Chapter 1

A Brutal Fact of Life

1.1 Causality and determinism

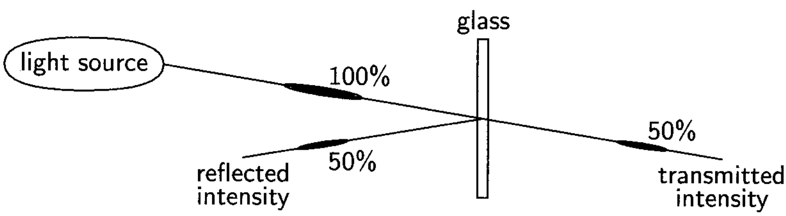
Before their first encounter with the quantum phenomena that govern the realm of atomic physics and sub-atomic physics, students receive a training in classical physics, where Isaac Newton's mechanics of massive bodies and James C. Maxwell's electromagnetism — the physical theory of the electromagnetic field and its relation to the electric charges — give convincingly accurate accounts of the observed phenomena. Indeed, almost all experiences of physical phenomena that we are conscious of without the help of refined instruments fit perfectly into the conceptual and technical framework of these classical theories. It is instructive to recall two characteristic features that are equally possessed by Newton's mechanics and Maxwell's electromagnetism: *Causality* and *Determinism*.

Causality is inference in time: Once you know the state of affairs — physicists prefer to speak more precisely of the "state of the system" — you can predict the state of affairs at any later time, and often also retrodict the state of affairs at earlier times. Having determined the relative positions of the sun, earth and moon and their relative velocities, we can calculate highly precisely when the next lunar eclipse will happen (extreme precision on short time scales and satisfactory precision for long time scales also require good knowledge of the positions and velocities of the other planets and their satellites, but that is a side issue here) or when the last one occurred. Quite similarly, present knowledge of the strength and direction of the electric and magnetic fields together with knowledge about the motion of the electric charges enables us to calculate reliably the electromagnetic field configuration in the future, or the past.

Causality, as we shall see, is also a property of quantal evolution: Given the state of the system now, we can infer the state of the system later (but, typically, not earlier). Such as there are Newton's equation of motion in mechanics, and Maxwell's set of equations for the electromagnetic field, there are also equations of motion in quantum mechanics: Erwin Schrödinger's equation, which is more in the spirit of Maxwell's equations, and Werner Heisenberg's equation, which is more in Newton's tradition.

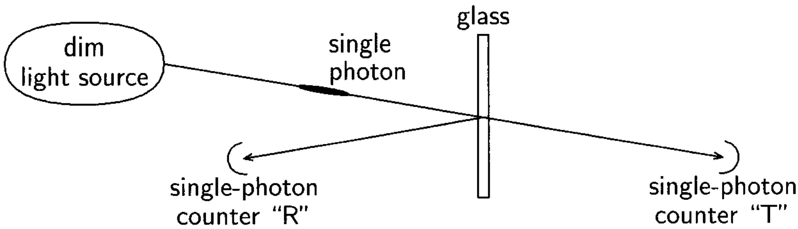
We say that the classical theories are *deterministic* because the state of the system uniquely determines all phenomena. When the positions and velocities of all objects are known in Newton's mechanics, also the results of all possible measurements are predictable, there is no room for any uncertainty in principle. Likewise, once the electromagnetic field is completely specified and the positions and velocities of all charges are known in Maxwell's electromagnetic theory, all possible electromagnetic phenomena are fully predictable.

Let us look at a somewhat familiar situation that illustrates this point and will enable us to establish the difference in situation that we encounter in quantum physics. You have all seen reflections of yourself in the glass of a shopping window, while at the same time having a good view of the goodies for sale. This is a result of the property of the glass sheet that it partly transmits light and partly reflects it. In a laboratory version we could have 50% probability each for transmission and reflection:



A light source emits pulses of light, which are split in two by such a half-transparent mirror, half of the intensity being transmitted, the other half reflected. Given the properties of pulses emitted by the source and the material properties of the glass, we can predict completely how much of the intensity is reflected, how much is transmitted, how the pulse shape is changed, and so forth — all these being implications of Maxwell's equations.

