3
Quantum Mechanics. Mixing

3.1 Introduction

The mechanics of elementary particles is different from that of classical objects such as tennis balls, or planets, or missiles. The movements of these are well described by Newton's laws of motion. The laws describing the motion of elementary particles are given by quantum mechanics. The laws of quantum mechanics are quite different from Newton's laws of motion; yet if a particle is sufficiently heavy the results of quantum mechanics are very close to those of Newtonian mechanics. So in some approximation elementary particles also behave much like classical objects, and for many purposes one may discuss their motion in this way. Nevertheless, there are very significant differences and it is necessary to have some feeling for these.

There are two concepts that must be discussed here. The first is that in quantum mechanics one can never really compute the trajectory of a particle such as one would do for a cannon ball; one must deal instead with probabilities. A trajectory becomes something that a particle may follow with a certain probability. And even that is too much: it is never really possible to follow a particle instant by instant (like one could follow a cannonball as it shoots through the air), all you can do is set it off and try to estimate where it will go to. The place where it will go to cannot be computed precisely; all one can do is compute a probability of where it will go, and then there may be some places where the probability of arrival is the highest. This must be explained, and it

Werner Heisenberg (1901–1976) published his paper introducing quantum mechanics in 1925. The unfamiliar mathematics (matrix calculus) made the paper difficult to read. In 1927 he published his famous uncertainty relations. He made further fundamental contributions to particle physics, for example he recognized that strong interactions are the same for proton and neutron and he found the correct mathematical way to formulate this. He really is one of the all-time greats of physics. In 1932 he received the Nobel prize.

His attitude towards the Nazi regime during World War II may be called ambiguous at best. During the war he was involved in a program aimed at studying uranium fission, but this did not lead to a nuclear bomb. Part of this failure was perhaps due to his poor experimental capabilities for which we may then be thankful.

After World War II Heisenberg was instrumental in the creation of the Max Planck Society with its series of Max Planck institutes. This method of creating centers of excellence has been very fruitful.

In his later years he tried to develop a "theory of everything". It was neither impressive nor successful, and in fact led to rather acerbic comments of Pauli, initially his collaborator.

Erwin Schrödinger (1887–1961) introduced his version of quantum mechanics in 1926. He formulated a wave theory for particles which to this day is the easiest and most often used tool for the quantum mechanical treatment of atoms and molecules. His fundamental equation, the Schrödinger equation, is valid only if the particles involved are not relativistic (speed much less than that of light), which is true for electrons in atoms and atoms in molecules. He received the Nobel prize in 1933.

Schrödinger conceived his ideas during an erotic outburst, spending a holiday in Arosa in Switzerland with an unknown lady. This escapade had apparently an enormous influence on his scientific creativity that for about 12 months remained at a stratospheric level. His life involved many women; his wife Anny maintained a (amorous) relationship with the famous mathematician Hermann Weyl.

The later part of his life, after 1939, was spent at the newly founded Institute for Advanced Studies in Dublin. Remarkably, there appeared to be little problem in this catholic country for him to live there with two women, his wife Anny and Mrs Hilde March (mother of his daughter Ruth).

is done using light as an example, which in pre-quantum physics is described quite accurately by electromagnetic waves. This must be re-examined with the knowledge that light consists of particles, the photons.

The second concept that must be introduced is the idea of an amplitude, a quantity that must be squared to obtain physical statements. That also may be understood by considering light.

3.2 The Two-Slit Experiment

Light, which we know to be nothing but electromagnetic fields, is well described by waves. This was first proposed by Huygens, while in that same period Newton advocated the particle idea. It would have been interesting to go back in time and organize a meeting with these two scientists. One can imagine them looking at a visitor from the future who knows all the answers. Thus, first question by Newton (or Huygens):

What is light: waves or particles?

The answer:

uuuh uuuh .... both.

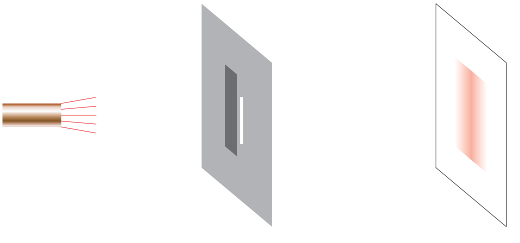
Probably Newton and Huygens would not be amused; one would have a hard time answering them, which would amount to teaching them quantum mechanics.

If one would want to give an answer that would be a bit more precise one could say the following. The trajectory that a particle is going to follow can approximately be found by doing a calculation with waves. That is what it is, a calculation. It is not true that the particle “is” a wave. It is just that to calculate where it goes one uses wave theory. That is the theory describing its motion. It is not the theory describing the particle itself. The particle remains a small, for all we know point-like, object of definite mass (the mass is zero for light). So the correct answer could have been: light is particles, but their laws of motion are those of waves.

So, light can be described by waves, like also sound is described by waves. Waves can interfere, and the classic experiment to see that is the two-slit experiment. The figure below shows the experimental set-up: a source shines light of a specific color onto a surface containing two openings, two slits. Laser light is excellently suited to this purpose. Further down there is a screen catching the light that passes through the slits. The fact that the light is of a specific color means that the frequency of the light is sharply defined, and hence all photons emitted from the source have the same energy, as given by the Einstein-Planck relation $E = h\nu$.

To avoid all possible confusion in the argument the source of light is supposed to be so weak that only one photon leaves the source every minute. Thus whenever a photon leaves the source the previous one has since a long time (for a photon) hit the screen. This very slow rate is to make sure that different photons in the beam do not bother each other. It is strictly a single photon experiment.

First one of the slits (call it the first) is kept open, the other is closed.



