

Chapter 2

Contextuality

We have seen in the previous chapter that the analysis of von Neumann has little impact on the question of whether a viable hidden variables theory may be constructed. However, further mathematical results were developed by Gleason [1] in 1957 and by Simon Kochen and Specker in 1967 [2], which were claimed by some¹ to imply the impossibility of hidden variables. In the words of Kochen and Specker [2, p. 73]: “If a physicist X believes in hidden variables. . . . the prediction of X contradicts the prediction of quantum mechanics”. The Gleason, and Kochen and Specker arguments are in fact, *stronger* than von Neumann’s in that they assume linearity only for *commuting* observables. Despite this, a close analysis reveals that the impossibility proofs of Gleason and of Kochen and Specker share with von Neumann’s proof the neglect of the possibility of a hidden variables feature called *contextuality* [5, 6]. We will find that this shortcoming makes these theorems inadequate as proofs of the impossibility of hidden variables.

We begin this chapter with the presentation of Gleason’s theorem and the theorem of Kochen and Specker. This will be followed by a discussion of contextuality and its relevance to these analyses. We will make clear in this discussion why the theorems in question fail as impossibility proofs. As far as the question of what conclusions *do* follow, we show that these theorems’ implications can be expressed in a simple and concise fashion, which we refer to as “spectral-incompatibility”. We conclude the chapter with the discussion of an experimental procedure first discussed by Albert² which provides further insight into contextuality.

¹ See [2–4].

² See Albert [7].

2.1 Gleason's Theorem

Von Neumann's theorem addressed the question of the form taken by a function $E(O)$ of the observables. Gleason's theorem essentially addresses the same question,³ the most significant difference being that the linearity assumption is relaxed to the extent that it is demanded that E be linear on only *commuting* sets of observables. Besides this, Gleason's theorem involves a function E on only the projection operators of the system, rather than on all observables. Finally, Gleason's theorem contains the assumption that the system's Hilbert space is at least three dimensional. As for the conclusion of the theorem, this is identical to von Neumann's: $E(P)$ takes the form $E(P) = \text{Tr}(UP)$ where U is a positive operator and $\text{Tr}(U) = 1$.

Let us make the requirement of linearity on the commuting observables somewhat more explicit. First we note that any set of projection $\{P_1, P_2, \dots\}$ onto mutually orthogonal subspaces $\{\mathcal{H}_1, \mathcal{H}_2, \dots\}$ will form a commuting set. Furthermore, if P projects onto the direct sum $\mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \dots$ of these subspaces, then $\{P, P_1, P_2, \dots\}$ will also form a commuting set. It is in the case of this latter type of set that the linearity requirement comes into play, since these observables obey the relationship

$$P = P_1 + P_2 + \dots. \quad (2.1)$$

The condition on the function E is then

$$E(P) = E(P_1) + E(P_2) + \dots. \quad (2.2)$$

The formal statement of Gleason's theorem is expressed as follows. For any quantum system whose Hilbert space is at least three dimensional, any expectation function $E(P)$ obeying the conditions (2.2), $0 \leq E(P) \leq 1$, and $E(\mathbf{1}) = 1$ must take the form

$$E(P) = \text{Tr}(UP), \quad (2.3)$$

where $\text{Tr}(U) = 1$ and U is a positive operator. We do not present the proof of this result⁴ here. In the next section, we present an outline the proof of Kochen and Specker's theorem. The same impossibility result derivable from Gleason's work also follows from this theorem.

³ The original form presented by A.M. Gleason referred to a probability measure on the subspaces of a Hilbert space, but the equivalence of such a construction with a value map on the projection operators is simple and immediate. This may be seen by considering that there is a one-to-one correspondence between the subspaces and projections of a Hilbert space and that the values taken by the projections are 1 and 0, so that a function mapping projections to their eigenvalues is a special case of a probability measure on these operators.

⁴ See Bell [5]. Bell proves that any function $E(P)$ satisfying the conditions of Gleason's theorem cannot map the projection operators to their eigenvalues.

It is straightforward to demonstrate that the function $E(P)$ considered within Gleason's theorem cannot be a value map function on these observables. To demonstrate this, one may argue in the same fashion as was done by von Neumann, since the form developed here for $E(P)$ is the same as that concluded by the latter. (See [Sect. 1.4.3](#)). We recall that if $E(O)$ is to represent a dispersion free state specified by ψ and λ , it *must* take the form of such a value map, and $E(O)$ evidently cannot be the expectation function for such a state. It is on this basis that the impossibility of hidden variables has been claimed to follow from Gleason's theorem.

2.2 Kochen and Specker's Theorem

As with Gleason's theorem, the essential assumption of Kochen and Specker's theorem is that the expectation function $E(O)$ must be linear on commuting sets of observables. It differs from the former only in the set of observables considered. Gleason's theorem was addressed to the projection operators on a Hilbert space of arbitrary dimension N . Kochen and Specker consider the squares $\{s_{\theta,\phi}^2\}$ of the spin components of a spin 1 particle. One may note that these observables are formally identical to projection operators on a three-dimensional Hilbert space. Thus, the Kochen and Specker observables are a subset⁵ of the " $N = 3$ " case of the Gleason observables. Among the observables $\{s_{\theta,\phi}^2\}$ any subset $\{s_x^2, s_y^2, s_z^2\}$ corresponding to mutually orthogonal component directions x, y, z will be a commuting set. Each such set obeys the relationship

$$s_x^2 + s_y^2 + s_z^2 = 2. \quad (2.4)$$

For every such subset, Kochen and Specker require that $E(s_{\theta,\phi}^2)$ must obey

$$E(s_x^2) + E(s_y^2) + E(s_z^2) = 2. \quad (2.5)$$

Kochen and Specker theorem states that there exists no function $E(s_{\theta,\phi}^2)$ on the squares of the spin components of a spin 1 particle which maps each observable to either 0 or 1 and which satisfies (2.5).

