

Introduction
To
Theoretical Physics

With Examples Of Solved Problems

by

E. Gozzi and D. Mauro

Department of Theoretical Physics

University of Trieste

Strada Costiera 11, Trieste, Italy

Preface

This is the text of a course delivered in the second semester of the second year to undergraduate students majoring in physics in Italy.

It should expose students for the first time to some aspects of theoretical physics. The aspects that we have chosen are Analytical Mechanics and Introductory Quantum Mechanics.

We used to tell the students to whom we deliver these lectures that “*this booklet is like the one they may have used to get the driving licence*”. The meaning of this sentence is that this booklet wants to give an *introductory working knowledge* of Classical Analytical Mechanics (CM) and of Quantum Mechanics (QM) like the introductory working knowledge that students get from the driving -licence booklet they buy when they enroll in a driving school. From that booklet they do not get too many technical details of how the engine, the clutch or the brakes work. They get some knowledge of these technical details but not to the point of becoming engineers able to design the engine, the clutch or the brakes. Analogously for the the readers of this book. They will get further skills and deeper knowledge , especially in QM, in more advanced courses where more mathematical structures will be presented and in a more rigorous way.

When a student learns how to drive a car he usually practices on his parents old car along some country lanes with the help of the parents or of older brothers, likewise here there are a lot of exercises worked out by us in class in order to help the student.He should anyhow practice later on at home by himself, and we listed below some very good exercise books.The most important part of the exam will be the written one with 2-3 problems to solve in 3-4 hours.

This book, besides being not too rigorous from the mathematical point of view, does not contain everything on CM and QM. In CM, for example, advanced problems on the Hamilton-Jacobi methods, integrability concepts, ergodicity, chaos, classical scattering theory, classical perturbation theory, and in QM , for example the full operator theory, angular momentum, spin, systems of identical particles, perturbation theory, variational methods, quantum scattering etc. are left for more advanced courses, but we think that the essence of CM and QM is nevertheless captured here together with a good working practice on basic problems. Most books on CM and QM are 400 pages long plus a second 300 page book on exercises for a total weight of at least 4 Kg. We wanted instead to create a book which was “lighter” in every sense..... Our ideal was the slim book by Landau on Classical Mechanics and the handwritten lecture notes on QM by E. Fermi (recently republished by the Chicago University Press: E. Fermi, “*Notes on Quantum Mechanics*” (Chicago, University Press, 1995)). For sure our result is not comparable

scientifically to these ones. In common we may only have the weight.

A lot of material, and especially the exercises, are taken from several books:

- L.D. Landau and E.M. Lifshitz, “*Course of Theoretical Physics, vol.1: Mechanics*”, (London, Pergamon Press,1976);
- H. Goldstein, “*Classical Mechanics*”, (Reading, MA, Addison-Wesley. Pub.Co. 1980);
- Y.K.Lim, “*Problems and solutions on Mechanics*”, Singapore, World Scientific, 1994;
- Dare Wells, “*Theory and Problems of Lagrangian Dynamics: with a treatment of Euler’s equations of motion, Hamilton’s equation and Hamilton’s principle.*” Schaum’s outline series, New York, MacGraw Hill 1967;
- R. Eisberg and R. Resnick, “*Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles*” (New York, Wiley, 1989);
- A. Messiah, “*Quantum Mechanics*” (Amsterdam, North-Holland, 1961);
- L.I. Schiff, “*Quantum Mechanics*” (New York, McGraw-Hill, 1968);
- V. Galitski, B. Karnakov and V. Kogan, “*Problemes de Mecanique Quantique*” (Moscow, MIR);
- Y.-K. Lim, “*Problems and Solutions on Quantum Mechanics*” (Singapore, World Scientific, 1997).

Mistakes that we may have done in adapting the material from the books above are entirely our responsibility. We hope anyhow to have assembled the material taken from those books in a manner useful for the student. Besides the notes presented here there is, on the same home-page, an appendix on the Noether theorem and one on the WKB method which is part of the course. In the future we may add other appendices.

The book is addressed not only to physics students who want to learn the basics of analytical CM and introductory QM but it is also addressed to engineering, chemistry and biology students for whom QM is becoming an increasingly important subject in their field of study and research.

We would like to thank all those from whom we learned CM and QM: they are M. Berry, S.Fubini, G.Furlan, G.C. Ghirardi, C.Orzalesi, M. Pauri, M. Reuter and B. Sakita.

This book is dedicated to all those future students who maybe, by having learned to drive the car and to bring it to its speed limits, will find out that the “engine” has some problems in the most extreme regimes. Maybe they will find a way to fix it or improve the “engine”, that means improve Quantum Mechanics, or find new experiments that will shed further light on QM. In that manner they will implement the dream of J. Bell who, for all his life, wanted to be a “*quantum engineer*”.

Contents

1	Classical mechanics	1
1.1	Least Action Principle	1
1.2	Lagrange equations	3
1.3	Hamilton equations of motion	6
1.4	The action functional	8
1.4.1	Use of the Hamilton-Jacobi equation to get the solution of the equations of motion	11
1.5	Poisson brackets	13
1.5.1	Constants of motion	14
1.6	Canonical transformations	15
1.6.1	Example of canonical transformations	17
1.6.2	Identity transformation	17
1.6.3	Invariance of the pb under canonical transformations	18
1.7	Liouville theorem	20
1.8	Symmetries and their generators (Hyper-simplified treatment)	23
2	Crisis of Classical Physics	26
2.1	Introduction	26
2.2	Black Body Radiation	27
2.3	Classical Derivation of the Black Body Radiation	29
2.4	Planck's Hypothesis	31
2.5	Photoelectric Effect	35
2.6	Compton Effect	38
2.7	Spectral Lines of an Atom; Rutherford's and Bohr's Models	42
2.8	De Broglie's idea	47
2.9	Wave-Particle Duality	49
2.10	Heisenberg's Uncertainty Principle	52
3	Schrödinger Equation	58
3.1	General Properties	58

3.2	Solution of the Time-Dependent Schrödinger Equation	60
3.3	Properties of the Solutions	62
3.4	Quantization of the Energy Levels	66
3.5	Schrödinger Equation for a Free Particle	70
3.6	One-dimensional Problems	72
3.7	Dirac Delta	88
	Bibliography	94

Chapter 1

Classical mechanics

1.1 Least Action Principle

According to Newton, the acceleration \vec{a}_i of a particle of mass m_i is given by the formula

$$m_i \vec{a}_i = \vec{F}_i^{\text{source}} + \vec{F}_i^{\text{const}} \quad (1.1.1)$$

where $\vec{F}_i^{\text{source}}$ are the forces exerted by internal sources like a potential (gravitational, electromagnetic etc.) and the \vec{F}_i^{const} are those exerted by constraints like for example a table on which the particle rests and so on.

It is often difficult to solve Eq. (1.1.1) because it is hard to figure out the \vec{F}_i^{const} . Usually constraints are given by relations like

$$C_a(\vec{r}_1, \vec{r}_2, \dots) = 0, \quad a = 1, \dots, k, \quad (1.1.2)$$

where C_a are functions of the positions $\vec{r}_1, \vec{r}_2, \dots$ of the various particles and to find out from (1.1.2) the \vec{F}_i^{const} may be hard. So people (Bernoulli, d'Alambert, Maupertuis, Lagrange, Hamilton) tried to develop methods to get the equations of motion without knowing the force exerted by the constraints. These methods are known as *variational methods* and work also for systems without constraints. We will work out first these systems and later on we shall show how the same methods work also in case constraints are present.

Let us introduce the following function of q, \dot{q} , known as Lagrangian, for a point particle of mass $m = 1$ in a potential $U(q)$

$$\mathcal{L}(q, \dot{q}) = \frac{\dot{q}^2}{2} - U(q). \quad (1.1.3)$$

Next let us introduce the following *functional* called the action

$$S[q(t)] = \int_{t_1 q_1}^{t_2 q_2} dt \mathcal{L}(q(t), \dot{q}(t)) \quad (1.1.4)$$

where $[q(t)]$ is a trajectory (any) between t_1, q_1 and t_2, q_2 :

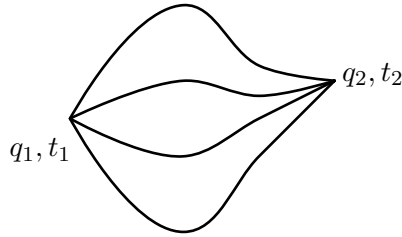


Fig. 1.1: Trajectories between q_1, t_1 and q_2, t_2 .

So $S[q(t)]$ gives you a number once you insert a “particular trajectory” in (1.1.4). A particular trajectory is a function $q(t)$, so once you insert it in (1.1.4), the integrand $\mathcal{L}(q(t), \dot{q}(t))$ becomes a function of t and you can perform explicitly the integration in t .

The *classical trajectory* is just a particular one among the many present in Fig. 1.1. which one? We know that to solve the equation of Newton

$$\ddot{q} = -\frac{\partial U}{\partial q} \quad (1.1.5)$$

we have to give *two* conditions that are either the initial position $q(t_1)$ and the initial velocity $\dot{q}(t_1)$ or the initial position $q(t_1)$ and the final one $q(t_2)$. Once these are given there is only one solution of the equation of motion (1.1.5), i.e. just one trajectory among those of Fig. 1.1 .

The principle of *least action* tells us that the classical trajectory is the one that “*minimize*” the action functional $S[q(t)]$.¹ Basically we look where the first variation is zero, or where the *functional* derivative of the functional $S[q(t)]$ is zero. It is the “analog” of looking where the first derivative of a function is zero. These points are “minima” if the second derivative is positive. This is the case in CM unless there are “focal points”, i.e. two trajectories with the same q_1, t_1 and q_2, t_2 .

Let us do a first variation of S , that means let us calculate S on $q(t) + \delta q(t)$ and on $q(t)$

$$\delta S = S[q(t) + \delta q(t)] - S[q(t)] \quad (1.1.6)$$

where the $\delta q(t)$ have the properties pictured in Fig. 1.1.6 that $\delta q(t_1) = \delta q(t_2) = 0$, i.e. the trajectories have the same initial and final points. We will expand (1.1.6) in powers of δq and put the first term to zero

$$\delta S = S[q(t) + \delta q(t)] - S[q(t)] = \frac{\delta S}{\delta q} \delta q + \frac{\delta^2 S}{\delta q^2} \delta^2 q + \dots$$

The variation of the action expressed in terms of the Lagrangian \mathcal{L} is given by:

$$\delta S = \delta \int_{t_1 q_1}^{t_2 q_2} dt \mathcal{L}(q, \dot{q}) = \int_{t_1 q_1}^{t_2 q_2} dt \left[\frac{\partial \mathcal{L}}{\partial q} \delta q(t) + \frac{\partial \mathcal{L}}{\partial \dot{q}} \delta \dot{q}(t) \right].$$

If we integrate by parts the last term of the previous equation we get

$$\delta S = \left[\frac{\partial \mathcal{L}}{\partial \dot{q}} \delta q \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} dt \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \delta q(t). \quad (1.1.7)$$

¹Actually one does not talk about “minimization” but “extremization” or “stationarity”

The first term on the RHS is zero because of the boundary conditions $\delta q(t_1) = \delta q(t_2) = 0$. As $\delta q(t)$ is arbitrary, for δS to be zero in the first order in δq , we need the integrand to be zero. This gives the following:

1.2 Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = 0. \quad (1.2.1)$$

Now if we use the explicit form of \mathcal{L} , i.e. $\mathcal{L} = \frac{\dot{q}^2}{2} - U(q)$, we get

$$-\frac{\partial U}{\partial q} - \frac{d}{dt} \dot{q} = 0 \implies \ddot{q} = -\frac{\partial U}{\partial q},$$

which are the standard equations of motion of classical mechanics.² Are these *minima* of the action or not? Up to now we have only proved that they are *extremals* of the action. We should prove that $\frac{\delta^2 S}{\delta q(t) \delta q(t)} > 0$. Actually it is possible to prove this if there is no *conjugate point* along the classical trajectory. A conjugate point is a point where two classical trajectories meet. It happens in fact that for *some system* it is not true anymore that giving an initial and a final point there is only one solution. It is true if you give the initial *position* and the initial *velocity* but not if you give the initial position and the final one. The point where two trajectories meet again is called *focal point*. For infinitesimal times $t_1 - t_2 \sim dt$ then it is *impossible* to get focal points and we are sure the trajectory is a minimum, see Schulman, “*Techniques and application of path integration*” (Wiley, 1981).

Why is the Lagrangian useful? Because for systems with constraints we do not need to insert in \mathcal{L} the forces exerted by the constraints but just add the constraints C^a via suitable Lagrangian multipliers λ_a :

$$\mathcal{L} \rightarrow \mathcal{L} + \sum_a \lambda_a C^a(r_1, r_2).$$

The variation with respect to λ_a gives the constraints:

$$C^a(r_1, r_2) = 0$$

and the equations of motion get modified by

$$\frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} + \sum_a \lambda_a \frac{\partial C^a}{\partial q} = 0.$$

The constraints which depend only on the configuration variables are called *holonomic*. The constraints which depend on the velocities are called *unholonomic*. The simple constraints like for example

$$a\dot{q}_1 = b\dot{q}_2 \quad (1.2.2)$$

²As we will prove at the end of this section the Lagrangian is not uniquely determined by this principle. We can in fact add the derivative of any function $F(q, t)$ to \mathcal{L} .

can be inserted directly in the variation of the action. In particular, the infinitesimal version of (1.2.2)

$$a \frac{dq_1}{dt} = b \frac{dq_2}{dt} \implies a \delta q_1 = b \delta q_2$$

can be inserted into

$$\delta S' = \delta S + \lambda [a \delta q_1 - b \delta q_2].$$

This is possible because δS contains only the variations of $\delta q_1, \delta q_2$ when we do

$$\frac{\delta S'}{\delta q_1} = \frac{\delta S}{\delta q_1} + \lambda a, \quad \frac{\delta S'}{\delta q_2} = \frac{\delta S}{\delta q_2} - \lambda b.$$

There is then a procedure to determine the Lagrange multipliers etc. (Dirac method) without even inserting the associated forces (see the problems).

Another approach is to pass from the $3N$ Cartesian coordinates to the $3N - k$ unconstrained variables q_i by solving the constraints:

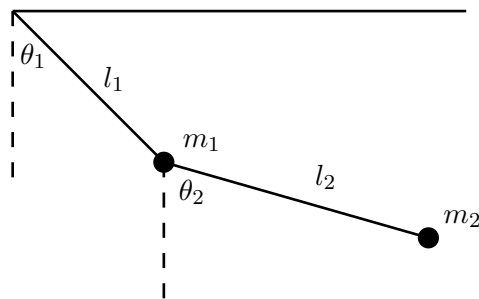
$$C^a(r_1, r_2, \dots) = 0, \quad a = 1, \dots, k.$$

We get

$$\begin{aligned} r_1 &= r_1(q_1, \dots, q_{3N-k}) \\ r_2 &= r_2(q_1, \dots, q_{3N-k}) \\ \dots &\quad \dots \\ r_{3N} &= r_{3N}(q_1, \dots, q_{3N-k}) \end{aligned}$$

and then write $\mathcal{L}(r, \dot{r})$ in terms of q_1, \dots, q_{3N-k} where in $\mathcal{L}(r, \dot{r})$ we consider only the external potential and not the constraints (because we have already taken care of the constraints by solving their equations).

Example. Study the motion of this double pendulum under the gravitational force.



There are constraints among the six degrees of freedom: the \vec{r}_1 describing the mass m_1 and the \vec{r}_2 describing the mass m_2 . If we are on a plane the six degrees of freedom becomes 4, then we have also 2 constraints $|r_1|^2 = l_1^2$ and $|r_2 - r_1|^2 = l_2^2$, so effectively there are just 2 degrees of freedom. The forces on m_1 are its gravitational pull, plus the

force of the ropes l_1 and l_2 . To calculate these forces is hard. Instead let us find out the 2 independent variables which are θ_1 and θ_2 . Now the kinetic energy for the particle 1 is

$$T_1 = \frac{1}{2}m_1\dot{r}_1^2 = \frac{1}{2}m_1l_1^2\dot{\theta}_1^2.$$

The potential energy for the particle 1 is $U_1 = -m_1gl_1 \cos \theta_1$. The Cartesian coordinates of the particle 2 are:

$$x_2 = l_1 \sin \theta_1 + l_2 \sin \theta_2, \quad y_2 = l_1 \cos \theta_1 + l_2 \cos \theta_2.$$

So the kinetic energy of the particle 2 is

$$\begin{aligned} T_2 &= \frac{1}{2}m_2(\dot{x}_2^2 + \dot{y}_2^2) \\ &= \frac{1}{2}m_2 \left[l_1^2\dot{\theta}_1^2 + l_2^2\dot{\theta}_2^2 + 2l_1l_2 \cos(\theta_1 - \theta_2)\dot{\theta}_1\dot{\theta}_2 \right]. \end{aligned}$$

The potential energy is

$$U_2 = -m_2gy_2 = -m_2g[l_1 \cos \theta_1 + l_2 \cos \theta_2].$$

Summing up everything we get

$$\begin{aligned} \mathcal{L} &= \frac{1}{2}(m_1 + m_2)l_1^2\dot{\theta}_1^2 + \frac{1}{2}m_2l_2\dot{\theta}_2^2 + m_2l_1l_2\dot{\theta}_1\dot{\theta}_2 \cos(\theta_1 - \theta_2) + \\ &\quad + (m_1 + m_2)gl_1 \cos \theta_1 + m_2gl_2 \cos \theta_2. \end{aligned}$$

So one sees that we avoided introducing the forces of the constraint but we consider only the external forces. The trick is

- Write the kinetic and potential terms in terms of the Cartesian coordinates without constraints.
- Write the Cartesian coordinates in terms of the unconstrained variables.

Question: Given the equations of motion is there one and only one Lagrangian which reproduces them?

No! Two Lagrangians which differ by a total derivative of a function of q, t , i.e. $F(q, t)$ gives the same equations of motion.

Proof. Let us define the following Lagrangian:

$$\mathcal{L}'(q, \dot{q}) \equiv \mathcal{L}(q, \dot{q}) + \frac{dF(q, t)}{dt}.$$

The action is

$$S' = \int_{t_1q_1}^{t_2q_2} dt \mathcal{L}'(q, \dot{q}) = \int_{t_1q_1}^{t_2q_2} dt \mathcal{L}(q, \dot{q}) + \int_{t_1q_1}^{t_2q_2} dt \frac{dF(q, t)}{dt},$$

which implies $S' = S + F(q_2t_2) - F(q_1t_1)$. Now when we do the variation of S' we get

$$\delta S' = \delta \int dt \mathcal{L} + \delta [F(q_1t_1) - F(q_2t_2)] = \delta \int dt \mathcal{L} + \frac{\partial F}{\partial q_1} \delta q_1 - \frac{\partial F}{\partial q_2} \delta q_2,$$

We know that the variation is such that the end points are fixed, so $\delta q_1 = \delta q_2 = 0$. so we get zero that means

$$\delta S' = \delta S$$

This implies $\delta S' = \delta S$ and if $\delta S = 0$ then the same is for $\delta S'$. So also from \mathcal{L}' we can reproduce the same equations of motion (1.2.1).

1.3 Hamilton equations of motion

The Lagrange equations are second order in the time derivative. One wonders if, like many second order equations, they can be turned into two equations each first order in the derivative w.r.t. t . The answer is yes. The procedure is the following. Define the momentum:

$$p(t) \equiv \frac{\partial \mathcal{L}(q, \dot{q})}{\partial \dot{q}} \quad (1.3.1)$$

which is a function of q and \dot{q} , i.e.: $p = F(\dot{q}, q)$. Eq. (1.3.1) can be inverted and we have

$$\dot{q} = \mathcal{G}(p, q). \quad (1.3.2)$$

Let us now build the function of q, p given by the Lagrangian with \dot{q} replaced by (1.3.2)

$$\mathcal{L}(q, \dot{q}) = \mathcal{L}(q, \mathcal{G}(q, p)) = \mathcal{L}'(q, p).$$

We can derive the following equations of motion:

$$\left\{ \begin{array}{l} \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} - \frac{\partial \mathcal{L}}{\partial q} = 0 \\ \frac{d}{dt} p - \frac{\partial \mathcal{L}}{\partial q} = 0 \end{array} \right.$$

The second equation implies

$$\dot{p} = \frac{\partial \mathcal{L}}{\partial q} = U(q, \dot{q}).$$

Now if we replace \dot{q} with (1.3.2) we have

$$\dot{p} = U(q, \mathcal{G}(q, p)).$$

We need also an equation for \dot{q} which is (1.3.2)

$$\dot{q} = \mathcal{G}(q, p).$$

Instead of this procedure, it is possible to get a simpler one introducing a function already of p and q , the Hamiltonian. We get it this way: consider the differential of \mathcal{L}

$$d\mathcal{L} = \frac{\partial \mathcal{L}}{\partial q} dq + \frac{\partial \mathcal{L}}{\partial \dot{q}} d\dot{q}$$

which can be written as

$$\begin{aligned} d\mathcal{L} &= \dot{p}dq + p d\dot{q} \\ &= \dot{p}dq + d(p\dot{q}) - \dot{q}dp \end{aligned}$$

which implies

$$d(p\dot{q} - \mathcal{L}) = -\dot{p}dq + \dot{q}dp. \quad (1.3.3)$$

On the LHS we have the differential of something which can be expressed in terms of the differentials of dq and dp , so it must be a function of q, p

$$p\dot{q} - \mathcal{L} \equiv H(q, p). \quad (1.3.4)$$

One obtains it by replacing the \dot{q} appearing on the LHS of (1.3.4) with $\dot{q} = \mathcal{G}(q, p)$ of (1.3.2). H is called the *Hamiltonian*. So (1.3.3) can be rewritten as

$$dH(q, p) = -\dot{p}dq + \dot{q}dp.$$

So we get from here the equations

$$\begin{cases} \dot{p} = -\frac{dH}{dq} \\ \dot{q} = \frac{dH}{dp} \end{cases} \quad (1.3.5)$$

These are the 2 Hamilton first order equations equivalent to the 1 second order equation of Lagrange. Replacing (1.3.2) into (1.3.4) and using the explicit expression of the Lagrangian (1.1.3) we get

$$H(q, p) = \frac{p^2}{2} + U(q)$$

and the equations (1.3.5) can be rewritten as:

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \dot{p} = -\frac{\partial H}{\partial q}$$

H can be identified with the energy and it is a conserved quantity:

$$\begin{aligned} \frac{dH}{dt} &= \frac{dH}{dq}\dot{q} + \frac{dH}{dp}\dot{p} \\ &= \frac{dH}{dq}\frac{dH}{dp} - \frac{dH}{dp}\frac{dH}{dq} = 0. \end{aligned}$$

We can derive the Hamilton equations from a variational principle. Let us start by rewriting the action in terms of the differentials in q and t :

$$\begin{aligned} S &= \int \mathcal{L}dt = \int (p\dot{q} - H)dt \\ &= \int p\frac{dq}{dt}dt - Hdt = \int (pdq - Hdt) \end{aligned}$$

Let us now make the variation in p, q with again $\delta q(t_1) = \delta q(t_2) = 0$. The δp instead can vary at both end points.

$$\delta S = \int \left(\delta pdq + p d(\delta q) - \frac{\partial H}{\partial q}\delta q dt - \frac{\partial H}{\partial p}\delta p dt \right).$$

Integrating by parts the piece $p\delta q$ we get

$$\delta S = \int \delta p \left[dq - \frac{\partial H}{\partial p} dt \right] - \int \delta q \left[dp + \frac{\partial H}{\partial q} dt \right] + p\delta q \Big|_{t_1}^{t_2}.$$

The last term is zero and so, as δp and δq are arbitrary in between, we get that, in order for δS to be zero, we need the integrands to be zero:

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial q}.$$

Minimal Coupling There are forces, like the Lorentz force, which depend on the velocity. There is a manner to provide a Hamiltonian formulation for them. The rule is to substitute \vec{p} in H with $\vec{p} - \frac{e}{c}\vec{A}$ where \vec{A} is the gauge vector potential associated to the fields present in the system

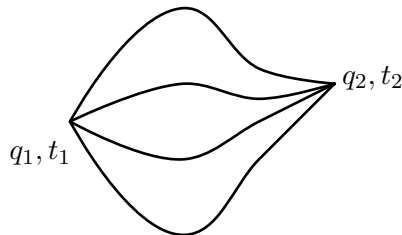
$$H = \frac{p^2}{2m} + U(q) \longrightarrow H' = \frac{\left(\vec{p} - \frac{e}{c}\vec{A}\right)^2}{2m} + U(q) + e\varphi(q)$$

where $\varphi(q)$ is the electric potential.

Exercise Prove that from H' one can get the usual Lorentz force generated by a magnetic field $\vec{F} = \frac{1}{c}\vec{v} \wedge \vec{B}$.

1.4 The action functional

Up to now we have defined the action functional $S[q(t)]$ on a particular subset of trajectories that are those with fixed end points (q_1, q_2) at initial and final times (t_1, t_2) . They are represented in the figure below:



The variation of the action δS on these set of paths was given in Eq. (1.1.7):

$$\delta S = \int_{t_1}^{t_2} dt \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \delta q(t). \quad (1.4.1)$$

We got this expression because the end points q_1, q_2 and times t_1, t_2 were fixed so

$$\begin{cases} \delta q_1 = \delta q_2 = 0 \\ \delta t_1 = \delta t_2 = 0. \end{cases} \quad (1.4.2)$$

Let us now enlarge the set of paths on which the action functional is defined as those for which t_1 and t_2 are fixed but q_1, q_2 are *not* fixed:

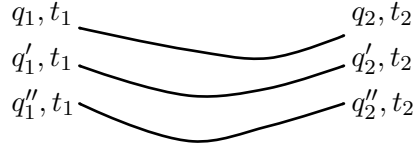


Fig. 1.2: Paths without fixed end points.