When the first photon passes through the open slit it will hit the screen somewhere, at a more or less unpredictable place. But sending on photons for hours a pattern develops: the photons will hit the screen in some area that is a widened, blurred image of the slit (the blurring is substantial only if the slit is not too wide). This is understood as diffraction (scattering) of the waves by the edges of the slit. If one knows diffraction theory the image can be computed accurately; the picture on the screen that is built up from individual photon hits will slowly fill out to a picture computed using wave diffraction theory. You may have to wait a few weeks at the rate mentioned, but that is what will happen.

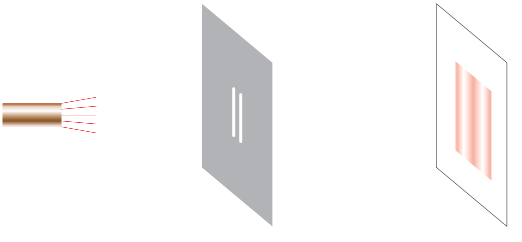
What can we learn from this curious behaviour? First, what is the meaning of the intensity of light for the case of particles? The answer seems obvious: the intensity is proportional to the number of photons. There where the light is intense there are many photons. That is also in line with the idea that the intensity of the light gives the energy density, since a photon has a definite energy. Now the photons are going to make a pattern. There will be many that hit the center, and less towards the edge of the image of the slit. Since the photons arrive one by one there is only one way to interpret this: the pattern on the screen describes the relative probability for the photons to hit the screen at some location. That probability is high where the picture is bright, lower towards the edges. Thus here is the idea: compute what light will do using the theory for the propagation of waves. This gives a pattern, a picture on the screen. That picture represents then the relative probabilities for the photons to hit the screen here or there.

This in a nutshell is quantum mechanics. Since the behaviour of waves is vastly different from classical trajectories of material objects it is not surprising that many have difficulties accepting these ideas. But in the end it is really not that complicated: use waves to compute patterns and that will then give us the probabilities for finding particles here or there.

It is when one tries to explicitly follow how a particle moves, from the source of light, through the slit to the screen that things become difficult. Since it is not the purpose of this book to create difficulties we will not occupy ourselves with questions concerning the whereabouts of the photon on its trajectory from source to screen. It is daydreaming. What counts is what you see on the screen. Do not ask if the particle did follow some continuous path. We do not know about that. Forget about it. For all we care the particle just skips the distance all together and will just hit the screen at some place with a certain probability. We have absolutely no idea if it ever passed through the slit, we never will have, and it cannot be established by any method. The only thing that we can do is compute the probability where it will hit the screen.

What happens if the experiment is repeated with the first slit closed and the other slit open? That is simple: exactly the same pattern will be observed except slightly displaced, because the second slit is slightly displaced relative to the first one.

Now open both slits. The naive person, assuming photons passing through the slits as particles, would say that the new pattern is simply the sum of the two, but that is not the case. There is interference, i.e. there are places where the waves from the first slit cancel out those of the second, and other places where they enhance each other. Using wave theory there is really no problem computing that. In the old days this constituted a convincing proof that Huygens was right and Newton wrong. It just goes to show how careful one must be.



How must this interference be understood? Well, there is nothing special. Compute the pattern to be expected using wave theory and that gives then the probability distribution for the photons such that precisely that pattern comes out in due time. That is the way it is. Individual particles move in unpredictable ways, but in the end, looking at many particles, a pattern forms, of which we can predict the precise form. It is like a roulette wheel: you never know (if you are in an honest place) where the ball will stop, but if you wait long enough it will distribute evenly over all holes. And even if the wheel was loaded there would be a pattern, peaking at some selected places.

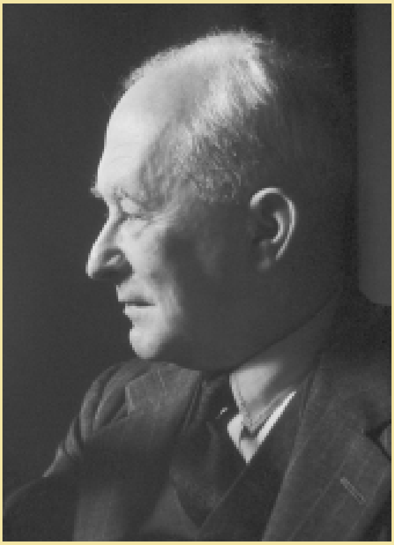
Remember now that the experiment was done with the photons strictly separated in time. To say it crudely: they do not interfere with each other, they interfere with themselves. An individual photon moves in a way such that the probability of arrival at some place includes the effect of interference. Of course, the idea of a material particle interfering with itself is quite lunatic, and you will save yourself a lot of headache not trying to visualize that. The interference is in the calculation trying to establish where the photon will go, or rather trying to compute the probability for arriving at a certain place.

3.3 Amplitude and Probability

There is an important consequence to draw from the two-slit experiment. In the calculation one uses waves, coming from both slits and canceling or amplifying each other. Waves may have a sign — think for example of waves on a water surface. Part of a wave is above the average surface (the surface if there was no wave), part is below. When two waves meet there will be interference: the result is that at certain places the water wave will move even more above or below the average surface, while at other places the waves may cancel each other. Now a probability is always positive and not larger than one; a negative probability or a probability larger than one is like saying that you are −20% or more than 100% sure of something. You cannot be less than 0% sure of something. That means already totally unsure. And you cannot be more sure than 100%.

The intensity of the wave is related to the amount the wave goes up or down, either plus or minus. The maximal deviation of the wave from the average (the deviation when on the top or in the valley) is called the amplitude. The intensity is given by the square of the amplitude of the wave, a fact which must now be made plausible.