We now make some comments regarding the nature of this theorem's proof. The problem becomes somewhat simpler to discuss when formulated in terms of a geometric model. Imagine a sphere of unit radius surrounding the origin in \mathbb{R}^3 . It is easy to see that each point on this sphere's surface corresponds to a direction in space, which implies that each point is associated with one observable of the set $\{s_{\theta,\phi}^2\}$. With this, E may be regarded as a function on the surface of the unit sphere. Since the eigenvalues of each of these spin observables are 0 and 1, it must be that $E(O)$ must take on these values, if it is to assume the form of a value map function. Satisfaction of (2.5) requires that for each set of mutually orthogonal directions, E must assign to

⁵ The set of projections on a three-dimensional space is actually a larger class of observables.

one of them 0 and to the other directions 1. To gain some understanding⁶ of why such an assignment of values must fail, we proceed as follows. To label the points on the sphere, we imagine that each point on the sphere to which the number 0 is assigned is painted red, and each point assigned 1 is painted blue. We label each direction as by the unit vector \hat{n} . Since one direction of every mutually perpendicular set is assigned red, then in total we require that one-third of the sphere is painted red. If we consider the components of the spin in *opposite* directions θ, ϕ , and $180^\circ - \theta, 180^\circ + \phi$, these values are always opposite, i.e., if s_x takes the value $+1$, then s_{-x} takes the value -1 . This implies that the values of $s_{\theta, \phi}^2$ and $s_{180^\circ - \theta, 180^\circ + \phi}^2$ will be *equal*. Therefore, we must have that points on the sphere lying directly opposite each other, i.e., the “antipodes”, must receive the same assignment from E . Suppose that one direction and its antipode are painted red. These points form the two poles of a great circle, and all points along this circle must then be painted blue, since all such points represent directions orthogonal to the directions of our two ‘red’ points. For every such pair of red points on the sphere, there must be many more blue points introduced, and we will find this makes it impossible to make one-third of the sphere red, as would be necessary to satisfy (2.5).

Suppose we paint the entire first octant of the sphere red. In terms of the coordinates used by geographers, this is similar to the region in the Northern hemisphere between 0 and 90° longitude. If the point at the north pole is painted red then the great circle at the equator must be blue. Suppose that the 0 and 90° meridians are also painted red. Then the octant which is the antipode of the first octant must also be painted red. This octant would be within the ‘southern hemisphere’ between 180 and 270° longitude. If we now apply the condition that for every point painted red, all points lying on the great circle defined with the point at its pole, we find that all remaining points of the sphere must then be colored blue, thereby preventing the addition of more red points. This assignment implies that more than two-thirds of the sphere is blue, therefore some sets of mutually perpendicular directions are all colored blue. An example of such a set of directions is provided by the points on the sphere’s surface lying in the mutually orthogonal directions represented by $(0.5, 0.5, -0.7071)$, $(-0.1464, 0.8535, 0.5)$, $(0.8535, -0.1464, 0.5)$. Each of these points lies in a quadrant to which we have assigned the color blue by the above scheme. This assignment of values to the spatial directions must therefore fail to meet the criteria demanded of the function $E(s_{\theta, \phi}^2)$ in Kochen and Specker’s premises: that $E(s_{\theta, \phi}^2)$ satisfies (2.5) and maps the observables to their values.

In their proof, Kochen and Specker show that for a discrete set of 117 different directions in space, it is impossible to give appropriate value assignments to the corresponding spin observables.⁷ Kochen and Specker then assert that hidden variables cannot agree with the predictions of quantum mechanics. Their conclusion is that if some physicist ‘X’, mistakenly decides to accept the validity of hidden variables then

⁶ We follow here the argument given in Belinfante [8, p. 38]

⁷ Since their presentation, the proof has been simplified by Peres in 1991 [9] whose proof is based on examination of 33 such \hat{n} vectors. We also note that a proof presented by Bell [5] may be shown to lead easily to a proof of Kochen and Specker. See Mermin in this connection [10].

“the prediction of X (for some measurements) contradicts the prediction of quantum mechanics” [2, p 73]. The authors cite a particular system on which one can perform an experiment they claim reveals the failure of the hidden variables prediction. We will demonstrate in the next section of this work that what follows from Kochen and Specker's theorem is only that a *non-contextual* hidden variables theory will conflict with quantum mechanics, so that the general possibility of hidden variables has not been disproved. Furthermore, we show that if the type of experiment envisioned by these authors is considered in more detail, it does not indicate where hidden variables must fail, but instead serves as an illustration indicating that the requirement of contextuality is a quite natural one.

More recent proofs have been offered involving smaller numbers of spin component observables rather than the 117 utilized by the original Kochen and Specker proof. In particular, Asher Peres has shown⁸ that value assignments to a particular set of 33 spin observables cannot be made such that quantum mechanical prescriptions are met.

2.3 Contextuality and Gleason's, and Kochen and Specker's Impossibility Proofs

We have seen that the theorems of Gleason, and Kochen and Specker each demonstrate the impossibility of value maps on some sets of observables such that the constraining relationships on each commuting set are obeyed. One might be at first inclined to conclude with Kochen and Specker that such results imply the impossibility of a hidden variables theory. However, if we consider that there exists a successful theory of hidden variables, namely Bohmian mechanics [12] (see Sect. 1.1.2), we see that such a conclusion is in error. Moreover, an explicit analysis of Gleason's theorem has been carried out by Bell [5, 6] and its inadequacy as an impossibility proof was shown. Bell's argument may easily be adapted⁹ to provide a similar demonstration regarding Kochen and Specker's theorem. The key concept underlying Bell's argument is that of *contextuality*, and we now present a discussion of this notion.

Essentially, contextuality refers to the dependence of measurement results on the detailed experimental arrangement being employed. In discussing this notion, we will find that an inspection of the quantum formalism suggests that contextuality is a natural feature to expect in a theory explaining the quantum phenomena. Furthermore we shall find that the concept is in accord with Niels Bohr's remarks regarding the fundamental principles of quantum mechanics. According to Bohr [13] “*a closer examination reveals that the procedure of measurement has an essential influence on the conditions on which the very definition of the physical quantities in question*

⁸ See [11].

⁹ As we have mentioned, since the observables are formally equivalent to projections on a three-dimensional Hilbert space, this theorem is actually a special case of Gleason's. Therefore, Bell's argument essentially addresses the Kochen and Specker theorem as well as Gleason's.

rests.” In addition, he stresses [14, p. 210] “*the impossibility of any sharp distinction between the behavior of atomic objects and the interaction with the measuring instruments which serve to define the conditions under which the phenomena appear.*” The concept of contextuality represents a concrete manifestation of the quantum theoretical aspect to which Bohr refers. We will first explain the concept itself in detail, and then focus on its relevance to the theorems of Gleason, and Kochen and Specker.

