which half of the ticket. That means that it is totally random what outcome Alice and Bob will find when they check their ticket stub for the value for the setting they decide to check. So our toy theory runs afoul of the “big” measurement problem. The ticket stubs, however, have values for all settings, so there is nothing like the “small” measurement problem in our toy theory. We will return to this issue in Section 6.5.

¹⁰ We are extremely grateful to Wayne Myrvold for identifying this third hurdle when we presented a preliminary version of the result in this section in Viterbo in May 2019 (see the preface).

¹¹ This response to Wayne Myrvold's request for clarification (see the preceding note) was inspired by a footnote in *Wahrscheinlichkeitstheoretischer Aufbau* (Von Neumann, 1927b, p. 249, note 9). In this footnote, von Neumann points out that the Hamiltonian for the harmonic oscillator \hat{H} has a discrete spectrum even though it is the sum of two terms, $\hat{p}^2/2m$ and $\alpha \hat{q}^2$, that both have a continuous spectrum. The value of \hat{H} will not be the sum of the values of $\hat{p}^2/2m$ and $\alpha \hat{q}^2$ even though the expectation value $\langle \hat{H} \rangle$ will be the sum of the expectation

Recall that the possible values of T_a^A are the eigenvalues of the operator $\hat{T}_a^A : \mathcal{H}_b \rightarrow \mathcal{H}_b$ (see Eq. (2.6.32)). It will be convenient to divide the components of the vector $\mathbf{v} = (v_1, v_2, v_3)$ we introduced in Eq. (3.1.10) by $\sigma = \frac{1}{2}$, the standard deviation for variables like T_a^A (see Eq. (3.2.3)). Introducing the quantities

$$\lambda_a \equiv v_1/\sigma, \quad \lambda_b \equiv v_2/\sigma, \quad \lambda_c \equiv v_3/\sigma, \quad (3.2.9)$$

we can write Eq. (3.1.10), applied to our banana peel-and-taste experiment, as

$$\langle (\lambda_a \hat{T}_a^A + \lambda_b \hat{T}_b^A + \lambda_c \hat{T}_c^A)^2 \rangle \geq 0. \quad (3.2.10)$$

As we noted in Section 2.6.2, our quantum bananas behave like spin- $\frac{1}{2}$ particles. In Section 4.1.1, we will introduce the spin vector $\hat{\mathbf{S}}$ (see Eq. (4.1.3)). We will also introduce the operators $(\hat{S}_a, \hat{S}_b, \hat{S}_c)$, defined as the inner products of $\hat{\mathbf{S}}$ with the unit vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$ (see Eq. (4.1.11)). These operators acting in the Hilbert space for a single spin- $\frac{1}{2}$ particle behave in the exact same way as the operators $(\hat{T}_a^A, \hat{T}_b^A, \hat{T}_c^A)$ acting in the one-banana Hilbert space \mathcal{H}_b that we are considering here. These operators thus satisfy an inequality of the exact same form as the one in Eq. (3.2.10), which can be written as

$$\langle (\hat{\mathbf{S}} \cdot (\lambda_a \mathbf{e}_a + \lambda_b \mathbf{e}_b + \lambda_c \mathbf{e}_c))^2 \rangle \geq 0. \quad (3.2.11)$$

The minimum value of 0 is reached whenever

$$\lambda_a \mathbf{e}_a + \lambda_b \mathbf{e}_b + \lambda_c \mathbf{e}_c = 0. \quad (3.2.12)$$

This can only happen when the three unit vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$ are coplanar. Recall that the anti-correlation coefficients $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ for our peel-and-taste experiments in the Mermin-style setup can be written as inner products of the unit vectors in the corresponding peeling directions (see Eq. (2.6.3)). As we saw in Section 2.6.2 (the paragraph before Eq. (2.6.14)), these vectors are coplanar for values of $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ on the surface of the ellipptope.

For the four vertices the ellipptope shares with the tetrahedron (see Figure 2.16), they are not just coplanar but collinear as illustrated in Table 3.1.

Figure 3.2 shows two examples of triplets of coplanar (but not collinear) unit vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$ such that some linear combination of them adds up to zero. The triplet on the left is for the situation in which the angles between

values $\langle \hat{p}^2/2m \rangle$ and $\langle \alpha \hat{q}^2 \rangle$. We are grateful to Christoph Lehner for alerting one of us (Janssen) to this footnote back in 2009.

vertex	$\cos \varphi_{ab}$	$\cos \varphi_{ac}$	$\cos \varphi_{bc}$	\mathbf{e}_a	\mathbf{e}_b	\mathbf{e}_c
(i)	+1	+1	+1	↑	↑	↑
(ii)	+1	-1	-1	↑	↑	↓
(iii)	-1	+1	-1	↑	↓	↑
(iv)	-1	-1	+1	↑	↓	↓

Table 3.1 Unit vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$, with \mathbf{e}_a chosen as \uparrow , for the triplets of peeling directions corresponding to the points labeled (i) through (iv) of the elliptope in Figure 2.16.

all three peeling directions is 120° (cf. Figure 2.1). This is the combination of peeling directions that gives rise to the Mermin correlation array in Figure 2.6. In this case, the vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$ form an equilateral triangle with sides of unit length. The vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$ on the right in Figure 3.2 are for a generic choice of three peeling directions such that a triangle with sides of length $(\lambda_a, \lambda_b, \lambda_c)$ in the directions of these three unit vectors can be formed. Eq. (3.2.12) is thus satisfied.

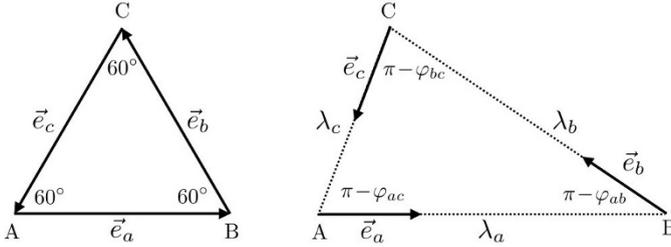


Fig. 3.2 Coplanar peeling directions $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$ resulting in triplets of correlation coefficients $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ coordinatizing points on the surface of the elliptope in Figure 2.16

As long as $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$ are coplanar but no two of them are collinear, it will be possible for any triplet of values for

$$\begin{aligned}
 \chi_{ab} &= \cos \varphi_{ab} = \mathbf{e}_a \cdot \mathbf{e}_b, \\
 \chi_{ac} &= \cos \varphi_{ac} = \mathbf{e}_a \cdot \mathbf{e}_c, \\
 \chi_{bc} &= \cos \varphi_{bc} = \mathbf{e}_b \cdot \mathbf{e}_c
 \end{aligned}
 \tag{3.2.13}$$

on the surface of the elliptope to construct a triangle with sides of length $(\lambda_a, \lambda_b, \lambda_c)$ in the directions of the triplet of corresponding coplanar unit

vectors $(\pm \mathbf{e}_a, \pm \mathbf{e}_b, \pm \mathbf{e}_c)$. If we need to flip one of the three unit vectors to form a triangle, we need to take minus the length of the corresponding side to satisfy Eq. (3.2.12). If we can use $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$ to form a triangle, the angles $(\varphi_{ab}, \varphi_{ab}, \varphi_{ab})$ will add up to 360° . If we have to flip one of the unit vectors to do so, one of the angles will be the sum of the other two (cf. our discussion of the angle inequalities in Eq. (2.6.24) at the end of Section 2.6.1).

This construction shows that by choosing the appropriate peeling directions in our quantum banana peel-and-taste experiment we can reach all points on the surface of the ellipptope. In particular we can reach the point $(\chi_{ab}, \chi_{ac}, \chi_{bc}) = (-1/2, -1/2, -1/2)$ at the center of the facet (ii)-(iii)-(iv) of the tetrahedron in Figure 2.17 corresponding to the ellipptope. For $v_a = v_b = v_c = 1$, Eq. (3.2.10) reduces to

$$\frac{1}{\sigma^2} \langle (\hat{T}_a^A + \hat{T}_b^A + \hat{T}_c^A)^2 \rangle \geq 0. \quad (3.2.14)$$

Expanding this expression, we find (cf. Eqs. (3.1.10)–(3.1.11))

$$\begin{aligned} \frac{1}{\sigma^2} \left(\langle (\hat{T}_a^A)^2 \rangle + \langle (\hat{T}_b^A)^2 \rangle + \langle (\hat{T}_c^A)^2 \rangle \right. \\ \left. + 2\langle \hat{T}_a^A \hat{T}_b^A \rangle + 2\langle \hat{T}_a^A \hat{T}_c^A \rangle + 2\langle \hat{T}_b^A \hat{T}_c^A \rangle \right) \geq 0. \end{aligned} \quad (3.2.15)$$

Since $\langle (\hat{T}_a^A)^2 \rangle = \langle (\hat{T}_b^A)^2 \rangle = \langle (\hat{T}_c^A)^2 \rangle = \sigma^2$ (see Eq. (3.2.3)) and

$$\frac{\langle \hat{T}_a^A \hat{T}_b^A \rangle}{\sigma^2} = \chi_{ab}, \quad \frac{\langle \hat{T}_a^A \hat{T}_c^A \rangle}{\sigma^2} = \chi_{ac}, \quad \frac{\langle \hat{T}_b^A \hat{T}_c^A \rangle}{\sigma^2} = \chi_{bc} \quad (3.2.16)$$

(see Eq. (3.2.5)), Eq. (3.2.15) simplifies to

$$3 + 2(\chi_{ab} + \chi_{ac} + \chi_{bc}) \geq 0, \quad (3.2.17)$$

in which we recognize the quantum counterpart to the CHSH-type Bell inequality for the Mermin-style setup (see Eq. (2.6.16)):

$$\chi_{ab} + \chi_{ac} + \chi_{bc} \geq -3/2. \quad (3.2.18)$$

The value $-3/2$, which is the Tsirelson bound for this setup, is reached at $(\chi_{ab}, \chi_{ac}, \chi_{bc}) = (-1/2, -1/2, -1/2)$.

Note that the vector $\mathbf{u} = (1, 1, 1)$ is an eigenvector with eigenvalue 0 of the anti-correlation matrix χ in Eq. (2.6.2) with $(\chi_{ab}, \chi_{ac}, \chi_{bc}) = (-1/2, -1/2, -1/2)$. If $\varphi_{ab} = \varphi_{ac} = \varphi_{bc} = 120^\circ$, then

$$\begin{aligned}
\chi &= \begin{pmatrix} 1 & \chi_{ab} & \chi_{ac} \\ \chi_{ab} & 1 & \chi_{bc} \\ \chi_{ac} & \chi_{bc} & 1 \end{pmatrix} = \begin{pmatrix} 1 & \cos \varphi_{ab} & \cos \varphi_{ac} \\ \cos \varphi_{ab} & 1 & \cos \varphi_{bc} \\ \cos \varphi_{ac} & \cos \varphi_{bc} & 1 \end{pmatrix} \\
&= \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix}.
\end{aligned} \tag{3.2.19}$$

Hence, for $\mathbf{u} = (1, 1, 1)$, $\chi \mathbf{u} = (0, 0, 0)$:

$$\begin{pmatrix} 1 & \chi_{ab} & \chi_{ac} \\ \chi_{ab} & 1 & \chi_{bc} \\ \chi_{ac} & \chi_{bc} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{3.2.20}$$

This is just one example of a general property. Let $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ be the coordinates of any point on the surface of the ellipsope that is not one of the four points the ellipsope shares with the tetrahedron. Let $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$ be a triplet of coplanar (but not collinear) unit vectors whose inner products give $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ (see Eq. (3.2.13)). Let the coefficients $(\lambda_a, \lambda_b, \lambda_c)$, chosen in such a way that Eq. (3.2.12) is satisfied, be the components of a vector \mathbf{u} . This vector will be an eigenvector with eigenvalue 0 of the anti-correlation matrix χ if its elements $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ are the coordinates of the point we chose on the surface of the ellipsope. This is a direct consequence of \mathbf{u} being an eigenvector with eigenvalue 0 of the matrix L introduced in Section 2.6 to write $\chi = L^\top L$ (see Eqs. (2.6.7)–(2.6.8)). That $L\mathbf{u}$ is the null vector, in turn, simply expresses the linear dependence of the three coplanar vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$. The columns of L are just the components of $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$. Using Eq. (2.6.4) for these components, one readily verifies that $L\mathbf{u}$ vanishes:

$$\begin{aligned}
\begin{pmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{pmatrix} \begin{pmatrix} \lambda_a \\ \lambda_b \\ \lambda_c \end{pmatrix} &= \lambda_a \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} + \lambda_b \begin{pmatrix} b_x \\ b_y \\ b_z \end{pmatrix} + \lambda_c \begin{pmatrix} c_x \\ c_y \\ c_z \end{pmatrix} \\
&= \lambda_a \mathbf{e}_a + \lambda_b \mathbf{e}_b + \lambda_c \mathbf{e}_c = 0,
\end{aligned} \tag{3.2.21}$$

where in the last step we used Eq. (3.2.12). It follows that $\chi \mathbf{u}$ also vanishes. Note that

$$\begin{pmatrix} \mathbf{e}_a \cdot \mathbf{e}_a & \mathbf{e}_a \cdot \mathbf{e}_b & \mathbf{e}_a \cdot \mathbf{e}_c \\ \mathbf{e}_b \cdot \mathbf{e}_a & \mathbf{e}_b \cdot \mathbf{e}_b & \mathbf{e}_b \cdot \mathbf{e}_c \\ \mathbf{e}_c \cdot \mathbf{e}_a & \mathbf{e}_c \cdot \mathbf{e}_b & \mathbf{e}_c \cdot \mathbf{e}_c \end{pmatrix} \begin{pmatrix} \lambda_a \\ \lambda_b \\ \lambda_c \end{pmatrix} \quad (3.2.22)$$

can be written as

$$\begin{pmatrix} \mathbf{e}_a \cdot (\lambda_a \mathbf{e}_a + \lambda_b \mathbf{e}_b + \lambda_c \mathbf{e}_c) \\ \mathbf{e}_b \cdot (\lambda_a \mathbf{e}_a + \lambda_b \mathbf{e}_b + \lambda_c \mathbf{e}_c) \\ \mathbf{e}_c \cdot (\lambda_a \mathbf{e}_a + \lambda_b \mathbf{e}_b + \lambda_c \mathbf{e}_c) \end{pmatrix}, \quad (3.2.23)$$

which vanishes on account of Eq. (3.2.21).

The result from quantum mechanics we invoked to get us over the third and final hurdle we faced when trying to apply the ellipptope inequality to our peel-and-taste experiment in the Mermin-style setup can also be used to avoid the counterfactual reasoning we resorted to to get over the first hurdle. In analogy with $\hat{\mathbf{S}}$, the spin vector, and $\hat{S}_a = \hat{\mathbf{S}} \cdot \mathbf{e}_a$, the spin component in the a -direction, we introduce $\hat{\mathbf{T}}$, the taste vector for one of our quantum bananas, and $\hat{T}_a = \hat{\mathbf{T}} \cdot \mathbf{e}_a$, its taste when peeled in the a -direction. More generally, we can define $\hat{T}_{\mathbf{w}} \equiv \hat{\mathbf{T}} \cdot \mathbf{w}$, where \mathbf{w} is an arbitrary vector (not necessarily a unit vector), as the taste of the banana (in units of \hbar) when peeled in the direction of \mathbf{w} multiplied by the length of \mathbf{w} . For \mathbf{w} we now choose (cf. Eqs. (3.2.12) and Eq. (3.2.21)):

$$\mathbf{w} = w_a \mathbf{e}_a + w_b \mathbf{e}_b + w_c \mathbf{e}_c, \quad (3.2.24)$$

with $w_a \equiv v_1/\sigma_a$, $w_b \equiv v_2/\sigma_b$ and $w_c \equiv v_3/\sigma_c$ (cf. Eq. (3.2.9)), where \mathbf{v} is an arbitrary vector with components (v_1, v_2, v_3) and $(\sigma_a, \sigma_b, \sigma_c)$ are the standard deviations of the taste variables (T_a, T_b, T_c) (these standard deviations are the same for Alice and Bob). If Alice measures $\hat{T}_{\mathbf{w}}^A$ and Bob measures $\hat{T}_{\mathbf{w}}^B$ in the same run, they are peeling in the same direction and are guaranteed to find opposite results. It follows that the covariance of these two variables must be less than or equal to zero:

$$\langle \hat{T}_{\mathbf{w}}^A \hat{T}_{\mathbf{w}}^B \rangle \leq 0. \quad (3.2.25)$$

Inserting

$$\begin{aligned} \hat{T}_{\mathbf{w}}^A &\equiv v_1 \frac{\hat{T}_a^A}{\sigma_a} + v_2 \frac{\hat{T}_b^A}{\sigma_b} + v_3 \frac{\hat{T}_c^A}{\sigma_c}, \\ \hat{T}_{\mathbf{w}}^B &\equiv v_1 \frac{\hat{T}_a^B}{\sigma_a} + v_2 \frac{\hat{T}_b^B}{\sigma_b} + v_3 \frac{\hat{T}_c^B}{\sigma_c} \end{aligned} \quad (3.2.26)$$

into Eq. (3.2.25) and proceeding along the exact same lines as in Eqs. (3.1.10)–(3.1.13), we can rewrite this inequality as

$$\begin{pmatrix} v_1, v_2, v_3 \end{pmatrix} \begin{pmatrix} \rho_{T_a^A T_a^B} & \rho_{T_a^A T_b^B} & \rho_{T_a^A T_c^B} \\ \rho_{T_b^A T_a^B} & \rho_{T_b^A T_b^B} & \rho_{T_b^A T_c^B} \\ \rho_{T_c^A T_a^B} & \rho_{T_c^A T_b^B} & \rho_{T_c^A T_c^B} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \leq 0. \quad (3.2.27)$$

The pairs of variables in the Pearson correlation coefficients in the matrix on the left-hand side can all actually be measured in different runs of our experiment, one variable by Alice, the other one by Bob. We thus directly use Bob’s results instead of using them as proxies for minus the results of measurements Alice does not and cannot perform.

Since Eq. (3.2.27) holds for arbitrary vectors $\mathbf{v} = (v_1, v_2, v_3)$, it follows that the determinant of this matrix of Pearson correlation coefficients must be less than or equal to zero:

$$\begin{vmatrix} -1 & \rho_{\hat{T}_a^A \hat{T}_b^B} & \rho_{\hat{T}_a^A \hat{T}_c^B} \\ \rho_{\hat{T}_a^A \hat{T}_b^B} & -1 & \rho_{\hat{T}_b^A \hat{T}_c^B} \\ \rho_{\hat{T}_a^A \hat{T}_c^B} & \rho_{\hat{T}_b^A \hat{T}_c^B} & -1 \end{vmatrix} = \begin{vmatrix} -1 & -\chi_{ab} & -\chi_{ac} \\ -\chi_{ab} & -1 & -\chi_{bc} \\ -\chi_{ac} & -\chi_{bc} & -1 \end{vmatrix} \leq 0, \quad (3.2.28)$$

where we used that $\rho_{T_a^A T_a^B} = \rho_{T_b^A T_b^B} = \rho_{T_c^A T_c^B} = -1$ for the diagonal components and Eq. (3.2.6) for the off-diagonal ones. Changing the minus signs to plus signs and the \leq -sign to a \geq -sign, we recover the elliptope inequality for the Mermin-style setup (cf. Eqs. (2.6.9) and (3.2.8)). Note that we did not need any counterfactual reasoning. The price we paid for avoiding counterfactuals is that we needed more input from quantum mechanics (in particular the inequality in Eq. (3.2.25)). This underscores that this derivation of the elliptope inequality is not entirely ‘from without’ but involves elements ‘from within’ (see Section 6.2 for further discussion).

One conclusion of the analysis in this section is that it is unsurprising that quantum mechanics does not allow any non-signaling correlations that violate the Tsirelson bound—or, more generally, any non-signaling correlations represented by points inside the non-signaling cube but outside the elliptope. This is not because of some elusive physical principle over and above non-signaling but simply because of the general constraint on (anti-)correlation coefficients given by the elliptope inequality in Eq. (3.1.15). What *is* surprising in light of this general constraint, is that the correlations allowed in our quantum banana

peel-and-taste experiment get beyond the classical tetrahedron and fill out the entire ellipptope. The taste of these bananas, after all, can only be $\pm\frac{1}{2}$ (in units of \hbar). Suppose we try simulate this experiment with a raffle of the kind we introduced in Section 2.5. Let (X_a^A, X_b^A, X_c^A) and (X_a^B, X_b^B, X_c^B) be the random variables mimicking the tastes of the bananas peeled by Alice and Bob in the directions labeled a, b and c . Since these variables can only take on the values $\pm\frac{1}{2}$, the linear combination $(X_a^A + X_b^A + X_c^A)/\sigma$, with $\sigma = 1/2$ can never be less than 1 (the same would obviously be true if we replace the variables pertaining to Alice’s half of the tickets by the variables pertaining to Bob’s). For these variables, Eq. (3.2.14) would thus become

$$\frac{1}{\sigma^2} \langle (X_a^A + X_b^A + X_c^A)^2 \rangle \geq 1. \quad (3.2.29)$$

If we parametrize the correlations between these three variables with the anti-correlation coefficients $(\chi_{ab}, \chi_{ac}, \chi_{bc})$, we would accordingly have to change Eq. (3.2.17) to

$$3 + 2(\chi_{ab} + \chi_{ac} + \chi_{bc}) \geq 1, \quad (3.2.30)$$

and Eq. (3.2.18) to

$$\chi_{ab} + \chi_{ac} + \chi_{bc} \geq -1. \quad (3.2.31)$$

This is just the CHSH-type Bell inequality for the Mermin-style setup that we found in Section 2.5 (see Eq. (2.5.1)). Quantum mechanics, to repeat, is less restrictive because it allows something no classical theory allows, namely to assign a value to the sum of three variables without assigning a value to all three of them individually.¹²

Bub (2007, 2008, 2010b, 2016, 2019b) and others have argued (see, e.g., Rieffel & Polak, 2011, pp. 326–327) that the speedup of a quantum computer comes from the ability to assign a truth value to a disjunction without having to assign a truth value to its disjuncts.¹³ Long before anybody was thinking about quantum computing, however, physicists had already run into a version of the conundrum encountered here and resolved by quantum mechanics. To

¹² Von Neumann’s example of the energy of an harmonic oscillator (see note 11) provides a simpler example. Quantum mechanics allows us to assign a value to the total energy of the harmonic oscillator, the sum of its kinetic and its potential energy, but not to the two terms in this sum separately, as the operators for momentum and position appearing in the kinetic and potential term, respectively, do not commute.

¹³ This is illustrated by a scenario involving a “Quantum Quasino” in *Totally Random* (Bub & Bub, 2018, pp. 186–215). This scenario, which strictly speaking is about quantum communication rather than quantum computation, follows Popescu (2014).

dispel any whiff of quantum-information parochialism (cf. our comments in Chapter 1), we remind the reader of a textbook example.

In their popular introductory textbook on quantum physics, [Eisberg & Resnick \(1985, p. 258\)](#) highlight the peculiar behavior of angular momentum in quantum mechanics. That the wave function for a one-electron atom “does not describe a state with a definite x and y component of orbital angular momentum,” they note, “is mysterious from the point of view of classical mechanics” (and is equally mysterious for intrinsic angular momentum, i.e., spin). In quantum mechanics, they continue, this is required by the uncertainty principle. If the z -component has a definite value, the x - and y -components cannot have definite values, as the corresponding operators do not commute (see Section 2.6.2).

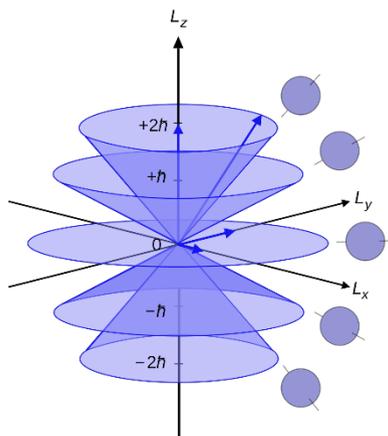


Fig. 3.3 Vector model of orbital angular momentum in the old quantum theory (Wikimedia Commons; cf. [Eisberg & Resnick, 1985, p. 258, Fig. 7-12](#)).

This behavior of angular momentum, [Eisberg & Resnick \(1985, pp. 258–259\)](#) note in the next paragraph, “can be conveniently represented by a *vector model*,” i.e., the one shown in Figure 3.3. In this model, the angular momentum vector precesses around the z -axis at a fixed angle determined by the value of its z -component. While the z -component remains fixed, the x - and y -components are constantly changing. This model, of course, still does not capture the true state of affairs in quantum mechanics where it is impossible to assign a definite value to all three components at any instant. The vector model is a left-over from the old quantum theory, where it was introduced to deal with problems posed

by multiplet spectra and the anomalous Zeeman effect (Duncan & Janssen, 2019, Sec. 1.3.6 and Ch. 7).

The quantum-mechanical treatment of angular momentum not only solved these problems in spectroscopy, it also restored order in a completely different field, the theory of electric and magnetic susceptibilities. As John H. Van Vleck, the author of an authoritative book on the subject, pointed out in the opening sentence of its preface: “The new quantum mechanics is perhaps most noted for its triumphs in the field of spectroscopy, but its less heralded successes in the theory of electric and magnetic susceptibilities must be regarded as one of its great achievements” (Van Vleck, 1932, p. vii). Since we inserted this digression to preempt charges of parochialism against the informational interpretation of quantum mechanics we are championing (see Chapter 1), it is amusing to note that Van Vleck in the aftermath of the quantum revolution of the mid-1920s berated himself and his colleagues for having been too focused on spectroscopy. In an article on the new quantum mechanics in a chemistry journal, he wryly observed that “[t]he chemist is apt to conceive of the physicist as some one who is so entranced in spectral lines that he closes his eyes to other phenomena” (Van Vleck, 1928, p. 493).¹⁴

To close this section, let us reiterate what, for our purposes, is its most important lesson. The ellipsope inequality, which can be derived *from within* quantum mechanics (from the geometry of Hilbert space) can also be derived *from without*, as a general constraint on correlations between three balanced random variables. In Chapter 6, we will return to this lesson and discuss its significance for our overall argument in this volume.

3.3 Why our raffles do not saturate the ellipsope

We can find CHSH-type inequalities like the one in Eq. (3.2.31) for raffles we will design in Section 4.2 trying to simulate correlations found in measurements on pairs of entangled particles of spin $s = 1, 3/2, 2, 5/2, \dots$ in the singlet state in the Mermin-style setup. The spin in any direction (corresponding to the taste of a banana peeled in that direction) can take on $2s+1$ different values $m\hbar$, where $m \in \{-s, -s+1, \dots, s-1, s\}$ and \hbar is Planck’s constant h divided

¹⁴ Both quotations are taken from Midwinter & Janssen (2013, p. 137). Van Vleck’s solution to the problem of susceptibilities hinges on the correct quantum-mechanical treatment of angular momentum, in this case of diatomic molecules such as hydrogen chloride (ibid., p. 199). We will return to this episode in Section 6.4 as an example of a problem solved by the new kinematics of quantum mechanics.

by 2π . We will analyze the correlations quantum mechanics predicts for such experiments in Section 4.1.

		Bob	
		\hat{a}	
Alice	\hat{a}		
	s	$-s$	
\hat{a}	s	0 $\frac{1}{2s+1}$ $\frac{1}{2s+1}$ 0	
	$-s$	$\frac{1}{2s+1}$ 0 0 $\frac{1}{2s+1}$	

Fig. 3.4 Cell along the diagonal of a non-signaling correlation array with $2s + 1$ outcomes per setting.

Figure 3.4 shows a cell along the diagonal of the correlation array for the quantum correlations for arbitrary spin s . It tells us that, when Alice and Bob use the same setting, all $2s+1$ ways in which they can find opposite results (including $m = 0$ if s is an integer) occur with equal probability. These quantum correlation arrays are non-signaling by virtue of having uniform marginals: in every row and every column of every cell, on or off the diagonal of the correlation array, the entries add up to $1/(2s+1)$.

In designing raffles for these higher-spin cases, we are immediately faced with a complication compared to the spin- $\frac{1}{2}$ case (see Chapter 4 for further discussion). Regardless of how many possible outcomes per setting there are, we can only put two outcomes per setting on any given ticket. This simple observation, unfortunately, has serious consequences for the design of our raffles. The reason we did not have to worry about this so far is that, for $s = 1/2$, there are only two possible outcomes per setting in the quantum experiment we are trying to simulate with our raffles. For $s > 1/2$, however, there are $2s + 1$ possible outcomes per setting. Here is why this is a problem. As we just saw, for any value of s , the quantum correlations are such that if Alice and Bob use the same setting, they are both expected to find all $2s + 1$ possible outcomes in equal proportion. There is no way we can simulate this feature of the quantum correlations with single-ticket raffles as these tickets can at most have two of these $2s + 1$ outcomes printed on them for each setting (for integer values of s it is possible that they have one and the same value, i.e., 0, on both sides). In other

words, for $s > 1/2$, the correlations we can produce with single-ticket raffles, while non-signaling by construction, do not give uniform marginals (see, e.g., Figure 4.8 in Section 4.2). To make sure that our raffles at least simulate the cells along the diagonal of the quantum correlation arrays correctly, we thus need to restrict ourselves to mixed raffles that do give uniform marginals. In Section 4.2, we will show how to construct those. All we need at this point is that any admissible raffle gives rise to a correlation array in which the cells along its diagonal have the form shown in Figure 3.4. Since the correlation arrays for our raffles will always be non-signaling, this implies that they give uniform marginals.

Consider an arbitrary (single-ticket or mixed) raffle with tickets with opposite outcomes on the left and the right side for the three settings \hat{a} , \hat{b} and \hat{c} and with $2s+1$ possible outcomes for each of these three settings. As in the case of two outcomes per setting considered so far, we can write (cf. Eqs. (3.2.14)–(3.2.18) and Eqs. (3.2.29)–(3.2.31)):

$$\begin{aligned} \langle (X_a^A + X_b^A + X_c^A)^2 \rangle &= \sigma_a^2 + \sigma_b^2 + \sigma_c^2 \\ &\quad - 2(\langle X_a^A X_b^B \rangle + \langle X_a^A X_c^B \rangle + \langle X_b^A X_c^B \rangle), \end{aligned} \quad (3.3.1)$$

where, as before, we use $-X_a^A X_b^B$ as a proxy for $X_a^A X_b^A$, etc.

The standard deviations are the square roots of variances computed for the cells along the diagonals of the relevant correlation arrays. Since these cells will all be of the form of the one in Figure 3.4, the three standard deviations in Eq. (3.3.1) will have the same value σ_s , given by

$$\sigma_s^2 = \sigma_a^2 = \langle (X_a^A)^2 \rangle|_{\text{UM}} = -\langle X_a^A X_a^B \rangle|_{\text{UM}}, \quad (3.3.2)$$

where $|_{\text{UM}}$ indicates that the (co-)variance be evaluated for a raffle giving uniform marginals. Inspection of Figure 3.4 and the well-known sum-of-squares formula gives

$$\sigma_s^2 = - \sum_{m=-s}^s \frac{(m\hbar)(-m\hbar)}{2s+1} = \frac{\hbar^2}{2s+1} \sum_{m=-s}^s m^2 = \frac{1}{3}s(s+1)\hbar^2. \quad (3.3.3)$$

For raffles giving uniform marginals, we have

$$\chi_{ab}|_{\text{UM}} = - \left(\frac{\langle X_a^A X_b^B \rangle}{\sigma_a \sigma_b} \right) \Big|_{\text{UM}} = - \frac{1}{\sigma_s^2} \langle X_a^A X_b^B \rangle|_{\text{UM}}. \quad (3.3.4)$$

Similar expressions obtain for χ_{ac} and χ_{bc} . For such raffles, Eq. (3.3.1) can be rewritten as

$$\langle (X_a^A + X_b^A + X_c^A)^2 \rangle_{\text{UM}} = \sigma_s^2 (3 + 2(\chi_{ab} + \chi_{ac} + \chi_{bc}))_{\text{UM}}. \quad (3.3.5)$$

For any half-integer value of s , $|X_a^A + X_b^A + X_c^A|$ cannot be made smaller than $\hbar/2$, hence

$$\langle (X_a^A + X_b^A + X_c^A)^2 \rangle \geq \frac{\hbar^2}{4} \quad \text{for half-integer } s. \quad (3.3.6)$$

For any integer value of s , $X_a^A + X_b^A + X_c^A$ can be made to vanish, hence

$$\langle (X_a^A + X_b^A + X_c^A)^2 \rangle \geq 0 \quad \text{for integer } s. \quad (3.3.7)$$

Restricting ourselves to raffles giving uniform marginals, in which case we can use Eq. (3.3.5), we thus find the following CHSH-type lower bounds on the sum of the anti-correlation coefficients in the Mermin-style setup:

$$(\chi_{ab} + \chi_{ac} + \chi_{bc})_{\text{UM}} \geq \frac{\hbar^2}{8\sigma_s^2} - \frac{3}{2} \quad \text{for half-integer } s, \quad (3.3.8)$$

$$(\chi_{ab} + \chi_{ac} + \chi_{bc})_{\text{UM}} \geq -\frac{3}{2} \quad \text{for integer } s. \quad (3.3.9)$$

For $s = 1/2$, $\sigma_s^2 = \hbar^2/4$ and Eq. (3.3.8) reduces to Eq. (3.2.31), the CHSH-like inequality for this setup (since all raffles for the spin- $1/2$ case give uniform marginals, the restriction $_{\text{UM}}$ can be dropped).

Eq. (3.3.9) tells us that, for integer values of s , we can (at least in principle) always reach the Tsirelson bound for a Mermin-style setup with $2s + 1$ outcomes per setting (in Section 4.2 we will design raffles that do indeed reach this bound). Eq. (3.3.8) tells us that, for half-integer values of s , this is true only in the limit that s goes to infinity, in which case the number of outcomes $2s + 1$ and the standard deviation σ_s also go to infinity.

As we will show in detail for the $s = 1$ case in Section 4.2, even though we *can* reach the Tsirelson bound on the sum of anti-correlation coefficients with our raffles, these raffles still *cannot* reproduce all individual entries in the correlation arrays for the quantum correlations they are supposed to simulate (see Eq. (4.2.23)). The reason for this will also become clear in Section 4. In Section 4.1, we will show that, regardless of the spin s of the two particles involved, the probabilities in any given cell in these quantum correlation arrays

can still be parametrized by the angle between measuring directions. In Section 4.2, we will see that in our raffles this is true only for the simple case $s = 1/2$ of two outcomes per setting that we have been considering so far. Even in the case of $s = 1$, we already need two parameters to specify the entries in any off-diagonal cell in our correlation arrays (see Figure 4.12).

To close this short section, we want to highlight one intriguing result of our analysis of the spin-1 case in Section 4.2.2. If we replace spin- $1/2$ particles by spin-1 particles in the Mermin-style setup, it would only take one particular outcome to rule out a local hidden-variable account of the statistics predicted by quantum mechanics. If the angles between the three measuring directions in this setup are all 120° , quantum mechanics predicts that there is a sizable probability (i.e., $1/8$) that Alice and Bob will find opposite results in runs in which they use different measuring directions (see Eq. (4.2.23)). In the raffle constructed to model a local hidden-variables account of this experiment, however, this probability has to be strictly zero, otherwise the raffle cannot reproduce other key features of the quantum statistics for the experiment (once again, see Eq. (4.2.23)). A local hidden-variable account of this experiment would thus be ruled out by a single run (or, allowing for the detectors not being 100% reliable, a few runs) in which Alice and Bob find opposite results when measuring spin in different directions. What makes this even more remarkable is that this is true even though the quantum correlations in this case do *not* violate any Bell inequality: the point representing these correlations in the non-signaling cube lies on the ellipsope and is a vertex of three facets of the relevant correlation polyhedron (see Figure 4.11).

3.4 The geometry of correlations: from Pearson and Yule to Fisher and De Finetti

In this section, we indicate how Yule (1897) found the general constraint on correlation coefficients in Eq. (3.1.15) in the context of regression theory (i.e., finding the straight lines best approximating correlations between variables) and how Fisher (1915, 1924) and De Finetti (1937) recovered the result Yule found algebraically by treating random variables as vectors and cashing out correlations in terms of angles between those vectors. The importance of this geometric approach was emphasized by Pearson:

It is greatly to be desired that the “trigonometry” of higher dimensional plane space should be fully worked out, for all our relations between multiple correlation and partial correlation coefficients of n variates are properties of the “angles,” “edges”

and “perpendiculars” of spheropolyhedra in multiple space. It would be a fine task for an adequately equipped pure mathematician to write a treatise on “spherical polyhedrometry”; he need not fear that his results would be without practical application for they embrace the whole range of problems from anatomy to medicine and from medicine to sociology and ultimately to the doctrine of evolution (Pearson, 1916, p. 237).

Linear algebra has proven to be much more convenient than spherical geometry for dealing with the relevant problems in statistics. As Aldrich (1998, p. 73) notes, “[t]he treatise on the trigonometry of correlations . . . never materialized. [Kendall, 1961] is a partial offering but it appeared just as a new approach was taking off. This drew on the Hilbert space theory developed in the early part of the century and assembled by [Stone, 1932].”

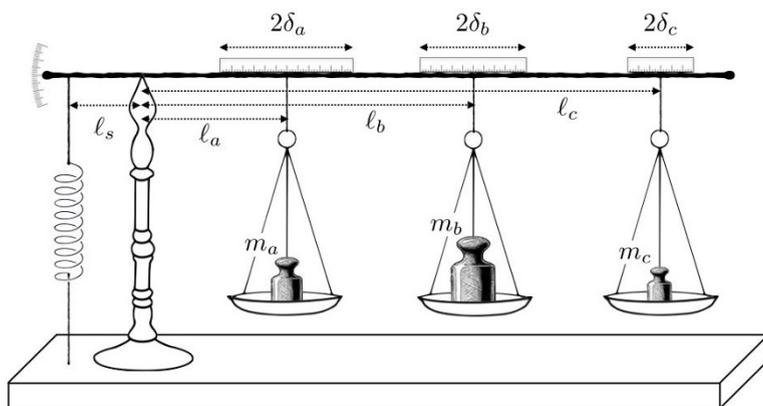


Fig. 3.5 *3m balance*: an elementary example of a correlation between three random variables $(X_a, X_b, X_c) \equiv (\Delta\ell_a, \Delta\ell_b, \Delta\ell_c)$ (with $|\Delta\ell_a| \leq \delta_a$, $|\Delta\ell_b| \leq \delta_b$ and $|\Delta\ell_c| \leq \delta_c$). The values for the displacements $(\Delta\ell_a, \Delta\ell_b, \Delta\ell_c)$ from the points (ℓ_a, ℓ_b, ℓ_c) are subject to the constraint that the torques coming from the three scale pans on the right with masses m_a , m_b and m_c exactly cancel the opposing torque coming from the spring on the left, so that the beam remains horizontal as it is for $(\Delta\ell_a, \Delta\ell_b, \Delta\ell_c) = (0, 0, 0)$.

Rather than discussing the geometry of correlations in the abstract, we use the concrete example of three balanced random variables (X_a, X_b, X_c) illustrated in Figure 3.5. This figure shows a device we will call a *3m balance*, so named for the masses m_a , m_b and m_c in the three scale pans hanging from its beam. In the configuration shown in the figure, the three pans are at distances ℓ_a , ℓ_b and ℓ_c from the beam’s pivot point. In this configuration, the torques coming from the three pans on the right exactly cancel the torque coming from the spring

on the left:

$$F_s \ell_s = g(m_a \ell_a + m_b \ell_b + m_c \ell_c). \quad (3.4.1)$$

Here F_s is the force exerted by the spring, ℓ_s its distance to the pivot point and g the acceleration of gravity. Now imagine we take these three pans off the beam and put them back on (with the same masses as before) such that (a) the system is once again perfectly balanced (i.e., the beam is horizontal) and (b) the pans with m_a , m_b and m_c are somewhere in the intervals $\ell_a \pm \delta_a$, $\ell_b \pm \delta_b$ and $\ell_c \pm \delta_c$, respectively. Suppose we end up putting the pans at the displaced positions

$$\ell_a + \Delta \ell_a, \quad \ell_b + \Delta \ell_b, \quad \ell_c + \Delta \ell_c, \quad (3.4.2)$$

where $|\Delta \ell_a| \leq \delta_a$, $|\Delta \ell_b| \leq \delta_b$ and $|\Delta \ell_c| \leq \delta_c$. The displacements $(\Delta \ell_a, \Delta \ell_b, \Delta \ell_c)$ that characterize this new perfectly balanced configuration of the system will satisfy the linear equation

$$m_a \Delta \ell_a + m_b \Delta \ell_b + m_c \Delta \ell_c = 0. \quad (3.4.3)$$

Suppose we repeat this many times. The triplets of displacements $(\Delta \ell_a, \Delta \ell_b, \Delta \ell_c)$ found in consecutive runs of this experiment, all satisfying Eq. (3.4.3), can then be treated as triplets of values for the three random variables

$$(X_a, X_b, X_c) \equiv (\Delta \ell_a, \Delta \ell_b, \Delta \ell_c). \quad (3.4.4)$$

Since we are as likely to find $\Delta \ell_a$ as $-\Delta \ell_a$, $\Delta \ell_b$ as $-\Delta \ell_b$ and $\Delta \ell_c$ as $-\Delta \ell_c$, these variables are balanced (see the definition labeled (3.1.1)).¹⁵ They are correlated: if we vary the position of one of the masses, we need to vary the position of at least one of the other two to satisfy Eq. (3.4.3). This is expressed by the linear relation

$$m_a X_a + m_b X_b + m_c X_c = 0 \quad (3.4.5)$$

between the three variables, which is just Eq. (3.4.3) with $\Delta \ell$ replaced by X .

Note that in this simple example there is no problem obtaining values for all three variables in every run. In our quantum banana peel-and-taste experiment,

¹⁵ Since X_a is balanced, its expectation value is zero and its variance is given by

$$\langle (X_a - \langle X_a \rangle)^2 \rangle = \langle X_a^2 \rangle = \frac{1}{2\delta_a} \int_{-\delta_a}^{\delta_a} X_a^2 dX_a = \frac{1}{6\delta_a} X_a^3 \Big|_{-\delta_a}^{\delta_a} = \frac{\delta_a^2}{3}.$$

The corresponding standard deviation is: $\sigma_a = \sqrt{\langle X_a^2 \rangle} = \delta_a/\sqrt{3}$. Similarly, $\sigma_b = \delta_b/\sqrt{3}$ and $\sigma_c = \delta_c/\sqrt{3}$.

we could only ascertain the values (tastes) for two of the three variables (peelings). The experiment with our $3m$ balance, however, does share several features with both our quantum banana experiment and the raffles meant to simulate them. As with any three balanced random variables the allowed values for the corresponding three Pearson correlation coefficients ($\rho_{ab}, \rho_{ac}, \rho_{bc}$) are bound by the ellipsope inequality. The correlation will be represented by a point on the surface of the ellipsope if some linear combination of the three variables (or, in the quantum case, *the operator representing* the three variables) gives zero. Eq. (3.4.5) gives this linear combination in the case of the $3m$ balance. This relation will hold as long as we can ignore errors in our measurement of the displacements ($\Delta\ell_a, \Delta\ell_b, \Delta\ell_c$) and as long as there are no other masses pulling on the balance's beam. We can represent both of these complicating factors by an additional pan containing some unknown mass pulling on the beam at some unknown location. In that case, the right-hand side of Eq. (3.4.5) is no longer zero. That, in turn, means that the correlation between these three variables will be represented by a point inside the ellipsope rather than on its surface.¹⁶

If we set $m_a = m_b = m_c = m$ and $\delta_a = \delta_b = \delta_c = \delta$ and only allow the values $\pm\delta/2$ for the variables X_a, X_b and X_c , their sum can no longer be made to vanish. To balance the beam in this case we need to put our thumb on the scale. This too can be represented by an additional pan on the beam. This pan must provide a torque of $mg\delta/2$ to compensate for the smallest torque possible coming from the other three pans combined. The quantity $(X_a + X_b + X_c)^2$ can never be less than $\delta^2/4$ in this case and we can no longer reach the point $(\rho_{ab}, \rho_{ac}, \rho_{bc}) = (-1/2, -1/2, -1/2)$ on the ellipsope, corresponding to the Tsirelson bound in our quantum banana experiment. Instead we find ourselves right back where we were when we tried to design a raffle to simulate the correlations found in our quantum banana peel-and-taste experiments (cf. Eq. (3.2.29)–(3.2.31) above).

In our discussion of the $3m$ balance so far, we have tacitly assumed that we know the masses in the three pans and thus the coefficients in the linear relation between X_a, X_b and X_c in Eq. (3.4.5). Typically we will not know those coefficients. Pearson, Yule and Fisher were especially interested in biological variables. These will seldom satisfy a simple linear relation such as the one

¹⁶ This provides a simple way to understand a comment by Richard Holt, the second H in CHSH, in an interview in 2001 about finding a result in an early test of the CHSH inequality that did not agree with the quantum-mechanical prediction: “whenever you’re looking for a stronger correlation, any kind of systematic error you can imagine typically weakens it and moves it toward the hidden-variable range” (Gilder, 2008, p. 286).

in Eq. (3.4.5). Suppose, however, that we have reason to believe that three balanced random variables do satisfy a linear relation of this form, with unknown coefficients and a non-zero right-hand side. In terms of our $3m$ balance this corresponds to a situation in which we do not know the masses in the three pans and cannot rule out that there is a fourth pan somewhere on the beam with another unknown mass. Taking the pans with the masses m_a , m_b and m_c off the beam of the $3m$ balance and putting them back on many times, making sure in each run of the experiment that the beam is balanced, and then subjecting the many triplets of values for the triplet of variables (X_a, X_b, X_c) found in this experiment to statistical analysis, we can make reliable estimates of these masses, which enter as coefficients of X_a , X_b and X_c in a linear relation between these variables.

One technique we can use is what is called *bivariate regression* of one variable, say X_c , on two other ones, say X_a and X_b , to determine the value of two of the coefficients in the suspected linear relation between these three variables given the value of the third coefficient. In the case of the experiment with the $3m$ balance, we start by rewriting Eq. (3.4.5) as

$$X_c = - \left(\frac{m_a}{m_c} \right) X_a - \left(\frac{m_b}{m_c} \right) X_b. \quad (3.4.6)$$

It follows that the ratios m_a/m_c and m_b/m_c satisfy the following pair of linear equations

$$\begin{aligned} \langle X_a X_c \rangle &= - \left(\frac{m_a}{m_c} \right) \langle X_a^2 \rangle - \left(\frac{m_b}{m_c} \right) \langle X_a X_b \rangle, \\ \langle X_b X_c \rangle &= - \left(\frac{m_a}{m_c} \right) \langle X_b X_a \rangle - \left(\frac{m_b}{m_c} \right) \langle X_b^2 \rangle. \end{aligned} \quad (3.4.7)$$

We then use the statistical data found in the $3m$ -balance experiment to evaluate variances and covariances, insert the results into Eq. (3.4.7) and solve for m_a/m_c and m_b/m_c . Given the value of m_c , this gives us the values of m_a and m_b .

We illustrate this procedure with a simple numerical example. To keep things as simple as possible we choose the same value δ for the three parameters δ_a , δ_b and δ_c and the same value m for the three masses m_a , m_b and m_c . We also restrict the spots on the balance beam where we can hang the three pans to just five spots per pan, one at the center of the ranges shown in Figure 3.5, two at the edges, and two exactly halfway between the center and the edges. So, in units of $\delta/2$, X_a , X_b and X_c can only take on the values $0, \pm 1$ and ± 2 . Suppose

run	X_a	X_b	X_c
1	2	-1	-1
2	-2	0	2
3	0	-2	2
4	-1	1	0
5	1	-2	1
6	-2	2	0
7	1	1	-2
8	2	0	-2
9	1	-1	0

run	X_a	X_b	X_c
10	-1	0	1
11	2	-2	0
12	0	1	-1
13	-1	2	-1
14	1	0	-1
15	0	-1	1
16	0	2	-2
17	-2	1	1
18	-1	-1	2

Table 3.2 Representative sample of 18 runs of an experiment with the $3m$ balance shown in Figure 3.5, in which the three pans with masses m_a , m_b and m_c can only be put on the beam at five spots each, at distances 0, 1 or 2, in units of $\delta_a/2$, $\delta_b/2$ and $\delta_c/2$, from the positions ℓ_a , ℓ_b and ℓ_c , respectively. The set of possible values for variables X_a , X_b and X_c (in their respective units) is thus $\{-2, -1, 0, 1, 2\}$. The table shows (in no particular order) all combinations of values of the triplet (X_a, X_b, X_c) for which the beam turns out to be balanced. We assume that the procedure followed to take pans of the beam and put them back on is such that all these combinations are all equiprobable. To keep things simple, we set $\delta_a = \delta_b = \delta_c = \delta$ and $m_a = m_b = m_c = m$, in which case the beam is balanced whenever $X_a + X_b + X_c = 0$.

the beam is balanced if and only if the triplet of variables (X_a, X_b, X_c) takes on one of the 18 triplets of values shown (in no particular order) in Table 3.2. In all 18 cases, these three numbers add up to 0, which, given that $\delta_a = \delta_b = \delta_c$, immediately tells us that the three masses must have the same value. We verify that this is also the upshot of a bivariate regression of the statistical data found in this simplified version of the experiment with the $3m$ balance. This will be true as long as all 18 triplets of values in Table 3.2 are equiprobable.

The most natural protocol for the experiment does, in fact, guarantee that these triplets will be equiprobable. In every run of the experiment, we begin by randomly choosing the order in which we are going to put the three pans back on the beam. For the first pan we can pick any one of the values 0, ± 1 and ± 2 that determines where we place it on the beam. However, if we pick, say, 2 for the first pan and then 1 or 2 for the second, we discover that there will be no

value for the third such that the beam is balanced. We discard all runs of the experiment in which we end up not being able to balance the beam. We also never simply put all three pans back at their original positions, i.e., we rule out picking 0 for all three variables. This leaves us with the 18 combinations of values for (X_a, X_b, X_c) in Table 3.2. A simple alternative protocol, readily seen to lead to the same result, is to randomly choose three (not necessarily different) numbers from the set $\{-2, -1, 0, 1, 2\}$, accept them if they add up to zero and reject them if they do not. Excluding $(0, 0, 0)$, we are, once again, left with the 18 combinations in Table 3.2. This alternative protocol (or direct inspection of Table 3.2) tells us that $X_a + X_b + X_c = 0$, which, given that $\delta_a = \delta_b = \delta_c = \delta$, once again tells us that $m_a = m_b = m_c$. We are now ready to check that bivariate regression reproduces this result.

First, we note that X_a, X_b and X_c are balanced. Given that their only possible values are 0, ± 1 and ± 2 , these variables clearly meet the first of the two conditions in the definition labeled (3.1.1) of a balanced random variable. They also meet the second condition. Using the relative frequency of an outcome in the representative sample in Table 3.2 as a measure of its probability (cf. note 5), we find for all three of these variables that

$$\Pr(2) = \Pr(-2) = \frac{1}{6}, \quad \Pr(1) = \Pr(-1) = \frac{2}{9}. \quad (3.4.8)$$

This means that their expectation values vanish ($\langle X_a \rangle = \langle X_b \rangle = \langle X_c \rangle = 0$) and that their variances are just the expectation value of their squares ($\langle X_a^2 \rangle, \langle X_b^2 \rangle$ and $\langle X_c^2 \rangle$). We can use the square averages $\overline{X_a^2}, \overline{X_b^2}$ and $\overline{X_c^2}$ in our sample as measures of these variances:

$$\langle X_a^2 \rangle = \langle X_b^2 \rangle = \langle X_c^2 \rangle = 4 \cdot \frac{1}{3} + 1 \cdot \frac{4}{9} = \frac{16}{9}. \quad (3.4.9)$$

The corresponding standard deviations are the square roots of these variances:

$$\sigma_a = \sigma_b = \sigma_c = \frac{4}{3}. \quad (3.4.10)$$

We can likewise use the average values of products of these variables in our sample as measures of their covariances. Looking at the first two columns in Table 3.2, we see that in 8 of the 18 entries either X_a or X_b equals zero. The contributions coming from the remaining 10 entries add up to

$$\overline{X_a X_b} = \frac{1}{18} \left(-2 - 1 - 2 - 4 + 1 - 1 - 4 - 2 - 2 + 1 \right) = -\frac{8}{9}. \quad (3.4.11)$$

Looking at other pairs of columns, we find the exact same results for $\overline{X_a X_c}$ and $\overline{X_b X_c}$. Hence:

$$\langle X_a X_b \rangle = \langle X_a X_c \rangle = \langle X_b X_c \rangle = -\frac{8}{9}. \quad (3.4.12)$$

Substituting these variances and covariances into Eq. (3.4.7), we find the following pair of equations

$$-\frac{8}{9} = -\left(\frac{m_a}{m_c}\right)\frac{16}{9} + \left(\frac{m_b}{m_c}\right)\frac{8}{9}, \quad -\frac{8}{9} = \left(\frac{m_a}{m_c}\right)\frac{8}{9} - \left(\frac{m_b}{m_c}\right)\frac{16}{9}. \quad (3.4.13)$$

It follows that $m_a/m_c = m_b/m_c = 1$. So bivariate regression of X_c on X_a and X_b does indeed confirm that the three masses are equal in this case.

Finally, we compute the Pearson correlation coefficients. Using Eqs. (3.4.10) and (3.4.12), we find

$$\begin{aligned} \rho_{ab} &\equiv \frac{\langle X_a X_b \rangle}{\sigma_a \sigma_b} = -\frac{1}{2}, \\ \rho_{ac} &\equiv \frac{\langle X_a X_c \rangle}{\sigma_a \sigma_c} = -\frac{1}{2}, \\ \rho_{bc} &\equiv \frac{\langle X_b X_c \rangle}{\sigma_b \sigma_c} = -\frac{1}{2}. \end{aligned} \quad (3.4.14)$$

These three coefficients satisfy the elliptope inequality in Eq. (3.1.15) with an equal sign, just as one would expect given that $X_a + X_b + X_c = 0$.