The variation of $S[q(t)]$ is now different and given by³

$$\delta S = \left[\frac{\partial \mathcal{L}}{\partial \dot{q}} \delta q \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} dt \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \delta q(t). \quad (1.4.3)$$

Finally we could ask what is the *variation* of S if we work on all possible paths, that means those for which even t_1 and t_2 are not fixed. Basically to the variation (1.4.3) we have to add the piece which comes from the variation of t_1 and t_2 . This is the following

$$\delta_{t_1} S + \delta_{t_2} S = \int_{t_1 + \Delta t_1}^{t_2} dt \mathcal{L} - \int_{t_1}^{t_2} dt \mathcal{L} + \int_{t_1}^{t_2 + \Delta t_2} dt \mathcal{L} - \int_{t_1}^{t_2} dt \mathcal{L} \quad (1.4.4)$$

$$= -\mathcal{L} \Delta t_1 + \mathcal{L} \Delta t_2 = \mathcal{L} \Delta t \Big|_{t_1}^{t_2}. \quad (1.4.5)$$

This is clearly the most general variation that S can undergo. We could write it in a slightly different form noting that q at the end points undergoes two types of variation, one indicated as δq that we performed without changing time, i.e.

$$\begin{cases} \delta q_1 = q'(t_1) - q(t_1) \\ \delta q_2 = q'(t_2) - q(t_2) \end{cases} \quad (1.4.6)$$

and another one due to the change of time $t_1 \rightarrow t_1 + \Delta t_1, t_2 \rightarrow t_2 + \Delta t_2$. This is $\dot{q} \Delta t$, so the overall change in q at the end points is

$$\begin{cases} \Delta q_1 \equiv \delta q_1 + \dot{q}_1 \Delta t_1 \\ \Delta q_2 \equiv \delta q_2 + \dot{q}_2 \Delta t_2. \end{cases} \quad (1.4.7)$$

We indicated it with Δ instead of δ . Δ is the following variation defined *not* at the same time

$$\begin{cases} \Delta q_1 = q'(t'_1) - q(t_1) \\ \Delta q_2 = q'(t'_2) - q(t_2). \end{cases} \quad (1.4.8)$$

Using (1.4.7) in the form

$$\begin{cases} \delta q_1 = \Delta q_1 - \dot{q}_1 \Delta t_1 \\ \delta q_2 = \Delta q_2 - \dot{q}_2 \Delta t_2 \end{cases}$$

we get

$$\delta S = \int_{t_1}^{t_2} dt \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \delta q + \left[\left(\mathcal{L} - \frac{\partial \mathcal{L}}{\partial \dot{q}} \dot{q} \right) \Delta t + \frac{\partial \mathcal{L}}{\partial \dot{q}} \Delta q \right]_{t_1}^{t_2}$$

³The first term on the RHS of (1.4.3) was zero in (1.4.1) because $\delta q_1 = \delta q_2 = 0$.

which can be rewritten as:

$$\delta S = \int_{t_1}^{t_2} dt \left(\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} \right) \delta q + (p\Delta q - H\Delta t) \Big|_{t_1}^{t_2}. \quad (1.4.9)$$

Note that if we restrict the paths in Fig. 1.2 to be the classical ones, i.e. those for which $\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = 0$ we get from⁴ (1.4.9)

$$\delta S_{\text{cl}} = (p\Delta q - H\Delta t) \Big|_{t_1}^{t_2}. \quad (1.4.10)$$

The “cl” is for classical. Note that this expression is not anymore a functional because there is no integration over the paths (while the δS in (1.4.9) was still a functional). From (1.4.10) if we restrict the paths further, to be the classical ones which start from q_1 at time t_1 (so $\delta q_1 = \delta t_1 = 0$) we get

$$\delta S_{\text{cl}} = p_2 \Delta q_2 - H \Delta t_2 \quad (1.4.11)$$

and from here we obtain

$$\begin{cases} \frac{\delta S_{\text{cl}}}{\Delta q_2} = p_2 \\ \frac{\delta S_{\text{cl}}}{\Delta t_2} = -H. \end{cases} \quad (1.4.12)$$

As the S_{cl} is now a function we can turn the δ into a partial derivative, and, calling p_2, q_2, t_2 as p, q, t , we can write (1.4.12) as

$$\begin{cases} \frac{\partial S_{\text{cl}}}{\partial q} = p \\ \frac{\partial S_{\text{cl}}}{\partial t} = -H. \end{cases} \quad (1.4.13)$$

From here we get that S_{cl} is a function of q, t . These are the only things that can change, as all the classical paths start from q_1, t_1 as in the figure below:

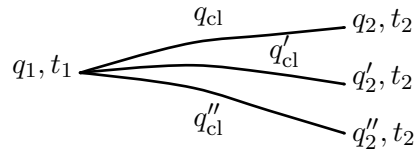


Fig. 1.3: Classical paths starting from q_1, t_1 .

From (1.4.13) we get also that

$$\begin{aligned} \frac{dS_{\text{cl}}(q, t)}{dt} &= \frac{\partial S_{\text{cl}}}{\partial t} + \frac{\partial S_{\text{cl}}}{\partial q} \dot{q} \\ &= -H + p\dot{q}, \end{aligned}$$

⁴“cl” stands for classical.

i.e.

$$\frac{dS_{\text{cl}}}{dt} = \mathcal{L}. \quad (1.4.14)$$

So the total derivative of S_{cl} with respect to t is the Lagrangian while the partial derivative is minus the Hamiltonian. Note from (1.4.14) that it *is not*: $\frac{dS}{dt} = \mathcal{L}$ with S the general functional we started from, but $\frac{dS_{\text{cl}}}{dt} = \mathcal{L}$ where S_{cl} is the “functional”

$$S_{\text{cl}} = \int_{t_1}^{t_2} \mathcal{L}[q_{\text{cl}}(t), \dot{q}_{\text{cl}}(t)] dt. \quad (1.4.15)$$

obtained by inserting in the integrand the *classical* trajectory which starts from (q_1, t_1) and ends up in (q, t) . S_{cl} can be calculated this way or by solving a differential equation because, after all, S_{cl} turns out to be a function. The differential equation can be built from (1.4.13):

$$\frac{\partial S_{\text{cl}}(q, t)}{\partial t} = -H = -\left(\frac{p^2}{2} + U(q)\right).$$

Replacing p above with $\frac{\partial S_{\text{cl}}}{\partial q}$ we get

$$\frac{\partial S_{\text{cl}}(q, t)}{\partial t} = -\left(\frac{\partial S_{\text{cl}}}{\partial q}\right)^2 / 2 - U(q). \quad (1.4.16)$$

This partial differential equation is called *Hamilton-Jacobi* (HJ) equation. It is possible to show that if we write the wave function of the Schroedinger equation as

$$\psi(q, t) = A(q, t)e^{i/\hbar S(q, t)} \quad (1.4.17)$$

then in the limit of $\hbar \rightarrow 0$ the Schroedinger equation goes into the Hamilton-Jacobi equation where the S in (1.4.17) becomes the S_{cl} of (1.4.16).

1.4.1 Use of the Hamilton-Jacobi equation to get the solution of the equations of motion

We said that the HJ is a third way to get the classical motion of a particle, the other two being the Lagrange equation and the Hamilton equation. But while in these two equations we get directly $q(t)$ and $\dot{q}(t)$ by solving the associated differential equation, i.e.:

$$\frac{\partial \mathcal{L}}{\partial q} - \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{q}} = 0$$

or

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} = -\frac{\partial H}{\partial q} \end{cases}$$

in the HJ equation we get the function $S(q, t)$ as solution of the equation

$$\frac{\partial S}{\partial t} + \frac{1}{2} \left(\frac{\partial S}{\partial q}\right)^2 + U(q) = 0. \quad (1.4.18)$$

How do we get the trajectories $q(t)$ and $\dot{q}(t)$? To solve (1.4.18) we have to give $S(q, 0)$. As $\frac{\partial S(q, 0)}{\partial q} = p(0)$ we immediately see that giving $S(q, 0)$ is equivalent to giving $p(0)$. If we then get a complete solution of (1.4.18) at any time t , i.e. $S(q, t)$ we know that

$$\frac{\partial S(q, t)}{\partial q} = p(t),$$

i.e.

$$\frac{\partial S(q, t)}{\partial q} = \dot{q}(t).$$

The LHS will be a function $F(q, t)$ i.e. $F(q, t) = \dot{q}(t)$. We can solve this equation by giving $q(0)$, as it is a first order equation. So we get the trajectory and all we have given is $q(0)$ and $p(0)$, like in the Hamilton equations.

The $S(q, t)$ actually must contain the constants $q(0)$ because S_{cl} was built from \mathcal{L} calculated along the classical solutions which started from $q(0)$ and ended up in a q which could change. So $S(q, t)$ should actually be $S(q, q(0), t)$. In general we can replace $q(0)$ with another constant α to get something like $S(q, \alpha, t)$. In fact, actually, the “complete solution” of the partial HJ differential equation for a theory of n degrees of freedom (q_1, q_2, \dots, q_n) is of the form $S(q_1, q_2, \dots, q_n; \alpha_1 \dots \alpha_{n+1}, t)$ with $n + 1$ constants. One of the constants is an additive constant. In fact if S is a solution also $S = W - \alpha t$ is a solution of the equation:

$$\frac{\partial S}{\partial t} + \left(\frac{\partial S}{\partial q} \right)^2 \frac{1}{2} + V(q) = 0$$

provided we fix the value of α as follows:

$$\alpha = \left(\frac{\partial W}{\partial q} \right)^2 \frac{1}{2} + V(q).$$

α is basically the constant energy. W is called restricted characteristic function and S characteristic function.

Let us represent in q space the surfaces $S(q, \alpha, t)$:

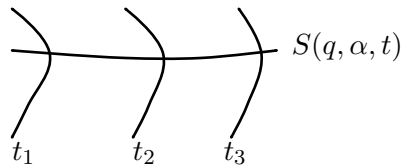


Fig. 1.4: Surfaces $S(q, \alpha, t)$ in q space.

The trajectories of the point particle have momenta $p = \frac{\partial S}{\partial q}$ or, in more than one dimension $\vec{p} = \vec{\nabla}_q S$. So \vec{p} is perpendicular to S , or, in other words, S describes a family of trajectories all perpendicular to S . The end point q can change, so the trajectories are more than one. S is like a *wave-front*. We will return on this at the end of the course once you have done *optics* in order to draw an analogy with that discipline.

1.5 Poisson brackets

If we take an observable $O(p, q, t)$ and make its evolution in time⁵

$$\frac{dO}{dt} = \frac{\partial O}{\partial t} + \frac{\partial O}{\partial q} \dot{q} + \frac{\partial O}{\partial p} \dot{p} \quad (1.5.1)$$

$$= \frac{\partial O}{\partial t} + \frac{\partial O}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial O}{\partial p} \frac{\partial H}{\partial q} \quad (1.5.2)$$

$$= \frac{\partial O}{\partial t} + \{O, H\}_{pb}, \quad (1.5.3)$$

where $\{O, H\}_{pb} = \frac{\partial O}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial O}{\partial p} \frac{\partial H}{\partial q}$. If O is just p or q we get the equations of motion for p and q :

$$\begin{cases} \dot{q} = \{q, H\}_{pb} \\ \dot{p} = \{p, H\}_{pb}. \end{cases} \quad (1.5.4)$$

One sees that they are identical in form and one should not worry where to put the - sign which appears in the Hamilton equations. In general, the pb between two functions f and g are defined as:

$$\{f, g\} \equiv \frac{\partial f}{\partial q} \frac{\partial g}{\partial p} - \frac{\partial f}{\partial p} \frac{\partial g}{\partial q}.$$

They have the following properties:

- $\{f, g\} = -\{g, f\}$
- If C is a constant than $\{f, C\} = 0$
- $\{f_1 + f_2, g\} = \{f_1, g\} + \{f_2, g\}$
- $\{f_1 f_2, g\} = f_1 \{f_2, g\} + f_2 \{f_1, g\}$
- $\frac{\partial}{\partial t} \{f, g\} = \left\{ \frac{\partial f}{\partial t}, g \right\} + \left\{ f, \frac{\partial g}{\partial t} \right\}$.
- **Constant of motion:** If O does not depend explicitly on t we have $\frac{dO}{dt} = \{O, H\}$.

So if $\frac{dO}{dt} = 0$ we get $\{O, H\} = 0$.

- $q_i, q_j = 0, \quad \{q_i, p_j\} = \delta_{ij}, \quad \{p_i, p_j\} = 0.$

$$\bullet \begin{cases} \{f, q\} = -\frac{\partial f}{\partial p} \\ \{f, p\} = \frac{\partial f}{\partial q} \end{cases} \Rightarrow \begin{cases} \{q, f\} = \frac{\partial f}{\partial p} \\ \{p, f\} = -\frac{\partial f}{\partial q}. \end{cases}$$

- **Jacobi identity:** As an exercise prove the following identity:

$$\{f, \{g, p\}\} + \{p, \{f, g\}\} + \{g, \{p, f\}\} = 0.$$

⁵pb stands for Poisson brackets.

- If O_1 and O_2 are constants of motion, also $\{O_1, O_2\}$ is a constant. In fact, using the Jacobi identity

$$\{H, \{O_1, O_2\}\} + \{O_2, \{H, O_1\}\} + \{O_1, \{O_2, H\}\} = 0$$

which implies $\{H, \{O_1, O_2\}\} = 0$, so $\{O_1, O_2\}$ is a constant of motion.

- On the space of functions $O(p, q)$ the Poisson brackets introduce the structure of an algebra.

1.5.1 Constants of motion

In the previous section we talked about constants of motion and we would like here to give some more details. These are quantities $O(p, q, t)$ which are constant in t once, and only once, we insert in them the classical trajectories for p and q . So a classical constant of motion is of the form $O(p_{\text{cl}}, q_{\text{cl}}(t), t)$ and not of the form $O(p(t), q(t), t)$ where $p(t)$ $q(t)$ are generic trajectories and not classical ones. In fact if O does not depend on t explicitly we do the following steps to check whether it is a constant of motion or not: start from $O(p(t), q(t))$, do the time derivative:

$$\frac{dO}{dt} = \frac{\partial O}{\partial q} \frac{\partial q}{\partial t} + \frac{\partial O}{\partial p} \frac{\partial p}{\partial t}.$$

Now we insert in place of $\frac{\partial q}{\partial t}$ and $\frac{\partial p}{\partial t}$ the Hamiltonian equations and we get

$$\frac{\partial O}{\partial q} \frac{\partial q}{\partial t} + \frac{\partial O}{\partial p} \frac{\partial p}{\partial t} = \frac{\partial O}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial O}{\partial p} \frac{\partial H}{\partial q}.$$

This step implies that the $q(t)$ and $p(t)$ in O are classical trajectories because they satisfy the equations of motion. This means that if we calculate O along a classical trajectory (see Fig. 1) at different instants of times, we get the same constant quantity, for example 25. This constant changes if we calculate it along a different classical trajectory:

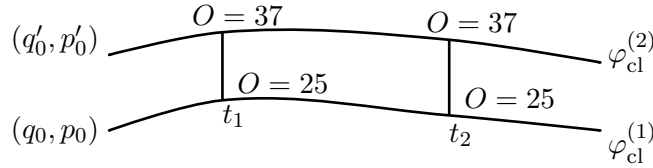


Fig. 1.5: Values of a constant along different classical trajectories.

For example it is a different number like for example 37, but it remains the same along this different trajectory $\varphi_{\text{cl}}^{(2)}$. One could ask why does it change the number in passing from one classical trajectory to another. The reason is the following: the only constant things we have on a trajectory are its initial positions and momenta q_0, p_0 and it is possible to prove that any other constant of motion $O(q(t), p(t))$ can be reduced to a function $\tilde{O}(q_0, p_0)$ of the initial positions and momenta

$$O(p_{\text{cl}}, q_{\text{cl}}) = \tilde{O}(q_0, p_0).$$

This happens at least for those systems which are called “exactly integrable” (on which we will give more details later in the course). Because of (1.5.1) it is clear that if we change the initial conditions (q_0, p_0) the value of $\tilde{O}(q_0, p_0)$ changes and as a consequence also the value of $O(p_{\text{cl}}(t), q_{\text{cl}}(t))$. This explains Fig. 1.5.

Let us instead consider a non-classical trajectory like in the figure below.

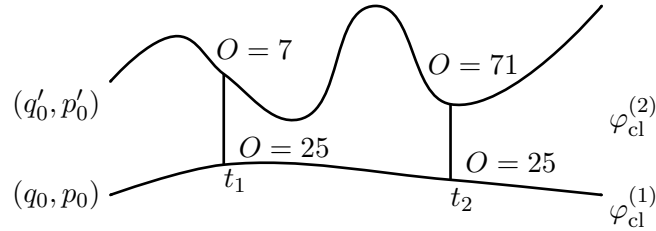


Fig. 1.6: Comparison between a classical and a non-classical trajectory.

The *non-classical* trajectory, indicated with $\varphi_{\text{non-cl}}$, which starts from (q'_0, p'_0) , has different values along its evolution: for example $O = 7$ at $t = t_1$, and $O = 22$ at time t_2 and $O = 71$ at time t_3 , while on the classical trajectory it is always $O = 25$. The non-classical trajectories play a role in quantum mechanics as we will see at the end of this course.

1.6 Canonical transformations

Usually one considers, at the *Lagrangian level*, only the point canonical transformations: $Q = Q(q, t)$ like for example a change of coordinates or similar. At the phase space level the number of coordinates is doubled (p, q) and so we could consider a generalized change of variables of the form

$$\begin{cases} Q = Q(q, p, t) \\ P = P(q, p, t). \end{cases} \quad (1.6.1)$$

The point is that we would like that the equations of motion retain the same form, i.e.

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} = -\frac{\partial H}{\partial q} \end{cases} \implies \begin{cases} \dot{Q} = \frac{\partial H(q(Q, P), p(Q, P))}{\partial P} \\ \dot{p} = -\frac{\partial H(q(Q, P), p(Q, P))}{\partial Q}. \end{cases} \quad (1.6.2)$$

This is not the feature of all set of transformations of the form (1.6.1) but only of a subset called *canonical transformations*. Actually we could be a little more *flexible* and require not that the new H is the old one with q and p replaced by their expression in terms of Q, P , but that is some function $K(Q, P)$

$$\begin{cases} \dot{q} = \frac{\partial H}{\partial p} \\ \dot{p} = -\frac{\partial H}{\partial q} \end{cases} \implies \begin{cases} \dot{Q} = \frac{\partial K}{\partial P} \\ \dot{p} = -\frac{\partial K}{\partial Q}. \end{cases} \quad (1.6.3)$$

Let us see which feature those transformations have. Let us remember that the two sets of equations can be derived from the variational principles

$$\delta \int_{t_1}^{t_2} dt (p\dot{q} - H) = 0, \quad \delta \int_{t_1}^{t_2} dt (P\dot{Q} - K) = 0.$$

In order to describe the same physics it is not necessary that the two integrands are the same but that they differ at most by a total derivative of a function F of the end points which are not varied (so of q, Q)

$$P\dot{Q} - K + \frac{dF(q, Q)}{dt} = p\dot{q} - H. \quad (1.6.4)$$

In that manner

$$\delta \left[\int_{t_1}^{t_2} dt \frac{dF}{dt} \right] = \delta [F(q_2, Q_2) - F(q_1, Q_1)]$$

but the last variation is zero and the total derivative does not contribute to the variation. From (1.6.4) we get

$$P\dot{Q} - K + \frac{\partial F}{\partial t} + \frac{\partial F}{\partial q} \dot{q} + \frac{\partial F}{\partial Q} \dot{Q} = p\dot{q} - H.$$

Since the old and new coordinates are independent, this equation holds if the coefficients of \dot{q} and \dot{Q} vanish

$$\begin{cases} \frac{\partial F}{\partial q} = p \\ \frac{\partial F}{\partial Q} = -P. \end{cases} \quad (1.6.5)$$

The other terms give:

$$H = K - \frac{\partial F}{\partial t}. \quad (1.6.6)$$

So one sees that K is not just H with q and p replaced by their expression in terms of Q and P , but it can have an extra piece. If we take F non explicitly dependent on t then we get $H = K$, which means that $K(Q, P)$ is obtained from H replacing q and p with their expression in terms of Q, P . $F(q, Q)$ is called the *generating function* of the canonical transformation. This is not the only type of generating function: we could build others depending on other sets of variables, like q and P . We can proceed as follows: start from (1.6.4) and write it in the form:

$$PdQ - Kdt + dF = pdq - Hdt. \quad (1.6.7)$$

Let us add to F the expression PQ we get from (1.6.7)

$$d(F + PQ) = pdq + QdP + (K - H)dt.$$

This means that the LHS is effectively a function Φ of q and P because its differential depends only on q and P so it can be written as

$$d\Phi(q, P, t) = pdq + QdP + (K - H)dt$$

which brings to

$$\frac{\partial\Phi}{\partial q}dq + \frac{\partial\Phi}{\partial P}dP + \frac{\partial\Phi}{\partial t}dt = pdq + QdP + (K - H)dt.$$

So the transformations are

$$\begin{cases} \frac{\partial\Phi}{\partial q} = p \\ \frac{\partial\Phi}{\partial P} = Q \\ \frac{\partial\Phi}{\partial t} = K - H. \end{cases} \quad (1.6.8)$$

Φ is another generating function different from F . In the literature Φ is also indicated with $F_2(q, P)$. All together, proceeding in the same manner, we can find two other types of generating functions $F_3(p, Q)$ and $F_4(p, P)$. Their manner to generate the “new” variables is different from (1.6.4) and (1.6.7). **Exercise** Find out the analog expressions to (1.6.4) and (1.6.7) for F_3 and F_4 .

1.6.1 Example of canonical transformations

Let us consider $F_1(q, Q) = qQ$. We get from (1.6.5) and (1.6.6)

$$\begin{aligned} \frac{\partial F}{\partial q} = p &\Rightarrow Q = p \\ \frac{\partial F}{\partial Q} = -P &\Rightarrow P = -q \\ H(p, q) &= K(Q, P). \end{aligned}$$

So one see that this canonical transformation exchange q with p ; what in one coordinate system was q becomes the momentum in the new system and vice versa. So, instead of talking about position and momenta we talk about *conjugate* variables.

1.6.2 Identity transformation

A generating functional of the type $F_2(q, P)$ can be used to generate the identity transformation. Take $F_2(q, P) = qP$ then from (1.6.8) we get

$$p = \frac{\partial F_2}{\partial q} = P, \quad Q = \frac{\partial F_2}{\partial P} = q.$$

A simple generalization of this is the infinitesimal transformation

$$F_2(q, P) = qP + \epsilon G(q, P)$$

with ϵ an infinitesimal parameter. The transformations analogous of (1.6.7) gives

$$\begin{cases} p = P + \epsilon \frac{\partial G}{\partial q}(q, P) \\ Q = q + \epsilon \frac{\partial G}{\partial P}(q, P) \end{cases}$$

which can be written, to order ϵ , as:

$$\begin{cases} P = p - \epsilon \frac{\partial G}{\partial q}(q, p) \\ Q = q + \epsilon \frac{\partial G}{\partial p}(q, p). \end{cases} \quad (1.6.9)$$

In (1.6.9) we have put in the argument of G the small p because we consider things at the first order in ϵ . (1.6.9) can also be written as

$$\begin{cases} P = p - \epsilon \frac{\partial G}{\partial q}(q, p) = p - \epsilon \{G, p\}_{pb} \\ Q = q + \epsilon \frac{\partial G}{\partial p}(q, p) = q - \epsilon \{G, q\}_{pb}. \end{cases} \quad (1.6.10)$$

From this we immediately derive that the time evolution is a particular canonical transformation, in fact take Eq. (1.4.14)

$$\begin{cases} \dot{q} = \{q, H\} \\ \dot{p} = \{p, H\} \end{cases} \Rightarrow \begin{cases} dq = dt\{q, H\} = -dt\{H, q\} \\ dp = dt\{p, H\} = -dt\{H, p\} \end{cases}$$

which implies

$$\begin{cases} q(t + dt) - q(t) = -dt\{H, q\} \\ p(t + dt) - p(t) = -dt\{H, p\} \end{cases} \Rightarrow \begin{cases} q(t + dt) = q(t) - dt\{H, q\} \\ p(t + dt) = p(t) - dt\{H, p\} \end{cases} \quad (1.6.11)$$

If we compare (1.6.11) with (1.6.10) we see that we can identify ϵ with dt and G with H . So the time evolution for (1.6.11) is a canonical transformation, like (1.6.10).