We begin by recalling a particular feature of the quantum formalism. In the presentation of this formalism given in chapter one, we discussed the representation of the system’s state, the rules for the state’s time evolution, and the rules governing the measurement of an observable. The measurement rules are quite crucial, since it is only through measurement that the physical significance of the abstract quantum state (given by ψ) is made manifest. Among these rules, one finds that any commuting set of observables may be measured simultaneously. With a little consideration, one is led to observe that the possibility exists for *a variety of different experimental procedures* to measure a single observable. Consider for example, an observable O which is a member of the commuting set $\{O, A_1, A_2, \dots\}$. We label this set as \mathcal{C} . A simultaneous measurement of the set \mathcal{C} certainly gives among its results a value for O and thus may be regarded as providing a measurement of O . It is possible that O may be a member of another commuting set of observables $\mathcal{C}' = \{O, B_1, B_2, \dots\}$, so that a simultaneous measurement of \mathcal{C}' also provides a measurement of O . Let us suppose further that the members of set $\{A_i\}$ fail to commute with those of $\{B_i\}$. It is then clear that experiments measuring \mathcal{C} and \mathcal{C}' are quite different, and hence must be distinct. A concrete difference appears for example, in the effects of such experiments on the system wave function. The measurement rules tell us that the wave function subsequent to an ideal measurement of a commuting set is prescribed by the Eq. 1.12, according to which the post-measurement wave function is calculated from the pre-measurement wave function by taking the projection of the latter into the joint-eigenspace of that set. Since the members of \mathcal{C} and \mathcal{C}' fail to commute, the joint-eigenspaces of the two are necessarily different, and the system wave function will not generally be affected in the same way by the two experimental procedures. Apparently the concept of ‘the measurement of an observable’ is *ambiguous*, since there can be distinct experimental procedures for the measurement of a single observable.

There are, in fact, more subtle distinctions between different procedures for measuring the same observable, and these also may be important. To introduce the experimental procedure of measurement into our formal notation, we shall write $\mathcal{E}(O)$, $\mathcal{E}'(O)$, etc., to represent experimental procedures used to measure the observable O . From what we have seen here, it is quite natural to expect that a hidden variables theory should allow for the possibility that *different experimental procedures*, e.g., $\mathcal{E}(O)$ and $\mathcal{E}'(O)$, for the measurement of an observable might yield *different results* on an individual system. This is contextuality.

Examples of observables for which there exist incompatible measurement procedures are found among the observables addressed in each of the theorems of Gleason, and Kochen and Specker. Among observables addressed by Gleason are the one-dimensional projection operators $\{P_\phi\}$ on an N -dimensional Hilbert space \mathcal{H}_N .

Consider a one-dimensional projection P_ϕ where ϕ belongs to two sets of orthonormal vectors given by $\{\phi, \psi_1, \psi_2, \dots\}$ and $\{\phi, \chi_1, \chi_2, \dots\}$. Note that the sets $\{\psi_1, \psi_2, \dots\}$ and $\{\chi_1, \chi_2, \dots\}$ are constrained only in that they span \mathcal{H}_ϕ^\perp (the orthogonal complement of the one-dimensional space spanned by ϕ). Given this, there exist examples of such sets for which some members of $\{\psi_1, \psi_2, \dots\}$ are distinct from and not orthogonal to the vectors in $\{\chi_1, \chi_2, \dots\}$. Since any distinct vectors that are not orthogonal correspond to projections which fail to commute, the experimental procedures $\mathcal{E}(P_\phi)$ measuring $\{P_\phi, P_{\psi_1}, P_{\psi_2}, \dots\}$ and $\mathcal{E}'(P_\phi)$ measuring $\{P_\phi, P_{\chi_1}, P_{\chi_2}, \dots\}$ are incompatible. The argument just given applies also to the Kochen and Specker observables (the squares of the spin components of a spin 1 particle), since these are formally identical to projections on a three-dimensional Hilbert space. To be explicit, the observable s_x^2 is a member of the commuting sets $\{s_x^2, s_y^2, s_z^2\}$ and $\{s_x^2, s_{y'}^2, s_{z'}^2\}$ where the y' and z' are oblique relative to the y and z axes. In this case, $s_{y'}^2, s_{z'}^2$ do not commute with s_y^2, s_z^2 . Thus, the experimental procedures to measure these sets are incompatible.

While it is true that the arguments against hidden variables derived from these theorems are superior to von Neumann's, since they require agreement only with operator relationships among commuting sets, these arguments nevertheless possess the following shortcoming. Clearly, the mathematical functions considered in each case, $E(P)$ and $E(s_{\theta, \phi}^2)$ do *not* allow for the possibility that the results of measuring each observable using different and possibly *incompatible* procedures may lead to different results. What the theorems demonstrate is that no hidden variables formulation *based on assignment of a unique value to each observable* can possibly agree with quantum mechanics. But this is a result we might well have expected from the fact that the quantum formalism allows the possibility of incompatible experimental procedures for the measurement of an observable. For this reason, neither of the theorems here considered—Gleason's theorem, and Kochen and Specker's theorem—imply the impossibility of hidden variables, since they fail to account for such a fundamental feature of the quantum formalism's rules of measurement.

2.3.1 Procedure to Measure the Kochen and Specker Observables

In a discussion of the implications of their theorem, Kochen and Specker mention a system for which well-known techniques of atomic spectroscopy may be used to measure the relevant spin observables. Although these authors mention this experiment to support their case against hidden variables, the examination of such an experiment actually reinforces the assertion that one should allow for contextuality—the very concept that refutes their argument against hidden variables.

Kochen and Specker note¹⁰ that for an atom of *orthohelium*¹¹ which is subjected to an electric field of a certain configuration, the first-order effects of this field on the electrons may be accounted for by adding a term of the form $aS_x^2 + bS_y^2 + cS_z^2$ to the electronic Hamiltonian. Here a, b, c are distinct constants, and S_x^2, S_y^2, S_z^2 are the squares of the components of the total spin of the two electrons with respect to the Cartesian axes x, y, z . The Cartesian axes are defined by the orientation of the applied external field. For such a system, an experiment measuring the energy of the electrons also measures the squares of the three spin components. To see this, note that the value of the perturbation energy will be $(a + b)$, $(a + c)$, or $(b + c)$ if the joint values of the set S_x^2, S_y^2, S_z^2 equal respectively $\{1, 1, 0\}$, $\{1, 0, 1\}$, or $\{0, 1, 1\}$.