But, we know that there is a different class of phenomena that reveal a certain graininess of light: the pulses consist of individual lumps of energy — "light quanta", or "photons". (We are a bit sloppy with the terminology here, at a more refined level, photons and light quanta are not the same, but that is irrelevant presently.) We become aware of the photons if we dim the light source by so much that there is only a single photon per pulse. We also register the reflected and transmitted light by single-photon counters:



What will be the fate of the next photon to come? Since it cannot split in two, either the photon is transmitted as a whole, or it is reflected as a whole, so that eventually *one* of the counters will register the photon. A single photon, so to say, makes one detector click: either we register a click of detector T or of detector R, but not of both.

What is important here is that we *cannot predict* which detector will click for the next photon, all we know is the history of the clicks of the photons that have already arrived. Perhaps a sequence such as

R R T R T T T R T R

was the case for the last ten photons. In a long sequence, reporting the detector clicks of very many photons, there will be about the same number of T clicks and R clicks, because it remains true that half of the intensity is reflected and half transmitted. On the single-photon level, this becomes a *probabilistic* fact: Each photon has a 50% chance of being reflected and an equal chance of being transmitted. And this is *all* we can say about the future fate of a photon approaching the glass sheet.

So, when repeating the experiment with another set of ten photons, we do not reproduce the above sequence of detector clicks, but rather get another one, perhaps

R T T R R R R T R T

And a third set of ten would give yet another sequence, all 2^{10} possible sequences occurring with the same frequency if we repeat the experiment very often.

Thus, although we know exactly all the properties of the incoming photon, we cannot predict which detector will click. We can only make statistical predictions that answer questions such as "How likely are four Ts and six Rs in the next sequence of ten?"

1-1 Answer this question.

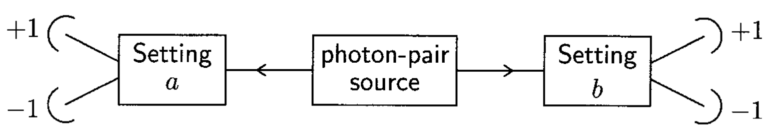
What we face here, in a simple but typical situation, is the *lack of determinism* of quantum phenomena. Complete knowledge of the state of affairs does not enable us to predict the outcomes of all measurements that could be performed on the system. In other words, the state does not determine the phenomena. There is a fundamental element of chance: The laws of nature that connect our knowledge about the state of the system with the observed phenomena are *probabilistic*, not deterministic.

1.2 Bell's inequality: No hidden determinism

Now, that raises the question about the origin of this probabilistic nature. Does the lack of determinism result from a lack of knowledge? Or, put differently, could we know more than we do, and then have determinism reinstalled? The answer is *No*. Even if we know everything that can possibly be known about the photon, we cannot predict its fate.

It is not simple to make this point for the example discussed above with simple photons incident on a half-transparent mirror. In fact, one can construct contrived formalisms in which the photons are equipped with internal clockworks of some sort that determine in a hidden fashion where each photon will go. But in more complicated situations, even the most ingenious deterministic mechanism cannot reproduce the observed facts in all respects. The following argument is a variant of the one given by John S. Bell in the 1960s.

Consider the more general scenario in which a photon-pair source always emits two photons, one going to the left, the other going to the right:



Each photon is detected by one of two detectors eventually — with measurement results +1 or −1 — and the devices allow for a number of parameter settings. We denote by symbol a the collection of parameters on the left, and by b those on the right. Details do not matter, all we need is that different settings are possible, that is: there is a choice between different measurements on both sides. The only restriction we insist upon is that there are only two possible outcomes for each setting, the abstract generalization of "transmission" and "reflection" in the single-photon plus glass sheet example above.

For any given setting, the experimental data is of this kind:

photon pair no.	1	2	3	4	5	6	7	8	...
on the left	+1	+1	−1	−1	+1	−1	+1	+1	...
on the right	+1	−1	−1	+1	−1	−1	−1	+1	...
product	+1	−1	+1	−1	−1	+1	−1	+1	...