We went over this example of bivariate regression because regression theory is the context in which Yule derived the elliptope inequality, as the title of his 1897 paper makes clear: “On the significance of Bravais’ [or Pearson’s, cf. note 2] formulae for regression, &c., in the case of skew correlation.” Instead of following Yule’s algebraic approach, however, we switch to the geometric approaches of Fisher and De Finetti.

For our use of Fisher (1915, 1924), we rely on the paper mentioned at the beginning of this section by Aldrich (1998, sec. 10, pp. 72–74; see also Kendall, 1961, pp. 55–57). Think of the values for the random variables X_a , X_b and X_c found in runs $1, 2, \dots, N$ of our experiment with the $3m$ balance as the components of the n -dimensional vectors

$$\mathbf{X}_a^{(N)} \equiv (x_a^1, \dots, x_a^N), \quad \mathbf{X}_b^{(N)} \equiv (x_b^1, \dots, x_b^N), \quad \mathbf{X}_c^{(N)} \equiv (x_c^1, \dots, x_c^N). \quad (3.4.15)$$

We will call such vectors *representative sample vectors*. A sample vector is representative if its components form a balanced sample, i.e., if the sample meets two conditions: (1) if the value x occurs in the sample, then $-x$ also occurs; (2) both x and $-x$ occur with the same frequency (cf. the definition labeled (3.1.1) of a balanced variable).¹⁷

The standard dot product of $\mathbf{X}_a^{(N)}$ with itself gives n times the square average of X_a in the representative sample which we can use as a measure for the variance of X_a (cf. note 5):

$$\mathbf{X}_a^{(N)} \cdot \mathbf{X}_a^{(N)} = \sum_{k=1}^N (x_a^k)^2 = N \langle X_a^2 \rangle = N \sigma_a^2, \quad (3.4.16)$$

where σ_a is the standard deviation. Similar results hold for the dot products of $\mathbf{X}_b^{(N)}$ and $\mathbf{X}_c^{(N)}$ with themselves. Hence, the norms of these vectors are

$$\|\mathbf{X}_a^{(N)}\| = \sqrt{N} \sigma_a, \quad \|\mathbf{X}_b^{(N)}\| = \sqrt{N} \sigma_b, \quad \|\mathbf{X}_c^{(N)}\| = \sqrt{N} \sigma_c. \quad (3.4.17)$$

The dot product of $\mathbf{X}_a^{(N)}$ and $\mathbf{X}_b^{(N)}$ gives N times the covariance of X_a and X_b :

$$\mathbf{X}_a^{(N)} \cdot \mathbf{X}_b^{(N)} = \sum_{k=1}^N x_a^k x_b^k = N \langle X_a X_b \rangle. \quad (3.4.18)$$

We can use these dot products to define angles between sample vectors. The cosines of these angles are just the Pearson correlation coefficients in Eq. (3.1.9) for the balanced variables X_a and X_b . We verify this for the cosine of the angle ϑ_{ab} between $\mathbf{X}_a^{(N)}$ and $\mathbf{X}_b^{(N)}$ (cf. Eq. (3.4.14)):

$$\cos \vartheta_{ab} = \frac{\mathbf{X}_a^{(N)} \cdot \mathbf{X}_b^{(N)}}{\|\mathbf{X}_a^{(N)}\| \|\mathbf{X}_b^{(N)}\|} = \frac{N \langle X_a X_b \rangle}{\sqrt{N} \sigma_a \sqrt{N} \sigma_b} = \frac{\langle X_a X_b \rangle}{\sigma_a \sigma_b} = \rho_{ab}. \quad (3.4.19)$$

Suppressing the superscripts (N), we now decompose sample vectors \mathbf{X}_a and \mathbf{X}_b into components parallel and perpendicular to sample vector \mathbf{X}_c :

¹⁷ The sample in Table 3.2 is an example of representative sample (with $N = 18$). Here is a general method for construction a representative sample vector of dimension $N = 2M$ if we only perform a relatively small number M runs of the experiment. For the first M components of the sample vectors ($\mathbf{X}_a^{(N)}, \mathbf{X}_b^{(N)}, \mathbf{X}_c^{(N)}$), we use the values for X_a, X_b and X_c found in these runs; for the last M components, we use *minus* these values. This construction guarantees that the samples from which the sample vectors are constructed are balanced.

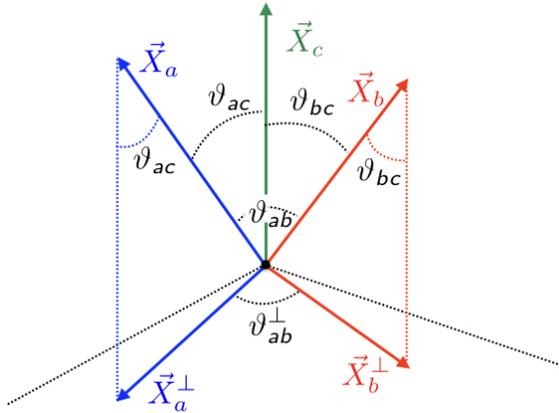


Fig. 3.6 Vectors representing balanced samples of triplets of values of the random variables (X_a, X_b, X_c) .

$$\mathbf{X}_a = \mathbf{X}_a^{\parallel} + \mathbf{X}_a^{\perp}, \quad \mathbf{X}_b = \mathbf{X}_b^{\parallel} + \mathbf{X}_b^{\perp} \quad (3.4.20)$$

This decomposition is illustrated in Figure 3.6. The parallel components can be seen as measures of the correlation between the random variables X_a and X_c and the correlation between X_b and X_c ; the perpendicular components as measures of the so-called partial or residual correlation between X_a and X_b (Aldrich, 1998, p. 73).¹⁸ As Fisher put it (translated into our notation):

[T]he correlation between $[X_a]$ and $[X_b]$. . . will be the cosine of the angle between $[\mathbf{X}_a]$ and $[\mathbf{X}_b]$. . . [T]he partial correlation between $[X_a]$ and $[X_b]$ is the cosine of the angle between the projections of $[\mathbf{X}_a]$ and $[\mathbf{X}_b]$ upon the region perpendicular to $[X_c]$ (Fisher, 1924, pp. 329–330).

The lengths of the parallel components of \mathbf{X}_a and \mathbf{X}_b are

$$\begin{aligned} \|\mathbf{X}_a^{\parallel}\| &= \|\mathbf{X}_a\| \cos \vartheta_{ac} = \sqrt{N} \sigma_a \cos \vartheta_{ac}, \\ \|\mathbf{X}_b^{\parallel}\| &= \|\mathbf{X}_b\| \cos \vartheta_{bc} = \sqrt{N} \sigma_b \cos \vartheta_{bc}, \end{aligned} \quad (3.4.21)$$

¹⁸ As we saw in Section 2.6, the tastes two bananas, one peeled by Alice in the a -direction, the other peeled by Bob in the b -direction, are completely *uncorrelated* if the peeling directions \mathbf{e}_a and \mathbf{e}_b are orthogonal. It follows that the geometry of these sample vectors differs from the geometry of state vectors in Hilbert space in quantum mechanics. We will return to this point in the context of our discussion of De Finetti (1937) below.

where we used Eq. (3.4.17) for $\|\mathbf{X}_a\|$ and $\|\mathbf{X}_b\|$. Similarly, the lengths of the perpendicular components of \mathbf{X}_a and \mathbf{X}_b are

$$\begin{aligned}\|\mathbf{X}_a^\perp\| &= \|\mathbf{X}_a\| \sin \vartheta_{ac} = \sqrt{N} \sigma_a \sin \vartheta_{ac}, \\ \|\mathbf{X}_b^\perp\| &= \|\mathbf{X}_b\| \sin \vartheta_{bc} = \sqrt{N} \sigma_b \sin \vartheta_{bc},\end{aligned}\tag{3.4.22}$$

We can rewrite the dot product in Eq. (3.4.19) as

$$\cos \vartheta_{ab} = \frac{(\mathbf{X}_a^\parallel + \mathbf{X}_a^\perp) \cdot (\mathbf{X}_b^\parallel + \mathbf{X}_b^\perp)}{N \sigma_a \sigma_b} = \frac{\mathbf{X}_a^\parallel \cdot \mathbf{X}_b^\parallel + \mathbf{X}_a^\perp \cdot \mathbf{X}_b^\perp}{N \sigma_a \sigma_b}.\tag{3.4.23}$$

With the help of Eq. (3.4.21), the parallel components of \mathbf{X}_a and \mathbf{X}_b can be written as

$$\begin{aligned}\mathbf{X}_a^\parallel &= \|\mathbf{X}_a\| \cos \vartheta_{ac} \frac{\mathbf{X}_c}{\|\mathbf{X}_c\|} = \frac{\sigma_a}{\sigma_c} \cos \vartheta_{ac} \mathbf{X}_c, \\ \mathbf{X}_b^\parallel &= \|\mathbf{X}_b\| \cos \vartheta_{bc} \frac{\mathbf{X}_c}{\|\mathbf{X}_c\|} = \frac{\sigma_b}{\sigma_c} \cos \vartheta_{bc} \mathbf{X}_c.\end{aligned}\tag{3.4.24}$$

Using these expressions along with $\mathbf{X}_c \cdot \mathbf{X}_c = \|\mathbf{X}_c\|^2 = N \sigma_c^2$, we can write the dot product of the parallel components of \mathbf{X}_a and \mathbf{X}_b as

$$\mathbf{X}_a^\parallel \cdot \mathbf{X}_b^\parallel = N \sigma_a \sigma_b \cos \vartheta_{ac} \cos \vartheta_{bc}.\tag{3.4.25}$$

The dot product of the perpendicular components can be written as

$$\mathbf{X}_a^\perp \cdot \mathbf{X}_b^\perp = \|\mathbf{X}_a^\perp\| \|\mathbf{X}_b^\perp\| \cos \vartheta_{ab}^\perp = N \sigma_a \sigma_b \sin \vartheta_{ac} \sin \vartheta_{bc} \cos \vartheta_{ab}^\perp,\tag{3.4.26}$$

where we used Eq. (3.4.22). Inserting Eqs. (3.4.25)–(3.4.26) into Eq. (3.4.23), we arrive at

$$\cos \vartheta_{ab} = \cos \vartheta_{ac} \cos \vartheta_{bc} + \sin \vartheta_{ac} \sin \vartheta_{bc} \cos \vartheta_{ab}^\perp.\tag{3.4.27}$$

Solving for $\cos \vartheta_{ab}^\perp$, we find that¹⁹

¹⁹ If all vectors in Figure 3.6 are turned into unit vectors so that their tips all lie on a unit sphere, one easily recognizes that Eq. (3.4.28) is nothing but the spherical law of cosines. This illustrates the observation by Pearson quoted at the beginning of this section. The relation between the “full” and the partial correlation between X_a and X_b is thus fairly trivial. Borrowing a comment by Aldrich (1998, p. 73) on a closely related aspect of Fisher’s

$$\cos \vartheta_{ab}^{\perp} = \frac{\cos \vartheta_{ab} - \cos \vartheta_{ac} \cos \vartheta_{bc}}{\sin \vartheta_{ac} \sin \vartheta_{bc}}. \quad (3.4.28)$$

If ρ 's are substituted for cosines of ϑ 's and expressions of the form $\sqrt{(1 - \rho^2)}$ for sines of ϑ , this turns into:

$$\cos \vartheta_{ab}^{\perp} = \frac{\rho_{ab} - \rho_{ac}\rho_{bc}}{\sqrt{(1 - \rho_{ac}^2)}\sqrt{(1 - \rho_{bc}^2)}}. \quad (3.4.29)$$

The quantity on the right-hand side can be found in Yule (1897, p. 485), who denotes it as ρ_{12} and calls it the “net coefficient of correlation between x_1 and x_2 ” (X_a and X_b in our notation; Yule's x_3 , likewise, is X_c in our notation). Squaring both sides of Eq. (3.4.29) we find:

$$\cos^2 \vartheta_{ab}^{\perp} = \frac{(\rho_{ab} - \rho_{ac}\rho_{bc})^2}{(1 - \rho_{ac}^2)(1 - \rho_{bc}^2)}. \quad (3.4.30)$$

Since $\cos^2 \vartheta_{ab}^{\perp} \leq 1$, we have the inequality (Yule, 1897, p. 486)

$$(\rho_{ab} - \rho_{ac}\rho_{bc})^2 \leq (1 - \rho_{ac}^2)(1 - \rho_{bc}^2); \quad (3.4.31)$$

or, expanding both sides,

$$\rho_{ab}^2 - 2\rho_{ab}\rho_{ac}\rho_{bc} + \rho_{ac}^2\rho_{bc}^2 \leq 1 - \rho_{ac}^2 - \rho_{bc}^2 + \rho_{ac}^2\rho_{bc}^2. \quad (3.4.32)$$

The terms $\rho_{ac}^2\rho_{bc}^2$ before and after the ‘ \leq ’ cancel. Reordering the remaining terms, we see that this is just the ellipptope inequality in Eq. (3.1.15) for the Pearson correlation coefficients for the variables X_a , X_b and X_c .

The construction above with sample vectors is problematic as probabilities only coincide with relative frequencies in the $N \rightarrow \infty$ limit. To remedy this, we consider the random variables themselves as vectors. This is actually a common approach in modern probability theory (see, e.g., Fristedt & Gray, 1997). An early instance of it can be found in a paper by De Finetti (1937).²⁰

De Finetti begins his geometric interpretation by observing that “since we can consider linear combinations of random variables, we can interpret them

analysis, we can say that Fisher's derivation of this formula for the partial correlation amounts to “something of a self-annihilating insight.”

²⁰ We used the translation by Luca Barone and Peter Laurence, who also co-authored a commentary on the paper (Laurence, Hwang, & Barone, 2008).

as vectors in an ‘abstract space’” (De Finetti, 1937, p. 5).²¹ That is to say, linear combinations of random variables are also random variables and so form a vector space. Standard deviations can be interpreted as the norm of vectors in this space (cf. Eqs. (3.4.16)–(3.4.17) for Fisher’s sample vectors),

$$\|X\| \equiv \sqrt{\langle X^2 \rangle}. \quad (3.4.33)$$

The distance (or metric) between two random variables X and Y can then be defined as:²²

$$d(X, Y) \equiv \|X - Y\|. \quad (3.4.34)$$

Similarly, covariances can be interpreted as inner products of two random variables in this vector space (p. 6; cf. Eq. (3.4.18) for Fisher’s sample vectors):²³

$$\langle X, Y \rangle \equiv \langle XY \rangle. \quad (3.4.35)$$

Hence De Finetti deduces that this vector space of random variables is not only a *metric space* but also a (normed) *inner product space*. One can go further and complete this space with respect to the above metric, resulting in a Hilbert space of random variables.

De Finetti in fact gives two such Hilbert space interpretations, one (on p. 6) where the inner product is the covariance $\langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle$, the other (on p. 8) where the inner product is $\langle XY \rangle$ (even if $\langle X \rangle$ and/or $\langle Y \rangle$ are non-zero). The first Hilbert space can be viewed as the subspace of the second, consisting of vectors orthogonal to all constant random variables, and so it is the second Hilbert space which is more common to the literature (see again Fristedt & Gray, 1997). For the purposes of this section, however, we only

²¹ Unless noted otherwise, page references in the remainder of this section are to De Finetti (1937)

²² One subtlety of this definition is that $d(X, Y) = 0$ when X and Y differ by a constant random variable (i.e., a random variable guaranteed to have the same value every time it is measured). Since we assume our random variables to have zero mean, this amounts to two random variables being equivalent if they are equal with probability 1, i.e., almost surely equal. This point is emphasized by De Finetti earlier in the paper, with him commenting that such ‘coinciding’ random variables are represented by the same vector. Since random variables that are almost surely equal agree in their expectation values, we will not distinguish them in our discussion of De Finetti’s paper (p. 5).

²³ The notation $\langle \cdot, \cdot \rangle$ is ours: De Finetti does not introduce a special notation for this inner product.

consider random variables with $\langle X \rangle = \langle Y \rangle = 0$ and so will not distinguish the two spaces.

De Finetti's purpose in his 1937 paper is not to define a Hilbert space but to use his vector space to bolster geometric intuition about random variables.²⁴ To this end, he uses his inner product to conclude—like Fisher (1915, 1924) before him but now treating the random variables themselves as vectors—that “the correlation coefficient is the cosine of the angle $\alpha(X, Y)$ between vectors X and Y ” (p. 6). He elaborates:

Zero correlation means orthogonality; positive or negative correlation means that α is acute or obtuse, respectively; for the extreme cases . . . $\alpha = 0$ and $\alpha = \pi$, respectively, the two vectors . . . only differ by a multiplicative constant, positive or negative, respectively (De Finetti, 1937, p. 6).

In our quantum banana peel-and-taste experiment the taste X_a^A found by Alice peeling \hat{a} is perfectly correlated to the taste X_b^B found by Bob peeling \hat{b} if the angle φ_{ab} between the unit vectors \mathbf{e}_a and \mathbf{e}_b for the corresponding peeling directions is 180° . As we noted in Section 2.6, the angle between the corresponding eigenvectors $|+\rangle_a$ and $|+\rangle_b$ in that case is 90° . De Finetti's angle α thus corresponds to the angle between the vectors \mathbf{e}_a and \mathbf{e}_b in ordinary space rather than to the angle between $|+\rangle_a$ and $|+\rangle_b$ in Hilbert space (note that the perfect correlation for $\varphi_{ab} = 180^\circ$ turns into a perfect anti-correlation once $-X_b^B$ is used as a proxy for X_b^A ; see Section 3.2). In De Finetti's formalism, we can thus think of (X_a^A/σ_a) as a vector that coincides with \mathbf{e}_a . This vector, in turn, is directly related to the operator for the spin component being measured, $\hat{S}_a \equiv \hat{S} \cdot \mathbf{e}_a$. Though we will not pursue this any further in this volume, we note that the Hilbert space encountered here is not the familiar Hilbert space of state vectors but a Hilbert space of (spin) operators (where the inner product is simply the expectation value of the product of the operators in the quantum state of the system).

Like Fisher (see Eq. (3.4.20)), De Finetti notes that “every random variable Y can be decomposed into two components, one correlated with X (the parallel

²⁴ While De Finetti does not consider issues of completeness in this paper, he cites a prior note of his (not on probability theory) in which he references Hilbert space and appears to directly address the issue of completeness (De Finetti, 1930, p. 248 and p. 254, respectively). It thus seems reasonable to assume that De Finetti understood perfectly well that his vector space of random variables is a Hilbert space. Joseph L. Doob (1934) appears to have been the first to use Hilbert space in the context of general probability theory. Applications to probability theory are not mentioned in the history of functional analysis by Birkhoff & Kreyszig (1984). A promising source for filling this gap in the literature on the history of mathematics is Bingham (2000).

component) and one uncorrelated with X (the orthogonal component)” (p. 6). Finally, De Finetti considers the angles between a triplet of random variables (X, Y, Z) .

If $\alpha(X, Y)$ is the angle between two random variables X and Y , a third random variable Z cannot form two arbitrary angles $\alpha(X, Z)$ and $\alpha(Y, Z)$, but we must have (obviously, if we think of the geometric picture) $\alpha(X, Y) \leq \alpha(X, Z) + \alpha(Y, Z) \leq 2\pi - \alpha(X, Y)$; we have the extreme case $\alpha(X, Z) + \alpha(Y, Z) = \alpha(X, Y)$ if and only if $Z = aX + bY$, $a > 0$, $b > 0$ (the coplanar vector, included in the concave angle between the two vectors), and the other $\alpha(X, Z) + \alpha(Y, Z) = 2\pi - \alpha(X, Y)$ if and only if $Z = -(aX + bY)$, $a > 0$, $b > 0$ (the aforementioned condition applied to minus the same vector). This shows that there are some constraints for the degrees of pairwise correlation among different random variables (De Finetti, 1937, p. 6).

These angle inequalities are exactly those in Eq. (2.6.24). De Finetti considers the following special case:

In particular, if three random variables are all equally correlated, since pairwise they are unable to form an angle bigger than $2\pi/3$ (i.e., 120°), the correlation coefficient cannot be smaller than $-1/2$ (De Finetti, 1937, pp. 6–7).

The observant reader will recognize the 120° case as the scenario where the Tsirelson bound is fulfilled. De Finetti identified this scenario more than forty years before either Cirel’son (1980) or Accardi & Fedullo (1982, see note 27 in Section 2.6) published on the subject!

Although it is beyond the scope of this volume to pursue the parallel any further, we hope that our brief coverage of some of the work of statisticians such as Pearson, Yule, Fisher and De Finetti on the geometry of correlations gives the reader at least a taste of the connection between this approach to probability theory and the Hilbert space formalism of quantum mechanics.



Chapter 4

Generalization to singlet state of two particles with higher spin

Quantum correlations for pairs of particles with higher spin in the singlet state • Designing raffles to simulate these quantum correlations • Classical polyhedra with more and more vertices and facets and getting closer and closer to the ellipsope.

In Chapter 2 we studied pairs of bananas that behave like pairs of spin- $\frac{1}{2}$ particles in the singlet state. This is just one type of banana in Bananaworld. There are many more. In Section 2.3, for instance, we came across Popescu-Rohrlich bananas. Bub (2016, Ch. 6) has made extensive study of more exotic species, such as Aravind-Mermin bananas and Klyachko bananas. One could add to this taxonomy by introducing pairs of bananas that behave like pairs of particles of arbitrary integer or half-integer spin $s \geq 1$ in the singlet state. For $s = 1$, bananas would taste yummy (+), nasty (-), or meh (0). As we move to higher spin, we would have to invent more refined taste palettes. The banana imagery thus starts to feel forced for higher spin and we will largely dispense with it in the remainder of this volume. Instead we will phrase our analysis directly in terms of spin. So rather than have Alice and Bob peel and taste pairs of bananas, we imagine them sending pairs of particles through Du Bois magnets (see Figure 6.1 and note 41 in Section 6.5) and measuring a component of their spin, choosing between three different directions, represented, as before, by unit vectors $(\mathbf{e}_a, \mathbf{e}_b, \mathbf{e}_c)$, corresponding to settings $(\hat{a}, \hat{b}, \hat{c})$. In Section 4.1 we present the quantum-mechanical analysis of the correlations found in these measurements, extending our discussion of the spin- $\frac{1}{2}$ case in Section 2.6 to higher-spin cases. In Section 4.2 we design raffles like the ones we introduced in Section 2.5 to simulate these correlations.

For the quantum-mechanical analysis we rely on the standard treatment of rotation in quantum mechanics (see, e.g., Messiah 1962, Vol. 2, Appendix C, or Baym 1969, Ch. 17). This will lead us to the so-called Wigner d-matrices (see Eq. (4.1.42)) with which we can readily compute the probabilities entering into the correlation arrays for measurements on the singlet state of two particles with arbitrary (half-) integer spin s . After showing how the results we found in Section 2.6 for the spin- $\frac{1}{2}$ case are recovered (and justified) in this more general formalism, we use it to find the entries of a typical cell in the correlation array

for the spin-1 and spin- $\frac{3}{2}$ cases (see Figures 4.2 and 4.3). We prove that the correlation arrays for higher-spin cases share some key properties with those for the spin- $\frac{1}{2}$ and spin-1 cases (Sections 4.1.1–4.1.3).

In Section 4.1.4, we show that all such correlation arrays have uniform marginals and are therefore non-signaling. In Section 4.1.5, we show that they can still be parametrized by the anti-correlation coefficients for three of their off-diagonal cells and that these coefficients are still given by the cosines of the angles between measuring directions,

$$\chi_{ab} = \cos \varphi_{ab}, \quad \chi_{ac} = \cos \varphi_{ac}, \quad \chi_{bc} = \cos \varphi_{bc}, \quad (4.0.1)$$

and subject to the same constraint we found in the spin- $\frac{1}{2}$ case (see Eq. (2.6.9)):

$$1 - \chi_{ab}^2 - \chi_{ac}^2 - \chi_{bc}^2 + 2\chi_{ab}\chi_{ac}\chi_{bc} \geq 0. \quad (4.0.2)$$

It follows that the class of correlations allowed by quantum mechanics in measurements on the singlet state of two particles with (half-)integer spin s can be represented by the ellipptope in Figure 2.16 regardless of the value of s .

In Section 4.1.6, we turn to a property of the correlation arrays for these higher-spin cases we did not pay much attention to in the spin- $\frac{1}{2}$ case: the symmetries of their cells. We show that, for arbitrary (half-)integer s , any cell in the correlation array for measurements on the singlet state of particles with (half-)integer spin s is *centrosymmetric*, *symmetric* and *persymmetric*, i.e., it is unchanged under reflection about its center, across its main diagonal and across its main anti-diagonal (see Eq. (4.1.85) for what this means in terms of the probabilities that form the entries of such cells).

In Section 4.2, we design raffles to simulate these quantum correlations. In preparation for this, we formalize the description and analysis of the raffles we used in Section 2.5 for the spin- $\frac{1}{2}$ case (Section 4.2.1). In Sections 4.2.2–4.2.4, we adapt the formalism developed for this case to design raffles that simulate—to the extent that this is at all possible—the correlation arrays for measurements on the singlet state of two particles with higher spin. In doing so, we run into four main complications compared to the spin- $\frac{1}{2}$ case.

First, as we already noted in Section 3.3, we can no longer admit single-ticket raffles. Our raffle tickets only have two outcomes per setting. If there are more than two possible outcomes (and for spin s there are $2s + 1$), it is therefore impossible to have diagonal cells of the form shown in Figure 3.4. Such single-ticket raffles not only fail to reproduce the quantum correlations; without identical diagonal cells, the anti-correlation coefficients for the off-

diagonal cells no longer suffice to characterize the correlations generated by these raffles. To get around this problem we need to restrict ourselves to mixed raffles that give uniform marginals. These consist of combinations of tickets such that all $2s + 1$ outcomes occur with the same frequency. This will guarantee that the diagonal cells in the resulting correlation arrays all have the form of the cell in Figure 3.4. Since our raffles are non-signaling by construction, this ensures uniform marginals for the entire correlation array.

The second complication has to do with the relationship between probabilities and expectation values (or anti-correlation coefficients). In the spin- $\frac{1}{2}$ case, the probabilities in any cell of the correlation array could be parametrized by the anti-correlation coefficient for that cell (see Figures 2.4 and 2.7 in Section 2.4). In the quantum correlation arrays, as noted above, this remains true for arbitrary (half-) integer spin s . As soon as there are more than two possible outcomes, however, it fails for our raffles. This severely limits our ability to simulate the quantum correlation arrays for higher-spin cases. We can design mixed raffles that simulate the diagonal cells of the quantum correlation arrays and that give the correct values for the *anti-correlation coefficients* of the off-diagonal cells; yet these raffles will, in general, still not give the right values for the *probabilities* in the off-diagonal cells. In Section 4.2.2, we will encounter a striking example of this complication. We will construct a raffle for the spin-1 case for which the sum of the anti-correlation coefficients is $-3/2$, the Tsirelson bound for this setup (see Eq. (3.3.9)). Yet the off-diagonal cells of the correlation array for this raffle are different from those in the quantum correlation array it was meant to simulate (see Eq. (4.2.23)).

The third (less serious) complication has to do with the symmetries of the off-diagonal cells in the correlation array. The design of our raffles guarantees that all cells in their correlation arrays are centrosymmetric. The condition for centrosymmetry of such cells, say the $\hat{a}\hat{b}$ one, is (see Eq. (4.1.85)):

$$\Pr(m_1 m_2 | \hat{a} \hat{b}) = \Pr(-m_1 - m_2 | \hat{a} \hat{b}). \quad (4.0.3)$$

This condition is automatically satisfied by our raffles: it simply expresses that Alice and Bob are as likely to get one side of any ticket as the other. The entries in cells of correlation arrays in the spin- $\frac{1}{2}$ case form 2×2 matrices. In that case, centrosymmetry trivially implies both symmetry and persymmetry. For the $(2s + 1) \times (2s + 1)$ matrices formed by the entries in cells of correlation arrays in the spin- s case with $s \geq 1$, this is no longer true (although any two of these symmetries still imply the third). Cells in the quantum correlation arrays, as noted above, have all three symmetries, regardless of the spin of the particles

in the singlet state on which Alice and Bob perform their measurements. To correctly simulate this feature of the quantum correlations we thus need to impose additional symmetry conditions on our raffles. Fortunately, this can be done without too much trouble.

The fourth complication is perhaps the most obvious one. As the number of outcomes increases, so does the number of different ticket types in our raffles. Figure 4.6 shows the $(3^3 + 1)/2 = 14$ different ticket types for raffles in the spin-1 case. Figures 4.14 and 4.15 show some of the $4^3/2 = 32$ different ticket types for the spin- $\frac{3}{2}$ case. In dealing with these higher-spin cases, we therefore turned to the computer for guidance.

As in the spin- $\frac{1}{2}$ case, we will represent the class of triplets of anti-correlation coefficients $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ for *admissible raffles* in the spin- s case (i.e., raffles that give uniform marginals and meet the symmetry requirements) by a polyhedron in the same 3-dimensional non-signaling cube as before. Henceforth, we will call this the *anti-correlation polyhedron*. In the spin- $\frac{1}{2}$ case, the anti-correlation polyhedron doubles as the local polytope (see Figures 2.5 and 2.8). In the spin- s case (with $s \geq 1$), the anti-correlation polyhedron is a particular (highly informative) projection of (a restricted version of) a now higher-dimensional local polytope (restricted by our admissibility conditions) to three dimensions (cf. the discussion above about complications in the relationship between probabilities and anti-correlation coefficients). The flowchart in Figure 4.7 shows how we get from the local polytope to the anti-correlation polyhedron in the spin-1 case. With considerable help from the computer, we were able to construct anti-correlation polyhedra for $s = 1, \frac{3}{2}, 2, \frac{5}{2}$ (see Figures 4.11, 4.13 and 4.17).

We pay special attention to admissible raffles for which the sum $\chi_{ab} + \chi_{ac} + \chi_{bc}$ takes on its minimum value (see Eqs. (3.3.8) and (3.3.9) in Section 3.3). In constructing these raffles, we take advantage of the insight that they will involve tickets for which the sum of the outcomes on both sides is either zero (for integer spin) or $\pm 1/2$ (for half-integer spin) (see Figures 4.9, 4.14, 4.15 and Table 4.4 for the tickets in such raffles). Comparing the anti-correlation polyhedra for higher spin values s to the tetrahedron for $s = 1/2$, we see that these polyhedra get closer and closer to the ellipsope as the number of outcomes $2s + 1$ increases (see Figure 4.18).¹ This is just what we would expect given what we learned in Section 3.3, viz. that Eq. (4.0.2) for the ellipsope determines the broadest conceivable class of triplets of (anti-)correlation coefficients.

¹ We have not been able to construct a formal proof of this convergence (see note 11).

4.1 The quantum correlations

4.1.1 Quantum formalism for one spin- s particle

We review the formalism for spin angular momentum of particles of arbitrary integer or half-integer spin s , starting with the one-particle case.² The state of a spin- s particle with component $m\hbar$ in the z -direction is represented by a state vector $|s, m\rangle_z$ in a one-particle Hilbert space (where $\hbar \equiv h/2\pi$ and h is Planck's constant). These vectors are simultaneous eigenvectors of the Hermitian operators

$$\hat{S}_z, \quad \hat{S}^2 \equiv \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2, \quad (4.1.1)$$

with eigenvalues $m\hbar$ (with $m \in \{-s, \dots, s\}$) and $s(s+1)\hbar^2$, respectively:

$$\hat{S}_z |s, m\rangle_z = m\hbar |s, m\rangle_z, \quad \hat{S}^2 |s, m\rangle_z = s(s+1)\hbar^2 |s, m\rangle_z. \quad (4.1.2)$$

We will follow the common practice of setting $\hbar = 1$. The operators \hat{S}_x , \hat{S}_y and \hat{S}_z can be thought of as components of a vector

$$\hat{\mathbf{S}} \equiv (\hat{S}_x, \hat{S}_y, \hat{S}_z). \quad (4.1.3)$$

The operator \hat{S}^2 represents the length of this vector.

The simultaneous eigenvectors of \hat{S}_z and \hat{S}^2 form an orthonormal basis of the $(2s+1)$ -dimensional one-particle Hilbert space,

$$\{|s, m\rangle_z\}_{m=-s}^s \quad \text{with} \quad {}_z\langle s, m | s, m'\rangle_z = \delta_{mm'}, \quad (4.1.4)$$

where $\delta_{mm'}$ is the Kronecker delta ($\delta_{mm'} = 1$ if $m = m'$ and $\delta_{mm'} = 0$ if $m \neq m'$). The operators \hat{S}_x , \hat{S}_y and \hat{S}_z satisfy the commutation relations

$$[\hat{S}_x, \hat{S}_y] = i\hat{S}_z, \quad [\hat{S}_y, \hat{S}_z] = i\hat{S}_x, \quad [\hat{S}_z, \hat{S}_x] = i\hat{S}_y. \quad (4.1.5)$$

We will also use the raising/lowering operators $\hat{S}_\pm \equiv \hat{S}_x \pm i\hat{S}_y$. The action of \hat{S}_\pm on one of the orthonormal basis vectors $|s, m\rangle_z$ is given by

$$\hat{S}_\pm |s, m\rangle_z = C_\pm(s, m) |s, m \pm 1\rangle_z. \quad (4.1.6)$$

² Our presentation roughly follows the treatment of angular momentum in [Messiah \(1962, Vol. 2, Appendix C\)](#).

Using that $|s, m\rangle_z$ and $|s, m \pm 1\rangle_z$ are unit vectors, we can find expressions for the constants $C_{\pm}(s, m)$. We do this for C_+ . We start from

$$C_+(s, m)^2 = C_+(s, m)^2 {}_z\langle s, m+1 | s, m+1 \rangle_z = {}_z\langle s, m | \hat{S}_+^\dagger \hat{S}_+ | s, m \rangle_z. \quad (4.1.7)$$

Since

$$\hat{S}_+^\dagger \hat{S}_+ = (\hat{S}_x - i\hat{S}_y)(\hat{S}_x + i\hat{S}_y) = \hat{S}_x^2 + \hat{S}_y^2 + i[\hat{S}_x, \hat{S}_y] = \hat{S}^2 - \hat{S}_z^2 - \hat{S}_z, \quad (4.1.8)$$

we can rewrite the right-hand side as

$${}_z\langle s, m | (\hat{S}^2 - \hat{S}_z^2 - \hat{S}_z) | s, m \rangle_z = (s(s+1) - m(m+1)) {}_z\langle s, m | s, m \rangle_z. \quad (4.1.9)$$

Choosing $C_+(s, m)$ to be positive and real (the so-called Condon-Shortley convention³), we conclude that

$$C_+(s, m) = \sqrt{s(s+1) - m(m+1)}. \quad (4.1.10)$$

A corollary of this phase convention is that the operator $i\hat{S}_y = \frac{1}{2}\hat{S}_+ - \frac{1}{2}\hat{S}_-$ has real matrix elements in the $|s, m\rangle_z$ basis.

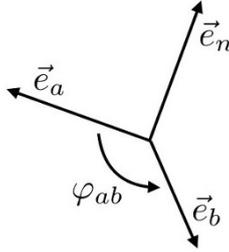


Fig. 4.1 Rotation by an angle φ_{ab} about the direction \mathbf{e}_n , mapping \mathbf{e}_a to \mathbf{e}_b .

As we already saw in Chapter 3, one can also define spin operators associated with directions other than the Cartesian axes. The spin operator associated with the direction given by the unit vector $\mathbf{e}_a = (a_x, a_y, a_z)$ is defined as

$$\hat{S}_a \equiv \hat{\mathbf{S}} \cdot \mathbf{e}_a = \hat{S}_x a_x + \hat{S}_y a_y + \hat{S}_z a_z. \quad (4.1.11)$$

³ Named for the authors of an influential text on spectroscopy (Condon & Shortley, 1935).

Such spin operators generate rotations in Hilbert space. Let \hat{S}_a and \hat{S}_b be spin operators associated with the directions given by the unit vectors \mathbf{e}_a and \mathbf{e}_b . Let \mathbf{e}_n be a unit vector in the direction of the cross product $\mathbf{e}_a \times \mathbf{e}_b$, so that we get from \mathbf{e}_a to \mathbf{e}_b by rotating around \mathbf{e}_n by the angle φ_{ab} between them (see Figure 4.1). This rotation is implemented in Hilbert space by the rotation operator $e^{-i\varphi_{ab}\hat{S}_n}$. It can be shown that the transformation from the spin operator \hat{S}_a associated with \mathbf{e}_a to the spin operator \hat{S}_b associated with \mathbf{e}_b is given by:

$$\hat{S}_b = e^{-i\varphi_{ab}\hat{S}_n}\hat{S}_a e^{i\varphi_{ab}\hat{S}_n} \quad (4.1.12)$$

(Messiah, 1962, Vol. 2, pp. 530–533; see also Baym 1969, pp. 305–307). The transformation of the eigenvectors of \hat{S}_a to the eigenvectors of \hat{S}_b is accordingly given by:⁴

$$|s, m\rangle_b = e^{-i\varphi_{ab}\hat{S}_n}|s, m\rangle_a. \quad (4.1.13)$$

We now show that, in the special case that $s = 1/2$ and \mathbf{e}_a to \mathbf{e}_b are in the xz -plane, the transformation law in Eq. (4.1.13) reduces to the one in Eqs. (2.6.39)–(2.6.40). For $s = 1/2$, the elements of the matrices representing the spin operators \hat{S}_x , \hat{S}_y and \hat{S}_z in the orthonormal basis in Eq. (4.1.4) of eigenvectors of \hat{S}_z are $1/2$ times the Pauli matrices (as long as the Condon-Shortley convention mentioned above is adopted):

$$\begin{aligned} \hat{S}_x &= 1/2 \hat{\sigma}_x = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \hat{S}_y &= 1/2 \hat{\sigma}_y = \frac{1}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \hat{S}_z &= 1/2 \hat{\sigma}_z = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned} \quad (4.1.14)$$

⁴ Whereas it takes some effort to prove the transformation law in Eq. (4.1.12), it is easy to see that this transformation law for operators entails the one for state vectors in Eq. (4.1.13). Using Eq. (4.1.12), we can show that, whenever $|s, m\rangle_a$ is an eigenvector of \hat{S}_a with eigenvalue m , $|s, m\rangle_b$ in Eq. (4.1.13) is, in fact, an eigenvector of \hat{S}_b with that same eigenvalue, as the notation suggests:

$$\begin{aligned} \hat{S}_b |s, m\rangle_b &= \left(e^{-i\varphi_{ab}\hat{S}_n} \hat{S}_a e^{i\varphi_{ab}\hat{S}_n} \right) \left(e^{-i\varphi_{ab}\hat{S}_n} |s, m\rangle_a \right) \\ &= e^{-i\varphi_{ab}\hat{S}_n} \hat{S}_a |s, m\rangle_a = e^{-i\varphi_{ab}\hat{S}_n} m |s, m\rangle_a = m |s, m\rangle_b. \end{aligned}$$

One readily verifies that these matrices satisfy the commutation relations in Eq. (4.1.5).

For $s = 1/2$, the rotation operator $e^{-i\vartheta\hat{S}_n}$ has a particularly simple form, which we can find using the completeness of the orthonormal basis of eigenvectors

$$\{|^{1/2}, ^{1/2}\rangle_n, |^{1/2}, -^{1/2}\rangle_n\} \quad (4.1.15)$$

of the spin operator \hat{S}_n associated with the unit vector \mathbf{e}_n . For the purposes of this calculation, we revert to the notation $|\pm\rangle_n$ of Section 2.6 for these eigenvectors. With the help of the resolution of unity in the basis of these eigenvectors,

$$\hat{1} = |+\rangle_n \langle +| + |-\rangle_n \langle -|, \quad (4.1.16)$$

we can write the spectral decomposition of the spin operator \hat{S}_n as

$$\hat{S}_n = \frac{1}{2} |+\rangle_n \langle +| - \frac{1}{2} |-\rangle_n \langle -|, \quad (4.1.17)$$

and the spectral decomposition of the rotation operator $e^{-i\vartheta\hat{S}_n}$ as

$$e^{-i\vartheta\hat{S}_n} = e^{-i\vartheta/2} |+\rangle_n \langle +| + e^{i\vartheta/2} |-\rangle_n \langle -|. \quad (4.1.18)$$

Collecting terms with cosines and sines and using Eqs. (4.1.16)–(4.1.17), we find that

$$\begin{aligned} e^{-i\vartheta\hat{S}_n} &= \left(\cos(\vartheta/2) - i \sin(\vartheta/2) \right) |+\rangle_n \langle +| \\ &\quad + \left(\cos(\vartheta/2) + i \sin(\vartheta/2) \right) |-\rangle_n \langle -| \\ &= \cos(\vartheta/2) \hat{1} - i \sin(\vartheta/2) 2\hat{S}_n. \end{aligned} \quad (4.1.19)$$

Note that we got from the angle ϑ to the half-angle $\vartheta/2$ because of the factor of $1/2$ in the eigenvalues $\hbar/2$ of the spin operator \hat{S}_n .

To recover Eqs. (2.6.39)–(2.6.40), we choose \mathbf{e}_n in the y -direction and \mathbf{e}_a in the z -direction. Hence, $\hat{S}_n = \hat{S}_y$ and $\hat{S}_z = \hat{S}_a$. In that case, as we saw in Eq. (4.1.14), the matrix representing $2\hat{S}_y$ in the basis $\{|+\rangle_a, |-\rangle_a\}$ is the Pauli matrix $\hat{\sigma}_y$. The matrix elements representing the rotation operator $e^{-i\vartheta\hat{S}_n}$ in this basis is then given by:

$$\begin{aligned} \cos\left(\frac{\vartheta}{2}\right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - i \sin\left(\frac{\vartheta}{2}\right) \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ = \begin{pmatrix} \cos(\vartheta/2) & -\sin(\vartheta/2) \\ \sin(\vartheta/2) & \cos(\vartheta/2) \end{pmatrix}. \end{aligned} \quad (4.1.20)$$

From this matrix we can read off the components of the eigenvectors of \hat{S}_b associated with a unit vector \mathbf{e}_b in the xz -plane obtained by rotating the eigenvectors of \hat{S}_a associated with a unit vector $\mathbf{e}_a = \mathbf{e}_z$ (cf. Eq. (4.1.13)):

$$|\pm\rangle_b = e^{-i\varphi_{ab}\hat{S}_y} |\pm\rangle_a. \quad (4.1.21)$$

Using the resolution of unity in the $\{|\pm\rangle_a\}$ basis, we find:

$$\begin{aligned} |\pm\rangle_b &= (|+\rangle_a \langle +| + |-\rangle_a \langle -|) e^{-i\varphi_{ab}\hat{S}_y} |\pm\rangle_a \\ &= \left(\langle +| e^{-i\varphi_{ab}\hat{S}_y} |\pm\rangle_a \right) |+\rangle_a + \left(\langle -| e^{-i\varphi_{ab}\hat{S}_y} |\pm\rangle_a \right) |-\rangle_a. \end{aligned} \quad (4.1.22)$$

Using Eq. (4.1.20) with $\vartheta = \varphi_{ab}$ for the matrix elements of the rotation operator in the $\{|\pm\rangle_a\}$ basis,

$${}_a\langle \pm | e^{-i\varphi_{ab}\hat{S}_y} |\pm\rangle_a = \begin{pmatrix} \cos(\varphi_{ab}/2) & -\sin(\varphi_{ab}/2) \\ \sin(\varphi_{ab}/2) & \cos(\varphi_{ab}/2) \end{pmatrix}, \quad (4.1.23)$$

we arrive at

$$\begin{aligned} |+\rangle_b &= \cos\left(\frac{\varphi_{ab}}{2}\right) |+\rangle_a + \sin\left(\frac{\varphi_{ab}}{2}\right) |-\rangle_a, \\ |-\rangle_b &= -\sin\left(\frac{\varphi_{ab}}{2}\right) |+\rangle_a + \cos\left(\frac{\varphi_{ab}}{2}\right) |-\rangle_a. \end{aligned} \quad (4.1.24)$$

This is just the inverse of the transformation from $|\pm\rangle_b$ to $|\pm\rangle_a$ in Eqs. (2.6.39)–(2.6.40). We used those equations to show that the singlet state for a pair of spin- $\frac{1}{2}$ particles has the same form in orthonormal bases $\{|\pm\rangle_a\}$ and $\{|\pm\rangle_b\}$ (see Eqs. (2.6.38)–(2.6.41)). We concluded that this singlet state has the same form in *any* orthonormal basis $\{|\pm\rangle_n\}$. Strictly speaking, our derivation of Eq. (4.1.24) only allows us to claim this for orthonormal bases of eigenvectors of spin operators associated with unit vectors in the xz -plane. However, since for

any two unit vectors \mathbf{e}_a and \mathbf{e}_b , we can choose the plane spanned by those two vectors to be the xz plane, this result holds for *any* orthonormal basis $\{|\pm\rangle_n\}$.

We now turn to the special case $s = 0$. Eq. (4.1.2) tells us that $\hat{S}^2|0, m\rangle_z = 0$. Using that $\hat{S}^2 = \hat{S}_y^2 + \hat{S}_x^2 + \hat{S}_z^2$, we thus have

$$\begin{aligned} 0 &= {}_z\langle 0, m | \hat{S}_x^2 | 0, m \rangle_z + {}_z\langle 0, m | \hat{S}_y^2 | 0, m \rangle_z + {}_z\langle 0, m | \hat{S}_z^2 | 0, m \rangle_z \\ &= |\hat{S}_x | 0, m \rangle_z|^2 + |\hat{S}_y | 0, m \rangle_z|^2 + |\hat{S}_z | 0, m \rangle_z|^2. \end{aligned} \quad (4.1.25)$$

This last expression, being a sum of squared absolute values, can only vanish if

$$\hat{S}_x | 0, m \rangle_z = \hat{S}_y | 0, m \rangle_z = \hat{S}_z | 0, m \rangle_z = 0. \quad (4.1.26)$$

This enforces that $m = 0$ (see Eq. (4.1.5)). More generally, it means that the singlet state has zero spin angular momentum along any direction. Hence we must have $\hat{S}_n | 0, 0 \rangle_z = 0$. This implies that the singlet state is invariant under rotation through an angle ϑ with respect to any direction \mathbf{e}_n ,

$$e^{-i\vartheta \hat{S}_n} | 0, 0 \rangle_z = \sum_{k=0}^{\infty} \frac{(-i\vartheta)^k}{k!} \hat{S}_n^k | 0, 0 \rangle_z = | 0, 0 \rangle_z, \quad (4.1.27)$$

as the only contribution to the sum comes from the $k = 0$ term.

4.1.2 Quantum formalism for two spin- s particles in the singlet state

In Section 2.6, we considered the singlet state for a pair of spin- $\frac{1}{2}$ particles (see Eq. (2.6.38)). For the rest of this section, we consider the singlet state for two particles of any (half-)integer spin s . Alice performs measurements on particle 1, with a one-particle Hilbert space spanned by $\{|s, m_1\rangle_{1z}\}_{m_1=-s}^s$ and one-particle spin operators $\hat{S}_{1x}, \hat{S}_{1y}, \hat{S}_{1z}$. Similarly, Bob performs measurements on particle 2, with a one-particle Hilbert space spanned by $\{|s, m_2\rangle_{2z}\}_{m_2=-s}^s$ and one-particle spin operators $\hat{S}_{2x}, \hat{S}_{2y}$ and \hat{S}_{2z} . The two-particle Hilbert space therefore has dimension $(2s + 1)^2$ and is spanned by the orthonormal basis

$$\{|s, m_1\rangle_{1z} |s, m_2\rangle_{2z}\}_{m_1, m_2=-s}^s. \quad (4.1.28)$$

The spin operators on this two-particle Hilbert space are

$$\hat{S}_n = \hat{S}_{1n} + \hat{S}_{2n}. \quad (4.1.29)$$

The action of these operators on the basis vectors in Eq. (4.1.28) is given by:

$$\hat{S}_n |s, m_1\rangle_{1z} |s, m_2\rangle_{2z} = \left(\hat{S}_{1n} |s, m_1\rangle_{1z} \right) |s, m_2\rangle_{2z} + |s, m_1\rangle_{1z} \left(\hat{S}_{2n} |s, m_2\rangle_{2z} \right). \quad (4.1.30)$$

We now show that the two-particle state

$$|0, 0\rangle_{12} = \sum_{m=-s}^s \frac{(-1)^{s-m}}{\sqrt{2s+1}} |s, m\rangle_{1z} |s, -m\rangle_{2z} \quad (4.1.31)$$

is the singlet state of two particles of any (half-)integer spin s .⁵ By definition, the singlet state must be annihilated by \hat{S}^2 . Given Eq. (4.1.8), to show that \hat{S}^2 annihilates $|0, 0\rangle_{12}$, it suffices to show that \hat{S}_z and \hat{S}_+ both annihilate $|0, 0\rangle_{12}$. For the former, we note that each product state in Eq. (4.1.31) has $m_1 = -m_2 = m$ and therefore $\hat{S}_z = \hat{S}_{1z} + \hat{S}_{2z}$ annihilates $|0, 0\rangle_{12}$ term by term. For the latter, we first need to write out how $\hat{S}_+ = \hat{S}_{1+} + \hat{S}_{2+}$ acts on $|0, 0\rangle_{12}$:

$$\begin{aligned} \hat{S}_+ |0, 0\rangle_{12} &= \sum_{m=-s}^s \frac{(-1)^{s-m}}{\sqrt{2s+1}} \left[(\hat{S}_{1+} |s, m\rangle_{1z}) |s, -m\rangle_{2z} + |s, m\rangle_{1z} (\hat{S}_{2+} |s, -m\rangle_{2z}) \right] \\ &= \sum_{m=-s}^s \frac{(-1)^{s-m}}{\sqrt{2s+1}} \left[C_+(s, m) |s, m+1\rangle_{1z} |s, -m\rangle_{2z} \right. \\ &\quad \left. + C_+(s, -m) |s, m\rangle_{1z} |s, -m+1\rangle_{2z} \right]. \end{aligned} \quad (4.1.32)$$

Shifting the summation index in the second term on the right-hand side and noting that

$$C_+(s, m) = C_+(s, -m-1), \quad (4.1.33)$$

which follows from the expression for $C_+(s, m)$ in Eq. (4.1.10), we readily verify that the sum cancels term by term. So both \hat{S}_z and \hat{S}_+ annihilate $|0, 0\rangle_z$. We conclude that $|0, 0\rangle_{12}$ is indeed the singlet state.

⁵ For $s = 0$, Eq. (4.1.31) trivially reduces to $|0, 0\rangle_{12} = |0, 0\rangle_{1z} |0, 0\rangle_{2z}$. For $s = \frac{1}{2}$, it reduces to

$$|0, 0\rangle_{12} = \frac{1}{\sqrt{2}} \left(\left| \frac{1}{2}, \frac{1}{2} \right\rangle_{1z} \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_{2z} - \left| \frac{1}{2}, -\frac{1}{2} \right\rangle_{1z} \left| \frac{1}{2}, \frac{1}{2} \right\rangle_{2z} \right)$$

which can be written more compactly in the familiar form given in Eq. (2.6.38) (with $z = a$, i.e., with \mathbf{e}_a in the z -direction).

Using the rotational invariance of the singlet state (see Eq. (4.1.27)), we can rewrite the action of any rotation operator as

$$\begin{aligned} |0,0\rangle_{12} &= e^{-i\vartheta\hat{S}_n}|0,0\rangle_{12} \\ &= \sum_{m=-s}^s \frac{(-1)^{s-m}}{\sqrt{2s+1}} e^{-i\vartheta\hat{S}_{1n}}|s,m\rangle_{1z} e^{-i\vartheta\hat{S}_{2n}}|s,-m\rangle_{2z}, \end{aligned} \quad (4.1.34)$$

which, given Eq. (4.1.13), reduces to:

$$|0,0\rangle_{12} = \sum_{m=-s}^s \frac{(-1)^{s-m}}{\sqrt{2s+1}} |s,m\rangle_{1b} |s,-m\rangle_{2b}. \quad (4.1.35)$$

It follows that the singlet state has the same form in all bases. In Section 2.6 we showed this for the special case of the singlet state of two spin- $\frac{1}{2}$ particles (see Eqs. (2.6.38)–(2.6.41) and the comment following Eq. (4.1.24)).