To understand why the external electric field affects the orthohelium electrons in this way, consider the ground state of orthohelium.¹² The wave function of this state is given by a spatial part $\phi(r_1, r_2)$, depending only on r_1, r_2 (the radial coordinates of the electrons), multiplied by the spin part, which is a linear combination of the eigenvectors $\psi_{+1}, \psi_0, \psi_{-1}$ of S_z , corresponding to $S_z = +1, 0, -1$, respectively. Thus, the ground state may be represented by any vector in the three-dimensional Hilbert space spanned by the vectors $\phi(r_1, r_2)\psi_{+1}, \phi(r_1, r_2)\psi_0, \phi(r_1, r_2)\psi_{-1}$. The external electric field will have the effect of “lifting the degeneracy” of the state, i.e., the new Hamiltonian will not be degenerate in this space, but its eigenvalues will correspond to three unique orthogonal vectors. Suppose that we consider a particular set of Cartesian axes x, y, z . We apply an electric field which is of orthorhombic symmetry¹³ with respect to these axes. It can be shown¹⁴ that the eigenvectors of the Hamiltonian

¹⁰ One can derive the analogous first-order perturbation term arising for a charged particle of orbital angular momentum $L = 1$ in such an electric field using the fact that the joint-eigenstates of L_x^2, L_y^2, L_z^2 are the eigenstates of the potential energy due to the field. This latter result is shown in Kittel [15, p. 427].

¹¹ Orthohelium and parahelium are two species of helium which are distinguished by the total spin S of the two electrons: for the former we have $S = 1$, and for the latter $S = 0$. There is a rule of atomic spectroscopy which prohibits atomic transitions for which $\Delta S = 1$, so that no transitions from one form to the other can occur spontaneously.

¹² Using spectroscopic notation, this state would be written as the ‘ $2^3 S$ ’ state of orthohelium. The ‘ 2 ’ refers to the fact that the principal quantum number n of the state equals 2, ‘ S ’ denotes that the total orbital angular momentum is zero, and the ‘ 3 ’ superscript means that it is a spin triplet state. Orthohelium has no state of principal quantum number $n = 1$, since the Pauli exclusion principle forbids the ‘ $1^3 S$ ’ state.

¹³ Orthorhombic symmetry is defined by the criterion that rotation about either the x or y axis by 180° would bring such a field back to itself.

¹⁴ A straightforward way to see this is by analogy with a charged particle of orbital angular momentum $L = 1$. The effects of an electric or magnetic field on a charged particle of spin 1 are analogous to the effects of the same field on a charged particle of orbital angular momentum 1. To calculate the first-order effects of an electric field of orthorhombic symmetry for such a particle, one can examine the spatial dependence of the $L_z = 1, 0, -1$ states $\psi_{-1}, \psi_0, \psi_{+1}$, together with the spatial dependence of the perturbation potential $V(\mathbf{r})$, to show that the states $1/\sqrt{2}(\psi_1 - \psi_{-1})$, $1/\sqrt{2}(\psi_1 + \psi_{-1})$, and ψ_0 are the eigenstates of such a perturbation. A convenient choice of V for this purpose is $V = Ax^2 + By^2 + Cz^2$. See Kittel in [15, p. 427].

due to this field are $v_1 = 1/\sqrt{2}(\psi_1 - \psi_{-1})$, $v_2 = 1/\sqrt{2}(\psi_1 + \psi_{-1})$ and $v_3 = \psi_0$. We drop the factor $\phi(r_1, r_2)$ for convenience of expression. These vectors are *also* the joint-eigenvectors of the observables S_x^2, S_y^2, S_z^2 , as we can easily show. When expressed as a matrix in the $\{\psi_{+1}, \psi_0, \psi_{-1}\}$ basis, the vectors v_1, v_2, v_3 take the form

$$v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \quad (2.6)$$

$$v_2 = \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \quad (2.7)$$

$$v_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}. \quad (2.8)$$

If we then express S_x^2, S_y^2, S_z^2 as matrices in terms of the same basis, then by elementary matrix multiplication, one can show that v_1 corresponds to the joint-eigenvalue $\mu = \{0, 1, 1\}$, v_2 corresponds to $\mu = \{1, 0, 1\}$, and v_3 corresponds to $\mu = \{1, 1, 0\}$. Thus, the eigenvectors v_1, v_2, v_3 of the Hamiltonian term H' which arises from a perturbing electric field (defined with respect to x, y, z) are also the joint-eigenvectors of the set $\{S_x^2, S_y^2, S_z^2\}$. This implies that we can represent H' by the expression $aS_x^2 + bS_y^2 + cS_z^2$, where H' 's eigenvalues are $\{(b+c), (a+c), (a+b)\}$.

All of this leads to the following conclusion regarding the measurement of the spin of the orthohelium ground state electrons. Let the system be subjected to an electric field with orthorhombic symmetry with respect to a given set of Cartesian axes x, y, z . Under these circumstances, the measurement of the total Hamiltonian will yield a result equal (to first-order approximation) to the (unperturbed) ground state energy plus one of the perturbation corrections $\{(b+c), (a+c), (a+b)\}$. If the measured value of the perturbation energy is $(a+b)$, $(a+c)$, or $(b+c)$ then the joint values of the set $\{S_x^2, S_y^2, S_z^2\}$ are given respectively by $\{1, 1, 0\}$, $\{1, 0, 1\}$, or $\{0, 1, 1\}$.

It is quite apparent from this example that it would be unreasonable to require that a hidden variables theory must assign a single value to S_x^2 , independent of the experimental procedure.

2.4 Contextuality Theorems and Spectral Incompatibility

We saw in our discussion of von Neumann's theorem that its implications toward hidden variables amounted to the assertion that there can be no mathematical function $E(O)$ that is linear on the observables and which maps them to their eigenvalues. This

was neither a surprising, nor particularly enlightening result, since it follows also from a casual observation of some example of linearly related non-commuting observables, as we saw in examining the observables $1/\sqrt{2}(\sigma_x + \sigma_y)$, σ_x , σ_y of a spin $\frac{1}{2}$ particle. The theorems of Gleason, and Kochen and Specker, imply a somewhat less obvious type of impossibility: there exists no function $E(O)$ mapping the observables to their eigenvalues, which obeys all relationships constraining *commuting* observables. What we develop here is a somewhat simpler expression of the implication of these theorems. We will find that they imply the *spectral-incompatibility* of the value map function: there exists no mathematical function that assigns to each commuting set of observables a joint-eigenvalue of that set.