Consider an idealized situation, where the images of both slits overlap. Then they will enhance each other in the middle,

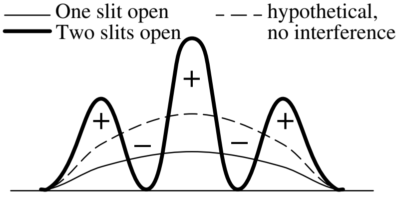


Max Born (1882–1970). While much less known to the general public than Heisenberg, Dirac or Schrödinger, Born must nonetheless be included as one of the founders of quantum mechanics. He was the one that made the link between the mathematics and physical observation by defining how probability relates to the wave function. That is, he found out that probability is obtained by squaring the amplitude. He received the Nobel prize twenty five years after that work, in 1954.

Born got into discussion with Einstein who refused to accept probability as a fundamental property of physics. It is in a letter to Born that Einstein wrote in 1926 his famous sentence: "Quantum mechanics is very impressive. But an inner voice tells me that it is not yet the real thing. The theory produces a good deal but hardly brings us closer to the secret of the Old One. I am at all events convinced that He does not play dice".

It is amusing to see that Einstein in fact admits that he has no hard arguments against quantum mechanics. He just does not want it. It may have been that he felt that there is something contradictory between quantum mechanics and his theory of gravitation. To this day there is a mystery there, and we do not have a good theory of quantum gravitation. For instance, black holes defy the basic laws of quantum mechanics, and no one has come up with a convincing way to handle that. What to do: disbelieve black holes or quantum mechanics?

interfere out a bit away from the middle, and further on again amplify each other, etc. In the figure above we tried to illustrate that. In the figure below the bold line shows how the intensity varies going through the area horizontally.



Very, very crudely this is what happens. If only one slit is open there will be some limited area where the light will hit. In the figure the thin line shows the intensity of the light on the screen for this case. If the other slit is open (and the first closed) the same result will be obtained (never mind the slight shift because the slits are slightly displaced with respect to each other). Now have both slits open. There will be light only in the same area as before. However, half the time the waves will compensate, the other half of the time they will enhance. Let us now consider the energy distribution in precise detail. As every photon carries a definite amount of energy that is also the distribution of the photons.

If there is only one slit open the smooth curve drawn with a thin line applies. The maximum amount of energy will be deposited in the centre tapering off towards the sides. The total amount of energy (the total number of photons) is proportional to the surface below that curve.

Open now both slits. If there were no interference then the hypothetical curve (the dashed curve) would apply. The total amount of energy is simply doubled, the surface below the dashed curve is twice the surface below the thin line curve. All that changes is that we get twice as many photons everywhere. In the centre the dashed curve is twice as high as the thin curve. There would be twice as many photons in the centre.

However, there is interference. The total number of photons will still be the same, twice as much as with only one slit open. However, their distribution is changed drastically. In the centre the light waves enhance each other, while slightly off centre they interfere destructively. Photons that (in the hypothetical case) went to the locations slightly off the centre now arrive in the centre. This is indicated by the + and – signs in the drawing. Extra photons in the central region have been taken from the off-centre regions. As a consequence there are twice as many photons in the central area as compared with the no interference case. That is four times as much as with one slit open. In the central region the bold curve is four times higher than the thin curve.

At this point consider the amplitudes of the light waves. There will be a certain amplitude in the centre if there is only one slit open. If there are two slits open, the waves arriving in the centre amplify each other and the result will be a wave with an amplitude twice as large. Think of waves on water. At the top of a wave the water particles are moved upwards by a certain amount. When two equal waves meet, and they are in phase (the tops coincide), the second wave will move the particles up by that same amount, so that all together the wave rises twice as high. Thus, comparing the one slit case with the two-slit case the amplitude in the centre will be twice as large for the two slits case. However, as argued above, the number of photons arriving in the centre is four times as much. What one sees is that if the amplitude doubles the number of photons quadruples. **The number of photons is proportional to the square of the amplitude**. This also cures the problem of a negative sign; even if the amplitude is negative, the probability, related to the square of the amplitude, will be positive.

The total amount of energy deposited on the screen does not change if there is interference. The distribution changes, but whatever there is extra in the centre has been taken away from the neighbouring regions.

A warning here: one must be very careful with arguments of the type given here. Interference is a complex phenomenon.

In quantum mechanics there is in this context a very important point: conservation of probability. The theory must be such that the total probability of a given particle to arrive somewhere on the screen should be 100%. It should go somewhere, and not disappear halfway, and all probabilities should add up to 100% according to the rule that if it does not hit here it must hit somewhere else. If the particle is unstable and if it can decay on its way to the screen then these decay configurations must be included in the total probability count: the probability of arriving at the screen and the probability of decaying somewhere in between should together add up to one.

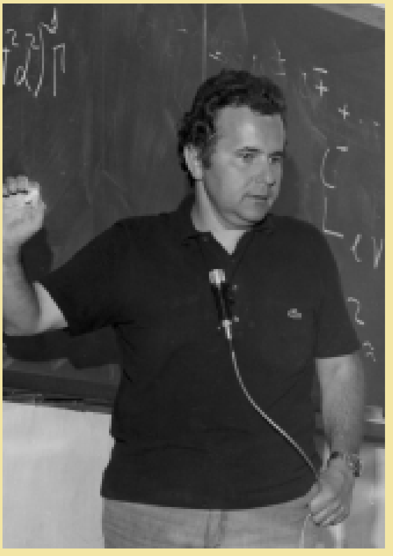
So here is the result: when there is more than one possible trajectory for a particle there is a wave of some amplitude associated with each of the possibilities. These waves must be superposed (which means addition or subtraction or something in between) producing a wave with another amplitude. The resultant amplitude must be squared and that gives the probability distribution.