1.6.3 Invariance of the pb under canonical transformations

In what follows we will prove that the Poisson brackets of two observables O_1 and O_2 are independent of which canonical variables we use to evaluate the pb, i.e.:

$$\{O_1, O_2\}_{pb}^{(q,p)} = \{O_1, O_2\}_{pb}^{(Q,P)}. \quad (1.6.12)$$

Let us restrict ourselves to *time-independent* canonical transformations, i.e.

$$\begin{cases} Q = Q(q, p) \\ P = P(q, p) \end{cases}$$

or the inverse

$$\begin{cases} q = q(Q, P) \\ p = p(Q, P). \end{cases}$$

Let us first introduce the symplectic matrix ω^{ab} . Let us introduce a compact notation for the phase space variables $\varphi^a = (q^1, \dots, q^n, p^1, \dots, p^n)$, where $a = 1 \dots 2n$. Then the Hamilton equations can be written as:

$$\dot{\varphi}^a = \omega^{ab} \frac{\partial H}{\partial \varphi^b}. \quad (1.6.13)$$

Let us now perform a canonical transformation independent of t

$$\varphi'^a = \varphi'^a(\varphi)$$

As it is independent of t , we know that the new Hamiltonian is

$$K(\varphi') = H(\varphi(\varphi')).$$

First of all let us prove that *the symplectic matrix is left invariant by the canonical transformation* $\varphi'^a = \varphi'^a(\varphi)$. As the transformation is canonical we must have the same equations of motion in the transformed variables:

$$\dot{\varphi}'^a = \omega^{ab} \frac{\partial H}{\partial \varphi'^b}. \quad (1.6.14)$$

The LHS of (1.6.14) can be rewritten as:

$$\dot{\varphi}'^a = C_b^a \dot{\varphi}^b.$$

where C_b^a is the following matrix:

$$C_b^a \equiv \frac{\partial \varphi'^a}{\partial \varphi^b}.$$

Using (1.6.13) we get:

$$\dot{\varphi}'^a = C_b^a \omega^{bd} \frac{\partial H}{\partial \varphi^d}. \quad (1.6.15)$$

Now

$$\frac{\partial H}{\partial \varphi^d} = \frac{\partial H}{\partial \varphi'^a} \frac{\partial \varphi'^a}{\partial \varphi^d} = \frac{\partial H}{\partial \varphi'^a} C_d^a = C^T \frac{\partial H}{\partial \varphi'}.$$

Replacing the previous equation into (1.6.15) we get

$$\dot{\varphi}'^a = C \omega C^T \frac{\partial H}{\partial \varphi'}. \quad (1.6.16)$$

Comparing this with (1.6.14) we get

$$\omega = C \omega C^T \Rightarrow C^{-1} \omega (C^T)^{-1} = \omega. \quad (1.6.17)$$

Let us now see the Poisson brackets of the observables $O_1 O_2$

$$\begin{aligned} \{O_1, O_2\}_{\text{pb}}^{(q,p)} &= \frac{\partial O_1}{\partial q} \frac{\partial O_2}{\partial p} - \frac{\partial O_1}{\partial p} \frac{\partial O_2}{\partial q} \\ &= \frac{\partial O_1}{\partial \varphi^a} \omega^{ab} \frac{\partial O_2}{\partial \varphi^b}. \end{aligned}$$

Analogously:

$$\{O_1, O_2\}_{\text{pb}}^{(Q,P)} = \frac{\partial O_1}{\partial \varphi'^a} \omega^{ab} \frac{\partial O_2}{\partial \varphi'^b}. \quad (1.6.18)$$

Now

$$\begin{aligned}\frac{\partial O_1}{\partial \varphi'^a} &= \frac{\partial O_1}{\partial \varphi^c} \frac{\partial \varphi^c}{\partial \varphi'^a} = \frac{\partial O_1}{\partial \varphi^c} (C^{-1})^c_a \\ \frac{\partial O_2}{\partial \varphi'^b} &= \frac{\partial O_2}{\partial \varphi^k} \frac{\partial \varphi^k}{\partial \varphi'^b} = \frac{\partial O_2}{\partial \varphi^k} (C^{-1})^k_b = (C^{-1})^T \frac{\partial O_2}{\partial \varphi}.\end{aligned}$$

So replacing in (1.6.18) we have

$$\{O_1, O_2\}_{\text{pb}}^{(Q,P)} = \frac{\partial O_1}{\partial \varphi} C^{-1} \omega (C^{-1})^T \frac{\partial O_2}{\partial \varphi}.$$

Using (1.6.17) we get:

$$\{O_1, O_2\}_{\text{pb}}^{(Q,P)} = \frac{\partial O_1}{\partial \varphi} \omega \frac{\partial O_2}{\partial \varphi} = \{O_1, O_2\}_{\text{pb}}^{(q,p)}$$

which proves (1.6.12).

Now the equations of motion can be written via Poisson brackets

$$\frac{dq}{dt} = \{q, H\} \quad \frac{dp}{dt} = \{p, H\}$$

and the RHS can be calculated in any system of canonical coordinates, so we can use the best one. Using the invariance we get also that the new canonical variables have the same brackets as the old ones

$$\{Q, P\}^{(q,p)} = \{Q, P\}^{(Q,P)}$$

but the RHS is one so also

$$\{Q, P\}^{(Q,P)} = 1.$$

The same holds for the other Poisson brackets:

$$\begin{aligned}\{Q, Q\}^{(q,p)} &= \{Q, Q\}^{(Q,P)} = 0 \\ \{P, P\}^{(q,p)} &= \{P, P\}^{(Q,P)} = 0.\end{aligned}$$

1.7 Liouville theorem

Let us take a volume Γ in phase space

$$\int_{\Gamma} \prod_{a=1}^{2n} d\varphi^a. \quad (1.7.1)$$

We will now prove that this volume remains invariant under a *canonical* change of variables

$$\varphi^a \longrightarrow \varphi'^a.$$

We have proved before, Eq. (1.6.17) that the symplectic matrix ω does not change as follows under a canonical transformation, i.e.:

$$\omega = C \omega C^T \quad (1.7.2)$$

where C is the matrix given by

$$C_b^a \equiv \frac{\partial \varphi'^a}{\partial \varphi^b}. \quad (1.7.3)$$

If we take the determinant on the right and left hand side of (1.7.2) we have

$$\begin{aligned} \det \omega &= \det(C\omega C^T) = \det C \det \omega \det C^T \\ &= (\det \omega)(\det C)^2 \end{aligned}$$

which implies

$$\det C = \pm 1. \quad (1.7.4)$$

In the above derivation we have used the fact that $\det C^T = \det C$.

Now, if the canonical transformation can be infinitesimally reduced to the identity with a continuous transformation, then $\det C = 1$. In fact, if we *continuously* deform the transformation to the identity, the determinant must also change continuously and as the identity transformation has $\det C = 1$, it means that also our canonical transformation has determinant 1. It cannot have $\det = -1$ otherwise it would have to change discontinuously in approaching the identity from -1 to +1.

Let us now go back to the volume of (1.7.1) $\int_{\Gamma} \prod_{a=1}^{2n} d\varphi^a$ and let us make a *canonical* change of variables

$$\varphi^a \rightarrow \varphi'^a(\varphi).$$

In the new variables φ'^a the volume becomes $\int_{\Gamma'} \prod_{a=1}^{2n} d\varphi'^a$. If we write this in terms of the old variables φ we get (Γ' is the transformed surface surrounding the volume)

$$\int_{\Gamma'} \prod_{a=1}^{2n} d\varphi'^a = \int_{\Gamma} \prod_{a=1}^{2n} d\varphi^a \left| \frac{\partial \varphi'^b}{\partial \varphi^a} \right|, \quad (1.7.5)$$

where $\left| \frac{\partial \varphi'^b}{\partial \varphi^a} \right|$ is the *determinant* of the matrix $\frac{\partial \varphi'^b}{\partial \varphi^a}$. This matrix is the C^T of (1.7.3).

We have proved before that $\det C^T = \det C = 1$. So (1.7.5) reduces to

$$\int_{\Gamma'} \prod_{a=1}^{2n} d\varphi'^a = \int_{\Gamma} \prod_{a=1}^{2n} d\varphi^a \det C = \int_{\Gamma} \prod_{a=1}^{2n} d\varphi^a \quad (1.7.6)$$

that is

$$\int_{\Gamma'} \prod_{a=1}^{2n} d\varphi'^a = \int_{\Gamma} \prod_{a=1}^{2n} d\varphi^a. \quad (1.7.7)$$

This means that phase space volumes are left invariant by canonical transformations. We have proved in (1.6.11) that the Hamiltonian evolution is a canonical transformation, so if the φ' variables in (1.7.7) are just the time-evolved variables with respect to φ

$$\varphi'(t) = \varphi(t + dt)$$

then from (1.7.7) we get:

$$\int_{\Gamma'} \prod_{a=1}^{2n} d\varphi^a(t + dt) = \int_{\Gamma} \prod_{a=1}^{2n} d\varphi^a(t).$$

So we can say that during the time evolution the volume can change its shape (from Γ to Γ') but not its total value. This is the Liouville theorem.

Another form of the Liouville theorem is the following one: let us suppose we introduce a “probability density in phase space” $\rho(\varphi^a)$. Let us suppose we evaluate

$$\int \rho(\varphi^a) \prod_{a=1}^{2n} d\varphi^a. \quad (1.7.8)$$

Next let us change variables from $\varphi^a(t)$ to $\varphi^a(t + dt)$

$$\int \rho(\varphi^a(t + dt)) \prod_{a=1}^{2n} d\varphi^a(t + dt).$$

As the variables are integrated over, we can change them without changing the value of the integral (1.7.8) so

$$\int \rho(\varphi^a(t)) \prod_{a=1}^{2n} d\varphi^a(t) = \int \rho(\varphi^a(t + dt)) \prod_{a=1}^{2n} d\varphi^a(t + dt).$$

As the volume is invariant we can write this as

$$\int \rho(\varphi^a(t)) \prod_{a=1}^{2n} d\varphi^a(t) = \rho(\varphi^a(t + dt)) \prod_{a=1}^{2n} d\varphi^a(t).$$

This means

$$\int [\rho(\varphi^a(t)) - \rho(\varphi^a(t + dt))] \prod_{a=1}^{2n} d\varphi^a = 0$$

which implies that

$$\rho(\varphi^a(t)) = \rho(\varphi^a(t + dt))$$

or

$$\frac{d\rho}{dt} = 0.$$

This can be written as

$$\begin{aligned} \frac{d\rho}{dt} &= \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial q} \frac{\partial q}{\partial t} + \frac{\partial \rho}{\partial p} \frac{\partial p}{\partial t} \\ &= \frac{\partial \rho}{\partial t} + \frac{\partial \rho}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial \rho}{\partial p} \frac{\partial H}{\partial q} = 0. \end{aligned}$$

This equation can be written as

$$\frac{\partial \rho}{\partial t} = \left(\frac{\partial H}{\partial q} \frac{\partial}{\partial p} - \frac{\partial H}{\partial p} \frac{\partial}{\partial q} \right) \rho$$

and is called the Liouville equation. It can also be written as

$$\frac{\partial \rho}{\partial t} = i\hat{L}\rho$$

where

$$\hat{L} \equiv i \left(\frac{\partial H}{\partial p} \frac{\partial}{\partial q} - \frac{\partial H}{\partial q} \frac{\partial}{\partial p} \right)$$

is called the *Liouville operator*.

1.8 Symmetries and their generators (Hyper-simplified treatment)

A symmetry is a set of transformations of q and p which leaves the equations of motion invariant. Associated to any symmetry of the system there is a *conserved quantity* (Noether theorem). Let us work out some simple examples. Suppose our Hamiltonian H does not depend on one of the generalized coordinates q_1, \dots, q_k , suppose q_3 . Then, if we do an infinitesimal transformation of q_3 , H does not change. That implies that the equations of motion remain invariant under an infinitesimal transformation of q_3 . The transformation is $\delta q_3 = \epsilon$. This can be put into the form (1.6.10) of a canonical transformation

$$\delta q_3 = \epsilon \{q_3, G\}_{pb}.$$

In our example G is nothing else than p_3 and it is called the generator of the transformation. We will now show that the conserved quantity under this symmetry is G , i.e. p_3 . In fact we have

$$\frac{\partial H}{\partial q_3} = 0$$

that can be written as

$$\frac{\partial H}{\partial q_3} = \{H, p_3\} = 0.$$

Now

$$\frac{dp_3}{dt} = \frac{\partial p_3}{\partial t} + \{p_3, H\} = 0$$

so p_3 is conserved.

In general we can prove that any generator of a symmetry transformation is conserved. The proof goes as follows: if $\delta\varphi = \epsilon\{\varphi, G\}$ is a symmetry it means that $\delta H = 0$ under that transformation i.e.

$$\begin{aligned} \frac{\partial H}{\partial \varphi} \delta\varphi = 0 &\implies \frac{\partial H}{\partial \varphi} \epsilon\{\varphi, G\} = 0 \implies \\ \epsilon \frac{\partial H}{\partial \varphi^a} \{\varphi^a, G\} &= \epsilon \frac{\partial H}{\partial \varphi^a} \omega^{ab} \frac{\partial G}{\partial \varphi^b} = \epsilon \{H, G\} = 0, \end{aligned}$$

which implies:

$$\frac{dG}{dt} = \frac{\partial G}{\partial t} + \{G, H\} = 0$$

Examples of generators. The generator of the translation in q is p . The generators of rotations are the angular momenta. The generator of the time translations is H . The first statement above was proved above. The third one was also prove, see Eq. (1.6.11). From this we derive that the conservation of energy is related to the symmetry of translation in time.

Let us now proceed to prove the second statement. Let us do a rotation of $d\theta$ along the z axis. We get

$$\begin{cases} \delta x = -y d\theta \\ \delta y = x d\theta \\ \delta z = 0 \end{cases} \quad (1.8.1)$$

$$\begin{cases} \delta p_x = -p_y d\theta \\ \delta p_y = p_x d\theta \\ \delta p_z = 0. \end{cases} \quad (1.8.2)$$

Let us now see if we can find a generator G such that (1.8.1) and (1.8.2) can be written as

$$\delta\varphi = d\theta\{\varphi, G\}.$$

It is easy to see that $G = xp_y - yp_x$, which is the angular momentum along z .

Problem. Find the generator of the boosts in the Galilean transformation.

Another important symmetry is given by the charge conservation which is implied by the gauge invariance.

Note: Most of the time the symmetries are not related to the fact that some variables are missing from H , like for q_3 in our first case. Most of the times there are combinations of variables like $(q_3 - q_2)$ which are missing and this kind of symmetries are more difficult to detect.

Chapter 2

Crisis of Classical Physics

2.1 Introduction

The interplay between theories and experiments in physics, as well as in all other fields of science, is the following:

1. The need to go beyond an established theory starts when such a theory cannot explain or justify some experimental data;
2. In this case it is necessary to look for a *new* theory able to explain the phenomena which cannot be set in the framework of the *old* theory;
3. The new theory must also predict new experimental facts;
4. Then such new experimental facts must be tested in laboratories.

Usually a lot of experimental data are already explained by old theories. Therefore it is crucial for the new theories to include the old ones as limiting cases. For example:

- 1) non-relativistic Newton's mechanics can be seen as a limiting case of relativistic mechanics when the speeds involved are very small compared with the speed of light, i.e. $v \ll c$;
- 2) geometrical optics can be seen as a limiting case of optics when the wavelengths involved are much smaller than the dimensions of the equipment used for their study.

The situation of physics at the end of the 19th century was more or less the following: there were three well-established theories, Newton's mechanics, statistical mechanics and thermodynamics, Maxwell's electromagnetism, which were able to describe a huge amount of experimental data and there was a clear distinction between the theory of particles on one side and the theory of waves on the other one. For example in those years experiments of diffractions by crystals proved the wavelike nature of x -rays. The particlelike nature of electrons instead emerged by the analysis of their trajectories in electric and magnetic fields which led Thompson to the well-known measurement of the ratio e/m between the electric charge and the mass of the electron. Nevertheless the problems began at the end of the 19th century when physicists realized that some new phenomena could not be explained by the three theories mentioned above. The main set

of these phenomena concerned the interaction of matter with radiation, e.g. the black body spectrum and the emission and absorption of radiation from atoms.

2.2 Black Body Radiation

Every body at temperature T emits radiation. The energy emitted per unit of time and surface, within a cone of solid angle $d\Omega$ whose axis forms an angle θ with the normal to the surface, and in the interval of frequencies $(\nu, \nu + d\nu)$, is given by $e(\nu, T, x) \cos\theta d\Omega d\nu$ where $e(\nu, T, x)$ is called the *rate of emission* of the body. Such a quantity changes from point to point of the body and depends on some parameters x of the body (material, form, internal structure, etc.). One can define also the *rate of absorption* of the body

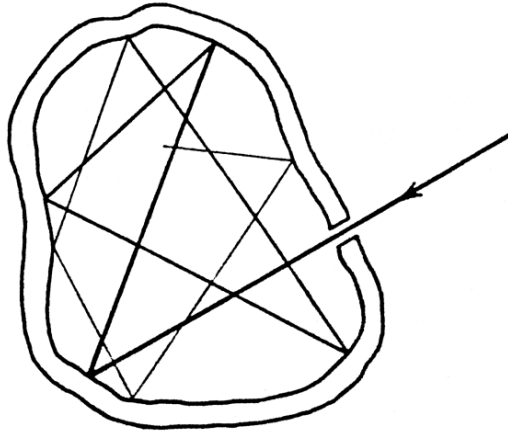


Fig. 2.1: A cavity in a body with a small hole. The hole emits like a black body

$a(\nu, T, x)$ as the ratio between the energy absorbed by the body (in a fixed interval of time and frequency and per unit of surface) and the associated incident energy. Also $a(\nu, T, x)$ depends on the parameters x of the body and from the definition itself it is easy to realize that $a(\nu, T, x) \leq 1$. The property which defines a black body is that its rate of absorption is just equal to one: $a(\nu, T, x) = 1$, i.e. a black body absorbs all the incident radiation at every frequency ν and at very temperature T . A typical example of a black body is given by a very small hole of a cavity heated to temperature T . We want to stress the fact that the hole itself, and not the cavity, has the property of being a black body.

From thermodynamics Kirchoff in 1859 proved that the ratio

$$\frac{e(\nu, T, x)}{a(\nu, T, x)} = \frac{c}{4\pi} u(\nu, T) \quad (2.2.1)$$

is a universal function, i.e. it does not depend on the variable x which is linked to particular features of the body. In particular in the case of a black body $a = 1$ and therefore Eq. (2.2.1) becomes

$$e(\nu, T) = \frac{c}{4\pi} u(\nu, T). \quad (2.2.2)$$

This equation tells us that the rate of emission of a black body $e(\nu, T)$ does not depend on the features x of the body and therefore can be identified with the universal function $u(\nu, T)$, modulo the conventional factor $\frac{c}{4\pi}$. If we heat the body, it will emit radiation via the hole and what will escape from the cavity is just the radiation inside which is in equilibrium with the walls. If we integrate the RHS of (2.2.2) over all the emission angles we find out the total energy irradiated from the hole per unit surface and frequency at temperature T :

$$E(\nu, T) = 2\pi \int_0^{\pi/2} d\theta \sin\theta \cos\theta \frac{c}{4\pi} u(\nu, T) = \frac{c}{4} u(\nu, T). \quad (2.2.3)$$

The plot of E as a function of the wave length $\lambda = c/\nu$ is given in Fig. 2.2 for different values of the temperature T . As it is clear from Fig. 2.2 the maximum in the emission of

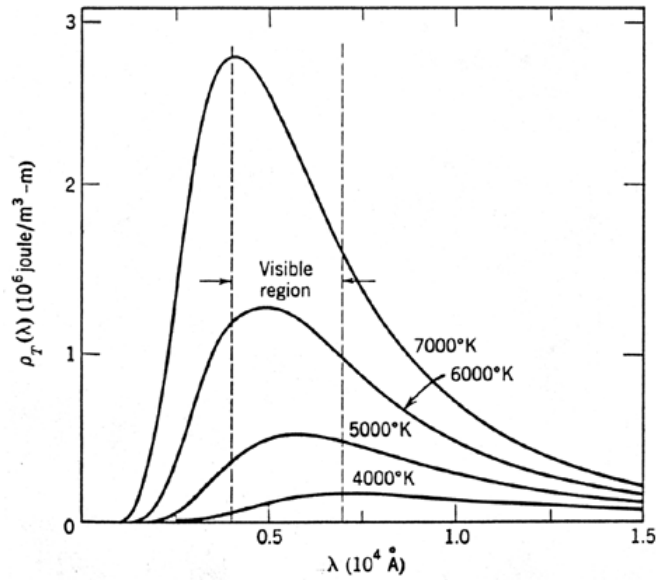


Fig. 2.2: Plot of the energy E emitted from the hole for different values of the temperature T .

the black body corresponds to a particular value of the wavelength λ_{\max} which changes with the temperature T . The position of this maximum obeys the following Wien's displacement law (which is a phenomenological law):

$$\lambda_{\max} T = \text{const} = 0.290 \text{ cm} \cdot \text{K}. \quad (2.2.4)$$

Therefore the hotter the black body is, the smaller the wavelength of the maximum is. This means that, by increasing the temperature, the colour of the black body shifts from red to blue.

The total energy irradiated from the hole per unit surface and time will be given by the integral of the function $E(\nu, T)$ over all the frequencies ν :

$$\int_0^{\infty} d\nu E(\nu, T) = \sigma T^4, \quad \sigma = 5.66 \cdot 10^{-5} \text{ erg} \cdot \text{cm}^{-2} \cdot \text{sec}^{-1}. \quad (2.2.5)$$

This is the Stefan-Boltzmann's law which was derived experimentally by Stefan in 1879 and derived from the thermodynamical laws five years later by Boltzmann. The Stefan-Boltzmann's law can be derived from the following phenomenological law discovered by Wien:

$$E(\nu, T) = \nu^3 F\left(\frac{\nu}{T}\right). \quad (2.2.6)$$

Even if the explicit form of the function F is not known the Stefan-Boltzmann's law (2.2.5) can be derived from (2.2.6). In fact:

$$\int_0^\infty d\nu E(\nu, T) = \int_0^\infty d\nu \nu^3 F\left(\frac{\nu}{T}\right) \quad (2.2.7)$$

and if we perform the following change of variables $\nu/T = x$ we obtain, for the total energy, the correct dependence on T :

$$\int d\nu E(\nu, T) = T^4 \int_0^\infty dx x^3 F(x) = T^4 \sigma. \quad (2.2.8)$$

2.3 Classical Derivation of the Black Body Radiation

If we want to derive from classical statistical mechanics the spectrum of a black body we have to calculate the energy density $E(\nu, T)$ within a cavity heated to temperature T . Since this energy density cannot depend on x we can choose a cubic cavity with metallic walls, like the one of Fig. 2.3. The thermal motion of the electrons of the walls causes

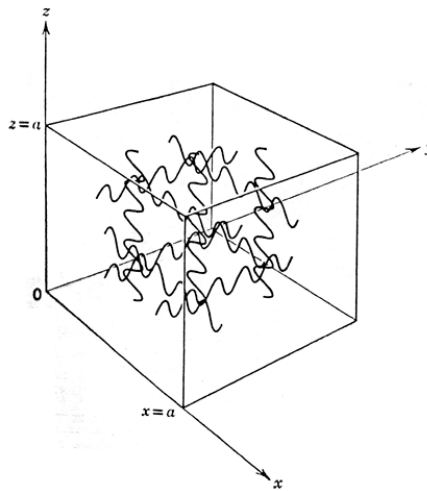


Fig. 2.3: A metallic walled cubic cavity filled with electromagnetic radiation.

the emission of electromagnetic waves. In particular the radiation inside the cavity is in form of standing waves. The electric field \mathbf{E} is perpendicular to the propagation direction of the wave and therefore it is parallel to the walls. At the walls \mathbf{E} must be zero because otherwise the flow of charges always neutralizes the electric field.

Let us limit ourselves to the one-dimensional case. The electric field will be described by the following function

$$E(x, t) = E_0 \sin\left(\frac{2\pi x}{\lambda}\right) \sin 2\pi \nu t, \quad \nu = \frac{c}{\lambda}. \quad (2.3.1)$$

Besides the origin, the electric field (2.3.1) is zero at every point x satisfying $\frac{2x}{\lambda} = n$ where $n = 1, 2, 3, \dots$. Now let us suppose the distance between the walls is a ; since $E(a, t)$ must be zero for every time t , the standing waves will be characterized by a wavelength given by $\frac{2a}{\lambda} = n$ or, equivalently, by a frequency $\nu = \frac{c}{\lambda} = \frac{cn}{2a}$.

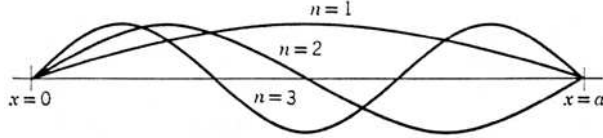


Fig. 2.4: Standing waves in a one-dimensional cavity.

Therefore there is a standing wave for each number $n \in \mathbb{N}$. Now the question is: how many waves are there between the frequencies ν and $\nu + d\nu$? Since $\frac{2a\nu}{c} = n$ in the interval $(\nu, \nu + d\nu)$ there are $N(\nu)d\nu = \frac{2a}{c}d\nu$ waves as it is clear from Fig. 2.5. If we take into account that for each frequency there are two independent polarizations of the electric field, the total number of waves per unit of frequency is given by:

$$N(\nu)d\nu = \frac{4a}{c}d\nu. \quad (2.3.2)$$

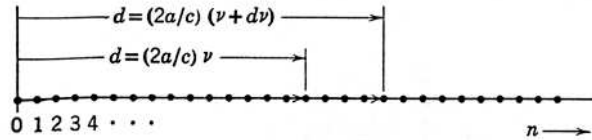


Fig. 2.5: Allowed values of frequency in a one-dimensional cavity.

In three dimensions Eq. (2.3.2) becomes:

$$N(\nu)d\nu = \frac{8\pi a^3}{c^3} \nu^2 d\nu. \quad (2.3.3)$$

Note that in (2.3.3) there appears a crucial factor ν^2 . It arises because, while in one dimension the number of waves with frequency between ν and $\nu + d\nu$ is proportional to the length of the interval $(\nu, \nu + d\nu)$ in three dimensions such a number will be proportional to the volume contained between the shells of radii $\nu + d\nu$ and ν . Now every wave has an energy proportional to the square of the amplitude of the electric field

E_0^2 . Nevertheless since there is a great number of waves in equilibrium at temperature T we have to use the laws of statistical mechanics in order to compute the energy density. In particular the theorem of equipartition of energy tells us that each molecule of a gas in thermal equilibrium at temperature T has an average kinetic energy $\bar{K} = \frac{kT}{2}$ where $k = 1.38 \cdot 10^{-23}$ joule/°K is the Boltzmann's constant. In our case the entities of the ensemble are not molecules of a gas but sinusoidal waves which behave like harmonic oscillators and whose energy is not only kinetic but also potential. Therefore the average total energy per standing wave is given by $\bar{E} = kT$ and the average energy density per unit volume becomes

$$u(\nu, T)d\nu = \frac{\bar{E}N(\nu)}{a^3}d\nu = \frac{8\pi kT}{c^3}\nu^2 d\nu. \quad (2.3.4)$$

Consequently the $E(\nu, T)$ of Eq. (2.2.3) is given by:

$$E(\nu, T) = \frac{c}{4}u(\nu, T) = \frac{2\pi}{c^2}kT\nu^2. \quad (2.3.5)$$

This is the Rayleigh-Jeans law for the black body radiation.