We begin by recalling the notions of joint eigenvectors and joint eigenvalues of a commuting set of observables (O^1, O^2, \dots). For a commuting set, the eigenvalue Eq. 1.5 $O|\psi\rangle = \mu|\psi\rangle$ is replaced by a set of relationships (1.7): $O^i|\psi\rangle = \mu^i|\psi\rangle$ $i = 1, 2, \dots$, one for each member of the commuting set. If a given $|\psi\rangle$ satisfies this relationship for *all* members of the set, it is referred to as a *joint-eigenvector*. The set of numbers (μ^1, μ^2, \dots) that allow the equations to be satisfied for this vector are collectively referred to as the *joint-eigenvalue* corresponding to this eigenvector, and the symbol $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots)$ is used to refer to this set. The set of all joint-eigenvalues $\{\boldsymbol{\mu}_a\}$ is given the name ‘joint-eigenspectrum’.

In general, the members of any given commuting set of observables might not be independent, i.e., they may be constrained by mathematical relationships. We label the relationships for any given commuting set $\{O^1, O^2, \dots\}$ as

$$\begin{aligned} f_1(O^1, O^2, \dots) &= 0 \\ f_2(O^1, O^2, \dots) &= 0 \\ &\vdots \end{aligned} \quad (2.9)$$

The Eqs. 2.1 in Gleason’s theorem, and 2.4 in Kochen and Specker’s theorem are just such relations. We now demonstrate the following two results. First, that every member of the joint-eigenspectrum must satisfy all relationships (2.9). Second, that any set of numbers ξ_1, ξ_2, \dots satisfying all of these relationships is a joint-eigenvalue.

To demonstrate the first of these, we suppose that $\boldsymbol{\mu} = (\mu^1, \mu^2, \dots)$ is a joint-eigenvalue of the commuting set $\{O^1, O^2, \dots\}$, with joint-eigenspace \mathcal{H} . We then consider the operation of $f_i(O^1, O^2, \dots)$ on a vector $\psi \in \mathcal{H}$ where $f_i(O^1, O^2, \dots) = 0$ is one of the relationships constraining the commuting set. We find

$$f_i(O^1, O^2, \dots)\psi = f_i(\mu_1, \mu_2, \dots)\psi = 0. \quad (2.10)$$

The second equality implies that $f_i(\mu_1, \mu_2, \dots) = 0$. Since $f_i(O^1, O^2, \dots) = 0$ is an arbitrary member of the relationships (2.9) for the commuting set $\{O^1, O^2, \dots\}$, it follows that every joint-eigenvalue $\boldsymbol{\mu}$ of the set must satisfy *all* such relationships.

We now discuss the demonstration of the second point. Suppose that the numbers $\{\xi_1, \xi_2, \dots\}$ satisfy all of (2.9) for some commuting set $\{O^1, O^2, \dots\}$. We consider

the following relation:

$$\left(\left[(O^1 - \mu_1^1)^2 + (O^2 - \mu_1^2)^2 + \dots \right] \left[(O^1 - \mu_2^1)^2 + (O^2 - \mu_2^2)^2 + \dots \right] \dots \right) \psi = 0. \quad (2.11)$$

Here, we operate on the vector ψ with a product whose factors each consist of a sum of various operators. The product is taken over all joint-eigenvalues $\{\mu_i\}$. We represent each joint-eigenvalue μ_i by a set $(\mu_i^1, \mu_i^2, \dots)$. The validity of (2.11) is easily seen. Since the joint-eigenspaces of any commuting set are complete, the vector ψ must lie within such a space. Suppose that $\psi \in \mathcal{H}_i$, where \mathcal{H}_i is the joint-eigenspace corresponding to μ_i . Then the i th factor in the product of operators in (2.11) must give zero when operating on ψ . Therefore, the entire product operating on ψ must also give zero. Since ψ is an arbitrary vector, it follows that

$$\left[(O^1 - \mu_1^1)^2 + (O^2 - \mu_1^2)^2 + \dots \right] \left[(O^1 - \mu_2^1)^2 + (O^2 - \mu_2^2)^2 + \dots \right] \dots = 0. \quad (2.12)$$

Note that (2.12) is itself a constraining relationship on the commuting set, so that it must be satisfied by the numbers (ξ_1, ξ_2, \dots) . This can only be true if these numbers form a joint-eigenvalue of $\{O^1, O^2, \dots\}$, and this is the result we were to prove.

From this, we can discern a simple way to understand the implications of the theorems of Gleason, and Kochen and Specker toward the question of a value map. The requirement (2.2) Gleason's theorem places on the function $E(P)$ can be restated as the requirement that for each commuting set, $E(P)$ must satisfy all the relationships constraining its members. From the above argument, it follows that this assumption is equivalent to the constraint that $E(P)$ must assign to each commuting set a joint-eigenvalue. The same is also true of assumption that $E(s_{\theta, \phi}^2)$ satisfy Eq. 2.5.

Thus, both of these theorems can be regarded as proofs of the impossibility of a function mapping the observables to their values such that each commuting set is assigned a joint-eigenvalue. An appropriate name for such a proof would seem to be 'spectral-incompatibility theorem.'

2.5 Albert's Example and Contextuality

In thinking about any given physical phenomenon, it is natural to try to picture to oneself the properties of the system being studied. In using the quantum formalism to develop such a picture, one may tend to regard the 'observables' of this formalism, i.e., the Hermitian operators (see Sect. 1.3), as representative of these properties. However, the central role played by the *experimental procedure* $\mathcal{E}(O)$ in the measurement of any given observable O seems to suggest that such a view of the operators may be untenable. We describe an experiment originally discussed by David Albert¹⁵ that indicates that this is indeed the case: the Hermitian operators cannot be regarded

¹⁵ See Albert in [7]. The experiment is also discussed by Ghirardi in [16].

as representative of the properties of the system.¹⁶ Albert considers two laboratory procedures that may be used to measure the z -component of the spin of a spin $\frac{1}{2}$ particle. Although the two procedures are quite similar to one another, they cannot be regarded as identical when considered in light of the hidden variables theory known as Bohmian mechanics. This is a particularly striking instance of contextuality, and it indicates the inadequacy of the conception that the spin operator σ_z represents an intrinsic property of the particle. From Albert's example, one can clearly see that the outcome of the σ_z measurement depends not only on the parameters of the particle itself, but also on the complete experimental setup.