The products in the last row distinguish the pairs with the same outcomes on the left and the right (product = +1) from those with opposite outcomes (product = −1). We use these products to define the *Bell correlation* $C(a,b)$ for the chosen setting specified by parameters a, b,

$$C(a,b) = \frac{(\text{number of } +1 \text{ pairs}) - (\text{number of } -1 \text{ pairs})}{\text{total number of observed pairs}}\,. \tag{1.2.1}$$

Clearly, we have $C(a,b) = +1$ if +1 on one side is always matched with a +1 on the other and −1 with −1, and we have $C(a,b) = -1$ if +1 is always paired with −1 and −1 with +1. In all other cases, the value of $C(a,b)$ is between these extrema, so that

$$-1 \le C(a,b) \le +1 \tag{1.2.2}$$

for any setting a, b.

Following Bell, let us now fantasize about a (hidden) mechanism that determines the outcome on each side. We conceive each pair as being characterizable by a set of parameters collectively called λ, and that the source realizes various λ with different relative frequencies. Thus, there is a positive weight function $\rho(\lambda)$, such that $\mathrm{d}\lambda\,\rho(\lambda)$ is the probability of having a λ value within a $\mathrm{d}\lambda$ volume around λ. These probabilities must be positive numbers that sum up to unity,

$$\rho(\lambda) \ge 0, \qquad \int \mathrm{d}\lambda\,\rho(\lambda) = 1\,. \tag{1.2.3}$$

We need not be more specific because further details are irrelevant to the argument — which is, of course, the beauty of it.

We denote by $A_\lambda(a)$ the measurement result on the left for setting a when the hidden control parameter has value λ, and by $B_\lambda(b)$ the corresponding measurement result on the right. Since all measurement results are either $+1$ or -1, we have

$$A_\lambda(a) = \pm 1, \qquad B_\lambda(b) = \pm 1 \qquad \text{for all} \quad a, b, \lambda \tag{1.2.4}$$

and also

$$A_\lambda(a) B_\lambda(b) = \pm 1 \qquad \text{for all} \quad a, b, \lambda\,. \tag{1.2.5}$$

This is then the product to be entered in the table above, before (1.2.1), for the pair that leaves the source with value λ and encounters the settings a and b. Upon summing over all pairs, we get

$$C(a, b) = \int \mathrm{d}\lambda\, \rho(\lambda) A_\lambda(a) B_\lambda(b) \tag{1.2.6}$$

for the Bell correlation, and all the rest follows from this expression.

Before proceeding, however, let us note that an important assumption has entered: We take for granted that the measurement result on the left does not depend on the setting of the apparatus on the right, and vice versa. This is an expression of *locality* as we naturally accept it as a consequence of Albert Einstein's observation that spatially well separated events cannot be connected by any causal links if they are simultaneous in one reference frame. Put differently, if the settings a and b are decided very late, just before the measurements actually take place, any influence of the setting on one side upon the outcome on the other side would be inconsistent with Einsteinian causality. With this justification there is no need to consider the more general possibility of having $A_\lambda(a, b)$ on the left and $B_\lambda(a, b)$ on the right. Such a b dependence of A_λ and an a dependence of B_λ are physically unacceptable, but of course it remains a mathematical possibility that cannot be excluded on purely logical grounds.

All together we now consider two settings on the left, a and a', and two on the right, b and b'. The difference between the Bell correlations for settings a, b and a, b' is then

$$\begin{aligned} C(a,b) - C(a,b') &= \int \mathrm{d}\lambda\, \rho(\lambda)[A_\lambda(a)B_\lambda(b) - A_\lambda(a)B_\lambda(b')] \\ &= \int \mathrm{d}\lambda\, \rho(\lambda)A_\lambda(a)B_\lambda(b)[1 \pm A_\lambda(a')B_\lambda(b')] \\ &\quad - \int \mathrm{d}\lambda\, \rho(\lambda)A_\lambda(a)B_\lambda(b')[1 \pm A_\lambda(a')B_\lambda(b)] \end{aligned} \tag{1.2.7}$$