Sometimes one reads about machines that create silence. This is the idea: if there is some noisy area (near a highway for example) then set up a speaker system that produces precisely the same sounds but in such a way as to cancel the original sounds. However, remember that energy must go somewhere. If there is somewhere a point where the waves interfere to zero then there is somwhere else another point where they amplify each other. It is not really possible to make a silencing machine. In the end you just add more noise, slightly differently distributed, depending on the wavelength of the sound waves. If you want to cancel out sounds it is necessary to go back to the source and create a situation where then no energy will be released.

In particle physics it is possible to have two amplitudes that cancel each other completely. One must always consider a process as a whole; if two amplitudes cancel completely then nothing can be emitted. For sound there is an explicit example of that, low frequency sound emitted by a loudspeaker not encased in a box. The sound waves emitted by the back of the speaker may go around and come out front, where they then interfere destructively with the waves produced by the front of the speaker cone. You will hear nothing. It becomes impossible to pump energy into the speaker. The cone will flop back and forth without giving off any substantial amount of energy to the surrounding air. Some air moves forth and back from the front to the back of the speaker. Thus some energy is pumped into the movement of the cone itself and in the movement of the thin layer of air around the speaker, but it is a minor amount. It is essential that the waves have low frequency (large wavelength), so that sound coming from the back, having to travel some distance, remains still out of phase with the waves from the front. So the effect disappears for wavelengths smaller than the diameter of the speaker. That is why speakers are put in boxes: to absorb the low frequency sound produced on the backside. You can also put the speaker at a hole in a soundproof wall. That gives quite a good reproduction even of low frequency sounds on both sides of the wall.

The feature that the energy in a wave is proportional to the square of the amplitude is quite universal. If the cone of the loudspeaker moves in and out twice as much (compared to some initial case) the energy emitted is four times as much. This is not an intuitively appealing result, but that is the way it is. You can easily confuse yourself by playing in your mind with speakers and imagining what they do. Do not forget that the sound of one of the speakers may reach the cone of the other speaker and so influence the movement of that cone. It tends to become complicated. Speaker technology is a complicated issue. Remember that above we were talking about monochromatic light. To have sound amplitudes of two speakers sum up the waves must also be monochromatic, that is of the same frequency. And then there is interference and arguments as given above apply.

Another example can be found in electricity (for those who know about circuits). If there is a current going through some circuit then the energy absorbed per second is the wattage, which is the product of voltage and current. The current itself



Nicola Cabibbo (1935). In 1963 the situation in particle physics was very confusing. There were many particles (now understood as bound states of quarks) that were unstable and decayed in a multitude of modes and strengths.

In a footnote in an article by Gell-Mann and Levy the idea of a fixed ratio between certain decay modes was mentioned. Moreover this was cast in the form of an angle, but no attempt was made to implement this idea. There is actually more to it than just an angle, but never mind.

It was Cabibbo's merit that he succeeded in implementing a complete scheme describing the relative strengths of many decay modes. For example, the angle could be fixed by considering the ratio of pion and kaon decay ($\rightarrow$ muon + antineutrino). Given then the angle he could precisely compute the decay of the muon ($\rightarrow$ electron + neutrino + antineutrino) from neutron decay. Many people including this author puzzled about these reactions; Cabibbo was at that time working in an office at CERN next to mine and at one point told me that he now understood the relation between neutron and muon decay. He said to me mysteriously, "it is an angle." I said: "Ha ha, I suppose we should call it the Cabibbo angle." The joke was in the end not funny. We now speak indeed of the Cabibbo angle.

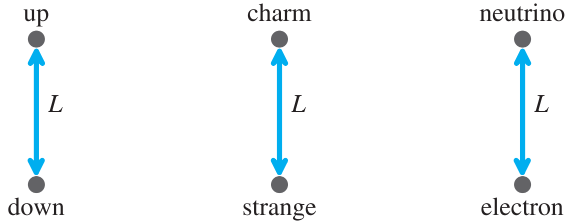
It was a revolution that brought order in a very confusing situation, and was of fundamental importance with respect to the further development of particle theory.

is proportional to the voltage, and therefore the energy absorbed is proportional to the square of the voltage (or the square of the current, make your choice). This works actually also for the speakers mentioned. The deviation of the speaker cone is proportional to the current that flows through the speaker coil, and the energy delivered is proportional to the square of the current (the energy is equal to I^2R where I is the current and R is the impedance of the speaker). For a speaker not in a box the impedance is for low frequencies largely inductive and no energy is absorbed by the speaker. On has then a situation analogous to a coil without any cone attached, moving freely in the magnetic field inside the speaker without absorbing any energy. A good speaker system behaves as a pure resistor all through the frequency spectrum.

3.4 Cabibbo and CKM Mixing

Now back to the particle families and their interactions with the three vector bosons, W^-, W^+ and Z^0. There is a small complication, yet with important consequences. First the difference between transition strength and coupling constant, mentioned before, must be emphasized. The coupling constant g involved in the up $\rightarrow$ down + W^+ transition has a certain magnitude. The transition strength, i.e. the transition probability for this reaction, which is what can be observed experimentally, is proportional to α_w which can be obtained by squaring g (and dividing by 4π). In other words, the coupling constant may have a sign (as for example, the electric charge of a particle has a sign), but the transition probability, being proportional to the square of the amplitude and hence to the square of the coupling constant, is of course always positive. In fact, this is basically the same squaring as mentioned in the previous section. The amplitude of the wave corresponding to the particle emitted (the W^+ etc.) is proportional to the coupling constant, and the probability is the square of that. That is not different from the emission of a photon by a charged particle: the electromagnetic field emitted is proportional to the charge of the particle (the coupling constant) and the probability will be proportional to the square of that.

