Chapter 14

Dependent (Contextual) Events

14.1 An Example

Consider two spin-half particles \(a\) and \(b\), and suppose that the corresponding Boolean algebra \(L\) of properties on the tensor product space \(A \otimes B\) is generated by a sample space of four projectors,

\[
\begin{align*}
[z_a^+ \otimes z_b^+], & \quad [z_a^+ \otimes z_b^-], & \quad [z_a^- \otimes z_b^+], & \quad [z_a^- \otimes x_b^-], \\
[z_a^+ \otimes x_b^+], & \quad [z_a^- \otimes x_b^+], & \quad [x_a^+ \otimes x_b^-], & \quad [x_a^- \otimes x_b^-],
\end{align*}
\]  

(14.1)

which sum to the identity operator \(I \otimes I\). Let \(A = [z_a^+]\) be the property that \(S_{az} = +1/2\) for particle \(a\), and its negation \(\tilde{A} = I - A = [z_a^-]\) the property that \(S_{az} = -1/2\). Likewise, let \(B = [z_b^+]\) and \(\tilde{B} = I - B = [z_b^-]\) be the properties \(S_{bz} = +1/2\) and \(S_{bz} = -1/2\) for particle \(b\). Together with the projectors \(AB\) and \(A\tilde{B}\), the first two items in (14.1), the Boolean algebra \(L\) also contains their sum

\[A = AB + A\tilde{B}\]  

(14.2)

and its negation \(\tilde{A}\). On the other hand \(L\) does not contain the projector \(B\) or its negation \(\tilde{B}\), as is obvious from the fact that these operators do not commute with the last two projectors in (14.1). Thus when using the framework \(L\) one can discuss whether \(S_{az}\) is \(+1/2\) or \(-1/2\) without making any reference to the spin of particle \(b\). But it only makes sense to discuss whether \(S_{bz}\) is \(+1/2\) or \(-1/2\) when one knows that \(S_{az} = +1/2\). That is, one cannot ascribe a value to \(S_{bz}\) in an absolute sense without making any reference to the spin of particle \(a\).

If it makes sense to talk about a property \(B\) when a system possesses the property \(A\) but not otherwise, we shall say that \(B\) is a contextual property: it is meaningful only within a certain context. Also we shall say that \(B\) depends on \(A\), and that \(A\) is the base of \(B\). (One might also call \(A\) the support of \(B\).) A slightly more restrictive definition is given in Sec. 14.3 below, and generalized to contextual events which do not have a base. It is important to notice that contextuality and the corresponding dependence is very much a function of the Boolean algebra \(L\) employed for constructing a quantum description. For example, the Boolean algebra \(L'\) generated by

\[
[z_a^+ \otimes z_b^+], \quad [z_a^+ \otimes z_b^-], \quad [z_a^- \otimes z_b^+], \quad [z_a^- \otimes x_b^-]
\]

(14.3)

contains both \(A = [z_a^+]\) and \(B = [z_b^+]\), and thus in this algebra \(B\) does not depend upon \(A\). And in the algebra \(L''\) generated by

\[
[z_a^+ \otimes z_b^+], \quad [z_a^- \otimes z_b^+], \quad [z_a^+ \otimes x_b^-], \quad [x_a^- \otimes z_b^-],
\]

(14.4)

167
the property \( A \) is contextual and depends on \( B \).

Since quantum theory does not prescribe a single “correct” Boolean algebra of properties to use in describing a quantum system, whether or not some property is contextual or dependent on another property is a consequence of the physicist’s choice to describe a quantum system in a particular way and not in some other way. In particular, when \( B \) depends on \( A \) in the sense we are discussing, one should not think of \( B \) as being caused by \( A \), as if the two properties were linked by a physical cause. The dependence is logical, not physical, and has to do with what other properties are or are not allowed as part of the description based upon a particular Boolean algebra.