The difference from the experimental data can be appreciated from Fig. 2.6. We can

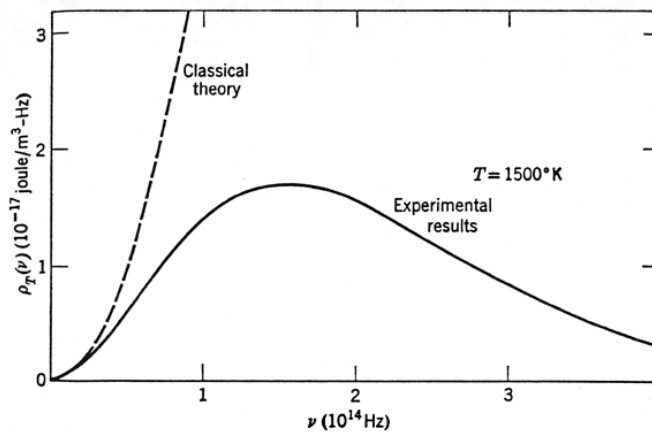


Fig. 2.6: Comparison of the Rayleigh-Jeans law with the experimental data.

see how the Rayleigh-Jeans law agrees with the experimental data for the low frequencies $\nu \approx 0$ but not for the high ones. Moreover if we integrate the $E(\nu, T)$ of (2.3.5) over all the frequencies we obtain $\int d\nu E(\nu, T) = \infty$ for every temperature T instead of the Stefan-Boltzmann's law (2.2.5). This is the so called *ultraviolet catastrophe*. Somehow this indicates that classical mechanics has problems in the high frequency regime which is also the regime of smaller and smaller intervals of time.

2.4 Planck's Hypothesis

A possible way to avoid the ultraviolet catastrophe mentioned above is to assume that the equipartition law is true only for low frequencies, i.e. $\bar{E} \xrightarrow{\nu \rightarrow 0} kT$, while, for higher

frequencies, the average energy of a standing wave varies with the frequency itself. In particular in the limiting case $\nu \rightarrow \infty$ the average energy \bar{E} must tend to zero in order to reproduce the experimental data of Fig. 2.6. Now the equipartition theorem was derived from the Boltzmann or canonical distribution. According to this distribution if a system contains a large number of entities in equilibrium at temperature T then the probability of finding a particular entity at energy E is:

$$p(E) = \frac{e^{-E/kT}}{kT} \quad (2.4.1)$$

from which we derive that the average value of energy is just given by

$$\bar{E} = \frac{\int_0^{\infty} dE E p(E)}{\int_0^{\infty} dE p(E)} = kT. \quad (2.4.2)$$

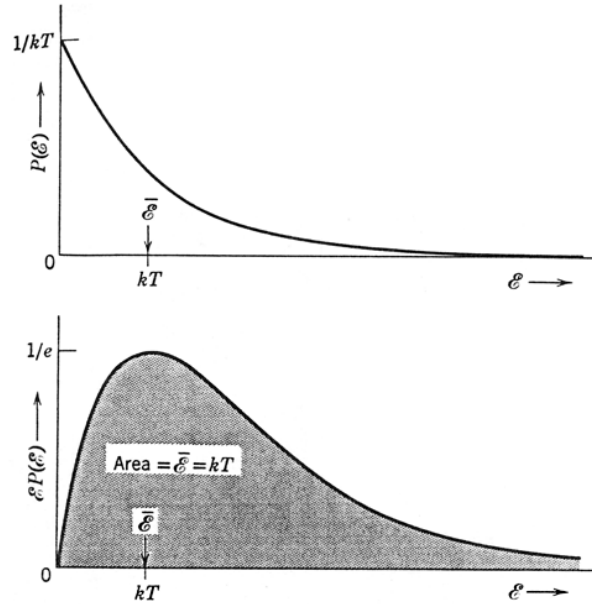


Fig. 2.7: Plot of the Boltzmann's distribution $p(E)$ and of the curve $Ep(E)$.

In December 1900 Planck presented at a meeting of the German Physical Society his idea or trick, as he called it, that solves the ultraviolet catastrophe. He assumed that, in calculating \bar{E} , the energy should have been treated as a discrete and not a continuous variable. In other words E can assume only some discrete values:

$$E = 0, \Delta E, 2\Delta E, \dots, n\Delta E, \dots \quad (2.4.3)$$

If energy is discretized when we calculate its average value we must replace in (2.4.2)

integrals with sums according to the formula:

$$\bar{E} = \frac{\int dE E p(E)}{\int dE p(E)} = \frac{\sum_n n \Delta E p(n \Delta E)}{\sum_n p(n \Delta E)}. \quad (2.4.4)$$

If Δ is sufficiently small, i.e. if $\Delta E \ll kT$, then the area under the curve $E p(E)$ is more or less equal to the sum of the areas of the rectangles $\sum_n n \Delta E p(n \Delta E)$, i.e. the average energy \bar{E} is more or less equal to kT , as predicted by the equipartition theorem, see Fig. 2.8.

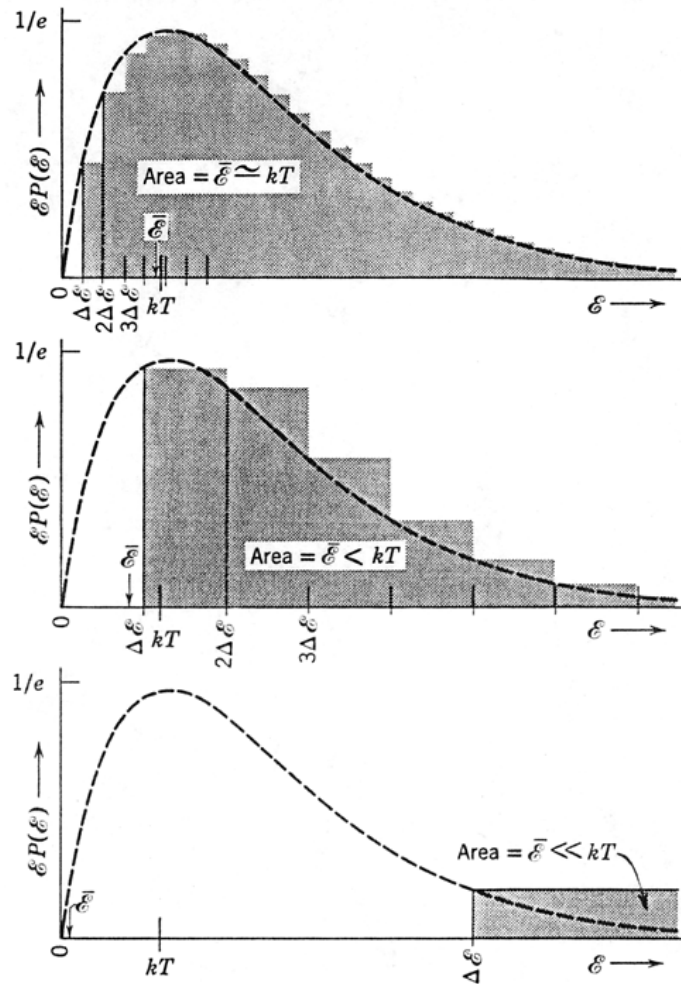


Fig. 2.8: Evaluation of the mean value of energy when energy can assume only discrete values in the three cases $\Delta E \ll kT$, $\Delta E \simeq kT$ and $\Delta E \gg kT$.

If instead the rectangles are larger because $\Delta E \sim kT$ we have that the average value

\bar{E} calculated via the sum is smaller than kT . Planck's choice was to identify $\Delta E = h\nu$ where $h = 6.63 \times 10^{-34}$ joule · sec is called Planck's constant. With this choice the only possible values of energy are given by $E_n = nh\nu$. If we replace Planck's choice into Eq. (2.4.4) we obtain that the average energy of a standing wave is:

$$\bar{E} = \frac{\sum_{n=0}^{\infty} E_n p(E_n)}{\sum_{n=0}^{\infty} p(E_n)} = \frac{\sum_{n=0}^{\infty} \frac{nh\nu}{kT} e^{-nh\nu/kT}}{\sum_{n=0}^{\infty} \frac{1}{kT} e^{-nh\nu/kT}} = kT \frac{\sum_{n=0}^{\infty} n\alpha e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}}. \quad (2.4.5)$$

where we have put $\alpha \equiv \frac{h\nu}{kT}$. Eq. (2.4.5) can be evaluated by noting that

$$-\alpha \frac{d}{d\alpha} \ln \sum_{n=0}^{\infty} e^{-n\alpha} = \frac{-\alpha \frac{d}{d\alpha} \sum_{n=0}^{\infty} e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}} = \frac{\sum_{n=0}^{\infty} n\alpha e^{-n\alpha}}{\sum_{n=0}^{\infty} e^{-n\alpha}}. \quad (2.4.6)$$

Therefore

$$\bar{E} = kT \left(-\alpha \frac{d}{d\alpha} \ln \sum_{n=0}^{\infty} e^{-n\alpha} \right) = -h\nu \frac{d}{d\alpha} \ln \sum_{n=0}^{\infty} e^{-n\alpha}. \quad (2.4.7)$$

The series in (2.4.7) is the geometric one and therefore it can be evaluated explicitly:

$$\sum_{n=0}^{\infty} e^{-n\alpha} = 1 + e^{-\alpha} + (e^{-\alpha})^2 + \dots = \frac{1}{1 - e^{-\alpha}}. \quad (2.4.8)$$

By replacing (2.4.8) into (2.4.7) we obtain, for the average energy, the following *distribution* which is called Planck distribution:

$$\bar{E} = -h\nu \frac{d}{d\alpha} \ln \frac{1}{1 - e^{-\alpha}} = \frac{h\nu}{e^{h\nu/kT} - 1}. \quad (2.4.9)$$

The energy density per unit of frequency would be given by

$$E(\nu, T) d\nu = \frac{c}{4} \frac{\bar{E} N(\nu)}{a^3} d\nu = \frac{2\pi}{c^2} \frac{h\nu^3}{e^{h\nu/kT} - 1} d\nu \quad (2.4.10)$$

or, using the wavelengths instead of the frequencies:

$$E(\lambda, T) d\lambda = \frac{2\pi hc^2}{\lambda^5} \frac{1}{e^{hc/\lambda kT} - 1} d\lambda \quad (2.4.11)$$

which fits exactly with the experimental data plotted in Fig. 2.2.

In a letter to R. W. Wood Planck called his act of discretizing energy “*an act of despair*” and for ten years he tried to set his idea within the framework of classical physics. In that same letter he wrote: “*I knew that the problem (of the equilibrium of matter and radiation) is of fundamental significance for physics; I knew the formula that*

reproduces the energy distribution in the normal spectrum; a theoretical interpretation had to be found at any cost, no matter how high.”

Homework: exercise 1. Prove that for $\nu \rightarrow 0$ Eq. (2.4.9) reproduces the equipartition law $\bar{E} = kT$ and that $\bar{E} \xrightarrow{\nu \rightarrow \infty} 0$.

Homework: exercise 2. Derive the Stefan’s law (2.2.5) and the Wien displacement’s law (2.2.4) from the Planck’s distribution (2.4.11).

Problem 1. A pendulum is made up of a mass $m = 0.01$ Kg hanged from a string of length $l = 0.1$ m. The amplitude of an oscillation is such that the string in its starting position forms an angle $\theta = 0.1$ rad with the vertical direction. The energy of the pendulum decreases because of the friction. Does the energy decrease continuously or not?

Solution. The oscillation frequency of the pendulum is:

$$\nu = \frac{1}{2\pi} \sqrt{\frac{g}{l}} = \frac{1}{2\pi} \sqrt{\frac{9.8 \text{ m/sec}^2}{0.1 \text{ m}}} = 1.6 \text{ sec}^{-1}. \quad (2.4.12)$$

The energy of the pendulum is its initial potential energy:

$$\begin{aligned} E &= mgl(1 - \cos \theta) = 0.01 \text{ Kg} \cdot 9.8 \text{ m/sec}^2 \cdot 0.1 \text{ m} \cdot (1 - \cos 0.1) \\ &= 5 \cdot 10^{-5} \text{ joule}. \end{aligned} \quad (2.4.13)$$

The gap between two consecutive energy levels is given by

$$\Delta E = h\nu = 6.63 \cdot 10^{-34} \text{ joule} \cdot \text{sec} \cdot 1.6 \text{ sec}^{-1} = 10^{-33} \text{ joule}. \quad (2.4.14)$$

Therefore the ratio between ΔE and E is $\frac{\Delta E}{E} = 2 \cdot 10^{-29}$. This means that in order to appreciate the discreteness of the quantum jumps we should measure the energy with a precision of at least two parts in 10^{29} . None of the most precise measurement instruments can give such a resolution. Therefore we cannot determine with a pendulum whether the Planck law is true or not. It is necessary to study regimes where $E \sim \Delta E$, i.e. $E \sim h\nu$. This will happen for very high frequencies ν , i.e. for very small wavelengths $\lambda = c/\nu$ (order of magnitude 10^{-8} cm).

2.5 Photoelectric Effect

Hertz, who had already shown the *electromagnetic* nature of light, performed, together with Lenard, the experiment which proved that light can have also a *particlelike* nature. Hertz realized that between the two electrodes of Fig. 2.10 a jump spark is more likely if we send ultraviolet radiation over the electrode *A*. This is so because, when the electromagnetic radiation arrives at the plate *A* of the electrode, it causes the emission of electrons from the surface. Those electrons are attracted towards *B* by the potential difference V present between *A* and *B*. The ammeter *G* will measure a current which is called *photoelectric current* and which has the features shown in Fig. 2.9.

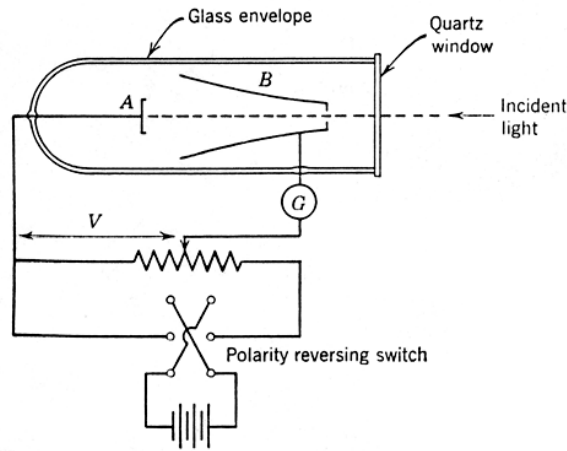


Fig. 2.9: The apparatus used to study the photoelectric effect.

By increasing the potential we soon obtain a *saturation current*, which means that all the electrons emitted from A reach B. If we change the sign of the potential V the current I does not tend to zero immediately. This suggests that the electrons emitted from the surface of A have a *kinetic energy* and some of them reach B even if there is a potential against them. If we increase the potential furthermore then the current becomes zero, i.e. all the electrons stop. The particular value of the potential V_0 which produces this situation is called *stopping potential*. If we multiply this quantity by e we can measure the *maximum kinetic energy* of the electrons:

$$K_{\max} = eV_0. \quad (2.5.1)$$

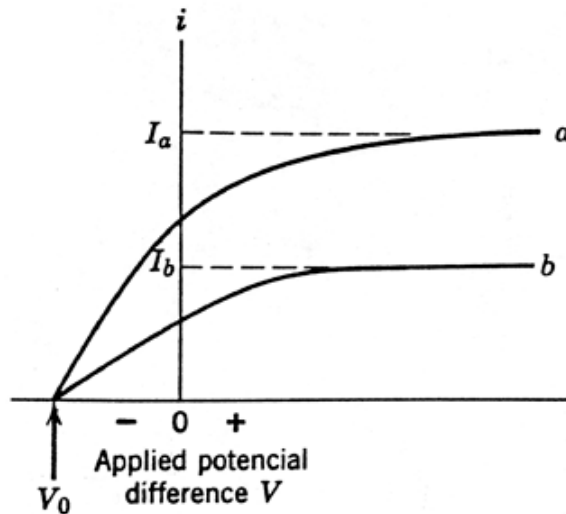


Fig. 2.10: Intensity of the photoelectric current as a function of the potential V .

In Fig. 2.10 we have plotted two different curves which correspond to two different intensities of the incident light. From this figure we can see that the value of the stopping potential V_0 is independent of the intensity of the radiation. Millikan, after having noticed that K_{\max} , or equivalently V_0 , did not depend on the *intensity*, tried to see whether it depended on the *frequency* of the incident radiation and he obtained the graphic plotted in Fig. 2.11 which shows that there exists a frequency ν_0 below which no electron is emitted.

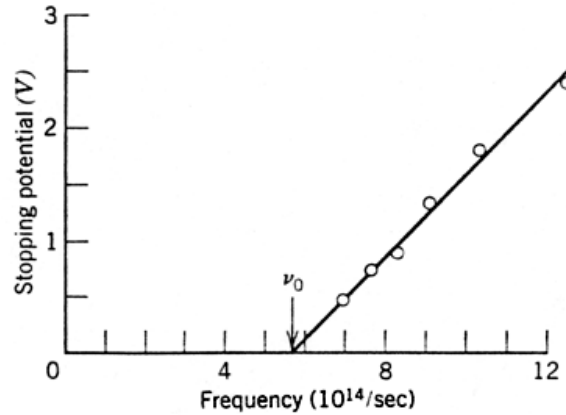


Fig. 2.11: Stopping potential as a function of the frequency.

This photoelectric effect cannot be explained via the wavelike and continuous theory of light. In fact

1. The classical theory predicts that if we increase the intensity of light, i.e. if we increase the electric field \vec{E} , the force $e\vec{E}$ acting over the electrons should increase as well as their kinetic energy and the stopping potential, but it is not so.
2. According to the wavelike nature of light there should be emission of electrons at *every* frequency provided the intensity is sufficiently large but it is not so.
3. The energy is uniformly spreaded along the wavefront and, if the intensity is low, the time necessary to extract the electrons is long, but it not so: in fact the emission of electrons is instantaneous.

Let us now see the explanation of the photoelectric effect given by Einstein. He proposed that the electromagnetic waves at microscopic scales appear as particles called *photons*. Interference and diffraction of electromagnetic waves can then be explained as statistical effects due to the presence of a large number of photons. In the case of the photoelectric effect every photon has an energy proportional to its frequency $E = h\nu$. Part of this energy is given to an electron of the plate A when the photon hits it. Consequently the kinetic energy of the electron reaches its maximum when all the energy of the photon is given to the electron. In such a case:

$$K_{\max} = h\nu - W_0 \quad (2.5.2)$$

where W_0 is the work required to extract the electron from the metal. This hypothesis explains the three open problems:

1. An increase of intensity of the incident radiation implies an increase of the number of photons and consequently of the electrons which are extracted from the metal. Therefore the current intensity I increases, but the kinetic energy and the stopping potential do not increase because they depend only on the frequency of the incident radiation: $K_{\max} = h\nu - W_0$.
2. There is a limiting frequency below which electrons cannot be extracted from the metal: this is the frequency for which $K_{\max} = 0$, i.e. $\nu_0 = W_0/h$.
3. Since the energy is not spreaded over a front but it appears in packets the emission is instantaneous and there is no response time.

Since $K_{\max} = eV_0$ from (2.5.2) we can derive that the stopping potential V_0 is given by:

$$V_0 = \frac{h}{e}\nu - \frac{W_0}{e}. \quad (2.5.3)$$

So V_0 as a function of ν is a straight line with slope h/e . In 1914 Millikan measured the charge of the electron e and therefore from Eq. (2.5.3) he could derive h . Not only, but the value he obtained agreed perfectly with the Planck's one from the black body spectrum. It is remarkable that two *different* phenomena give the *same* value of h .

Problem 2. Calculate the energy of a photon of yellow light and the number of photons which corresponds to an intensity of light $I = 5 \cdot 10^{17} \text{ eV/m}^2\text{sec}$.

Solution. The wavelength of a photon of yellow light is $\lambda = 5.89 \cdot 10^{-7} \text{ m} = 5890 \text{ \AA}$. Consequently the energy of the photon is:

$$\begin{aligned} E = h\nu = \frac{hc}{\lambda} &= \frac{6.63 \cdot 10^{-34} \text{ joule sec} \cdot 3 \cdot 10^8 \text{ m/sec}}{5.89 \cdot 10^{-7} \text{ m}} \\ &= 3.4 \cdot 10^{-19} \text{ joule} = 2.1 \text{ eV}. \end{aligned} \quad (2.5.4)$$

This energy is sufficient to win the work function necessary to extract electrons from the surface of a metal, e.g. of potassium. If we used microwaves the wavelength would be 10 cm and the energy of the photons $E = 10^{-5} \text{ eV}$ would be too low to extract the electrons from the surface. Even if the light intensity is $I = 5 \cdot 10^{17} \text{ eV/m}^2\text{sec}$ (which is a low intensity of the order of 0.1 Watt/m^2), the number N of photons per unit time and surface is very large:

$$N = \frac{I}{h\nu} = \frac{5 \cdot 10^{17}}{2.1} = 2.4 \cdot 10^{17} \text{ photons/m}^2 \cdot \text{sec}. \quad (2.5.5)$$

2.6 Compton Effect

In 1923 Compton performed an experiment which confirmed the particlelike nature of electromagnetic radiation. In an experiment of scattering of x -rays by a plate of graphite he noticed the presence of two different final frequencies for the x -rays.

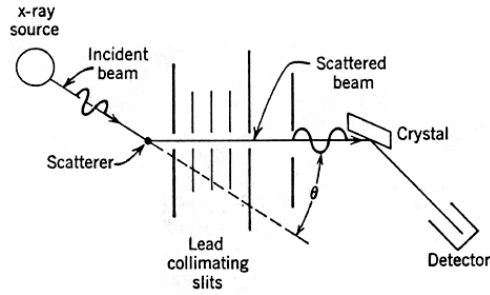


Fig. 2.12: Compton's experimental arrangement.

According to the wavelike classical theory there should be only one frequency because electrons in graphite oscillate with the same frequency of the incident photons and emit at the same frequency. If instead we consider photons as particles we can explain the phenomenon as an elastic collision between electrons and photons, dealing with them just as they were billiard balls, see Fig. 2.13.

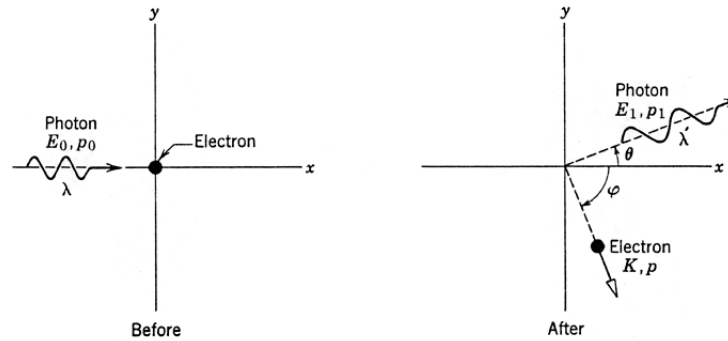


Fig. 2.13: Compton's interpretation.

From the *conservation of the momentum* we obtain:

$$\begin{cases} p_0 = p_1 \cos \theta + p \cos \varphi \\ p_1 \sin \theta = p \sin \varphi \end{cases} \quad (2.6.1)$$

where p_0 (p_1) is the initial (final) momentum of the photon, p is the final momentum of the electron, θ is the angle of deflection of the photon and φ is the recoil angle of the electron. The square of the two equations of (2.6.1) gives:

$$\begin{cases} (p_0 - p_1 \cos \theta)^2 = p^2 \cos^2 \varphi \\ p_1^2 \sin^2 \theta = p^2 \sin^2 \varphi. \end{cases} \quad (2.6.2)$$

By summing the two equations above we obtain

$$p_0^2 + p_1^2 - 2p_0 p_1 \cos \theta = p^2. \quad (2.6.3)$$

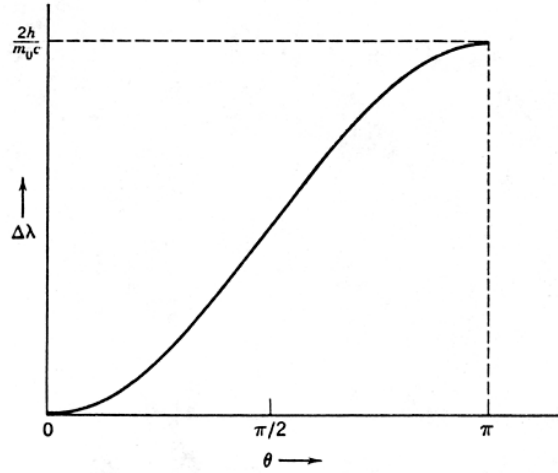


Fig. 2.14: Compton's wavelength shift.

For the *conservation of energy* we have instead

$$E_0 + m_0c^2 = E_1 + K + m_0c^2 \quad (2.6.4)$$

where E_0 (E_1) is the initial (final) energy of the photon, m_0c^2 is the energy of the electron at rest and K is its kinetic energy. From (2.6.4) we derive that

$$E_0 - E_1 = K. \quad (2.6.5)$$

Remember that for a relativistic particle

$$E^2 = c^2p^2 + (m_0c^2)^2 \quad (2.6.6)$$

but in the case of photons $m_0 = 0$ and therefore

$$E = cp \implies p = \frac{E}{c} = \frac{h\nu}{c} = \frac{h}{\lambda}. \quad (2.6.7)$$

Using (2.6.7) we can rewrite Eq. (2.6.5) as:

$$c(p_0 - p_1) = K. \quad (2.6.8)$$

Since the photon gives some of its energy to the electron it changes its frequency from $\nu_0 = E_0/h$ to $\nu_1 = E_1/h$. The scattered frequency ν_1 is independent of the material. This confirms that the process of scattering does not involve the atoms of the material but only its electrons. Now replacing $E = K + m_0c^2$ into (2.6.6) we obtain

$$(K + m_0c^2)^2 = c^2p^2 + (m_0c^2)^2 \quad (2.6.9)$$

which gives $K^2 + 2Km_0c^2 = c^2p^2$ or

$$\frac{K^2}{c^2} + 2Km_0 = p^2. \quad (2.6.10)$$

Now we can take p^2 from (2.6.3) and K from (2.6.8) and put them in (2.6.10). What we obtain is:

$$(p_0 - p_1)^2 + 2m_0c(p_0 - p_1) = p_0^2 + p_1^2 - 2p_0p_1 \cos \theta \quad (2.6.11)$$

which reduces to

$$m_0c(p_0 - p_1) = p_0p_1(1 - \cos \theta) \Rightarrow \frac{1}{p_1} - \frac{1}{p_0} = \frac{1}{m_0c}(1 - \cos \theta), \quad (2.6.12)$$

which, written in terms of the wavelengths $\lambda = h/p$, becomes:

$$\Delta\lambda = \lambda_1 - \lambda_0 = \lambda_C(1 - \cos \theta) \quad (2.6.13)$$

where $\lambda_C \equiv h/m_0c$ is the Compton wavelength. Let us notice that the Compton shift $\Delta\lambda$ depends on the scattering angle θ but not on λ_0 .