4.1.3 Wigner d -matrices and correlation arrays in the spin- s case

Suppose that Alice measures the spin component of particle 1 in the direction \mathbf{e}_a and that Bob measures the spin component of particle 2 in the direction \mathbf{e}_b . Using the Born rule, we can compute the probability that Alice finds m_1 and Bob finds m_2 by taking the absolute square of the inner product of the singlet state with the product state $|s,m_1\rangle_{1a}|s,m_2\rangle_{2b}$ (cf. Eq. (2.6.43)):

$$\Pr(m_1 m_2 | \hat{a} \hat{b}) = \left| \left({}_{1a}\langle s, m_1 | {}_{2b}\langle s, m_2 | \right) |0,0\rangle_{12} \right|^2. \quad (4.1.36)$$

Using the expansion of the singlet state in Eq. (4.1.31), we can write the inner product on the right-hand side of Eq. (4.1.36) as

$$\sum_{m=-s}^s \frac{(-1)^{s-m}}{\sqrt{2s+1}} {}_{1a}\langle s, m_1 | s, m \rangle_{1a} {}_{2b}\langle s, m_2 | s, -m \rangle_{2a}. \quad (4.1.37)$$

Using that ${}_{1a}\langle s, m_1 | s, m \rangle_{1a} = \delta_{m_1 m}$, we arrive at

$$\left({}_{1a}\langle s, m_1 | {}_{2b}\langle s, m_2 | \right) |0,0\rangle_{12} = \frac{(-1)^{s-m_1}}{\sqrt{2s+1}} {}_{2b}\langle s, m_2 | s, -m_1 \rangle_{2a}. \quad (4.1.38)$$

To evaluate ${}_{2b}\langle s, m_2 | s, -m_1 \rangle_{2a}$ we use the invariance of the singlet state under rotation to choose, without loss of generality, the unit vectors \mathbf{e}_a and \mathbf{e}_b as

$$\mathbf{e}_a = \mathbf{e}_z, \quad \mathbf{e}_b = \mathbf{e}_z \cos \varphi_{ab} + \mathbf{e}_x \sin \varphi_{ab}. \quad (4.1.39)$$

The unit vector \mathbf{e}_b is obtained from \mathbf{e}_z by rotating around the y-axis through the angle φ_{ab} (see Figure 4.1 with \mathbf{e}_n and \mathbf{e}_a relabeled \mathbf{e}_y and \mathbf{e}_z , respectively). So we can express the states of the second particle as

$$|s, m_2\rangle_{2b} = e^{-i\varphi_{ab}\hat{S}_{2y}} |s, m_2\rangle_{2z}. \quad (4.1.40)$$

As such, the desired probability may be written as

$$\begin{aligned} \Pr(m_1 m_2 | \hat{a} \hat{b}) &= \frac{1}{2s+1} \left| {}_{2b}\langle s, m_2 | s, -m_1 \rangle_{2a} \right|^2 \\ &= \frac{1}{2s+1} \left| {}_{2a}\langle s, -m_1 | s, m_2 \rangle_{2b} \right|^2 \\ &= \frac{1}{2s+1} \left| {}_{2z}\langle s, -m_1 | e^{-i\varphi_{ab}\hat{S}_{2y}} | s, m_2 \rangle_{2z} \right|^2. \end{aligned} \quad (4.1.41)$$

The inner product in the final expression in Eq. (4.1.41) is an element of the so-called *Wigner d-matrix*, introduced in [Wigner \(1931, Ch. XV\)](#):⁶

⁶ The ‘d’ in d-matrix stands for *Darstellung* (representation) rather than *Drehung* (rotation). [Messiah \(1962, Vol. 2, p. 1070, Eq. C55\)](#) uses the notation $r^{(s)}(\beta)$ for this matrix and writes its elements as

$$r_{mm'}^{(s)}(\beta) \equiv {}_z\langle s, m | e^{-i\beta\hat{S}_y} | s, m' \rangle_z.$$

The only difference between our expressions and Messiah’s is that he is considering the total angular momentum, the sum of intrinsic and orbital angular momentum, whereas we are focusing on spin, i.e., intrinsic angular momentum. As Messiah notes, the elements of the Wigner d-matrix are always real (*ibid.*, p. 1071).

Philip W. Anderson recalls taking a class on group theory in the late 1940s with John H. Van Vleck, using Wigner’s book in the original German ([Midwinter & Janssen, 2013, p. 148](#)). During World War II, it had been reprinted in facsimile by, as it says on the title page, “Authority of the Alien Property Custodian.”

Those who have ever taken a course on quantum mechanics covering elements of group theory may vaguely recognize our Wigner d-matrices. They appear on the same page as Clebsch-Gordan coefficients in the Particle Data Group’s *Review of Particle Physics* ([Tanabashi et al., 2018, p. 564](#)). This page makes for a convenient formula sheet for exams.

$$d_{mm'}^{(s)}(\vartheta) \equiv {}_z\langle s, m | e^{-i\vartheta \hat{S}_y} | s, m' \rangle_z. \quad (4.1.42)$$

Setting $m = -m_1$, $m' = m_2$ and $\vartheta = \varphi_{ab}$, we obtain Eq. (4.1.36) in terms of these Wigner d-matrix elements:

$$\begin{aligned} \Pr(m_1 m_2 | \hat{a} \hat{b}) &= \left| \left({}_1a\langle s, m_1 | {}_2b\langle s, m_2 | | 0, 0 \rangle_{12} \right) \right|^2 \\ &= \frac{1}{2s+1} \left(d_{-m_1 m_2}^{(s)}(\varphi_{ab}) \right)^2. \end{aligned} \quad (4.1.43)$$

We examine this result for the special cases of $s = \frac{1}{2}$ and $s = 1$. In Sections 4.1.4–4.1.6 we will do so for arbitrary (half-)integer values of s . We actually already encountered the Wigner d-matrix for $s = \frac{1}{2}$ in Section 4.1.1 (see Eq. (4.1.20) with $\vartheta = \varphi_{ab}$):

$$d^{(\frac{1}{2})}(\varphi_{ab}) = \begin{pmatrix} \cos(\varphi_{ab}/2) & -\sin(\varphi_{ab}/2) \\ \sin(\varphi_{ab}/2) & \cos(\varphi_{ab}/2) \end{pmatrix}. \quad (4.1.44)$$

Squaring these matrix elements, we recover the probabilities in the $\hat{a}\hat{b}$ cell of the correlation array for the spin- $\frac{1}{2}$ case in Figure 2.19.

For $s = 1$, the Wigner d-matrix is

$$d^{(1)}(\varphi_{ab}) = \begin{pmatrix} \frac{1}{2}(1 + \cos \varphi_{ab}) & -\frac{1}{2}\sqrt{2} \sin \varphi_{ab} & \frac{1}{2}(1 - \cos \varphi_{ab}) \\ \frac{1}{2}\sqrt{2} \sin \varphi_{ab} & \cos \varphi_{ab} & -\frac{1}{2}\sqrt{2} \sin \varphi_{ab} \\ \frac{1}{2}(1 - \cos \varphi_{ab}) & \frac{1}{2}\sqrt{2} \sin \varphi_{ab} & \frac{1}{2}(1 + \cos \varphi_{ab}) \end{pmatrix} \quad (4.1.45)$$

(Messiah, 1962, Vol. 2, p. 1073, Eq. (C.75) with $\alpha = \gamma = 0$ and $\beta = \varphi_{ab}$; see also Wigner, 1931, p. 182, Eq. 29). The elements of this matrix are $d_{ij}^{(1)}(\varphi_{ab})$ where $i = 1, 0, -1$ labels the rows and $j = 1, 0, -1$ labels the columns.

Squaring these matrix elements and introducing (with malice aforethought) the notation

$$\chi_{ab} \equiv \cos \varphi_{ab} \quad (4.1.46)$$

(and similarly $\chi_{bc} \equiv \cos \varphi_{bc}$ and $\chi_{ac} \equiv \cos \varphi_{ac}$) we find the probabilities in the cell of a correlation array in Figure 4.2. We verify this for two of these probabilities. The probability that Alice (using setting \hat{a}) and Bob (using setting \hat{b}) both find the outcome 1 is given by Eq. (4.1.43) with $s = 1$ and the Wigner

d-matrix in Eq. (4.1.45):

$$\begin{aligned}
 \Pr(m_1 = 1, m_2 = 1 | \hat{a} \hat{b}) &= \frac{1}{3} \left(d_{-11}^{(1)}(\varphi_{ab}) \right)^2 \\
 &= \frac{1}{12} (1 - \cos \varphi_{ab})^2 \\
 &= \frac{1}{12} (1 - \chi_{ab})^2. \tag{4.1.47}
 \end{aligned}$$

Similarly, the probability that Alice finds 1 for \hat{a} and Bob finds 0 for \hat{b} is given by:

$$\begin{aligned}
 \Pr(m_1 = 1, m_2 = 0 | \hat{a} \hat{b}) &= \frac{1}{3} \left(d_{-10}^{(1)}(\varphi_{ab}) \right)^2 \\
 &= \frac{1}{6} \sin^2 \varphi_{ab} \\
 &= \frac{1}{6} (1 - \chi_{ab}^2). \tag{4.1.48}
 \end{aligned}$$

		Bob		
		\hat{b}		
Alice		1	0	-1
	1	$\frac{1}{12}(1 - \chi_{ab})^2$	$\frac{1}{6}(1 - \chi_{ab}^2)$	$\frac{1}{12}(1 + \chi_{ab})^2$
	0	$\frac{1}{6}(1 - \chi_{ab}^2)$	$\frac{1}{3}\chi_{ab}^2$	$\frac{1}{6}(1 - \chi_{ab}^2)$
-1	$\frac{1}{12}(1 + \chi_{ab})^2$	$\frac{1}{6}(1 - \chi_{ab}^2)$	$\frac{1}{12}(1 - \chi_{ab})^2$	

Fig. 4.2 Cell in a correlation array given by quantum mechanics for measurements on the singlet state of two spin-1 particles ($\chi_{ab} = \cos \varphi_{ab}$).

The $\hat{a} \hat{c}$ and $\hat{b} \hat{c}$ cells of the correlation arrays will have the exact same structure as the $\hat{a} \hat{b}$ one, with φ_{ab} replaced by φ_{ac} and φ_{bc} . The correlation arrays for measurements on the singlet state of two spin- $\frac{1}{2}$ or two spin-1 particles can thus be parametrized by the cosines of the angles φ_{ab} , φ_{ac} and

φ_{bc} between the directions \mathbf{e}_a , \mathbf{e}_b and \mathbf{e}_c in which the spin is being measured. This is also true for the singlet state of two particles with higher spin (Messiah, 1962, Vol. 2, p. 1072, Eq. (3.72)). In Section 4.1.5, we will show that these cosines can be interpreted as anti-correlation coefficients, thereby justifying the notation introduced in Eq. (4.1.46).

4.1.4 Non-signaling in the spin- s case

We now show that the correlations found in measurements on the singlet state of two particles of any (half-)integer spin s have uniform marginals and are therefore non-signaling. Consider the correlation array in Figure 4.2 and the marginal probability of Alice finding the outcome $m_1 = 1$ when she uses setting \hat{a} and Bob uses setting \hat{b} :

$$\begin{aligned} \Pr(+1_A|\hat{a}\hat{b}) &= \sum_{m_2=-1}^1 \Pr(+1 m_2|\hat{a}\hat{b}) \\ &= \frac{1}{12}(1 + \chi_{ab})^2 + \frac{1}{6}(1 - \chi_{ab}^2) + \frac{1}{12}(1 - \chi_{ab})^2 = \frac{1}{3}. \end{aligned} \quad (4.1.49)$$

For $m_1 = 0$ and $m_1 = -1$ we similarly find that

$$\Pr(0_A|\hat{a}\hat{b}) = \Pr(-1_A|\hat{a}\hat{b}) = \frac{1}{3}. \quad (4.1.50)$$

None of these marginal probabilities—and this observation is key—depend on φ_{ab} . They are thus unaltered if Bob's settings are changed from \hat{b} to \hat{a} or \hat{c} . The same is true for the marginal probabilities of Bob measuring $m_2 = (1, 0, -1)$ using any of these three settings. In every cell of the correlation array that the cell in Figure 4.2 is part of, all three rows and all three columns add up to $1/3$. Like the singlet state for a pair of spin- $\frac{1}{2}$ particles, the singlet state for a pair of spin-1 particles thus cannot be used for superluminal signaling.

The same is true for the singlet state of two particles of higher spin. The marginal probability of Alice finding m_1 for \hat{a} when Bob uses \hat{b} is:

$$\Pr(m_1|\hat{a}\hat{b}) = \sum_{m_2=-s}^s \Pr(m_1 m_2|\hat{a}\hat{b}). \quad (4.1.51)$$

Using the first line of Eq. (4.1.41), we find that

$$\begin{aligned}
\Pr(m_1|\hat{a}\hat{b}) &= \frac{1}{2s+1} \sum_{m_2=-s}^s |{}_{2b}\langle s, m_2|s, -m_1\rangle_{2a}|^2 \\
&= \frac{1}{2s+1} \sum_{m_2=-s}^s {}_{2a}\langle s, -m_1|s, m_2\rangle_{2b} {}_{2b}\langle s, m_2|s, -m_1\rangle_{2a}. \quad (4.1.52)
\end{aligned}$$

Evaluating this sum, using the completeness relation

$$\hat{1}_2 = \sum_{m=-s}^s |s, m\rangle_{2b} {}_{2b}\langle s, m|, \quad (4.1.53)$$

we arrive at

$$\Pr(m_1|\hat{a}\hat{b}) = \frac{1}{2s+1} {}_{2a}\langle s, -m_1|s, -m_1\rangle_{2a} = \frac{1}{2s+1}. \quad (4.1.54)$$

This same formula holds if we substitute m_2 for m_1 or any two of the triplet of settings $(\hat{a}, \hat{b}, \hat{c})$ for $\hat{a}\hat{b}$. It follows that the correlations found in the measurements on the singlet state of two spin- s particles are indeed non-signaling. In every cell of the corresponding correlation array, all $2s+1$ rows and all $2s+1$ columns sum to $1/(2s+1)$.

4.1.5 Anti-correlation coefficients in the spin- s case

We now turn our attention to the quantities χ_{ab} introduced in Eq. (4.1.46) and the analogous quantities χ_{ac} and χ_{bc} . We show that these can be interpreted as anti-correlation coefficients just as in the case of $s = 1/2$ (see Eq. (2.6.48)). Consider the expectation value

$$\langle \hat{S}_{1a}\hat{S}_{2b} \rangle_{00} \equiv {}_{12}\langle 0, 0 | \hat{S}_{1a}\hat{S}_{2b} | 0, 0 \rangle_{12}. \quad (4.1.55)$$

Recalling our choice of \mathbf{e}_a and \mathbf{e}_b in Eq. (4.1.39), we have

$$\hat{S}_{1a} = \hat{S}_{1z}, \quad \hat{S}_{2b} = \hat{S}_{2z} \cos \varphi_{ab} + \hat{S}_{2x} \sin \varphi_{ab}. \quad (4.1.56)$$

Inserting these expressions into Eq. (4.1.55), we arrive at:

$$\langle \hat{S}_{1a}\hat{S}_{2b} \rangle_{00} = \langle \hat{S}_{1z}\hat{S}_{2z} \rangle_{00} \cos \varphi_{ab} + \langle \hat{S}_{1z}\hat{S}_{2x} \rangle_{00} \sin \varphi_{ab}. \quad (4.1.57)$$

The quantity $\langle \hat{S}_{1z}\hat{S}_{2z} \rangle_{00}$ in the first term on the right-hand side is minus the square of the standard deviation σ_s (see Eq. (3.3.2) in Chapter 3). This quantity

is thus given by

$$\langle \hat{S}_{1z} \hat{S}_{2z} \rangle_{00} = \sigma_s^2 = -\frac{1}{3} s(s+1). \quad (4.1.58)$$

This same result can be derived directly from properties of the singlet state. The expectation value of the product of \hat{S}_{1z} and \hat{S}_{2z} in the singlet state is given by

$$\langle \hat{S}_{1z} \hat{S}_{2z} \rangle_{00} = {}_{12} \langle 0, 0 | \hat{S}_{1z} \hat{S}_{2z} | 0, 0 \rangle_{12}. \quad (4.1.59)$$

Using that $\hat{S}_{2z} = \hat{S}_z - \hat{S}_{1z}$ and that $\hat{S}_z | 0, 0 \rangle_{12} = 0$, we can rewrite this as

$$\langle \hat{S}_{1z} \hat{S}_{2z} \rangle_{00} = -{}_{12} \langle 0, 0 | \hat{S}_{1z}^2 | 0, 0 \rangle_{12}. \quad (4.1.60)$$

Rotational invariance requires that

$${}_{12} \langle 0, 0 | \hat{S}_{1x}^2 | 0, 0 \rangle_{12} = {}_{12} \langle 0, 0 | \hat{S}_{1y}^2 | 0, 0 \rangle_{12} = {}_{12} \langle 0, 0 | \hat{S}_{1z}^2 | 0, 0 \rangle_{12}. \quad (4.1.61)$$

Hence

$$\langle \hat{S}_{1z} \hat{S}_{2z} \rangle_{00} = -\frac{1}{3} {}_{12} \langle 0, 0 | (\hat{S}_{1x}^2 + \hat{S}_{1y}^2 + \hat{S}_{1z}^2) | 0, 0 \rangle_{12}. \quad (4.1.62)$$

Substituting \hat{S}_1^2 for $\hat{S}_{1x}^2 + \hat{S}_{1y}^2 + \hat{S}_{1z}^2$ and using that $\hat{S}_1^2 | 0, 0 \rangle_{12} = s(s+1) | 0, 0 \rangle_{12}$, we recover Eq. (4.1.58):

$$\langle \hat{S}_{1z} \hat{S}_{2z} \rangle_{00} = -\frac{1}{3} {}_{12} \langle 0, 0 | \hat{S}_1^2 | 0, 0 \rangle_{12} = -\frac{1}{3} s(s+1). \quad (4.1.63)$$

Again using the rotational invariance of the singlet state, we can show that the second term in Eq. (4.1.57) vanishes. Consider a rotation of the singlet state over 180° around the z -axis. Since the singlet state is invariant under arbitrary rotation, the action of the operator $e^{-i\pi \hat{S}_z}$ implementing this rotation (see Eq. (4.1.12)) on the singlet state simply reproduces the singlet state. It follows that

$$\langle \hat{S}_{1z} \hat{S}_{2x} \rangle_{00} = {}_{12} \langle 0, 0 | \hat{S}_{1z} \hat{S}_{2x} | 0, 0 \rangle_{12} = {}_{12} \langle 0, 0 | e^{i\pi \hat{S}_z} \hat{S}_{1z} \hat{S}_{2x} e^{-i\pi \hat{S}_z} | 0, 0 \rangle_{12}. \quad (4.1.64)$$

Inserting $\hat{S}_z = \hat{S}_{1z} + \hat{S}_{2z}$ and using that \hat{S}_{1z} commutes with both \hat{S}_{2z} and \hat{S}_{2x} , we can rewrite this as:

$$\langle \hat{S}_{1z} \hat{S}_{2x} \rangle_{00} = {}_{12} \langle 0, 0 | \hat{S}_{1z} e^{i\pi \hat{S}_{2z}} \hat{S}_{2x} e^{-i\pi \hat{S}_{2z}} | 0, 0 \rangle_{12}. \quad (4.1.65)$$

Recalling the transformation law for spin operators in Eq. (4.1.12), we note that:

$$e^{i\pi\hat{S}_{2z}}\hat{S}_{2x}e^{-i\pi\hat{S}_{2z}} = -\hat{S}_{2x}. \quad (4.1.66)$$

Inserting this expression into Eq. (4.1.65), we conclude that

$$\langle\hat{S}_{1z}\hat{S}_{2x}\rangle_{00} = -{}_{12}\langle 0,0|\hat{S}_{1z}\hat{S}_{2x}|0,0\rangle_{12} = -\langle\hat{S}_{1z}\hat{S}_{2x}\rangle_{00} = 0. \quad (4.1.67)$$

So the only contribution to Eq. (4.1.57) comes from the first term:

$$\langle\hat{S}_{1a}\hat{S}_{2b}\rangle_{00} = \langle\hat{S}_{1a}\hat{S}_{2a}\rangle_{00} \cos\varphi_{ab} = -\frac{1}{3}s(s+1)\chi_{ab}, \quad (4.1.68)$$

where we used Eq. (4.1.63) for $\langle\hat{S}_{1a}\hat{S}_{2a}\rangle_{00}$ and the definition of χ_{ab} in Eq. (4.1.46). Using that the standard deviations

$$\sigma_{1a} \equiv \sqrt{\langle\hat{S}_{1a}^2\rangle_{00}} \quad \text{and} \quad \sigma_{2b} \equiv \sqrt{\langle\hat{S}_{2b}^2\rangle_{00}} \quad (4.1.69)$$

are both given by Eq. (4.1.58), we can rewrite Eq. (4.1.68) as

$$\chi_{ab} = -\frac{\langle\hat{S}_{1a}\hat{S}_{2b}\rangle_{00}}{\sigma_a\sigma_b} \quad (4.1.70)$$

which we recognize as the definition of the anti-correlation coefficient χ_{ab} (see Eqs. (3.2.5) and (3.2.6)). This justifies the use of the χ_{ab} notation in Eq. (4.1.46). This identification moreover means that the main conclusion of our examination of the special case of spin- $\frac{1}{2}$ particles in Section 2.6 carries over to the general case considered in this section: the set of values for $(\chi_{ab}, \chi_{ac}, \chi_{bc})$ that can be obtained by measurements on the singlet state of two particles of higher spin saturates the ellipsope in Figure 2.16.

4.1.6 Cell symmetries in the spin- s case

In this subsection, we will consider various symmetries of cells in correlation arrays for measurements on particles of arbitrary (half-)integer spin s . Figures 2.19 and 4.2 show such cells for the $s = 1/2$ and $s = 1$ cases. Figure 4.3 shows it for the $s = 3/2$ case.

Cells on the diagonals in all these cases are particularly simple. Since for those cells $\chi_{ab} = \cos\varphi_{ab} = 1$, elements on the skew-diagonal are equal and sum to 1 while the other elements are 0. This is true not just for $s = 1/2$

		Bob			
		\hat{b}			
Alice	\hat{a}	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$
	$\frac{3}{2}$	$\frac{1}{32}(1-\chi)^3$	$\frac{3}{32}(1-\chi)^2(1+\chi)$	$\frac{3}{32}(1+\chi)^2(1-\chi)$	$\frac{1}{32}(1+\chi)^3$
	$\frac{1}{2}$	$\frac{3}{32}(1-\chi)^2(1+\chi)$	$\frac{1}{32}(1-\chi)(1+3\chi)^2$	$\frac{1}{32}(1+\chi)(1-3\chi)^2$	$\frac{3}{32}(1+\chi)^2(1-\chi)$
	$-\frac{1}{2}$	$\frac{3}{32}(1+\chi)^2(1-\chi)$	$\frac{1}{32}(1+\chi)(1-3\chi)^2$	$\frac{1}{32}(1-\chi)(1+3\chi)^2$	$\frac{3}{32}(1-\chi)^2(1+\chi)$
$-\frac{3}{2}$	$\frac{1}{32}(1+\chi)^3$	$\frac{3}{32}(1+\chi)^2(1-\chi)$	$\frac{3}{32}(1-\chi)^2(1+\chi)$	$\frac{1}{32}(1-\chi)^3$	

Fig. 4.3 Cell in a correlation array given by quantum mechanics for measurements on the singlet state of two spin- $\frac{3}{2}$ particles ($\chi = \cos \varphi_{ab}$).

and $s = 1$, but for any (half-)integer value of s . Whenever the angle between the measuring directions used by Alice and Bob is zero, Eq. (4.1.42) for the Wigner d-matrix element reduces to

$$d_{-m_1 m_2}^{(s)}(0) = {}_z \langle s, -m_1 | s, m_2 \rangle_z = \delta_{-m_1 m_2}. \quad (4.1.71)$$

Hence, the probability for Alice to measure m_1 and Bob to measure m_2 along a common direction, say \mathbf{e}_a , is given by (cf. Eq. (4.1.43)):

$$\Pr(m_1 m_2 | \hat{a} \hat{a}) = \frac{1}{2s+1} \left(d_{-m_1 m_2}^{(s)}(0) \right)^2 = \frac{1}{2s+1} \delta_{-m_1 m_2}. \quad (4.1.72)$$

Any cell on the diagonal of a correlation array for measurements on the singlet state of two particles of any integer or half-integer spin s thus has values $1/(2s+1)$ on the skew-diagonal and zeros everywhere else (see Figure 3.4 in Chapter 3).

The off-diagonal cells, while not as simple as the diagonal ones, also exhibit features that are the same for all values of s . Note, for instance, that the cells in Figure 2.19 (for $s = 1/2$) and Figures 4.2–4.3 (for $s = 1$ and $s = 3/2$) are all

symmetric across both the diagonal and the skew-diagonal. This is true not just for $s = 1/2$ and $s = 1$ but for any (half-)integer s . This follows directly from the following three symmetry properties of the Wigner d-matrix in Eq. (4.1.42):

$$d_{mm'}^{(s)}(\vartheta) = (-1)^{m-m'} d_{m'm}^{(s)}(\vartheta) \quad (4.1.73)$$

$$= (-1)^{m-m'} d_{-m-m'}^{(s)}(\vartheta) \quad (4.1.74)$$

$$= d_{-m'-m}^{(s)}(\vartheta). \quad (4.1.75)$$

To establish the first of these symmetries, consider a rotation of the one-particle system through an angle of 180° around the z -axis, implemented as $e^{-i\pi\hat{S}_z}$. Such a rotation leaves the z -axis unchanged and as such the states are only changed up to an overall phase factor, i.e.,

$$e^{-i\pi\hat{S}_z} |s, m\rangle_z = e^{-i\pi m} |s, m\rangle_z. \quad (4.1.76)$$

By contrast, this rotation flips the y -axis, which means that the operator \hat{S}_y transforms as

$$e^{-i\pi\hat{S}_z} \hat{S}_y e^{i\pi\hat{S}_z} = -\hat{S}_y. \quad (4.1.77)$$

The same is true for any function of \hat{S}_y :

$$e^{-i\pi\hat{S}_z} f(\hat{S}_y) e^{i\pi\hat{S}_z} = f(-\hat{S}_y). \quad (4.1.78)$$

Hence, we can rewrite the Wigner d-matrix element in Eq. (4.1.42) as

$$d_{mm'}^{(s)}(\vartheta) = {}_z\langle s, m | e^{-i\vartheta\hat{S}_y} |s, m'\rangle_z = {}_z\langle s, m | e^{-i\pi\hat{S}_z} e^{i\vartheta\hat{S}_y} e^{i\pi\hat{S}_z} |s, m'\rangle_z. \quad (4.1.79)$$

On account of Eq. (4.1.76), this reduces to:

$$d_{mm'}^{(s)}(\vartheta) = {}_z\langle s, m | e^{-i\pi m} e^{i\vartheta\hat{S}_y} e^{i\pi m'} |s, m'\rangle_z = (-1)^{m-m'} d_{mm'}^{(s)}(-\vartheta). \quad (4.1.80)$$

Given the Condon-Shortley phase convention, the operator $i\hat{S}_y$ has real matrix elements in the $|s, m\rangle_z$ basis (see Eqs. (4.1.6)–(4.1.10)). Since the elements of the Wigner d-matrix are matrix elements of a function of $i\hat{S}_y$ in this basis, it follows that they too must be real. We can thus rewrite $d_{mm'}^{(s)}(-\vartheta)$ in Eq. (4.1.80) as:

$$\begin{aligned}
d_{mm'}^{(s)}(-\vartheta) &= d_{mm'}^{(s)}(-\vartheta)^* \\
&= {}_z\langle s, m | e^{i\vartheta \hat{S}_y} | s, m' \rangle_z^* \\
&= {}_z\langle s, m' | e^{-i\vartheta \hat{S}_y} | s, m \rangle_z \\
&= d_{m'm}^{(s)}(\vartheta).
\end{aligned} \tag{4.1.81}$$

Inserting this expression in Eq. (4.1.80), we arrive at the symmetry stated in Eq. (4.1.73).

To establish the second symmetry of the Wigner d-matrix, the one in Eq. (4.1.74), we consider a rotation through an angle of 180° around the y-axis, as implemented by $e^{-i\pi \hat{S}_y}$. This leaves \hat{S}_y unchanged but flips \hat{S}_z . Thus the action of this rotation operator on the state $|s, m\rangle_z$ —aside from an overall phase factor⁷—is to replace m by $-m$:

$$e^{-i\pi \hat{S}_y} |s, m\rangle_z = (-1)^{s-m} |s, -m\rangle_z. \tag{4.1.82}$$

Inserting the operators $e^{-i\pi \hat{S}_y}$ and $e^{i\pi \hat{S}_y}$ in the expression for the Wigner d-matrix in Eq. (4.1.42),

$$d_{mm'}^{(s)}(\vartheta) = {}_z\langle s, m | e^{-i\pi \hat{S}_y} e^{-i\vartheta \hat{S}_y} e^{i\pi \hat{S}_y} | s, m' \rangle_z, \tag{4.1.83}$$

and using Eq. (4.1.82), we arrive at the symmetry in Eq. (4.1.74):

$$\begin{aligned}
d_{mm'}^{(s)}(\vartheta) &= {}_z\langle s, -m | (-1)^{s-m} e^{-i\vartheta \hat{S}_y} (-1)^{s-m'} | s, -m' \rangle_z \\
&= (-1)^{m-m'} d_{-m-m'}^{(s)}(-\vartheta),
\end{aligned} \tag{4.1.84}$$

where in the last step we used that $(-1)^{s-m'} = (-1)^{m'-s}$.

The third symmetry of the Wigner d-matrix, the one in Eq. (4.1.75), follows as a corollary of the first two, the ones in Eqs. (4.1.73) and (4.1.74), though it can also be established directly through an argument similar to those in Eqs.

⁷ This phase factor appears in Messiah (1962, Vol. 2, p. 1071, Eq. C65) and may be derived, for instance, by appeal to Wigner's explicit formula (see p. 1072, Eq. C72 in Messiah) for the d-matrix elements. However, this phase factor does not enter into the probabilities and so we do not derive its value here.

(4.1.76)–(4.1.81) and Eqs. (4.1.82)–(4.1.82), involving a rotation around the x -axis.

These three symmetries of the Wigner d -matrix translate into three symmetries of the probabilities in Eq. (4.1.43):

$$\begin{aligned} \Pr(m_1 m_2 | \hat{a} \hat{b}) &= \Pr(-m_2 -m_1 | \hat{a} \hat{b}) && \text{persymmetry} \\ &= \Pr(-m_1 -m_2 | \hat{a} \hat{b}) && \text{centrosymmetry} \\ &= \Pr(m_2 m_1 | \hat{a} \hat{b}) && \text{symmetry.} \end{aligned} \quad (4.1.85)$$

This tells us that the probability is unchanged if we either swap m_1 and m_2 or flip both of their signs. These symmetries, in turn, translate into symmetries of any cell in a correlation array for measurements on the singlet state of two particles with spin s . The first line in Eq. (4.1.85) expresses that any such cell is *persymmetric*, i.e., symmetric across its main anti-diagonal; the second that it is *centrosymmetric*, i.e., symmetric about its center; the third that it is *symmetric*, i.e., symmetric across its main diagonal. Any two of these symmetries imply the third. As we noted in the introduction to Chapter 4, in the spin- $\frac{1}{2}$ case, centrosymmetry implies both symmetry and persymmetry. Since the design of our raffle tickets guarantees centrosymmetry, we did not have to impose any conditions on our raffles to ensure all three symmetries. As we will see in Section 4.2, however, such conditions are needed for higher-spin cases if we want our raffles to give correlation arrays with the same symmetries as the quantum correlation arrays they are supposed to simulate.

4.2 Designing raffles to simulate the quantum correlations

We will now design raffles to simulate the quantum correlations found in measurements on the singlet state of two particles with spin $s \geq \frac{1}{2}$ that we investigated in Section 4.1. In Section 4.2.1, to set the stage and fix notation, we give a more formal analysis of the raffles for the spin- $\frac{1}{2}$ case we introduced in Section 2.5. In Sections 4.2.2–2.2.4 we gradually work our way up to higher-spin cases.

4.2.1 Spin- $\frac{1}{2}$

The raffles we considered in Section 2.5 all involved baskets containing a mixture of the four ticket types shown in Figure 4.4 (see the tree structure in Figure 2.11 for their numbering). Let f^k denote the fraction of tickets of type

(k) in such a basket (with $k = i, ii, iii, iv$). These ticket fractions evidently are non-negative and normalized:

$$f^k \geq 0 \quad (k = i, ii, iii, iv) \quad \text{and} \quad \sum_{k=i}^{iv} f^k = 1. \quad (4.2.1)$$

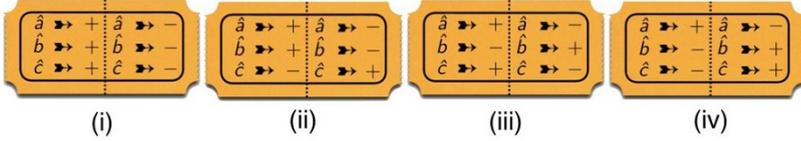


Fig. 4.4 The four different raffle tickets for three settings and two outcomes (cf. Figure 2.11).

As we observed in Section 2.5 (see note 17), the imagery of a basket with a mix of tickets restricts us to values for f^k that are rational numbers. We will continue to discuss our raffles in terms of baskets of tickets, but we do want to point out that with a simple change of imagery we can accommodate real values as well. Instead of baskets with tickets, consider wheels of fortune such as those shown in Figure 4.5. These wheels have pie charts printed on them showing the mix of tickets in a particular raffle, each ticket of type (k) occurring with a fraction f^k in the raffle corresponding to a segment $f^k \times 100\%$ of the pie chart. Instead of randomly drawing a ticket from a basket, we would spin the wheel of fortune and pick a ticket of the type the pointer points to when the wheel of fortune comes to rest.

When we think of

$$\mathbf{f} \equiv (f^i, f^{ii}, f^{iii}, f^{iv}) \quad (4.2.2)$$

as a point in \mathbb{R}^4 , the constraints in Eq. (4.2.1) define what is conventionally known as the *3D standard simplex* or *3-simplex* in \mathbb{R}^4 . Equivalently, this simplex is the convex hull of the four points

$$\mathbf{f}_i = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{f}_{ii} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{f}_{iii} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{f}_{iv} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad (4.2.3)$$

each corresponding to a raffle with just one type of ticket in the basket. These single-ticket raffles are *basic* in the sense that any raffle can be obtained as a mix of them. As the notation already suggests, Eq. (4.2.3) can also be seen as giving the unit vectors of an *orthonormal basis* in which we can expand

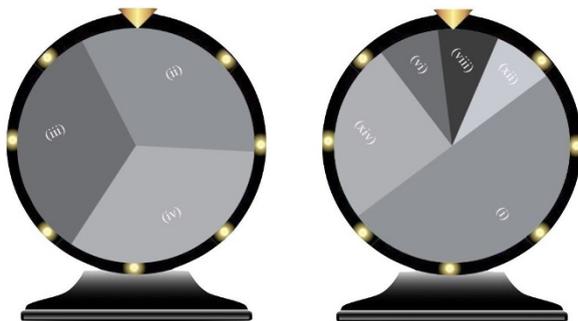


Fig. 4.5 Wheels of fortune for two raffles. The one on the left is for a raffle with a mix of $1/3$ each of tickets of type (ii), (iii) and (iv) in Figure 4.4. This is raffle (b) in Figure 2.14 in Section 2.5, our optimal simulation of the Mermin correlation array in Figure 2.6. The one on the right is for a 75%/25% mix of the admissible raffles (α) and (β) in Figure 4.9 in Section 4.2.2.

vectors characterizing arbitrary raffles. The quantity \mathbf{f} in Eq. (4.2.2) can then be thought of as the vector

$$\mathbf{f} = \sum_{k=i}^{iv} f^k \mathbf{f}_k \quad (4.2.4)$$

and the ticket fractions f^k as its components in this basis.

The 3-simplex is a polytope in 4D Euclidean space and as such cannot be visualized. To circumvent this problem we consider triplets of anti-correlation coefficients rather than quartets of ticket fractions. The anti-correlation coefficient $\chi_{ab}|\mathbf{f}$ for an arbitrary (mixed or single-ticket) raffle characterized by some vector \mathbf{f} is defined as (cf. Eqs. (3.2.5)–(3.2.6)):

$$\chi_{ab}|\mathbf{f} \equiv - \left. \frac{\langle X_a^A X_b^B \rangle}{\sigma_a^A \sigma_b^B} \right|_{\mathbf{f}}. \quad (4.2.5)$$

The anti-correlation coefficient for this raffle is equal to the weighted average of the anti-correlation coefficients for the four single-ticket raffles in Eq. (4.2.3) with the weights given by the ticket fractions f^k . In Section 2.5, we already used this property of our raffles. We will now give a formal proof of it.

We start by evaluating the covariance in the numerator on the right-hand side of Eq. (4.2.5) (cf. Eq. (3.1.7)). By definition,

$$\langle X_a^A X_b^B \rangle|_{\mathbf{f}} = \sum_{m_1, m_2} m_1 m_2 \Pr(m_1 m_2 | \hat{a} \hat{b})|_{\mathbf{f}}, \quad (4.2.6)$$

where the outcomes m_1 and m_2 can only take on the values $\pm 1/2$ (if we set \hbar and \hbar equal to 1). It immediately follows from the design of our raffles that the probability of finding some combination of outcomes for some combination of settings in a raffle with a mix of tickets characterized by \mathbf{f} is the weighted average of those same probabilities in the four basic single-ticket raffles with the weights given by the ticket fractions:

$$\Pr(m_1 m_2 | \hat{a} \hat{b}) \Big|_{\mathbf{f}} = \sum_{k=i}^{iv} f^k \Pr(m_1 m_2 | \hat{a} \hat{b}) \Big|_{\mathbf{f}_k}. \quad (4.2.7)$$

The probabilities on the right-hand side take on a very simple form: they are $1/2$ if the ticket has the combination of outcomes ($X_a = m_1, X_b = m_2$) and zero if it does not. For example, the tickets in Figure 4.4 tell us that

$$\begin{aligned} \Pr\left(\frac{1}{2}, -\frac{1}{2} \Big| \hat{a} \hat{b}\right) \Big|_{\mathbf{f}_i} &= \Pr\left(\frac{1}{2}, -\frac{1}{2} \Big| \hat{a} \hat{b}\right) \Big|_{\mathbf{f}_{ii}} = \frac{1}{2}, \\ \Pr\left(\frac{1}{2}, -\frac{1}{2} \Big| \hat{a} \hat{b}\right) \Big|_{\mathbf{f}_{iii}} &= \Pr\left(\frac{1}{2}, -\frac{1}{2} \Big| \hat{a} \hat{b}\right) \Big|_{\mathbf{f}_{iv}} = 0. \end{aligned} \quad (4.2.8)$$

For this combination of outcomes, Eq. (4.2.7) thus gives

$$\Pr\left(\frac{1}{2}, -\frac{1}{2} \Big| \hat{a} \hat{b}\right) \Big|_{\mathbf{f}} = \frac{1}{2} (f^i + f^{ii}). \quad (4.2.9)$$

As this example illustrates, Eq. (4.2.7) is a more formal expression of another feature of our raffles we repeatedly made use of in Section 2.5: the entries in the correlation array for a mixed raffle are weighted averages of the corresponding entries in the correlation arrays for single-ticket raffles.

Inserting Eq. (4.2.7) into Eq. (4.2.6) and changing the order of the summations, we see that the covariance $\langle X_a^A X_b^B \rangle$ in a mixed raffle is likewise the weighted average of that same covariance in the four single-ticket raffles:

$$\langle X_a^A X_b^B \rangle \Big|_{\mathbf{f}} = \sum_{k=i}^{iv} f^k \sum_{m_1, m_2} m_1 m_2 \Pr(m_1 m_2 | \hat{a} \hat{b}) \Big|_{\mathbf{f}_k} = \sum_{k=i}^{iv} f^k \langle X_a^A X_b^B \rangle \Big|_{\mathbf{f}_k}. \quad (4.2.10)$$

Because the diagonal cells in the correlation arrays for raffles with any mix of tickets of types (i) through (iv) in Figure 4.4 are the same, the standard deviations in the expression on the right-hand side of Eq. (4.2.5) are also the same for all these raffles. For any \mathbf{f} , they are given by Eq. (3.3.3) in Chapter 3

for $s = 1/2$:

$$\sigma_a|_{\mathbf{f}} = \sigma_b|_{\mathbf{f}} = \sigma_{s=1/2} = \sqrt{\frac{s(s+1)}{3}} = \frac{1}{2}. \quad (4.2.11)$$

Substituting Eq. (4.2.10) into Eq. (4.2.5) and using Eq. (4.2.11), we arrive at

$$\chi_{ab}|_{\mathbf{f}} = -\frac{\langle X_a^A X_b^B \rangle|_{\mathbf{f}}}{\sigma_{s=1/2}^2} = \sum_{k=i}^{iv} f^k \left(-\frac{\langle X_a^A X_b^B \rangle}{\sigma_a \sigma_b} \Big|_{\mathbf{f}_k} \right) = \sum_{k=i}^{iv} f^k \chi_{ab}|_{\mathbf{f}_k}, \quad (4.2.12)$$

which is what we set out to prove.

Similar relations hold for χ_{ac} and χ_{bc} . We combine these three relations and write them in matrix form:

$$\begin{pmatrix} \chi_{ab} \\ \chi_{ac} \\ \chi_{bc} \end{pmatrix} \Big|_{\mathbf{f}} = \begin{pmatrix} \chi_{ab}|_{\mathbf{f}_i} & \chi_{ab}|_{\mathbf{f}_{ii}} & \chi_{ab}|_{\mathbf{f}_{iii}} & \chi_{ab}|_{\mathbf{f}_{iv}} \\ \chi_{ac}|_{\mathbf{f}_i} & \chi_{ac}|_{\mathbf{f}_{ii}} & \chi_{ac}|_{\mathbf{f}_{iii}} & \chi_{ac}|_{\mathbf{f}_{iv}} \\ \chi_{bc}|_{\mathbf{f}_i} & \chi_{bc}|_{\mathbf{f}_{ii}} & \chi_{bc}|_{\mathbf{f}_{iii}} & \chi_{bc}|_{\mathbf{f}_{iv}} \end{pmatrix} \begin{pmatrix} f^i \\ f^{ii} \\ f^{iii} \\ f^{iv} \end{pmatrix}. \quad (4.2.13)$$

The 3×4 matrix M on the right-hand side thus serves as a map from the 3-simplex of possible ticket fractions to the *anti-correlation polyhedron* (see the introduction of Chapter 4 for a definition). Denoting the vector on the left-hand side as $\chi|_{\mathbf{f}}$, we can write Eq. (4.2.13) more compactly as

$$\chi|_{\mathbf{f}} \equiv M\mathbf{f}. \quad (4.2.14)$$

The matrix elements of M can be read off of Table 2.1 in Section 2.5 (or directly from the tickets in Figure 4.4: note that $1/\sigma_{s=1/2}^2 = 4$ cancels the 4 in the denominators of the covariances):

$$M = \begin{pmatrix} M_{ab}^i & M_{ab}^{ii} & M_{ab}^{iii} & M_{ab}^{iv} \\ M_{ac}^i & M_{ac}^{ii} & M_{ac}^{iii} & M_{ac}^{iv} \\ M_{bc}^i & M_{bc}^{ii} & M_{bc}^{iii} & M_{bc}^{iv} \end{pmatrix} = \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}. \quad (4.2.15)$$

M will map the basis vectors in Eq. (4.2.3) for the four single-ticket raffles onto vectors whose components are the columns of M :

$$\chi|_{\mathbf{f}_i} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \chi|_{\mathbf{f}_{ii}} = \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, \quad \chi|_{\mathbf{f}_{iii}} = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}, \quad \chi|_{\mathbf{f}_{iv}} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}. \quad (4.2.16)$$

These vectors correspond to the vertices (i) through (iv) of the tetrahedron we found in Section 2.5 (see Figure 2.15).

4.2.2 Spin-1

Having reviewed the spin- $\frac{1}{2}$ case, we now move to new territory and consider the spin-1 case. The first change is to the number of tickets. There are now three possible outcomes for the three settings: 1, 0 and -1 (again setting $\hbar = \bar{\hbar} = 1$). This means that there are now $3^3 = 27$ ways to specify the outcomes on one side of the ticket, which, as before, determine the outcomes on the other side. Since it is totally random which side of the ticket goes to Alice and which side goes to Bob, it suffices to consider the 14 ticket types labeled (i) through (xiv) shown in Figure 4.6 (to number these we used a tree structure similar to the one in Figure 2.11).

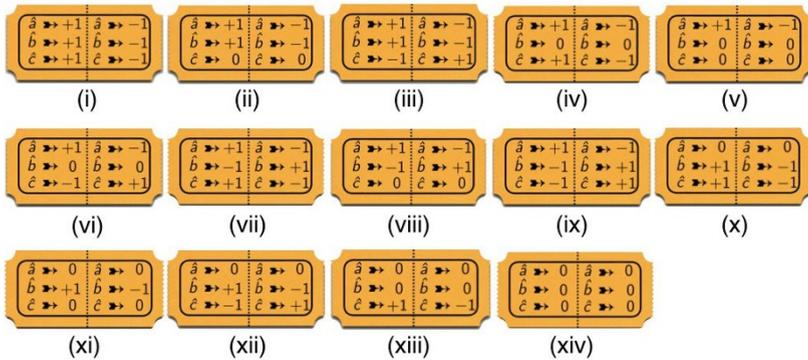


Fig. 4.6 The fourteen different types of raffle tickets for three settings and three outcomes that are left once the condition is imposed that, when Alice and Bob use the same setting, they should find opposite results if the outcome is $+1$ or -1 but the same result if the outcome is 0.

A generic raffle is some mix of these 14 ticket types. Geometrically this corresponds to a point in the 13D simplex, defined as the convex hull of the 14 points corresponding to the standard unit basis vectors in \mathbb{R}^{14} (cf. Eqs. (4.2.1)–(4.2.4), with the index k now running from i to xiv). The diagram in Figure 4.7 provides a flow chart for how to get from this simplex in \mathbb{R}^{14} to a polyhedron

in \mathbb{R}^3 characterizing the class of quantum correlations found in measurements on the singlet state of two spin-1 particles that can be simulated—albeit, as we will see, imperfectly—with these raffles.

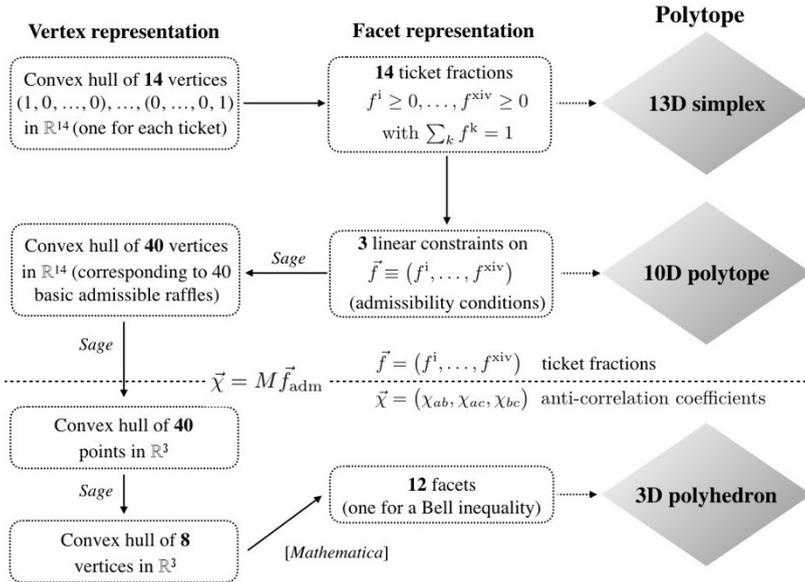


Fig. 4.7 Flow chart illustrating the construction of the polyhedron characterizing the set of correlations that (a) can be generated with raffles mixing tickets of types (i) through (xiv) in Figure 4.6 and (b) replicate key features of the quantum correlations for measurements on the singlet state of two spin-1 particles.

A polytope can be represented either in terms of its vertices or in terms of its facets. These are called the *V-representation* and the *H-representation*, respectively. The *H-representation* is given in terms of a set of inequalities restricting points to be on one side of some (hyper-)plane (hence the “*H*”, which stands for “half-space”). As our flow chart illustrates, we switch back and forth between these two representations as we go from the 13D simplex to the 3D polyhedron. For several computation-intensive steps we used the open-source mathematical software system SageMath.

The tickets for the spin-1 case immediately reveal one key difference with the spin- $\frac{1}{2}$ case that we already drew attention to in Chapter 3. Our raffle tickets can only have two outcomes for each setting. As soon as there are more than

two possible outcomes (and for spin s there are $2s + 1$), it is thus impossible for all outcomes to occur in equal proportion in a single-ticket raffle. Yet this is what Alice and Bob find if they use the same setting in the quantum experiment these raffles are supposed to simulate. In terms of probabilities, the problem is that for spins greater than $1/2$, single-ticket raffles—while non-signaling by construction—do not give uniform marginals, whereas measurements on two particles of arbitrary spin entangled in the singlet state do (as we showed in Section 4.1.4). The correlation array in Figure 4.8 for a single-ticket raffle illustrates the problem. The array is non-signaling but the marginals take on three different values: 0, $1/2$ and 1. The solution to this problem is to allow only mixed raffles that give uniform marginals.

		Bob									
		\hat{a}			\hat{b}			\hat{c}			
Alice		+1	0	-1	+1	0	-1	+1	0	-1	
	\hat{a}	+1	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	$\frac{1}{2}$	0
	0	0	0	0	0	0	0	0	0	0	
	-1	$\frac{1}{2}$	0	0	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	
\hat{b}	+1	$\frac{1}{2}$	0	0	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	
0	0	0	0	0	0	0	0	0	0		
-1	0	0	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0	$\frac{1}{2}$	0	
\hat{c}	+1	0	0	0	0	0	0	0	0	0	
0	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	0	1	0		
-1	0	0	0	0	0	0	0	0	0		

Fig. 4.8 Correlation array for a single-ticket raffle with tickets of type (viii) (see Figure 4.6).

Before we show how to construct such admissible raffles, we remind the reader of another property of the correlation array in Figure 4.8: the 9×9

matrix formed by its entries is symmetric. This is a direct consequence of the design of our raffles (for any number of outcomes per setting). Since the correlation array obviously stays the same if Alice and Bob swap both the ticket stubs they receive and the settings they use, its entries for any one of our raffles satisfy

$$\Pr(m_1 m_2 | \hat{a} \hat{b}) = \Pr(m_2 m_1 | \hat{b} \hat{a}). \quad (4.2.17)$$

In our raffles for the spin- $\frac{1}{2}$ case, it followed directly from the symmetry of the 6×6 matrix for the correlation array as a whole that the 2×2 matrices formed by the entries of its individual cells are also symmetric and that cells on opposite sides of the diagonal are the same (see, e.g., Figure 2.12). For the correlation arrays for the quantum correlations that we are trying to simulate with our raffles, both claims are true for arbitrary spin s (see Figure 4.2 for spin-1; recall that $\chi_{ab} = \cos \varphi_{ab} = \cos \varphi_{ba} = \chi_{ba}$). Neither one is true, however, for the cells in the correlation array for the single-ticket raffle in Figure 4.8. In four of the nine cells, the 3×3 matrices formed by its entries are not symmetric and the matrices for cells on opposite sides of the diagonal are each other's transpose.

Fortunately, these discrepancies are easily dealt with. First note that the transpose of a symmetric matrix is that matrix itself. Demanding individual cell symmetry thus suffices to ensure that cells on opposite sides of the diagonal are identical. Second, with raffles for the spin-1 case, as we will see, demanding uniform marginals suffices to ensure cell symmetry. In raffles for higher-spin cases, however, cell symmetry calls for additional admissibility conditions (see Section 4.2.3).

There is one more difference between the spin- $\frac{1}{2}$ and the spin-1 case that we want to highlight. In the spin- $\frac{1}{2}$ case, as we saw in Section 2.5, the probabilities in a cell of a correlation array are fully determined by the anti-correlation coefficient for that cell. In quantum mechanics, this remains true for spin $s > \frac{1}{2}$ even though the dependence of the probabilities on the anti-correlation coefficient is no longer linear (see, e.g., Figure 4.2 in Section 4.1 for $s = 1$). As soon as $s > \frac{1}{2}$, however, it is no longer true for the raffles meant to simulate these quantum correlations. In the spin-1 case, as we will see in Section 4.2.3, it takes two parameters to fix the entries in any off-diagonal cell of a correlation array for an admissible raffle. In these higher spin cases, the same triplet of anti-correlation coefficients will therefore in general correspond to more than one admissible raffle. This should not surprise us. We should not expect the projection from higher-dimensional polytopes of admissible raffles to anti-correlation polyhedra to be injective.

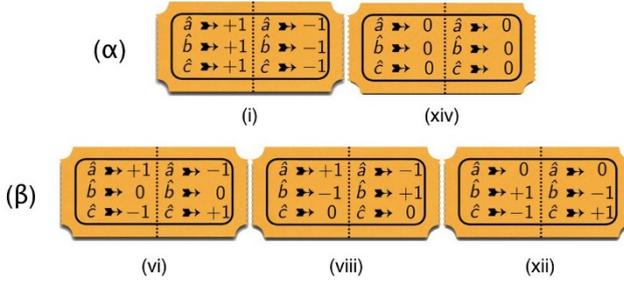


Fig. 4.9 Two mixed raffles giving uniform marginals: in (α) $2/3$ of the tickets are of type (i) and $1/3$ is of type (xiv); (β) has $1/3$ each of types (vi), (viii) and (xii).

We are now ready to start constructing raffles for the spin-1 case that give uniform marginals. Figure 4.9 shows two examples of such raffles. The first, labeled (α) and characterized by the vector

$$\mathbf{f}_\alpha = \frac{2}{3}\mathbf{f}_i + \frac{1}{3}\mathbf{f}_{xiv}, \quad (4.2.18)$$

is probably the easiest way to ensure uniform marginals. It uses tickets with the same outcomes for all three settings. With two tickets of type (i) and one of type (xiv), all three outcomes occur six times for all three settings. That means that, for all three settings, all three outcomes will be found with equal probability by both Alice and Bob. In the raffle labeled (β) and characterized by the vector

$$\mathbf{f}_\beta = \frac{1}{3}(\mathbf{f}_{vi} + \mathbf{f}_{viii} + \mathbf{f}_{xii}), \quad (4.2.19)$$

the outcomes for setting \hat{a} are the same as those in raffle (α) while the outcomes for settings \hat{b} and \hat{c} are permutations of those in raffle (α) . These permutations were chosen so as to make the sum $X_a + X_b + X_c$ vanish on both sides of all three tickets. As we saw in Chapter 3, this means that the (anti-)correlations produced by this raffle are represented by a point on the ellipsope in Figure 2.16.