14.2 Classical Analogy

It is possible to construct an analogy for quantum contextual properties based on purely classical ideas. The analogy is somewhat artificial, but even its artificial character will help us understand better why dependency is to be expected in quantum theory, when it normally does not show up in classical physics. Let \( x \) and \( y \) be real numbers which can take on any values between 0 and 1, so that pairs \((x, y)\) are points in the unit square, Fig. 14.1. In classical statistical mechanics one sometimes divides up the phase space into nonoverlapping cells (Sec. 5.1), and in a similar way we shall divide up the unit square into cells of finite area, and regard each cell as an element of the sample space of a probabilistic theory. The sample space corresponding to the cells in Fig. 14.1(a) consists of 4 mutually exclusive properties

\[
\{0 \leq x < 1/2, 0 \leq y < 1/2\}, \quad \{0 \leq x < 1/2, 1/2 \leq y \leq 1\},
\{1/2 \leq x \leq 1, 0 \leq y < 1/2\}, \quad \{1/2 \leq x \leq 1, 1/2 \leq y \leq 1\}. \tag{14.5}
\]

Let \( A \) be the property \( 0 \leq x < 1/2 \), so its complement \( \bar{A} \) is \( 1/2 \leq x \leq 1 \), and let \( B \) be the property \( 0 \leq y < 1/2 \), so \( \bar{B} \) is \( 1/2 \leq y \leq 1 \). Then the four sets in (14.5) correspond to the properties \( A \wedge B \), \( A \wedge \bar{B} \), \( \bar{A} \wedge B \), \( \bar{A} \wedge \bar{B} \). It is then obvious that the Boolean algebra of properties generated by (14.5) contains both \( A \) and \( B \), so (14.5) is analogous in this respect to the quantum sample space (14.3).

Figure 14.1: Unit square in the \( x, y \) plane: (a) shows the set of cells in (14.5), (b) the set of cells in (14.6), and (c) the cells in a common refinement (see text). Property \( A \) is represented by the vertical rectangular cell on the left, and \( B \) by the horizontal rectangular cell (not present in (b)) on the bottom. The gray region represents \( A \wedge B \).

An alternative choice for cells is shown in Fig. 14.1(b), where the four mutually exclusive
14.3. CONTEXTUAL PROPERTIES AND CONDITIONAL PROBABILITIES

If $A$ and $B$ are defined in the same way as before, the new algebra of properties generated by (14.6) contains $A$ and $A \wedge B$, but does not contain $B$. In this respect it is analogous to (14.1) in the quantum case, and $B$ is a contextual or dependent property: it only makes sense to ask whether the system has or does not have the property $B$ when the property $A$ is true, i.e., when $x$ is between 0 and 1/2, but the same question does not make sense when $x$ is between 1/2 and 1, that is, when $A$ is false.

Isn’t this just some sort of formal nitpicking? Why not simply refine the sample space of Fig. 14.1(b) by using the larger collection of cells shown in Fig. 14.1(c)? The corresponding Boolean algebra of properties includes all those in (14.6), so we have not the lost the ability to describe whatever we would like to describe, and now $B$ as well as $A$ is part of the algebra of properties, so dependency is no longer of any concern. Such a refinement of the sample space can always be employed in classical statistical mechanics. However, a similar type of refinement may or may not be possible in quantum mechanics. There is no way to refine the sample space in (14.1), for the four projectors in that list already project onto one dimensional subspaces, which is as far as a quantum refinement can go. The move from (b) to (c) in Fig. 14.1, which conveniently gets rid of contextual properties in a classical context, will not work in the case of (14.1); the latter is an example of an irreducible contextuality.

To be more specific, the refinement in Fig. 14.1(c) is obtained by forming the products of the indicators for $B$, $\bar{B}$, $B'$, and $\bar{B}'$ with one another and with $A$ and $\bar{A}$, where $B'$ is the property $0 \leq y < 2/3$. The analogous process for (14.1) would require taking products of projectors such as $[z_b^+]$ and $[x_b^+]$, but since they do not commute with each other, their product is not a projector. That non-commutativity of the projectors is at the heart of the contextuality associated with (14.1) can also be seen by considering two classical spinning objects $a$ and $b$ with angular momenta $L_a$ and $L_b$, and interpreting $[z_a^+]$ and $[z_b^+]$ in (14.1) as $L_{az} \geq 0$ and $L_{az} < 0$, etc. In the classical case there is no difficulty refining the sample space of (14.1) to get rid of dependency, for $[z_b^+][x_b^+]$ is the property $L_{bz} \geq 0 \wedge L_{bx} \geq 0$, which makes perfectly good (classical) sense. But its quantum counterpart for a spin-half particle has no physical meaning.