Suppose for the moment that there are only two families, the up-down and charm-strange families. Consider the transitions among the quarks caused by the charged spin 1 particles, W^+ and W^-. These transitions specified above would be strictly a "family business", but the actual situation is different. Earlier it was stated that the up quark can become a down quark, emitting a W^+, and the charmed quark can become a strange quark, emitting a W^+. The negative vector-boson W^- is involved in the opposite transition, like down $\rightarrow$ up + W^-. The strength of these transitions is the same as among the leptons, like for example neutrino $\rightarrow$ electron + W^+. In other words, the coupling constant for all these couplings is the same, denoted above by g. This coupling constant universality is an important property that plays a large role in theoretical considerations. The figure below shows the transition amplitudes; they have magnitude L and they are proportional to the universal coupling constant g. The transition probability L^2 is proportional to $\alpha_w = g^2/4\pi$.



In actual fact the quark transitions are slightly rotated with respect to the family structure. One has that

$$\text{up} \rightarrow \text{down} + W^+$$

goes with a probability slightly less than the lepton transition

$$\text{neutrino} \rightarrow \text{electron} + W^+$$

but the difference equals the probability of a new transition,

$$\text{up} \to \text{strange} + W^+$$

We ignored energy considerations which actually forbid the reactions as shown. For example, a massless neutrino cannot decay into an electron and a W^+. However, reactions derived from the above by crossing may be possible. Thus the sum of the transition probabilities of the actually observable processes

$$W^- \to \text{anti-up} + \text{down} \quad \text{and} \quad W^- \to \text{anti-up} + \text{strange}$$

is equal to the transition probability of

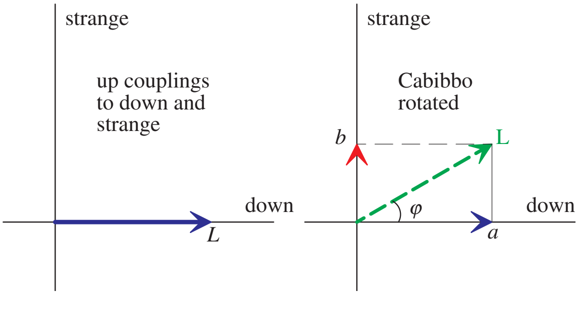
$$W^- \to \text{antineutrino} + \text{electron}$$

Similarly the sum of the transition probabilities for

$$\text{charm} \to \text{strange} + W^+ \quad \text{and} \quad \text{charm} \to \text{down} + W^+$$

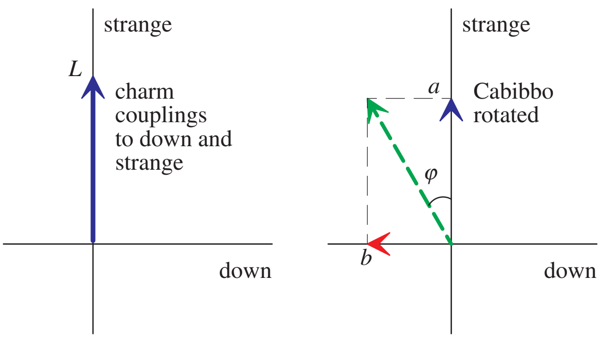
is equal to the leptonic transition probability ($\nu \to e^- + W^+$). This whole affair can be viewed as a rotation of the quantity L over an angle, the Cabibbo angle. To explain this consider the figure below, the left part.

The bold line represents the coupling constant for the coupling of the up quark to down and strange quark (plus emission of a W^+). The projection of the bold line on the horizontal axis gives the amount for the down quark coupling, the projection on the vertical axis similarly gives the coupling to the strange quark. As the fat line is horizontal, the coupling to the strange quark is zero.



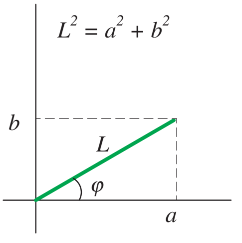
Now rotate the bold line by an angle φ. That rotation is the "Cabibbo rotation". The horizontal projection (indicated by a) is slightly less than in the original figure (where it was equal to L), while there is now a non-zero value for the up to strange transition (b).

A similar situation holds for the coupling of the charmed quark to down and strange quark ($+W^+$). This is shown in the two figures below. Originally there is no charm to down coupling (the bold line is strictly vertical, no projection on the horizontal axis), after rotation over the same Cabibbo angle there is an amount b for that transition, while the transition to a strange quark is slightly diminished from L to the value a.

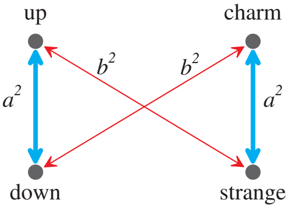


The experimentally determined value for the Cabibbo angle is about 12.7 degrees. The idea of an angle, implying that the probability of some reaction diminishes but that a new reaction takes that up, has been a very fruitful one. At once a lot of poorly understood experimental data started to make sense. In 1963 it was seen that neutron decay (due to the decay $d \rightarrow u$ + electron + antineutrino) proceeded with a coupling constant that was slightly less than that for muon decay ($\mu \rightarrow \nu_\mu$ + electron + antineutrino). The Cabibbo theory explained that, in perfect agreement with experiment.

Now the question of total probability. It is a property of rotations that the sum of the squares of the components remains the same: the total probability is unchanged. This is a consequence of the well-known theorem of Pythagoras.

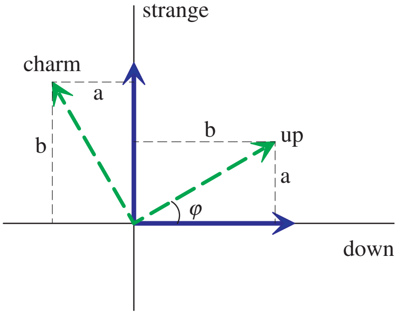


Consider a stick of a certain length. In the figure it is the bold line of length L. From the projections along mutually perpendicular directions the length of the stick can be obtained by using the Pythagorean equation. The sum of the squares of the projections must be calculated, and the length is the square root of that sum. This length L, the length of the stick, is always the same, independent of the angle of rotation, denoted by φ in the picture, and it is directly related to the sum of the squares of the individual components.