Figure 4.10 shows the correlation array for the mixed raffle (β) . Unlike the correlation array in Figure 4.8 (for a single-ticket raffle), it has uniform marginals, as we demanded. The diagonal cells in correlation arrays for raffles that give uniform marginals are all the same. The standard deviations σ_a , σ_b and σ_c for such raffles are equal to $\sigma_{s=1} = \sqrt{2/3}$, in accordance with the general formula $\sigma_s^2 = \frac{1}{3}s(s+1)$ (see Eq. (3.3.3) in Chapter 3). This is not true

Bob \ Alice		\hat{a}			\hat{b}			\hat{c}		
		+1	0	-1	+1	0	-1	+1	0	-1
\hat{a}	+1	0	0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{6}$	0
	0	0	$\frac{1}{3}$	0	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{6}$
	-1	$\frac{1}{3}$	0	0	0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{6}$
\hat{b}	+1	$\frac{1}{6}$	$\frac{1}{6}$	0	0	0	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{6}$	0
	0	$\frac{1}{6}$	0	$\frac{1}{6}$	0	$\frac{1}{3}$	0	$\frac{1}{6}$	0	$\frac{1}{6}$
	-1	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$	0	0	0	$\frac{1}{6}$	$\frac{1}{6}$
\hat{c}	+1	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{6}$	0	0	0	$\frac{1}{3}$
	0	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{6}$	0	$\frac{1}{3}$	0
	-1	0	$\frac{1}{6}$	$\frac{1}{6}$	0	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{3}$	0	0

Fig. 4.10 Correlation array for a raffle with equal numbers of tickets of types (vii), (viii) and (xii) (see Figure 4.9, (β)).

for single-ticket raffles or raffles with an arbitrary mix of ticket types. This provides another reason for restricting ourselves to raffles that give uniform marginals. The projection from \mathbb{R}^{14} to \mathbb{R}^3 we are after only works if we can assume that the diagonal cells of the correlation arrays and thus the standard deviations σ_a , σ_b and σ_c are the same for all raffles considered.

With the help of the correlation array in Figure 4.10, we find that

$$\langle X_a^A X_b^B \rangle_{\mathbf{f}_\beta} = \Pr(11|\hat{a}\hat{b})_{\mathbf{f}_\beta} + \Pr(-1-1|\hat{a}\hat{b})_{\mathbf{f}_\beta} = \frac{1}{3}. \tag{4.2.20}$$

Inserting Eq. (4.2.20) and $\sigma_a^A \sigma_b^A = \sigma_{s=1}^2 = 2/3$ into Eq. (4.2.5), we arrive at

$$\chi_{ab}|_{\mathbf{f}_\beta} = -\frac{\langle X_a^A X_b^B \rangle_{\mathbf{f}_\beta}}{\sigma_a^A \sigma_b^B}|_{\mathbf{f}_\beta} = -1/2. \tag{4.2.21}$$

We similarly find that $\chi_{ac}|_{\mathbf{f}_\beta} = \chi_{bc}|_{\mathbf{f}_\beta} = -1/2$. The raffle (β) is thus represented by the point

$$(\chi_{ab}, \chi_{ac}, \chi_{bc})|_{\mathbf{f}_\beta} = (-1/2, -1/2, -1/2) \quad (4.2.22)$$

on the ellipsope, where the sum $\chi_{ab} + \chi_{ac} + \chi_{bc}$ has its absolute minimum value of $-3/2$ (cf. Chapter 3). With raffle (β) we have thus constructed a toy example of a local hidden-variable theory with which we can reach the Tsirelson bound for this setup. This raffle, however, still fails to reproduce all features of the correlation array for the corresponding quantum case.

In quantum mechanics, Eq. (4.2.22) holds for the correlations found in measurements on the singlet state of two particles of arbitrary (half-)integer spin s if each of the angles φ_{ab} , φ_{ac} and φ_{bc} between the three measuring directions is equal to 120° . In that case the anti-correlation coefficients χ_{ab} , χ_{ac} and χ_{bc} are all equal to $\cos 120^\circ = -1/2$. By demanding uniform marginals, we made sure that the diagonal cells of the correlation array for raffle (β) in Figure 4.10 are the same as those of the quantum correlation array for $s = 1$ we are trying to simulate. The six off-diagonal cells, however, while identical to each other, differ from the six identical off-diagonal cells in this quantum correlation array. Below we put the entries of two of these off-diagonal cells side-by-side, for raffle (β) on the left, for the quantum correlations on the right (we obtain the latter by substituting $\chi_{ab} = -1/2$ in the $\hat{a}\hat{b}$ cell in Figure 4.2 in Section 4.1):

$$\begin{aligned} \text{Raffle } (\beta): \quad & \frac{1}{6} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \frac{1}{48} \begin{pmatrix} 8 & 8 & 0 \\ 8 & 0 & 8 \\ 0 & 8 & 8 \end{pmatrix}, \\ \text{Quantum:} \quad & \frac{1}{48} \begin{pmatrix} 9 & 6 & 1 \\ 6 & 4 & 6 \\ 1 & 6 & 9 \end{pmatrix}. \end{aligned} \quad (4.2.23)$$

In both these cells, rows and columns sum to $1/3$ and the anti-correlation coefficient is equal to $-1/2$. Both cells are symmetric, persymmetric and centrosymmetric. In the full correlation array, $\chi_{ab} + \chi_{ac} + \chi_{bc} = -3/2$ in both cases. Yet, despite reproducing all these features of the quantum correlations, our raffle still fails to fully simulate the quantum correlations. As Eq. (4.2.23) shows, it is impossible in our raffle for Alice and Bob to find opposite results

when using different measurement settings (that would be incompatible with $X_a^A + X_b^A + X_c^A = 0$), whereas in measurements on the singlet state of two spin-1 particles there is a small probability that they do: $1/48$ for one of them finding $+1$ and the other one finding -1 ; $1/12$ for both of them finding 0 . Finding one such outcome in an actual experiment would thus disprove our local hidden-variable theory even though this theory does not put a tighter bound on the sum of the anti-correlation coefficients than quantum mechanics!

Raffles (α) and (β) are just two examples of raffles that give uniform marginals. We now determine what conditions the vector \mathbf{f} for some mixed raffle has to satisfy to ensure uniform marginals. As noted above, in the spin-1 case this is the only requirement for a raffle to be *admissible*. Consider the vector giving the ticket fractions for such an admissible raffle:

$$\mathbf{f}_{\text{adm}} = \sum_{k=1}^{\text{xiv}} f_{\text{adm}}^k \mathbf{f}_k. \quad (4.2.24)$$

To ensure that some mixed raffle gives uniform marginals it suffices to require the diagonal cells in its correlation array to have the form

$$\begin{pmatrix} 0 & 0 & 1/3 \\ 0 & 1/3 & 0 \\ 1/3 & 0 & 0 \end{pmatrix}. \quad (4.2.25)$$

Since the correlations produced by our raffles are, by construction, non-signaling, this will automatically take care of the off-diagonal cells. In principle, we thus need to impose the following nine conditions

$$\Pr(mm|\hat{a}\hat{a}) = \Pr(mm|\hat{b}\hat{b}) = \Pr(mm|\hat{c}\hat{c}) = 1/3, \quad (4.2.26)$$

for $m = 0, \pm 1$. However, we know from normalization and centrosymmetry that

$$\begin{aligned} 1 &= \sum_{m_1, m_2} \Pr(m_1 m_2 | \hat{a} \hat{a}) \\ &= \Pr(11|\hat{a}\hat{a}) + \Pr(00|\hat{a}\hat{a}) + \Pr(-1-1|\hat{a}\hat{a}) \\ &= 2\Pr(11|\hat{a}\hat{a}) + \Pr(00|\hat{a}\hat{a}), \end{aligned} \quad (4.2.27)$$

and similarly for the diagonal cells with settings $\hat{b}\hat{b}$ and $\hat{c}\hat{c}$. Hence it suffices to impose

$$\Pr(00|\hat{a}\hat{a}) = \Pr(00|\hat{b}\hat{b}) = \Pr(00|\hat{c}\hat{c}) = 1/3. \quad (4.2.28)$$

These probabilities can be expressed in terms of ticket fractions (exactly how can be seen upon inspection of the tickets in Figure 4.6):

$$\begin{aligned} \Pr(00|\hat{a}\hat{a}) &= f^x + f^{xi} + f^{xii} + f^{xiii} + f^{xiv}, \\ \Pr(00|\hat{b}\hat{b}) &= f^{iv} + f^v + f^{vi} + f^{xiii} + f^{xiv}, \\ \Pr(00|\hat{c}\hat{c}) &= f^{ii} + f^v + f^{viii} + f^{xi} + f^{xiv}. \end{aligned} \quad (4.2.29)$$

The admissibility conditions in this case thus boil down to three linear constraints on the ticket fractions (f^i, \dots, f^{xiv}). These will restrict the 13D simplex of raffles to a 10D convex polytope of admissible raffles. We used SageMath to compute its vertices. This yields 40 vertices in \mathbb{R}^{14} , which we do not reproduce for reasons of space. These corresponds to 40 *basic admissible raffles* (i.e., any admissible raffle can be obtained by mixing these 40).

Like the 13D simplex of arbitrary mixed raffles, the 10D polytope of admissible raffles cannot be visualized as such. As we did in the spin- $\frac{1}{2}$ case (see Eqs. (4.2.5)–(4.2.16)), we therefore switch from ticket fractions to anti-correlation coefficients. In other words, we map the 10D polytope to a 3D polyhedron. As before (see Eq. (4.2.14)), the mapping can compactly be written as

$$\chi|_{\mathbf{f}_{\text{adm}}} = M \mathbf{f}_{\text{adm}}. \quad (4.2.30)$$

What complicates matters in this case is that M is no longer given by $\chi|_{\mathbf{f}_k}$, as it was in the spin- $\frac{1}{2}$ case (see Eq. (4.2.12)). This is because the standard deviations σ_a , σ_b and σ_c are different for different single-ticket raffles in the spin-1 case. However, the components of $\chi|_{\mathbf{f}_{\text{adm}}}$ can still be written as a sum of covariances for single-ticket raffles. The ab component, for instance, can be written as (cf. Eq. (4.2.12)):

$$\chi_{ab}|_{\mathbf{f}_{\text{adm}}} = -\frac{1}{\sigma_{s=1}^2} \sum_{k=1}^{xiv} f_{\text{adm}}^k \langle X_a^A X_b^B \rangle|_{\mathbf{f}_k}. \quad (4.2.31)$$

Similar results hold for the other components of $\chi|_{\mathbf{f}_{\text{adm}}}$. Comparing Eq. (4.2.30) and Eq. (4.2.31) and using that $\sigma_{s=1}^2 = 2/3$, we see that the components of M in this case are given by

$$M_{ab}^k = -\frac{3}{2} \langle X_a^A X_b^B \rangle \Big|_{\mathbf{f}_k}, \quad \text{with } k = \text{i} \dots \text{xiv}, \quad (4.2.32)$$

and similar expressions with ab replaced by ac or bc . The covariances for single-ticket raffles on the right-hand side of these expressions are collected in Table 4.1.

ticket	$\langle X_a^A X_b^B \rangle$	$\langle X_a^A X_c^B \rangle$	$\langle X_b^A X_c^B \rangle$
(i) _[+++]	-1	-1	-1
(ii) _[++0]	-1	0	0
(iii) _[++-]	-1	1	1
(iv) _[+0+]	0	-1	0
(v) _[+00]	0	0	0
(vi) _[+0-]	0	1	0
(vii) _[+-+]	1	-1	1
(viii) _[+-0]	1	0	0
(ix) _[+--]	1	1	-1
(x) _[0++]	0	0	-1
(xi) _[0+0]	0	0	0
(xii) _[0+-]	0	0	1
(xiii) _[00+]	0	0	0
(xiv) _[000]	0	0	0

Table 4.1 Covariances for single-ticket raffles with tickets (i)–(xiv) in Figure 4.6. The subscript on each ticket number gives the values for the settings \hat{a} , \hat{b} and \hat{c} on the left side of a ticket of that type.

Using this table and Eq. (4.2.32) we can write out the 14×3 matrix M in Eq. (4.2.30). Rows in the table multiplied by $-3/2$ turn into columns of M :

$$M = -\frac{3}{2} \begin{pmatrix} -1 & -1 & -1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & -1 & -1 & 0 & 1 & 0 & 0 \end{pmatrix}. \quad (4.2.33)$$

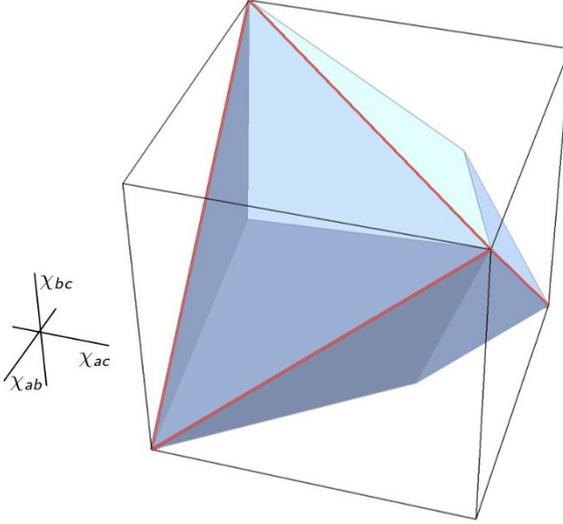


Fig. 4.11 Anti-correlation polyhedron for raffles simulating quantum correlations for spin-1 (cf. the classical tetrahedron in Figure 2.15).

To illustrate how this mapping works, consider the admissible raffle (β) introduced above. This raffle is represented by the point

$$(0, 0, 0, 0, 0, f_{\beta}^{\text{vi}}, 0, f_{\beta}^{\text{viii}}, 0, 0, 0, f_{\beta}^{\text{xii}}, 0, 0) \quad (4.2.34)$$

of the 13D simplex, with $f_{\beta}^{\text{vi}} = f_{\beta}^{\text{viii}} = f_{\beta}^{\text{xii}} = 1/3$ (see Eq. (4.2.19)). We designed this raffle so that it is represented by the point $(-1/2, -1/2, -1/2)$ of the anti-correlation polyhedron (see Eq. (4.2.22)). We verify that M correctly maps the point in Eq. (4.2.34) to the point in Eq. (4.2.22). For $\mathbf{f}_{\text{adm}} = \mathbf{f}_{\beta}$, the mapping in Eq. (4.2.30) reduces to:

$$\left. \begin{pmatrix} \chi_{ab} \\ \chi_{ac} \\ \chi_{bc} \end{pmatrix} \right|_{\mathbf{f}_{\beta}} = \begin{pmatrix} M_{ab}^{\text{vi}} & M_{ab}^{\text{viii}} & M_{ab}^{\text{xii}} \\ M_{ac}^{\text{vi}} & M_{ac}^{\text{viii}} & M_{ac}^{\text{xii}} \\ M_{bc}^{\text{vi}} & M_{bc}^{\text{viii}} & M_{bc}^{\text{xii}} \end{pmatrix} \left. \begin{pmatrix} f^{\text{vi}} \\ f^{\text{viii}} \\ f^{\text{xii}} \end{pmatrix} \right|_{\mathbf{f}_{\beta}}. \quad (4.2.35)$$

Using Eq. (4.2.33) for the matrix elements of M and setting these three components of \mathbf{f}_{β} equal to $1/3$, we confirm that

$$\left. \begin{pmatrix} \chi_{ab} \\ \chi_{ac} \\ \chi_{bc} \end{pmatrix} \right|_{\mathbf{f}_\beta} = \begin{pmatrix} 0 & -3/2 & 0 \\ -3/2 & 0 & 0 \\ 0 & 0 & -3/2 \end{pmatrix} \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \\ -1/2 \end{pmatrix}. \quad (4.2.36)$$

More generally, Eq. (4.2.30) maps the 10D polytope of admissible raffles to the 3D anti-correlation polyhedron in Figure 4.11. This polyhedron is obtained by projecting the 40 vertices MathSage found for us and taking their convex hull. Once again using MathSage, we found that of the 40 points so projected, only 8 are vertices of the polyhedron. Four of them are the vertices of the tetrahedron in Figure 2.15. The other four are the points $(\pm 1/2, \pm 1/2, \pm 1/2)$ in which an odd number of minus signs occur. Raffle (β) is represented by the one with three minus signs (see Eqs. (4.2.22) and (4.2.36)). To construct raffles represented by the three other points, we change the sign of the values ± 1 for one of the three settings for all three ticket types in raffle (β) . If this results in a ticket that is not among the fourteen tickets in Figure 4.6, we simply switch the left and the right side of the ticket. The proportions of the tickets are kept the same. If we do this for setting \hat{a} , we get a new raffle (β') (with tickets of type (ii), (iv) and (xii)) for which

$$\langle X_a^A X_b^B \rangle|_{\mathbf{f}_{\beta'}} = -\langle X_a^A X_b^B \rangle|_{\mathbf{f}_\beta} = 1/2 \quad (4.2.37)$$

$$\langle X_a^A X_c^B \rangle|_{\mathbf{f}_{\beta'}} = -\langle X_a^A X_c^B \rangle|_{\mathbf{f}_\beta} = 1/2 \quad (4.2.38)$$

$$\langle X_b^A X_c^B \rangle|_{\mathbf{f}_{\beta'}} = \langle X_b^A X_c^B \rangle|_{\mathbf{f}_\beta} = -1/2 \quad (4.2.39)$$

If we do the same thing for setting \hat{b} , we get a second new raffle (β'') (with tickets of type (ii), (vi) and (x)), in which $\langle X_a^A X_c^B \rangle$ will be the same as in raffle (β) while the other two change sign. Finally, if do this for setting \hat{c} , we get a third new raffle (β''') (with tickets of type (iv), (viii) and (x)), in which $\langle X_a^A X_b^B \rangle$ will be the same as in raffle (β) and the other two change sign. In the polyhedron in Figure 4.11, the raffles (β') , (β'') and (β''') are thus represented by the points $(1/2, 1/2, -1/2)$, $(1/2, -1/2, 1/2)$ and $(-1/2, 1/2, 1/2)$, respectively.

4.2.3 Spin- $\frac{3}{2}$

The flow chart in Figure 4.7 that we used to deal with raffles for the spin-1 case will also guide us in our analysis of raffles for the spin- $\frac{3}{2}$ case. The number of tickets now jumps from $(3^3 + 1)/2 = 14$ to $4^3/2 = 32$. We will only display a

subset of these tickets (see Figures 4.14 and 4.15). We numbered them using the same convention as before (cf. the tree structure in Figure 2.11 in Section 2.5). With 32 tickets, the number of vertices and facets to keep track of is starting to get unwieldy but the spin- $\frac{3}{2}$ case is still tractable. And there are at least two good reasons for examining it in some detail: it is the simplest case in which cell symmetry calls for separate admissibility conditions; it nicely illustrates the difference between integer and half-integer spin cases in terms of the bound on the sum of the anti-correlation coefficients $\chi_{ab} + \chi_{bc} + \chi_{ac}$ (see Chapter 3, Eqs. (3.3.6)–(3.3.9)).

		Bob		
		\hat{b}		
Alice		1	0	-1
\hat{a}	1	α	β	γ
	0	δ	ε	δ
	-1	γ	β	α

		Bob			
		\hat{b}			
Alice		$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$
\hat{a}	$\frac{3}{2}$	α	β	γ	δ
	$\frac{1}{2}$	ε	ζ	η	ϑ
	$-\frac{1}{2}$	ϑ	η	ζ	ε
	$-\frac{3}{2}$	δ	γ	β	α

Fig. 4.12 Cell in the correlation arrays for the spin-1 and spin- $\frac{3}{2}$ cases showing centrosymmetry.

Figure 4.12 shows a generic off-diagonal cell in the correlation array for two of our raffles, the one on the left for the spin-1 case, the one on the right for the spin- $\frac{3}{2}$ case. As noted in the introduction to Chapter 4, the design of our raffles guarantees that these cells are centrosymmetric (see Eq. (4.0.3)). Hence, 5 parameters (labeled α through ε) suffice to fix the 9 entries in the cell on the left and 8 parameters (labeled α through ϑ) suffice to fix the 16 entries in the one on the right (these parameters have to satisfy the obvious requirement that the sum of all entries in a cell equals 1).

Requiring uniform marginals further reduces the number of parameters needed to determine the entries in the cells in Figure 4.12. The conditions ensuring uniform marginals in the spin- $\frac{3}{2}$ case are a straightforward generalization of those in the spin-1 case (cf. Eq. (4.2.28)):

$$\begin{aligned}
 \Pr\left(\frac{3}{2} - \frac{3}{2} \mid \hat{a}\hat{a}\right)_{\mathbf{f}_{\text{adm}}} &= \Pr\left(\frac{1}{2} - \frac{1}{2} \mid \hat{a}\hat{a}\right)_{\mathbf{f}_{\text{adm}}} = \\
 \Pr\left(-\frac{1}{2} \frac{1}{2} \mid \hat{a}\hat{a}\right)_{\mathbf{f}_{\text{adm}}} &= \Pr\left(-\frac{3}{2} \frac{3}{2} \mid \hat{a}\hat{a}\right)_{\mathbf{f}_{\text{adm}}} = \frac{1}{4},
 \end{aligned} \tag{4.2.40}$$

and similarly for the diagonal cells with settings $\hat{b}\hat{b}$ and $\hat{c}\hat{c}$. As in the spin-1 case, it suffices to impose one of these conditions for each diagonal cell:

$$\Pr\left(\frac{3}{2} - \frac{3}{2} \middle| \hat{a}\hat{a}\right) \Big|_{\mathbf{f}_{\text{adm}}} = \Pr\left(\frac{3}{2} - \frac{3}{2} \middle| \hat{b}\hat{b}\right) \Big|_{\mathbf{f}_{\text{adm}}} = \Pr\left(\frac{3}{2} - \frac{3}{2} \middle| \hat{c}\hat{c}\right) \Big|_{\mathbf{f}_{\text{adm}}} = \frac{1}{4}. \quad (4.2.41)$$

Once the conditions for uniform marginals have been imposed, the entries in the cell for the spin-1 case will satisfy

$$\alpha + \beta + \gamma = \alpha + \delta + \gamma = 2\beta + \varepsilon = 1/3. \quad (4.2.42)$$

It follows that

$$\gamma = 1/3 - \alpha - \beta, \quad \delta = \beta, \quad \varepsilon = 1/3 - 2\beta. \quad (4.2.43)$$

It thus only takes two independent parameters, α and β , to fix all entries in the cell.

Note that this is still one more parameter than is needed to fix all entries in a cell in a correlation array for measurements on the singlet state of two particles of arbitrary (half-)integer spin s . The entries in those cells are fixed by (a highly non-linear function of) a single parameter, the angle φ_{ab} between the measuring directions. It thus need not surprise us that, for $s \geq 1$, we can no longer perfectly simulate the quantum correlations (see Eq. (4.2.23)).

Inserting the expressions for γ , δ and ε in Eq. (4.2.43) in the cell for the spin-1 case in Figure 4.12, we see that this cell is now symmetric as well as centro-symmetric. As we noted in Section 4.2.2, requiring uniform marginals thus automatically ensures cell symmetry in the spin-1 case.

In the spin- $\frac{3}{2}$ case, this is no longer true. Instead of Eq. (4.2.42), the requirement of uniform marginals now gives

$$\alpha + \beta + \gamma + \delta = \alpha + \varepsilon + \vartheta + \delta = 1/4, \quad (4.2.44)$$

from which it follows only that $\beta + \gamma = \varepsilon + \vartheta$. To ensure cell symmetry we need to impose an extra condition. We will set $\beta = \varepsilon$. Since Eq. (4.2.44) then entails $\gamma = \vartheta$, this suffices to make the cell on the right in Figure 4.12 is symmetric.

The condition $\beta = \varepsilon$ translates into

$$\Pr\left(\frac{3}{2} \frac{1}{2} \middle| \hat{a}\hat{b}\right) \Big|_{\mathbf{f}_{\text{adm}}} = \Pr\left(\frac{1}{2} \frac{3}{2} \middle| \hat{a}\hat{b}\right) \Big|_{\mathbf{f}_{\text{adm}}}. \quad (4.2.45)$$

We need three such conditions, one for each of the three off-diagonal cells in the correlation array (recall that the cell symmetry and the symmetry of the correlation array as a whole guarantee that cells on opposite sides of the diagonal are the same). The probabilities in these admissibility conditions are given by half the sum of the ticket fractions f_{adm}^k for those tickets that have the relevant combination of outcomes (cf. Eqs. (4.2.7)–(4.2.8)). Like the conditions for uniform marginals, these symmetry conditions thus take the form of linear constraints on the ticket fractions f^k (with k now running from i through xxxii).

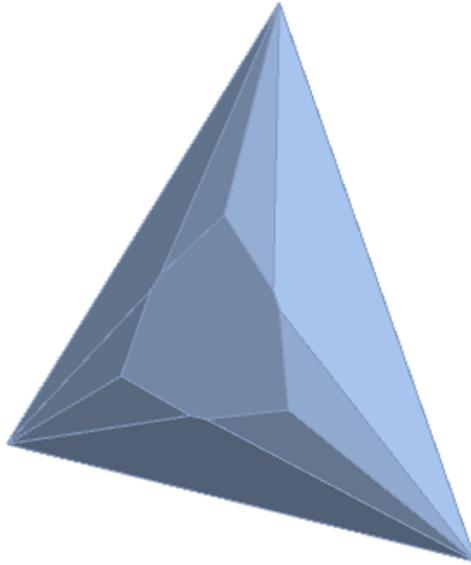


Fig. 4.13 Facet of the anti-correlation polyhedron for $\text{spin-}\frac{3}{2}$. Imagine this figure ‘glued onto’ the four facets of the classical tetrahedron in Figure 2.15.

For the $\text{spin-}\frac{3}{2}$ case we have a total of 6 admissibility conditions, 3 to ensure uniform marginals, 3 to ensure cell symmetry. Since there are 32 different tickets, the polytope we start from is the 31D standard simplex in \mathbb{R}^{32} (cf. the flow chart in Figure 4.7). Upon imposing our 6 admissibility conditions, we arrive at a 25D polytope of admissible raffles in \mathbb{R}^{32} . With the help of Sagemath we find that this polytope has a total of 450 (!) vertices. We can map this polytope in \mathbb{R}^{32} to the 3D anti-correlation polyhedron using a 32×3 matrix M with elements of the same form as those of the 14×3 matrix for

raffles in the spin-1 case (see Eqs. (4.2.30)–(4.2.33) and Table 4.1). For our purposes we only need a subset of the 96 elements of this matrix (see Tables 4.2 and 4.3). Applying the mapping given by the matrix M to the 450 vertices of our 25D polytope and using Sagemath to determine which of the resulting 450 points are the vertices of its 3D image, we found a polyhedron in \mathbb{R}^3 with 40 vertices. We then used the program Mathematica to find the facets of this polyhedron and generate the picture in Figure 4.13. This picture shows the facets we need to “glue onto” each of the four facets of the tetrahedron in Figure 2.15 to get the full anti-correlation polyhedron in this case.

The polyhedron for raffles (imperfectly) simulating the quantum correlations in the spin- $\frac{3}{2}$ case picks up 36 extra vertices compared to the tetrahedron for raffles for the spin- $\frac{1}{2}$ case, 9 for each facet of the tetrahedron. Figure 4.13 shows those 9 extra vertices for one of these four facets, 6 of them in pairs that lie so close together that it may look as if there are only 6 rather than 9 points. All 9 vertices lie in the same plane. In the case of the facet of the tetrahedron closest to the point $(-1, -1, -1)$ of the non-signaling cube (see Figure 2.15) this is the plane where the sum $\chi_{ab} + \chi_{ac} + \chi_{bc}$ has its minimum value. Eq. (3.3.8) in Chapter 3 tells us that the minimum value of this quantity for $s = 3/2$ is

$$\frac{1}{8\sigma_{3/2}^2} - \frac{3}{2} = \frac{1}{10} - \frac{3}{2} = -\frac{7}{5}. \quad (4.2.46)$$

where we used Eq. (3.3.3) to set $\sigma_{3/2}^2 = 5/4$.

With a little help from the computer, we were able to construct raffles represented by these 9 vertices. Figures 4.14 and 4.15 show the mix of tickets for two of them, labeled (μ) and (ν) . We obtain raffles represented by the other seven vertices by permutations of the outcomes for the three settings \hat{a} , \hat{b} and \hat{c} on all tickets in these two raffles (switching left and right sides of tickets if necessary). By changing the sign of the values for one of the settings on all tickets in these nine raffles (again switching sides of tickets if necessary), we can construct raffles for the 27 vertices obtained if the facets in Figure 4.13 are “glued on” to the other three facets of the tetrahedron (this is the same procedure that we followed for the spin-1 case in Section 4.2.2).

Consider Figure 4.14 for raffle (μ) characterized by the vector

$$\mathbf{f}_\mu = \frac{1}{4}(\mathbf{f}_{xii} + \mathbf{f}_{xv} + \mathbf{f}_{xxiv} + \mathbf{f}_{xxx}). \quad (4.2.47)$$

It is easy to see that this raffle yields uniform marginals: For all three settings, all four outcomes occur twice in every set of these four tickets. It

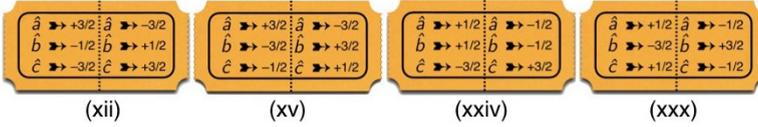


Fig. 4.14 Admissible raffle (μ) for the spin- $\frac{3}{2}$ case with equal numbers of tickets of type (xii), (xv), (xxiv) and (xxx).

is also easy to see that it will be represented by a point on the plane where $\chi_{ab} + \chi_{ac} + \chi_{bc}$ has its minimum value: On both sides of all four tickets, the sum of outcomes for the three settings is $\pm 1/2$ resulting in the minimum value of $1/4$ of the expectation value of $(X_a^A + X_b^A + X_c^A)^2$ (cf. Chapter 3). To make sure that the off-diagonal cells in the correlation array for raffle (μ) are symmetric, we check that Eq. (4.2.45) is satisfied, not just for the $\hat{a}\hat{b}$ cell but for the $\hat{a}\hat{c}$ and $\hat{b}\hat{c}$ cells as well. Since

$$\begin{aligned} \Pr\left(\frac{3}{2} \frac{1}{2} \middle| \hat{a}\hat{b}\right) \Big|_{\mathbf{f}_\mu} &= \frac{1}{2} f^{\text{xii}}, & \Pr\left(\frac{1}{2} \frac{3}{2} \middle| \hat{a}\hat{b}\right) \Big|_{\mathbf{f}_\mu} &= \frac{1}{2} f^{\text{xxx}}, \\ \Pr\left(\frac{3}{2} \frac{1}{2} \middle| \hat{a}\hat{c}\right) \Big|_{\mathbf{f}_\mu} &= \frac{1}{2} f^{\text{xv}}, & \Pr\left(\frac{1}{2} \frac{3}{2} \middle| \hat{a}\hat{c}\right) \Big|_{\mathbf{f}_\mu} &= \frac{1}{2} f^{\text{xxiv}}, \\ \Pr\left(\frac{3}{2} \frac{1}{2} \middle| \hat{b}\hat{c}\right) \Big|_{\mathbf{f}_\mu} &= \frac{1}{2} f^{\text{xxx}}, & \Pr\left(\frac{1}{2} \frac{3}{2} \middle| \hat{b}\hat{c}\right) \Big|_{\mathbf{f}_\mu} &= \frac{1}{2} f^{\text{xxiv}}, \end{aligned} \quad (4.2.48)$$

and all four ticket fractions are equal, these symmetry conditions are indeed satisfied.

Using Eq. (4.2.30) for the mapping from ticket fractions of admissible raffles to triplets of allowed anti-correlation coefficients, we can find the components of χ for raffle (μ). In this case, Eq. (4.2.30) reduces to

$$\begin{pmatrix} \chi_{ab} \\ \chi_{ac} \\ \chi_{bc} \end{pmatrix} \Big|_{\mathbf{f}_\mu} = \begin{pmatrix} M_{ab}^{\text{xii}} & M_{ab}^{\text{xv}} & M_{ab}^{\text{xxiv}} & M_{ab}^{\text{xxx}} \\ M_{ac}^{\text{xii}} & M_{ac}^{\text{xv}} & M_{ac}^{\text{xxiv}} & M_{ac}^{\text{xxx}} \\ M_{bc}^{\text{xii}} & M_{bc}^{\text{xv}} & M_{bc}^{\text{xxiv}} & M_{bc}^{\text{xxx}} \end{pmatrix} \begin{pmatrix} f^{\text{xii}} \\ f^{\text{xv}} \\ f^{\text{xxiv}} \\ f^{\text{xxx}} \end{pmatrix} \Big|_{\mathbf{f}_\mu}. \quad (4.2.49)$$

The ticket fractions are all equal to $1/4$. The elements of the matrix M are given by

$$M_{ab}^k = -\frac{1}{\sigma_{s=3/2}^2} \langle X_a^A X_b^B \rangle, \quad \text{with } k = \text{i} \dots \text{xxxii}, \quad (4.2.50)$$

and similar expressions with ab replaced by ac or bc (cf. Eq. (4.2.32)). These matrix elements are given by $-1/\sigma_{s=3/2}^2 = -4/5$ times the relevant entries in Table 4.2.

ticket	$\langle X_a^A X_b^B \rangle$	$\langle X_a^A X_c^B \rangle$	$\langle X_b^A X_c^B \rangle$
(xii) _{$[\frac{3}{2} - \frac{1}{2} - \frac{3}{2}]$}	3/4	9/4	-3/4
(xv) _{$[\frac{3}{2} - \frac{3}{2} - \frac{1}{2}]$}	9/4	3/4	-3/4
(xxiv) _{$[\frac{1}{2} \frac{1}{2} - \frac{3}{2}]$}	-1/4	3/4	3/4
(xxx) _{$[\frac{1}{2} - \frac{3}{2} \frac{1}{2}]$}	3/4	-1/4	3/4

Table 4.2 Covariances for single-ticket raffles for the four ticket types in admissible raffle (μ) in Figure 4.14 (cf. Table 4.1).

For the components of χ for raffle (μ) we then find:

$$\begin{pmatrix} \chi_{ab} \\ \chi_{ac} \\ \chi_{bc} \end{pmatrix} \Big|_{\mathbf{f}_\mu} = \begin{pmatrix} -3/5 & -9/5 & 1/5 & -3/5 \\ -9/5 & -3/5 & -3/5 & 1/5 \\ 3/5 & 3/5 & -3/5 & -3/5 \end{pmatrix} \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix} = \begin{pmatrix} -7/10 \\ -7/10 \\ 0 \end{pmatrix}, \quad (4.2.51)$$

which confirms that $\chi_{ab} + \chi_{ac} + \chi_{bc} = -7/5$, as it should be given the way we constructed this raffle. Raffle (μ) is represented by one of the 9 vertices in the plane of the polyhedron where $\chi_{ab} + \chi_{ac} + \chi_{bc}$ has its minimum value. Through suitable permutation of the outcomes for settings \hat{a} , \hat{b} and \hat{c} , we can create raffles similar to raffle (μ) represented by two other vertices in that same plane: $(0, -7/10, -7/10)$ and $(-7/10, 0, -7/10)$.

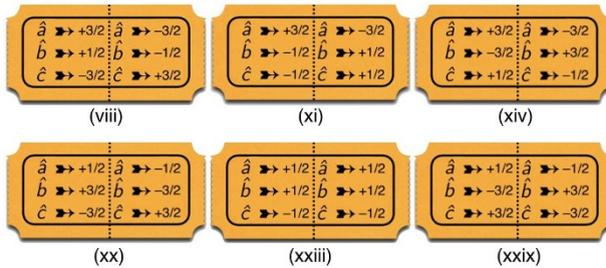


Fig. 4.15 Admissible raffle (ν) for the spin- $\frac{3}{2}$ case: a basket with equal numbers of tickets of type (viii), (xi), (xiv), (xx), (xxiii) and (xxix).

Figure 4.15 shows the tickets for a raffle represented by one of the remaining six points in the plane where $\chi_{ab} + \chi_{ac} + \chi_{bc} = -7/5$. This raffle is represented by the vector

$$\mathbf{f}_v = \frac{1}{6} (\mathbf{f}_{viii} + \mathbf{f}_{xi} + \mathbf{f}_{xiv} + \mathbf{f}_{xx} + \mathbf{f}_{xxiii} + \mathbf{f}_{xxix}), \quad (4.2.52)$$

As with raffle (μ), the outcomes of both sides of all tickets add up to $\pm 1/2$ and both sets of admissibility conditions (uniform marginals and cell symmetry) are met. These last two properties can be verified the same way as in the case of raffle (μ).

ticket	$\langle X_a^A X_b^B \rangle$	$\langle X_a^A X_c^B \rangle$	$\langle X_b^A X_c^B \rangle$
(viii) $_{[\frac{3}{2} \frac{1}{2} - \frac{3}{2}]}$	$-3/4$	$9/4$	$3/4$
(xi) $_{[\frac{3}{2} - \frac{1}{2} - \frac{1}{2}]}$	$3/4$	$3/4$	$-1/4$
(xiv) $_{[\frac{3}{2} - \frac{3}{2} \frac{1}{2}]}$	$9/4$	$-3/4$	$3/4$
(xx) $_{[\frac{1}{2} \frac{3}{2} - \frac{3}{2}]}$	$-3/4$	$3/4$	$9/4$
(xxiii) $_{[\frac{1}{2} \frac{1}{2} - \frac{1}{2}]}$	$-1/4$	$1/4$	$1/4$
(xxix) $_{[\frac{1}{2} - \frac{3}{2} \frac{3}{2}]}$	$3/4$	$-3/4$	$9/4$

Table 4.3 Covariances for single-ticket raffles for the six ticket type in admissible raffle (v) in Figure 4.15 (cf. Table 4.2).

To find the components of χ for this raffle, we once again use Eq. (4.2.30), which in this case reduces to:

$$\begin{pmatrix} \chi_{ab} \\ \chi_{ac} \\ \chi_{bc} \end{pmatrix} \Big|_{\mathbf{f}_v} = \begin{pmatrix} M_{ab}^{viii} & M_{ab}^{xi} & M_{ab}^{xiv} & M_{ab}^{xx} & M_{ab}^{xxiii} & M_{ab}^{xxix} \\ M_{ac}^{viii} & M_{ac}^{xi} & M_{ac}^{xiv} & M_{ac}^{xx} & M_{ac}^{xxiii} & M_{ac}^{xxix} \\ M_{bc}^{viii} & M_{bc}^{xi} & M_{bc}^{xiv} & M_{bc}^{xx} & M_{bc}^{xxiii} & M_{bc}^{xxix} \end{pmatrix} \begin{pmatrix} f^{viii} \\ f^{xi} \\ f^{xiv} \\ f^{xx} \\ f^{xxiii} \\ f^{xxix} \end{pmatrix} \Big|_{\mathbf{f}_v}. \quad (4.2.53)$$

Evaluating the relevant elements of the matrix M with the help of Table 4.3 and setting all ticket fractions equal to $1/6$, we arrive at

$$\left(\begin{array}{c} \chi_{ab} \\ \chi_{ac} \\ \chi_{bc} \end{array} \right) \Big|_{\mathbf{f}_v} = \begin{pmatrix} 3/5 & -3/5 & -9/5 & 3/5 & 1/5 & -3/5 \\ -9/5 & -3/5 & 3/5 & -3/5 & -1/5 & 3/5 \\ -3/5 & 1/5 & -3/5 & -9/5 & -1/5 & -9/5 \end{pmatrix} \begin{pmatrix} 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \\ 1/6 \end{pmatrix}, \quad (4.2.54)$$

i.e.,

$$\left(\begin{array}{c} \chi_{ab} \\ \chi_{ac} \\ \chi_{bc} \end{array} \right) \Big|_{\mathbf{f}_v} = \begin{pmatrix} -4/15 \\ -5/15 \\ -12/15 \end{pmatrix}. \quad (4.2.55)$$

Note that, once again, $\chi_{ab} + \chi_{ac} + \chi_{bc} = -7/5$, as it should be given the way we constructed this raffle. Through suitable permutation of the outcomes for settings \hat{a} , \hat{b} and \hat{c} , we can create raffles similar to raffle (v) represented by five other vertices of the facet of the anti-correlation polyhedron where $\chi_{ab} + \chi_{ac} + \chi_{bc} = -7/5$. Since $4/15$ and $1/3$ only differ by $1/15$, these six vertices can naturally be grouped into three pairs of neighboring points that are hard to tell apart in Figure 4.13:

$$\begin{aligned} & \{ (-4/15, -1/3, -4/5), (-1/3, -4/15, -4/5) \}, \\ & \{ (-4/5, -4/15, -1/3), (-4/5, -1/3, -4/15) \}, \\ & \{ (-4/15, -4/5, -1/3), (-1/3, -4/5, -4/15) \}. \end{aligned} \quad (4.2.56)$$

The analysis of raffles (μ) and (v) allowed us to understand the structure of the facets in Figure 4.13 in full detail. For higher-spin cases, this becomes impractical. In the next subsection, we will therefore explore alternative methods for dealing with these higher-spin cases.

4.2.4 Spin- s ($s \geq 2$)

To simulate the correlations found by measurements on a singlet state of two spin-2 particles, we use tickets with five outcomes $\pm 2, \pm 1, 0$. This results in 63 relevant ticket types, corresponding to a 62-dimensional simplex of raffles. The conditions for uniform marginals and cell symmetry yield six linear constraints, yielding a 50-dimensional polytope of admissible raffles in 63-dimensional space. Unfortunately, the number of vertices determined

by these conditions is daunting: If we consider only the six conditions for uniform marginals, computer calculations produce 553,664 vertices. This is already orders of magnitude larger than the vertex sets for the earlier polytopes. The full case, obtained by including the six conditions for cell symmetry, is presumably even larger but we have not been able to run the vertex enumeration algorithm to completion on a personal computer. Hence we do not know how many vertices the spin-2 admissible polytope has, much less the full list of such. This computational obstacle only grows worse as the spin increases, as is evident from the numbers in Figure 4.16. The enumeration of basic admissible raffles thus becomes intractable for spin $s \geq 2$. As a consequence, the flowchart in Figure 4.7 comes to a halt and we cannot hope to compute the anti-correlation polyhedron in the way it depicts.

	# tickets	admissibility conditions		admissible raffle polytope		correlation polyhedron	
		uniform marginals	cell symmetry	dimension	# vertices	# vertices	# facets
spin-1/2	4	0	0	3	4	4	4
spin-1	14	3	0	10	40	8	12
spin-3/2	32	3	3	25	450	40	40
spin-2	63	6	6	50	?	116	120
spin-5/2	108	6	12	89	?	424	520

Fig. 4.16 Number of ticket types, admissibility conditions, and facet/vertex counts for polytopes and polyhedra up to spin- $\frac{5}{2}$.

Fortunately, there is an alternative approach which—while it does not yield the full polytope—does allow us to characterize the polyhedron of anti-correlation coefficients. As a first step we observe that, for every spin considered so far, the polyhedron of anti-correlation coefficients always at least contained the spin- $\frac{1}{2}$ classical tetrahedron. This is not a coincidence. For instance, consider the point $\chi = (1, 1, 1)$ where all outcomes are perfectly anti-correlated. There is actually a simple construction of an admissible spin- s raffle with such behavior. To do so, it is convenient to introduce the following notation for the remainder of this section: Let $[m_a, m_b, m_c]$ denote the ticket

type for which the outcomes m_a, m_b, m_c for settings $\hat{a}, \hat{b}, \hat{c}$ appear on one side. As usual, we regard $[-m_a, -m_b, -m_c]$ as equivalent to $[m_a, m_b, m_c]$.

With this notation in mind, consider the spin- s raffle (with $2s + 1$ outcomes $m = -s$ to s) containing one of each of the tickets

$$[s, s, s], [s - 1, s - 1, s - 1], \dots, [-s + 1, -s + 1, -s + 1], [-s, -s, -s].$$

Since each of these ticket types appears once, the raffle has uniform marginals. If we swap any two settings on every ticket, all tickets are unchanged and therefore the raffle is cell-symmetric. Hence this raffle is admissible. As every outcome on one side of each ticket is strictly anti-correlated with every outcome on the other side, we obtain $\chi_{ab} = \chi_{ac} = \chi_{bc} = 1$.

Thus for any spin there is an admissible raffle for which $\chi = (1, 1, 1)$. The other three points then follow by symmetry: We take each ticket type and swap all outcomes for a given setting (e.g., take the outcomes for setting \hat{a} to range from $-s$ to $+s$). This generates three more admissible raffles, characterized by $\chi = (-1, -1, 1), (-1, 1, -1), (1, -1, -1)$, respectively. Hence all four vertices of the classical tetrahedron can be produced by admissible raffles regardless of the number of outcomes. By convexity, it follows that the anti-correlation polyhedron always includes the entire classical tetrahedron.

In general, however, the classical tetrahedron is not the full polyhedron. To establish this, we again recall that the points $\chi = (-1, -1, 1), (-1, 1, -1), (1, -1, -1)$ generate a facet of the classical tetrahedron. This would correspond to the Bell inequality $\chi_{ab} + \chi_{ac} + \chi_{bc} \geq -1$. But, as we have shown for spin 1 through $\frac{3}{2}$, there exist admissible raffles that violate this inequality.

Ticket fractions	Ticket types
f^1	$[0, 0, 0]$
f^2	$[0, 1, -1] [1, 0, -1] [1, -1, 0]$
f^3	$[0, 1, -1] [1, 0, -1] [1, -1, 0]$
f^4	$[0, 1, -1] [1, 0, -1] [1, -1, 0]$

Table 4.4 Ticket types and ticket fractions for the four ticket groups used to construct admissible raffles which maximally violate the Bell inequality. Each ticket type in a given group occurs with the same frequency in the full raffle.

Suppose we now focus on finding admissible raffles which maximally violates this inequality. One approach exploits the observation made in Section 3: To violate the Bell inequality as strongly as possible, we should use ticket types which render $(X_a^A + X_b^A + X_c^A)^2$ as small as possible. For integer spin, this suggests restricting attention to those ticket types whose outcomes sum to zero on the left side (and so also the right side). In the case of spin 2, only 10 out of the 63 tickets to fulfill this criterion. Enforcing cell symmetry on this limited subset of tickets results in the four groups of ticket types shown in Table 4.4. For a raffle to be cell-symmetric requires that the fraction of tickets of a given type be the same for all members of a group. We denote these common ticket fractions as f^1, f^2, f^3, f^4 for the four respective ticket groups. Enforcing uniform marginals then yields the conditions (cf. Eqs. (4.2.25)-(4.2.29))

$$f^1 + f^2 + f^3 = f^2 + f^4 = f^3 + 1/2f^4 = 1/5.$$

These four equations in three unknowns have a one-dimensional solution set. It can then be shown that the subset of non-negative solutions is generated by convex combinations of the following:

- $(f^1, f^2, f^3, f^4) = (1/10, 0, 1/10, 2/10)$, corresponding to a raffle with tickets

$$\begin{array}{cccccc} [0, 0, 0] & [2, 0 - 2] & [1, 1, -2] & [1, -2, 1] & [2, -1, -1] \\ [0, 2, -2] & [2, -2, 0] & [1, 1, -2] & [1, -2, 1] & [2, -1, -1] \end{array}$$

- $(f^1, f^2, f^3, f^4) = (0, 1/15, 2/15, 1/15)$, corresponding to a raffle with tickets

$$\begin{array}{cccccc} [0, 1, -1] & [0, 2, -2] & [2, 0, -2] & [1, 1, -2] & [1, -2, 1] \\ [1, 0, -1] & [0, 2, -2] & [2, -2, 0] & [1, 1, -2] & [2, -1, -1] \\ [1, -1, 0] & [2, 0 - 2] & [2, -2, 0] & [1, -2, 1] & [2, -1, -1] \end{array}$$

One may confirm that, in keeping with the ticket outcomes all summing to zero, both raffles map to the point $\chi = (-1/2, -1/2, -1/2)$. We have thus gone from a 50-dimensional polytope of admissible raffles to a 1-dimensional subspace of such raffles, all of which maximally violate the Bell inequality.⁸

⁸ These raffles validate that, for spin up to 2, we can always find an admissible raffle while using only tickets which minimize the magnitude of $X_a^A + X_b^A + X_c^A$. This is actually true in general: For any spin, there is a procedure to construct an admissible raffle using only tickets which minimize $(X_a^A + X_b^A + X_c^A)^2$. This construction, however, is somewhat involved and moreover beyond the scope of the present work, so we do not pursue this line further.

We have thus obtained another vertex of our spin-2 polyhedron, one which lies beyond the facet of the classical tetrahedron opposite the point $\chi = (1, 1, 1)$. By exploiting the tetrahedral symmetry of the anti-correlation coefficients, we obtain three other such vertices that lie beyond the other three facets. Invoking convexity again, we obtain another polyhedron of admissible anti-correlation coefficients which is larger than the classical tetrahedron (though still necessarily bounded by the ellipsope) and therefore more closely ‘approximates’ the full anti-correlation polyhedron.

At this point, there is nothing in principle to stop us from pursuing the following strategy, which is essentially the convex hull algorithm of [Lassez & Lassez \(1992\)](#). We start with some ‘approximate’ polyhedron, such as the classical tetrahedron, which the true anti-correlation polyhedron contains. We pick one of the facets of the approximate polyhedron, and determine whether any admissible raffles exist which violate the corresponding linear inequality on anti-correlation coefficients. If no such raffle exists, we conclude that the facet under consideration is indeed a facet of the anti-correlation polyhedron and we move on to another facet. If such a raffle does exist, however, we determine one which maximally violates the corresponding inequality. The resulting set of anti-correlation coefficients will be an extreme point of the anti-correlation polyhedron. We then invoke convexity and enlarge our approximate polyhedron to include this new extreme point, thereby obtaining a better approximation of the anti-correlation polyhedron. This algorithm terminates when every facet of the approximate polyhedron is verified to be a facet of the anti-correlation polyhedron, at which point we conclude that the entire polyhedron has been generated.

The hardest step in this method is to determine whether any admissible raffles map to points beyond a given facet. One approach would be to imitate the strategy outlined for the case of $\chi_{ab} + \chi_{ac} + \chi_{bc}$: For each linear inequality, we determine a corresponding linear relation $v_a X_a^A + v_b X_b^A + v_c X_c^A$ and look for admissible raffles using only tickets for which this quantity has small magnitude, thereby ensuring that the facet inequality is violated as strongly as possible. Take, for example, the spin-2 case. We have shown that its anti-correlation polyhedron contains the points $(-1, 1, -1)$, $(1, -1, -1)$, and $(-1/2, -1/2, -1/2)$. The plane through these three points is given by $\chi_{ab} + \chi_{ac} + 2\chi_{bc} = -2$. We want to determine whether this gives a facet. Observing that

$$\langle (X_a^A + 2X_b^A + 2X_c^A)^2 \rangle = 18 + 8(\chi_{ab} + \chi_{ac} + 2\chi_{bc}), \quad (4.2.57)$$

we are led to consider tickets for which $(X_a^A + 2X_b^A + 2X_c^A)^2$ is small in order to minimize $\chi_{ab} + \chi_{ac} + 2\chi_{bc}$. Proceeding in this way, we ultimately obtain all admissible raffles which are characterized by the seven-sided facet at the bottom of Figure 4.17.

There are, however, several problems with this method. The first is that it is necessarily somewhat tedious: The linear relation to be minimized must be computed for each facet under consideration, and therefore will lead to different sets of tickets in each case. Second, it is not obvious how many tickets one will need to successfully produce admissible raffles. In the case of spin 2, there are 7 tickets for which this expression is zero and 9 for which it has magnitude 1; it turns out that we need to use tickets of both magnitudes to get an admissible raffle. This is still far smaller than the 63 ticket types total, but it is not nearly as attractive as the 10 tickets with outcomes summing to zero in the $\chi_{ab} + \chi_{ac} + \chi_{bc}$ case. Finally, and most importantly, our iterative strategy only needs one admissible raffle that maximally violates the relevant facet inequality. It is therefore altogether excessive to characterize the entire set of such admissible raffles for a given facet.

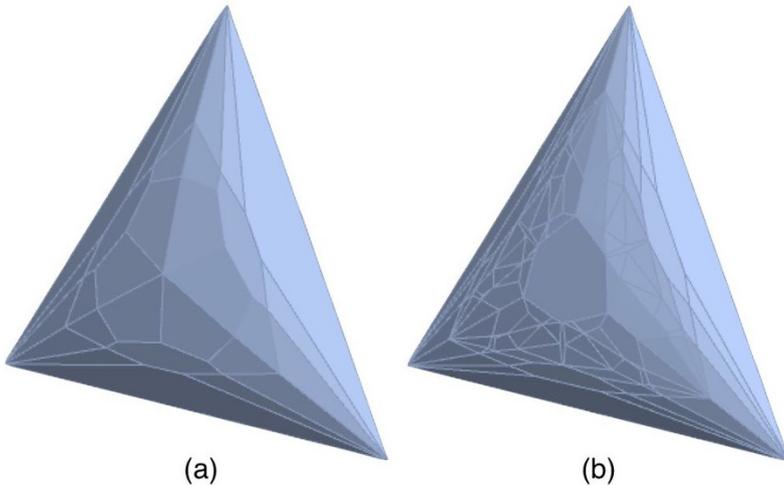


Fig. 4.17 Facet of anti-correlation polyhedron for (a) spin-2 and (b) spin- $\frac{5}{2}$ (cf. Figure 4.13).