The Albert example is concerned with the measurement of spin¹⁷ as performed using a Stern–Gerlach magnet. The schematic diagram given in Fig. 2.1. exhibits the configuration used in both of the measurement procedures to be described here. Note that we use a Cartesian coordinate system for which the x -axis lies along the horizontal direction with positive x directed toward the right, and the z -axis lies along the vertical direction with positive z directed upward. The y -axis (not shown) is perpendicular to the plane of the figure, and—since we use a right-handed coordinate system—positive y points into this plane. The long axis of the Stern–Gerlach magnet system is oriented along the x -axis, as shown. The upper and lower magnets of the apparatus are located in the directions of positive z and negative z . We define the Cartesian system further by requiring that x -axis (the line defined by $y = 0, z = 0$) passes through the center of the Stern–Gerlach magnet system.

In each experiment, the spin $\frac{1}{2}$ particle to be measured is incident on the apparatus along the positive x -axis. In the region of space the particle occupies before entering the Stern–Gerlach apparatus, its wave function is of the form

$$\psi_t(\mathbf{r}) = \varphi_t(\mathbf{r})(|\uparrow\rangle + |\downarrow\rangle), \quad (2.13)$$

where the vectors $|\uparrow\rangle$ and $|\downarrow\rangle$ are the eigenvectors of σ_z corresponding to eigenvalues $+\frac{1}{2}$ and $-\frac{1}{2}$, respectively. Here $\varphi_t(\mathbf{r})$ is a localized wave packet moving in the positive x direction toward the magnet.

We wish to consider two experiments that differ only in the orientation of the magnetic field inside the Stern–Gerlach apparatus. In experiment 1, the upper magnet has a strong magnetic north pole toward the region of particle passage, while the lower has a somewhat weaker magnetic south pole toward this region. In experiment 2, the magnets are such that the gradient of the field points in the *opposite* direction, i.e., the upper magnet has a strong magnetic *south* pole toward the region of passage, while the lower has a weak magnetic north pole towards it. After passing the Stern–Gerlach apparatus, the particle will be described by a wave function of one of the following forms:

¹⁶ This idea has been propounded by Daumer et al. in [17]. See also Bell in [18].

¹⁷ As is usual in discussions of Stern–Gerlach experiments, we consider only those effects relating to the interaction of the magnetic field with the magnetic moment of the particle. We consider the electric charge of the particle to be zero.

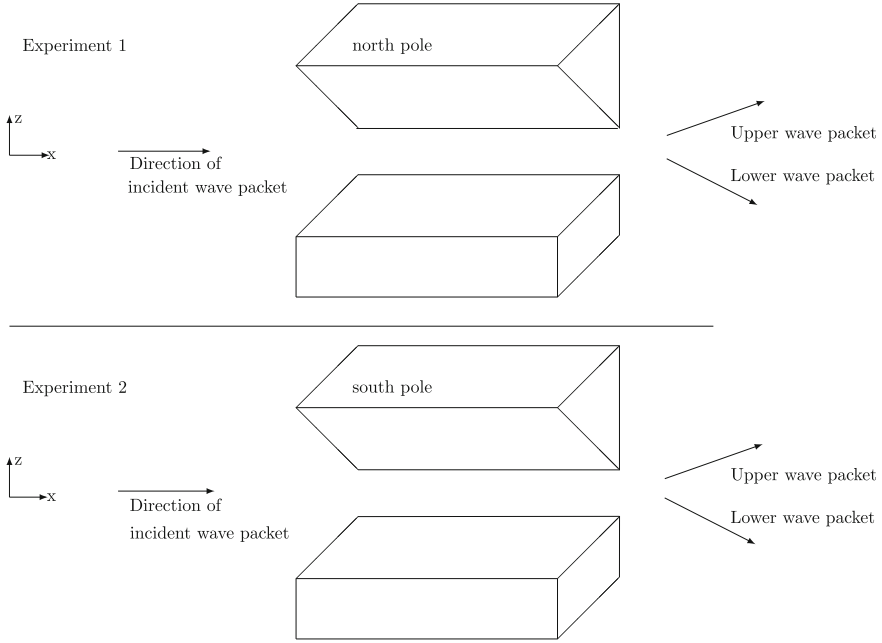


Fig. 2.1 Geometry of the Stern–Gerlach Experiments

$$\begin{aligned}\psi_t^1(\mathbf{r}) &= \frac{1}{\sqrt{2}}(\phi_t^+(\mathbf{r})|\uparrow\rangle + \phi_t^-(\mathbf{r})|\downarrow\rangle) \\ \psi_t^2(\mathbf{r}) &= \frac{1}{\sqrt{2}}(\phi_t^-(\mathbf{r})|\uparrow\rangle + \phi_t^+(\mathbf{r})|\downarrow\rangle).\end{aligned}\tag{2.14}$$

Here $\psi_t^1(\mathbf{r})$ corresponds to experiment 1, and $\psi_t^2(\mathbf{r})$ corresponds to experiment 2. In both cases, the function $\phi_t^+(\mathbf{r})$ represents a localized wave packet moving obliquely upward and $\phi_t^-(\mathbf{r})$ represents a localized wave packet moving obliquely downward. To measure σ_z , one places detectors in the paths of these wave packets. Examination of the first equation of (2.14) shows that for experiment 1, if the particle is detected in the upper path, the result of our σ_z measurement is $+\frac{1}{2}$. If the particle is detected in the lower path, the result is $-\frac{1}{2}$. For experiment 2, the second equation of (2.14) leads to the conclusion that similar detections are associated with results opposite in sign to those of experiment 1. Thus, for experiment 2, detection in the upper path implies $\sigma_z = -\frac{1}{2}$, while detection in the lower implies $\sigma_z = +\frac{1}{2}$.