where the "$\pm$" terms compensate for each other provided we take either both upper signs or both lower signs. Now, since

$$\rho(\lambda) \geq 0, \qquad |A_\lambda(a)B_\lambda(b)| = 1, \qquad 1 \pm A_\lambda(a)B_\lambda(b) \geq 0 \tag{1.2.8}$$

for both signs and all values of λ, a, b, repeated applications of the triangle inequality gives

$$\begin{aligned} |C(a,b) - C(a,b')| &\leq \int \mathrm{d}\lambda\, \rho(\lambda)[1 \pm A(a')B(b)] \\ &\quad + \int \mathrm{d}\lambda\, \rho(\lambda)[1 \pm A(a')B(b')] \\ &= 2 \pm [C(a',b) + C(a',b')]\,. \end{aligned} \tag{1.2.9}$$

Consequently, the left-hand side cannot exceed the smaller one of the two right-hand sides (one for $+$ and one for $-$ in $\pm$),

$$|C(a,b) - C(a,b')| \leq 2 - |C(a',b) + C(a',b')| \tag{1.2.10}$$

or, after rearranging,

$$|C(a,b) - C(a,b')| + |C(a',b) + C(a',b')| \leq 2\,. \tag{1.2.11}$$

This is (a variant of) the so-called *Bell inequality*. Given the very simple argument and the seemingly self-evident assumptions entering at various stages, one should confidently expect that it is generally obeyed by the correlations observed in any experiment of the kind described on page 4. Anything else would defy common sense, would it not? But the fact is that rather strong violations *are* observed in real-life experiments in which the left-hand side substantially exceeds 2, getting very close indeed to $2\sqrt{2}$, the maximal value allowed for Bell correlations in quantum mechanics (see Section 2.18).

Since we cannot possibly give up our convictions about locality, and thus about Einsteinian causality, the logical conclusion must be that there just is no such hidden deterministic mechanism [characterized by $\rho(\lambda)$ as well as $A_\lambda(a)$ and $B_\lambda(b)$]. We repeat:

> There is no mechanism that decides the outcome of a quantum measurement.

What is true for such correlated pairs of photons is, by inference, also true for individual photons. There is no mechanism that decides whether the photon is transmitted or reflected by the glass sheet, it is rather a *truly probabilistic* phenomenon.

This is a brutal fact of life. In a very profound sense, quantum mechanics is about learning to live with it.

1.3 Remarks on terminology

We noted the fundamental lack of determinism at the level of quantum phenomena and the consequent inability to predict the outcome of all experiments that could be performed. It may be worth emphasizing that this lack of predictive power is of a very different kind than, say, the impossibility of forecasting next year's weather.

The latter is a manifestation of the chaotic features of the underlying dynamics, frequently referred to as *deterministic chaos*. In this context, "deterministic" means that the equations of motion are differential equations that have a unique solution for given initial values — the property that we called "causal" above. There is a clash of terminology here if one wishes to diagnose one.

The deterministic chaos comes about because the solutions of the equations of motion depend extremely sensitively on the initial values, which in turn are never known with utter precision. This sensitivity is a generic feature of nonlinear equations, and not restricted to classical phenomena. Heisenberg's equations of motion of an interacting quantum system are just as nonlinear as Newton's equations for the corresponding classical system if there is one.

In classical systems that exhibit deterministic chaos, our inability to make reliable predictions concerns phenomena that are sufficiently far away in the future — a weather forecast for the next three minutes is not such a challenge. In the realm of quantum physics, however, the lack of determinism is independent of the time elapsed since the initial conditions were established. Even perfect knowledge of the state of affairs immediately before a measurement is taken does not enable us to predict the outcome; at best we can make a probabilistic prediction, a statistical prediction.

Of course, matters tend to be worse whenever one extrapolates from the present situation, which is perhaps known with satisfactory precision, to future situations. The knowledge may not be accurate enough for a long-term extrapolation. In addition to the fundamental lack of determinism in quantum physics — the nondeterministic link between the state and the phenomena — there is then a classical-type vagueness of the probabilistic predictions, rather similar to the situation of classical deterministic chaos.