Rather than pursue this approach further, we will instead exploit the well-known connection between polytopes and *linear programming* (Dantzig & Thapa, 1997, 2003). We consider a spin- s raffle and let Δ^{n-1} be the appropriate $(n-1)$ -dimensional simplex of raffles, where n is the number of relevant ticket

types. Since the admissibility conditions are all linear in the ticket fractions, we may conveniently express them in the form $B\mathbf{f}_{\text{adm}} = \mathbf{b}$; the matrix B and the vector \mathbf{b} will have as many rows as there are (linearly independent) constraints. Finally, the linear function to be minimized can be expressed as $\mathbf{c} \cdot \chi|_{\mathbf{f}_{\text{adm}}}$ for some real 3-vector \mathbf{c} . This *objective function* is in terms of the anti-correlation coefficients, but as before the mapping of ticket fractions to anti-correlation coefficients is expressed as $\chi|_{\mathbf{f}_{\text{adm}}} = M\mathbf{f}_{\text{adm}}$ where M is a $3 \times n$ matrix. Hence the objective function may be written in terms of ticket fractions:

$$\mathbf{c} \cdot \chi|_{\mathbf{f}_{\text{adm}}} = \mathbf{c} \cdot (M\mathbf{f}_{\text{adm}}) = (M^T \mathbf{c}) \cdot \mathbf{f}_{\text{adm}}. \quad (4.2.58)$$

To maximally violate a particular facet inequality, then, is equivalent to minimizing this objective function over the set of $\mathbf{f}_{\text{adm}} \in \Delta^{n-1}$ satisfying $B\mathbf{f}_{\text{adm}} = \mathbf{b}$. The problem of finding such an admissible raffle therefore takes the form of a particular *linear program*, and thus our iterative strategy requires the solution of a finite number of linear programs. Such linear programs can be solved via the so-called simplex method at relatively low computational cost. In this way, we implemented our iterative strategy in Mathematica and so obtained an algorithm to compute all vertices of the anti-correlation polyhedron. The resulting local polyhedra, in the case of spin-2 as well as spin- $\frac{5}{2}$, appear in Figure 4.17. As in Figure 4.13, this picture shows the facets we need to “glue onto” each of the four facets of the classical tetrahedron to get the full polyhedron.

	polyhedron volume (% of the ellipsope)	minimum value of $\chi_{ab} + \chi_{ac} + \chi_{bc}$
spin-1/2	54.0%	-1
spin-1	81.1%	$-1 \frac{1}{2}$
spin-3/2	90.4%	$-1 \frac{2}{5}$
spin-2	94.5%	$-1 \frac{1}{2}$
spin-5/2	96.1%	$-1 \frac{16}{35}$

Fig. 4.18 Polytope volumes and minimal values of the sum of anti-correlation coefficients

The anti-correlation polyhedra presented thus far (Figures 2.15, 4.11, and 4.13) suggests that, as the number of outcomes increases, the anti-correlation polyhedron converges to the full ellipsope. This is not really surprising, for if we were to allow for a continuous range of outcomes rather than a discrete set then the full ellipsope of correlation matrices is certainly generated. Indeed, for the case of a multivariate Gaussian distribution, the probability distribution is parametrized by the choice of a positive-definite correlation matrix.⁹ Thus the failure to obtain the full ellipsope rests on the discrete nature of the outcomes.¹⁰ As numerical evidence for convergence we consider the volume of our polyhedra, which can be computed from the list of vertices. These volumes are listed in Figure 4.18 in terms of their fraction of the ellipsope volume (which may be shown to be exactly $\pi^2/2$). These volumes are seen to increase monotonically to that of the full ellipsope as spin increases, in agreement with the convergence we are seeing in the figures.¹¹

⁹ A simpler example is provided by the *3m balance* discussed in Section 3.4 (see Figure 3.5).

¹⁰ One could say that the *discreteness* introduced by quantum mechanics puts some restrictions on the ellipsope but that those restrictions are lifted again by the *contextuality* it introduces.

¹¹ The admissibility conditions on the correlations represented by points in our correlation polyhedra form the main obstacle to a formal proof of this convergence. As we saw in Section 4.2.3, these conditions are of two kinds: the correlations should have uniform marginals and their correlation arrays should have the same symmetries as the quantum-mechanical ones they are meant to simulate. The convergence issue can meaningfully be studied without these symmetry conditions. The uniform-marginals condition, however, is critical for the construction of our correlation polyhedra. For one thing, as already mentioned in the introduction to Chapter 4, we need this condition to ensure that the correlation arrays in some allowed class differ only in their off-diagonal cells.



Chapter 5

Correlation arrays, polytopes and the CHSH inequality

Raffles and correlation arrays for experiments to test the CHSH inequality • Deriving the CHSH inequality and the Tsirelson bound for this setup.

In the preceding chapters we studied the Mermin setup, which involves two parties and three settings per party. In Chapter 2, we analyzed the case with two outcomes per setting; in Chapter 4, we extended our analysis to three and more outcomes per setting. In this chapter, we return to the simple case of two outcomes per setting. However, the two parties now get to choose from two different pairs of settings, rather than from the same triplet of settings. In other words, we replace the Mermin setup by the more common setup for which the CHSH inequality (Clauser et al., 1969) was formulated and tested. As in Chapter 2, we focus on correlations found in measurements performed on a pair of spin- $\frac{1}{2}$ particles in the singlet state and on raffles designed to simulate these correlations.

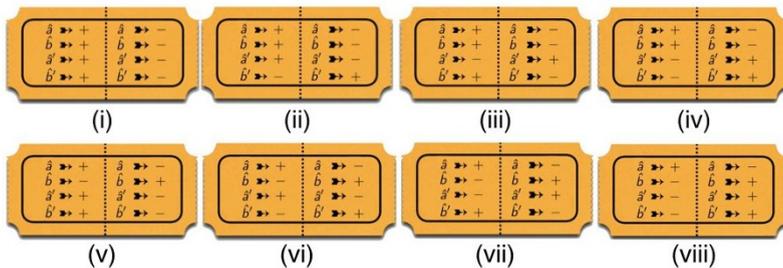


Fig. 5.1 Raffle tickets for four settings and two outcomes.

Our goal in this chapter is to recover the CHSH inequality and the corresponding Tsirelson bound in the Bub-Pitowsky-inspired framework developed in Chapters 2–4. The key to achieving this objective is to note that the CHSH setup, in which Alice and Bob pick from two *different* pairs of settings, (\hat{a}, \hat{b}) and (\hat{a}', \hat{b}') , can be treated as a straightforward generalization of the Mermin setup in which they pick from the *same quartet* of settings

$(\hat{a}, \hat{b}, \hat{a}', \hat{b}')$. The special case of this generalized Mermin setup is that Alice never actually uses the settings (\hat{a}', \hat{b}') and that Bob never actually uses the settings (\hat{a}, \hat{b}) . Nothing prevents us, however, from adding cells for the unused combinations of settings to the correlation arrays for the CHSH setup. In this way, the 2×2 correlation arrays for the CHSH setup turn into 4×4 correlation arrays (see, e.g., Figure 5.2) that are similar to the 3×3 ones for the Mermin setup (see, e.g., Figures 2.6, 2.12 and 2.14). The off-diagonal cells of these 4×4 correlation arrays can be parametrized by six anti-correlation coefficients, two of which (χ_{ab} and $\chi_{a'b'}$) do not play a role in the CHSH setup. The CHSH inequality and the Tsirelson bound in this case are conditions on the remaining four, $\chi_{aa'}$, $\chi_{ab'}$, $\chi_{ba'}$ and $\chi_{bb'}$. To derive the CHSH inequality, we use the kind of raffles we introduced in Section 2.5. As in Section 2.6, we derive the corresponding Tsirelson bound from the positive semi-definiteness of the anti-correlation matrix χ , which in this case is a symmetric 4×4 matrix with 1's on the diagonal and the six anti-correlation coefficients as its off-diagonal elements.

		Bob		\hat{a}		\hat{b}		\hat{a}'		\hat{b}'		
		+	-	+	-	+	-	+	-	+	-	
Alice	\hat{a}	+	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0
	-	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
\hat{b}	+	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
	-	0	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
\hat{a}'	+	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
	-	0	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
\hat{b}'	+	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	$\frac{1}{2}$	0	$\frac{1}{2}$
	-	0	$\frac{1}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0

Fig. 5.2 Correlation array for a single-ticket raffle with tickets of type (viii) (see Figure 5.1). Blue-on-white cells show perfect anti-correlation; white-on-blue cells show perfect correlation. The four cells in the upper-right corner are the ones actually probed in tests of the CHSH inequality.

Figure 5.1 shows the eight different types of raffle tickets for the CHSH setup. Figure 5.2 shows the correlation array for a raffle with a basket containing only type-(viii) tickets. Six of the off-diagonal cells in this correlation array show a perfect correlation, the other six show a perfect anti-correlation. This means that the values of the six anti-correlation coefficients parametrizing these twelve cells are:

$$\chi_{ab}^{(viii)} = \chi_{aa'}^{(viii)} = \chi_{ab'}^{(viii)} = -1, \quad \chi_{ba'}^{(viii)} = \chi_{bb'}^{(viii)} = \chi_{a'b'}^{(viii)} = 1. \quad (5.1)$$

These values can also be read directly off ticket (viii).

ticket	$\chi_{aa'}^{(k)}$	$\chi_{ab'}^{(k)}$	$\chi_{ba'}^{(k)}$	$\chi_{bb'}^{(k)}$	$\chi_{aa'}^{(k)} + \chi_{ab'}^{(k)} + \chi_{ba'}^{(k)} - \chi_{bb'}^{(k)}$
(i)	+1	+1	+1	+1	2
(ii)	+1	-1	+1	-1	2
(iii)	-1	+1	-1	+1	-2
(iv)	-1	-1	-1	-1	-2
(v)	+1	+1	-1	-1	2
(vi)	+1	-1	-1	+1	-2
(vii)	-1	+1	+1	-1	2
(viii)	-1	-1	+1	+1	-2

Table 5.1 Values of four of the anti-correlation coefficients for tickets (i)–(viii) in Figure 5.1. The final column shows the values for a linear combination of these coefficients.

In the CHSH setup, data are taken only for the four combinations of settings corresponding to the four cells in the upper-right corner of these 4×4 correlation arrays. These cells are characterized by four of the six anti-correlation coefficients in Eq. (5.1). Table 5.1 lists their values for all eight tickets in Figure 5.1. The final column gives the value for a linear combination of these four anti-correlation coefficients. Note that for tickets (i), (ii), (v) and (vii), this quantity is equal to 2, while for tickets (iii), (iv), (vi) and (viii), it is equal to -2 . For a raffle with any mix of tickets of type (i) through (viii), this quantity must therefore lie between -2 and 2:

$$-2 \leq \chi_{aa'} + \chi_{ab'} + \chi_{ba'} - \chi_{bb'} \leq 2. \quad (5.2)$$

This is the CHSH inequality (Bub, 2016, p. 68).

We now want to connect the raffles with tickets for the present case of four settings to the raffles in Section 2.5 for the case of three settings. An obvious way to do this is to ignore one of the four settings and have Alice and Bob choose between the remaining three. In that case, there are only three anti-correlation coefficients, which will satisfy the inequality in Eq. (2.5.1) for the Mermin setup. One could object, however, that this way of recovering Eq. (2.5.1) requires that we consider results found for combinations of settings Alice and Bob are not actually using in the CHSH setup. To avoid this objection, suppose Bob's setting \hat{b}' is the same as Alice's setting \hat{b} . This means that we restrict ourselves to tickets with the same values for \hat{b} and \hat{b}' . These are the tickets (i), (iii), (vi) and (viii) in Figure 5.1. Focusing on the rows for those four tickets in Table 5.1 and using that, in those rows, $\chi_{bb'} = 1$ and $\chi_{ab'} = \chi_{ab}$, we see that the inequality in Eq. (5.2) reduces to:

$$-1 \leq \chi_{aa'} + \chi_{ab} + \chi_{ba'} \leq 3. \quad (5.3)$$

If \hat{a}' is relabeled \hat{c} , this inequality turns into the one in Eq. (2.5.1) for the Mermin setup. Note that this is the form in which Bell (1964) originally derived the Bell inequality (see Section 2.1).

As was stressed in Section 2.5 for the case of three settings, however, the CHSH inequality is a necessary but not a sufficient condition for classical correlations. To obtain a complete characterization of the correlations allowed classically, we once again examine their geometrical representation.

The tickets in Table 5.1 give the coordinates of eight points in the four-dimensional space of anti-correlation coefficients $\chi_{aa'}, \chi_{ab'}, \chi_{ba'}, \chi_{bb'}$. Since the anti-correlation coefficients of a classical mixed state will be a weighted average of those for the classical pure states, we may interpret those eight points as vertices of some convex hull. This is the local polytope, i.e., the set of anti-correlation coefficients $\chi_{aa'}, \chi_{ab'}, \chi_{ba'}, \chi_{bb'}$ that can be simulated by such raffles.

While we cannot visualize this four-dimensional polytope, some geometric observations can be made. The eight vertices may be viewed as four pairs of antipodal points, with the four line segments between the pairs being mutually orthogonal. As was noted by Pitowsky (2008, p. 5), this polytope is the 4-dimensional octahedron or hyperoctahedron (Bub, 2016, p. 112). It has a total of 16 facets, half of which are given by the inequalities

$$-2 \leq \chi_{aa'} + \chi_{ab'} + \chi_{ba'} - \chi_{bb'} \leq 2, \quad (5.4)$$

$$-2 \leq -\chi_{aa'} + \chi_{ab'} - \chi_{ba'} - \chi_{bb'} \leq 2, \quad (5.5)$$

$$-2 \leq \chi_{aa'} - \chi_{ab'} - \chi_{ba'} - \chi_{bb'} \leq 2, \quad (5.6)$$

$$-2 \leq -\chi_{aa'} - \chi_{ab'} + \chi_{ba'} - \chi_{bb'} \leq 2. \quad (5.7)$$

Eq. (5.4) is just the CHSH inequality stated above when Alice and Bob use settings (\hat{a}, \hat{b}) and (\hat{a}', \hat{b}') , respectively. We may obtain the others similarly by reversing some of the settings. For instance, Eq. (5.5) is the CHSH inequality if Bob were to use setting $-\hat{b}$ instead of \hat{b} . For our purposes, we will regard each of them as a CHSH inequality.

For the remaining eight facets, note that all coordinates of the eight vertices are ± 1 , with $+1$ occurring an even number of times. Hence these are eight of the sixteen vertices for a four-dimensional hypercube, which is the non-signaling polytope in the CHSH setup. It bounds the local polytope, thus providing the remaining eight facets which Pitowsky (2008, p. 3, Eq. (2)) refers to as “trivial”:

$$-1 \leq \chi_{aa'} \leq 1, \quad -1 \leq \chi_{ba'} \leq 1, \quad -1 \leq \chi_{ab'} \leq 1, \quad -1 \leq \chi_{bb'} \leq 1. \quad (5.8)$$

In the case of the Mermin example, the four facets of the classical tetrahedron already sufficed to restrict the correlations to the non-signaling cube. This is not sufficient in the present setup. For instance, the values $\chi_{aa'} = 2, \chi_{ab'} = \chi_{ba'} = \chi_{bb'} = 0$ are allowed by the CHSH inequalities but lie outside the non-signaling hypercube. It is thus necessary to include the hypercube facets explicitly.

The parts of the local polytope which lie on one of these hypercube facets are particularly notable. If we restrict our attention to the facet $\chi_{bb'} = 1$, for instance, then we are looking at what part of the local polytope falls in the region parametrized by $\chi_{aa'}, \chi_{ab'}$ and $\chi_{ba'}$. Since each of these coefficients has magnitude 1, we are working with a copy of the non-signaling cube for three outcomes. The inequalities in Eqs. (5.4)–(5.7) then simplify to

$$-1 \leq \chi_{aa'} + \chi_{ab'} + \chi_{ba'} \leq 3, \quad (5.9)$$

$$-1 \leq -\chi_{aa'} + \chi_{ab'} - \chi_{ba'} \leq 3, \quad (5.10)$$

$$-1 \leq \chi_{aa'} - \chi_{ab'} - \chi_{ba'} \leq 3, \quad (5.11)$$

$$-1 \leq -\chi_{aa'} - \chi_{ab'} + \chi_{ba'} \leq 3. \quad (5.12)$$

Aside from having \hat{a}', \hat{b}' instead of \hat{c}, \hat{b} , this system of inequalities is identical to those that generated the classical tetrahedron in the Mermin setup (see Eqs. (2.5.2)–(2.5.5)). Hence the local polytope in the CHSH setup contains a copy of the classical tetrahedron, occurring on the $\chi_{bb'} = 1$ facet of the non-signaling hypercube; this facet is itself a copy of the non-signaling cube.

We now consider the analogous quantum story, i.e., the correlations found in measurements on the singlet state for the CHSH setup. Cells in the correlation array for the CHSH setup are no different from cells in the correlation array for the Mermin setup given in Section 2.6. The anti-correlation coefficients for any combination of measurement directions \mathbf{e}_a and \mathbf{e}_b are given by

$$\chi_{ab} = -\frac{\langle \hat{S}_{1a} \hat{S}_{2b} \rangle_{00}}{\sigma_{1a} \sigma_{1b}} = \cos \varphi_{ab} \quad (5.13)$$

(cf. Eqs. (3.2.5)–(3.2.6) and Eq. (4.1.70)) with (cf. Eqs. (3.2.2)–(3.2.3))

$$\sigma_{1a} = \sigma_{1b} = \sigma_{s=1/2} = \frac{\hbar}{2}. \quad (5.14)$$

We now introduce the anti-correlation matrix χ :

$$\begin{pmatrix} \chi_{aa} & \chi_{ab} & \chi_{aa'} & \chi_{ab'} \\ \chi_{ba} & \chi_{bb} & \chi_{ba'} & \chi_{bb'} \\ \chi_{a'a} & \chi_{a'b} & \chi_{a'a'} & \chi_{a'b'} \\ \chi_{b'a} & \chi_{b'b} & \chi_{b'a'} & \chi_{b'b'} \end{pmatrix} = \begin{pmatrix} 1 & \cos \varphi_{ab} & \cos \varphi_{aa'} & \cos \varphi_{ab'} \\ \cos \varphi_{ab} & 1 & \cos \varphi_{ba'} & \cos \varphi_{bb'} \\ \cos \varphi_{aa'} & \cos \varphi_{ba'} & 1 & \cos \varphi_{a'b'} \\ \cos \varphi_{ab'} & \cos \varphi_{bb'} & \cos \varphi_{a'b'} & 1 \end{pmatrix}.$$

Using that $\cos \varphi_{ab} = \mathbf{e}_a \cdot \mathbf{e}_b$ etc. and that $\mathbf{e}_a = (a_x, a_y, a_z)$, etc., we can factorize this matrix as (cf. Eqs. (2.6.5)–(2.6.7)):

$$\chi = \begin{pmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ a'_x & a'_y & a'_z \\ b'_x & b'_y & b'_z \end{pmatrix} \begin{pmatrix} a_x & b_x & a'_x & b'_x \\ a_y & b_y & a'_y & b'_y \\ a_z & b_z & a'_z & b'_z \end{pmatrix} \equiv L^\top L. \quad (5.15)$$

This factorization implies that, given any vector $\mathbf{v} = (v_a, v_b, v_c)^\top$, one has

$$\mathbf{v}^\top \chi \mathbf{v} = \mathbf{v}^\top L^\top L \mathbf{v} = (L \mathbf{v})^\top L \mathbf{v} \geq 0, \quad (5.16)$$

where in the last step we used that this quantity is the length squared of the vector Lv . Hence χ is positive semi-definite (cf. Eqs. (3.1.13)–(3.1.15)).

The set of such 4×4 matrices—that is, those which are symmetric, have 1's on the diagonal and are positive semi-definite—is conventionally known as the 4-elliptope. The terminology is an obvious generalization of the elliptope for the 3×3 case and for the purposes of this chapter we refer to the latter as the 3-elliptope. The 4-elliptope, being parametrized by six anti-correlation coefficients, is a six-dimensional set; like the 3-elliptope, it is moreover convex (since any weighted average of two positive semi-definite matrices is itself positive semi-definite). As noted before, however, the values of χ_{ab} and $\chi_{a'b'}$ are not provided in the context of the CHSH inequality. As such we must project this convex set to the four-dimensional space of the coefficients $\chi_{a'b'}$, $\chi_{aa'}$, $\chi_{ab'}$ and $\chi_{ba'}$. The resulting four-dimensional convex set, a shadow of the 4-elliptope, represents the class of correlations that Alice and Bob can obtain by performing measurements on the singlet state using the settings (\hat{a}, \hat{b}) and (\hat{a}', \hat{b}') , respectively.¹

Like the local polytope, neither the 4-elliptope nor its projected shadow can be visualized in its totality, yet some geometrical conclusions can still be drawn. For instance, since these anti-correlation coefficients are cosines, their magnitude does not exceed 1. Hence, the set of quantum correlations is contained within the non-signaling hypercube. Moreover, this set contains the local polytope. To see this, consider a configuration in which the unit vectors for all four measurement settings lie along the same line, i.e., are either parallel or anti-parallel. This configuration can be realized in eight ways. Then all the cosines are ± 1 , with $+1$ occurring an even number of times. This corresponds directly to the eight vertices of the local polytope (i.e., the rows of Table 5.1). Since the 4-elliptope (and therefore its shadow) is convex, we conclude that the set of quantum correlations does indeed include the local polytope.

It remains to show that the set of quantum correlations is larger than the local polytope. Recall that we can use $-\hat{S}_{2b}$ as a stand-in for \hat{S}_{1b} when evaluating the expectation value $\langle \hat{S}_{1a} \hat{S}_{1b} \rangle_{00}$ (see Section 3.2). That means that the anti-correlation coefficient χ_{ab} can be written as

¹ In the present case, we can further note that the four columns of L are unit vectors in three-dimensional space. As such they must be linearly dependent, which implies that zero is an eigenvalue of L (and thus of χ as well). Hence χ is not positive definite and must lie in the five-dimensional boundary of the 4-elliptope. However, this restriction plays no role upon projecting to the shadow of the 4-elliptope and so will not be elaborated upon further.

$$\chi_{ab} = -\frac{\langle \hat{S}_{1a} \hat{S}_{2b} \rangle_{00}}{\sigma_{s=1/2}^2} = \frac{\langle \hat{S}_{1a} \hat{S}_{1b} \rangle_{00}}{\sigma_{s=1/2}^2}. \quad (5.17)$$

Now note that

$$\begin{aligned} \frac{1}{\sigma_{s=1/2}^2} \left\langle \left(\hat{S}_{1a'} - \frac{1}{\sqrt{2}} \hat{S}_{1a} - \frac{1}{\sqrt{2}} \hat{S}_{1b} \right)^2 \right\rangle_{00} \\ = 2 - \sqrt{2} \chi_{aa'} - \sqrt{2} \chi_{a'b} + \chi_{ab} \geq 0. \end{aligned} \quad (5.18)$$

Similarly,

$$\begin{aligned} \frac{1}{\sigma_{s=1/2}^2} \left\langle \left(\hat{S}_{1b'} - \frac{1}{\sqrt{2}} \hat{S}_{1a} + \frac{1}{\sqrt{2}} \hat{S}_{1b} \right)^2 \right\rangle_{00} \\ = 2 - \sqrt{2} \chi_{ab'} + \sqrt{2} \chi_{bb'} - \chi_{ab} \geq 0. \end{aligned} \quad (5.19)$$

Adding these two inequalities, we arrive at

$$2 - \sqrt{2} \chi_{aa'} - \sqrt{2} \chi_{a'b} + 2 - \sqrt{2} \chi_{ab'} + \sqrt{2} \chi_{bb'} \geq 0, \quad (5.20)$$

which can be rewritten as

$$\chi_{aa'} + \chi_{a'b} + \chi_{ab'} - \chi_{bb'} \leq 2\sqrt{2}. \quad (5.21)$$

This is the Tsirelson bound for the CHSH setup (Bub, 2016, p. 68).

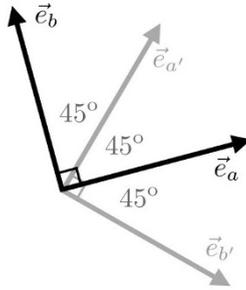


Fig. 5.3 Measurement directions for maximum violation of the CHSH inequality.

To reach the Tsirelson bound we need both of the expectation values in Eqs. (5.18)–(5.19) to vanish. This occurs when the unit vectors for the measurement settings of Alice and Bob are related as

$$\mathbf{e}_{a'} = \frac{1}{\sqrt{2}}\mathbf{e}_a + \frac{1}{\sqrt{2}}\mathbf{e}_b, \quad \mathbf{e}_{b'} = \frac{1}{\sqrt{2}}\mathbf{e}_a - \frac{1}{\sqrt{2}}\mathbf{e}_b. \quad (5.22)$$

Squaring the first of these relations, we find that

$$1 = (\mathbf{e}_{a'})^2 = \left(\frac{1}{\sqrt{2}}\mathbf{e}_a + \frac{1}{\sqrt{2}}\mathbf{e}_b \right)^2 = \frac{1}{2} + \frac{1}{2} + \mathbf{e}_a \cdot \mathbf{e}_b, \quad (5.23)$$

from which it follows that $\mathbf{e}_a \perp \mathbf{e}_b$. The situation is thus as shown in Figure 5.3. It follows that $\varphi_{ab} = \varphi_{a'b'} = 90^\circ$, $\varphi_{aa'} = \varphi_{ab'} = \varphi_{ba'} = 45^\circ$ and $\varphi_{bb'} = 135^\circ$. Hence

$$\begin{aligned} \chi_{ab} &= \chi_{a'b'} = \cos 90^\circ = 0, \\ \chi_{aa'} &= \chi_{ba'} = \chi_{ab'} = \cos 45^\circ = \frac{1}{\sqrt{2}}, \\ \chi_{bb'} &= \cos 135^\circ = -\frac{1}{\sqrt{2}}. \end{aligned} \quad (5.24)$$

In this case we therefore have

$$\chi_{aa'} + \chi_{ab'} + \chi_{ba'} - \chi_{bb'} = 2\sqrt{2}. \quad (5.25)$$

This represents the maximum violation of the CHSH inequality and we see once again that the set of quantum correlations is larger than the local polytope.

As was noted in Section 2.5 for the three-setting case, however, Tsirelson's bound is a necessary but not a sufficient condition for quantum correlations. To see this presently, note that *every* possible expectation value of the form Eqs. (5.18)–(5.19) ought to be non-negative. But each such expectation gives rise to another linear inequality on the six anti-correlation coefficients. For instance, if Bob were to instead use $(-\hat{a}', -\hat{b}')$ then the inequality obtained is

$$\chi_{aa'} + \chi_{ab'} + \chi_{ba'} - \chi_{bb'} \geq -2\sqrt{2}. \quad (5.26)$$

The anti-correlation coefficients therefore satisfy not one but infinitely many linear bounds, all of which must be satisfied if the coefficients arise quantum-mechanically. Hence Tsirelson's bound is evidently not sufficient to characterize the set of quantum correlations in the CHSH setup and moreover no finite list of such inequalities will suffice either. The set of quantum correlations allowed in the CHSH setup is therefore not the four-dimensional analogue of a polyhedron but a compact convex body like the ellipsope in three dimensions.

As was the case for the 3-elliptope in relation to the three-setting case, however, this still leaves the possibility to characterize the set of quantum correlations by *nonlinear* inequalities. To find these it is useful to relate the set of quantum correlations for the four-setting case to those for the three-setting case. For instance, suppose Alice and Bob are allowed to use three of the four settings, e.g., Alice and Bob measure using settings $(\hat{a}, \hat{b}, \hat{a}')$. The results of Section 2.6 then apply and therefore the anti-correlation coefficients $(\chi_{ab}, \chi_{aa'}, \chi_{ba'})$ must lie within the corresponding 3-elliptope:

$$1 - \chi_{ab}^2 - \chi_{aa'}^2 - \chi_{ba'}^2 + 2\chi_{ab}\chi_{aa'}\chi_{ba'} \geq 0. \quad (5.27)$$

In mathematical terms, this is an application of Sylvester's criterion: A matrix is positive semi-definite if and only if none of its principal minors are negative. It is then useful to rewrite the 3-elliptope in the equivalent form²

$$|\chi_{aa'}\chi_{ba'} - \chi_{ab}| \leq \sqrt{1 - \chi_{aa'}^2} \sqrt{1 - \chi_{ba'}^2}. \quad (5.28)$$

The coefficient χ_{ab} cannot be deduced from Alice and Bob's measurements in the CHSH setup. Nevertheless this quantity does exist and the last inequality signifies that it cannot differ too much from $\chi_{aa'}\chi_{ba'}$. A similar calculation shows that, if Alice and Bob instead only made measurements on the settings $(\hat{a}, \hat{b}, \hat{b}')$, then

$$|\chi_{ab'}\chi_{bb'} - \chi_{ab}| \leq \sqrt{1 - \chi_{ab'}^2} \sqrt{1 - \chi_{bb'}^2}. \quad (5.29)$$

We thus have bounds on how much χ_{ab} can differ from the values of $\chi_{aa'}\chi_{ba'}$ and $\chi_{ab'}\chi_{bb'}$. The triangle inequality then bounds how much these two quantities can differ from each other:

$$\begin{aligned} |\chi_{aa'}\chi_{ba'} - \chi_{ab'}\chi_{bb'}| &= |(\chi_{aa'}\chi_{ba'} - \chi_{ab}) + (\chi_{ab} - \chi_{ab'}\chi_{bb'})| \\ &\leq |\chi_{aa'}\chi_{ba'} - \chi_{ab}| + |\chi_{aa'}\chi_{ba'} - \chi_{ab}| \\ &\leq \sqrt{1 - \chi_{aa'}^2} \sqrt{1 - \chi_{ba'}^2} + \sqrt{1 - \chi_{ab'}^2} \sqrt{1 - \chi_{bb'}^2}. \end{aligned} \quad (5.30)$$

Any set of anti-correlation coefficients obtained from measurements on the singlet state in the CHSH setup must satisfy this inequality, first obtained by

² Aside from an overall square root, this form of the inequality (for three rather than four random variables) can be found in Yule (1897, p. 485). Cf. Eq. (3.4.29) in Chapter 3.

Landau (1988). The 3-elliptope therefore proves useful even when characterizing quantum correlations in the non-visualizable case of four settings.³

As an application of these results we consider again the special case where Alice and Bob share one setting, e.g. $\hat{b}' = \hat{b}$. Then $\chi_{bb'} = 1$ and $\chi_{ab'} = \chi_{ab}$, so the second 3-elliptope inequality is then fulfilled trivially (both sides vanish identically) and Landau's result collapses to the 3-elliptope inequality. If we further relabel Bob's setting $\hat{a}' \rightarrow \hat{c}$, then this 3-elliptope takes the form

$$|\chi_{ac}\chi_{bc} - \chi_{ab}| \leq \sqrt{1 - \chi_{ac}^2} \sqrt{1 - \chi_{bc}^2}. \quad (5.31)$$

But this is the same 3-elliptope as considered originally in the Mermin example. Geometrically this means that the shadow of the 4-elliptope contains a copy of the 3-elliptope, occurring on the $\chi_{bb'} = 1$ facet of the non-signaling hypercube. This is as it should be: The scenario where Alice and Bob, respectively, use settings (\hat{a}, \hat{b}) and (\hat{b}, \hat{c}) is exactly the setup originally employed by Bell (1964). Combining this with the corresponding results for the local polytope and the non-signaling hypercube, we conclude that we recover the Mermin setup of Section 2.5–2.6 (including the entirety of Figure 2.16) when $\hat{b}' = \hat{b}$.

To concretely illustrate how the non-signaling hypercube, the set of quantum correlations and the local polytope are related, we consider the family of anti-correlation coefficients given by

$$\chi_{aa'} = \chi_{ab'} = \chi_{ba'} = -\chi_{bb'} = -t \quad (5.32)$$

for $0 \leq t \leq 1$. The case $t = 0$, where all correlation coefficients vanish, can be simulated classically (and therefore quantum mechanically), for instance by a raffle with basket containing all ticket types in equal proportion. By contrast, the case $t = 1$ corresponds to the PR box shown in Figure 2.3 in Section 2.3 and should not be realizable in quantum mechanics despite being non-signaling. More precisely, it is one of the vertices of the non-signaling hypercube which is not also a vertex of the local polytope.

The questions are then for what range of t can such correlation coefficients be simulated classically or quantum-mechanically. For the classical case we

³ It should be noted that we have only established that this is a necessary condition for quantum correlations, i.e., if $\chi_{aa'}$, $\chi_{ab'}$, $\chi_{ba'}$ and $\chi_{bb'}$ do not satisfy Landau's inequality then one cannot find χ_{ab} , $\chi_{a'b'}$ such that χ is positive semi-definite (and therefore these are not quantum correlations). The converse claim, i.e., that Landau's inequality is also a sufficient condition for quantum correlations, is beyond the scope of this work and will not be addressed further.

observe that

$$|\chi_{aa'} + \chi_{ab'} + \chi_{ba'} - \chi_{bb'}| = 2t. \quad (5.33)$$

Since the CHSH inequality bounds this magnitude by 2, we conclude that $t \leq 1/2$ is the classical bound. For the quantum case we appeal to Landau's inequality, which takes the form

$$\begin{aligned} |\chi_{aa'}\chi_{ba'} - \chi_{ab'}\chi_{bb'}| &= 2t^2 \\ &\leq \sqrt{1 - \chi_{aa'}^2} \sqrt{1 - \chi_{ba'}^2} + \sqrt{1 - \chi_{ab'}^2} \sqrt{1 - \chi_{bb'}^2} \\ &= 2(1 - t^2). \end{aligned} \quad (5.34)$$

We therefore conclude that $t \leq 1/\sqrt{2}$, corresponding to Tsirelson's bound, i.e., the maximal violation of the CHSH inequality, marks the boundary between quantum and non-quantum correlations along this family of anti-correlation coefficients.

Despite the evident parallels of the quantum set for the three- and four-setting cases, there is one novelty which we have not yet addressed (Goh et al., 2018). In the three-setting case, the four vertices and the four curved faces of the 3-elliptope had two characteristics. First, none of these points can be expressed as convex combinations of other points on the 3-elliptope; as such, they are said to be *extreme* points of this convex set. Second, for each of these extreme points there is a plane passing through it such that the rest of the 3-elliptope is on one side of the plane. Such points are said to be *exposed*, and so we conclude that every extreme point of the 3-elliptope is also exposed (see Goh et al., 2018, Appendix A: convex sets, for formal definitions of extreme and exposed points).

As emphasized by Goh et al. (2018, sec. III.C.1), however, extreme points need not be exposed (though the converse is true). In particular, they explicitly note that, for the CHSH setup,

$$\chi_{aa'} = \chi_{ab'} = \chi_{a'b} = -\frac{1}{2}, \quad \chi_{bb'} = 1 \quad (5.35)$$

is an extreme non-exposed point.⁴ This arises because the CHSH quantum set includes curved and flat boundary regions which, crucially, join smoothly

⁴ Instead of our notation (\hat{a}, \hat{a}') and (\hat{b}, \hat{b}') , Goh et al. (2018) use the notation (A_0, A_1) and (B_0, B_1) for the settings of Alice and Bob, respectively. They also label the point in Eq. (5.35) by values of correlation rather than anti-correlation coefficients. In their notation, Eq.

at their common border. This ensures that any point which would expose the border point would also expose the rest of the flat boundary region. Hence these border points, despite being extreme, cannot be exposed points.

(5.35) thus becomes $\langle A_0 B_0 \rangle = \langle A_0 B_1 \rangle = \langle A_1 B_0 \rangle = 1/2$, $\langle A_1 B_1 \rangle = -1$ (see [Goh et al., 2018](#), p. 022104-7, Eq. (24) and the unnumbered equation above it).



Chapter 6

Interpreting quantum mechanics

Detailed summary of Chapters 2–5 • The interplay between principle-theoretic and constructive approaches to physics • The new kinematics of quantum theory • Examples of problems solved by the new kinematics • Measurement

6.1 The story so far

In Chapter 2 we introduced the concept of a correlation array—a concise representation of the statistical correlations between the outcomes of measurements performed by two separated parties in the context of a given experimental setup. We focused primarily on the Mermin-style setup involving parties named Alice and Bob, each of whom are given one of two correlated systems and are asked to measure them using one of the three settings \hat{a} , \hat{b} and \hat{c} . Such a setup can be characterized using a 3×3 correlation array in which each cell corresponds to one of the nine possible combinations for Alice’s and Bob’s setting choices. In Section 2.4 we showed how to parameterize the cells in such a correlation array by means of an anti-correlation coefficient, which, as we showed in Chapter 3 (see Eqs. (3.2.5)–(3.2.6)), is the negative of the expectation value of the product of Alice’s and Bob’s random variables, divided by the product of their standard deviations. In particular, when there are two possible outcomes per measurement, a symmetric 3×3 correlation array with zeroes along the diagonal can be parameterized using three anti-correlation coefficients χ_{ab} , χ_{ac} and χ_{bc} , as depicted in Figure 2.7. The Mermin correlation array given in Figure 2.6 is an example of a correlation array that can be parametrized in this way.

We considered local hidden-variable models for 3×3 correlation arrays of this kind in Section 2.5. We imagined, in particular, modeling such arrays with mixtures of raffle tickets like the ones in Figure 2.11, and for such models we derived the following constraints on the anti-correlation coefficients χ_{ab} , χ_{ac} and χ_{bc} :¹

¹ These are identical to the inequalities given in Eqs. (2.5.2)–(2.5.5).

$$-1 \leq \chi_{ab} + \chi_{ac} + \chi_{bc} \leq 3, \quad (6.1.1)$$

$$-1 \leq \chi_{ab} - \chi_{ac} - \chi_{bc} \leq 3, \quad (6.1.2)$$

$$-1 \leq -\chi_{ab} + \chi_{ac} - \chi_{bc} \leq 3, \quad (6.1.3)$$

$$-1 \leq -\chi_{ab} - \chi_{ac} + \chi_{bc} \leq 3. \quad (6.1.4)$$

Together these four linear inequalities are necessary and sufficient to characterize the space of possible statistical correlations realizable in any such model. This space can be visualized as the tetrahedron in Figure 2.15; i.e., for any given point $(\chi_{ab}, \chi_{ac}, \chi_{bc})$, it is contained in the convex set represented by the tetrahedron if and only if it satisfies all four of Eqs. (6.1.1)–(6.1.4). In Section 2.6 we showed that the convex set characterizing the allowable *quantum* correlations for 3×3 two-outcome setups, in the case of the singlet state of a pair of entangled spin- $\frac{1}{2}$ particles, is a superset of those allowed in a local hidden-variables model. It can be characterized by the *non-linear inequality*²

$$1 - \chi_{ab}^2 - \chi_{ac}^2 - \chi_{bc}^2 + 2\chi_{ab}\chi_{ac}\chi_{bc} \geq 0, \quad (6.1.5)$$

whose associated inflated tetrahedron or ellipsope is shown in Figure 2.16. We called this inequality the *ellipsope inequality*.

Our work is both continuous with and extends that of Pitowsky. Pitowsky, in turn building on the work of George Boole (Pitowsky, 1994), also considers the distinction between quantum and classical theory in light of the inequalities that characterize the possibility space of relative frequencies for a given classical event space. Pitowsky describes a general algorithm for determining these inequalities: Given the logically connected events E_1, \dots, E_n , write down the propositional truth table corresponding to them and then take each row to represent a vector in an n -dimensional space. Their convex hull yields a polytope, and the sought-for inequalities characterize the facets of this polytope. Alternately, if we already know the inequalities we can then determine the polytope associated with them.

In our own case the event space associated with a 3×3 correlation array for a setup involving two possible outcomes per measurement yields an easily visualizable three-dimensional representation of possible correlations between events for both a quantum and a local hidden-variables model, when the former is of a pair of entangled spin- $\frac{1}{2}$ particles in the singlet state. More-

² This equation is identical to Eq. (2.6.9).

over in the quantum case we showed that the resulting representation remains three-dimensional even when we transition to setups involving more than two outcomes; i.e., to setups involving a pair of spin- m particles, where m is an integer or half-integer $> \frac{1}{2}$, entangled in the singlet state. Indeed, we showed in Section 4.1.5 that the resulting representation is in every case (i.e., for all such values of m) the very same ellipsope as the one we derived in Section 2.6 for two outcomes (i.e., for spin- $\frac{1}{2}$) and which we depicted in Figure 2.16. In the local hidden-variables case (where we model correlations with raffles) the local polytopes characterizing the space of possible correlations for setups with more than two possible outcomes per measurement are of much higher dimension than three. In part through considering only those raffles that have a hope of recovering the quantum set, we showed in Section 4.2 how to project these higher-dimensional polytopes down to three-dimensional anti-correlation polyhedra (see the flowchart in Figure 4.7).³ We showed that with increasing spin these polyhedra become further and further faceted and correspondingly more and more closely approximate the full quantum ellipsope⁴—though actually computing these polyhedra becomes more and more intractable as the number of possible outcomes per setting increases. Finally, in addition to providing an easily visualizable representation in three dimensions of the quantum and local hidden-variable correlations associated with a 3×3 Mermin-style setup, we showed how the correlation array formalism for this case can be straightforwardly extended so as to provide useful insight into the more familiar correlational space associated with CHSH-style setups, if the latter are characterized using 4×4 correlation arrays and parameterized with six anti-correlation coefficients (see Chapter 5).

As Pitowsky observes (1989b, p. IV),⁵ linear inequalities such as those characterizing the facets of our polytopes have been an object of study for probability theorists since at least the 1930s. And although they were (re)discovered in a context far removed from these abstract mathematical investigations, the various versions of Bell's inequality are all inequalities of just this kind. *Non-linear* inequalities such as the ellipsope inequality in Eq. (6.1.5), in contrast, are not. Nevertheless, equations like this one have also been an object of study for probability theorists. Drawing directly on their work, we showed in Section 3.1 how one can derive an equation analogous to Eq. (6.1.5) characterizing the quantum ellipsope from general statistical considerations concerning three

³ We call them polyhedra rather than polytopes since they are always three-dimensional.

⁴ See the table in Figure 4.18 and the facets of the polyhedra in Figures 4.11, 4.13 and 4.17.

⁵ We previously noted Pitowsky's observation in Chapter 1, where we quoted him.

balanced random variables X , Y and Z (for the meaning of *balanced*, see the definition numbered (3.1.1)). Specifically, we derived a constraint on the correlation coefficients ρ_{XY} , ρ_{XZ} and ρ_{YZ} that is of exactly the same form as Eq. (6.1.5) (which, recall, constrains the *anti*-correlation coefficients χ_{ab} , χ_{ac} and χ_{bc}):^{6,7}

$$1 - \rho_{XY}^2 - \rho_{XZ}^2 - \rho_{YZ}^2 + 2\rho_{XY}\rho_{XZ}\rho_{YZ} \geq 0. \quad (6.1.6)$$

In Sections 3.2 and 3.3 we took up the questions, respectively, of how to model this general statistical constraint quantum-mechanically and in a local hidden-variables model, noting that the general derivation of Eq. (6.1.6) relies essentially on the fact that we can consider a linear combination of the three random variables X , Y and Z in order to determine the expectation value of its square:⁸

$$\left\langle \left(v_1 \frac{X}{\sigma_X} + v_2 \frac{Y}{\sigma_Y} + v_3 \frac{Z}{\sigma_Z} \right)^2 \right\rangle \geq 0. \quad (6.1.7)$$

To model such a relation with local hidden-variables, however, we require a joint probability distribution over X , Y and Z . This in turn actually entails a tighter bound on the correlation coefficients than the one given by Eq. (6.1.6). Namely, it entails the analogue of the CHSH inequality for our setup, which should be unsurprising given the classical assumptions we began with. Thus, while the ellipsope equation given by Eq. (6.1.6) indeed constrains correlations between local hidden-variables in the setups we are considering, those correlations do not saturate that ellipsope. In the case where there are only two possible values corresponding to each of the three random variables, the subset of the ellipsope achievable is just the tetrahedron given in Figure 2.15. For more than two values per variable the situation is more complicated: When the number of possible values, n , per variable is odd, one can actually reach the Tsirelson bound for this setup—the minimum value of 0 in Eq. (6.1.6)—while when the number of possible values, n , per variable is even, one reaches the bound only in the limit as $n \rightarrow \infty$ (see Eqs. (3.3.8)–(3.3.9)).⁹ But in either case—whether one reaches the Tsirelson bound or not—it appears that one

⁶ If X is the taste of Alice’s banana when peeled in the a -direction and Y is the taste of Bob’s banana (from the same pair) when peeled in the b -direction, the (Pearson) correlation coefficient ρ_{XY} is just the negative of its corresponding anti-correlation coefficient, χ_{ab} (cf. Eq. (3.2.6)).

⁷ The following equation is identical to Eq. (3.1.15).

⁸ This equation is identical to Eq. (3.1.10).

⁹ A local hidden-variables theory with an odd number of possible values per variable is analogous to the case of integer spin in quantum mechanics, and a local hidden-variables

requires a number of possible values $n \rightarrow \infty$ per random variable in order to saturate the volume of the ellipsope in its entirety.¹⁰

From a slightly different point of view we can understand this as follows. Think of arbitrary linear combinations of three vectors \mathbf{X}_a , \mathbf{X}_b and \mathbf{X}_c in some vector space, and let φ_{ab} , φ_{ac} and φ_{bc} represent the angles between these vectors (cf. De Finetti, 1937, sec. 4). The correlation coefficient $\rho_{\alpha\beta}$ (where α and β can be any one of a , b or c) may then be defined as the inner product of the vectors \mathbf{X}_α and \mathbf{X}_β , yielding (for instance) the natural property that two vectors are uncorrelated whenever they are orthogonal. As we explained in Section 3.4, from this point of view we can interpret Eq. (6.1.6) geometrically as a constraint on the angles $\varphi_{\alpha\beta}$.

To express this mathematically is one thing. It is another to give a model for it. Note that such a model *need not be classical*. De Finetti's own interpretation of the probability calculus, for instance, was not.¹¹ Any underlying model for these correlations that is classical, however, presupposes the existence of a joint distribution over the individual vectors \mathbf{X}_a , \mathbf{X}_b and \mathbf{X}_c . From this it follows that the correlations realizable in such a model cannot saturate the full volume of the ellipsope expressed by Eq. (6.1.6) except in the limit as the number of possible values associated with each vector goes to infinity.

As we explained in Section 3.2, there are a number of challenges that need to be overcome in order to provide a quantum-mechanical model for the general statistical constraint expressed in Eq. (6.1.6).¹² The most important of these is that in quantum mechanics one cannot consistently assume a joint probability distribution over incompatible observables, such as one would

theory with an even number of possible values per variable is analogous to the case of half-integer spin.

¹⁰ For further discussion, see Section 4.2.4 and in particular the caveat contained in note 11 of that section.

¹¹ De Finetti distinguished between coherent degrees of belief in—and therefore probabilities associated with—verifiable as opposed to unverifiable events. This has consequences for his theory of probability. For instance if A and B are verifiable but not jointly verifiable they are not subject to the inequality $P(A) + P(B) - P(A \& B) \leq 1$. See Berkovitz (2012, 2019) for further discussion.

¹² The word *model*, in our usage, plays the same role in the sentence 'quantum-mechanical model' as it does in the sentence 'classical model'; i.e., it signifies a concrete system or collection of systems that satisfies the constraints imposed on physical systems by quantum mechanics and classical mechanics, respectively. An example of a classical model is a basket of raffle tickets. An example of a quantum-mechanical model is an ensemble of spin- $\frac{1}{2}$ systems. We will explore the analogies and disanalogies between these particular models further in Section 6.5.

have to do in order to non-ambiguously define a vector \mathbf{X} by taking a linear combination of the quantum analogues of \mathbf{X}_a , \mathbf{X}_b and \mathbf{X}_c . Since the sum of any two Hermitian operators is also Hermitian, however, then given three observables represented by, say, the operators \hat{S}_a , \hat{S}_b and \hat{S}_c , one can always also consider the observable represented by the operator $\hat{S} \equiv \hat{S}_a + \hat{S}_b + \hat{S}_c$. As von Neumann observed already in 1927,¹³ quantum mechanics allows us to assign in this way a value to the sum of three variables without assigning values to all of them individually. From this it follows, not only that the ellipsope equation constrains the possible correlations in the setups we are considering, but also that it tightly constrains them. The quantum-mechanical correlations in these setups, that is, saturate the full volume of the ellipsope. Moreover we saw in Section 3.2 how, in virtue of certain other assumptions we needed to model the constraint quantum-mechanically, Eq. (6.1.6)—the equation we derived *from without*—reduces to Eq. (6.1.5)—the equation we derived *from within* quantum mechanics.

The remainder of this chapter is devoted to the philosophical conclusions that can be drawn from the foregoing. In Section 6.2 we will comment on the nature of our derivation of the space of quantum correlations for the setups we have considered. We will note that our derivation evinces aspects of both the principle-theoretic and the constructive approaches to physics, and that in our own derivation and generally in the practice of theoretical physics, both work together to yield understanding of the physical world. In Section 6.3 we will argue that our derivation yields the insight that the fundamental novelty of the quantum mode of description is in the kinematical and not the dynamical part of quantum theory. This distinction—between the kinematical and dynamical parts of a theory—is one we take to be of far more significance than the distinction between principle-theoretic and constructive approaches that has been the object of so much recent attention. In Section 6.4 we consider examples, from the history of quantum theory, of physical problems that were resolved by considering the changes to the kinematical framework of fundamental physics introduced by quantum mechanics. We close, in Section 6.5, with the topic of measurement. Though we locate them elsewhere than is standardly done, we conclude that philosophical puzzles yet remain in relation to the physics of measurement in quantum mechanics. The mere fact that puzzles exist, be they philosophical or physical, does not warrant our viewing the account of a measurement provided by quantum mechanics to be deficient, however. Every candidate fundamental theoretical framework

¹³ See note 11 in Section 3.2.

has its puzzles, otherwise it would be more than just a *candidate* fundamental theoretical framework.¹⁴ What is the most important of these for quantum mechanics as regards the question of measurement? As we will expound upon in Section 6.5, it is *not* what has been referred to elsewhere as the “big” measurement problem (cf. [Bub & Pitowsky, 2010](#), p. 438):¹⁵ the problem that stems from the fact, that of the various possible values obtainable through the measurement of a selected observable, the one that will actually be obtained is, according to quantum mechanics, fundamentally a matter of chance. In Section 6.5 we will argue that that problem is a superficial one. The more important puzzle of measurement, what has elsewhere been called the “small” measurement problem (cf. *ibid.*), is that within quantum mechanics it is not even possible to describe all of our experience consistently in accordance with a (classical) probability distribution at all. For selecting a particular observable represented by a particular Hermitian operator in the context of a measurement entails that only the observables that are compatible (i.e., whose Hermitian operators commute) with that of the selected observable get to be assigned definite values over which a (classical) probability distribution can be defined in the context of the measurement. In Section 6.5 we will see both that the physical significance of the puzzle of measurement as well as the physical account quantum mechanics provides of particular measurements flow naturally from the constraints that quantum mechanics’ kinematical core imposes on our representations of quantum systems. Along the way, which begins in the next section and continues through to the end of this chapter, we will be reflecting on the conception of the world that we take to be suggested by these constraints.

Before moving on to Section 6.2 we want to comment on the interpretation of the distinction between principle-theoretic and constructive approaches that figures prominently within it. The idea of such a distinction dates back to a popular article Einstein published in the *London Times* ([Einstein, 1919](#)) shortly after the Eddington-Dyson eclipse expeditions had (practically overnight) turned him into an international celebrity. The distinction Einstein drew there has since taken on a life of its own, both in the historical and in the foundational physics literature. The account of this distinction, which we give in the next section, is

¹⁴ Quantum mechanics is a candidate fundamental theoretical framework in the sense that it constitutes a fundamental set of rules and constraints in accordance with which particular physical theories can be constructed (cf. [Nielsen & Chuang, 2016](#), p. 2), for instance quantum field theories.

¹⁵ The distinction between a big and a small measurement problem first appeared in [Pitowsky \(2006\)](#).

meant to more closely reflect the latter literature (especially the literature on quantum foundations). It is not meant to reflect what Einstein intended by the distinction either in 1919 or in his later career.¹⁶

The account that we give of the distinction is also different from certain others whose interpretations of quantum mechanics are close to ours on the phylogenetic tree we mentioned in Chapter 1. For instance, on our reading of him (based on his unpublished monograph), Bill Demopoulos uses the label “constructive” to refer to particular dynamical hypotheses concerning the micro-constituents of matter, and uses the label “principle-theoretic” to refer to the structural constraints that a theory imposes on the representations it allows. By contrast, our own way of using the label “constructive” is broader than this; a constructive characterization may involve the kinematical features of a theory (cf. Janssen, 2009), and a principle-theoretic characterization may include dynamical posits (see Koberinski & Müller, 2018, especially sec. 12.4). In the next section we will be speaking about constructive and principle-theoretic *derivations* in particular. What is essential about the former kind of derivation is that it begins from an internal perspective—it is a derivation *from within* quantum theory characterizing some aspect of the world that is described by it, while what is essential about the latter kind of derivation is that it begins from an external perspective—it is a derivation *from without*, i.e., from a more general mathematical framework, aiming to characterize some aspect of the quantum world.