We make here a few remarks regarding the symmetry of the system. We constrain the form of the wave packet $\varphi_t(\mathbf{r})$ of (2.13), by demanding that it has no dependence on y , and that it exhibits reflection symmetry through the plane defined by $z = 0$, i.e., $\varphi_t(x, z) = \varphi_t(x, -z)$. Moreover, the vertical extent of this wave packet is to be the same size as the vertical spacing between the upper and lower magnets of the

apparatus. As regards the wave packets $\phi_t^+(\mathbf{r})$ and $\phi_t^-(\mathbf{r})$ of five, if the magnetic field within the apparatus is such that $\partial B_z/\partial z$ is constant¹⁸ then for both experiments, these packets move at equal angles above and below the horizontal (See Fig. 2.1). Thus, the particle is described both before and after it passes the Stern–Gerlach magnet, by a wave function which has reflection symmetry through the plane defined by $z = 0$.

2.5.1 Bohmian Mechanics and Albert’s Example

We have mentioned that the hidden variables theory developed by David Bohm [12] gives an explanation of quantum phenomena which is empirically equivalent to that given by the quantum formalism. Bohmian mechanics allows us to regard any given system as a set of particles having well-defined (but distinctly non-Newtonian [20]) trajectories. Within Bohmian mechanics, it is the *configuration* of the system $\mathbf{q} = (q_1, q_2, q_3, \dots)$ which plays the role of the hidden variables parameter λ . Thus, the state description in this theory consists of both ψ and \mathbf{q} . Bohmian mechanics does not involve a change in the mathematical form of ψ : just as in the quantum formalism, ψ is a vector in the Hilbert space associated with the system, and it evolves with time according to the Schrödinger equation:

$$i\hbar \frac{\partial \psi}{\partial t} = H\psi. \quad (2.15)$$

The system configuration \mathbf{q} is governed by the equation:

$$\frac{d\mathbf{q}}{dt} = (\hbar/m) \text{Im} \left(\frac{\psi^* \nabla \psi}{\psi^* \psi} \right). \quad (2.16)$$

In the case of a particle with spin, we make use of the spinor inner-product in this equation. For example, in the case of a spin $\frac{1}{2}$ particle whose wave function is $\psi = \chi_+(\mathbf{r})|\uparrow\rangle + \chi_-(\mathbf{r})|\downarrow\rangle$, the Eq. 2.16 assumes the form

¹⁸ The term added to the particle’s Hamiltonian to account for a magnetic field is $g\mathbf{s} \cdot \mathbf{B}$, where \mathbf{s} is the spin, \mathbf{B} is the magnetic field and g is the gyromagnetic ratio. To determine the form of this term in the case of a Stern–Gerlach apparatus, we require the configuration of the magnetic field. A Stern–Gerlach magnet apparatus has a “long axis” which for the example of Fig. 2.1 lies along the x -axis. Since the component of the magnetic field along this axis will *vanish* except within a small region before and after the apparatus, the effects of B_x may be neglected. Furthermore, B_y and B_z within the apparatus may be regarded as being independent of x . The magnetic field in the x, z plane between the magnets lies in the z -direction, i.e., $\mathbf{B}(x, 0, z) = B_z(z)\hat{k}$. Over the region of incidence of the particle, the field is such that $\frac{\partial B_z}{\partial z}$ is constant. See for example, Weidner and Sells [19] for a more detailed discussion of the Stern–Gerlach apparatus. The motion of the particle in the y -direction is of no importance to us, and so we do not consider the effects of any Hamiltonian terms involving only y dependence. The results we discuss in the present section are those which arise from taking account of the magnetic field by adding to the Hamiltonian term of the form $g\sigma_z B_z(z=0) + g\sigma_z \left(\frac{\partial B_z}{\partial z}(z=0) \right) z$.

$$\frac{d\mathbf{q}}{dt} = (\hbar/m)\text{Im}\left(\frac{\chi_+^*(\mathbf{r})\nabla\chi_+(\mathbf{r}) + \chi_-^*(\mathbf{r})\nabla\chi_-(\mathbf{r})}{\chi_+^*(\mathbf{r})\chi_+(\mathbf{r}) + \chi_-^*(\mathbf{r})\chi_-(\mathbf{r})}\right). \quad (2.17)$$

As we expect from the fact that this theory is in empirical agreement with quantum theory, Bohmian mechanics does *not* generally provide, given ψ and \mathbf{q} , a mapping from the observables to their values. In other words, it does not provide a non-contextual value map for each state. As we shall see, the choice of *experimental procedure* plays such a pronounced role in the Bohmian mechanics description of Albert's spin measurements, that one cannot possibly regard the spin operator as representative of an objective property of the particle.

We first discuss the Bohmian mechanics description of the Albert experiments. Since the wave function and its evolution are the same as in quantum mechanics, the particle's wave function ψ is taken to be exactly as described above. As far as the configuration \mathbf{q} is concerned, there are two important features of the Bohmian evolution equations to be considered: the uniqueness of the trajectories and the *equivariance* of the time evolution. The first feature refers to the fact that each initial ψ and \mathbf{q} leads to a *unique* trajectory. Since the particle being measured has a fixed initial wave function, its initial conditions are defined solely by its initial position. The equivariance of the system's time evolution is a more complex property. Suppose that at some time t , the probability that the system's configuration is within the region $d\mathbf{q}$ about \mathbf{q} obeys the relationship:

$$P(\mathbf{q}' \in d\mathbf{q}) = |\psi(\mathbf{q})|^2 d\mathbf{q}. \quad (2.18)$$

According to equivariance, this relationship will continue to hold for all later times $t' : t' > t$. In considering the Bohmian mechanics description of a system, we assume that the particle initially obeys (2.18). By equivariance, we then have that for all later times the particle will be guided to “follow” the motion of the wave function. Thus, after it passes through the Stern–Gerlach apparatus, the particle will enter either the upward or downward moving packet. From consideration of the uniqueness of the trajectory and the equivariance of the time evolution, it follows that the question of *which branch* of the wave function the particle enters *depends solely on its initial position*.

If we consider the situation in a little more detail, we find a simple criterion on the initial position of the particle that determines which branch of the wave function it will enter. Recall that the initial (2.13) and final (2.14) wave function have no dependence on the y coordinate, and that they exhibit reflection symmetry through the $z = 0$ plane. From this symmetry together with the uniqueness of Bohmian trajectories, it follows that the particle cannot cross the $z = 0$ plane. In conjunction with equivariance, this result implies that if the particle's initial z coordinate is greater than zero, it must enter the upper branch, and if its initial z is less than zero the particle must enter the lower branch.