The figure summarizes explicitly the effect of the Cabibbo rotation. Before rotation the transition probability up $\rightarrow$ down is L^2 (with L equal to that found in muon decay). After the rotation the transition probability of the up quark to go to a down quark is a^2 and to the strange quark b^2, with the sum remaining the same: $a^2 + b^2 = L^2$. Similarly for the charmed quark. The attentive reader may note that in the figure there are arrows on both ends of the lines. This is done to include also the inverse transitions, such as down $\to$ up + W^-. It makes the figure inversion invariant, that is, if you turn it upside down it looks the same. The Cabibbo rotation can equally well be discussed considering these inverse reactions.

The rotation may be visualized in a figure, see below. Originally there are two bold lines of equal lengths orthogonal to each other. The Cabibbo rotation rotates these bold lines to the dashed ones. The projections from the dashed line marked with up gives the transitions from the up quark to down and strange quark, and similarly for the charm quark, represented by the dashed line marked charm.



The Cabibbo rotation is experimentally well established, but its origin remains a mystery. The value of the Cabibbo angle, 12.7 degrees, is another number for which we have no explanation, just like for the masses of the various particles. Theoretically there is a relationship to the Higgs particle, but that relationship clarifies nothing. Once more one might hope to understand more if this Higgs particle shows up in the detectors at future machines.

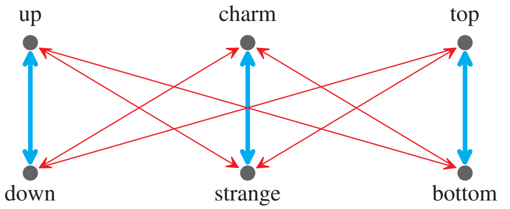
The actual situation is even more complicated because there are three families. There are many more transitions, shown in the figure below. It requires a lot of experimental effort to measure all these transitions and that work is far from completed. Also in this figure we again included the inverse transitions, by attaching arrows to both ends, making the figure invariant under inversion, i.e. turning it upside down.

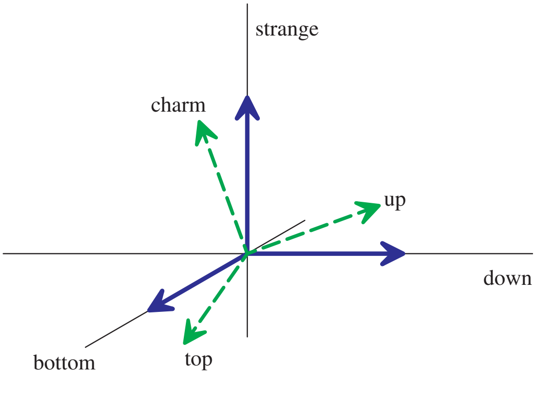
Makoto Kobayashi (1944) and **Toshihide Maskawa** (1940). In 1973 Kobayashi and Maskawa extended the Cabibbo idea of mixing to three families. At the time there was not even suspicion for the existence of a third family; they did it because in the case of two families they did not have the freedom to accommodate certain data. This concerns the imaginary part of a coupling constant, observed experimentally through the existence of certain reactions. The subject is not discussed here simply because it would require a lot of elaboration.

Anyway, Kobayashi and Maskawa saw that having only two families resulted in a scheme that was too narrow to accommodate all experimental data. In a bold move they assumed the existence of a third family yet to be discovered. In the mood of those days suggesting the existence of a new particle was just "not done". Today many irresponsible people do it. The tau, discovered by Martin Perl in 1975, was the first member of the third family observed experimentally, and gradually the rest of the family was discovered, with at last the top quark being established in 1995.

The story is not finished. A considerable amount of experimental effort is being made to measure and understand that complex coupling constant. At SLAC the B-factory (an accelerator producing lots of bound states of the bottom quark) is at this time running very satisfactorily, giving new information on the subject. The mystery of the complex coupling constant relates to the Higgs particle. Again!



The rotation involves now another axis, the bottom axis. So the figure showing the rotation has become three-dimensional. The next figure is an attempt to visualize this. The bottom axis is assumed out of the paper. The rotation becomes much more complicated: the charm axis moves to the left and slightly forward (out of the paper), and then there is yet another rotation in the up-top plane.



The projections of the bold dashed lines marked up, charm and top onto the third axis (the one sticking out of the paper) give the strengths of the transitions of the up, charm and top quark to the bottom quark. This generalization of the Cabibbo rotation to a rotation of three mutually perpendicular (bold) lines was done by two Japanese physicists, Kobayashi and Maskawa; one hence speaks of the CKM rotation. The remarkable thing is that they did this even without knowing about the third family! They anticipated the existence of the third family on the basis of certain esoteric arguments. We shall discuss some aspects of the CKM rotation. There is no real need to delve into it here, but the facts must be mentioned. You may skip the next two paragraphs.

A rotation in three dimensions is described by three angles: the charm axis is rotated to the left, then rotated forward and finally there is a rotation of the up-top plane, keeping the charm axis fixed. Thus the CKM rotation involves three angles, one of which is the Cabibbo angle. Now here comes something which is truly a matter of quantum mechanics. In quantum mechanics the coupling constants are not just positive or negative, they can be complex, having an imaginary part. If you do not know what complex and imaginary means then that is just too bad, there is really no easy way to explain it. The closest analogy comes from AC electric currents. For an AC current positive or negative is meaningless (except momentarily), but if one considers two currents one can compare them. They can be in the same direction or opposite, but more generally may have a certain phase with respect to each other. Such a phase may be represented by a complex number.