Jeff Bub and Itamar Pitowsky also distinguish principle-theoretic from constructive approaches in their 2010 paper. In their case it is actually not clear to us which of the two senses of the distinction given above is the one they really intend, and at various times they seem to be appealing to both (see especially sec. 2 of their paper), although in fairness they appear to do so consistently. This slippage is in any case understandable: The idea that the kinematical core of a theory constrains all of its representations is easily mistaken for the idea that this core constitutes a characterizing principle for the theory. In our own discussion we will endeavor to be careful in distinguishing the former from the latter. But regardless of what one makes of the distinction between constructive and principle-theoretic approaches, we take this distinction to be of relatively minor importance. As we will see further below, the more important distinction to bear in mind when interpreting a physical theory, as

¹⁶ For the views of one of us on what Einstein meant by this distinction and how it captures Einstein’s own scientific methodology, see Janssen (2009, sec. 3.5, pp. 38–41), Janssen & Lehner (2014, p. 16, pp. 26–28) and Duncan & Janssen (2019, Ch. 3, especially p. 102 and pp. 119–120).

one of us has pointed out in the context of special relativity, is the distinction between the kinematics and the dynamics of that theory (Janssen, 2009, p. 38).

6.2 From within and from without

Our derivation of the space of possible quantum correlations in the 2-party, 3-setting, Mermin-style setup illustrates the interplay between principle-theoretic and constructive approaches that is typical of the actual practice and methodology of theoretical physics (compare, e.g., Sainz, Guryanova, Acín, & Navascués, 2018). Our goal was to carve out the space of quantum correlations so as to gain insight into what distinguishes quantum from classical theory. Accordingly, guided by the work of probability theorists and statisticians like De Finetti, Fisher, Pearson and Yule, we associated vectors with random variables and derived a constraint on the angles between such vectors, Eq. (6.1.6), which has the same form as the constraint that characterizes the quantum correlational space of our Mermin-style setup. But it would be wrong to stop here. In and of itself Eq. (6.1.6) is just an abstract equation; it neither explains the space of quantum correlations, nor what distinguishes that space from the corresponding classical space. To gain insight into these matters we needed to model the elliptope inequality both in a quantum and in a local hidden-variables model.

In the case of a local hidden-variables model, the classical assumptions that underlie the vectors constrained by Eq. (6.1.6) entail a tighter bound on the correlations between them than what is given by the inequality itself. Specifically, assigning a value to the sum of three variables classically requires that we assign a value to all three of them individually. And because of this, the correlations in a local hidden-variables model cannot saturate the full space described by Eq. (6.1.6), unless the number n of possible values for a random variable goes to infinity—unless, that is, the range of possible values for a random variable is actually continuous (see Section 4.2.4 for further discussion, as well as note 10 in the same section).

In quantum theory, in contrast, this classical presupposition regarding a sum of random variables does not apply. We can indeed still take a sum of three random variables in quantum theory, but we do not need to assign a value to each of them individually in order to do so. As a result, the constraint expressed by the quantum version of the inequality turns out to be tight—quantum correlations, that is, saturate the full volume of the elliptope—regardless of the number of possible values we can assign to the random variables in a particular setup. In this sense Eq. (6.1.5)—a constraint on expectation values—expresses an essential structural aspect of the quantum probability space. Moreover a

visual comparison of the quantum ellipsope with the various polyhedra we derived for local hidden-variable models vividly demonstrates the way that their respective probabilistic structures differ. This, finally, motivates us to think of quantum mechanics as a theory that is, at its core, *about* probabilities. But this should not be misunderstood. What is being expressed here is the thought that the conceptual *novelty* of quantum theory consists precisely in the way that it departs from the assumptions that underlie classical probability spaces.

One of the strengths of principle-theoretic approaches to physics is that they give us insight into the multi-faceted nature of the objects of a theory.¹⁷ A formal framework is set up, for example the C^* -algebraic framework of Clifton et al. (2003), one of the minimalist operationalist frameworks of states, transformations and effects discussed in Myrvold (2010), “general probabilistic” frameworks (Koberinski & Müller, 2018), “informational” and/or “computational” frameworks (Dakic & Brukner, 2011, Chiribella & Ebler, 2019), “operator tensor” formulations (Hardy, 2012) and so on.¹⁸ Each such framework focuses on a particular aspect of quantum phenomena, for example on distant quantum correlations, quantum measurement statistics, quantum dynamics and so on. In the language of a given framework one then posits a principle (or a small set of them), e.g., “no signaling” (Popescu & Rohrlich, 1994), “no restriction” (Chiribella, D’Ariano, & Perinotti, 2010), “information causality” (Pawłowski, Paterek, Kaszlikowski, Scarani, Winter, & Żukowski, 2009), “purification” (D’Ariano, Chiribella, & Perinotti, 2017) or what have you. These principles carve up the conceptual space of a given framework into those theories that satisfy them with respect to the phenomena considered, and those that do not. The correlations predicted by quantum theory, for instance, satisfy the information causality principle, but any theory that allows correlations above the Tsirelson bound corresponding to the CHSH inequality does not (for discussion, see Cuffaro, 2020).

It may sometimes even be possible to uniquely characterize a theory in a given context—to fix the point in a framework’s conceptual space that is occupied by the theory—and if the principles from which such a unique characterization follows are sufficiently compelling in that context, then situating the theory within it adds to our understanding both of the theory and of the

¹⁷ See the end of Section 6.1 for a discussion of the way that our characterization of the principle-theoretic and constructive approaches differs from other ways in which they have been characterized in the literature.

¹⁸ For more on all of these and other related topics, see the collection of essays edited by Chiribella & Spekkens (2016).

phenomena described by it (cf., e.g., Masanes, Galley, & Müller, 2019).¹⁹ We are not, of course, claiming that this or that abstract characterizing principle exhausts all that there is to say about quantum theory. But by situating quantum theory within the abstract space provided by a formal framework we subject it to a kind of “theoretical experiment.”

In the course of an actual experiment that has been set up to determine this or that property of a physical system, we control for (i.e., in our lab) parameters that we deem irrelevant to or that interfere with our determination of the particular property of interest. Likewise, in our theoretical experiments we abstract away from features of quantum theory that are irrelevant to or obfuscate our characterization of it *as a theory of information processing of a particular sort*, or *as a particular kind of C^* algebra*, or *as a theory of probabilities* and so on. Quantum theory can be thought of as each of these things. Insofar as it occupies a particular position (or region) within the conceptual space of these respective frameworks, it can be characterized from each of these points of view. And for each perspective within which it can be so characterized, there are constraints on what a quantum system can be from that perspective. It is these constraints which our theoretical experiments set out to discover. And it is these constraints which convey to us information about what quantum theory *is* and how the systems it describes *actually behave* under that mode of description.

The value of the principle-theoretic approach is, moreover, not limited to this descriptive role. Principle-theoretic approaches are also instrumental for the purposes of theory development. For instance in the course of setting up a conceptual framework in which to situate quantum theory, we might consider it more natural to relax rather than maintain one or more of the principles that characterize quantum theory in that framework (cf. Hardy, 2007). In this way we feel our way forward to new physics. Even, that is, if we do not expect that they themselves will constitute new physical theories, the formal frameworks we set up enable theoretical progress by helping us grasp the descriptive limits of our existing theories and get a sense of what lies beyond them.

And yet, earlier we stated our conviction that, “at its core,” quantum mechanics is fundamentally *about* probabilities. How can this univocal statement be consistent with the claim we have just been making regarding the essentially perspectival nature of the insights obtainable through a principle-theoretic ap-

¹⁹ One of us has expressed previously in print the contention that only constructive approaches to physics can yield explanatory content (Balashov & Janssen, 2003, Janssen, 2009). All three of us are now of the opinion that both principle-theoretic and constructive approaches can be explanatory.

proach? In fact it would be wrong to describe the interpretation of quantum theory that we have been advancing in this volume as a principle-theoretic one.²⁰ As we have described them above, principle-theoretic approaches offer perspectives on quantum theory (or on some aspect of it) that are essentially external: One first sets up a formal framework which in itself has little to do with quantum theory; next one seeks to motivate and define a principle or set of them with which to pinpoint quantum theory within that framework. But what does it mean to pinpoint quantum theory within a framework? Generally this means matching the set of phenomena circumscribed by the principle(s) with the set of phenomena predicted by quantum mechanics, i.e., with those obtained via a derivation *from within*. In this way one tests that the set of phenomena captured by a set of principles really is the one predicted by quantum mechanics, that these characterizing principles really do constitute a perspective on the theory. A principle-theoretic approach to understanding quantum theory, therefore, is not wholly external. But on the approach just outlined the internal perspective only becomes relevant at the end of the procedure, as a way to gauge the success of one's theoretical experiment.

For us this latter step was very far from trivial. Indeed it was only through it that we were able to gain full insight into the aspect of quantum phenomena that we were seeking to understand. To recapitulate: We first set up a generalized framework for characterizing correlations and within this framework we considered the elliptope inequality relating correlation coefficients for linear combinations of random variables expressed by Eq. (6.1.6). We thus began our derivation *from without*. We then asked whether one could view this as an expression of the fundamental nature of the correlations between random variables in either a local hidden-variables model such as our raffles, or in a quantum model. That is, we asked whether the correlations in either case saturate the elliptope described by Eq. (6.1.6). To answer this question we then took a constructive step in both cases: We gave both a local hidden-variables and a quantum model for the general constraint expressed by Eq. (6.1.6). And by proceeding in this way *from within* both frameworks we were able to show that, as a consequence of the assumptions underlying the framework of classical probability theory, the elliptope inequality cannot be seen as a fundamental expression of the nature of correlations in a local hidden-variables model, for

²⁰ Ours is not a principle-theoretic interpretation in the way that we have expounded that term here. As discussed at the end of Section 6.1, our own usage of the term is intended to reflect the way it is used in the contemporary literature on quantum foundations. Our interpretation could, though, be seen as a principle-theoretic one in the sense in which (for instance) Bill Demopoulos uses that term.

additional inequalities need to be satisfied in such a model and these further constrain the allowable correlations in the model to polyhedra inscribed within the ellipsope. As for the quantum model, we saw how the constraints imposed on it by the mathematical framework of quantum theory enable it to succeed, where a local hidden-variables model cannot, in entirely filling up the ellipsope. Finally by considering *how* it is capable of doing this we are able to understand what the essential distinction between quantum and classical theory is.

6.3 The new kinematics of quantum theory

What, then, is the essential distinction between quantum and classical theory? In the end we saw that the key assumption we needed to derive the quantum version of the ellipsope inequality is one that follows straightforwardly from quantum theory's kinematical core: from the Hilbert space formalism of quantum mechanics that applies universally to all of the dynamical systems described by it.²¹ Our case studies were limited to a relatively small number of particular experimental setups—the Mermin-style setups we considered in Chapters 2–4 and the CHSH-like setup we considered in Chapter 5. They were also limited in terms of the quantum states measured in those setups. But we see now that the wider significance of our analyses of these case studies is not likewise limited. For the key feature of the quantum formalism that these special but informative case studies point us to is in fact a fully general one.

As mentioned in Chapter 1, our interpretation of quantum theory owes much to the work of Jeff Bub, Bill Demopoulos, Itamar Pitowsky and others who have proceeded from similar motivations. In their 2010 paper, Bub and Pitowsky characterize their interpretation of quantum theory as both principle- and information-theoretic (pp. 445–446), arguing both that the Hilbert space structure of quantum theory is derivable on the basis of information-theoretic constraints, and that quantum theory should in this sense be thought of as being all about information (cf. Bub, 2004). Misinterpreted as some sort of ontological claim, the latter is surely false. If, instead, one correctly interprets this as a claim about where the conceptual novelty of quantum theory is located (cf. Demopoulos, 2018), namely in the structural features of its kinematical core, then we take this claim to be correct, even if we prefer to speak of *probability* rather than of information (see Chapter 1).

There is a common viewpoint on interpretation that holds that what it means to interpret a theory is to ask the question: “What would the world be like if the

²¹ See Janssen (2009) for similar observations about the kinematical core of special relativity.

theory were literally true of it?” (cf., for instance, Caulton, 2015, p. 153). We reject this (cf. also Curriel, 2019) as exhaustive of what it can mean to interpret a theory, and rather affirm that often the more interesting interpretational question is the one which asks what the world must be like in order for a given theory to be of use to us;²² i.e., to be effective in describing and structuring our experience and in enabling us to speak objectively about it to one another. Note, on the one hand, the realist commitment implicit in this question. But note, on the other hand, that the question does not presuppose the literal or even the approximately literal truth of the theory being considered. For even classical mechanics, superseded as it has been by quantum mechanics, is of use to us in this sense. And it is a meaningful question to ask how this constrains our possible representations of the world.

Such a question can be answered in a number of ways. One might begin, for instance, by positing a priori constraints on what an underlying ontological picture of the world must be like in a general sense, e.g., that it must be some kind of particle ontology (cf. Albert, 2018). The descriptive success of quantum mechanics (and, correspondingly, the descriptive failure of classical mechanics) would then entail a number of constraints on this general ontological picture, in particular that it must be fundamentally non-local. Alternately (i.e., rather than positing a general ontological picture of the world a priori) one might choose instead to focus more directly on the relation between the formalism of the theory and the phenomena it describes. What aspect of the formalism, one might ask from this point of view, is key to enabling quantum theory to be successful in describing phenomena and coordinating our experience, and what does that tell us about the world? A natural way of illuminating this question is to compare quantum with classical modes of description—to consider what is *novel* in the quantum as compared with the classical mode of description—and to consider how this allows quantum theory to succeed where classical theory fails. We take the investigations in the prior chapters of this volume to have shown that this novel content can be located in the kinematical core of quantum theory, in the structural constraints that quantum mechanics places on our representations of the physical systems it describes.²³

²² This is a point that Wayne Myrvold (in personal correspondence) has emphasized on a few different occasions.

²³ Cf. Bohr (1948, p. 316), who writes: “In representing a generalization of classical mechanics suited to allow for the existence of the quantum of action, Quantum mechanics offers a frame *sufficiently wide* to account for empirical regularities *which cannot* be comprised in the classical way of description” (our emphasis). To use a literary metaphor, we can think of quantum mechanics as a new way of writing that allows us to express nuances we could

In classical mechanics, an observable A is represented by a function on the phase space of a physical system: $A = f(q_i, p_i)$ where the q_i and p_i are the system's generalized coordinates and conjugate momenta, respectively, on its phase space. Points on this space can be thought of as “truthmakers” (Bub & Pitowsky, 2010) for the occurrence or non-occurrence of events related to the system in the sense that assigning particular values to the q_i and p_i fixes the values of every other observable defined with respect to the system in question. With each observable A one can associate a Boolean algebra representing the possible yes or no questions that can be asked concerning that observable in relation to the system. And because one can simultaneously assign values to every observable given the state specification (q_i, p_i) , one can embed the Boolean algebras corresponding to each of them within a global Boolean algebra that is the union of them all. In general there is no reason to think of observables as representing the properties of a physical system within this framework. But because we can fix the value of every observable associated with the system *in advance* given a specification of the system's state—because the union of the Boolean algebras corresponding to these observables is itself representable as a Boolean algebra—it is in this case conceptually unproblematic to treat these observables as though they do represent the properties of the system, properties that are possessed by that system irrespective of how we interact (or not) with it.

In quantum mechanics an observable, A , is represented by a Hermitian operator, \hat{A} (whose spectrum can be discrete, continuous or a combination of both) acting on the Hilbert space associated with a physical system, with the possible values for A given by the eigenvalues of \hat{A} . Unlike the case in classical mechanics, the quantum state specification for a physical system, $|\psi\rangle$, cannot be thought of as the truthmaker for the occurrence or non-occurrence of events related to it, for specifying the state of a system at a given moment in time does not fix in advance the values taken on at that time by every observable associated with the system. First, the state specification of a system yields, in general, only the probability that a given observable associated with it will take on a particular value when selected. Second, and more importantly, the Boolean algebras corresponding to the observables associated with the system cannot be embedded into a larger Boolean algebra comprising them all. Thus

not express before. We are reminded of what E. M. Forster (1942, p. 28) once said about Virginia Woolf: “She pushed the light of the English language a little further against the darkness.” Thinking about Forster's observation in the context of quantum mechanics, we can imagine “the darkness” to stand for the intrinsic randomness all our quantum-mechanical descriptions of the world eventually come up against.

one can only say that *conditional* upon the selection of the observable A , there will be a particular probability for that observable to take on a particular value. At the same time no one of the individual Boolean sub-algebras of this larger non-Boolean structure yields what would be regarded, from a classical point of view, as a complete characterization of the properties of the system in question. As we will see in Section 6.5, this does not preclude a different kind of completeness from being ascribed to the quantum description of a system. But because our characterization is not *classically* complete, it is no longer unproblematic to take the observable A as a stand-in for one of the underlying properties of the system, even in the case where quantum mechanics predicts a particular value with certainty conditional upon a particular measurement.²⁴

To put it a different way: Because classical-mechanical observables can be set down in advance, irrespective of the nature of the interaction with the system from which they result, they can straightforwardly be taken to represent “beables” (see Bell, 1984, sec. 2) with respect to a given state specification. Quantum-mechanical observables cannot be, or at any rate there can be no direct, unproblematic, inference from observable to beable within quantum theory—something more, some further argument must be given.²⁵ As for us, we have yet to see a convincing argument to this effect.²⁶ We rather take quantum theory to be telling us that there is no ground in the classical sense of a fully determinate globally Boolean noncontextual assignment of values to all of the observable parameters relevant to a given system (cf. Pitowsky, 1994, sec. 9). *In this specific sense*, then, we take quantum state specifications to be non-representational (cf. Bub & Pitowsky, 2010, p. 433). What *is* represented by a quantum state specification is a particular space of possibilities for a given system conditional upon its interactions with external systems.²⁷ This is in fact no different than in classical mechanics. But whereas

²⁴ Contrast this statement with Einstein et al.’s famous criterion for physical reality: “If, without in any way disturbing a system, we can predict with certainty (i.e., with probability equal to unity) the value of a physical quantity, then there exists an element of physical reality corresponding to this physical quantity” (1935, p. 777, emphasis in original).

²⁵ Arguably this is the real significance of von Neumann’s much-maligned proof of the impossibility of hidden-variables theories; i.e., that the beables of a deterministic hidden-variables theory cannot be represented by Hermitian operators in Hilbert space. For further discussion, see Bub (2010a) and Dieks (2017).

²⁶ An example of something further that could be appealed to will be discussed in the next chapter when we briefly compare our informational interpretation with the Everett family of views.

²⁷ We will come back to this question again in Section 6.5.

the structure of the classical-mechanical possibility space is such that it invites the inference to a posited underlying physical system as the bearer of a globally Boolean collection of properties, the more complex structure of the quantum-mechanical possibility space does not similarly invite the inference to an underlying system in that sense.²⁸

6.4 Examples of problems solved by the new kinematics

In the context of space-time theories, Minkowski space-time encodes generic constraints on the space-time configurations allowed by any specific relativistic theory compatible with its kinematics. These constraints are satisfied as long as all of the observables are represented by mathematical objects that transform as tensors (or spinors) under Lorentz transformations. Analogously, in quantum mechanics, Hilbert space encodes generic constraints on the possible values of observables as well as on the correlations between such values that are allowed within any specific quantum theory compatible with its kinematics. These constraints are satisfied as long as all of the observables are represented by Hermitian operators acting on Hilbert space. In the case of Minkowski space-time, the determination of the particular tensor (or spinor) representative of a given observable is the province of the dynamics, not the kinematics, of the specific relativistic theory in question. Likewise, determining the particular operator representative of a given action on a system is a province of the dynamics, not the kinematics, of the specific quantum theory in question. Just as in special relativity, the kinematical part of quantum theory is a comparatively small one. The lion's share of the practice of quantum theory is concerned with determining the particular dynamical aspects of particular systems of interest. Yet, conceptually, the kinematics of quantum theory may justifiably be regarded as its most important part; it constitutes the "operating system" upon which the dynamics of particular systems can be seen as "applications" being run (Aaronson, 2013, p. 110; Nielsen & Chuang, 2016, p. 2).

²⁸ Cf. Bernard d'Espagnat, who writes that "classical physics, considered as a universal theory, was ontologically interpretable. This does not mean that such an interpretation was logically necessary. It was not. But it does mean that it was admissible . . . quantum mechanics as we know it is not ontologically interpretable. This is not necessarily to be considered as a defect but it implies that, in the realm of interpretational problems such as the one here on hand, we should not argue as if it were. In particular the 'collapse riddle' should prevent us from tacitly assuming that the wave function possesses in every circumstances [sic.] all the attributes of reality" (d'Espagnat, 2001, pp. 7–8).

As in the transition from 19th-century ether theory to special relativity (see [Janssen, 2009](#)), one can find in the transition from the old to the new quantum theory examples of puzzles solved as a direct result of considering the changes introduced via the latter's novel kinematics. It should come as no surprise, given how we characterized Heisenberg's and Schrödinger's respective "big discoveries" in [Chapter 1](#), that these examples are easier to come by in the early history of matrix mechanics than in the early history of wave mechanics, but they can be found in both.

The basic idea of the paper with which [Heisenberg \(1925\)](#) laid the foundation of matrix mechanics was not to repeal the laws of classical mechanics but to reinterpret them ([Janssen, 2019](#), p. 139). This is clearly expressed in the title of the paper: "Quantum-theoretical reinterpretation (*Umdeutung*) of kinematical and mechanical relations." Heisenberg replaced the real numbers p and q by non-commuting arrays of numbers soon to be recognized as matrices and then as operators. These operators, \hat{p} and \hat{q} , satisfy the same relations as p and q (e.g., the functional dependence of the Hamiltonian on these variables will remain the same) but they are subject to the commutation relation, $[\hat{q}, \hat{p}] = i\hbar$, the quantum analogue, as [Dirac \(1926\)](#) realized early on, of Poisson brackets in classical mechanics.

In the final section of the *Dreimännerarbeit*, the joint effort of Max Born, Heisenberg and Jordan that consolidated matrix mechanics, the authors (or rather Jordan who was responsible for this part of the paper) showed that the new formalism automatically yields both terms of a famous formula for energy fluctuations in black-body radiation ([Born, Heisenberg, & Jordan, 1926](#), pp. 375–385).²⁹ [Einstein \(1909a,b\)](#) had derived this formula from little more than the connection between entropy and probability expressed in the formula $S = k \ln W$ carved into Boltzmann's tombstone and Planck's law for black-body radiation. One of its two terms suggested waves, the other particles. Einstein had argued in 1909 that the latter called for a modification of Maxwell's equations ([Duncan & Janssen, 2019](#), pp. 120–126). He had contemplated such drastic measures before when faced with the tension between Maxwell's equations and the relativity principle. The new kinematics of special relativity had resolved that tension. Jordan now showed that the tension between Maxwell's equations and Einstein's fluctuation formula could also be resolved by a change in the kinematics.

²⁹ For a detailed reconstruction of Jordan's argument, see [Duncan & Janssen \(2008\)](#). The ensuing debate over this reconstruction (see, especially, [Bacciagaluppi, Crull, & Maroney, 2017](#)) does not, as far as we can tell, affect our use of this example in the present context.

Instead of a cavity with electromagnetic waves obeying Maxwell's equations, Jordan considered a simple model, due to Paul Ehrenfest (1925), of waves in a string fixed at both ends. This string can be replaced by an infinite number of uncoupled harmonic oscillators, one for every mode of the string. Quantizing those oscillators, using the basic commutation relation $[\hat{q}, \hat{p}] = i\hbar$, and calculating the fluctuation of the energy in a small segment of the string in a narrow frequency interval, Jordan recovered both the wave and the particle term of Einstein's formula. Using classical kinematics, one only finds the wave term. As Jordan concluded:

The reasons for the occurrence of a term not delivered by the classical theory are obviously closely related to the reasons for the occurrence of the zero-point energy [of the harmonic oscillator, which itself follows directly from the commutation relation for position and momentum]. In both cases, the basic difference between the theory attempted here and the one attempted so far [i.e., classical theory with the restrictions imposed on it in the old quantum theory] *lies not in a disparity of the mechanical laws but in the kinematics characteristic for this theory*. One could even see in [this fluctuation formula], into which no mechanical principles whatsoever even enter, one of the most striking examples of the difference between quantum-theoretical kinematics and the one used hitherto (Born et al., 1926, p. 385; our emphasis and our translation, quoted in part in Janssen, 2009, p. 50).

Our second example turns on the quantum-mechanical treatment of orbital angular momentum, which proceeds along the exact same lines as the treatment of intrinsic or spin angular momentum underlying the quantum-mechanical analysis of the experiments we have been studying in Chapters 2–5. We already alluded to this example at the end of Section 3.2. It is the problem of the electric susceptibility of diatomic gases such as hydrogen chloride.³⁰ One of the two terms in the so-called Langevin-Debye formula for this quantity comes from the alignment of the molecule's permanent dipole with the external field. This term decreases with increasing temperature as the thermal motion of these dipoles frustrates their alignment. This makes it at least intuitively plausible that only the lowest energy states of the molecule contribute to the susceptibility. This is indeed what the classical theory predicts. In the old quantum theory, however, this feature was lost. This is a direct consequence of the way in which angular momentum was quantized. The length L of the angular momentum vector could only take on values $l\hbar$ in the old quantum theory, where l is an integer greater than 1. The value $l = 0$ was ruled out for the same reason that it was ruled out for the hydrogen atom: an orbit with zero angular momentum would have to be a straight line going back and forth through the nucleus! Hence

³⁰ For a detailed analysis of this episode, see Midwinter & Janssen (2013).

$l \geq 1$ for all states contributing to the susceptibility. This led to the strange situation, as Pauli (1921, p. 325) noted in one of his early papers, that there are “*only such orbits present that according to the classical theory do not give a sizable contribution to the electrical polarization*” (emphasis in the original). Fortunately, the allowed orbits (or energy states) with $l \geq 1$ do give a sizable contribution. Unfortunately, this contribution is almost five times too large.

The quantization of angular momentum in the new quantum mechanics was worked out in the *Dreimännerarbeit* mentioned above (Born et al., 1926, pp. 364–374). The upshot was that the correct quantization of angular momentum leads to the eigenvalues $l(l+1)$ for \hat{L}^2 , where the allowed integer values of l start at 0 rather than 1 (cf. Eq. (4.1.2)). This new quantization rule for angular momentum follows directly from the basic commutation relation for position and momentum.

Pauli and his former student Lucy Mensing showed how this new quantization rule solved the puzzle of the electric susceptibility of diatomic gases.³¹ As in classical theory, only the lowest ($l = 0$) state contributes to the susceptibility, the contributions of all other terms sum to zero (and this depends delicately on the exact quantization rule). As Mensing & Pauli (1926, p. 512) noted with palpable relief: “*Only the molecules in the lowest state will therefore give a contribution to the temperature-dependent part of the dielectric constant*” (emphasis, once again, in the original). The new quantum theory thus reverted to the classical theory in this respect. In a note to *Nature* on the topic, Van Vleck made the same point: “The remarkable result is obtained that only molecules in the state of lowest rotational energy make a contribution to the polarisation. This corresponds very beautifully to the fact that in the classical theory only molecules with [the lowest energy] contribute to the polarisation” (Van Vleck, 1926, p. 227).

Van Vleck expanded on this comment when interviewed in 1963 by his former PhD student Thomas S. Kuhn for the Archive for History of Quantum Physics (AHQP):

I showed that [the Langevin-Debye formula for susceptibilities] got restored in quantum mechanics, whereas in the old quantum theory, it had all kinds of horrible oscillations . . . you got some wonderful nonsense, whereas it made sense with the new quantum mechanics. I think that was one of the strong arguments for quantum mechanics. One always thinks of its effect and successes in connection with spectroscopy, but I remember Niels Bohr saying that one of the great arguments

³¹ Mensing gave up her career in physics when she married a fellow physicist in 1930 (Münster, 2020).

for quantum mechanics was its success in these non-spectroscopic things such as magnetic and electric susceptibilities.³²

Van Vleck was so taken with this result that it features prominently in his Nobel lecture in 1977 (Midwinter & Janssen, 2013, p. 138). The important point for our purpose is that this is another example of a problem that was solved by a change in the kinematics rather than the dynamics.

The two examples given so far both turned on the commutation relation $[\hat{q}, \hat{p}] = i\hbar$ at the heart of matrix mechanics. Our third and last example turns on a key feature of wave mechanics. As we noted in Chapter 1, Schrödinger, unlike Heisenberg, may not have emphasized that his new theory provided a new framework for doing physics but this is, of course, as true for wave mechanics as it is for matrix mechanics. One obvious example of a change in the basic framework for doing physics that emerged from the development of wave mechanics rather than matrix mechanics is the introduction of quantum statistics, especially Bose-Einstein statistics, which preceded the formulation of wave mechanics. We close this section with a somewhat less obvious, but particularly informative, further example.³³

In the same year that saw the appearance of Bohr's atomic model, Johannes Stark (1913) discovered the effect named after him, the splitting of spectral lines due to an external electric field, the analogue of the effect discovered by Pieter Zeeman in 1896, the splitting of spectral lines due to a magnetic field. It was not until two key contributions—one by a physicist, Arnold Sommerfeld, in late 1915; one by an astronomer, Karl Schwarzschild, in early 1916—that there was any hope of accounting for the Stark effect on the basis of the old quantum theory. Sommerfeld's key contribution to the explanation of the Stark effect was to introduce (even though he did not call it that) *degeneracy*, the notion that the same energy level can be obtained with different combinations of quantum numbers. External fields will lift this degeneracy and result in a splitting of the spectral lines associated with transitions between these energy levels. Schwarzschild's key contribution was to bring the advanced techniques developed in celestial mechanics to bear on the analysis of the miniature planetary systems representing atoms in the old quantum theory. Once those two ingredients were available, Schwarzschild (1916) and Paul Epstein (1916),

³² Cf. the opening sentence of the preface of his classic text on magnetic and electric susceptibilities (Van Vleck, 1932) quoted in note 14 in Section 3.2, the book that earned him the informal title of “father of modern magnetism” (Midwinter & Janssen, 2013, p. 139).

³³ For a detailed analysis of this episode, see Duncan & Janssen (2014, 2015, 2019, sec. 6.3 and Appendix A).

an associate of Sommerfeld, quickly and virtually simultaneously derived formulas for the line splittings in the Stark effect in hydrogen that were in excellent agreement with the experimental data.

Even though some energy states and some transitions between them had to be ruled out rather arbitrarily and even though there was no convincing explanation for the polarizations and relative intensities of the components into which the Stark effect split the spectral lines, this was seen as a tremendous success for the old quantum theory. As Sommerfeld exulted in the conclusion of the first edition of *Atombau und Spektrallinien* (Atomic structure and spectral lines), which became known as the “the bible of atomic theory” (Eckert, 2013, pp. 255–256): “the theory of the Zeeman effect and especially the theory of the Stark effect belong to the most impressive achievements of our field and form a beautiful capstone on the edifice of atomic physics” (Sommerfeld, 1919, pp. 457–458).

Even in the case of the Stark effect (to say nothing of the Zeeman effect), Sommerfeld’s jubilation would prove to be premature. In addition to the limitations mentioned above, there was a more subtle but insidious difficulty with Schwarzschild and Epstein’s result. To find the line splittings of the Stark effect, they had to solve the so-called Hamilton-Jacobi equation, familiar from celestial mechanics, for the motion of an electron around the nucleus of a hydrogen atom immersed in an external electric field. This could only be done in coordinates in which the Hamilton-Jacobi equation for this problem is *separable*, i.e., in coordinates in which the equation splits into three separate equations, one for each of the three degrees of freedom of the electron. Similar problems in celestial mechanics made it clear that they needed so-called parabolic coordinates for this purpose. These then were also the coordinates in which Schwarzschild and Epstein imposed the quantum conditions to select a subset of the orbits allowed classically. As long as there is no external electric field, it was much simpler to do the whole calculation in polar coordinates. Letting the strength of the external field go to zero, one would expect that the quantized orbits found in parabolic coordinates reduce to those found in polar coordinates. This turns out not to be the case. The energy levels are the same in both cases but the orbits are not. Both Sommerfeld and Epstein recognized that this is a problem (Schwarzschild died the day his paper appeared in the proceedings of the Berlin academy). As Epstein (1916, p. 507) put it:

Even though this does not lead to any shifts in the line series, the notion that a preferred direction introduced by an external field, no matter how small, should drastically alter the form and orientation of stationary orbits seems to me to be unacceptable (quoted in Duncan & Janssen, 2015, p. 251).

The old quantum theory, however, simply did not have the resources to tackle this problem and nothing was done about it.

The Stark effect in hydrogen was one of the first applications of Schrödinger's new wave mechanics. The calculation is actually very similar to the one in the old quantum theory. This is no coincidence. An important inspiration for Schrödinger's wave mechanics was Hamilton's optical-mechanical analogy (Joas & Lehner, 2009). So it is not terribly surprising that Hamilton-Jacobi theory informed the formalism Schrödinger came up with. The time-independent Schrödinger equation was actually modeled on the Hamilton-Jacobi equation. The time-independent Schrödinger equation for an electron in a hydrogen atom in an external electric field is, once again, most easily solved in parabolic coordinates. Independently of one another, Schrödinger (1926), in Part III of his famous four-part paper, "Quantization as an eigenvalue problem," and Epstein (1926) did this calculation. To first order in the strength of the electric field, it yields the same splittings as the old quantum theory. However, as both Schrödinger and Epstein emphasized, no additional restrictions on states or transitions between states are necessary and the theory also correctly predicts the polarizations and intensities of the various Stark components. Epstein did not address the problem of the non-uniqueness of orbits of the old quantum theory but Schrödinger clearly realized that wave mechanics solves this problem too. In Part II of his paper, he had already put his finger on the key to the solution:

We recognize here a complete parallel to the familiar circumstances we encounter with the old quantization method in the case of *degeneracy*. There is only one not unwelcome difference. If we applied the Sommerfeld-Epstein quantum conditions *without* taking into account the possibility of degeneracy, we did, it is true, always get the same energy levels, but, depending on which coordinates we chose, we arrived at different results for the allowed orbits. Now that is *not* the case here. It is true that we arrive at a completely different set of eigenfunctions if, for example, we treat the vibration problem corresponding to the unperturbed Kepler motion in *parabolic* coordinates instead of the polar coordinates we used in Part I. A possible *vibration state*, however, is not just a *single proper mode* [*Eigenschwingung*] but an arbitrary finite or infinite *linear aggregate* of these. And the eigenfunctions found one way [i.e., using one coordinate system] can always be expressed as linear aggregates of the eigenfunctions found any other way [i.e., using a different coordinate system], as long as these eigenfunctions form a *complete* set (Schrödinger, 1926, pp. 33–34; our translation; italics in the original).

Translated into the language of modern quantum mechanics, what Schrödinger is saying here is that the problematic non-uniqueness of orbits in the old quantum theory turns into a completely innocuous non-uniqueness of bases of

eigenstates in the new quantum theory (Duncan & Janssen, 2014, sec. 5, pp. 76–77).³⁴

Both Heisenberg and Schrödinger recognized the problematic nature of the old quantum theory's electron orbits, which had been imported from celestial mechanics along with the mathematical machinery to analyze atomic structure and atomic spectra (Janssen, 2019, p. 171). An area in which the trouble with orbits had become glaringly obvious by the early 1920s was optical dispersion, the study of the dependence of the index of refraction on the frequency of the refracted light. Heisenberg's *Umdeutung* paper builds on a paper he co-authored with Hans Kramers, Bohr's right-hand man in Copenhagen, on Kramers's new quantum theory of dispersion (Kramers & Heisenberg, 1925). Taking his cue from this theory, Heisenberg steered clear of orbits altogether in his *Umdeutung* paper and focused instead on observable quantities such as frequencies and intensities of spectral lines (Duncan & Janssen, 2007, Janssen, 2019, pp. 134–142). The quantities with which he replaced position and momentum were not, in his original scheme, themselves observable. Instead they functioned as auxiliary quantities that allowed him to calculate the values of (indirectly) observable quantities such as energy levels and transition probabilities. Schrödinger did not get rid of orbits as radically as Heisenberg. His wave functions can be seen as a new way to characterize atomic orbits once we have come to recognize that they are the manifestation of an underlying wave phenomenon. Comparing these different responses to the trouble with orbits in the old quantum theory, we see the beginnings of the two main lineages of the genealogy we proposed in Chapter 1 to classify different interpretations of quantum mechanics. This will figure importantly in the next and last section of this chapter.

6.5 Measurement

We began this chapter in Section 6.1 with a survey of the ground that we covered in Chapters 2–5. Then in Section 6.2 we reflected in particular on the

³⁴ It was only after this paper was published that the authors realized that Schrödinger had essentially already found this solution to the problem. In the case of special relativity, it took longer for physicists to recognize that some puzzles had been resolved by the new kinematics. In the case of the Trouton-Noble experiment, for instance, Butler (1968) was the first to show that the torque on a moving capacitor that the experimenters had been looking for in 1903 was nothing but an artifact of how one slices Minkowski space-time when defining the momentum and angular momentum of spatially extended systems (Janssen, 2009, p. 45; see also Teukolsky, 1996).

nature of our derivation, summarized in Section 6.1, of the space of possible quantum correlations in the Mermin-style setups we considered in Chapters 2–4, noting how our derivation evinces aspects of both the principle-theoretic and constructive approaches to physics. In Section 6.3 we argued that our derivation yields the insight that the fundamental novelty of the quantum mode of description is in the kinematical and not in the dynamical part of quantum theory. In Section 6.4 we considered examples, from the history of quantum theory, of physical problems that were resolved by considering the changes to the kinematical framework of fundamental physics introduced by quantum mechanics. In this final section of this chapter we consider the topic of measurement.

Though we will locate them elsewhere than is standardly done, we will conclude that philosophical puzzles yet remain in relation to the physics of measurement in quantum mechanics. In itself this is not a reason to reject quantum mechanics, for every candidate fundamental theoretical framework has its puzzles (otherwise it would be more than just a candidate).³⁵ As for the most important of these puzzles in relation to quantum mechanics' account of a measurement, we will argue below that it does *not* stem from the fact, that of the various possible values obtainable through the measurement of a selected observable, the one that will actually be obtained is, according to quantum mechanics, fundamentally a matter of chance. The more important puzzle of measurement is that within quantum mechanics it is not even possible to describe all of our experience consistently in accordance with a (classical) probability distribution at all. We will see below, both that the physical significance of this puzzle of measurement, as well as the physical account quantum mechanics provides of particular measurements, flow naturally from the constraints that quantum mechanics' kinematical core imposes on our representations of quantum systems, and we will reflect on the conception of the world that we take these constraints to suggest.

Consider a measurement device that has been set up to assess the spin state of electrically neutral spin- $\frac{1}{2}$ particles, such as silver atoms, that have been prepared in a particular way. For instance, imagine preparing a number of silver atoms all in the superposition of z -basis states given by (cf. Section 2.6):

$$|\varphi\rangle = \alpha|+\rangle_z + \beta|-\rangle_z, \quad (6.5.1)$$

³⁵ See note 14 above for an explanation of what we mean when we say that quantum mechanics is a candidate fundamental theoretical framework.

where the coefficients α and β are complex numbers subject to the constraint $|\alpha|^2 + |\beta|^2 = 1$. We direct the silver atoms one at a time toward the device, which we have prepared so that it will “measure spin in the z -direction.” We will unpack the physical meaning of the just-quoted sentence in a few paragraphs. For now (but only for now) we treat the measurement device as essentially a black box. Upon observing the result, m_z , of any one of our experiments, we note that the spin state of a silver atom is always recorded as either ‘up’ ($m_z = +$) or ‘down’ ($m_z = -$),³⁶ and further that the outcomes are statistically distributed so that the relative frequencies of up and down tend toward $|\alpha|^2$ and $|\beta|^2$, respectively. What, one might ask, is the explanation for these facts?

Here is an attempt. If S is an electrically neutral spin- $\frac{1}{2}$ particle in the state given by Eq. (6.5.1), then according to the Born rule the probability assigned to a particular measurement outcome $m_z \in \{+, -\}$ is given (in the case of a projective measurement in the z -basis) by:

$$\Pr(m_z|\hat{z}) = \langle \varphi | \hat{P}_{m_z} | \varphi \rangle, \quad (6.5.2)$$

where

$$\hat{P}_{m_z} \equiv |m_z\rangle_z \langle m_z| \quad (6.5.3)$$

is the *projection operator* or *projector* corresponding to the outcome m_z . Eq. (6.5.2) can be re-expressed as follows:

$$\Pr(m_z|\hat{z}) = \langle \varphi | \hat{P}_{m_z} | \varphi \rangle = \langle \varphi | m_z \rangle_z \langle m_z | \varphi \rangle = |{}_z \langle m_z | \varphi \rangle|^2. \quad (6.5.4)$$

For a spin- $\frac{1}{2}$ particle in the state given by Eq. (6.5.1) in particular, we see that:

$$\Pr(+|\hat{z}) = |{}_z \langle + | \varphi \rangle|^2 = |\alpha|^2, \quad \Pr(-|\hat{z}) = |{}_z \langle - | \varphi \rangle|^2 = |\beta|^2 \quad (6.5.5)$$

in accordance with how we have been using the Born rule in Chapters 2 and 4 (see, e.g., Eq. (2.6.30) and Eq. (4.1.36)). With high probability (which increases as we increase the number of systems assessed) we will find $\Pr(+|\hat{z})$ and $\Pr(-|\hat{z})$ to be in accord with the statistics that we actually observe when we assess the spin states of a number of silver atoms that have all been prepared in the state given by Eq. (6.5.1).

There are parallels between the silver atoms we are considering now and the various baskets of raffle tickets that we used to model classical statistical ensembles in Sections 2.5 and 4.2. To bring out what these parallels are,

³⁶ “+” and “-”, as they pertain to the spin- $\frac{1}{2}$ particle S , should be read as $\pm\hbar/2$.

we follow [Von Neumann \(1927b\)](#) (and indirectly, von Mises, whose work influenced von Neumann; see [Duncan & Janssen, 2013](#)) in introducing the concept of a *quantum statistical ensemble*, or what we have colorfully referred to in the title of this book as a “quantum raffle” (see also Chapter 1): a large number of similar quantum systems S_i for which the state vector corresponding to the i^{th} system, S_i , is $|\psi_i\rangle$. Mathematically, a quantum statistical ensemble is characterized by what we now call a *density operator*, commonly denoted by $\hat{\rho}$.

It is useful to think of each individual system S_i in an ensemble as a ‘copy’ of some reference system S , if one understands that ‘copy’ in this instance does not mean clone. A system, S_i , is like S only in the sense that it is a system of the same type. In our own example we are dealing with an ensemble of silver atoms. As for what is actually reproduced in every case, this is a particular preparation procedure through which the ensemble is progressively generated. Such a preparation procedure can be probabilistic. For instance we might agree to prepare a given S_i in the state $|\psi_i\rangle$ if the result of the i^{th} flip of a coin is heads, and in the state $|\psi_i'\rangle$ if it is tails. A procedure of this sort generates a *non-uniform* ensemble, characterized by a *mixed* state—specifically a *properly* mixed state. A non-uniform ensemble can be regarded as composed of a number of distinct sub-ensembles, each of which has been generated using a different preparation procedure.

The procedure for generating an ensemble in an *improperly* mixed state is more general than the one given above. A concrete example of such a procedure is (i) preparing an ensemble of entangled pairs of spin- $\frac{1}{2}$ particles in the singlet state, and then (ii) choosing the left-hand particle in each pair. The empirical difference between a proper and an improper mixture will become evident when we consider the interaction between a measurement device and a system in more detail below.³⁷ At this stage what is most pertinent is that the statistics conditional upon particular measurements on the members of an ensemble in an improperly mixed state are not distinguishable from the statistics conditional upon those same measurements on an ensemble of systems that has been properly mixed in a particular way. Thus, conditional on a particular measurement, both a properly and an improperly mixed state can effectively be considered as grounded in an ensemble that is non-uniform; i.e., one in which some definite proportion of its members have been prepared

³⁷ For a somewhat differently focused discussion of these issues, see sec. 1 of [Timpson & Brown \(2005\)](#). The distinction between proper and improper mixtures was first introduced in [d’Espagnat \(1976\)](#).

in a certain way, and other members have been prepared in other, different ways. In the case both of a proper and an improper mixture we will represent, mathematically, the state of the ensemble by a density operator $\hat{\rho}^{\text{mix}}$.

The ensemble of silver atoms we have been considering so far is an example of a *uniform* ensemble; i.e., one that cannot be regarded as being composed of distinct proper sub-ensembles. In the context of our example this means that each spin- $\frac{1}{2}$ particle has physically interacted with our preparation device in exactly the same way. A uniform ensemble is characterized by a *pure* state, represented by a density operator $\hat{\rho}^{\text{pure}}$, which we can construct as the projection onto the unique state vector $|\psi\rangle$ that characterizes each and every one of the ensemble's members; i.e.,

$$\hat{\rho}^{\text{pure}} \equiv |\psi\rangle\langle\psi|. \quad (6.5.6)$$

Our classical analogue of a uniform quantum ensemble is a single-ticket raffle. Single-ticket raffles like the ones we considered in Sections 2.5 and 4.2, and the quantum ensemble we are now considering, all illustrate that even when an ensemble is uniform and the measurements performed on its members are all identical, there can be dispersion in the statistics that these measurements generate. In uniform classical ensembles, such as our single-ticket raffles, the dispersion can be eliminated by including additional variables in our model of a system (which in that case is a ticket). In uniform quantum ensembles, such additional variables are not to be had.³⁸

³⁸ In the case of the single-ticket raffles we considered in Sections 2.5 and 4.2, the dispersion (for each measurement setting they choose, Alice and Bob will find both outcomes listed on the ticket with roughly equal frequency) comes from the fact that a given ticket does not specify which of its halves should go to Alice and which to Bob. We can eliminate the dispersion, and in that sense turn the ticket into a truthmaker for the outcome of a draw (see Section 6.3), simply by labeling one half “Alice” and the other half “Bob” or by stipulating that Alice always gets the left and Bob always gets the right side of the ticket. Other examples of classical uniform ensembles with dispersion are ensembles of identical fair coins or identical fair dice. In those cases, we can eliminate the dispersion in the outcomes (heads or tails, 1 through 6) by replacing the uniform ensemble of coins or dice by a non-uniform ensemble for which each member does not consist just of a coin or a die but of a coin or a die plus the person or contraption flipping or rolling it, where for every member of that new ensemble a careful characterization is provided of how exactly the coin is flipped or the die is thrown. In quantum mechanics, none of these strategies to eliminate dispersion in uniform ensembles will work. We note that on our view Einstein was right to comment that *He* does not play dice (Einstein to Born, December 4, 1926 (Born, 1971, p. 91)), even if the reason for it is likely not what Einstein had in mind: According to quantum mechanics no object can actually play the role of a classical die—i.e., to be such

Quantum mechanics, it will be recalled, is all about probabilities. Using the concept of a statistical ensemble, we can cash out the probability $\Pr(m_z|\hat{z})$ in Eq. (6.5.4) in terms of the relative frequency with which we expect to find the result m_z when we randomly select members from the ensemble and measure their spin in the z -direction, analogously to the way we might randomly select tickets from a basket and examine a particular setting-to-outcome mapping on each.³⁹

To begin to make this all more explicit, we once again rewrite the Born rule, this time in terms of the density operator $\hat{\rho}^{\text{pure}}$ introduced in Eq. (6.5.6). Using the resolution of unity $\hat{1} = \sum_k |k\rangle\langle k|$ in some arbitrary orthonormal basis $\{|k\rangle\}$ (with $k = 1, 2$) of the Hilbert space of S , we can rewrite Eq. (6.5.2), the Born rule for our setup, as

$$\begin{aligned}\Pr(m_z|\hat{z}) &= \sum_k \langle \varphi | \hat{P}_{m_z} | k \rangle \langle k | \varphi \rangle \\ &= \sum_k \langle k | \varphi \rangle \langle \varphi | \hat{P}_{m_z} | k \rangle \\ &= \text{Tr} \left(|\varphi\rangle\langle\varphi| \hat{P}_{m_z} \right),\end{aligned}\tag{6.5.7}$$

where in the last step we used the definition of the trace of a matrix as the sum of its diagonal elements. This enables us to rewrite Eq. (6.5.7) as

$$\Pr(m_z|\hat{z}) = \text{Tr} \left(\hat{\rho}^{\text{pure}} \hat{P}_{m_z} \right),\tag{6.5.8}$$

where $\hat{\rho}^{\text{pure}} \equiv |\varphi\rangle\langle\varphi|$ in accordance with the construction given in Eq. (6.5.6).

We can do something similar when we are dealing with a non-uniform ensemble. Recall that a non-uniform ensemble can be thought of as one for which some of its members have been prepared in a certain way while others have been prepared in different ways, similarly to the way one might conduct a raffle using a basket with a mix of different types of tickets (see Sections 2.5 and 4.2). Consider now an ensemble consisting of a large number of similar quantum systems S_i which have been prepared in one of a number of different states $|\psi_j\rangle$ with relative frequencies given by non-negative real numbers a_j such that $\sum_j a_j = 1$. In that case, the probability of finding m_z when we measure spin in the z -direction on a randomly selected member of the ensemble will be

that the outcome of throwing it could be perfectly predicted given a specification of how it is thrown. We will come back to this issue again below.

³⁹ We do not see this as committing ourselves to a frequentist interpretation of probability. But cashing out probabilities in terms of this von Mises-von Neumann scheme clearly brings out the difference between classical and quantum probabilities.

the weighted average, with weights given by the a_j , of the probabilities, given by the Born rule, of obtaining m_z in the course of a z -direction measurement on a system in the state $|\psi_j\rangle$:

$$\begin{aligned}\Pr(m_z|\hat{z}) &= \sum_j a_j \left| {}_z\langle m_z|\psi_j\rangle \right|^2 \\ &= \sum_j a_j \langle \psi_j|m_z\rangle {}_z\langle m_z|\psi_j\rangle \\ &= \sum_j a_j \langle \psi_j|\hat{P}_{m_z}|\psi_j\rangle.\end{aligned}\tag{6.5.9}$$

Substituting $|\varphi_j\rangle$ for $|\psi_j\rangle$ in Eq. (6.5.9), and proceeding similarly to the way we derived Eq. (6.5.7), we then obtain, for our example:

$$\begin{aligned}\Pr(m_z|\hat{z}) &= \sum_k \sum_j a_j \langle \varphi_j|\hat{P}_{m_z}|k\rangle \langle k|\varphi_j\rangle \\ &= \sum_k \langle k|\left(\sum_j a_j |\varphi_j\rangle \langle \varphi_j|\right)\hat{P}_{m_z}|k\rangle.\end{aligned}\tag{6.5.10}$$

Introducing the density operator

$$\hat{\rho}^{\text{mix}} \equiv \sum_j a_j |\varphi_j\rangle \langle \varphi_j|\tag{6.5.11}$$

to represent the mixed state characterizing our ensemble, we can write Eq. (6.5.10) in the exact same form as Eq. (6.5.8):

$$\Pr(m_z|\hat{z}) = \text{Tr}\left(\hat{\rho}^{\text{mix}} \hat{P}_{m_z}\right).\tag{6.5.12}$$

We note, finally, that Gleason's theorem (1957) tells us that quantum mechanics' assignment of probabilities is *complete* in the sense that every probability measure on the Boolean sub-algebras associated with the observables of a system is representable by means of a density operator in the manner just described.⁴⁰

The account of a quantum-mechanical measurement given above will be criticized. What has been given, it will be objected, is merely a recipe for

⁴⁰ Gleason's proof assumes that measurements are represented as projections and is valid for Hilbert spaces of dimension ≥ 3 . Busch (2003) proves an analogous result for the more general class of positive operator valued measures (POVMs, or "effects") which is valid for Hilbert spaces of dimension ≥ 2 . An extended discussion of the issue of completeness in relation to Gleason's theorem may be found in Demopoulos (2018).

recovering the statistics associated with such a measurement. All we learn from this recipe is that, and how, the quantum formalism may be used to predict the relative frequencies one observes upon having an ensemble of similarly prepared systems interact with a device that has been set up to measure one of those systems' dynamical parameters. No account has been given here of how the measurement interaction itself allows for this. But this is what is demanded by our objector.

Consider again a projective measurement in the z -basis on a spin- $\frac{1}{2}$ particle that is part of a uniform ensemble of systems that have each been prepared in the state described by Eq. (6.5.1). This state description is quantum. Nevertheless it follows from our foregoing discussion that, conditional upon measuring spin in the z -direction, we can effectively describe the observed statistics as having arisen, not from the uniform ensemble that was actually prepared, but from a non-uniform ensemble of spin- $\frac{1}{2}$ particles for which the relative proportions of its systems in the states $|+\rangle_z$ and $|-\rangle_z$ are $|\alpha|^2$ and $|\beta|^2$, respectively. That is, conditional upon measuring spin in the z -direction, the observed statistics will be indistinguishable from those that would be observed from a measurement of spin in the z -direction on a non-uniform ensemble characterized by the mixed density operator

$$\hat{\rho}^{\text{mix}} = |\alpha|^2 |+\rangle_z \langle +| + |\beta|^2 |-\rangle_z \langle -|. \quad (6.5.13)$$

Because of this we can simulate the observed statistics, conditional upon such a measurement, with a local hidden-variables model similar to the raffles we discussed in Sections 2.5 and 4.2. Unlike those raffles, the phenomena we are simulating here are not correlations, thus our tickets will not need to have two halves like the ones depicted, e.g., in Figure 2.11. In the current scenario we can make do with a basket of raffle tickets inscribed with a single symbol, either “ $|+\rangle_z$ ” or “ $|-\rangle_z$ ”, whose relative proportions in the basket are $|\alpha|^2$ and $|\beta|^2$, respectively. Thus, we have here an account of how, through measuring a system in a given basis, our characterization of the system transitions from a quantum to a classical description. Moreover, if one considers the relative frequencies that, with high probability, will be observed upon measuring, in other measurement bases besides z , systems in an ensemble whose members have all been prepared in the state given by Eq. (6.5.1), then one can convince oneself that similar (though incompatible) local hidden-variable characterizations can be given of the observed statistics conditional upon each one of those measurements.