The experiments showed also the appearance, together with the scattered photons of wavelength given by (2.6.13), of some photons with the same frequency of the incident radiation, see Fig. 2.15.

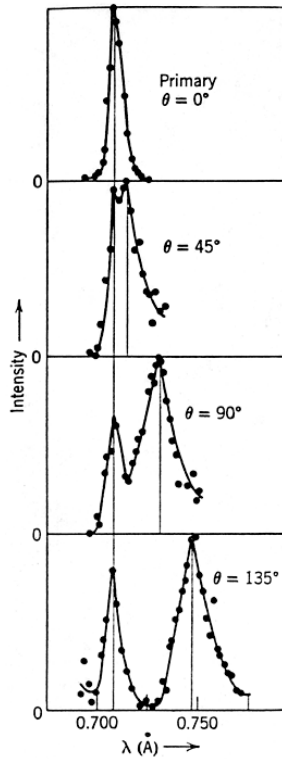


Fig. 2.15: Compton's experimental results.

This phenomenon is due to the so-called *Rayleigh scattering*: some photons are scattered by atoms or heavy ions which do not recoil at all. In this case the mass m_0 of the targets of the scattering process is very large, the associated Compton wavelength is approximately zero and, from Eq. (2.6.13), the Compton shift is $\Delta\lambda = 0$. The Rayleigh

scattering is the most common process for very large wavelengths $\lambda \rightarrow \infty$, while Compton effects make their appearance at low wavelengths $\lambda \rightarrow 0$, i.e. at high frequencies $\nu \rightarrow \infty$. In this regime the energy of the x -rays is sufficiently large to free the electrons in the collisions. It is just in this regime that we can observe quantum effects such as the granularity of energy.

2.7 Spectral Lines of an Atom; Rutherford's and Bohr's Models

The dimensions of the atoms ($r \sim 10^{-8}\text{m}$) were first calculated starting from the measurement of the mean free path λ in experiments of diffusion of molecules in a gas. In experiments such as the photoelectric effect and the x -rays diffusion it was seen that inside the atoms there were electrons with a negative charge. Nevertheless the atoms were neutral and therefore there should be also positive charges. Thomson proposed the so-called plum pudding model of atom in which electrons are uniformly distributed inside a sphere of positive charge.

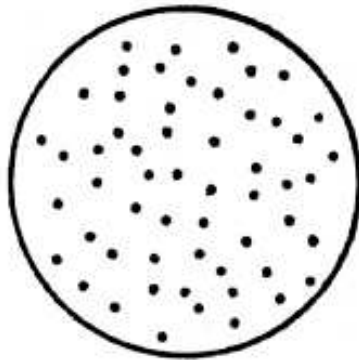


Fig. 2.16: Thomson's model of the atom—a sphere of positive charge embedded with electrons.

According to this classical model the electron of a hydrogen atom would move like a harmonic oscillator with only one frequency associated with a ultraviolet wavelength $\lambda = 1200 \text{ \AA}$, see Problem 3. Unfortunately it is possible to show experimentally that the spectra of the hydrogen atoms have many lines.

Problem 3. *Prove that the electrons in Thomson's model of the atom move like harmonic oscillators.*

Solution. If ρ is the density of the positive charge of the atom, then the total positive charge inside a sphere of radius a is given by $e = \frac{4}{3}\pi a^3 \rho$. Therefore the Coulomb force felt by the electron is

$$F \cong -\frac{1}{4\pi\epsilon_0} \left(\frac{4}{3}\pi a^3 \rho \right) \frac{e}{a^2} = -\frac{\rho e a}{3\epsilon_0}. \quad (2.7.1)$$

So the force F over the electron of the hydrogen atom is proportional to the distance $F = -ka$. The electrons moves like a harmonic oscillator with $k = \frac{\rho e}{3\epsilon_0}$. If the total positive charge is given by the charge of an electron e and it is distributed over a sphere of radius $r' = 10^{-10}$ m then the charge density is $\rho = \frac{e}{\frac{4}{3}\pi r'^3}$ and the frequency of the

harmonic oscillations becomes:

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{m}} = \frac{1}{2\pi} \sqrt{\frac{\rho e}{3\epsilon_0 m}} = \frac{1}{2\pi} \sqrt{\frac{e^2}{4\pi\epsilon_0 r'^3 m}} = 2.5 \cdot 10^{15} \text{ sec}^{-1}. \quad (2.7.2)$$

The associated wavelength becomes:

$$\lambda = \frac{c}{\nu} = 1200 \text{ \AA} \quad (2.7.3)$$

which corresponds to the wavelength of ultraviolet radiation.

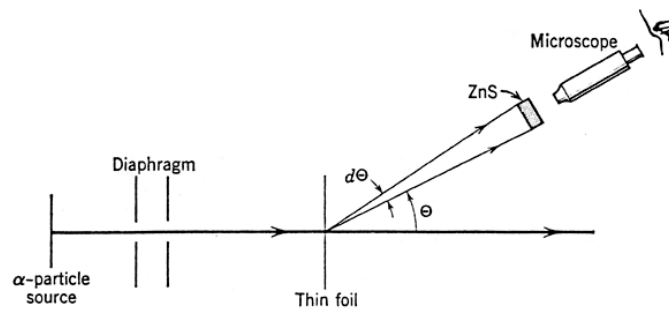


Fig. 2.17: Arrangement of an α -particle scattering experiment.

Rutherford in 1911 performed his well-known experiment of scattering of α -particles by a thin golden foil. He realized that the motion of the α -particles could be explained by a Coulomb scattering with pointlike sources of positive charges. Let us suppose the beam has an intensity of n α -particles per second and unit surface and that N is the number of atoms reached by the beam. Then the number of particles dn deflected with angles between θ and $\theta + d\theta$ is:

$$dn = nN \frac{Z^2 e_0^4}{4E^2} \frac{1}{\sin^4 \frac{\theta}{2}} d\Omega \quad (2.7.4)$$

where $d\Omega = 2\pi \sin\theta d\theta$, Ze_0 is the charge of the nucleus and E is the energy of the incident α -particle. Eq. (2.7.4) holds if particles are pointlike and the interaction is Coulombian.

Nevertheless in his experiment Rutherford saw that when the impact parameter was $b \sim 10^{-13}$ cm there were deviations from the law (2.7.4) and there were even α -particles rebounding towards the source. This can be justified by admitting that the positive charge was concentrated in an almost pointlike region of radius 10^{-13} cm.

After Rutherford's experiment the picture of the atom was the following one: there is a central nucleus of radius 10^{-13} cm and positive charge Ze_0 and there are Z electrons

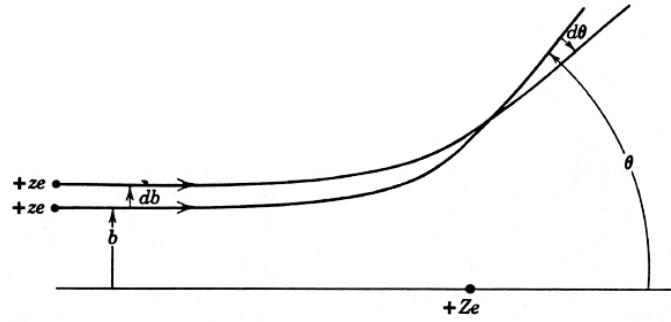


Fig. 2.18: Two hyperbolic Rutherford trajectories for two different values of the impact parameters.

at a mean distance of about 10^{-8} cm. The electrons have a kinetic energy and move following an orbit like planets around the Sun. Also this picture of the atom suffered a big problem: in fact the electrons are charged objects and, consequently, they irradiate energy while accelerating. The energy irradiated per unit time is given by the so-called Hertz's formula:

$$\frac{dE}{dt} = \frac{2}{3} \frac{e^2}{c^3} \ddot{x}^2. \quad (2.7.5)$$

So electrons should lose their energy and in about 10^{-8} sec they should fall down on the nucleus.

According to classical mechanics the radiation emitted from the electrons should have the same frequency of the electrons themselves. If we equate the Coulomb force to the centrifugal one:

$$m_e \frac{v^2}{r} = \frac{e_0^2}{r^2} \quad (2.7.6)$$

we obtain that the angular frequency is:

$$\omega \equiv \frac{v}{r} = \frac{e_0}{\sqrt{m_e}} \frac{1}{r^{3/2}}. \quad (2.7.7)$$

Therefore the wavelength of the emitted radiation is given by:

$$\lambda = \frac{2\pi c}{\omega} = \frac{2\pi}{(e_0^2/m_e c^2)^{1/2}} r^{3/2}. \quad (2.7.8)$$

If the electron changes the radius of its orbit from $r \approx 5 \cdot 10^{-8}$ cm to $r \approx 10^{-10}$ cm the emitted radiation is between $\lambda = 10^4 \text{ \AA}$ (infrared radiation) and $\lambda = 1 \text{ \AA}$ (*x*-rays). Therefore according to classical mechanics the spectrum of emission should contain all the possible frequencies. Now the continuous spectrum of emission and the non-stability of the atoms predicted by classical mechanics did not agree at all with experiments.

In fact in the absorption and emission spectra of the hydrogen atom there were not all the frequencies but only some *fixed* and *discrete* ones. The emitted lines had the following wave numbers $\tilde{\nu} \equiv \frac{\nu}{c} = \frac{1}{\lambda}$:

$$\tilde{\nu} = R_H \left(\frac{1}{n'^2} - \frac{1}{n^2} \right) \quad (2.7.9)$$

where

$$R_H = 109\,678 \text{ cm}^{-1} \quad (2.7.10)$$

is the Rydberg constant. By varying n and n' we obtain the following series:

- $n' = 1, n = 2, 3, 4, \dots$, Lyman series (ultraviolet radiation);
- $n' = 2, n = 3, 4, 5, \dots$, Balmer series (visible radiation);
- $n' = 3, n = 4, 5, 6, \dots$, Paschen series (infrared radiation);
- $n' = 4, n = 5, 6, 7, \dots$, Brackett series (infrared radiation).

By increasing the number n the lines are thickening around the line $\frac{R_H}{n'^2}$. Therefore every series have its limiting line. Behind this regularity there should be a *beautiful model* or *theory*.

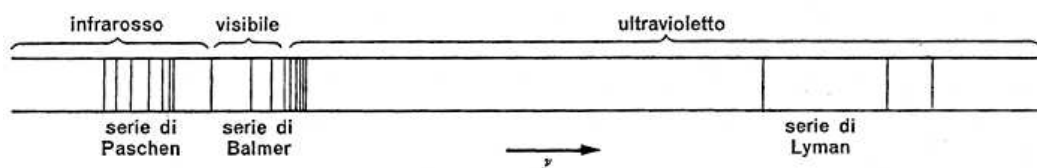


Fig. 2.19: Spectrum of the hydrogen atom.

A first attempt to explain such a regularity is given by the Bohr's model of atom. This is an abstract model with ad hoc rules which violate the principles of classical mechanics. These rules are:

1. The electrons choose some *allowed* orbits on which they do not irradiate.
2. The emission and absorption happen when there is a transition of an electron from an allowed orbit to another one.
3. The frequency ν of the emitted and absorbed radiation is given by the Bohr's formula:

$$E_n - E'_n = h\nu \quad (2.7.11)$$

which agrees with the quantization rule for the energy given by Planck.

It remains the problem of finding the rule which selects the allowed orbits. If energy is quantized also other physical quantities, like for example the angular momentum, can be quantized. Therefore for a hydrogen atom Bohr postulated that the angular momentum is quantized according to the rule:

$$m_e v r = n \hbar \quad (2.7.12)$$

where $\hbar = h/2\pi$. Therefore if we put the centrifugal force equal to the Coulomb one:

$$m_e \frac{v^2}{r} = \frac{e_0^2}{r^2} \quad (2.7.13)$$

and we eliminate the velocity v using (2.7.12), we obtain that:

$$r_n = n^2 \frac{h^2}{4\pi^2 e_0^2 m_e}. \quad (2.7.14)$$

Now, using Eq. (2.7.13), the total energy becomes:

$$E_n = \frac{1}{2} m_e v^2 - \frac{e_0^2}{r_n} = -\frac{1}{2} \frac{e_0^2}{r_n} \quad (2.7.15)$$

and, if we use Eq. (2.7.14) for r_n , we get that the energy levels are related with the quantum numbers $n = 1, 2, 3, \dots$ by:

$$E_n = \frac{-Rhc}{n^2} \quad (2.7.16)$$

where $R = \frac{2\pi^2 e_0^4 m_e}{h^3 c} = 109\,700 \text{ cm}^{-1}$ which is in quite good agreement with the experimental value given by (2.7.10). This agreement can be improved further by taking into account the mass of the nucleus m_N and by replacing in all the formulas the mass of the electron m_e with its reduced mass $\frac{m_e m_N}{m_e + m_N}$.

Using Eq. (2.7.16) in (2.7.11) we can reproduce the Balmer's formula (2.7.9):

$$h\nu = E_{n'} - E_n = -Rhc \left[\frac{1}{n'^2} - \frac{1}{n^2} \right]. \quad (2.7.17)$$

Bohr's model is an ad hoc phenomenological model which reproduces all the experimental data. Just to give some numerical values, the first orbit has energy

$$E_1 = -Rhc = -13.6 \text{ eV} \quad (2.7.18)$$

and radius:

$$r_1 = \frac{h^2}{4\pi^2 e_0^2 m_e} = 0.5 \text{ \AA}. \quad (2.7.19)$$

This is the Bohr's radius which is usually indicated with a_0 .

Actually the Bohr's rules worked well only in the case of hydrogen; for more complicated atoms Sommerfeld imposed other quantization rules such as, for example:

$$J_i(E) \equiv \oint p_i(E) dr_i = n_i h \quad (2.7.20)$$

where the variables J_i are called action variables and can be defined every time we have a periodic orbit. Since there is an action variable for every degree of freedom there are $3k$ quantum numbers where k is the total number of degrees of freedom. Also this model did not work very well. Furthermore was it possible to justify the Bohr's or Sommerfeld's quantization rules? How was it possible that a particle did not irradiate while accelerating? Why only some particular orbits were allowed?

2.8 De Broglie's idea

A first attempt to answer the previous questions was given in 1924 by Louis de Broglie (who had done undergraduate studies in history and whose brother, Maurice, was an experimental physicist). In his Ph.D. thesis (which was sent by Langevin to Einstein and accepted by him) de Broglie proposed that, as light has a particlelike nature besides the wavelike one, so matter should have a wavelike nature besides the particlelike one. For a photon, since $E = h\nu = pc$, the momentum p can be written as:

$$p = h\frac{\nu}{c} = \frac{h}{\lambda}. \quad (2.8.1)$$

De Broglie proposed to use Eq. (2.8.1) also for particles, that means he associated to every particle of momentum p a wavelength:

$$\lambda = \frac{h}{p} \quad (2.8.2)$$

which is called de Broglie wavelength.

Problem 4. Calculate the de Broglie wavelength for a ball of mass $m = 1$ Kg and speed $v = 10$ m/sec. Calculate the wavelength of an electron of energy $E = 100$ eV.

Solution. It is sufficient to apply Eq. (2.8.2):

$$\begin{aligned} \lambda &= \frac{h}{p} = \frac{h}{mv} = \frac{6.6 \times 10^{-34} \text{ joule} \cdot \text{sec}}{1 \text{ kg} \times 10 \text{ m/sec}} \\ &= 6.6 \times 10^{-35} \text{ m} \sim 6.6 \times 10^{-25} \text{ \AA}. \end{aligned} \quad (2.8.3)$$

For an electron of energy $E = 100$ eV we have instead:

$$\begin{aligned} \lambda &= \frac{h}{p} = \frac{h}{\sqrt{2mE}} = \frac{6.6 \times 10^{-34} \text{ joule} \cdot \text{sec}}{\sqrt{2 \times 9.1 \times 10^{-31} \times 100 \cdot 1.6 \cdot 10^{-19}} \text{ Kg} \cdot \text{m/sec}} \\ &= 1.2 \text{ \AA}. \end{aligned} \quad (2.8.4)$$

Just to give a term of comparison for the results of Problem 4, the wavelength of x -rays is 1 \AA , while the one of γ -rays is $10^{-12} \text{ cm} = 10^{-2} \text{ \AA}$. Remember that in order to see diffraction phenomena the slits must have a width $a \sim \lambda$. In the case of the ball of Problem 4 the wavelength is $\lambda \approx 10^{-25} \text{ \AA}$. This means that in order to see diffraction with such a macroscopic baseball we should use slits of width $a \sim 10^{-25} \text{ \AA}$!

Elsasser realized in 1926 that it should be possible to observe the wavelike nature of electrons with the same kind of experiments used to show the wavelike nature of x -rays, which have approximately the same wavelength, i.e. by diffraction from a crystalline solid. In 1927 Davisson and Germer in the USA and G. P. Thomson (son of J. J. Thomson who had proved in 1897 the particlelike nature of electrons) in Scotland obtained a figure of diffraction by experiments with electrons scattered by golden crystals, confirming in this way the de Broglie's hypothesis.

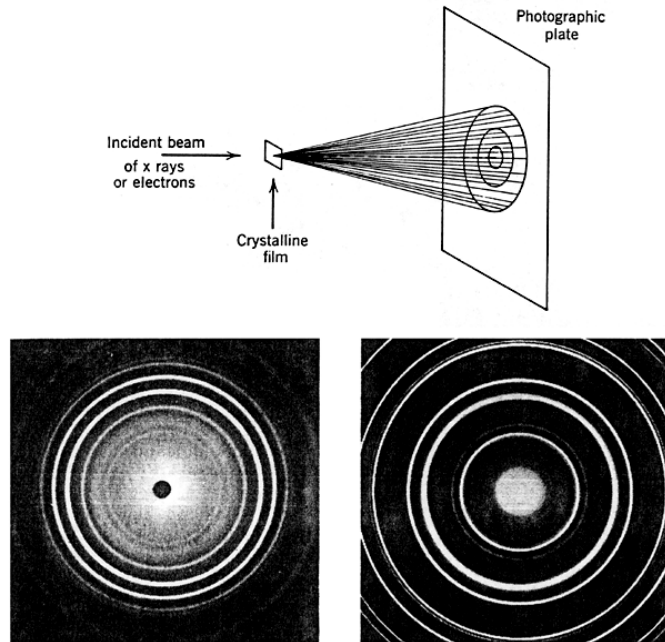


Fig. 2.20: Diffraction of x -rays or electrons by a polycrystalline material.

These experiments were difficult to realize because the intensities of the atomic rays were very small. In order to describe these diffraction effects physicists began to associate to every particle of energy E and momentum p a wave:

$$\psi(x, t) = Ae^{\frac{2\pi i}{h}(px - Et)}. \quad (2.8.5)$$

It is interesting to note that Bohr's rules can be obtained from de Broglie's hypothesis.

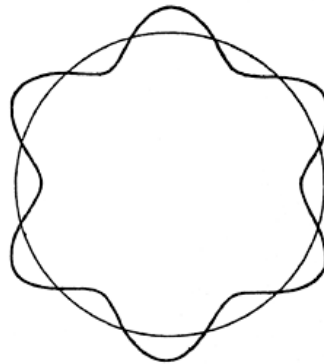


Fig. 2.21: A standing de Broglie wave in a Bohr orbit.

From Fig. 2.21 it is easy to get convinced that, in order to have a periodic wave, it is necessary for the length of the circular orbit of radius r to be an integer multiple of the

wavelength:

$$2\pi r = n\lambda = n \frac{h}{p_e} \quad (2.8.6)$$

from which we derive that

$$p_e r = n\hbar \quad (2.8.7)$$

which is just the Bohr's quantization rule (2.7.12).

As we said before the first experiments confirming the wavelike nature of electrons were obtained by diffraction from a slit or from a crystal lattice. This last experiment pictured in Fig. 2.22 had already been used by Bragg in order to prove the wavelike nature of x -rays. Note that there is a constructive interference between the two diffracted waves if $2d \sin \theta = n\lambda$. In the case of electrons this interference is between different parts of the same wave (2.8.5) describing the system. Diffraction effects with molecules were seen for the first time by Esterman, Stern e Frisch; diffraction effects with slow neutrons were seen for the first time by Fermi. If instead we want to see the particlelike nature of matter we have to decrease the intensity of the beam (it is even possible to have the emission of only one electron at once) and to use a detector.

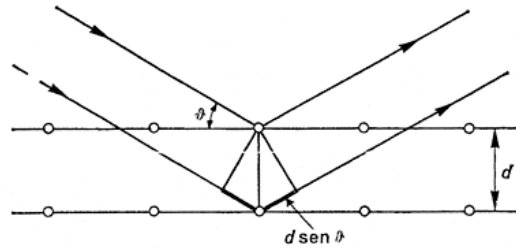


Fig. 2.22: Diffraction from a crystal lattice.

2.9 Wave-Particle Duality

In the previous sections we have shown how the quantum description of nature is characterized by a sort of *wave-particle duality* in the sense that both matter and radiation have both a wavelike and a particlelike nature.

1. For what concerns *particles* we have seen that every particle of momentum p has an associated wavelength given by $\lambda = h/p$. Now the Planck's constant is small, so if we want to observe the wavelike features of matter it is necessary to have particles with small momenta p . The momentum of the electrons is usually sufficiently small in order to see, e.g., interference and diffraction phenomena. For macroscopic objects instead the masses m and the momenta p are so large that the associated wavelength is practically zero. We can also say that, since in *classical mechanics* $h = 0$, classical particles cannot show their wavelike nature.

2. *Waves* show their particlelike nature when they are emitted or absorbed. Their wavelike nature instead manifests itself when waves are transmitted. This phenomenon occurs also for matter: for example in the detectors we can see the particlelike nature of electrons.

The wavelike and the particlelike features are complementary. This means that every experiment showing the wavelike nature never shows the particlelike one and vice versa. Therefore there is no contradiction¹. This principle stated by N. Bohr is called *complementary principle* and can be justified by the probabilistic interpretation of the wave-particle duality. Such a probabilistic interpretation was pioneered by Einstein in the case of waves. Since the light from a source of an electric bulb contains a very large number of photons we can give, as in the kinetic theory of gases, only a *statistical* explanation of the phenomena. In the wave picture of light the intensity I is proportional to $\overline{E^2}$ which is the average value over one temporal cycle of the square of the electric field. In a particle picture the intensity of light is given by $I = Nh\nu$ where N is the average number of photons crossing unit surface perpendicular to the direction of propagation per unit time. The two pictures are consistent. In fact if we have a pointlike source of radiation emitting randomly in all the directions then the number of photons emitted per unit time is fixed. The photons are spreaded over a sphere whose surface increases with the square of the distance r from the source. Consequently the number of photons N crossing the unit surface decreases as $1/r^2$ and the intensity of light, which is proportional to N , decreases as $1/r^2$. This result is the same as in the wave picture where the spherical wave spreads from the source and the intensity decreases with $1/r^2$. Therefore we can say that the two different pictures of light are consistent.

If the intensity of the light is large, i.e. if the number of photons N is large, then the *fluctuations* are small in the sense that the ratio $\frac{\Delta N}{N}$ is approximately zero. Therefore in this case fluctuations are not important and we can use an effective non-granular theory. Vice versa if the intensity is low the ratio $\frac{\Delta N}{N}$ is large, fluctuations become relevant; this confirms the granular nature of light which can be directly observed in these regimes of low intensities.

Since the complementary principle can be explained by the probabilistic interpretation of the wave-particle duality, our theory must give the *probability* $P(r, t)$ of having a photon in r at time t . We want to prove that the knowledge of this probability is *not* sufficient in order to describe the situation. Suppose we have a source of photons. If we reduce the intensity of emission we can emit only one photon at once. The photons go through a two-slit plate and arrive on the final plate of Fig. 2.23. Since for low intensities the particlelike nature of light should emerge, we expect to obtain on the final plate a total probability like the one of Fig. 2.23-(b). In particular, because of the particlelike nature of the radiation, we should have a large probability in the region opposite the two slits and a low probability in the middle region. We should also have that the total probability P_T must be given by the sum $P_T = P_1 + P_2$ of the probabilities P_1 and P_2 that we would obtain by closing the hole 2 and the hole 1 separately. What we observe instead on the final plate is an interference figure like in Fig. 2.23-(c) which implies that the *two probabilities must not be summed*: $P_T \neq P_1 + P_2$.

¹After all physicists must give explanations of the experiments and not of “nature” of things.

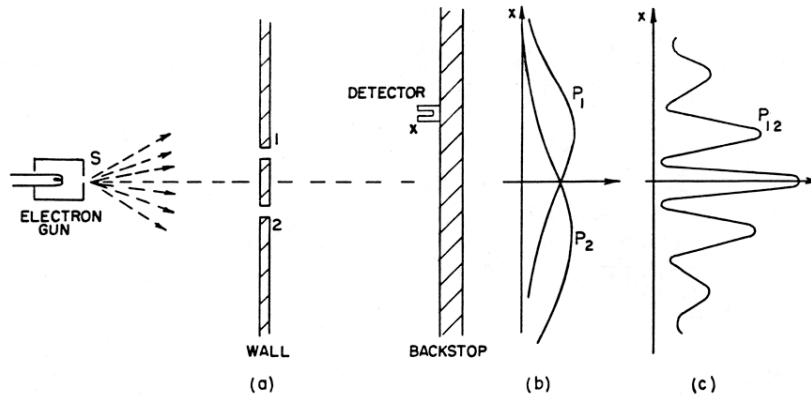


Fig. 2.23: Arrangement of a two-slit experiment.