In practice, for the CKM rotation, this means that there is a fourth "angle", and it determines the relative phase of the coupling constants. This angle can be measured in a quite distinct way, it is related to what is called CP violation. But it is outside the scope of this book to explain that in detail.

At this point one may ask if there is also a rotation among the Z^0 couplings analogous to that among the W couplings. In the description given earlier the Z^0 coupling to the down quark for example is

$$\text{down} \rightarrow \text{down} + Z^0$$

One could image that there is mixing between the families, meaning that there could be a coupling such as

$$\text{down} \rightarrow \text{strange} + Z^0$$

However, this is not the case. Such transitions are experimentally seen to be absent to a very high degree, a fact which caused quite some confusion among theorists. Theoretically this is now understood, but explaining that is again outside the scope of this book. Physicists have a way of speaking about the absence of this last reaction: "the absence of neutral strangeness-changing currents". The word neutral refers to the charge of the Z^0. The change of the down quark to the strange quark is referred to as a change of strangeness. The word current refers to the detailed way the Z^0 is coupled to the quarks.

3.5 Neutrino Mixing

One may ask: why is there no mixing among the leptons of the three generations? The answer to that is that we do not know whether there is or not. From the theory it is known that this mixing becomes unobservable if the neutrino masses are all zero. So far the measurements only provide us with upper limits for these masses, and the theory has nothing to say about their possible values. But if the neutrino masses were non-zero there could be something like CKM mixing for neutrinos, and these days a quite large amount of experimental effort is directed towards investigating this issue. Here follows a very simplified discussion.

Consider a solar process involving the emission of a neutrino. That is always due to a transition of the type

$$\text{electron} \rightarrow \text{neutrino} + W^-$$

and crossed versions of this.

Solar neutrino experiments are designed to detect the neutrinos coming from such reactions. If there is no mixing then the neutrino is always an electron-neutrino. If there is mixing, the neutrino emitted in this reaction is some mixture of electron-neutrino, muon-neutrino and tau-neutrino, and that could be observed by considering the reactions induced by these neutrinos. Experimentally that is far from easy, but observations seem to indicate some mixing. We shall have to await more detailed experimental results. Solar neutrino experiments are among the general class of experiments over which the experimenter has only limited control, and for a truly unequivocal proof we will have to wait for accelerator experiments, of which there are several being built.

Some indication of why masses play a role here may be useful. Imagine the production of a neutrino as in the reaction mentioned above. On the detection side precisely the inverse transition is looked for in the detectors. Since mixing is the same on both sides you would never know that there is any, one would still obtain an electron from precisely the mixture emitted primarily.

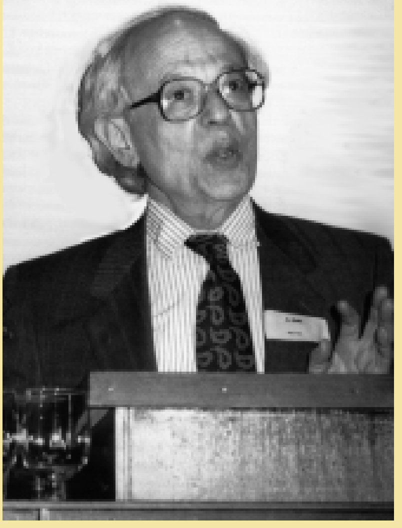
However, the neutrino must cross some distance from emission to detection, such as from the sun to the earth. The neutrinos have a certain energy, and if the masses of the electron-neutrino and the muon-neutrino (or the tau-neutrino) were different then they would travel at slightly different velocities. In other words, the neutrino mixture will change while traveling, and the mixture observed at detection is no more the same as the one emitted.

The reader may be warned that the above argument is a very simplified one and should be understood only as an indication why the distance between emission and detection and the values of the masses are of importance when observing neutrinos. Quantum mechanics tells us that the propagation of particles has much to do with the propagation of waves, and that plays an important role in these discussions. Even so, there is much truth in the argument.

3.6 Particle Mixing

The strange phenomenon of particle mixing is another exclusively quantum mechanical effect. Some discussion is in order.

Cabibbo mixing is thought to be the result of a particle mixing process, so let us take that as an example. Forget for the moment about the top and bottom quarks. Consider first the case before rotation. The up quark goes exclusively into the down quark, the charm quark exclusively into the strange quark.



Abraham Pais (1918–2000). Pais, the author of the books mentioned in the introduction, was a very accomplished physicist. Together with Gell-Mann he published a paper introducing the idea of particle mixing. This was in connection with K^0–$\overline{K}^0$ mixing, a very curious system indeed. When producing a K_0 it would after a while become an $\overline{K}^0$ and the other way around. In the end this resolved itself into a combination of two mixtures, called K_S and K_L. They have very different properties; K_S decays quite quickly, while K_L lives much longer.

Pais introduced the idea of associated production, which is in fact the idea of a new quantum number now called strangeness which had to be conserved in all but weak interactions. Actually, several Japanese physicists published similar ideas at about the same time. This rule explains why certain particles were always produced in pairs (one with strangeness +1, the other −1, so that the sum was 0), given that the initial particles would have no strangeness. This was generally the case, because proton and neutron have strangeness zero, and the new particles were seen in collisions of protons with the protons or neutrons in a nucleus.

Pais, Jewish, living in the Netherlands during World War II, barely survived. He was released from jail just before the end of the war, after an appeal by a very courageous lady armed with a letter from Kramers to Heisenberg (who did not intervene). Perhaps the commanding officer saw the end coming, reason for a leniency extremely rarely seen. A friend of Pais, arrested at the same time, was shot.