Again this will be criticized. This explanation of the measurement process, it will be objected, is no explanation at all. On the account of a measurement given above, it functions essentially as a black box. But the interests of physics demand that we open all such boxes, and it will be demanded of us that we open this one as well. In the present instance this is a completely legitimate demand, of course, for so far we have not actually said anything about the details of the measurement interaction that we outlined above. Obviously there are many good reasons to want to know about these details. If actual measurement statistics diverge from our expectations, for instance, examining the physical details of the measurement interaction involved may be required to ensure that our device is working. We will need to have a good, detailed, understanding of the physics of a particular type of measurement interaction, to take another example, if we need to reproduce it in another physical location with a different piece of equipment. Maybe we want to understand the details of a particular physical interaction simply for understanding's sake. These are all legitimate reasons to demand a deeper explanation of the measurement interaction than the one that we gave above.

In fact, within quantum theory it is always possible (in principle) to describe, to any desired level of detail, how a particular measurement apparatus dynamically interacts with a given system of interest and thereby gives rise to an entangled state of the system and apparatus that yields probabilities for the state of the apparatus that will be found upon its being assessed. In this way we move back “the cut” (see Bub, 2016, sec. 10.4): the dividing line between, on the one hand, our quantum description of a measurement interaction, and on the other hand, our effectively classical description of whatever instrument we are using to determine a system's state. The part of the phenomenon that, on our earlier analysis, was the instrument of measurement is, on this more detailed analysis, represented in our quantum description of its physical interaction with the system. Moving back “the cut” in this way allows us to see, in more detail, how a particular measurement interaction invites one of the possible effectively classical characterizations of the statistics that will be observed.

To come back to our running example, consider again a measurement of spin in the z -direction on an electrically neutral spin- $\frac{1}{2}$ particle S such as a silver atom. We did not explicitly model the measuring device in our earlier analyses, but let us now do this. Let our measuring device, M_z , be a Du Bois magnet with its axis (the line going from its south to its north pole) pointing

in the positive z -direction.⁴¹ A Du Bois magnet pointing in the z -direction will deflect a silver atom sent through the magnetic field between its two poles either up or down, corresponding to the two possible values of the atom's spin in the z -direction, $m_z = +$ and $m_z = -$.

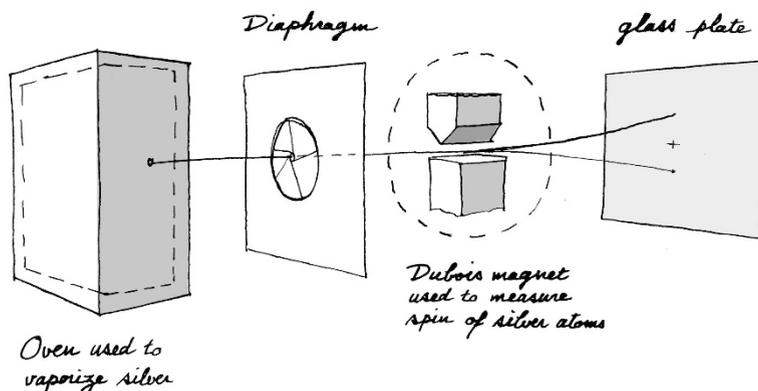


Fig. 6.1 Schematic drawing of the setup in a Stern-Gerlach experiment. A beam of silver atoms sent through a Du Bois magnet all hit a glass plate behind the magnet in one of two spots. Drawing: Laurent Taudin.

⁴¹ A Du Bois magnet is a ring or half-ring magnet used to produce intense inhomogeneous magnetic fields (Glazebrook, 1922, pp. 319–320), named after H. Du Bois, a Dutch applied physicist known for his work in electromagnetism and optics (see C. G. K., 1919) and for his book, *The Magnetic Circuit in Theory and Practice* (1894). Variations on the Stern-Gerlach experiment, in which particles are sent through multiple Du Bois magnets with axes pointing in different directions before being made to impact on a screen (see Eisberg & Resnick, 1985, sec. 8-3), have been used many times before, by philosophers and textbook writers alike, to introduce the counterintuitive nature of quantum mechanics. Two well-known examples from the philosophy of quantum mechanics literature are Hughes (1989, Introduction, pp. 1–8), who cites *The Feynman Lectures* in this context (Feynman, Leighton, & Sands, 1965, Vol. 3, Ch. 5), and Albert (1992, Ch. 1, pp. 1–16), who uses special “boxes” measuring “color” and “hardness” instead of Du Bois magnets measuring spin in different directions. Examples of modern physics textbooks taking this approach are Sakurai & Napolitano (2017, sec. 1.1, pp. 1–6), Townsend (2012, Ch. 1, pp. 1–23), who, like Hughes, cites *The Feynman Lectures* and Sakurai as his inspiration for this approach (ibid., p. xii and p. 6); and McIntyre (2012, Prologue, p. xxi; sec. 1.1, pp. 1–10), who (ibid., p. xv and p. 33, “Further Reading”) credits the approach to Feynman, Sakurai, Townsend, Cohen-Tannoudji & Laloë (1977) and Styer (2000). For examples of this approach in the quantum computing and information literature, see Bernhardt (2019, pp. 1–11) and Nielsen & Chuang (2016, sec. 1.5.1, pp. 43–46).

The Du Bois magnet is only one component in a Stern-Gerlach experiment to determine a silver atom's spin (see Figure 6.1). It plays (arguably) the essential part in the dynamics of such an experiment, which justifies our focusing primarily on it in the following account of a measurement interaction, but of course it does not suffice, for we still need some way to assess the result of the Du Bois magnet's action on a given atom. If, for instance, we send a stream of silver atoms that have all been prepared in the state $|\varphi\rangle$ given by Eq. (6.5.1) through the magnet, then we will need to determine, somehow, what the relative frequencies of up and down deflections of silver atoms by the magnet were. Typically we do this by covering a screen behind the magnet with some emulsion that is sensitive to the impact of silver atoms. If a stream of silver atoms all in the state $|\varphi\rangle$ is then sent through the magnet, the result will be two dots on the screen connected by a line segment pointing in the z -direction. The relative intensities of these two dots, which we may determine by eyeballing them, or if necessary with the aid of some additional instrument, correspond to the relative frequencies of up and down. The process through which a silver atom produces a dot on a screen has as little to do with a measurement of spin, of course, as the process through which those dots are subsequently registered on our retinas. We could, all the same, consider the screen, any further instruments and in principle even our eyes as part of a quantum description of the interaction. But that would add little to our understanding of the relevant dynamical interaction involved in a spin measurement.

In any case, suppose that our measuring device, the Du Bois magnet pointing in the z -direction, M_z , is working properly. That means that if we prepare a stream of spin- $\frac{1}{2}$ particles all in the state $|+\rangle_{S_z}$, send them through the magnet and then let them hit the screen, we will find that they all land at the top spot, which corresponds to the state $|+\rangle_{M_z}$ of the magnet. And if we prepare a stream of spin- $\frac{1}{2}$ particles all in the state $|-\rangle_{S_z}$, we will find that they all land at the bottom spot, which corresponds to the magnet's state, $|-\rangle_{M_z}$. These two outcomes, top-top and bottom-bottom, correspond to two of the possible states for the compound system $S + M_z$ consequent to an interaction between S and M_z : $|+\rangle_{M_z}|+\rangle_{S_z}$ and $|-\rangle_{M_z}|-\rangle_{S_z}$, respectively. The linearity of quantum mechanics (the "superposition principle") tells us that, if we prepare a particle in the state $\alpha|+\rangle_{S_z} + \beta|-\rangle_{S_z}$ (cf. Eq. (6.5.1)), the state of $S + M_z$ after a measurement interaction will be the entangled superposition:

$$|\Phi\rangle_{S+M_z} = \alpha|+\rangle_{M_z}|+\rangle_{S_z} + \beta|-\rangle_{M_z}|-\rangle_{S_z}. \quad (6.5.14)$$

The density operator $\hat{\rho}_{S+M_z}^{\text{pure}}$ for a uniform ensemble consisting of many copies of the compound system $S+M_z$ in this state will be (suppressing the subscripts) the projector $|\Phi\rangle\langle\Phi|$ onto $|\Phi\rangle$. This density operator can be written as:

$$\begin{aligned}\hat{\rho}_{S+M_z}^{\text{pure}} &\equiv \left(\alpha |+\rangle_{M_z} |+\rangle_{S_z} + \beta |-\rangle_{M_z} |-\rangle_{S_z} \right) \\ &\quad \left(\bar{\alpha} \langle +|_{M_z} \langle +|_{S_z} + \bar{\beta} \langle -|_{M_z} \langle -|_{S_z} \right) \\ &= |\alpha|^2 \left(|+\rangle_{M_z} \langle +|_{M_z} \right) \left(|+\rangle_{S_z} \langle +|_{S_z} \right) \\ &\quad + \alpha \bar{\beta} \left(|+\rangle_{M_z} \langle -|_{M_z} \right) \left(|+\rangle_{S_z} \langle -|_{S_z} \right) \\ &\quad + \beta \bar{\alpha} \left(|-\rangle_{M_z} \langle +|_{M_z} \right) \left(|-\rangle_{S_z} \langle +|_{S_z} \right) \\ &\quad + |\beta|^2 \left(|-\rangle_{M_z} \langle -|_{M_z} \right) \left(|-\rangle_{S_z} \langle -|_{S_z} \right),\end{aligned}\quad (6.5.15)$$

where $\bar{\alpha}$ and $\bar{\beta}$ are the complex conjugates of α and β . To obtain the probability of a particular outcome for the combined system, given a measurement in the z -direction, we apply the same trace formula we derived for the system S , Eq. (6.5.8), to the compound system $S+M_z$. Substituting $\hat{\rho}_{S+M_z}^{\text{pure}}$ for $\hat{\rho}^{\text{pure}}$ and $\hat{P}_{m_z}^{S+M_z}$ for \hat{P}_{m_z} in that equation, we arrive at:

$$\Pr(m_z|\hat{z}) = \text{Tr} \left(\hat{\rho}_{S+M_z}^{\text{pure}} \hat{P}_{m_z}^{S+M_z} \right) \quad (6.5.16)$$

where, as the superscript indicates, $\hat{P}_{m_z}^{S+M_z}$ is a projector in the Hilbert space of the compound system $S+M_z$, the tensor product of the Hilbert spaces of S and M_z :

$$\hat{P}_{m_z}^{S+M_z} \equiv \left(|m_z\rangle_{M_z} \langle m_z| \right) \left(|m_z\rangle_{S_z} \langle m_z| \right). \quad (6.5.17)$$

Suppose that we randomly draw a copy of the compound system $S+M_z$ from the statistical ensemble characterized by $\hat{\rho}_{S+M_z}^{\text{pure}}$. Eq. (6.5.16) then gives the expected relative frequency with which a measurement of spin in the z -direction will result in the outcome m_z for both S and M_z (i.e., S has spin m_z in the z -direction, and M_z indicates that S has spin m_z in the z -direction). We should not forget, though, that the magnet's action on a given silver atom must still be assessed, something that, as we discussed earlier, is typically done by

placing a screen with some emulsion sensitive to the impact of a silver atom behind M_z . We will return to this point again soon.

If, as is often the case, we are only interested in the system S , we should calculate the marginals for S of the probability distribution over $S + M_z$. To find those we trace out the degrees of freedom of M_z in $\hat{\rho}_{S+M_z}^{\text{pure}}$ to obtain the *reduced density operator* $\hat{\rho}_S^{\text{red}}$:

$$\hat{\rho}_S^{\text{red}} \equiv \text{Tr}_{M_z}(\hat{\rho}_{S+M_z}^{\text{pure}}) = \sum_l M_z \langle l | \hat{\rho}_{S+M_z}^{\text{pure}} | l \rangle_{M_z}, \quad (6.5.18)$$

where in the last step we introduced an arbitrary orthonormal basis $\{|l\rangle_{M_z}\}$ with $l = 1, 2, \dots$ of the Hilbert space of M_z . Let us now derive a concrete expression for the reduced density operator $\hat{\rho}_S^{\text{red}}$ corresponding to the density operator $\hat{\rho}_{S+M_z}^{\text{pure}}$ given in Eq. (6.5.15). The first of the four terms on the right-hand side of this equation gives:

$$|\alpha|^2 \sum_l M_z \langle l | \left(|+\rangle_{M_z} M_z \langle + | |+\rangle_{S_z} S_z \langle + | \right) | l \rangle_{M_z}, \quad (6.5.19)$$

which can be rewritten as

$$|\alpha|^2 \sum_l M_z \langle l | + \rangle_{M_z} M_z \langle + | l \rangle_{M_z} | + \rangle_{S_z} S_z \langle + |. \quad (6.5.20)$$

Using the completeness of $\{|l\rangle_{M_z}\}$ to set

$$\begin{aligned} \sum_l M_z \langle l | + \rangle_{M_z} M_z \langle + | l \rangle_{M_z} &= \sum_l M_z \langle + | l \rangle_{M_z} M_z \langle l | + \rangle_{M_z} \\ &= M_z \langle + | + \rangle_{M_z} = 1, \end{aligned} \quad (6.5.21)$$

we see that this reduces to:

$$|\alpha|^2 | + \rangle_{S_z} S_z \langle + |. \quad (6.5.22)$$

The three other terms in Eq. (6.5.15) can similarly be worked out to yield:

$$\begin{aligned} \alpha \bar{\beta} M_z \langle - | + \rangle_{M_z} | + \rangle_{S_z} S_z \langle - |, \\ \beta \bar{\alpha} M_z \langle + | - \rangle_{M_z} | - \rangle_{S_z} S_z \langle + |, \\ |\beta|^2 | - \rangle_{S_z} S_z \langle - |. \end{aligned} \quad (6.5.23)$$

For the device M_z to be *effective* in measuring the spin of S in the z -direction, the states $|+\rangle_{M_z}$ and $|-\rangle_{M_z}$ should be clearly distinguishable and the absolute value of their inner product, $M_z\langle-|+\rangle_{M_z}$, should therefore be small. Introducing the (in general complex) coefficients

$$\varepsilon \equiv (\overline{\beta/\alpha})_{M_z}\langle-|+\rangle_{M_z}, \quad \eta \equiv (\overline{\alpha/\beta})_{M_z}\langle+|-\rangle_{M_z}, \quad (6.5.24)$$

we can thus assume that $|\varepsilon| \ll 1$ and $|\eta| \ll 1$. With the help of Eqs. (6.5.22)–(6.5.24), we can rewrite Eq. (6.5.18) as:

$$\begin{aligned} \hat{\rho}_S^{\text{red}} = |\alpha|^2 |+\rangle_{S_z} \left(s_z \langle+| + \varepsilon s_z \langle-| \right) \\ |\beta|^2 |-\rangle_{S_z} \left(s_z \langle-| + \eta s_z \langle+| \right). \end{aligned} \quad (6.5.25)$$

This is the reduced density operator, with respect to S , for a uniform ensemble of copies of the compound system $S + M_z$ in the pure state $|\Phi\rangle_{S+M_z}$ given by Eq. (6.5.14). As we will explicitly show below, $\hat{\rho}_S^{\text{red}}$ will yield a probability distribution over the outcomes $+$ and $-$ resulting from a measurement of the spin of S in the z -direction that (a) does not depend on the values of ε and η in Eq. (6.5.25) at all and (b) is exactly the same as the probability distribution yielded by the density operator $\hat{\rho}^{\text{mix}}$ in Eq. (6.5.13) for a measurement of the spin of S in the z -direction. This density operator, which we will denote more explicitly as

$$\hat{\rho}_{S_z}^{\text{mix}} \equiv |\alpha|^2 |+\rangle_{S_z} s_z \langle+| + |\beta|^2 |-\rangle_{S_z} s_z \langle-|, \quad (6.5.26)$$

describes a *non-uniform* ensemble of copies of the system S for which the proportions of those copies that are in the state $|+\rangle_{S_z}$ and $|-\rangle_{S_z}$, respectively, are $|\alpha|^2$ and $|\beta|^2$. *For the purposes of computing the probability of finding $+$ or $-$ when measuring the spin of S in the z -direction*, we are thus justified in replacing the uniform ensemble described by a density operator of the form given in Eq. (6.5.6), in which all copies of S are in the same entangled superposition given in Eq. (6.5.14), with the non-uniform ensemble described by the density operator given in Eqs. (6.5.13) and (6.5.26).

In fact, these two ensembles will also give indistinguishable (though generally not exactly the same) probability distributions for measurements of the spin of S in other directions besides the z -direction. Suppose we draw a copy of the compound system $S + M_z$ from the uniform ensemble described by the density operator $\hat{\rho}_{S+M_z}^{\text{pure}}$ given in Eq. (6.5.15) and measure the spin of S in some arbitrary direction d . We can use the reduced density operator $\hat{\rho}_S^{\text{red}}$ given in Eq.

(6.5.25) to compute the probabilities of finding $+$ or $-$ in this measurement. We could also draw a copy of the system S from the non-uniform ensemble described by the mixed density operator $\hat{\rho}_{S_z}^{\text{mix}}$ given in Eq. (6.5.26) and measure the spin of S in that same direction d , using $\hat{\rho}_{S_z}^{\text{mix}}$ to compute the probabilities of finding $+$ or $-$ in this measurement. As long as $|\varepsilon| \ll 1$ and $|\eta| \ll 1$, these two probability distributions will be indistinguishable. If $\varepsilon = \eta = 0$ (i.e., if $|+\rangle_{M_z}$ and $|-\rangle_{M_z}$ are orthogonal), they will be exactly the same.

Let us now take a closer look at how the coefficients ε and η do and do not affect the probability distributions we expect to find in various experiments. We begin by using $\hat{\rho}_S^{\text{red}}$ given in Eq. (6.5.25) to compute the marginal probability of finding $+$ in a measurement of the spin of S in the z -direction:

$$\begin{aligned} \Pr(m_z = + | \hat{z}) &= \text{Tr}\left(\hat{\rho}_{S_z}^{\text{red}} |+\rangle_{S_z} \langle +|\right) \\ &= \sum_k {}_S\langle k | \left(\hat{\rho}_{S_z}^{\text{red}} |+\rangle_{S_z} \langle +|\right) |k\rangle_S \\ &= \sum_k {}_{S_z}\langle + | k\rangle_S {}_S\langle k | \hat{\rho}_{S_z}^{\text{red}} |+\rangle_{S_z} \\ &= {}_{S_z}\langle + | \hat{\rho}_{S_z}^{\text{red}} |+\rangle_{S_z}, \end{aligned} \quad (6.5.27)$$

where we used an arbitrary orthonormal basis $\{|k\rangle\}$ (with $k = 1, 2$) of the Hilbert space of S to evaluate the trace and the resolution of unity in that basis. The only term in $\hat{\rho}_{S_z}^{\text{red}}$ in Eq. (6.5.25) that contributes to this probability is $|+\rangle_{S_z} \langle +|$. Similarly, the only term in $\hat{\rho}_{S_z}^{\text{red}}$ contributing to the marginal probability of finding $-$ in this measurement is $|-\rangle_{S_z} \langle -|$. Hence,

$$\Pr(m_z = + | \hat{z}) = {}_{S_z}\langle + | \left(|\alpha|^2 |+\rangle_{S_z} \langle +|\right) |+\rangle_{S_z} = |\alpha|^2, \quad (6.5.28)$$

$$\Pr(m_z = - | \hat{z}) = {}_{S_z}\langle - | \left(|\beta|^2 |-\rangle_{S_z} \langle -|\right) |-\rangle_{S_z} = |\beta|^2, \quad (6.5.29)$$

which confirms that, for this particular experiment, the reduced density operator $\hat{\rho}_{S_z}^{\text{red}}$ in Eq. (6.5.25) and the mixed density operator $\hat{\rho}_{S_z}^{\text{mix}}$ in Eq. (6.5.26) give the exact same probabilities.

Note that the quantities ε and η played no role in the derivation of Eqs. (6.5.28)–(6.5.29). This may seem puzzling. On the one hand, ε and η have been defined explicitly in Eq. (6.5.24) in terms of inner products of the states of

the Du Bois magnet, $|+\rangle_{M_z}$ and $|-\rangle_{M_z}$, that we used to assess the two possible states of S in the z -direction. On the other hand, for the statistics generated by an ensemble constructed specifically to analyze this measurement, it makes no difference whether these two states of the Du Bois magnet are distinguishable or not! The probabilities in Eqs. (6.5.28)–(6.5.29) remain the same if we replace the two vectors $|+\rangle_{M_z}$ and $|-\rangle_{M_z}$, which ideally should be orthogonal to one another, by one and the same vector!

The key to resolving this puzzle is to remind ourselves that, although the Du Bois magnet plays the essential part in the dynamics of a measurement of an atom's spin, it does not complete the measurement. We still need a screen if we want to assess the relative frequency of silver atoms that were deflected up or down in a given direction. We only left the screen out of our quantum description of the measurement because the process by which a dot is created on a screen by a particle is not something we take to be controversial in this context. But of course the completion of the measurement through the particular expedient of the screen only works if the Du Bois magnet is functioning properly, because this implies that the results of its interaction with an atom will be effectively distinguishable on the screen. Suppose we replaced the Du Bois magnet by a plastic dummy of one. According to Eqs. (6.5.28)–(6.5.29) this would not change the probabilities of finding $+$ and $-$ when measuring the spin of S in the z -direction. But if M_z were a dummy we would no longer be able to assess the result of its interaction with the system simply by having that interaction take place in front of a screen: A dummy would not deflect the particles and they would all end up landing at the same spot. To properly complete the measurement in this particular case, we would have to send the particles coming out of the dummy Du Bois magnet through a second *functioning* Du Bois magnet M'_z pointing in the z -direction, and only then let them hit the screen. The probabilities for this modified version of the experiment will still be given by Eqs. (6.5.28)–(6.5.29), but if M_z is a dummy (in which case $M_z\langle-|\rangle_{M_z} = 1$), the ensemble of systems $S + M_z$ gives us no insight into the dynamical process we wish to analyze.

Let us return to the case where M_z is a functioning Du Bois magnet pointing in the z -direction and to the ensemble of copies of the compound system $S + M_z$ characterized by the density operator $\hat{\rho}_{S+M_z}^{\text{pure}}$ given in Eq. (6.5.15), the ensemble we introduced to analyze the measurement of the spin of S in the z -direction. Suppose we draw a copy of $S + M_z$ from this ensemble but then decide to assess the spin of S in the x -direction rather than the z -direction by placing a second Du Bois magnet, M'_x , right behind the first with its axis pointing in the x -direction. After being deflected up or down by the first magnet, M_z , particles

are deflected left or right by the second, M'_x . We can make the second Du Bois magnet much bigger than the first to make sure that all of the particles travel through the gap between the poles of the second magnet regardless of whether they were deflected up or down by the first magnet. After the particles have passed through both Du Bois magnets, we let them hit a screen behind the second, where we can observe the formation of two dots separated by a line segment in the x -direction.

We can still use the reduced density operator $\hat{\rho}_{S_z}^{\text{red}}$ given in Eq. (6.5.25) to calculate the probability of finding $+$ or $-$ in this new experiment,

$$\Pr(m_x = + | \hat{x}) = \text{Tr}\left(\hat{\rho}_{S_z}^{\text{red}} |+\rangle_{S_x} \langle +|\right) = {}_{S_x}\langle + | \hat{\rho}_{S_z}^{\text{red}} |+\rangle_{S_x}. \quad (6.5.30)$$

Inserting Eq. (6.5.25) for $\hat{\rho}_{S_z}^{\text{red}}$ —and temporarily changing the subscripts S_x and S_z to x and z for convenience—we find four terms. Two of these are given by

$${}_x\langle + | \left(|\alpha|^2 |+\rangle_z \langle +| + |\beta|^2 |-\rangle_z \langle -| \right) |+\rangle_x, \quad (6.5.31)$$

which can be rewritten as

$$|\alpha|^2 |{}_x\langle + | + \rangle_z|^2 + |\beta|^2 |{}_x\langle + | - \rangle_z|^2. \quad (6.5.32)$$

These are the two terms that would have appeared in our characterization of the probability distribution over $+$ and $-$ in the x -direction had we used the mixed density operator $\hat{\rho}_{S_z}^{\text{mix}}$ in Eq. (6.5.26) instead of the reduced density operator $\hat{\rho}_{S_z}^{\text{red}}$. The remaining two terms are what are commonly called “interference terms”:

$$|\alpha|^2 \epsilon {}_x\langle + | + \rangle_z \langle - | + \rangle_x + |\beta|^2 \eta {}_x\langle + | - \rangle_z \langle + | + \rangle_x. \quad (6.5.33)$$

The second of these is the complex conjugate of the first.⁴² Their sum is thus twice the real part of one of them:

$$2 \text{Re}\left(\alpha \bar{\beta} {}_{M_z}\langle - | + \rangle_{M_z} {}_x\langle + | + \rangle_z \langle - | + \rangle_x\right). \quad (6.5.34)$$

Combining Eqs. (6.5.32) and (6.5.34) (and changing the subscripts x and z back to S_x and S_z), we can write the marginal probability in Eq. (6.5.30) as:

⁴² First, note that $|\alpha|^2 \epsilon = \alpha \bar{\alpha} (\bar{\beta} / \alpha) {}_{M_z}\langle - | + \rangle_{M_z} = \alpha \bar{\beta} {}_{M_z}\langle - | + \rangle_{M_z}$, and that $|\beta|^2 \eta = \beta \bar{\beta} (\alpha / \beta) {}_{M_z}\langle + | - \rangle_{M_z} = \beta \bar{\alpha} {}_{M_z}\langle + | - \rangle_{M_z}$, which means that $|\beta|^2 \eta = \overline{|\alpha|^2 \epsilon}$. Second, note that ${}_x\langle + | - \rangle_z \langle + | + \rangle_x = {}_x\langle + | + \rangle_z \langle - | + \rangle_x$.

$$\begin{aligned} \Pr(m_x = +|\hat{x}) &= |\alpha|^2 |s_x\langle +|+\rangle_{s_z}|^2 + |\beta|^2 |s_x\langle +|-\rangle_{s_z}|^2 \\ &\quad + 2\operatorname{Re}\left(\alpha\bar{\beta} {}_{M_z}\langle -|+\rangle_{M_z} s_x\langle +|+\rangle_{s_z} s_z\langle -|+\rangle_{s_x}\right). \end{aligned} \quad (6.5.35)$$

A similar expression gives $\Pr(m_x = -|\hat{x})$.

To find the marginal probability of obtaining $+$ for a measurement of the spin of S in *any* direction given that S has first gone through a Du Bois magnet pointing in the z -direction, we simply replace the subscript x in Eq. (6.5.35) with a subscript d referring to the direction chosen. For one thing, this shows that, despite the presence of the complex coefficients ε and η in Eq. (6.5.25) for $\hat{\rho}_{S_z}^{\text{red}}$, the probabilities calculated with this reduced density operator are always real. For $d = z$, Eq. (6.5.35) reduces to Eq. (6.5.28) regardless of the value of ${}_{M_z}\langle -|+\rangle_{M_z}$ since $s_z\langle +|+\rangle_{s_z} = 1$ and $s_z\langle +|-\rangle_{s_z} = 0$. For all directions other than the z -direction, however, the probabilities of finding $+$ or $-$ in that direction will depend on ${}_{M_z}\langle -|+\rangle_{M_z}$, i.e., on the effectiveness of the device M_z in measuring spin in the z -direction. Given the discussion following Eqs. (6.5.28)–(6.5.29), this is just what we would expect. The statistics found when measuring the spin of S in the x -direction (by sending particles through a Du Bois magnet pointing in the x -direction) will depend on whether these particles first went through a functioning Du Bois magnet pointing in the z -direction or through a dummy version of same. In the latter case, ${}_{M_z}\langle -|+\rangle_{M_z} = 1$ and there will be “interference” effects; in the former case, ${}_{M_z}\langle -|+\rangle_{M_z} = 0$ and there will be no “interference” effects. In that case, the probabilities found with the reduced density operator $\hat{\rho}_S^{\text{red}}$ in Eq. (6.5.25) will be the same as those found with the mixed density operator $\hat{\rho}_{S_z}^{\text{mix}}$ in Eq. (6.5.26). If the first Du Bois magnet only works part of the time, we expect to find statistics somewhere in between these two extreme cases.

The experiments with Du Bois magnets considered here can be seen as variations on the famous double-slit experiment, in which the interference pattern registered on a screen behind two slits when both of them are open for the duration of the experiment is different from the pattern that gets registered when, at any given time, only one of the two is. To see this analogy between these variations on the Stern-Gerlach and the double-slit experiment, imagine three photon detectors, one at each slit and one near the screen.⁴³ If the detectors at the slits are turned off, the detector at the screen will detect the pattern

⁴³ Strictly speaking, we only need a detector at one of the slits. If that detector clicks, we know that a photon went through that slit. If it does not click, we know the photon went through the other one.

corresponding to two open slits. If the detectors at the slits are turned on and detect the photons with 100% efficiency, the detector at the screen will register the pattern corresponding to the case where only one slit is open at a time. If the efficiency of the detectors at the slits is slowly increased from 0% to 100%, the former pattern will gradually change into the latter (see Schlosshauer, 2007, pp. 63–65, for a detailed analysis of the two-slit experiment along these lines). The value $M_z \langle -|+\rangle_{M_z} = 1$ in our example corresponds to the case of 0% efficiency in the two-slit example, the value $M_z \langle -|+\rangle_{M_z} = 0$ to 100% efficiency.

In the context of the double-slit experiment, talk of “interference” (like talk of “superposition”) conjures up images of Schrödinger wave functions interacting with one another like waves of light.⁴⁴ But the occurrence of “interference” in experiments involving spin- $\frac{1}{2}$ particles passing through Du Bois magnets, rather than photons passing through slits, should serve as a reminder that this terminology is to be taken metaphorically. The salient point is that the statistics we find upon performing a measurement on a system depend, in ways that cannot be captured by classical theory, on what other systems or devices this system has interacted with before.

The analysis above of a measurement of the spin of S in the z -direction with a measuring device M_z underscores one of the central points we want to make in this section: Including M_z in the quantum description of a measurement interaction by considering, as we did, an ensemble of compound systems $S + M_z$ uniformly in the pure state $|\Phi\rangle_{S+M}$ given in Eq. (6.5.14), gives us a dynamical account of the physics of that interaction but (a) this does not dictate which observable we should actually measure in that setup (even though it stands to reason that we should choose to measure the spin of S in the z -direction

⁴⁴ In the famous paper on quantum collisions for which he won the Nobel Prize, Born (1926, p. 804) talked about “interference of . . . ‘probability waves’.” In *Neue Begründung*, Jordan (1927a, p. 812) used the phrase “interference of probabilities” to describe the relevant phenomena, crediting Pauli rather than Born with this terminology (Duncan & Janssen, 2013, pp. 184–185). Inspired by this notion of “interference of probabilities,” Jordan argued that, fundamentally, quantum mechanics amounts to a new probability theory characterized by new versions of the addition and multiplication rules that apply not to the probabilities themselves but to probability amplitudes. In *Mathematische Begründung*, the first installment of his 1927 trilogy, Von Neumann (1927a, p. 46) followed Jordan in this. However, in the second installment, *Wahrscheinlichkeitstheoretischer Aufbau*, Von Neumann (1927b, p. 246) rejected this idea, probably under the influence of criticism of Jordan on this score in the uncertainty paper of Heisenberg (1927, pp. 183–184, p. 196; cf. Duncan and Janssen, 2013, p. 187, note 39, and Janssen, 2009, p. 159). When modern commentators argue that quantum mechanics is essentially a new theory of probability (see, e.g., Aaronson, 2013, p. xxvii, p. 110), they recognize that this new theory is *not* obtained the way Jordan envisioned by changing the addition and multiplication rules.

as we introduced this ensemble to analyze that particular measurement) and (b) it still does not give us a definite outcome of our measurement, only a (classical) probability distribution over its possible outcomes. In other words, this analysis solves neither the “small” nor the “big” measurement problem mentioned at the beginning of this section: it does not tell us which observable will acquire a definite value, nor what that definite value will be.

Corresponding to any physical interaction through which we assess the state of a given system of interest is a Boolean frame over which a classical probability distribution effectively characterizing the possible observational results conditional upon that interaction can be defined. Conversely, associated with every physically meaningful Boolean frame is a physical interaction (cf. Bohr, 1958, pp. 392–393).⁴⁵ A given Boolean frame, however, will not be compatible with every other Boolean frame one might choose to impose upon a system of interest (cf. Frauchiger & Renner (2018), and see the responses by Brukner (2018) and Bub (2020)). For example, had we decided to measure spin in the x -direction, rather than in the z -direction, on the same system S in the same state $|\varphi\rangle = \alpha|+\rangle_z + \beta|-\rangle_z$ given by Eq. (6.5.1), we would have rewritten $|\varphi\rangle$ as $\alpha'|+\rangle_x + \beta'|- \rangle_x$,⁴⁶ changed the orientation of the Du Bois magnet so that its axis pointed in the x -direction, and repeated the steps that led us to Eq. (6.5.25), replacing M_z by M_x everywhere. We would then have replaced the resulting reduced density operator $\hat{\rho}_{S_x}^{\text{red}}$ with the mixed-state density operator,

$$\hat{\rho}_{S_x}^{\text{mix}} = |\alpha'|^2 |+\rangle_{S_x} \langle +| + |\beta'|^2 |-\rangle_{S_x} \langle -|, \quad (6.5.36)$$

which is obviously different from the one in Eq. (6.5.26). These are not only different; they are incompatible: Eq. (6.5.26) and Eq. (6.5.36) will give different

⁴⁵ Bohr writes: “In the treatment of atomic problems, actual calculations are most conveniently carried out with the help of a Schrödinger state function, from which the statistical laws governing observations obtainable under specified conditions can be deduced by definite mathematical operations. It must be recognized, however, that we are here dealing with a purely symbolic procedure, the *unambiguous physical interpretation of which in the last resort requires a reference to a complete experimental arrangement*. Disregard of this point has sometimes led to confusion, and in particular the use of phrases like ‘disturbance of phenomena by observation’ or ‘creation of physical attributes of objects by measurements’ is hardly compatible with common language and practical definition” (our emphasis).

⁴⁶ The coefficients in the x -basis are related to those in the z -basis via:

$$\begin{pmatrix} \alpha' \\ \beta' \end{pmatrix} = \begin{pmatrix} {}_x\langle +|+\rangle_z & {}_x\langle +|-\rangle_z \\ {}_x\langle -|+\rangle_z & {}_x\langle -|-\rangle_z \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}.$$

predictions, for instance, for the relative frequencies of finding $+$ or $-$ when we measure the spin of S in the z -direction (cf. d’Espagnat, 2001, sec. 2).

The kinematics of quantum mechanics presents us with a fundamentally *non-Boolean* probabilistic global structure. Upon this structure, in a given measurement context, we impose a particular Boolean frame. We do this to express our experience of the results of measurements conducted in that context—an experience of events that either do or do not occur, and that together fit into a consistent picture, in that measurement context, of the phenomenon being described (cf. Bohr, 1935, p. 701).⁴⁷ In this way we partition the quantum-theoretical description of a phenomenon into its quantum (non-Boolean) part, and what can effectively be regarded as its classical (Boolean) part. This classical description, which will in general be incompatibly different across different measurement contexts, is what we leave out of the quantum description of the phenomenon. But it is left out by stipulation, for ultimately it is something that we impose upon our description of nature in the context of a given measurement.⁴⁸

The analysis that we gave in Eqs. (6.5.14)–(6.5.26) for the measurement of a spin- $\frac{1}{2}$ particle could of course be given in even more detail. More of the components of the measuring devices being used and of the dynamical interactions occurring between them and between them and a spin- $\frac{1}{2}$ particle can be included in our description of the experiment—in fact we can include as many of these components as we like. A likely candidate is the screen behind the Du Bois magnet(s) but in principle even your nose could be included in a dynamical description of the measurement interaction if it were somehow relevant to assessing its result. Indeed, quantum mechanics can be used to analyze the interaction between *any* two systems, one of which is to be called the “system of interest,” the other the “measuring device,” *irrespective* of the level of internal complexity of either of them. This analysis will proceed in essentially the same way that it proceeded above. And in all cases the quantum

⁴⁷ Bohr writes: “While, however, in classical physics the distinction between object and measuring agencies does not entail any difference in the character of the description of the phenomena concerned, its fundamental importance in quantum theory, as we have seen, has its root in the *indispensable use of classical concepts* in the interpretation of all proper measurements, *even though the classical theories do not suffice* in accounting for the new types of regularities with which we are concerned in atomic physics” (our emphasis). Compare also Bohr (1937, p. 293): “the requirement of communicability of the circumstances and results of experiments implies that we can speak of *well defined experiences* only within the framework of ordinary concepts” (our emphasis).

⁴⁸ Our way, here, of characterizing the distinction between the classical and quantum subdivisions of a phenomenon is heavily indebted to informal correspondence with Jeff Bub.

description of a measurement interaction will tell us how, conditional upon it, the observed statistics can effectively be treated as classical.

Two objections might be raised to the foregoing. The first is that what we have described above is merely the recipe one should follow if one is to give an account of a particular measurement in quantum mechanics. One might demand, however, of a proper theory of measurement that it provide more than just an account of this or that particular measurement (however detailed it may be), and insist that it should tell us how measurement happens *in general*. Second, it could be objected that, while we have seen above how to include the measurement apparatus in one's description of a measurement interaction, on the account of a measurement given above it is still the case that the effectiveness of a particular probability distribution in describing the results of that interaction is always conditional upon the particular (classical) assessment that we make of it, a choice that is arbitrary from the point of view of the theory. But this, it will be objected, is unacceptable; it cannot constitute a total description of reality.

The first demand—that an account of measurement must take the form of a general dynamical account (cf. Ghirardi, 2018)—is a demand that we reject. There is no dynamical process of measurement in general. There are only particular measurements. And in every particular case quantum mechanics provides, as we have illustrated, the general scheme through which a dynamical account of that measurement process can be given. Quantum mechanics provides, that is, the tools we need in order to give an account of how the particular measurement apparatus in question dynamically interacts with a particular system of interest so as to give rise to a combined system in an entangled superposition yielding probabilities for the state of the measuring device that will be found upon assessing it.

As for the second objection: This, we maintain, is based on a misunderstanding of the nature of the distinction upon which the quantum-mechanical assignment of conditional probabilities is based. For asserting the necessity of this distinction does not amount to the claim that measurement must necessarily involve an interaction with some 'classical physical system', where by this we imagine something large or heavy or both.⁴⁹ Indeed, an *atomic* system

⁴⁹ Contrast this with Chris Fuchs's characterization (which we strongly disagree with) of Bohr's view: "Bohr says 'the observational problem is free of any special intricacy,' but the question is, where does the measuring device stop and the observing agent begin? For Bohr, the measuring devices can be (and should be!) excluded from a quantum description when they are heavy enough and capable of obtaining an irreversible mark upon them" (Fuchs, 2017, p. 30). We take Bohr to be telling a different story about the dividing line

can in many cases serve very well as a measuring apparatus (see, for instance, [Bacciagaluppi, 2017](#)). The claim being made here, rather, is an epistemological one. Specifically, the claim is that in order to represent the assessment of a system's state, one needs to be able to distinguish that assessment from the system being assessed. This is true regardless of the measurement interaction in question, and indeed it is true even if the measurement scenario imagined is one in which the state of the entire universe is being assessed, say, by a supreme being. It is still the case that this supreme being must be able to distinguish, in its description of the situation, its measurement of the system from the system it is measuring. And there is no reason to stop there; for there is nothing to stop one from considering the supreme being and the universe as together comprising a single physical system (supposing that the supreme being exists somehow in space and time); and in that case, as in every case, if one is to assess the state of that combined system, one will still need to be able to distinguish that assessment from the system one is assessing. There is no "view from nowhere" within quantum mechanics with respect to its account of a measurement.

Einstein famously asked Abraham Pais, during one of their walks together, whether Pais believed that the moon is there even when nobody looks ([Pais, 1979](#), p. 907). For many, Einstein's question to Pais has amounted to a *reductio ad absurdum* of "orthodox" interpretations such as ours. Insofar as we do not exactly answer "yes" to Einstein's question, many see this as the final and conclusive word against us all. Be that as it may. In the end the answer to Einstein's question is that if, by the moon, one means the classical object whose logic of observables is globally Boolean—the object that we normally think of when we think of the moon—then the answer is: *No* the moon is *not* there when nobody looks, at least not literally, for the world that we inhabit

between classical and quantum description. Consider: "This necessity of discriminating in each experimental arrangement between those parts of the physical system considered which *are to be treated* as measuring instruments and those which constitute the objects under investigation may indeed be said to form a principal distinction between classical and quantum-mechanical description of physical phenomena. It is true that the place *within each measuring procedure* where this discrimination is made is in both cases *largely a matter of convenience*" ([Bohr, 1935](#), p. 701, our emphasis). Consider also: "After all, the concept of observation *is in so far arbitrary* as it depends upon which objects are included in the system to be observed. Ultimately every observation can of course be reduced to our sense perceptions. The circumstance, however, that in interpreting observations use has always to be made of theoretical notions, *entails that for every particular case* it is a *question of convenience* at what point the concept of observation involving the quantum postulate with its inherent 'irrationality' is brought in" ([Bohr, 1928](#), p. 580, our emphasis).

is *not* a globally Boolean one. As the incompatibility (for an ensemble of systems all in the state given by Eq. (6.5.1)) between the classical probability distributions described by Eqs. (6.5.26) and (6.5.36) emphatically affirms: The classical objects one can construct, on the basis of measuring a given set of observables—on the basis of the catalogs of values yielded by those measurements—will in general be different from and incompatible with the classical objects that might have been constructed had we only measured other, different, observables. This does not mean that nothing would exist if we were not around to measure it. It only means that the observable structure of what *does* exist is not describable by a globally Boolean algebra.

This is not anti-realism but a kind of critical realism that affirms that the way that we conceive of our interactions with the world has a bearing on the possible conceptions of the world that can be constructed on their basis. There can be no mistake: There is a real world and further there are properties of objects like mass and spin and charge that are not dynamical in nature, that we do not directly measure and that do not depend upon any particular measurement one might make—our assignments of values to these quantities are taken to be valid in any measurement context.⁵⁰ Properties ascribed to a system by its dynamical state description are not like this. And yet there is nevertheless something in relation to the dynamical state that we may call perspective-independent. This is the Hilbert space that it inhabits: The probabilistic structure of the relations between the dynamical observables that we take to represent a given measurement context; the constraints this structure places on the catalogs of values that we can simultaneously assign to observables; the constraints that, by extension, are finally placed on our possible representations of the systems we (effectively) can consider those observables to be properties of.

Out of all of the individual somewheres from which we prod and poke at the world, we build up a picture of the globally non-Boolean constraints that quantum mechanics imposes on our locally Boolean representations of the results of those interactions, as well as a picture of how these representations may evolve through time in accordance with those constraints. The language that we use to describe the observational possibilities related by this global structure is one that ineliminably includes ‘ourselves’, i.e., it involves the concept of a Boolean frame and its associated measurement context, conditional upon which, and as long as our measuring device is working properly, we can effectively describe our experience in a classically probabilistic way. But

⁵⁰ This is an unpacking (though only a slight one) of a point also made by Demopoulos in his unpublished book.

the structure we progressively uncover in this way exists independently of our particular interactions with the world even if, like the kinematical core of spacetime, we do not feel in any way tempted to conceive of it as some sort of substance (see Chapter 1).⁵¹ For what the kinematical core of quantum mechanics tells us is that the world—the world that we inhabit—is a *quantum*, not a classical world: The structure manifested by this world does not invite us to infer from it to the existence of a posited underlying physical system as the bearer of a globally Boolean collection of properties.

In Chapter 1, we argued for the usefulness of organizing the various major interpretations of quantum mechanics into a phylogenetic tree. In part to motivate this we argued that the usual way in which interpretations are classified, as ‘epistemic’ or ‘ontic’, is of rather limited use. But we see now that even labels such as ‘representational’ and ‘non-representational’ (Wallace, 2019) are apt to be misleading. The issue is not whether or not quantum mechanics represents. The issue that divides, on the one hand, us and those nearby to us on the phylogenetic tree from, on the other hand, those with whom all of us disagree, is the issue of *what* quantum-mechanical structures represent. The quantum state description, on our view, is *not* a window into a globally Boolean structure, and thus not a window into a posited underlying physical system that is the bearer of the properties related by that structure. It is a window into the underlying *non-Boolean* structure of the world, into the objective probabilistic space of possibilities associated with the *observables* related by this structure—the propensities that, when actualized, form the basis for the theoretical conception of the world that we build up from them. This theoretical conception, despite being non-Boolean, is such that an effectively Boolean description of phenomenal reality can be constructed on its basis. And through phenomenal experience, effectively described in this way, we eventually discover that when, in particular, the relations between our phenomenal experiences are scrutinized, they are found to be in tension with a globally Boolean underlying structure, for the structure of the world of our experience is not globally Boolean but given by Hilbert space.

⁵¹ Cf. Bohr, who writes, “The description of atomic phenomena has in these respects a perfectly objective character, in the sense that no explicit reference is made to *any individual observer* and that therefore, with proper regard to relativistic exigencies, no ambiguity is involved in the communication of information” (Bohr, 1958, p. 390, emphasis added). D’Espagnat calls this a *weakly objective* view, and contrasts it with a *strongly objective* view, according to which certain entities appearing in a theory’s formulae are “really existing entities” (d’Espagnat, 2001, p. 8). See also note 28 above.

The conception of the world invited by Hilbert space contrasts starkly with the conception invited by classical phase space (see Section 6.3). But we see now that the classical worldview is not inevitable. Nor had it always been the all-dominant conception it was to eventually become during physics' classical period. In "Science's disappearing observer," the first chapter of their book *Baroque Science*, a revisionist take on the Scientific Revolution, Ofer Gal and Raz Chen-Morris (2013) describe how the central focus in prominent optical theories from antiquity through to the early modern period had been sight rather than light.⁵² While light had had a place in traditional optics, and indeed had been held to be a necessary condition for vision, it was never identified with vision. Vision, in traditional optics, had been characterized as "a self-authenticating process of communication between object and reason through the eye" (p. 27), that occurs as the result of something being propagated between a visual object and a subject. But what this something is that is propagated is not light. It is "an entity that is both indubitably authentic to the object and immediately transparent to the intellect; a 'form'" (ibid., p. 23). The object of sight was held in this way to be in direct communication with the seer (ibid., p. 18, 21), and the image of an object was held to be an exact representation of it. This is the fact upon which traditional optical theories are constructed: "a theory of visual perception, and any such theory that failed to account for the adequacy of the seen image [to its object] is ipso facto false" (ibid., p. 22). Sight on the traditional conception is in fact no different than the other senses: "For Aristotle . . . the wax metaphor stresses the direct contact of the object, through the medium, with the sense organ, and reinforces the teleology, immediacy, and veridicality of sense perception" (ibid., p. 44).

Gal and Chen-Morris describe how, in contrast, in Johannes Kepler's optical theory images were considered to be generated by light, the result of a natural causal process that we can learn about and develop a theory of through experiment, a theory which then furnishes a criterion for determining the extent to which a given image can be trusted (ibid., p. 24). Observation through instruments such as the *camera obscura* had illustrated, for Kepler, the general lesson that images do not belong to any particular observer: "The image of the moon [on the paper, generated via the *camera obscura*,] is not the culmination of a cognitive process. It does not require an observer; a piece of paper is

⁵² One can consider David C. Lindberg's account of the development of optics during this period as the canonical foil for Gal and Chen-Morris's account. See Lindberg (1992, pp. 307–315) for a brief discussion of the development of optics through Aristotle, Euclid, Ptolemy, Alhazen, Kepler and others, and also Lindberg (1976) for his monograph on the later stage of this development.

enough. In fact, even the paper is not necessary: it can be moved around without affecting the production of the image” (ibid., p. 16). Nor does it belong to the moon: “[The] essential relation between source and image completely disappears from Kepler’s account, together with the exactness of representation it ensures” (p. 18). This mystifies the relation between object and observer at the heart of the concept of vision. And this is what Gal and Chen-Morris call “the optical paradox”: “the naturalization of the eye estranges the observer, and a deeper understanding of optics turns vision into a mystery” (p. 26).⁵³ Optics is no longer fundamentally about *vision* at all: “[For Kepler] [o]ur very visual conception of the object—three-dimensional, smooth-contoured, upright—is a construct; a product of the ‘conceits of the mind’” (p. 28). The new optics is a theory of the production of images through light (p. 20).

The new optics soon proved its usefulness. Besides providing a natural explanation of how instruments such as the *camera obscura* work and how they can be progressively improved upon, what was eventually to become known as classical optics explained, in a straightforward way, phenomena such as the colors of the rainbow (ibid., 34–42) that had always eluded traditional optics. As for the epistemological consequences of the new optics, these received their full elaboration in the work of Descartes. As Gal and Chen-Morris explain:

[In Descartes,] [t]he process by which images are created belongs to light, not to the eyes or to the objects. It owes no inherent allegiance to either; both are just accidental points that light happens to bounce off of. This is not a random or capricious process; the same image on the retina may be assumed to be the outcome of the same process, and therefore represent the same object, just as we can expect a word to always signify the same object. But this precarious uniformity is the only anchor for our trust in our perceptions, and in itself it is nothing more than the regularity of cause and effect: from the epistemological foundations of all science, vision had become dependent on science as the guarantor of its limited reliability (p. 46).

In his *Dioptrique*, Descartes absorbed even the eye itself into his mechanistic account of nature; “[the eye] is no longer the end of the visual process, merely an arbitrary point of reference, an unprivileged station in the natural process” (Gal & Chen-Morris, 2013, p. 49; see also Figure 6.2).

This conceptual “detachment” of the physical process of image formation from the concept of seeing, and the abandonment of the latter in the new theories of optics, illustrates the foundational methodological strategy that lies at the heart of what was eventually to become known as classical physics, the classical physics that presupposes, as we discussed above, that nature is to

⁵³ It would be interesting to consider this in light of the more general epistemological thesis defended in Nagel (1989).

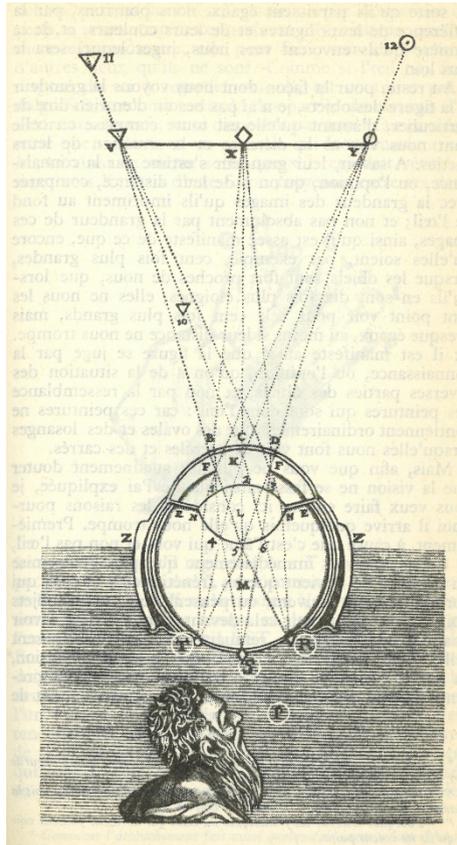


Fig. 6.2 Descartes' depiction of an observer observing an ox's eye. Reproduced with permission of the Rare Books and Special Collections Library, The University of Sydney. This figure is also included and discussed in [Gal & Chen-Morris \(2013, p. 50\)](#) and in [Darrigol \(2012, p. 43\)](#).

be framed within a globally Boolean structure, and through this conceives of nature as something that exists in a particular way independently of how we interact with it.⁵⁴ What quantum mechanics shows us is that as enormously use-

⁵⁴ [Fix \(2019\)](#) gives a similar account of the history of theories of hearing and sound. [Müller \(2020\)](#) describes the classical worldview as the “current conception of physics” before laying the groundwork for an alternative observer-centric conception of physical theory cast in the framework of algorithmic information theory.

ful as this strategy has been, and continues to be in non-fundamental physics, it eventually brings us to a point beyond which the usefulness of pursuing it diminishes.⁵⁵ To make this worldview work for elementary phenomena we need to twist and contort nature, and hide, as if behind a curtain, the parts of it that are responsible for troubling the worldview. In hindsight it was inevitable, we maintain, that the classical worldview should eventually have found a limit to its usefulness. For ultimately, physics must grapple with the fact that we are in the world, after all. This fact, for us, is not the end but the *beginning* of our interpretation.

With regard to the universe *as a whole*, the important question to ask, in light of this, is not whether we can describe, in the absence of any perspective, the universe's dynamical state. The important question to ask is whether and how our individual, local, measurements can actually function in the way that the theory says that they must—whether there actually exist measurement instruments that can play the role required of them, and how we can use them to progressively build up our shared (and in general, non-local) conception of the world. Consider, by way of analogy, the claim one might make in the context of classical physics that one can measure the length of a given body with a rod, or the lifetime of a given particle with a clock. To conduct an accurate measurement, the rod must be rigid, the clock ideal. And a legitimate demand one might make in this instance is that the existence of such rigid rods and ideal clocks be substantiated. Einstein accepted this, and in a debate with Weyl in 1918 related to this issue appealed to the identical spectral lines manifested by atoms of the same kind as compelling evidence for the existence of such ideal instruments (see Giovanelli, 2014). Inquiring into the existence and nature of ideal measurements is something one does in the context of quantum mechanics as well (see, for instance, Cabello, 2019, Pokorný, Zhang, Higgins, Cabello, Kleinmann, & Hennrich, 2019). Moreover quantum mechanics provides criteria for determining the degree to which a given measurement has been successful (see Eq. 6.5.35 above and the surrounding discussion).