14.3. CONTEXTUAL PROPERTIES AND CONDITIONAL PROBABILITIES

If $A$ and $B$ are elements of a Boolean algebra $\mathcal{L}$ for which a probability distribution is defined, then

$$\Pr(B \mid A) = \frac{\Pr(AB)}{\Pr(A)}$$

(14.7)

is defined provided $\Pr(A)$ is greater than zero. If, however, $B$ is not an element of $\mathcal{L}$, then $\Pr(B)$ is not defined and, as a consequence, $\Pr(A \mid B)$ is also not defined. In view of these remarks it makes sense to define $B$ as a contextual property which depends upon $A$, $A$ is the base of $B$, provided $\Pr(B \mid A)$ is positive (which implies $\Pr(AB) > 0$), whereas $\Pr(B)$ is undefined. This definition is stricter than the one in Sec. 14.1, but the cases it eliminates—those with $\Pr(B \mid A) = 0$—are
in practice rather uninteresting. In addition, one is usually interested in situations where the
dependence is irreducible, i.e., it cannot be eliminated by appropriately refining the sample space,
unlike the classical example in Sec. 14.2.

One can extend this definition to events which depend on other contextual events. For example,
let \( A, B, \) and \( C \) be commuting projectors, and suppose \( A, AB, \) and \( ABC \) belong to the Boolean
algebra, but \( B \) and \( C \) do not. Then as long as

\[
\Pr(C | AB) = \frac{\Pr(ABC)}{\Pr(AB)} \tag{14.8}
\]

is positive, we shall say that \( C \) depends on \( B \) (or on \( AB \)), and \( B \) depends on \( A \). Note that if (14.8)
is positive, so is \( \Pr(AB) \), and thus \( \Pr(B | A) \), (14.7), is also positive.

There are situations in which the properties \( A \) and \( B \), represented by commuting projectors,
are contextual even though neither can be said to depend upon or be the base of the other. That
is, \( AB \) belongs to the Boolean algebra \( \mathcal{L} \) and has positive probability, but neither \( A \) nor \( B \) belongs
to \( \mathcal{L} \). In this case neither \( \Pr(A | B) \) nor \( \Pr(B | A) \) is defined, so one cannot say that \( B \) depends
on \( A \) or \( A \) on \( B \), though one might refer to them as “codependent”. As an example, let \( A \) and \( B \)
be two Hilbert spaces of dimension 2 and 3, respectively, with orthonormal bases \( \{|0a\}, |1a\} \) and
\( \{|0b\}, |1b\}, |2b\} \). In addition, define

\[
|+b\rangle = (|0b\rangle + |1b\rangle)/\sqrt{2}, \quad |-b\rangle = (|0b\rangle - |1b\rangle)/\sqrt{2}, \tag{14.9}
\]

and \( |+a\rangle \) and \( |-a\rangle \) in a similar way. Then the six kets

\[
|0a\rangle \otimes |0b\rangle, \quad |1a\rangle \otimes |+b\rangle, \quad |+a\rangle \otimes |2b\rangle,
|0a\rangle \otimes |1b\rangle, \quad |1a\rangle \otimes |-b\rangle, \quad |-a\rangle \otimes |2b\rangle, \tag{14.10}
\]

form an orthonormal basis for \( A \otimes B \), and the corresponding projectors generate a Boolean algebra
\( \mathcal{L} \). If \( A = |0a\rangle \otimes I \) and \( B = I \otimes |0b\rangle \), then \( \mathcal{L} \) contains \( AB \), corresponding to the first ket in
(14.10), but neither \( A \) nor \( B \) belongs to \( \mathcal{L} \), since \( |0a\rangle \) does not commute with \(|+a\rangle \), and \( |0b\rangle \) does
not commute with \(|+b\rangle \). More complicated cases of “codependency” are also possible, as when \( \mathcal{L} \)
contains the product \( ABC \) of three commuting projectors, but none of the six projectors \( A, B, C, \)
\( AB, BC, \) and \( AC \) belong to \( \mathcal{L} \).