Now imagine that there is some special process, some interaction, that causes the strange quark to go over into a down quark, and a similar interaction making the down quark become a strange quark. These things are quite possible, there is nothing that says that particle processes must involve three particles only (such as for example in the process up $\rightarrow$ down + W^+). In fact, one may have transitions involving four particles, or only two particles, and yes, even stranger, only one particle. The latter is really strange, it is like a particle that just stops to exist. Because energy must be conserved that particle must then have zero energy to begin with, but that is sometimes possible. Anyway, let us turn back to the case of two-particle transitions, namely down $\rightarrow$ strange and strange $\rightarrow$ down. Let us suppose that they occur with a certain strength. The reader may ask how it is possible that particles of different mass go over into each other, and indeed that is not possible except for very short times. That will be discussed in Chapter 9, about particle theory. Just do not worry about that aspect now.

Consider now again the process up $\rightarrow$ down + W^+. Since the down quark may now change through this special process into a strange quark we might in the end observe the process up $\rightarrow$ strange + W^+. That would precisely produce the process described through the Cabibbo rotation, and indeed the current philosophy is that this is the mechanism. The situation is slightly more complicated than stated here, because nothing prevents the strange quark from turning into a down quark again and so on. There is a lengthy set of possibilities and it is up to the theorists to figure out what happens in the end. One must consider chains of transitions.

The way these things work out is that there are two very special combinations of the down and strange quark such that they do not change under such a chain of transitions. Let us call these special combinations the Down and Strange quarks. First consider the Down quark, a combination of down and strange quark. What happens is that the down quark in this Down quark can become a strange quark, but on the other hand the strange quark (in this

Down quark) may turn into a down quark. You can imagine that things are such that the net effect is zero, i.e. that there is no change in the total amount of strange quark inside the Down.

Let us give a very crude example. Image a person, Mr A, a dress artist, capable of changing his clothes very quickly. Assume then that he has two sets of clothes, one red, the other green. Suppose further there is a second person, Mr B, capable of the same quick change of dress. He will dress up in whatever is not used by A. If now A changes from red to green, B must give up his green dress and quickly change into the red one.

Assuming that they change clothes quicker than the eye can see what you will observe are two persons with clothes of a color that you can get by mixing red and green. The precise color depends on how Mr A divides up his time in green and red. If Mr A stays, say, for 4 millisec. in red clothes, changes, and stays in green clothes for 2 millisec. etc. he will look some shade of orange. Mr B, staying longer in green, will show a lemon type color. In other words, we will see two persons in a definite complementary color depending on the time distribution of the clothes.

The Down and Strange are the two complementary combinations. In the experiments we will see the Down and the Strange quark, not down and strange. The process whereby two particles turn into certain mixtures because of particle-particle transitions happens just about everywhere where it is possible. An example where no mixing can occur is this: there can never be a transition mixing the up and the down quark. That would involve a change of charge, which nature is careful not ever to do. So, conservation of quantum numbers may prevent certain mixtures. But in general, if two particles have the same quantum numbers (including spin) then they will mix. For example, in principle the up quark could mix with the charm quark, but while that is true it is not observable because the effects of that cannot be distinguished from the effects of down-strange mixing. Cabibbo mixing can be seen as down-strange mixing or up-charm mixing or even a combination of the two, the net result is the same. This is of course why we emphasized earlier the invariance with respect to figure inversion (upside down flipping). In the CKM rotations shown the last figure above you can rotate the bold lines or keep the bold lines fixed but rotate the coordinate system drawn with thin lines. Physicists have opted for the down-strange mixing convention.

Theoretically, the quark mixing described above is thought to be due to the Higgs particle. It may interact with the quarks in a way that produces this mixing. Other interactions never produce the type of particle-particle transitions needed for mixing. This of course is not an explanation, it just shifts the mystery of the CKM rotation to the mystery of the Higgs couplings. When speaking of the theory it is the specific theoretical construction involving this mysterious Higgs particle. It may not be true. So using the word "theoretically" in this Chapter means that it cannot be explained simply, and that it may be wrong.

Another case of mixing concerns the photon and the Z^0. They have the same quantum numbers and they are indeed the final product of some mixing. There is another angle here, called the weak mixing angle, and one speaks of electroweak mixing. The Z^0 couples to the neutrino's, the photon does not as it indeed should not because the neutrino's carry no charge. Here the mixing corrects a potential problem: the photon is precisely that mixture that has no coupling to the neutrino. That is one of the strange effects of mixing: while two particles may both couple to something, it is quite possible that a certain mixture of the two does not. The various possibilities may cancel. Apparently there is a link between electromagnetic and Higgs interactions. A lot of dirt is swept under the Higgs rug!

Theoretically, the Higgs particle is thought to be largely responsible for the CKM mixing, although also other interactions play a role. From the actual mixing as deduced experimentally one may draw important conclusions concerning the Higgs particle and its interactions. The Cabibbo angle can be measured by comparing the processes

$$\text{up} \rightarrow \text{down} + W^+ \qquad\qquad \text{neutrino} \rightarrow \text{electron} + W^+$$

(and crossed versions). Similarly one may compare the processes

$$\text{up} \to \text{down} + W^+ \qquad\qquad \text{up} \to \text{up} + Z^0$$

If there were no mixing they would go at equal strength. By measuring the strength of these transitions the weak mixing angle can be determined.

Earlier some remarks were made concerning gluons of the "diagonal" type. That are gluons whose two colors are each other's anti-color, such as the red/antired gluon. Also here there are mixing possibilities. A red/antired gluon can become a green/antigreen or a blue/antiblue gluon without any violation of quantum numbers. Therefore the actual combinations that propagate are mixtures of these. One of these combinations (one that might be called the "white gluon") is such that in the end it couples to nobody. Since it would never take part in any reaction we may just as well postulate that it does not exist. For the white gluon to play any role would require a new interaction besides the existing quark-gluon couplings.