The previous problem can be solved by considering, instead of the probability $P(r, t)$ of finding a particle in r at time t , its “square root” $\psi(r, t)$:

$$|\psi(r, t)|^2 = P(r, t) \quad (2.9.1)$$

which is called *wave function* or *probability amplitude*. In general $\psi(r, t)$ is a *complex function* and from

$$1 = \int d^3r P(r, t) = \int d^3r |\psi|^2 \quad (2.9.2)$$

we have that $\psi(r, t) \in L^2$. In order to reproduce the interference figure we *must sum the probability amplitudes* instead of the probabilities:

$$\begin{aligned} P_T &= |\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + \psi_1^* \psi_2 + \psi_2^* \psi_1 \\ &= P_1^2 + P_2^2 + \psi_1^* \psi_2 + \psi_2^* \psi_1. \end{aligned} \quad (2.9.3)$$

It is just the extra term $\psi_1^* \psi_2 + \psi_2^* \psi_1$ appearing in (2.9.3) which causes the appearance of the interference figure. On the space of wave functions the following superposition principle holds: if $\psi_1(r)$ and $\psi_2(r)$ are two allowed “configurations” then also every linear combination with complex coefficients

$$\psi_{\alpha\beta}(r) = \alpha\psi_1(r) + \beta\psi_2(r) \quad (2.9.4)$$

is an allowed configuration. Only by allowing such a combination it is possible to sum configurations which are *mutually exclusive*, like the passage through the two slits in an interference experiment. Since according to de Broglie’s idea there were waves associated not only with radiation but also with *particles*, Born proposed that also for particles the $\psi(r)$ represents the *probability amplitude* of finding a particle in r . Therefore the ψ is not a material wave but a *probability wave*. The interference effects of matter are due both to the complex character of the wave functions and to the fact the probability amplitudes and not the probabilities must be summed. In quantum mechanics moreover the probabilistic ingredient does not come, like in statistical mechanics, from our ignorance

about the initial conditions but it is somehow intrinsic of the formalism itself of quantum mechanics.

Is it possible to relate the probability amplitudes $\psi(r)$ with the concept of possible paths followed by the particles? Let us consider the situation pictured in Fig. 2.24 where we have a source of radiation and a series of slits between the source and the final plate. Which is the probability amplitude of finding a particle in a certain point x on the final plate? This amplitude will be given by the sum of the amplitudes each corresponding to a configuration in which only one slit per plate is open. We indicate this as:

$$\psi(r) = \sum \psi_{A^1 B^3 C^5 D^2} + \psi_{A^1 B^2 C^5 D^2} + \psi_{A^1 B^1 C^5 D^2} + \dots \quad (2.9.5)$$

where $\psi_{A^1 B^3 C^5 D^2}$ means that the slit 1 is open in plate A , the slit 3 is open in plate B , etc.

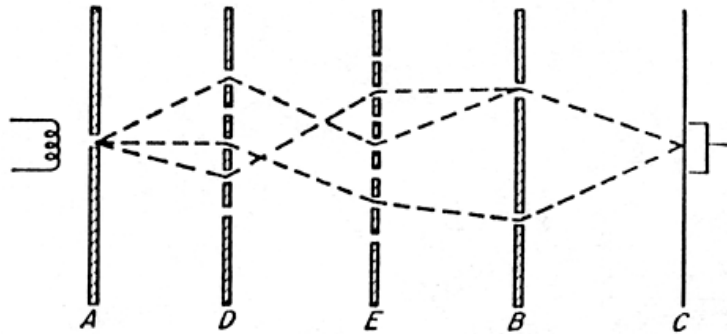


Fig. 2.24: A generalization of the two-slit experiment.

Finally, with a little bit of imagination, we can consider the full open space between the source and the screen as made up of an infinite number of plates with infinite slits. Notice from Fig. 2.24 that in such a way every *amplitude* $\psi_{i,j,k,\dots}$ can be put in *one-to-one correspondence* with a *path* and the total wave function:

$$\psi(r) \longleftrightarrow \sum_{\text{paths}} A \quad (2.9.6)$$

can be put in correspondence with a sum of the amplitudes over all the possible paths connecting the initial and the final configuration. It is important to stress that every path which appears in (2.9.6) must be intended only as a series of indices, corresponding to the numbers of slits which are left open, and not as a real path followed by the quantum particle. Nevertheless this idea is the starting point for the well-known and useful path integral formulation of quantum mechanics proposed by R. P. Feynman.

2.10 Heisenberg's Uncertainty Principle

As we have seen in the previous section the probabilistic features of quantum mechanics are somehow intrinsic in the formalism of the theory. In *classical mechanics* the situation

is very different: if we give the initial conditions $q(0)$ and $p(0)$, it is possible to reconstruct the trajectory $(q(t), p(t))$ and find the point it reaches at every later instant of time t . Even if we have a system with an infinite number of particles like in statistical mechanics, we could measure in principle all their positions and velocities and know everything about the system. We will not have the time to do it but in principle nothing forbids us to measure with infinite precision all the coordinates of the particles in the phase space and to know everything about their future. In *quantum mechanics* instead there is something in principle which forbids us to know $q(0)$ and $p(0)$ with infinite precision even for a single particle. Basically this is the origin of all the intrinsic statistical features of quantum mechanics and it is related to the fact that a quantum measurement disturbs the physical system because of the interaction between the classical instrument and the measured quantum object. This effect is small for macroscopic objects but becomes relevant in quantum mechanics.

The situation at the beginning of quantum mechanics was similar to the one existing at the beginning of special relativity, when physicists spoke about simultaneity of events without asking themselves how simultaneity could be physically established. Then Einstein showed that simultaneity of events is a relative concept: two events which are simultaneous in a certain system can be not simultaneous in another one. In quantum mechanics instead the problem was to give a set of operative rules in order to establish how position and momentum could be measured. In particular the question was: can we measure with absolute precision both the position and the momentum of a particle? The answer given by quantum mechanics is: No! The uncertainties in the measurement of positions and momenta satisfy the following inequality:

$$\Delta x \Delta p_x \geq \frac{\hbar}{2} \quad (2.10.1)$$

where $\hbar = h/2\pi$. So the precision in the measurement of the position x is inversely proportional to the precision in the measurement of the associated momentum p_x . In particular if we measure x with infinite precision then $\Delta x = 0$ and therefore $\Delta p_x = \infty$, i.e. the momentum p_x is totally undetermined. This is the well-known *Heisenberg's uncertainty principle*. This principle involves only canonically conjugate variables, so for example $\Delta x \Delta p_y = 0$, i.e. a measurement of x does not influence p_y while instead

$$\Delta E \Delta t \geq \frac{\hbar}{2} \quad (2.10.2)$$

which means that also the precision in a measurement of energy is inversely proportional to the time required to perform the measurement. For example we can identify the ΔE of (2.10.2) with the energy spread of the photons emitted in a time interval Δt by an atom. Since Heisenberg's relations can be derived from the wavelike nature of matter postulated by de Broglie, they are confirmed by experiments just as the de Broglie's postulate. If $\hbar = 0$ from (2.10.1) we would obtain $\Delta x \Delta p_x \geq 0$, i.e. it would be possible to measure with infinite precision the position and the momentum of a particle. We can say that is just the smallness of \hbar which implies the absence of an uncertainty principle in classical mechanics.

In order to verify (2.10.1) we can use a thought experiment proposed by N. Bohr and which is known as the Heisenberg's microscope. Suppose we want to measure the

position and the momentum of an electron by illuminating it with a source of light or photons and by using a microscope, like it is shown in Fig. 2.25.

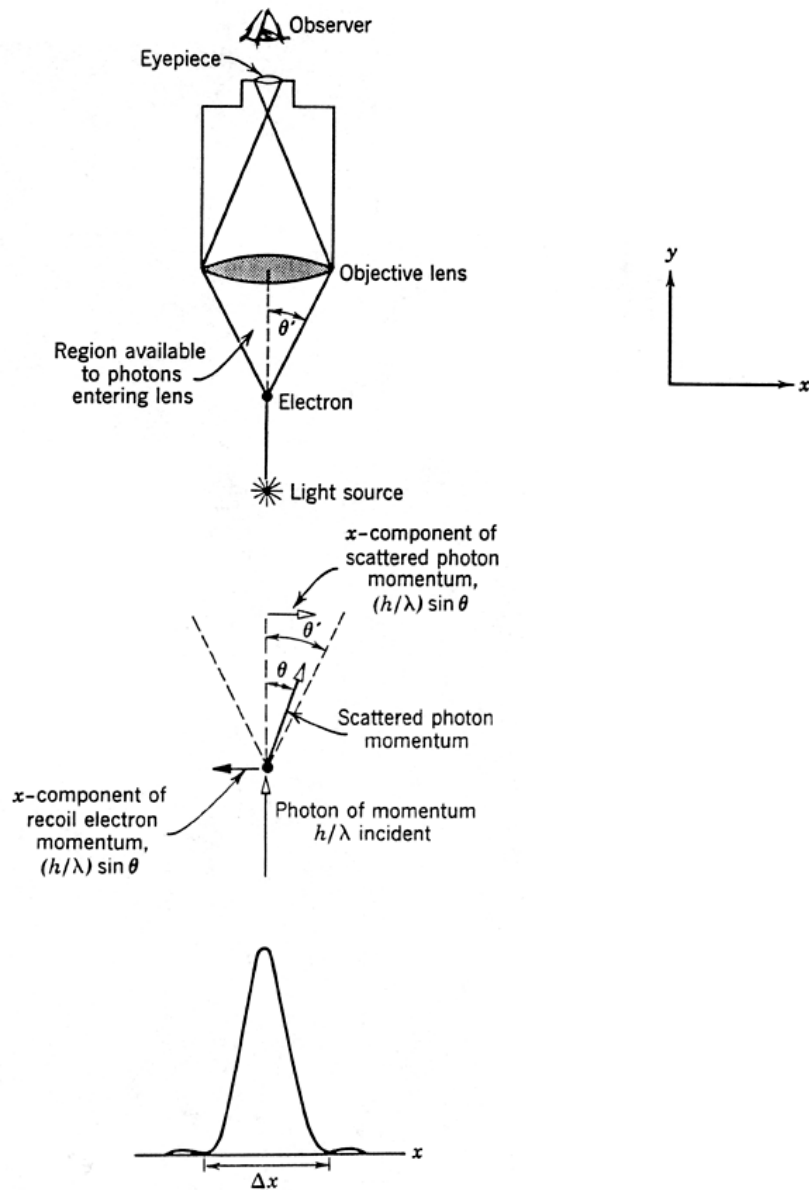


Fig. 2.25: Bohr's microscope thought experiment.

In order to reduce the disturbance of the measurement we can imagine of using only one photon; we cannot do better because at least one photon is necessary in order to view the electron. The interaction between the photon and the electron must be treated as the Compton effect: if the incident photon has a momentum $p = h/\lambda$ then the photon which is scattered by an angle θ has an x component of the momentum given by $p_x = \frac{h}{\lambda} \sin \theta$

which is equal and opposite to the x component of the momentum of the recoiling electron. If $2\theta'$ is the angle subtended by the lens at the electron the x component of the momentum can vary from $p \sin \theta'$ to $-p \sin \theta'$. Therefore the uncertainty in the measurement of the momentum of the photon in the x direction is:

$$\Delta p_x = 2p \sin \theta' = 2 \frac{h}{\lambda} \sin \theta'. \quad (2.10.3)$$

The electron recoils in the opposite direction and, for the conservation of the momentum, also the x component of the electron must be uncertain by the quantity (2.10.3). Now the image of a point at a microscope is not a point but a diffraction pattern, like shown in Fig. (2.25). The resolving power of the microscope determines the maximum accuracy we can reach in the measurement of the position x of the electron:

$$\Delta x = \frac{\lambda}{\sin \theta'}. \quad (2.10.4)$$

Therefore the product of the two uncertainties (2.10.3) and (2.10.4) is given by:

$$\Delta x \Delta p_x = \frac{\lambda}{\sin \theta'} \frac{2h}{\lambda} \sin \theta' = 2h \quad (2.10.5)$$

which is bigger than $\hbar/2 \approx h/12$. The crucial concept behind this thought experiment is that, due to the particlelike nature of the photon and the Compton effect, there is a disturbance effect of the "measurement instrument" (the photon) and the observed system (the electron). This effect is minimum in the case of the Heisenberg's microscopic because there is only one photon but in any case such an effect cannot be avoided. For example if we used light with small values of λ from (2.10.4) it is clear that we would have a better resolution $\Delta x \sim 0$, but at the same time Δp_x would increase because λ appears at the denominator of (2.10.3). In classical mechanics instead photons do not show their particlelike features and so we can decrease the intensity of the light down to zero or to arbitrary small values in order not to disturb the observed system and at the same time we can decrease λ in order to improve the resolution. In classical mechanics there cannot be any uncertainty relation like the one of Eq. (2.10.5).

Before concluding this section we want to prove how the uncertainty relation between time and energy (2.10.2) can be derived from (2.10.1). If we consider a free particle the uncertainty in the measurement of its energy can be related with the uncertainty in the measurement of its momentum. In fact:

$$\Delta E = \Delta \left(\frac{p_x^2}{2m} \right) = \frac{p_x}{2m} 2\Delta p_x = v_x \Delta p_x. \quad (2.10.6)$$

The uncertainty in the position Δx can be related with the time interval Δt required for the measurement: $\Delta x = v_x \Delta t$. Consequently $v_x = \frac{\Delta x}{\Delta t}$ which, replaced into (2.10.6), gives:

$$\Delta E = v_x \Delta p_x = \frac{\Delta x}{\Delta t} \Delta p_x \implies \Delta E \Delta t = \Delta x \Delta p_x \geq \frac{\hbar}{2}. \quad (2.10.7)$$

Problem 5. *The velocity of a bullet ($m = 50$ g) and of an electron ($m = 9.1 \times 10^{-28}$ g) are measured to be the same, 300 m/sec, with an uncertainty of 0.01%. Which is the accuracy in the measurement of their positions?*

Solution. For the electron:

$$p = mv = 9.1 \times 10^{-31} \text{ Kg} \times 300 \text{ m/sec} = 2.7 \times 10^{-28} \frac{\text{Kg} \cdot \text{m}}{\text{sec}}.$$

Its uncertainty is given by:

$$\Delta p = m\Delta v = 0.0001 \times 2.7 \times 10^{-28} \text{ Kg} \cdot \text{m/sec} = 2.7 \times 10^{-32} \text{ Kg} \cdot \text{m/sec}.$$

Therefore in this case the uncertainty principle $\Delta x \Delta p \geq \frac{\hbar}{2} \geq \frac{h}{4\pi}$ implies immediately:

$$\Delta x \geq \frac{h}{4\pi\Delta p} = \frac{6.6 \times 10^{-34} \text{ joule} \cdot \text{sec}}{4\pi \times 2.7 \times 10^{-32} \text{ Kg} \cdot \text{m/sec}} = 2 \cdot 10^{-3} \text{ m} = 0.2 \text{ cm}.$$

For the bullet instead:

$$p = mv = 0.05 \text{ Kg} \times 300 \text{ m/sec} = 15 \text{ Kg} \cdot \text{m/sec}$$

$$\Delta p = 0.0001 \times 15 \text{ Kg} \cdot \text{m/sec} = 1.5 \times 10^{-3} \text{ Kg} \cdot \text{m/sec}.$$

Therefore the uncertainty in the measurement of the position is given by:

$$\Delta x \geq \frac{h}{4\pi\Delta p} = \frac{6.6 \times 10^{-34} \text{ joule} \cdot \text{sec}}{4\pi \times 1.5 \times 10^{-3} \text{ Kg} \cdot \text{m/sec}} = 3 \times 10^{-32} \text{ m}.$$

In this case the uncertainty in the measurement of the position is about 10^{-17} times the diameter of the nucleus (which is approximately 10^{-15} cm) and so there is practically no limit in the measurement precision of x for a microscopic bullet. In the case of the electron instead we have an uncertainty in x of 0.2 cm which is 10^7 times the diameter of an atom (10^{-8} cm).

Chapter 3

Schrödinger Equation

3.1 General Properties

In 1926 Schrödinger asked himself which was the equation of motion for the probability amplitudes or wave functions $\Psi(x)$. This equation had to satisfy the following properties:

1. the so called Planck-de Broglie's relation: $\lambda = \frac{h}{p}$, $\nu = \frac{E}{h}$;
2. the conservation of energy $E = \frac{p^2}{2m} + V$;
3. the linearity: if Ψ_1 and Ψ_2 are two different solutions of the equation of motion then also every linear combination $\Psi = c_1\Psi_1 + c_2\Psi_2$ must be a solution;
4. if the potential $V(x, t)$ is identically zero then the solutions of the equation must be given by the following plane waves:

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} e^{i2\pi(x/\lambda - \nu t)} = \frac{1}{\sqrt{2\pi}} e^{i\frac{2\pi}{h}(px - Et)} = \frac{1}{\sqrt{2\pi}} e^{i\frac{1}{h}(px - Et)}. \quad (3.1.1)$$

Alternatively we can use the wave number $k = 2\pi/\lambda$ and the angular frequency $\omega = 2\pi\nu$ and the plane wave (3.1.1) can be written as:

$$\Psi(x, t) = \frac{1}{\sqrt{2\pi}} e^{i(kx - \omega t)}. \quad (3.1.2)$$

Let us consider the plane waves (3.1.2) and let us build their derivatives; we are interested in the *derivatives* because the wave functions $\Psi(x, t)$ will satisfy a *differential* equation, just as the point particles in classical mechanics satisfy a differential equation of motion given by $m\ddot{x} = -\frac{\partial V}{\partial x}$. If we derive (3.1.2) with respect to x we obtain:

$$\frac{\partial}{\partial x} \Psi(x, t) = ik\Psi(x, t) = i\frac{2\pi}{\lambda} \Psi(x, t) = i\frac{2\pi p}{h} \Psi(x, t) = \frac{ip}{\hbar} \Psi(x, t) \quad (3.1.3)$$

which implies:

$$-i\hbar \frac{\partial}{\partial x} \Psi(x, t) = p\Psi(x, t). \quad (3.1.4)$$

Let us instead derive (3.1.2) with respect to t :

$$\frac{\partial \Psi}{\partial t} = -i\omega\Psi = -i2\pi\nu\Psi = -i2\pi\frac{E}{h}\Psi = -i\frac{E}{\hbar}\Psi$$

from which we obtain

$$i\hbar\frac{\partial}{\partial t}\Psi = E\Psi. \quad (3.1.5)$$

Multiplying the definition of energy for a free particle $\frac{p^2}{2m} = E$ by the wave function Ψ and using (3.1.4)-(3.1.5), we obtain the equation of motion for the amplitudes Ψ in the case of free particles:

$$\begin{aligned} \frac{p^2}{2m}\Psi = E\Psi &\Rightarrow \frac{1}{2m}\left(-i\hbar\frac{\partial}{\partial x}\right)^2\Psi = i\hbar\frac{\partial}{\partial t}\Psi \\ &\Rightarrow -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t). \end{aligned} \quad (3.1.6)$$

Therefore we can say that the equation of motion of the wave functions reflects somehow the law of conservation of energy. The most general solution Ψ of Eq. (3.1.6) will be given by a superposition of plane waves and, since the equation holds for a particular plane wave, then it must hold also for a superposition of plane waves:

$$\Psi(x,t) = \int \frac{dk}{\sqrt{2\pi}} [e^{ikx-i\omega t}] \tilde{\Psi}(k). \quad (3.1.7)$$

This is a linear combination of plane waves with coefficients $\tilde{\Psi}(k)$ determined by the initial condition $\Psi(x,0) = \Psi_0(x)$. Since a linear combination of solutions of the Schrödinger equation must be itself a solution, also the wave function (3.1.7) has to satisfy the Schrödinger equation (3.1.6). It represents a non-sinusoidal wave, i.e. a wave with a wavelength λ which is not constant but changes with x .

Schrödinger postulated that, starting from the conservation of energy, a similar equation for $\Psi(x,t)$ should be derived also for a particle in a potential $V(x)$. First of all he multiplied the definition of the conserved energy by the wave function Ψ :

$$\frac{p^2}{2m} + V(x) = E \implies \frac{p^2}{2m}\Psi + V(x)\Psi = E\Psi. \quad (3.1.8)$$

Then Schrödinger postulated that also in this case p and E have to be replaced by the following operators:

$$\begin{aligned} p &\Rightarrow -i\hbar\frac{\partial}{\partial x} \\ E &\Rightarrow i\hbar\frac{\partial}{\partial t}. \end{aligned} \quad (3.1.9)$$

By using (3.1.9) into (3.1.8) Schrödinger obtained the equation of motion for a quantum particle in a potential $V(x)$:

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\Psi(x,t) + V(x)\Psi(x,t) = i\hbar\frac{\partial}{\partial t}\Psi(x,t) \quad (3.1.10)$$

which is a differential equation with partial derivatives, linear in the wave function $\Psi(x, t)$ and satisfying all the four properties indicated at the beginning of this section.

The Schrödinger equation is a non-relativistic equation. In the relativistic case the definition of energy is the following one:

$$E^2 = p^2c^2 + m_0^2c^4 \implies E = \pm\sqrt{p^2c^2 + (m_0c^2)^2}. \quad (3.1.11)$$

The big problem with the relativistic equation is to perform and to give a sense to the operator associated with the square root appearing in (3.1.11). This was done by Dirac and brought to the discovery of the antiparticles.

3.2 Solution of the Time-Dependent Schrödinger Equation

In this section we want to find a method to solve the time-dependent Schrödinger equation:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2} + V(x)\Psi(x, t) = i\hbar \frac{\partial \Psi(x, t)}{\partial t}. \quad (3.2.1)$$

We will use the method of separation of variables, by looking for solutions of the form:

$$\Psi(x, t) = \psi(x)\phi(t). \quad (3.2.2)$$

By replacing (3.2.2) into (3.2.1) we obtain:

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} [\psi(x)\phi(t)] + V(x)[\psi(x)\phi(t)] = i\hbar \frac{\partial}{\partial t} [\psi(x)\phi(t)].$$

Since

$$\frac{\partial^2}{\partial x^2} [\psi(x)\phi(t)] = \phi(t) \frac{d^2\psi(x)}{dx^2}, \quad \frac{\partial}{\partial t} [\psi(x)\phi(t)] = \psi(x) \frac{d\phi(t)}{dt}$$

the Schrödinger equation becomes the following differential equation:

$$-\frac{\hbar^2}{2m} \phi(t) \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x)\phi(t) = i\hbar\psi(x) \frac{d\phi(t)}{dt}$$

without partial derivatives. If we divide the previous equation by $\psi(x)\phi(t)$ we get:

$$\frac{1}{\psi(x)} \left[-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) \right] = i\hbar \frac{1}{\phi(t)} \frac{d\phi(t)}{dt}. \quad (3.2.3)$$

Now the RHS of (3.2.3) depends only on t while the LHS depends only on x . Since the two sides are the same they must be equal to a constant, so:

$$i\hbar \frac{1}{\phi(t)} \frac{d\phi(t)}{dt} = E = \text{const} \quad (3.2.4)$$

$$\frac{1}{\psi(x)} \left[-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) \right] = E. \quad (3.2.5)$$

From (3.2.4) we get:

$$\phi(t) = e^{-iEt/\hbar} \quad (3.2.6)$$

while from (3.2.5) we obtain:

$$-\frac{\hbar^2}{2m} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x). \quad (3.2.7)$$

This is the so called time-independent Schrödinger equation. Let us note that on the LHS of (3.2.7) we have both the kinetic and the potential terms, so the constant E appearing on the RHS is just the energy of the system. The wave functions which solve Eq. (3.2.7) are called energy eigenfunctions. The reason is that we can introduce an operator $\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x)$ and have it act on ψ giving the equation: $\hat{H}\psi(x) = E\psi(x)$ which is the same as (3.2.7). If we look at the operator \hat{H} as a matrix then the equation $\hat{H}\psi = E\psi$ is like $H_i^j v_j = E v_i$ which is the equation which diagonalizes H . The reader may wonder in which sense we can look at $\psi(x)$ as a vector like the v_i appearing in $H_i^j v_j = E v_i$ and how we can look at the differential operator \hat{H} as a matrix. Here we will suggest an intuitive but not rigorous explanation of this fact. Let us suppose we discretize the x -space. Then the $\psi(x)$ becomes a collection of numbers

$$\psi(x) = \left[\psi(x_1), \psi(x_2), \dots, \psi(x_n), \dots \right]$$

and we can consider the numbers $\psi(x_k)$ as the components of a vector with ∞ components. We can apply the same discretization to any differential operator; for example the first order derivative in x_j becomes the following

$$\left. \frac{d}{dx} \psi(x) \right|_{x=x_j} = \frac{\psi(x_j) - \psi(x_{j-1})}{x_j - x_{j-1}} = -\frac{\psi(x_{j-1})}{x_j - x_{j-1}} + \frac{\psi(x_j)}{x_j - x_{j-1}}.$$

If now we want to consider the derivative at every point acting on the full vector $(\psi(x_1), \psi(x_2), \dots)$ we get:

$$\sum_j \begin{pmatrix} 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -\frac{1}{x_j - x_{j-1}} & \frac{1}{x_j - x_{j-1}} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \dots \\ \psi(x_{j-1}) \\ \psi(x_j) \\ \dots \\ 0 \end{pmatrix}. \quad (3.2.8)$$

So we can say that in the discretized version of the x -space the differential operator $\frac{d}{dx}$ acts as a matrix and its expression from (3.2.8) is:

$$A_{ij} = -\delta_{i,j-1} \frac{1}{x_j - x_{j-1}} + \delta_{i,j} \frac{1}{x_j - x_{j-1}}.$$

The same can be done for the differential operator $\hat{H} = -\hbar^2 \frac{d^2}{dx^2} + V(x)$ which will become an ∞ matrix. In this sense the equation $\hat{H}\psi = E\psi$ is analogous to the equation which diagonalizes H : $H_i^j v_j = E v_i$.

Homework: exercise 3. Calculate the discretized form of the second derivative with respect to x .

If the reader does not like the intuitive, but non-rigorous, reasoning we have presented above to explain why $\psi(x)$ can be considered an ∞ vector, he could instead use the following idea. The wave function $\psi(x)$ can be expanded on a “basis”, given for example by the plane waves

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int dk e^{ikx} \tilde{\psi}(k). \quad (3.2.9)$$

In Eq. (3.2.9) the $\tilde{\psi}(k)$ are the coefficients of the expansion of $\psi(x)$ on the basis of the plane waves $\frac{1}{\sqrt{2\pi}} e^{ikx}$. In this sense $\psi(x)$ can be considered an ∞ -dimensional vector.