### 14.4 Dependent Events in Histories

In precisely the same way that quantum properties can be dependent upon other quantum properties
of a system at a single time, a quantum event—a property of a quantum system at a particular
time—can be dependent upon a quantum event at some different time. That is, in the family of
consistent histories used to describe the time development of a quantum system, it may be the
case that the projector \( B \) for an event at a particular time does not occur by itself in the Boolean
algebra \( \mathcal{L} \) of histories, but is only present if some other event \( A \) at some different time is present
in the same history. Then \( B \) depends on \( A \), or \( A \) is the base of \( B \), using the terminology introduced
earlier. And there are situations in which a third event \( C \) at still another time depends on \( B \), so
that it only makes sense to discuss \( C \) as part of a history in which both \( A \) and \( B \) occur. Sometimes
this contextuality can be removed by refining the history sample space, but in other cases it is
irreducible, either because a refinement is prevented by non-commuting projectors, or because it would result in a violation of consistency conditions.

Families of histories often contain contextual events that depend upon a base that occurs at an earlier time. Such a family is said to show “branch dependence”. A particular case is a family of histories with a single initial state $\Psi_0$. If one uses the Boolean algebra suggested for that case in Sec. 11.5, then all the later events in all the histories of interest are (ultimately) dependent upon the initial event $\Psi_0$. This is because the only history in which the negation $\Psi_0 = I - \Psi_0$ of the initial event occurs is the history $Z$ in (11.14), and in that history only the identity occurs at later times. It may or may not be possible to refine such a family in order to remove some or all of the dependence upon $\Psi_0$.

Figure 14.2: Upper and lower beams emerging from a Stern-Gerlach magnet $SG$. An atom in the lower beam passes through an additional region of uniform magnetic field $M$. The square boxes indicate regions in space, and the time when the atom will pass through a given region is indicated at the bottom of the figure.

An example of branch dependence involving something other than the initial state is shown in Fig. 14.2. A spin-half particle passes through a Stern-Gerlach magnet (Sec. 17.2) and emerges moving at an upwards angle if $S_z = +1/2$, or a downwards angle if $S_z = -1/2$. Let $E$ and $F$ be projectors on two regions in space which include the upward- and downward-moving wave packets at time $t_1$, assuming a state $|\Psi_0\rangle$ (space and spin wave function of the particle) at time $t_0$. In the interval between $t_1$ and $t_2$ the downward-moving wave packet passes through a region $M$ of uniform magnetic field which causes the spin state to rotate by $90^\circ$ from $S_z = -1/2$ to $S_z = +1/2$. This situation can be described using a consistent family whose support is the two histories

$$\Psi_0 \circ E \circ [z^+], \quad \Psi_0 \circ F \circ [x^+]$$

which can also be written in the form

$$\Psi_0 \circ \begin{cases} E \circ [z^+] \\ F \circ [x^+] \end{cases}$$

(14.12)

where the initial element common to both histories is indicated only once. Consistency follows from the fact that the spatial wave functions at the final time $t_2$ have negligible overlap, even though they are not explicitly referred to in (14.12). Whatever may be the zero-weight histories, it is at once evident that neither of the two histories

$$\Psi_0 \circ I \circ [z^+], \quad \Psi_0 \circ I \circ [x^+]$$

(14.13)
can occur in the Boolean algebra, since the projector for the first history in (14.13) does not commute with that for the second history in (14.12), and the second history in (14.13) is incompatible with the first history in (14.12). Consequently, in the consistent family (14.12) \([z^+]\) at \(t_2\) depends upon \(E\) at \(t_1\), and \([x^+]\) at \(t_2\) depends upon \(F\) at \(t_1\). Furthermore, as the necessity for this dependency can be traced to non-commuting projectors, the dependency is irreducible: one cannot get rid of it by refining the consistent family.