In the context of special relativity, the further objection was not made, though it conceivably could have been, that in connecting the theory up with our experience in this way, it is still presupposed, in every particular case, that somehow a rod or a clock has been determined to be a suitable one, and to complete the theory we require an account of exactly how that determination

⁵⁵ An early discussion of what it means to say that the classical worldview is, on the one hand, useful, and on the other hand, useful only up to a certain limit, can be found in sec. 16 of Grete Hermann's 1935 essay on quantum mechanics. For a recent commentary on Hermann's paper, including on this section, see Cuffaro (2021a).

has arisen. In the context of special relativity this objection is easily dismissed as an extra-physical concern. And yet, the analogous question in the case of quantum mechanics is not so readily dismissed. The reason for this, it seems, is the intrinsic randomness of the theory. The dynamical account of a given measurement that is provided by quantum mechanics ultimately ends in probabilities; it does not end in definite outcomes. And yet when one assesses the state of a given system the result is in every case a definite outcome. What, one will ask, is to be made of the definite character of these particular outcomes as contrasted with the apparent indefiniteness we attach to the description of a quantum state, and how is it that the former can be seen as arising from the latter? For someone motivated by this worry, appealing to the quantum-mechanical account of the dynamics of a particular measurement, as we did above, is a *non sequitur*; the quantum-mechanical account of a measurement, no matter how deep or encompassing one makes it, in the end can only yield indefiniteness; it can in general only assign a probability to a particular measurement outcome. But it is an account of the mechanism through which a particular definite outcome emerges from this indefiniteness which is now being demanded.

Bub & Pitowsky (2010) refer—with irony—to this as the “big” measurement problem, for in truth the problem is a superficial one. If we compare on the one hand, measurements in the z -basis on a uniform ensemble of entangled systems in the state given by Eq. (6.5.15), whose outcomes are distributed according to Eq. (6.5.16), and on the other hand drawings from a basket of raffle tickets in which the proportions of “ $|+\rangle_{M_z}|+\rangle_{S_z}$ ” and “ $|-\rangle_{M_z}|-\rangle_{S_z}$ ” outcomes are respectively $|\alpha|^2$ and $|\beta|^2$, the important conceptual difference between the two models is *not* that the outcome obtained in a particular run in one but not the other model is determined stochastically. This is in fact true of both the quantum ensemble and the raffle. To be sure, in the case of a raffle, we can always eliminate the indeterminism in the outcome for a particular run by further distinguishing the tickets from one another in some way.⁵⁶ In quantum mechanics, in contrast, the theory does not make available further parameters that could give us the answer to which of the two outcomes will ensue when we measure the spin in the z -direction of a system in the state given by Eq. (6.5.1). A quantum-mechanical state assignment fixes in advance only the probability that a selected observable will take on a particular value when we query the

⁵⁶ A concrete example of how to do this with the raffles we used to simulate correlation experiments in Sections 2.5 and 4.2 is given in this chapter in note 38 above. To apply that discussion to the present example it may be helpful to change Alice’s and Bob’s names to Maggie and Sid.

system concerning it (i.e., when the operator representing the observable is applied to the state vector describing the system). Nothing, of course, bars us from modifying quantum theory by adding further elements to its formalism. By enhancing our interrogation techniques in this way we can certainly force it to yield up answers to our questions. But rather than do this, we elect to take as true what the kinematical core of quantum theory is telling us of its own accord: that the world is fundamentally nondeterministic, that there is no further story to tell about how a particular definite outcome emerges as the result of a given measurement; that measurement outcomes are intrinsically random—in general only determinable probabilistically.

The more important problem of measurement is not this but rather what [Bub & Pitowsky \(2010\)](#) refer to as the “small” measurement problem: the problem in accounting for the effective emergence of what appears as a globally Boolean ‘macro-structure’ of events out of a globally non-Boolean ‘micro-structure’ underlying it.⁵⁷ The problem is that all of the classical probability distributions (i.e., all of the Boolean frames) that can be used to describe the statistics emerging from our interactions with a quantum system are such that their effectiveness in describing those statistics is conditional upon the choices we make from among the many possible measurements performable on the system. An ensemble of quantum systems all prepared in the same state $|\varphi\rangle$ characterized by the density operator $\hat{\rho}_S^{\text{pure}} = |\varphi\rangle\langle\varphi|$, for example, yields statistics that can be simulated by a particular classical probability distribution over the outcomes $m_z = \pm$ when the systems from the ensemble interact with a Stern-Gerlach apparatus whose Du Bois magnet is oriented along the z -direction. If we allow the ensemble to interact with an apparatus whose magnet is oriented along the x -direction, however, we require a different probability distribution to effectively describe the measurement statistics that ensue, which is in general incompatible with the first. Reality—the reality that we construct on the basis of our measurements, seems then to be contingent upon our choices. This is a problem if our goal is to construct a conception of reality wherein we play no role.

Note that the problem is not solved by including aspects of the measuring apparatus (or indeed all of it) in our quantum-mechanical description of the experimental setup as we did above. For given the entangled superposition in Eq. (6.5.14), we are still left with the choice of whether to measure one set of observables on this combined system or to measure some other set.

⁵⁷ We write ‘macro-structure’ and ‘micro-structure’ in scare-quotes to emphasize that the important distinction being pointed to here is not one of scale but of logic.

Quantum mechanics does not make this choice for us, nor does it explain it away. It is up to us. The picture of the macroworld that we construct out of the consequences of our choices in this sense depends upon them. The picture that we construct for a physical system, in other words, will be different depending on whether we decide to assess this or that aspect of it, and it will in general be different in incompatible ways. This is the “small” problem of measurement.⁵⁸ And this is where the essential difference between the quantum-mechanical model and our raffle model is located. In the raffle model, all of the tickets are marked with the outcomes corresponding to the different possible choices of measurement setting in advance. We may thus ascribe definite values to all of them regardless of which setting we happen to look at in a given run. This is not the case in the quantum model.

And yet it is wrong to think of this as a problem pertaining to the quantum-mechanical account of a measurement. For given a particular measurement context, quantum mechanics provides us with all of the resources we need to account for the dynamics of the measurement interaction between the system of interest and measurement device, and explain in this way why a particular classical probability distribution can effectively be used to describe the statistics characterizing the observations emerging out of that measurement context. Actually carrying out such an analysis, in full detail, in the context of a real physical interaction is no small feat. This is a hard dynamical problem. But whatever its solution in a particular measurement context happens to be, quantum mechanics does not impose that particular measurement context upon us. From the point of view of the theory the measurement choices we make or do not make are up to us.

⁵⁸ For the Everettian this is known as the preferred basis problem.



Chapter 7

Conclusion

Our interpretation of quantum mechanics in a nutshell • Comparison with the Everett family of interpretations

We noted in Chapter 1 that for Heisenberg, quantum mechanics provided us with a new general framework for doing physics, one that was sorely needed in light of the persistent failures of classical mechanics and the old quantum theory of Bohr and Sommerfeld to deal with the puzzling (mostly spectroscopic) experimental data it was confronted with in the first two decades of the last century (Duncan & Janssen, 2019). Heisenberg's core insight into quantum mechanics' significance is one that we and those close to us on the phylogenetic tree of interpretations share. In the body of this volume we saw a number of concrete examples vividly illustrating the essential differences between the quantum and the classical kinematical framework, how those differences are manifested in the correlations between and in the dynamics of quantum systems, and finally how the quantum-kinematical framework enables us to learn about the specifics of particular systems through measurement. In this final chapter we present our view in a nutshell.

Quantum mechanics is about probabilities. The kinematical framework of the theory is probabilistic in the sense that the state specification of a given system yields, in general, only the probability that a selected observable will take on a particular value when we query the system concerning it. Quantum mechanics' kinematical framework is also non-Boolean: The Boolean algebras corresponding to the individual observables associated with a given system cannot be embedded into a globally Boolean algebra comprising them all, and thus the values of these observables cannot (at least not straightforwardly) be taken to represent the properties possessed by that system in advance of their determination through measurement (see Section 6.3). It is in this latter—non-Boolean—aspect of the probabilistic quantum-kinematical framework that its departure from classicality can most essentially be located.

Despite this character, we have seen above (see Section 6.5) how the quantum-mechanical framework provides a recipe through which one can effectively acquire information concerning particular systems by classical means. Given an ensemble of quantum systems either prepared uniformly

in a particular state $|\psi\rangle$ or as a mixture of states $|\psi_i\rangle$ (described by the density operators $\hat{\rho} = |\psi\rangle\langle\psi|$ and $\hat{\rho} = \sum_i \alpha_i |\psi_i\rangle\langle\psi_i|$, respectively), and conditional upon a particular classically describable assessment of one of the parameters of the systems in that ensemble—conditional, that is, upon a particular Boolean frame that we impose on those systems—the information we obtain from our assessment can always be effectively (re)described as having arisen from an ensemble of classical systems (like the raffles in our examples) with a certain distribution of values for the parameter in question.

This recipe does not solve the “small” measurement problem, i.e., the problem to account for how it is that only some of the classical probability distributions implicit in the quantum state description are actualized in the context of a given measurement. And yet we have seen above both that the physical significance of this puzzle of measurement, as well as the physical account quantum mechanics provides of particular measurements, flow naturally from the constraints quantum mechanics’ kinematical core imposes on our representations of quantum systems. Quantum mechanics provides us, moreover, with all of the tools we need to give as precise an account as we would like to give of the dynamics of a particular measurement interaction, and through this explain why a particular classical probability distribution (i.e., a particular Boolean frame) can be used to characterize the statistics observed within that measurement context, despite the non-classical nature of the quantum state description.

Our goal in this book has been to present our own particular take on the informational interpretation of the general framework of quantum mechanics. It has not been part of that goal to show that ours is the only possible interpretation, or to highlight the deficiencies in other families of interpretations of quantum mechanics. Detailed comparisons of our own view with these others would have been inappropriate given this goal, even though such comparisons could have been made at several junctures as we progressed through our discussion in Chapters 1–6. As we bring this book to a close, however, it will be useful to consider our view in the light of the Everett family of interpretations in particular, which like ours add nothing to the quantum-mechanical formalism.

In the pages of this book we have argued that the fundamental novelty of quantum mechanics is to be located in its kinematical core: in the structural constraints that the theory imposes on all of the dynamical systems it describes. And we have seen that this kinematical framework is, unlike the kinematical framework of classical mechanics, fundamentally non-Boolean. The conception of the physical world invited by classical mechanics—the so-

called classical conception of physics—the metaphysical worldview which, by the turn of the last century, had come to dominate physical science, is one in which observer-independent *properties* are primary. The values of dynamical quantities revealed in experiments are understood to originate in those properties. And for anyone who cannot conceive of reality in anything but this way,¹ and yet agrees with us that quantum mechanics describes reality completely, the Everett interpretation of quantum mechanics is the natural choice. Since the values of quantum-mechanical observable quantities cannot be consistently interpreted to represent the observer-independent properties of a single classically describable physical system, Everett takes the lesson of quantum mechanics to be that what would seem in the context of a given interaction to be the appearance of one is in reality the appearance of (in principle exponentially) many more. As for the globally Boolean description of the real physical situation concerning a system: Take this to be of its Hilbert space as a whole. All of the possibilities are realized. The universe becomes a multiverse.

On our view, in contrast, *values* come first. Not just the values of directly observable dynamical quantities but the values of fixed, non-dynamical, quantities as well. *A priori*, on a picture in which values come first, it remains open whether the mathematical description of reality that we manage to construct, on their basis, is most naturally interpreted in an observer-independent or an observer-dependent way. We take the former to be inadequate to the reality actually described by quantum mechanics, even though we recognize that it is still possible to make such a picture work if one really wants to do so. What quantum mechanics has shown us, on our interpretation, is that as enormously useful as pursuing the ideal of an observer-independent conception of reality has been, and continues to be in non-fundamental physics, there is a point beyond which the usefulness of pursuing it diminishes. In particular we remain unconvinced that any Boolean redescription of the phenomena described by quantum mechanics can yield anything other than the satisfaction of a pre-existing metaphysical urge to cast the world in accordance with the ideal of the detached observer.² Metaphysical urges are useful, sometimes.

¹ One of the contributions of Bub & Bub's (2018) "serious comic on entanglement" (see the essay review by Cuffaro & Doyle 2021) is to provide a really insightful analogy (the "choose your own adventure story" on pp. 98–101) through which to visualize the kind of world that Bubism takes to be suggested by quantum theory.

² In the context of quantum computing, Deutsch (2010, p. 542) has claimed that only the Everett interpretation is able to make sense of how quantum computers work. This claim does not stand up to scrutiny. See Cuffaro (2012, sec. 4), Cuffaro (2021b, sec. 3) and Duwell (2007) for reasons to be skeptical of the Everettian explanation of quantum computing, and

The classical worldview motivated by the ideal of the detached observer was tremendously successful for hundreds of years. But an observer-independent reality is no longer suggested by the formalism of fundamental physics. Despite this, quantum mechanics provides us with all of the tools that we need to give a description of the actual physical situation of the actual physical world that we are living in: a world that is at bottom objectively random in ways that are incompatible with a globally Boolean probability distribution over the values of the dynamical properties of systems; a world that we can acquire information about using classical means despite that (cf. Bohr, 1937, p. 293).³ There is nothing else to say. The change to the kinematical framework of fundamental physics brought about by the discovery of quantum mechanics may or may not be the last of such changes. But there is no going back.

Quantum mechanics is telling us, on our interpretation, that the choices that we make or do not make at a given moment in time actually have a bearing on the kinds of choices that we can and cannot make in the future. This should not sound that unfamiliar. It is characteristic, after all, both of the practice of physics and everyday life, that our actual choices have an impact on the actual world. Quantum mechanics goes beyond classical mechanics insofar as it shows us that the relation between our choices and the world runs deeper than one would have thought on the basis of the earlier theory. According to quantum mechanics, our choices affect the world at the level of possibilities (see Chapter 1). This motivates us to deny the ontological force of the classical ideal of the detached observer. But what is preserved and indeed affirmed, on our interpretation, is what we take to be the far more important strategy behind the success of classical physics: the *methodology* that begins with the values of fixed or dynamical quantities, considered under the (probabilistic) theoretical constraints that the theory places on such quantities, and yields up the objects of the actual world of our experience.

The world of our experience does not consist in probability distributions. Its objects include this table, that banana and the other dynamical objects we observe and interact with, both in the kitchens of the world and outside

see Hagar & Cuffaro (2019, sec. 5.1) for a summary of some of the other ways to explain the power of quantum computers. We noted Jeff Bub's explanation of the power of quantum computers in Chapter 3.

³ Bohr writes: "the proper rôle of the indeterminacy relations consists in *assuring quantitatively the logical compatibility* of apparently contradictory laws which appear when we use two different experimental arrangements, of which only one permits an unambiguous use of the concept of position, while only the other permits the application of the concept of momentum" (emphasis ours).

of them, every day. These objects will not be found within the quantum-kinematical framework, nor will the recipe we mentioned above yield them up in and of itself. As we emphasized in Section 6.3, the kinematical framework of quantum mechanics does not in itself provide us with specific observables. It encodes generic constraints, both on the possible values of observables as well as on the possible correlations between these values, that are satisfied for observables representable as Hermitian operators acting on Hilbert space. Determining the particular Hermitian operator representative of a given action on a quantum system is a province of the dynamics, not the kinematics, of the specific quantum theory of that system. And conditional upon a given measurement, the quantum-mechanical recipe allows one to transition from the quantum description of an interaction to the effectively classical description of the observations that ensue. From there we already know how to construct, on the basis of these observations, the familiar objects of our world.

As our examples have demonstrated, quantum theory is successful where classical theory fails in describing physical phenomena. But besides the particular things that it teaches us about the world there is a wider moral that we glean from the changes introduced to the kinematics of fundamental physics by quantum mechanics. The logical framework of classical physics is a globally Boolean structure, which invites us to speak of a world that exists in a particular way irrespective of our particular interactions with it. Quantum mechanics shows us that this classical conception is useful only up to a certain point, and that the logical structure of the world as it presents itself to us is globally non-Boolean, in ways that are in fact quantifiable. Whatever else we may discover in the course of the future development of physical theory, this is a non-trivial fact that we have discovered about our world. Moreover it is a fact that will remain with us (cf. Pitowsky, 1994, p. 98). It is, further, a non-trivial fact that we can effectively learn about our world, despite this non-Boolean character, through classical means. It will be objected that what we have just called “facts about the world” are really only relational facts about our connection to the world. This is entirely correct. But that, we maintain, is how it should be. For we are entangled with the world, and our concepts both of the world and of ourselves are only marginals of that true entangled description. That description, along with its many seemingly incompatible aspects, arises out of and is made possible through the non-Boolean probabilistic structure of the quantum-mechanical kinematical core.

There is more to atomic reality than the values of dynamical quantities. That an electron has, for instance, a rest mass and that it carries a particular charge are both examples of objectively true statements that we can make on the basis

of quantum theory which do not depend for their validity on any particular perspective. As for a system's dynamical state: Here too, quantum theory provides us with an objective description of the physical situation concerning that system that is valid irrespective of one's particular choices and irrespective of one's particular interests in making those choices. But the description that quantum theory provides to us of a given system's dynamical state *is unlike* the corresponding description given to us by classical theory. In quantum theory, what is exhibited to us through the quantum state description *is not the set of observer-independent properties* of the system of interest. What is exhibited, rather, is the structure of, interrelations between, and interdependencies among *the possible perspectives that one can impose upon that system* (see Section 6.5). In this way quantum theory informs us regarding the structure of the world—a world that *includes* ourselves—and our place within that structure.

Bibliography

- Aaronson, S. (2013). *Quantum Computing Since Democritus*. Cambridge: Cambridge University Press.
- Accardi, L., & Fedullo, A. (1982). On the statistical meaning of complex numbers in quantum mechanics. *Lettere al Nuovo Cimento*, *34*, 161–172.
- Acuña, P. (2014). On the empirical equivalence between special relativity and Lorentz's ether theory. *Studies in History and Philosophy of Modern Physics*, *46*, 283–302.
- Albert, D. Z. (1992). *Quantum Mechanics and Experience*. Cambridge, MA: Harvard University Press.
- Albert, D. Z. (2018). Quantum's leaping lizards (review of Becker (2018)). *The New York Review of Books*, *65*. No. 7 (April 19, 2018).
- Aldrich, J. (1998). Doing least squares: perspectives from Gauss and Yule. *International Statistical Review*, *66*, 61–81.
- Bacciagaluppi, G. (2017). Bohr's slit and Hermann's microscope. In Crull & Bacciagaluppi (2017), pp. 239–278.
- Bacciagaluppi, G., Crull, E., & Maroney, O. J. E. (2017). Jordan's derivation of blackbody fluctuations. *Studies in History and Philosophy of Modern Physics*, *60*, 23–34.
- Bacciagaluppi, G., & Valentini, A. (2009). *Quantum Theory at the Crossroads. Reconsidering the 1927 Solvay Conference*. Cambridge: Cambridge University Press.
- Balashov, Y., & Janssen, M. (2003). Presentism and relativity. *British Journal for the Philosophy of Science*, *54*, 327–346.
- Ball, P. (2018). *Beyond Weird: Why Everything You Thought You Knew About Quantum Physics is Different*. London: The Bodley Head.
- Baym, G. (1969). *Lectures on Quantum Mechanics*. Reading, MA: Addison-Wesley.
- Becker, A. (2018). *What is Real? The Unfinished Quest for the Meaning of Quantum Physics*. New York: Basic Books.
- Bell, J. S. (1964). On the Einstein-Podolsky-Rosen paradox. *Physics*, *1*, 195–200. Reprinted in Bell (1987, pp. 14–21) and (in facsimile) in Wheeler & Zurek (1983, pp. 403–408).
- Bell, J. S. (1984). Beables for quantum field theory. Tech. Rep. CERN-TH.4035/84, CERN. Reprinted in Bell (1987, pp. 159–166).

- Bell, J. S. (1987). *Speakable and Unspeakable in Quantum Mechanics*. Cambridge: Cambridge University Press. Page references to second edition (2004).
- Bell, J. S. (1990). Against measurement. *Physics World*, 8, 33–40.
- Ben-Menahem, Y. (2017). The PBR theorem: Whose side is it on? *Studies in History and Philosophy of Modern Physics*, 57, 80–88.
- Ben-Menahem, Y., & Hemmo, M. (Eds.) (2012). *Probability in Physics*. Heidelberg: Springer.
- Berkovitz, J. (2012). The world according to de Finetti. In [Ben-Menahem & Hemmo \(2012\)](#), pp. 249–280.
- Berkovitz, J. (2019). On de Finetti's instrumentalist philosophy of probability. *European Journal for the Philosophy of Science*, 9, 25 (48 pages).
- Bernhardt, C. (2019). *Quantum Computing for Everyone*. Cambridge, MA: The MIT Press.
- Bingham, N. H. (2000). Studies in the history of probability and statistics XLVI. Measure into probability: From Lebesgue to Kolmogorov. *Biometrika*, 87, 145–156.
- Birkhoff, G., & Kreyszig, E. (1984). The establishment of functional analysis. *Historia Mathematica*, 11, 258–321.
- Bohr, N. (1928). The quantum postulate and the recent development of atomic theory. *Nature*, 121, 580–590.
- Bohr, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 48, 696–702.
- Bohr, N. (1937). Causality and complementarity. *Philosophy of Science*, 4, 289–298.
- Bohr, N. (1948). On the notions of causality and complementarity. *Dialectica*, 2, 312–319.
- Bohr, N. (1958). Quantum physics and philosophy. In R. Klibansky (Ed.) *Philosophy in the Mid-Century: A Survey*, (pp. 308–314). Firenze: La Nuova Italia Editrice. Page reference to reprint in: J. Kalckar, Ed., *Niels Bohr: Collected Works*, Vol. 7 (pp. 385–394). Amsterdam: Elsevier.
- Born, M. (1926). Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik*, 38, 803–827.
- Born, M. (Ed.) (1971). *The Born-Einstein Letters*. Glasgow: MacMillan.
- Born, M., Heisenberg, W., & Jordan, P. (1926). Zur Quantenmechanik II. *Zeitschrift für Physik*, 35, 557–615. English translation in [Van der Waerden \(1968\)](#), pp. 321–385.
- Bowler, P. J. (2003). *Evolution. The History of an Idea*. Berkeley, Los Angeles, London: University of California Press. 3rd revised and expanded edition.

- Bravais, A. (1846). Analyse mathématique sur les probabilités des erreurs de situation d'un point. *Mémoires de l'Académie Royale des Sciences de l'Institut de France*, 9, 255–332.
- Brown, H. R. (2005). *Physical Relativity. Space-time Structure from a Dynamical Perspective*. Oxford: Oxford University Press.
- Brown, H. R., & Timpson, C. G. (2006). Why special relativity should not be a template for a fundamental reformulation of quantum mechanics. In W. Demopoulos, & I. Pitowsky (Eds.) *Physical Theory and its Interpretation*, (pp. 29–42). Berlin: Springer.
- Brukner, Č. (2017). On the quantum measurement problem. In *Quantum [Un] Speakables II*, (pp. 95–117). Springer.
- Brukner, Č. (2018). A no-go theorem for observer-independent facts. *Entropy*, 20, 350–1–350–10.
- Bub, J. (2004). Why the quantum? *Studies in History and Philosophy of Modern Physics*, 35, 241–266.
- Bub, J. (2007). Quantum computation from a quantum logical perspective. *Quantum Information and Computation*, 7, 281–296.
- Bub, J. (2008). Quantum computation and pseudotelepathic games. *Philosophy of Science*, 75, 458–472.
- Bub, J. (2010a). Von Neumann's 'no hidden variables' proof: A re-appraisal. *Foundations of Physics*, 40, 1333–1340.
- Bub, J. (2010b). Quantum computation: where does the speed-up come from? In A. Bokulich, & G. Jaeger (Eds.) *Philosophy of Quantum Information and Entanglement*, (pp. 231–246). Cambridge: Cambridge University Press.
- Bub, J. (2016). *Bananaworld. Quantum Mechanics for Primates*. Oxford: Oxford University Press. Slightly revised paperback edition: 2018.
- Bub, J. (2017). Why Bohr was (mostly) right. arXiv:1711.01604v1.
- Bub, J. (2019a). Interpreting the quantum world: old questions and new answers (essay review of Freire (2015), Becker (2018) and Ball (2018)). *Historical Studies in the Natural Sciences*, 49, 226–239.
- Bub, J. (2019b). Quantum entanglement and information. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2019 ed.
- Bub, J. (2020). In defense of a "single-world" interpretation of quantum mechanics. *Studies in History and Philosophy of Modern Physics*, 72, 251–255.
- Bub, J., & Demopoulos, W. (2010). Itamar Pitowsky 1950–2010. *Studies in History and Philosophy of Modern Physics*, 41, 85.

- Bub, J., & Pitowsky, I. (2010). Two dogmas about quantum mechanics. In [Saunders et al. \(2010\)](#), pp. 433–459.
- Bub, T., & Bub, J. (2018). *Totally Random. Why Nobody Understands Quantum Mechanics. A Serious Comic on Entanglement*. Princeton: Princeton University Press.
- Busch, P. (2003). Quantum states and generalized observables: A simple proof of Gleason's theorem. *Physical Review Letters*, *91*, 120403.
- Butler, J. W. (1968). On the Trouton-Noble experiment. *American Journal of Physics*, *36*, 936–941.
- C. G. K. (1919). Dr. H. E. J. Du Bois. *Nature*, *102*, 408–409.
- Cabello, A. (2019). Quantum correlations from simple assumptions. *Physical Review A*, *100*, 032120.
- Caulton, A. (2015). The role of symmetry in the interpretation of physical theories. *Studies in History and Philosophy of Modern Physics*, *52*, 153–162.
- Chiribella, G., D'Ariano, G. M., & Perinotti, P. (2010). Probabilistic theories with purification. *Physical Review A*, *81*, 062348.
- Chiribella, G., & Ebler, D. (2019). Quantum speedup in the identification of cause-effect relations. *Nature Communications*, *10*, 1472.
- Chiribella, G., & Spekkens, R. (2016). *Quantum Theory: Informational Foundations and Foils*. Dordrecht: Springer.
- Cirel'son, B. S. (1980). Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, *4*, 93–100.
- Clauser, J. F., Horne, M. A., Shimony, A., & Holt, R. A. (1969). Proposed experiment to test local hidden-variable theories. *Physical Review*, *23*, 880–884.
- Clifton, R., Bub, J., & Halvorson, H. (2003). Characterizing quantum theory in terms of information-theoretic constraints. *Foundations of Physics*, *33*, 1561–1591.
- Cohen-Tannoudji, B., C. Diu, & Laloë, F. (1977). *Quantum Mechanics*. New York: Wiley.
- Condon, E. U., & Shortley, G. H. (1935). *The Theory of Atomic Spectra*. Cambridge: Cambridge University Press.
- Crull, E., & Bacciagaluppi, G. (Eds.) (2017). *Grete Hermann: Between Physics and Philosophy*. Dordrecht: Springer.
- Cuffaro, M. E. (2010). The Kantian framework of complementarity. *Studies in History and Philosophy of Modern Physics*, *41*, 309–317.

- Cuffaro, M. E. (2012). Many worlds, the cluster-state quantum computer, and the problem of the preferred basis. *Studies in History and Philosophy of Modern Physics*, 43, 35–42.
- Cuffaro, M. E. (2020). Information causality, the Tsirelson bound, and the ‘being-thus’ of things. *Studies in History and Philosophy of Modern Physics*, 72, 266–277.
- Cuffaro, M. E. (2021a). Grete Hermann, quantum mechanics, and the evolution of Kantian philosophy. In J. Peijnenburg, & A. A. Verhaegh (Eds.) *Women in the History of Analytic Philosophy*. Dordrecht: Springer. Forthcoming.
- Cuffaro, M. E. (2021b). The philosophy of quantum computing. In E. Miranda (Ed.) *Quantum Computing in the Arts and Humanities*. Dordrecht: Springer. Forthcoming.
- Cuffaro, M. E., & Doyle, E. P. (2021). Essay review of Tanya and Jeffrey Bub’s *Totally Random: Why Nobody Understands Quantum Mechanics: A Serious Comic on Entanglement*. *Foundations of Physics*, 51, 28:1–28:16.
- Cuffaro, M. E., & Fletcher, S. C. (Eds.) (2018). *Physical Perspectives on Computation, Computational Perspectives on Physics*. Cambridge: Cambridge University Press.
- Curiel, E. (2019). Schematizing the observer and the epistemic content of theories. arXiv:1903.02182v3.
- Dakic, A., & Brukner, Č. (2011). Quantum theory and beyond: Is entanglement special? In H. Halvorson (Ed.) *Deep Beauty: Understanding the Quantum World through Mathematical Innovation*, (pp. 365–392). Cambridge: Cambridge University Press.
- Dantzig, G. B., & Thapa, M. N. (1997). *Linear programming 1: Introduction*. New York: Springer.
- Dantzig, G. B., & Thapa, M. N. (2003). *Linear programming 2: Theory and Extensions*. New York: Springer.
- D’Ariano, G. M., Chiribella, G., & Perinotti, P. (2017). *Quantum Theory from First Principles: An Informational Approach*. Cambridge: Cambridge University Press.
- Darrigol, O. (2012). *A History of Optics From Greek Antiquity and the Nineteenth Century*. Oxford: Oxford University Press.
- Darrigol, O. (2014). *Physics & Necessity. Rationalist Pursuits from the Cartesian Past to the Quantum Present*. Oxford: Oxford University Press.
- De Finetti, B. (1930). Spazi astratti metrici (\mathcal{D}_m). *Atti della Pontificia Accademia delle Scienze Nuovi Lincei*, 83, 248–256.
- De Finetti, B. (1937). A proposito di “correlazione”. *Supplemento Statistico ai Nuovi problemi di Politica, Storia ed Economia*, 3, 41–57. Page reference

- to English translation in *Journal Électronique d'Histoire des Probabilités et de la Statistique* (www.jehps.net) 4 (4) (2008): 1–15.
- Demopoulos, W. (2010). Effects and propositions. *Foundations of Physics*, 40, 368–389.
- Demopoulos, W. (2012). Generalized probability measures and the framework of effects. In [Ben-Menahem & Hemmo \(2012\)](#), pp. 201–217.
- Demopoulos, W. (2018). *On Theories*. Unpublished book manuscript.
- d'Espagnat, B. (1976). *Conceptual Foundations of Quantum Mechanics*. New York: Addison-Wesley. 2nd ed. (1st ed.: 1971).
- d'Espagnat, B. (2001). Reply to K A Kirkpatrick. arXiv:quant-ph/0111081v1.
- Deutsch, D. (2010). Apart from universes. In [Saunders et al. \(2010\)](#), pp. 542–552).
- Deza, M. M., & Laurent, M. (1997). *Geometry of Cuts and Metrics*. Berlin: Springer.
- Dieks, D. (2017). Von Neumann's impossibility proof: Mathematics in the service of rhetorics. *Studies in History and Philosophy of Modern Physics*, 60, 136–148.
- Dirac, P. A. M. (1926). The fundamental equations of quantum mechanics. *Proceedings of the Royal Society of London*, A109, 642–653. Reprinted in [Van der Waerden \(1968\)](#), pp. 307–320).
- Dirac, P. A. M. (1927). The physical interpretation of the quantum dynamics. *Proceedings of the Royal Society of London. Series A*, 113, 621–641.
- Dirac, P. A. M. (1930). *Principles of Quantum Mechanics*. Oxford: Clarendon.
- Doob, J. L. (1934). Probability and statistics. *Transactions of the American Mathematical Society*, 36, 759–775.
- Du Bois, H. (1894). *Magnetische Kreise, deren Theorie und Anwendung*. Berlin: Springer. English translation: *The Magnetic Circuit in Theory and Practice*. London: MacMillan, 1896.
- Duncan, A., & Janssen, M. (2007). On the verge of *Umdeutung* in Minnesota: Van Vleck and the correspondence principle. 2 pts. *Archive for History of Exact Sciences*, 61, 553–624, 625–671.
- Duncan, A., & Janssen, M. (2008). Pascual Jordan's resolution of the conundrum of the wave-particle duality of light. *Studies in History and Philosophy of Modern Physics*, 39, 634–666.
- Duncan, A., & Janssen, M. (2013). (Never) mind your p 's and q 's: Von Neumann versus Jordan on the foundations of quantum theory. *The European Physical Journal H*, 38, 175–259.

- Duncan, A., & Janssen, M. (2014). The trouble with orbits: the Stark effect in the old and the new quantum theory. *Studies in History and Philosophy of Modern Physics*, 48, 68–83.
- Duncan, A., & Janssen, M. (2015). The Stark effect in the Bohr-Sommerfeld theory and in Schrödinger's wave mechanics. In F. Aaserud, & H. Kragh (Eds.) *One Hundred Years of the Bohr Atom*, (pp. 217–271). Copenhagen: Det Kongelige Danske Videnskabernes Selskab.
- Duncan, A., & Janssen, M. (2019). *Constructing Quantum Mechanics, Vol. 1, The Scaffold, 1900–1923*. Oxford: Oxford University Press.
- Duwell, A. (2007). The many-worlds interpretation and quantum computation. *Philosophy of Science*, 74, 1007–1018.
- Earman, J. (1989). *World Enough and Space-Time: Absolute versus Relational Theories of Space and Time*. Cambridge, MA: The MIT Press.
- Eckert, M. (2013). *Arnold Sommerfeld. Science, Life and Turbulent Times*. New York: Springer.
- Ehrenfest, P. (1925). Energieschwankungen im Strahlungsfeld oder Kristallgitter bei Superposition quantisierter Eigenschwingungen. *Zeitschrift für Physik*, 34, 362–373.
- Einstein, A. (1909a). Zum gegenwärtigen Stand des Strahlungsproblems. *Physikalische Zeitschrift*, 10, 185–193.
- Einstein, A. (1909b). Über die Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung. *Physikalische Zeitschrift*, 10, 817–825.
- Einstein, A. (1919). Time, space, and gravitation. *The London Times*, November 28. Reprinted under a different title (What is the theory of relativity?) in A. Einstein, *Ideas and Opinions* (pp. 227–232). New York: Crown Publishers, 1954.
- Einstein, A., Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47, 777–780. Reprinted in facsimile in Wheeler & Zurek (1983, pp. 138–141).
- Eisberg, R., & Resnick, R. (1985). *Quantum Physics of Atoms, Molecules, Solids, Nuclei, and Particles*. New York: Wiley. 2nd ed. (1st ed.: 1974).
- Epstein, P. (1916). Zur Theorie des Starkeffektes. *Annalen der Physik*, 50, 489–521.
- Epstein, P. (1926). The Stark effect from the point of view of Schrödinger's quantum theory. *Physical Review*, 28, 695–710.
- Feynman, R. P., Leighton, R. B., & Sands, M. L. (1965). *The Feynman Lectures on Physics*. Reading, MA: Addison-Wesley. 3 Vols.

- Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika*, *10*, 507–521.
- Fisher, R. A. (1924). The distribution of the partial correlation coefficient. *Metron*, *3*, 329–332.
- Fix, A. (2019). *Sensible Mathematics: The Science of Music in the Age of the Baroque*. Ph.D. thesis, University of Minnesota.
- Forster, E. M. (1942). *Virginia Woolf. The Rede Lecture 1941*. Cambridge: Cambridge University Press.
- Franklin, A., Edwards, A. W. F., Fairbanks, D. J., Hartl, D. L., & Seidenfeld, T. (2008). *Ending the Mendel-Fisher Controversy*. Pittsburgh: University of Pittsburgh Press.
- Frauchiger, D., & Renner, R. (2018). Quantum theory cannot consistently describe the use of itself. *Nature Communications*, *9*, 3711.
- Freedman, S. J., & Clauser, J. F. (1972). Experimental test of local hidden-variable theories. *Physical Review Letters*, *28*, 938–941. Reprinted in facsimile in Wheeler & Zurek (1983, pp. 414–417).
- Freire, O., Jr (2015). *The Quantum Dissidents: Rebuilding the Foundations of Quantum Mechanics (1950–1990)*. Heidelberg: Springer.
- Fristedt, B. E., & Gray, L. F. (1997). *A Modern Approach to Probability Theory*. Boston, Basel, Berlin: Birkhäuser.
- Fuchs, C. A. (2017). Notwithstanding Bohr, the reasons for QBism. *Mind and Matter*, *15*, 245–300.
- Fuchs, C. A., & Stacey, B. C. (2019). Are non-Boolean event structures the precedence or consequence of quantum probability? arXiv:1912.10880v1.
- Gal, O., & Chen-Morris, R. (2013). *Baroque Science*. Chicago: University of Chicago Press.
- Ghirardi, G. (2018). Collapse theories. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2018 ed.
- Gilder, L. (2008). *The Age of Entanglement. When Quantum Physics Was Reborn*. New York: Knopf. Reprinted: New York: Vintage, 2009.
- Giovanelli, M. (2014). ‘But one must not legalize the mentioned sin’: Phenomenological vs. dynamical treatments of rods and clocks in Einstein’s thought. *Studies in History and Philosophy of Modern Physics*, *48*, 20–44.
- Glazebrook, R. (1922). *A Dictionary of Applied Physics*. Vol. II. London: MacMillan.
- Gleason, A. M. (1957). Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics*, *6*, 885–893.

- Goh, K. T., Kaniewski, J., Wolfe, E., Vértesi, T., Wu, X., Cai, Y., Liang, Y.-C., & Scarani, V. (2018). Geometry of the set of quantum correlations. *Physical Review A*, *97*, 022104–1–022104–20.
- Greenberger, D. M., Horne, M. A., & Zeilinger, A. (1989). Going beyond Bell's theorem. In Kafatos (1989, 69–72).
- Hagar, A., & Cuffaro, M. E. (2019). Quantum computing. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2019 ed.
- Hald, A. (1998). *A History of Mathematical Statistics from 1750 to 1930*. New York: Wiley.
- Handsteiner, J., Friedman, A. S., Rauch, D., Gallicchio, J., Liu, B., Hosp, H., Kofler, J., Bricher, D., Fink, M., Leung, C., et al. (2017). Cosmic Bell test: measurement settings from milky way stars. *Physical Review Letters*, *118*, 060401–1–7.
- Hardy, L. (2001). Quantum theory from five reasonable axioms. arXiv:quant-ph/0101012v4.
- Hardy, L. (2007). Towards quantum gravity: A framework for probabilistic theories with non-fixed causal structure. *Journal of Physics A*, *30*, 3081.
- Hardy, L. (2012). The operator tensor formulation of quantum theory. *Philosophical Transactions of the Royal Society A*, *370*, 3385–3417.
- Harrigan, N., & Spekkens, R. W. (2010). Einstein, incompleteness, and the epistemic view of quantum states. *Foundations of Physics*, *40*, 125–157.
- Healey, R. (2017). *The Quantum Revolution in Philosophy*. Oxford: Oxford University Press.
- Healey, R. (2019). The mysteries of quantum mechanics in graphic-novel form (review of Bub and Bub, 2018). *Physics Today*, *72*, 56–57.
- Heisenberg, W. (1925). Über die quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. *Zeitschrift für Physik*, *33*, 879–893. English translation in Van der Waerden (1968, pp. 261–276).
- Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift für Physik*, *43*, 172–198.
- Hermann, G. (1935). Die naturphilosophischen Grundlagen der Quantenmechanik. *Abhandlungen der Fries'schen Schule*, *6*, 75–152. English translation in Crull & Bacciagaluppi (2017, pp. 239–278).
- Hughes, R. I. G. (1989). *The Structure and Interpretation of Quantum Mechanics*. Cambridge, MA: Harvard University Press.
- Janssen, M. (2002). COI stories: Explanation and evidence in the history of science. *Perspectives on Science*, *10*, 457–522.

- Janssen, M. (2009). Drawing the line between kinematics and dynamics in special relativity. *Studies in History and Philosophy of Modern Physics*, 40, 26–52.
- Janssen, M. (2019). Arches and scaffolds: Bridging continuity and discontinuity in theory change. In A. C. Love, & W. C. Wimsatt (Eds.) *Beyond the Meme. Development and Structure in Cultural Evolution*, (pp. 95–199). Minneapolis: University of Minnesota Press.
- Janssen, M., & Lehner, C. (Eds.) (2014). *The Cambridge Companion to Einstein*. Cambridge: Cambridge University Press.
- Janssen, M., & Pernice, S. (2020). Sleeping Beauty on Monty Hall. *Philosophies*, 5, 15.
- Joas, C., & Lehner, C. (2009). The classical roots of wave mechanics: Schrödinger's transformations of the optical-mechanical analogy. *Studies in History and Philosophy of Modern Physics*, 40, 338–351.
- Jordan, P. (1927a). Über eine neue Begründung der Quantenmechanik. *Zeitschrift für Physik*, 40, 809–838.
- Jordan, P. (1927b). Über eine neue Begründung der Quantenmechanik II. *Zeitschrift für Physik*, 44, 1–25.
- Kafatos, M. (Ed.) (1989). *Bell's Theorem, Quantum Theory and Conceptions of the Universe*. Dordrecht: Kluwer.
- Kaiser, D. (2011). *How the Hippies Saved Physics. Science, Counterculture, and the Quantum Revival*. New York: Norton.
- Kakalios, J. (2009). *The Physics of Superheroes*. New York: Gotham Books. 2nd ed.
- Kakalios, J. (2018). A graphic tale of entanglement (review of Bub and Bub, 2018). *Physics World*, 31, 37–38.
- Kendall, M. G. (1961). *A Course in the Geometry of n Dimensions*. New York: Hafner.
- Kevles, D. J. (1985). *In the Name of Eugenics*. New York: Alfred A. Knopf. Reprint (with a new preface): Cambridge: Harvard University Press, 1995.
- Koberinski, A., & Müller, M. P. (2018). Quantum theory as a principle theory: Insights from an information theoretic reconstruction. In Cuffaro & Fletcher (2018), pp. 257–281.
- Kramers, H. A., & Heisenberg, W. (1925). Über die Streuung von Strahlung durch Atome. *Zeitschrift für Physik*, 31, 681–707. English translation in Van der Waerden (1968), pp. 223–252.
- Landau, L. J. (1988). Empirical two-point correlation functions. *Foundations of Physics*, 18, 440–460.

- Lassez, C., & Lassez, J.-L. (1992). Quantifier elimination for conjunctions of linear constraints via a convex hull algorithm. In B. R. Donald, D. Kapur, & J. L. Mundy (Eds.) *Symbolic and Numerical Computation for Artificial Intelligence*, (pp. 103–119). San Diego: Academic Press.
- Laurence, P., Hwang, T.-H., & Barone, L. (2008). Geometric properties of multivariate correlation in de Finetti's approach to insurance theory. *Journal Électronique d'Histoire des Probabilités et de la Statistique (www.jehps.net)*, 4, 1–13.
- Lindberg, D. C. (1976). *Theories of Vision from Al-Kindi to Kepler*. Chicago: University of Chicago Press.
- Lindberg, D. C. (1992). *The Beginnings of Western Science: The European Scientific Tradition in Philosophical, Religious, and Institutional Context, 600 BC to AD 1450*. Chicago: University of Chicago Press.
- Martínez, A. A. (2009). *Kinematics. The Lost Origins of Einstein's Relativity*. Baltimore: The Johns Hopkins University Press.
- Masanes, L., Galley, T. D., & Müller, M. P. (2019). The measurement postulates of quantum mechanics are operationally redundant. *Nature Communications*, 10, 1361.
- McGrayne, S. B. (2011). *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, & Emerged Triumphant from Two centuries of Controversy*. New Haven: Yale University Press.
- McIntyre, D. H. (2012). *Quantum Mechanics. A Paradigms Approach*. San Francisco: Pearson. 1st ed., 2nd ed.: McIntyre, Manogue, & Tate (2013).
- McIntyre, D. H., Manogue, C. A., & Tate, J. (2013). *Quantum Mechanics*. San Francisco: Pearson.
- McMahon, D. (2008). *Quantum Computing Explained*. Hoboken, NJ: Wiley.
- Mensing, L., & Pauli, W. (1926). Über die Dielektrizitätskonstante von Dipolgasen nach der Quantenmechanik. *Physikalische Zeitschrift*, 27, 509–512.
- Mermin, N. D. (1981). Quantum mysteries for everyone. *Journal of Philosophy*, 78, 397–408. Page references to reprint in Mermin (1990, pp. 81–94).
- Mermin, N. D. (1988). Spooky actions at a distance: mysteries of the quantum theory. In *The Great Ideas Today*. Encyclopaedia Britannica. Page references to reprint in Mermin (1990, pp. 110–176).
- Mermin, N. D. (1990). *Boojums All the Way Through. Communicating Science in a Prosaic Age*. Cambridge: Cambridge University Press.
- Mermin, N. D. (2007). *Quantum Computer Science. An Introduction*. Cambridge: Cambridge University Press.

- Messiah, A. (1962). *Quantum Mechanics*. New York: Wiley. 2 Vols.
- Midwinter, C., & Janssen, M. (2013). Kuhn losses regained: Van Vleck from spectra to susceptibilities. In M. Badino, & J. Navarro (Eds.) *Research and Pedagogy: A History of Early Quantum Physics Through Its Textbooks*, (pp. 137–205). Berlin: Edition Open Access.
- Müller, M. P. (2020). Law without law: from observer states to physics via algorithmic information theory. *Quantum*, 4, 301.
- Münster, G. (2020). (K)eine klassische Karriere? *Physik Journal*, 19, pp. 30–34. June 2020.
- Myrvold, W. (2010). From physics to information theory and back. In A. Bokulich, & G. Jaeger (Eds.) *Philosophy of Quantum Information and Entanglement*, (pp. 181–207). Cambridge: Cambridge University Press.
- Nagel, T. (1989). *The View from Nowhere*. Oxford: Oxford University Press.
- Nielsen, M. A., & Chuang, I. L. (2016). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
- Pais, A. (1979). Einstein and the quantum theory. *Reviews of Modern Physics*, 51, 863–914.
- Pauli, W. (1921). Zur Theorie der Dielektrizitätskonstante zweiatomiger Dipolgase. *Zeitschrift für Physik*, 6, 319–327.
- Pauli, W. (1994). *Writings on Physics and Philosophy*. Springer: Berlin. C. P. Enz and K. von Meyenn (Eds.).
- Pawłowski, M., Paterek, T., Kaszlikowski, D., Scarani, V., Winter, A., & Żukowski, M. (2009). Information causality as a physical principle. *Nature*, 461, 1101–1104.
- Pearson, K. (1896). Mathematical contributions to the theory of evolution III. Regression, heredity and panmixia. *Philosophical Transactions of the Royal Society, London, Series A*, 186, 343–414.
- Pearson, K. (1916). On some novel properties of partial and multiple correlation in a universe of manifold characteristics. *Biometrika*, 11, 231–238.
- Pitowsky, I. (1986). The range of quantum probabilities. *Journal of Mathematical Physics*, 27, 1556–1566.
- Pitowsky, I. (1989a). from George Boole to John Bell: The origins of Bell's inequalities. In Kafatos (1989), pp. 37–49).
- Pitowsky, I. (1989b). *Quantum Probability—Quantum Logic*. Heidelberg: Springer.
- Pitowsky, I. (1991). Correlation polytopes, their geometry and complexity. *Mathematical Programming A*, 50, 395–414.

- Pitowsky, I. (1994). George Boole's 'conditions of possible experience' and the quantum puzzle. *The British Journal for the Philosophy of Science*, 45, 95–125.
- Pitowsky, I. (2006). Quantum mechanics as a theory of probability. In W. Demopoulos, & I. Pitowsky (Eds.) *Physical Theory and its Interpretation. Essays in Honor of Jeffrey Bub*, (pp. 213–240). New York: Springer.
- Pitowsky, I. (2008). Geometry of quantum correlations. *Physical Review A*, 77, 062109.
- Pokorny, F., Zhang, C., Higgins, G., Cabello, C., Kleinmann, M., & Hennrich, M. (2019). Tracking the dynamics of an ideal quantum measurement. arXiv:1903.10398v1.
- Popescu, S. (2014). Nonlocality beyond quantum mechanics. *Nature Physics*, 10, 264–270.
- Popescu, S. (2016). Foreword. In Bub (2016, pp. v–vi).
- Popescu, S., & Rohrlich, D. (1994). Quantum nonlocality as an axiom. *Foundations of Physics*, 24, 370–385.
- Pusey, M. F., Barrett, J., & Rudolph, T. (2012). On the reality of the quantum state. *Nature Physics*, 8, 475–478.
- Quine, W. V. O. (1951). Two dogmas of empiricism. *The Philosophical Review*, 60, 20–43.
- Rastall, P. (1995). Locality, Bell's theorem, and quantum mechanics. *Foundations of Physics*, 15, 963–972.
- Rieffel, E., & Polak, W. (2011). *Quantum Computing: A Gentle Introduction*. Cambridge, MA: The MIT Press.
- Rovelli, C. (1996). Relational quantum mechanics. *International Journal of Theoretical Physics*, 35, 1637–1678.
- Sainz, A. B., Guryanova, Y., Acín, A., & Navascués, M. (2018). Almost-quantum correlations violate the no-restriction hypothesis. *Physical Review Letters*, 120, 200402.
- Sakurai, J. J., & Napolitano, J. (2017). *Modern Quantum Mechanics*. Cambridge: Cambridge University Press. 2nd Ed. (first ed.: 1994).
- Saunders, S., Barrett, J., Kent, A., & Wallace, D. (Eds.) (2010). *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford: Oxford University Press.
- Schlosshauer, M. (2007). *Decoherence and the Quantum-to-Classical Transition*. Berlin: Springer.
- Schrödinger, E. (1926). Quantisierung als Eigenwertproblem. 4 pts. *Annalen der Physik*, (pp. 79: 361–376, 489–527; 80: 437–490; 81: 109–139). Page reference to English translation in: *Collected Papers on Wave Mechanics*. 3rd (augmented) ed. New York: Chelsea Publishing Company, 1982.

- Schrödinger, E. (1935). Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31, 555–563.
- Schwarzschild, K. (1916). Zur quantenhypothese. *Königlich Preussischen Akademie der Wissenschaften (Berlin)*, Sitzungsberichte, 548–568.
- Sommerfeld, A. (1919). *Atombau und Spektrallinien*. Braunschweig: Vieweg.
- Stachel, J. (1994). Changes in the concepts of space and time brought about by relativity. In C. C. Gould, & R. S. Cohen (Eds.) *Artifacts, Representation and Social Practice*, (pp. 141–162). Dordrecht: Kluwer.
- Stark, J. (1913). Beobachtungen über den Effekt des elektrischen Feldes auf Spektrallinien. *Königlich Preussischen Akademie der Wissenschaften (Berlin)*, Sitzungsberichte, 932–946.
- Stone, M. H. (1932). *Linear Transformations in Hilbert Space*. New York: American Mathematical Society.
- Styer, D. F. (2000). *The Strange World of Quantum Mechanics*. Cambridge: Cambridge University Press.
- Tanabashi, M., Hagiwara, K., Hikasa, K., Nakamura, K., Sumino, Y., Takahashi, F., Tanaka, J., Agashe, K., Aielli, G., Amsler, C., et al. (2018). Review of particle physics (particle data group). *Physical Review D*, 98, 030001.
- Teukolsky, S. A. (1996). The explanation of the Trouton-Noble experiment revisited. *American Journal of Physics*, 64, 1104–1106.
- Timpson, C. G. (2010). Rabid dogma? Comments on Bub and Pitowsky [2010]. In Saunders et al. (2010), pp. 460–466.
- Timpson, C. G. (2013). *Quantum Information Theory and the Foundations of Quantum Mechanics*. Oxford: Oxford University Press.
- Timpson, C. G., & Brown, H. R. (2005). Proper and improper separability. *International Journal of Quantum Information*, 3, 679–690.
- Townsend, J. S. (2012). *A Modern Approach to Quantum Mechanics*. Herndon, VA: University Science Books. 2nd ed. (1st ed.: 1992).
- Van der Waerden, B. L. (Ed.) (1968). *Sources of Quantum Mechanics*. New York: Dover.
- Van Fraassen, B. C. (1991). *Quantum Mechanics. An Empiricist View*. Oxford: Clarendon.
- Van Vleck, J. H. (1926). Magnetic susceptibilities and dielectric constants in the new quantum mechanics. *Nature*, 118, 226–227.
- Van Vleck, J. H. (1928). The new quantum mechanics. *Chemical Reviews*, 5, 467–507.
- Van Vleck, J. H. (1932). *The Theory of Electric and Magnetic Susceptibilities*. Oxford: Oxford University Press.

- Von Mises, R. (1928). *Wahrscheinlichkeit, Statistik und Wahrheit*. Vienna: Springer.
- Von Neumann, J. (1927a). Mathematische Begründung der Quantenmechanik. *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten*, 1–57.
- Von Neumann, J. (1927b). Wahrscheinlichkeitstheoretischer Aufbau der Quantenmechanik. *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten*, 245–272.
- Von Neumann, J. (1927c). Thermodynamik quantenmechanischer Gesamtheiten. *Königliche Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-physikalische Klasse. Nachrichten*, 273–291.
- Von Neumann, J. (1932). *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer.
- Wallace, D. (2016). What is orthodox quantum mechanics? arXiv:1604.05973v1.
- Wallace, D. (2019). On the plurality of quantum theories: Quantum theory as a framework, and its implications for the quantum measurement problem. In S. French, & J. Saatsi (Eds.) *Realism and the Quantum*, (pp. 78–102). Oxford: Oxford University Press.
- Wheeler, J. H., & Zurek, W. H. (1983). *Quantum Theory and Measurement*. Princeton: Princeton University Press.
- Wigner, E. (1931). *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren*. Braunschweig: Vieweg.
- Yule, G. U. (1897). On the significance of Bravais' formulae for regression, &c., in the case of skew correlation. *Proceedings of the Royal Society of London*, 60, 477–489.