3.3 Properties of the Solutions

Linearity. As we said in Sec 2.1 the Schrödinger equation is linear in the wave function: this means that if Ψ_1 and Ψ_2 are two solutions of the Schrödinger equation (3.1.10) then also

$$\Psi(x, t) = c_1 \Psi_1 + c_2 \Psi_2$$

is a solution of the same equation. In fact

$$\begin{aligned} & -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x)\Psi - i\hbar \frac{\partial \Psi}{\partial t} = \\ & = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} [c_1 \Psi_1 + c_2 \Psi_2] + V(x) [c_1 \Psi_1 + c_2 \Psi_2] - i\hbar \frac{\partial}{\partial t} [c_1 \Psi_1 + c_2 \Psi_2] = \\ & = c_1 \left[-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi_1}{\partial x^2} + V\Psi_1 - i\hbar \frac{\partial \Psi_1}{\partial t} \right] + c_2 \left[-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi_2}{\partial x^2} + V\Psi_2 - i\hbar \frac{\partial \Psi_2}{\partial t} \right] = 0 \end{aligned}$$

where in the last step we have used the fact both Ψ_1 and Ψ_2 are solutions of the Schrödinger equation (3.1.10).

Constant phases. What we want to prove now is that if $\Psi(x, t)$ is a solution of the Schrödinger equation then also $\Psi_\alpha(x, t) = \Psi(x, t)e^{i\alpha}$ where α is a constant phase solves the same equation. The hypothesis is:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x)\Psi = i\hbar \frac{\partial \Psi}{\partial t}. \quad (3.3.1)$$

By replacing Ψ with $\Psi_\alpha = \Psi e^{i\alpha}$ into (3.3.1) we obtain that:

$$\begin{aligned} & -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi_\alpha}{\partial x^2} + V(x)\Psi_\alpha - i\hbar \frac{\partial \Psi_\alpha}{\partial t} = \\ & e^{i\alpha} \left[-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x)\Psi - i\hbar \frac{\partial \Psi}{\partial t} \right] = 0. \end{aligned}$$

Consequently, if Ψ is a solution of the Schrödinger equation, then also all the functions of the form $\Psi_\alpha = \Psi e^{i\alpha}$ are solutions of the same equation. This set of solutions make

a set called “ray”. Usually one takes into account only one element Ψ_α for each *ray*. In making a sum, the representatives of the rays must not be changed anymore:

$$\Psi = \Psi_1 e^{i\alpha_1} + \Psi_2 e^{i\beta_2}. \quad (3.3.2)$$

In fact if we modify by a phase the states in a sum like the one appearing in (3.3.2), we obtain a different wave function:

$$\Psi' = \Psi_1 e^{i\alpha_1} e^{i\gamma} + \Psi_2 e^{i\beta_2} e^{i\delta}.$$

In general the new function Ψ' does not belong to the same ray of Ψ , i.e. there does *not* exist any ϵ such that $\Psi' \neq \Psi e^{i\epsilon}$. So the sum of two rays may not be a ray while the sum of two states is a state.

Properties of the wave functions. Since the wave function $\Psi(x, t)$ is related with the probability density of finding a particle in x at time t : $P(x, t) = |\Psi(x, t)|^2$ we must require that Ψ be normalized, i.e. $\int dx |\Psi|^2 = \int dx P(x, t) = 1$, and also that:

1. $\Psi(x, t)$ be finite for every value of x .
2. $\Psi(x, t)$ be single-valued.
3. $\Psi(x, t)$ be continuous.

This is due to the fact that if $\Psi(x, t)$ does not satisfy the properties above then also the probability will not satisfy such properties but this cannot happen because the probabilities must be finite in every point, single-valued and continuous.

Derivative of Ψ and probability current. The three properties mentioned above must be satisfied also by the derivative $\frac{\partial \Psi}{\partial x}$. Such a quantity in fact is related with the *probability current*. To show that let us take again the Schrödinger equation for Ψ :

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + V(x)\Psi. \quad (3.3.3)$$

and the one for the complex conjugate Ψ^* :

$$-i\hbar \frac{\partial \Psi^*}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi^*}{\partial x^2} + V(x)\Psi^*. \quad (3.3.4)$$

If we multiply (3.3.3) by Ψ^* and (3.3.4) by Ψ and we make the difference of the results we get:

$$\begin{aligned} i\hbar \Psi^* \frac{\partial \Psi}{\partial t} + i\hbar \Psi \frac{\partial \Psi^*}{\partial t} &= -\frac{\hbar^2}{2m} \Psi^* \frac{\partial^2 \Psi}{\partial x^2} + \frac{\hbar^2}{2m} \Psi \frac{\partial^2 \Psi^*}{\partial x^2} \\ \implies i\hbar \frac{\partial}{\partial t} (\Psi^* \Psi) &= -\frac{\hbar^2}{2m} \frac{\partial}{\partial x} \left[\Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x} \right] \\ \implies i\hbar \frac{\partial}{\partial t} P(x, t) + \frac{\hbar^2}{2m} \frac{\partial}{\partial x} \left[\Psi^* \frac{\partial \Psi}{\partial x} - \Psi \frac{\partial \Psi^*}{\partial x} \right] &= 0. \end{aligned}$$

This equation of evolution for the probability density $P(x, t)$ can be written as:

$$\frac{\partial}{\partial t} P(x, t) + \frac{\partial}{\partial x} J(x, t) = 0 \quad (3.3.5)$$

where $J(x) = -\frac{i\hbar}{2m} \left[\Psi^* \frac{\partial \Psi}{\partial x} - \left(\frac{\partial \Psi^*}{\partial x} \right) \Psi \right]$ is the probability current. Since the derivative $\frac{\partial \Psi}{\partial x}$ enters into the current $J(x)$ which, due to its physical meaning, is finite, continuous and single-valued, we must require that also $\frac{\partial \Psi}{\partial x}$ be finite, continuous and single-valued. Sometimes however this condition is relaxed, e.g. when we have Dirac-delta kind of potentials.

In three dimensions Eq. (3.3.5) becomes:

$$\frac{\partial}{\partial t} P(\vec{r}, t) + \text{div} \vec{J} = 0, \quad \vec{J} = -\frac{i\hbar}{2m} [\Psi^* \vec{\nabla} \Psi - (\vec{\nabla} \Psi^*) \Psi] = 0. \quad (3.3.6)$$

This is the same equation satisfied by the charge density $\rho(\vec{r}, t)$ in electrodynamics. In that case, when we integrate the continuity equation over a volume V , we obtain just the equation of evolution for the charge Q contained in the volume:

$$\frac{d}{dt} \int_V d^3x \rho(\vec{r}, t) + \int_V d^3x \text{div} \vec{J} = 0 \Rightarrow \frac{dQ}{dt} + \oint_S \vec{J} \cdot \vec{n} = 0. \quad (3.3.7)$$

With this analogy in mind Schrödinger first interpreted $|\Psi(\vec{r}, t)|^2$ as a charge density. However this interpretation is not correct because during the evolution the wave function Ψ can spread considerably while such a phenomenon was never observed for electric charges. This is the reason why Schrödinger's interpretation of the wave function was soon abandoned in favour of the one of Born.

Problem 6. Prove that the following wave function

$$\Psi(x, t) = A e^{-(\sqrt{Cm}/2\hbar)x^2} e^{-i/2\sqrt{C/mt}} \quad (3.3.8)$$

solves the Schrödinger equation for a harmonic oscillator with potential $V(x) = Cx^2/2$. Then calculate the probability density of finding a particle in x if the system is described by the wave function (3.3.8).

Solution. We must verify that:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial x^2} + \frac{C}{2} x^2 \Psi = i\hbar \frac{\partial \Psi}{\partial t}. \quad (3.3.9)$$

First of all let us consider the wave function (3.3.8) and calculate its derivatives. Since (3.3.8) is an exponential in t its derivative with respect to time gives:

$$\frac{\partial \Psi}{\partial t} = -\frac{i}{2} \sqrt{\frac{C}{m}} \Psi. \quad (3.3.10)$$

Analogously the derivative with respect to x gives:

$$\frac{\partial \Psi}{\partial x} = -\frac{\sqrt{Cm}}{\hbar} x \Psi \implies \frac{\partial^2 \Psi}{\partial x^2} = -\frac{\sqrt{Cm}}{\hbar} \Psi + \frac{Cm}{\hbar^2} x^2 \Psi. \quad (3.3.11)$$

By replacing (3.3.10) and (3.3.11) into (3.3.9) we get:

$$\begin{aligned} \frac{\hbar^2}{2m} \frac{\sqrt{Cm}}{\hbar} \Psi - \frac{\hbar^2}{2m} \frac{Cm}{\hbar^2} x^2 \Psi + \frac{C}{2} x^2 \Psi &= i\hbar \left(-\frac{i}{2}\right) \sqrt{\frac{C}{m}} \Psi \\ &\Downarrow \\ \frac{\hbar}{2} \sqrt{\frac{C}{m}} \Psi - \frac{C}{2} x^2 \Psi + \frac{C}{2} x^2 \Psi &= \frac{\hbar}{2} \sqrt{\frac{C}{m}} \Psi \end{aligned}$$

which is identically satisfied.

If the wave function describing the system is (3.3.8), then the probability density of finding a particle in x is:

$$P(x, t) = \Psi^* \Psi = A^2 \exp \left[-\frac{\sqrt{Cm}}{\hbar} x^2 \right]. \quad (3.3.12)$$

Such a probability density does not depend on t ; this is a common feature of all the wave functions corresponding to fixed values of energy. So if our system is described by the state (3.3.8) then the position of the particle is spread. We cannot know with absolute precision the position of the particle but only the probability of finding it at a certain point if we perform a measurement.

Let us briefly compare the classical and the quantum harmonic oscillators. In the classical case the probability of finding a particle in a point x is inversely proportional to the velocity of the particle in that point:

$$P(x) dx \propto \frac{dx}{v(x)}. \quad (3.3.13)$$

If the oscillator moves according to the law $x(t) = x_0 \sin(\omega t)$ where $x_0 = \sqrt{\frac{2E}{C}}$ then its velocity is

$$v(t) = x_0 \omega \cos(\omega t) = x_0 \omega \sqrt{1 - \sin^2 \omega t} \implies v(x) = \omega \sqrt{x_0^2 - x^2} \quad (3.3.14)$$

and the probability density of finding the particle in x is given by:

$$P(x) \propto \frac{1}{\sqrt{x_0^2 - x^2}} \quad (3.3.15)$$

which is plotted in Fig. 3.1.

The velocity of the classical particle is maximum when it goes through the origin and therefore the probability of finding the particle near the origin is minimum; the probability instead is maximum at the extrema of the oscillation range $\pm \sqrt{2E/C}$, where the velocity is practically zero. In the quantum case the situation is just the opposite: the particle can be found also outside the extrema of the oscillation range and the probability is maximum just near the origin, see Fig. 3.1.

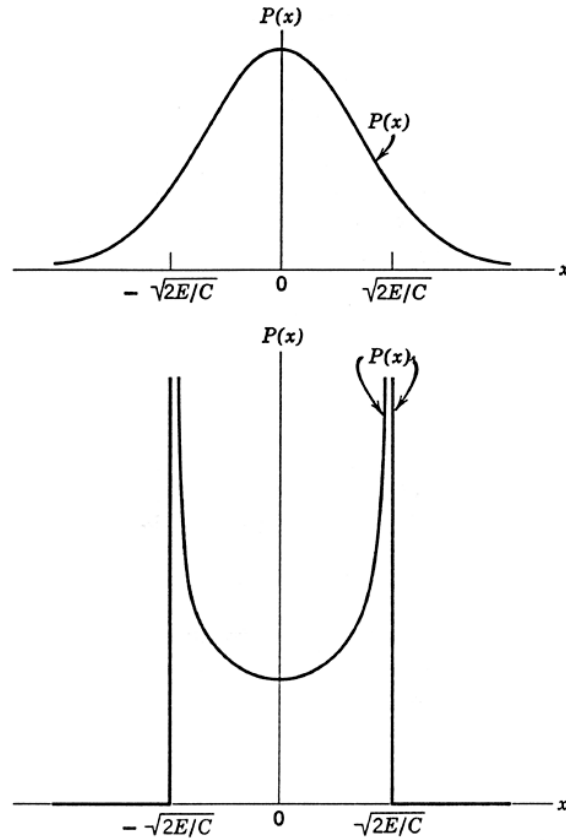


Fig. 3.1: Quantum and classical probability densities for a harmonic oscillator.

3.4 Quantization of the Energy Levels

What we want to do now is to solve the time-independent Schrödinger equation (3.2.7) and to prove that, for some particular values of the potential $V(x)$, the solutions exist only for certain values of E (energy is discretized). First of all let us make a qualitative analysis of the solutions of the eigenvalue equation:

$$\frac{d^2\psi}{dx^2} = \frac{2m}{\hbar^2}[V(x) - E]\psi \quad (3.4.1)$$

for the potential plotted in Fig 3.2. . As boundary conditions we give the value of $\psi(x)$ and of its derivative $\psi'(x)$ in $x = x_0$ with $x' < x_0 < x''$.

Let us suppose that $\psi(x) > 0$ for every point x in the interval (x_0, x'') . Therefore from (3.4.1) we derive that in such an interval the function is concave downwards. After the point x'' the wave function goes into a region where $V(x) - E > 0$, and so from (3.4.1) it becomes concave upwards and it remains concave upwards up to $+\infty$. As we see from the curve 1 of Fig. 3.3 in this case the wave function diverges towards $+\infty$. If we change the value of the derivative $\psi'(x)$ in x_0 we can change the initial slope of the curve. If we decrease the slope too much we obtain the curve 2 of Fig. 3.3. In this case the wave function ψ intersects the x axis and it becomes negative; after the point of intersection we have that $\frac{d^2\psi}{dx^2} < 0$, i.e. the wave function is concave downwards and it diverges

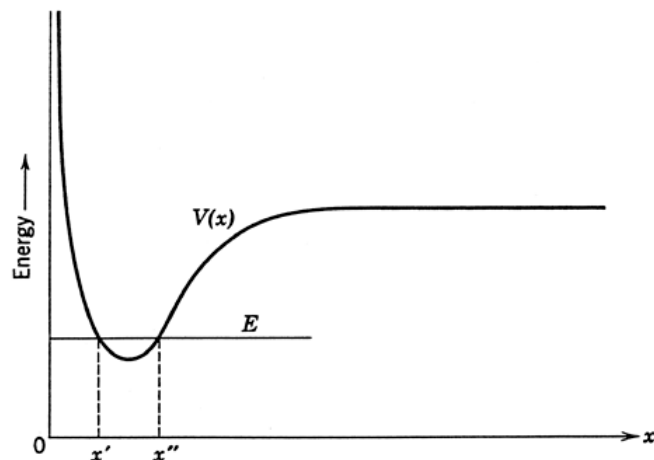


Fig. 3.2: The potential energy $V(x)$ used in the qualitative analysis of the Schrödinger equation, and the total energy E chosen for the analysis.

towards $-\infty$. For a suitable initial value of the derivative $\psi'(x)$ it is possible to have the wave function 3 of Fig. 3.3 which is concave upwards, it remains above the x axis and tends to zero asymptotically at $+\infty$. Such a wave function has a behavior at $+\infty$ that allows it to be normalizable. Nevertheless if we plot with a dashed curve the behavior of the curve 3 for decreasing values of x , we discover from Eq. (3.4.1) that such a curve diverges at $-\infty$, see Fig. 3.3. Therefore we must conclude that for the particular choice of the energy E that we made there are no acceptable solutions of the time-independent Schrödinger equation (3.4.1).

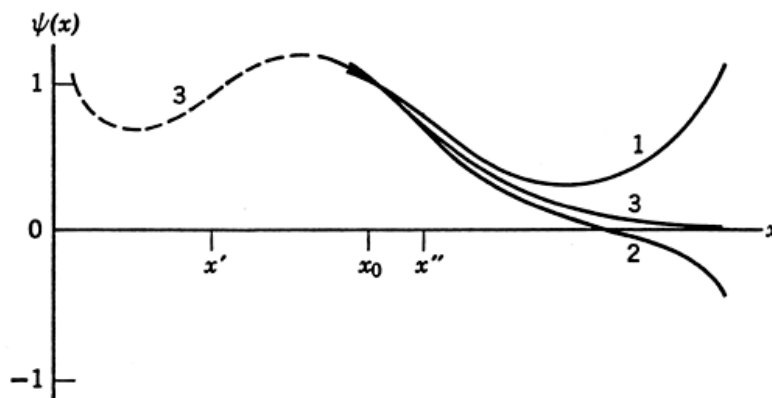


Fig. 3.3: Three attempts of finding an acceptable solution to the time-independent Schrödinger equation.

By repeating this kind of reasoning for all the possible values of E we will discover that only a discrete set of values of the energy E gives origin to acceptable solutions for the time-independent Schrödinger equation. So we can say that the discretization of the energy levels is just a consequence of the request of normalizability of the wave function at $\pm\infty$ and of the continuity of the wave function and its derivative.

Let us now give a more quantitative analysis of the eigenvalue equation. First of all

let us write the Schrödinger equation (3.4.1) as:

$$y'' + [\tilde{E} - U(x)]y = 0 \quad (3.4.2)$$

where $\tilde{E} = 2mE/\hbar^2$ and $U(x) = 2mV(x)/\hbar^2$. Let us take $U(x)$ bounded from below and continuous. Given two solutions $y_1(x)$ and $y_2(x)$ we can define the following object known as Wronskian:

$$W(y_1, y_2) \equiv y_1 y_2' - y_2 y_1'. \quad (3.4.3)$$

If at a certain point $x = a$ the Wronskian (3.4.3) is zero then $y_1(a)y_2'(a) = y_2(a)y_1'(a)$.

This means that $\frac{y_2'(a)}{y_2(a)} = \frac{y_1'(a)}{y_1(a)}$, i.e. the logarithmic derivatives of the two solutions are the same.

Wronskian theorem. If $z_1(x)$ and $z_2(x)$ are solutions of the following equations:

$$z_1'' + F_1(x)z_1 = 0 \quad (3.4.4)$$

$$z_2'' + F_2(x)z_2 = 0 \quad (3.4.5)$$

then the variation of the associated Wronskian in the interval (a, b) is given by:

$$W(z_1, z_2) \Big|_a^b = \int_a^b dx [F_1(x) - F_2(x)] z_1 z_2. \quad (3.4.6)$$

Proof. If we multiply (3.4.4) by z_2 and (3.4.5) by z_1 and calculate the difference of the two results we obtain that:

$$[z_2 z_1'' - z_1 z_2''] + (F_1 - F_2) z_1 z_2 = 0. \quad (3.4.7)$$

The first term is, except for a sign, the derivative with respect to x of the Wronskian $W(z_1, z_2)$. So if we integrate (3.4.7) in x we obtain just (3.4.6).

Corollary I. If z_1 and z_2 are two solutions of the Schrödinger equation (3.4.2) corresponding to two different energies \tilde{E}_1 and \tilde{E}_2 then $F_1(x) = \tilde{E}_1 - U(x)$; $F_2(x) = \tilde{E}_2 - U(x)$ and (3.4.6) becomes:

$$W(z_1, z_2) \Big|_a^b = (\tilde{E}_1 - \tilde{E}_2) \int_a^b dx z_1 z_2. \quad (3.4.8)$$

Corollary II. If the solutions are associated with the same energy, i.e. $\tilde{E}_1 = \tilde{E}_2$, then we get from (3.4.8) that the Wronskian is independent of x : $W(z_1, z_2) = \text{const.}$

Corollary III. If $y(x, \tilde{E})$ is a solution of the Schrödinger equation (3.4.2) and F defined as

$$F(x, \tilde{E}) \equiv \frac{y'(x, \tilde{E})}{y(x, \tilde{E})}$$

is its logarithmic derivative assuming a particular value f_a at the point a , then $F(x, \tilde{E})$ is a monotonic function of \tilde{E} , increasing if $x < a$ and decreasing if $x > a$. In particular its derivative is given by:

$$\frac{\partial F(x, \tilde{E})}{\partial \tilde{E}} = -\frac{1}{y^2(x; \tilde{E})} \int_a^x d\xi y^2(\xi; \tilde{E}). \quad (3.4.9)$$

From (3.4.9) we have that the derivative is negative if $x > a$ and positive if $x < a$. This confirms that F is a monotonic function of \tilde{E} .

Proof of Corollary III. Once we fix \tilde{E} a solution of (3.4.2) is determined by its value and the value of its derivative at a given point $x = a$. Let us fix as boundary conditions $y(a, \tilde{E}) = y_a$ and $y'(a, \tilde{E}) = y'_a$ and replace \tilde{E} with $\tilde{E} + \delta\tilde{E}$. There will be two nearby solutions y and $y + \delta y$ satisfying the same boundary conditions. According to Corollary I at first order in δ we obtain:

$$W(y, y + \delta y) \Big|_a^b = -\delta\tilde{E} \int_a^b dx y^2. \quad (3.4.10)$$

In $x = a$ the Wronskian is $W(y, y + \delta y) = 0$ because the two solutions satisfy the same boundary conditions. From (3.4.3) we have that for every other value of $x \in (a, b)$:

$$W(y, y + \delta y) = W(y, \delta y) = y\delta y' - y'\delta y = y^2\delta(y'/y) = y^2\delta F. \quad (3.4.11)$$

By comparing (3.4.10) and (3.4.11) we get:

$$-y^2\delta F \Big|_{x=b} = \delta\tilde{E} \int_a^b dx y^2. \quad (3.4.12)$$

Consequently:

$$\frac{\delta F}{\delta\tilde{E}} \Big|_{x=b} = -\frac{1}{y^2(b)} \int_a^b dx y^2(x). \quad (3.4.13)$$

Since b is arbitrary we can conclude that F is a monotonic (increasing or decreasing) function of \tilde{E} . The properties of the solution of the Schrödinger equation contained in the three corollaries above are completely independent of the potential $U(x)$.

Let us now take the potential $U(x)$ plotted in Fig. 3.4 with $U_- > U_+$.

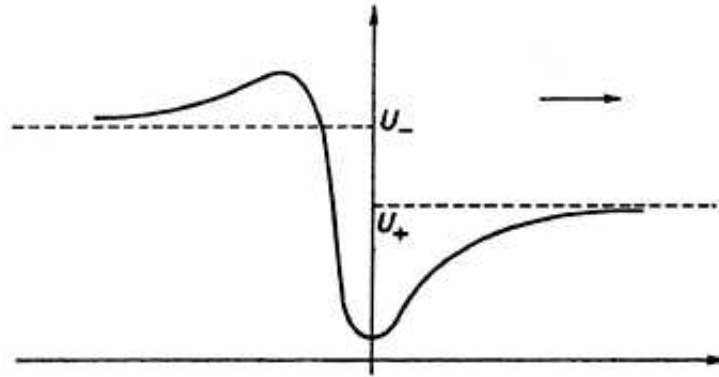


Fig. 3.4: Potential $U(x)$ with asymptotic behavior at $\pm\infty$.

I case. If $\tilde{E} > U_- > U_+$ then $\tilde{E} - U$ is positive at the two extrema and constant. Therefore at $+\infty$ Eq. (3.4.2) becomes:

$$y'' + [\tilde{E} - U_+]y = 0 \implies y'' = -ky \quad (3.4.14)$$

where $k = \tilde{E} - U_+ > 0$ and the asymptotic behavior of y at $+\infty$ is given by $y \sim e^{i\sqrt{k}x}$. At $-\infty$ the Schrödinger equation becomes:

$$y'' + [\tilde{E} - U_-]y = 0 \implies y'' = -\tilde{k}y \quad (3.4.15)$$

with $\tilde{k} = \tilde{E} - U_- > 0$. So for $x \rightarrow -\infty$ the solution y is given by $y \sim e^{i\sqrt{\tilde{k}}x}$. We can conclude that the solutions are bounded and oscillating everywhere. From the asymptotic behavior of the solutions there is no constraint on energy, *every value of energy* is allowed and the spectrum is continuous.

II case. If $U_- > \tilde{E} > U_+$ then $\tilde{E} - U$ is negative for $x \rightarrow -\infty$ and

$$y'' + [\tilde{E} - U]y = 0 \implies y'' = -ky. \quad (3.4.16)$$

Since $k = \tilde{E} - U_- < 0$ the solution of (3.4.16) is:

$$y'' = |k|y \implies y \sim e^{\sqrt{|k|x}}, \quad (3.4.17)$$

which tends to zero for $x \rightarrow -\infty$ and therefore can be accepted. For $x \rightarrow +\infty$ we have instead:

$$y'' + [\tilde{E} - U]y = 0 \implies y'' = -\tilde{k}y. \quad (3.4.18)$$

Since in this case $\tilde{k} = \tilde{E} - U_+$ is positive we will have

$$y'' = -|\tilde{k}|y \implies y \sim e^{i\sqrt{\tilde{k}}x} \quad (3.4.19)$$

which is an oscillating function. Also in this case the spectrum of energy is continuous and there is no restriction on the possible values of \tilde{E} .

III case. If $U_- > U_+ > \tilde{E}$ we have at both infinities that

$$y'' + (\tilde{E} - U_{\pm})y = 0 \implies y'' = -k^{\pm}y \implies y^{\pm}(x) \sim e^{\mp\sqrt{k^{\pm}}x}. \quad (3.4.20)$$

So in this case the solutions of the Schrödinger equation are bounded and dumped at both the extrema. In order to get the spectrum of energy we have to match the two functions of Eq. (3.4.20). Let us consider the logarithmic derivatives of the two solutions $F_+(x, \tilde{E})$ and $F_-(x, \tilde{E})$ and let us put them equal at a certain point $x = a$ (this is so because we want to have only one solution $y(x)$ defined everywhere). Now, according to Corollary III, for $x > a$ we have that F_+ is a decreasing function of \tilde{E} . For $x < a$ instead F_- is an increasing function of \tilde{E} . Now F_- and F_+ must be equal in $x = a$; this can happen only for some particular values \tilde{E}_n of \tilde{E} . This is the proof that the eigenvalues of energy in this case are only certain fixed and isolated numbers \tilde{E} and the spectrum is *discrete*.