An alternative way of thinking about the same gedanken experiment is to note that at \(t_2\) the wave packets do not overlap, so we can find mutually orthogonal projectors \(E\) and \(F\) on non-overlapping regions of space, Fig. 14.2, which include the upward- and downward-moving parts of the wave packet at this time. Consider the consistent family whose support is the two histories

\[
\Psi_0 \otimes I \otimes \{[z^+] \otimes E, [x^+] \otimes F\},
\]

where the notation is a variant of that in (14.12): the two events inside the curly brackets are both at the time \(t_2\), so one history ends with the projector \([z^+] \otimes E\), the other with the projector \([x^+] \otimes F\). Once again, the final spin states \([z^+]\) and \([x^+]\) are dependent events, but now \([z^+]\) depends upon \(E\) and \([x^+]\) upon \(F\), so the bases occur at the same time as the contextual events which depend on them. This is a situation which resembles (14.1), with \(E\) and \(F\) playing the roles of \([z_a^+]\) and \([z_a^-]\), respectively, while the spin projectors in (14.14) correspond to those of the \(b\) particle in (14.1). One could also move the regions \(E\) and \(F\) further to the right in Fig. 14.2, and obtain a family of histories

\[
\Psi_0 \otimes I \otimes \begin{cases} [z^+] \otimes E, \\ [x^+] \otimes F, \end{cases}
\]

for the times \(t_0 < t_1 < t_2 < t_3\), in which \([z^+]\) and \([x^+]\) are dependent on the later events \(E\) and \(F\).

Dependence on later events also arises, for certain families of histories, in the next example we shall consider, which is a variant of the toy model discussed in Sec. 13.5. Figure 14.3 shows a device which is like a Mach-Zehnder interferometer, but the second beam splitter has been replaced by a weakly-coupled measuring device \(M\), with initial ("ready") state \(|M\rangle\). The relevant unitary transformations are

\[
|\Psi_0\rangle = |0a\rangle \otimes |M\rangle \rightarrow (|1c\rangle + |1d\rangle)/\sqrt{2} \otimes |M\rangle
\]

for the time interval \(t_0\) to \(t_1\), and

\[
\begin{align*}
|1c\rangle \otimes |M\rangle &\rightarrow |2f\rangle \otimes (|M\rangle + |M^c\rangle)/\sqrt{2}, \\
|1d\rangle \otimes |M\rangle &\rightarrow |2e\rangle \otimes (|M\rangle + |M^d\rangle)/\sqrt{2}
\end{align*}
\]

for \(t_1\) to \(t_2\). Here \(|0a\rangle\) is a wave packet approaching the beam splitter in channel \(a\) at \(t_0\), \(|1c\rangle\) is a wave packet in the \(c\) arm at time \(t_1\), and so forth. The time \(t_1\) is chosen so that the particle is inside the device, somewhere between the initial beam splitter and the detector \(M\), whereas at \(t_2\) it has emerged in \(e\) or \(f\). The states \(|M\rangle\), \(|M^c\rangle\), and \(|M^d\rangle\) of the detector are mutually orthogonal and normalized. Combining (14.16) and (14.17) yields a unitary time development

\[
|\Psi_0\rangle \rightarrow (|2e\rangle \otimes |M^d\rangle + |2f\rangle \otimes |M^c\rangle + \sqrt{2}|2s\rangle |M\rangle)/2
\]

from \(t_0\) to \(t_2\), where

\[
|2s\rangle = (|2e\rangle + |2f\rangle)/\sqrt{2}
\]
is a superposition state of the final particle wave packets.

Consider the consistent family for $t_0 < t_1 < t_2$ whose support is the three histories

$$\Psi_0 \otimes I \otimes \left\{ [2e] \otimes M^d, [2f] \otimes M^c, [2s] \otimes M \right\}. \quad (14.20)$$

Since the projector $[2s]$ does not commute with the projectors $[2e]$ and $[2f]$, it is clear that $[2e]$, $[2f]$, and $[2s]$ are dependent upon the detector states $M^d$, $M^c$, and $M$ at the (same) time $t_2$, and one has conditional probabilities

$$\Pr(2e \mid M^d_2) = \Pr(2f \mid M^c_2) = \Pr(2s \mid M_2) = 1. \quad (14.21)$$

On the other hand, $\Pr(M^d_2 \mid 2e)$, $\Pr(M^c_2 \mid 2f)$, and $\Pr(M_2 \mid 2s)$ are not defined. (Following our usual practice, $\Psi_0$ is not shown explicitly as one of the conditions.) One could also say that $2e$ and $2f$ are both dependent upon the state $M^t$ with projector $M^c + M^d$, corresponding to the fact that the detector has detected something.