If we now consider the above described spin measurements, we find a somewhat curious situation. For any given initial configuration \mathbf{q} , the question of whether the measurement result is $\sigma_z = +\frac{1}{2}$ or $\sigma_z = -\frac{1}{2}$ depends on the configuration of

the experimental apparatus. Suppose that the particle has an initial $z > 0$, so that according to the results just shown, it will enter the upper branch of the wave function. If the magnetic field inside the Stern–Gerlach apparatus is such that $\partial B_z / \partial z < 0$, then the particle’s final wave function is given by the first equation in (2.14), and its detection in the upper branch then implies that $\sigma_z = +\frac{1}{2}$. If, on the other hand, the magnetic field of the Stern–Gerlach magnet has the *opposite* orientation, i.e., $\partial B_z / \partial z > 0$, then the second equation in (2.14) obtains and the detection in the upper branch implies that $\sigma_z = -\frac{1}{2}$. Thus, we arrive at the conclusion that the “measurement of σ_z ” gives a *different* result for two situations that differ only in the experimental configuration.

The quantum formalism’s rules of measurement strongly suggest that the Hermitian operators represent objective properties of the system. Moreover, such a conception is a common element of the expositions given in quantum mechanics textbooks. On the other hand, the fact that the result of the “measurement” of σ_z can depend on properties of *both* system *and* apparatus contradicts this conception. In general, one must consider the results of the “measurement of an observable” to be a joint-product of system and measuring apparatus. Recall Niels Bohr’s comment that [13] “a closer examination reveals that the procedure of measurement has an essential influence on the conditions on which the very definition of the physical quantities in question rests.” For further discussion of the role of Hermitian operators in quantum theory, the reader is directed to Daumer, Dürr, Goldstein, and Zanghì in [17]. According to these authors: “*the basic problem with quantum theory . . . more fundamental than the measurement problem and all the rest, is a naive realism about operators . . . by (this) we refer to various, not entirely sharply defined, ways of taking too seriously the notion of operator-as-observable, and in particular to the all too casual talk about ‘measuring operators’ which tends to occur as soon as a physicist enters quantum mode.*”

References

1. Gleason, A.M.: Measures on the closed subspaces of a Hilbert space. J. Math. Mech. **6**, 885–893 (1957)
2. Kochen, S., Specker, E.P.: The problem of hidden variables in quantum mechanics. J. Math. Mech. **17**, 59 (1967)
3. Mermin, N.D.: Simple unified form for the major no-hidden-variables theorems. Phys. Rev. Lett. **65**, 3373 (1990)
4. Pagels, H.R.: The Cosmic Code. Simon and Schuster, New York (1982)
5. Bell, J.S.: On the problem of hidden variables in quantum mechanics. Rev. Mod. Phys. **38**, 447–452 (1966). Reprinted in [21, p. 1] and [22, p. 397]
6. Bell, J.S.: On the impossible pilot wave. Found. Phys. **12**, 989–999 (1982). Reprinted in [1, p. 159]
7. Albert, D.: Quantum Mechanics and Experience. Harvard University Press, Cambridge (1992)
8. Belinfante, F.J.: A Survey of Hidden-Variables Theories. Pergamon Press, New York (1973)
9. Peres, A.: Incompatible results of quantum measurements. Phys. Lett. A **151**, 107 (1990)

10. Mermin, N.D.: Hidden variables and the two theorems of John Bell. *Rev. Mod. Phys.* **65**, 803 (1993)
11. Peres, A.: *Quantum Theory: Concepts and Methods*. Springer, Heidelberg (1995)
12. Bohm, D.: A suggested interpretation of the quantum theory in terms of “Hidden” variables. *Phys. Rev.* **85**, 166, 180 (1952)
13. Bohr, N.: Quantum mechanics and physical reality. *Nature* **136**, 65 (1935). Reprinted in [22, p. 144]
14. Schilpp, P.A.: *Albert Einstein: Philosopher-Scientist*. Harper and Row, New York (1949)
15. Kittel, C.: *Introduction to Solid State Physics*. 7th edn., Wiley, New York (1996)
16. Ghirardi, G.C.: *Sneaking a Look at God’s Cards*. Princeton University Press, Princeton (2005)
17. Daumer, M., Dürr, D., Goldstein, S., Zanghí, N.: Naive realism about operators. In: *Erkenntnis*, 1996, Special issue in honor of Prof. Jeffrey, R., Costantini, D., Gallavotti, M.C. (eds.) *Proceedings of the International Conference “Probability, Dynamics and Causality”*, Luino, Italy, 15–17 June 1995.
18. Bell, J.S.: Against ‘Measurement’. *Phys. World* **3**, 33 (1990). This is reprinted in [23, p. 902]
19. Weidner, R.T., Sells, R.L.: *Elementary Modern Physics*, 2nd edn., Allyn and Bacon, Boston (1980)
20. Dürr, D., Goldstein, S., Zanghí, N.: Quantum equilibrium and the origin of absolute uncertainty. *J. Statis. Phys.* **67**, 843–907 (1992)
21. Bell, J.S.: *Speakable and Unspeakable in Quantum Mechanics*. Cambridge University Press, Cambridge (1987). (1987) Many of the works by Bell which are of concern to us may be found in this reference. See also [23] and [24]. The latter two are complete collections containing all of Bell’s papers on quantum foundations
22. Wheeler, J.A., Zurek, W.H. (eds.): *Quantum Theory and Measurement*. Princeton University Press, Princeton (1983)
23. Bell, M., Gottfried, K., Veltman, M. (eds.): *Bell on the Foundations of Quantum Mechanics, High Energy Physics and Accelerators: Selected Papers of John S. Bell (with commentary)*. World Scientific Publishing Company, Singapore (1995)
24. Bell, M., Gottfried, K., Veltman, M. (eds.): *Bell on the Foundations of Quantum Mechanics*. World Scientific Publishing Company, Singapore (2001)

Bell's Theorem and Quantum Realism

Reassessment in Light of the Schrödinger Paradox

Hemmick, D.L.; Shakur, A.M.

2012, XII, 97 p. 1 illus., Softcover

ISBN: 978-3-642-23467-5