Some understanding of the physical significance of this dependency can be obtained by supposing that later experiments are carried out to confirm (14.21). One can check that the particle emerging from $M$ is in the $e$ channel if the detector state is $M^d$, or in $f$ if the detector is in $M^c$, by placing detectors in the $e$ and $f$ channels. One could also verify that the particle emerges in the superposition state $s$ in a case in which it is not detected (the detector is still in state $M$ at $t_2$) by the strategy of adding two more mirrors to bring the $e$ and $f$ channels back together again at a beam splitter which is followed by detectors. Of course, this last measurement cannot be carried out if there are already detectors in the $e$ and $f$ channels, reflecting the fact that the property $2s$ is incompatible with $2e$ and $2f$. (A similar pair of incompatible measurements is discussed in Sec. 18.4, see Fig. 18.3.)

An alternative consistent family for $t_0 < t_1 < t_2$ has support

$$\Psi_0 \otimes \left\{ [1c] \otimes M^c, [1d] \otimes M^d, [1r] \otimes M \right\}. \quad (14.22)$$

Figure 14.3: Mach-Zehnder interferometer with the second beam splitter replaced by a measuring device $M$. 
where

\[ |1r\rangle = \left( |1c\rangle + |1d\rangle \right) / \sqrt{2} \]  \hspace{1cm} (14.23)

is a superposition state of the particle before it reaches \( M \). From the fact that \(|1r\rangle\) does not commute with \(|1c\rangle\) or \(|1d\rangle\), it is obvious that the particle states at the intermediate time \( t_1 \) in (14.22) must depend upon the later detector states: \(|1c\rangle\) upon \( M^c \), \(|1d\rangle\) upon \( M^d \), and \(|1r\rangle\) upon \( M \). Indeed,

\[ \Pr(1c \mid M_2^c) = \Pr(1d \mid M_2^d) = \Pr(1r \mid M_2) = 1, \]  \hspace{1cm} (14.24)

whereas \( \Pr(M_2^c \mid 1c) \), \( \Pr(M_2^d \mid 1d) \) and \( \Pr(M_2 \mid 1r) \), the conditional probabilities with their arguments in reverse order, are not defined. A very similar dependence upon later events occurs in the family (13.46) associated with weak measurements in the arms of a Mach-Zehnder interferometer, Sec. 13.5.

It may seem odd that earlier contextual events can depend on later events. Does this mean that the future is somehow influencing the past? As already noted in Sec. 14.1, it is important not to confuse the term *depends on*, used to characterize the logical relationship among events in a consistent family, with a notion of *physical influence* or *causality*. The following analogy may be helpful. Think of a historian writing a history of the French revolution. He will not limit himself to the events of the revolution itself, but will try and show that these events were preceded by others which, while their significance may not have been evident at the time, can in retrospect be seen as useful for understanding what happened later. In selecting the type of prior events which enter his account, the historian will use his knowledge of what happened later. It is not a question of later events somehow “causing” the earlier events, at least as causality is ordinarily understood. Instead, those earlier events are introduced into the account which are useful for understanding the later events. While classical histories cannot provide a perfect analogy with quantum histories, this example may help in understanding how the earlier particle states in (14.22) can be said to “depend on” the later states of \( M \) without being “caused by” them.

To be sure, one often encounters quantum descriptions in which earlier events, such as the initial state, are the bases of later dependent events, and it is rather natural in such cases to think of (at least some of) the later events as actually caused by the earlier events. This may be why later contextual events that depend on earlier events somehow seem more intuitively reasonable than the reverse. Nonetheless, the ideas of causation and contextuality are quite distinct, and confusing the two can lead to paradoxes.