3.5 Schrödinger Equation for a Free Particle

As a first application of the Schrödinger equation let us consider the case of a free particle. In this case the potential is identically zero: $V(x) = 0$ and the Schrödinger equation (3.2.1) becomes:

$$i\hbar \frac{\partial \Psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \Psi(x, t)}{\partial x^2}.$$

If we use the method of separation of the variables analyzed in Sec. 2.2, i.e. if we put $\Psi(x, t) = \psi(x)\phi(t)$, then the Schrödinger equation decouples into an equation for ψ and an equation for ϕ . The equation for $\phi(t)$ is easily solved and, according to (3.2.6), gives:

$$\phi(t) = e^{-iEt/\hbar} = e^{-i\omega t} \quad (3.5.1)$$

where the angular frequency ω is given by: $\omega = E/\hbar$. The $\psi(x)$ satisfies instead the following equation:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} = E\psi \implies \frac{\partial^2 \psi}{\partial x^2} = -\frac{2mE}{\hbar^2} \psi = -\tilde{E}\psi. \quad (3.5.2)$$

The solutions of (3.5.2) are:

$$\psi_1(x) = e^{i\sqrt{\tilde{E}}x}, \quad \psi_2(x) = e^{-i\sqrt{\tilde{E}}x}$$

and all their linear combinations:

$$\psi(x) = Ae^{i\sqrt{\tilde{E}}x} + Be^{-i\sqrt{\tilde{E}}x}. \quad (3.5.3)$$

Now $\sqrt{\tilde{E}}$ is just the wave number k . In fact:

$$\sqrt{\tilde{E}} = \frac{\sqrt{2mE}}{\hbar} = \frac{1}{\hbar} \sqrt{2m \frac{p^2}{2m}} = p/\hbar = 1/\lambda = k.$$

Therefore the general solution (3.5.3) of the Schrödinger equation for a free particle becomes:

$$\psi(x) = Ae^{ikx} + Be^{-ikx}. \quad (3.5.4)$$

If we include also the dependence (3.5.1) on time we get:

$$\Psi(x, t) = Ae^{ikx-i\omega t} + Be^{-ikx-i\omega t}. \quad (3.5.5)$$

Let us analyze separately the two terms appearing on the RHS of (3.5.5). The first one represents a plane wave propagating towards $+x$. In fact it can be rewritten as:

$$\Psi(x, t) = A[\cos(kx - \omega t) + i \sin(kx - \omega t)].$$

From the previous equation it is easy to see that the nodes of the real part, which are the solutions of the equation:

$$kx_n - \omega t = \left(n + \frac{1}{2}\right)\pi \implies x_n = \left(n + \frac{1}{2}\right)\frac{\pi}{k} + \omega \frac{t}{k},$$

propagate in time towards increasing values of x . For the same value of energy $E = \hbar\omega$ we have also the wave propagating in the opposite direction given by the second term on the RHS of (3.5.5):

$$\Psi(x, t) = B e^{i(-kx - \omega t)}.$$

Therefore we have two *different* plane waves, corresponding to two *different* solutions of (3.5.2) associated with the *same* value of the energy E . This phenomenon is called *degeneracy*.

Let us also notice that in this case the energy E can assume every *continuous value*, i.e. the boundary conditions cannot limit the values of energy. In fact let us give the value of the function and of its derivative at the point a :

$$\psi(a) = \psi_a, \quad \psi'(a) = \psi'_a. \quad (3.5.6)$$

From the expression (3.5.4) of the wave function we obtain that at the point a :

$$\begin{aligned} \psi_a &= Ae^{ika} + Be^{-ika} \\ \psi'_a &= ikAe^{ika} - ikBe^{-ika}. \end{aligned}$$

Let us see whether we can express A and B as functions of ψ_a and ψ'_a :

$$\begin{aligned} \psi_a &= Ae^{ika} + Be^{-ika} \\ \frac{\psi'_a}{ik} &= Ae^{ika} - Be^{-ika}. \end{aligned}$$

By making the sum and the difference of the two equations above we can find A and B in terms of the boundary conditions (3.5.6) independently of the values of k and, consequently, of the energy $E = \frac{\hbar^2 k^2}{2m}$:

$$\begin{aligned} A &= \left[\psi_a + \frac{\psi'_a}{ik} \right] \frac{e^{-ika}}{2} \\ B &= \left[\psi_a - \frac{\psi'_a}{ik} \right] \frac{e^{ika}}{2}. \end{aligned} \quad (3.5.7)$$

So in the case of a free particle there is no constraint in the value of energy.

3.6 One-dimensional Problems

Problem 8. Solve the time-independent Schrödinger equation for the infinite square well potential $V(x) = \begin{cases} \infty & x < -a/2 \text{ or } x > a/2 \\ 0 & -a/2 < x < a/2 \end{cases}$ plotted in Fig. 3.5.

Solution. At the *classical* level a particle in a box can assume *every value of the energy*. When the particle arrives at one of the two walls it bounces back. For example at the wall on the right the momentum changes from $+p$ to $-p$ in an interval of time which is infinitely small.

Let us instead analyze the situation at the *quantum* level. Between the two walls the particle behaves like if it were free and so we have the same wave function given in (3.5.4):

$$\psi(x) = A'e^{ikx} + B'e^{-ikx} = A \sin kx + B \cos kx \quad (3.6.1)$$

with $k = \sqrt{2mE}/\hbar$. The difference is that in this case the potential outside the walls is ∞ . This means that the particle cannot go out the walls and therefore the wave function

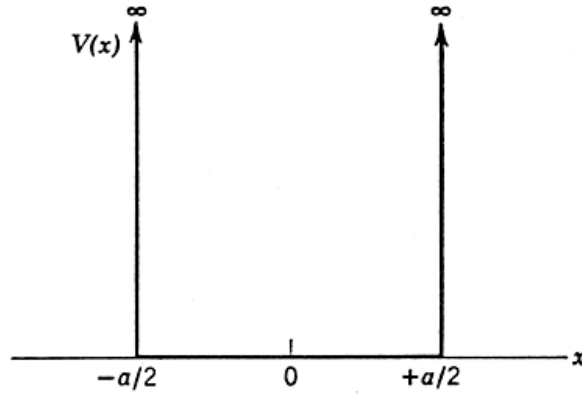


Fig. 3.5: Infinite square well potential.

must be identically zero for $x < -a/2$ and $x > a/2$. Since the wave function must be continuous we request it to be zero not only outside but also at the walls:

$$\begin{cases} \psi(a/2) = 0 \\ \psi(-a/2) = 0. \end{cases} \quad (3.6.2)$$

If we impose the boundary conditions (3.6.2) on the wave function (3.6.1) we obtain:

$$\begin{cases} A \sin \frac{ka}{2} + B \cos \frac{ka}{2} = 0 \\ A \sin(-\frac{ka}{2}) + B \cos(-\frac{ka}{2}) = 0 \end{cases}$$

which implies

$$\begin{cases} A \sin \frac{ka}{2} + B \cos \frac{ka}{2} = 0 \\ -A \sin \frac{ka}{2} + B \cos \frac{ka}{2} = 0. \end{cases}$$

By making the sum and the difference of the two equations above we obtain:

$$\begin{cases} 2B \cos \frac{ka}{2} = 0 \\ 2A \sin \frac{ka}{2} = 0. \end{cases} \quad (3.6.3)$$

There is no value of k which makes both the sine and the cosine equal to zero. We cannot put $A = B = 0$ because in this case the wave function would be identically zero. Then the only possible solutions of (3.6.3) are:

$$A = 0, \quad \cos \frac{ka}{2} = 0 \quad (3.6.4)$$

and

$$B = 0, \quad \sin \frac{ka}{2} = 0. \quad (3.6.5)$$

Eq. (3.6.4) implies:

$$\psi(x) = B \cos kx$$

and

$$\frac{ka}{2} = \frac{\pi}{2}(2\bar{n} + 1) \implies k_{\bar{n}} = (2\bar{n} + 1)\frac{\pi}{a}, \quad \bar{n} = 0, 1, 2, \dots \quad (3.6.6)$$

Eq. (3.6.5) instead leads to:

$$\psi(x) = A \sin kx$$

and

$$\frac{ka}{2} = 2\tilde{n}\frac{\pi}{2} \implies k_{\tilde{n}} = 2\tilde{n}\frac{\pi}{a}, \quad \tilde{n} = 1, 2, \dots \quad (3.6.7)$$

In the previous case $\tilde{n} = 0$ would imply $k = 0$, i.e. an identically zero wave function which is a non-sense. By putting (3.6.6) and (3.6.7) together we have that all the possible values of k are of the form $k_n = n\pi/a$ with $n = 1, 2, \dots$. Since $k = \sqrt{2mE}/\hbar$ the allowed values of the energy are given by:

$$E_n = \frac{\hbar^2 k_n^2}{2m} = \frac{\hbar^2 \pi^2}{2ma^2} n^2. \quad (3.6.8)$$

This proves that the energy is quantized. The ground state, given by the lowest of the allowed energies, is:

$$E_1 = \frac{\pi^2 \hbar^2}{2ma^2}. \quad (3.6.9)$$

Therefore the particle cannot have zero energy (as we have seen before in such a case the wave function would be zero and the wave function would not exist). The fact that the zero point energy is not zero is just a consequence of the *uncertainty principle*. In fact, since the particle is confined by the potential to be between $-a/2$ and $a/2$, it means that we know its position with an error $\Delta x \approx a$. As a consequence the uncertainty in the momentum is at least $\Delta p \approx \frac{\hbar}{2\Delta x} \approx \frac{\hbar}{2a}$. The uncertainty principle does not allow the particle to have zero total energy since that would mean that the uncertainty in the momentum is zero violating the bound above. For the particular case of the first energy level (3.6.9) the momentum is given by:

$$p_1 = \sqrt{2mE_1} = \frac{\pi\hbar}{a}. \quad (3.6.10)$$

Since the particle can move in both directions the uncertainty on the momentum is about $\Delta p = 2p_1 \approx 2\pi\hbar/a$ and therefore $\Delta x \Delta p \approx a 2\pi\hbar/a = 2\pi\hbar$ which is in quite good agreement with the lowest limit, $\hbar/2$, set by the uncertainty principle. The existence of a zero-point energy has a lot of important physical consequences: for example the helium remains liquid and does not solidify at very low temperatures (about 0.001 °K) just because of the zero-point motion.

The first three eigenfunctions of the infinite square well potential are plotted in Fig. 3.6. Let us notice that, since the number of half wavelengths of each eigenfunction is given by the quantum number n , the number of nodes is $n + 1$. It is like having the vibrating strings of a violin, fixed at the two extrema. The difference is that in the case of the infinite square well, since the energies are given by (3.6.8), the allowed frequencies increase with the square of n : $\nu = E/h \sim n^2$. In the violin instead the frequencies follow

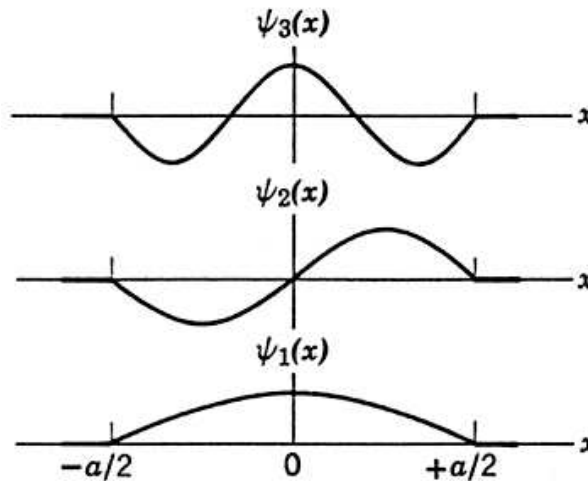


Fig. 3.6: First eigenfunctions of the infinite square well potential.

a “harmonic progression”: $\nu \sim n$. This is due to the fact that the wave equations in the two cases are different.

Problem 9. Before the discovery of the neutron it was thought that a nucleus of atomic number Z and atomic weight A was composed of A protons and $A - Z$ electrons (total charge: $A - (A - Z) = Z$). The electrons were imagined as confined by a potential in a very small region like the one occupied by the nucleus. Calculate the zero point energy for the electrons, considering as potential the one analyzed in the previous problem.

Solution. Let us put the mass of the electron $m \sim 10^{-30}$ Kg and the width of the well $a \sim 10^{-14}$ m. The zero point energy becomes:

$$\begin{aligned} E_1 &= \frac{\pi^2 \hbar^2}{2ma^2} \approx \frac{10 \times 10^{-68} \text{ joule}^2 \cdot \text{sec}^2}{2 \times 10^{-30} \text{ Kg} \times 10^{-28} \text{ m}^2} \\ &\simeq \frac{10^{-9}}{2} \text{ joule} \sim 10^9 \text{ eV} = 10^3 \text{ MeV}. \end{aligned} \quad (3.6.11)$$

The zero point energy becomes very large because of the smallness of the mass m of the electron. Let us see whether the binding potential is greater than the zero point energy: only in this case an electron could be found in a nucleus with such a zero point energy. The potential is given by the electrostatic attraction electron-protons. If we take, as number of protons, $A = 100$ and as mean distance between the electron and the protons $r = 10^{-14}$ m we obtain:

$$\frac{Q_1 Q_2}{4\pi\epsilon_0 r} = -\frac{Ae^2}{4\pi\epsilon_0 r} \simeq -10 \text{ MeV} \quad (3.6.12)$$

which is one hundred times smaller than the required binding energy. So the picture of a nucleus composed of electrons cannot be correct. Because of its zero point energy an electron cannot be bound in a nucleus by the electrostatic potential.

Homework: exercise 4. Check the numerical value of the potential (3.6.12).

Problem 10. Analyze a quantum particle in a gravitational field. In particular calculate the energy eigenfunctions and eigenvalues for a particle which is so near the surface of the earth, that the gravitational potential can be approximated with:

$$U(z) = \begin{cases} mgz, & z \geq 0 \\ \infty, & z < 0. \end{cases} \quad (3.6.13)$$

The time-independent Schrödinger equation for the potential (3.6.13) becomes:

$$\ddot{\psi}(z) + \frac{2m}{\hbar^2} [E - mgz] \psi(z) = 0. \quad (3.6.14)$$

Such an equation can be rewritten as:

$$\ddot{\psi}(z) + \frac{2m^2g}{\hbar^2} \left(\frac{E}{mg} - z \right) \psi(z) = 0. \quad (3.6.15)$$

Let us define a new variable:

$$x \equiv \left(\frac{2m^2g}{\hbar^2} \right)^{1/3} \left(z - \frac{E}{mg} \right). \quad (3.6.16)$$

With this replacement the wave function can be written as $\tilde{\psi}(x) = \psi(z(x))$ and its second derivative becomes:

$$\begin{aligned} \ddot{\psi}(z) &= \frac{d^2\psi(z)}{dz^2} = \frac{d}{dz} \left(\frac{d\tilde{\psi}}{dx} \frac{dx}{dz} \right) = \\ &= \frac{dx}{dz} \frac{d}{dx} \left(\frac{d\tilde{\psi}}{dx} \frac{dx}{dz} \right) = \frac{d^2\tilde{\psi}}{dx^2} \left(\frac{dx}{dz} \right)^2. \end{aligned} \quad (3.6.17)$$

From (3.6.16) we get

$$\left(\frac{dx}{dz} \right)^2 = \left(\frac{2m^2g}{\hbar^2} \right)^{2/3} \implies \ddot{\psi}(z) = \left(\frac{2m^2g}{\hbar^2} \right)^{2/3} \frac{d^2\tilde{\psi}}{dx^2}. \quad (3.6.18)$$

By inserting (3.6.16) and (3.6.18) into (3.6.15) we have:

$$\left(\frac{2m^2g}{\hbar^2} \right)^{2/3} \frac{d^2\tilde{\psi}}{dx^2} - \left(\frac{2m^2g}{\hbar^2} \right)^{2/3} x \tilde{\psi}(x) = 0 \implies \frac{d^2\tilde{\psi}(x)}{dx^2} - x \tilde{\psi}(x) = 0. \quad (3.6.19)$$

The solution of this equation which tends to zero for $x \rightarrow \infty$ is the Airy function plotted in Fig. 3.7; we will indicate it as $Ai(x)$.

Consequently:

$$\tilde{\psi}(x) = Ai(x) = Ai \left[\left(\frac{2m^2g}{\hbar^2} \right)^{1/3} \left(z - \frac{E}{mg} \right) \right].$$

How can we find the energy levels? For $z < 0$ there is an ∞ potential, so the wave function must be zero also at $z = 0$:

$$\psi(z = 0) = 0 \implies Ai \left[\left(\frac{2m^2g}{\hbar^2} \right)^{1/3} \left(-\frac{E}{mg} \right) \right] = 0.$$

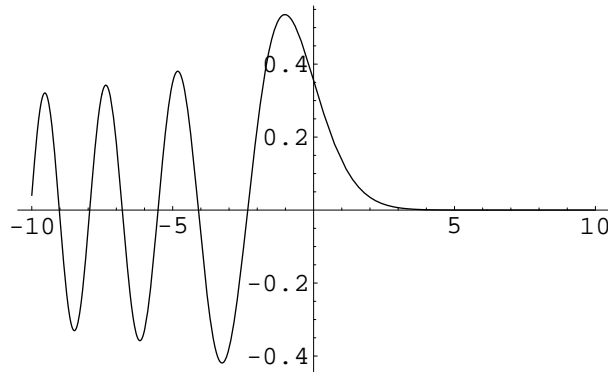


Fig. 3.7: Plot of the Airy function $Ai(x)$.

So to determine the energy levels E we must find the zeros of the Airy function Ai . These zeros are known and classified in the literature. Let us indicate them with $-\alpha_n$ where $n = 1, 2, \dots$:

$$\left(\frac{2m^2g}{\hbar^2}\right)^{1/3} \left(-\frac{E}{mg}\right) = -\alpha_n \implies E_n = \left(\frac{mg^2\hbar^2}{2}\right)^{1/3} \alpha_n.$$

The first energy level is associated to the first zero of the Airy function $\alpha_1 \cong 2.34$. Therefore:

$$E_1 = 2.34 \left(\frac{mg^2\hbar^2}{2}\right)^{1/3} \sim 10^{-12} \text{ eV}.$$

Such a number has been measured very recently in the Netherlands. Just to give an idea of the smallness of such a number remember that the energy of an electron in the hydrogen atom is about 13 eV. So we can conclude that also in the case of a particle in the gravitational field there is a finite zero-point energy. The particle cannot be at rest on the ground like a classical particle. This is due to the uncertainty principle.

Problem 11. *Show that in a one-dimensional problem the bound states are always non-degenerate.*

Solution. Let us prove it *ab absurdo* and suppose the opposite is true. Let $\psi_1(x)$ and $\psi_2(x)$ be linearly independent solutions with the same energy E :

$$\psi_1'' + \frac{2m}{\hbar^2}(E - V)\psi_1 = 0, \quad \psi_2'' + \frac{2m}{\hbar^2}(E - V)\psi_2 = 0. \quad (3.6.20)$$

Let us divide the first equation by ψ_1 and the second one by ψ_2 ; what we obtain is:

$$\frac{\psi_1''}{\psi_1} = \frac{2m}{\hbar^2}(V - E), \quad \frac{\psi_2''}{\psi_2} = \frac{2m}{\hbar^2}(V - E). \quad (3.6.21)$$

This means that:

$$\frac{\psi_1''}{\psi_1} = \frac{\psi_2''}{\psi_2} = \frac{2m}{\hbar^2}(V - E) \quad (3.6.22)$$

which implies:

$$\psi_1''\psi_2 - \psi_2''\psi_1 = (\psi_1'\psi_2)' - (\psi_2'\psi_1)' = 0.$$

By integrating the previous equation we get:

$$\psi_1' \psi_2 - \psi_2' \psi_1 = \text{const.} \quad (3.6.23)$$

Since the equation above is true for every x and at ∞ the wave function ψ of a bound state must be zero, the constant appearing in (3.6.23) must also be zero, which implies:

$$\psi_1' \psi_2 - \psi_2' \psi_1 = 0 \implies \frac{\psi_1'}{\psi_1} = \frac{\psi_2'}{\psi_2}. \quad (3.6.24)$$

By integration we get:

$$\log \psi_1 = \log \psi_2 + \log c \implies \psi_1 = C \psi_2. \quad (3.6.25)$$

Therefore the two states are linearly dependent and there is no degeneracy (there can be degeneracy only if the states are linearly independent). This is true only in one dimension. In three dimensions the theorem above does not hold and there can be degeneracy.

Homework: exercise 5. *Prove that, since the eigenfunctions ψ corresponding to bound states are non-degenerate, they can be chosen as real.*

Problem 12. *Show that the first derivative of the stationary wave functions are continuous even when the potential $V(x)$ has a finite discontinuity.*

Solution. Let us suppose the discontinuity is at the point x_0 and replace $V(x)$ in the interval $(x_0 - \epsilon, x_0 + \epsilon)$ with a potential $V_1(x)$ linearly interpolating between $V(x_0 - \epsilon)$ and $V(x_0 + \epsilon)$. In such a way we obtain a continuous potential whose eigenfunctions are given by the solutions of:

$$\psi_1'' + \frac{2m}{\hbar^2} [E - V_1(x)] \psi_1(x) = 0. \quad (3.6.26)$$

By integrating it between $x_0 - \epsilon$ and $x_0 + \epsilon$ we get:

$$\int_{x_0 - \epsilon}^{x_0 + \epsilon} dx \psi_1''(x) = \frac{2m}{\hbar^2} \int_{x_0 - \epsilon}^{x_0 + \epsilon} dx [V_1(x) - E] \psi_1(x)$$

which is equivalent to

$$(\psi_1') \Big|_{x_0 + \epsilon} - (\psi_1') \Big|_{x_0 - \epsilon} = \frac{2m}{\hbar^2} \int_{x_0 - \epsilon}^{x_0 + \epsilon} dx [V_1(x) - E] \psi_1(x). \quad (3.6.27)$$

In the limit $\epsilon \rightarrow 0$ we have that $V_1(x) \rightarrow V(x)$ and $\psi_1(x) \rightarrow \psi(x)$. Since the integrand function on the RHS of (3.6.27) is finite, it remains finite also in the limit $\epsilon \rightarrow 0$. According to the mean value theorem:

$$\int_a^b dx f(x) = \bar{f} [b - a] \quad (3.6.28)$$

where \bar{f} is the mean value. If \bar{f} remains finite and the interval $b - a$ tends to zero also the LHS of (3.6.28) must tend to zero. So in our case the mean value theorem implies

that the RHS of (3.6.27) tends to zero for $\epsilon \rightarrow 0$. Consequently also the LHS will go to zero. This means that:

$$\psi' \Big|_{x_0+0} = \psi' \Big|_{x_0-0}, \quad (3.6.29)$$

i.e. the derivatives of the eigenfunctions are continuous. If the potential has an ∞ discontinuity, then we cannot use the mean value theorem to say that, for $\epsilon \rightarrow 0$, the integral in (3.6.27) is zero, so in this case the derivative ψ' can have a finite gap. We will see an example of this phenomenon when we will analyze the Dirac delta potentials.

The Theorem of the Zeros. *The “theorem of the zeros” states that if the discrete eigenvalues of a one-dimensional Schrödinger equation are ordered from the smallest to the largest one: $E_1 < E_2 < \dots < E_n$, then the associated eigenfunctions have an increasing number of zeros. To prove this let us show that, between two consecutive zeros of the n -th eigenfunction, there is at least one zero of the $n + 1$ -th eigenfunction.*

Proof. Consider the eigenfunctions $\psi_n(x)$ and $\psi_{n+1}(x)$ with eigenvalues $E_n < E_{n+1}$. From the equations:

$$\psi_n''(x) + \frac{2m}{\hbar^2}(E_n - V)\psi_n = 0, \quad \psi_{n+1}''(x) + \frac{2m}{\hbar^2}(E_{n+1} - V)\psi_{n+1} = 0$$

we get that:

$$\frac{\psi_n''}{\psi_n} = \frac{2m}{\hbar^2}(V - E_n), \quad \frac{\psi_{n+1}''}{\psi_{n+1}} = \frac{2m}{\hbar^2}(V - E_{n+1}).$$

By taking the difference of the two equations above we get:

$$\frac{\psi_n''}{\psi_n} - \frac{\psi_{n+1}''}{\psi_{n+1}} = \frac{2m}{\hbar^2}(E_{n+1} - E_n) \quad (3.6.30)$$

which implies:

$$\begin{aligned} \psi_n''\psi_{n+1} - \psi_{n+1}''\psi_n &= \frac{2m}{\hbar^2}(E_{n+1} - E_n)\psi_n\psi_{n+1} \\ &\Downarrow \\ \frac{d}{dx}[\psi_n'\psi_{n+1} - \psi_{n+1}'\psi_n] &= \frac{2m}{\hbar^2}(E_{n+1} - E_n)\psi_n\psi_{n+1}. \end{aligned} \quad (3.6.31)$$

Let us integrate Eq. (3.6.31) between two consecutive zeros α and β of the wave function ψ_n . Since the function ψ_n is continuous, it does not change sign between α and β . Then we can suppose $\psi_n > 0$ which implies $\psi_n'(\alpha) > 0$ and $\psi_n'(\beta) < 0$. After the integration between α and β the LHS of (3.6.31) becomes:

$$\begin{aligned} \psi_n'(\beta)\psi_{n+1}(\beta) - \psi_{n+1}'(\beta)\psi_n(\beta) - \psi_n'(\alpha)\psi_{n+1}(\alpha) + \psi_{n+1}'(\alpha)\psi_n(\alpha) &= \\ = \psi_n'(\beta)\psi_{n+1}(\beta) - \psi_n'(\alpha)\psi_{n+1}(\alpha) &= \\ = -\psi_{n+1}(\beta)|\psi_n'(\beta)| - \psi_{n+1}(\alpha)|\psi_n'(\alpha)| \end{aligned} \quad (3.6.32)$$

where in the first step we used the fact that α and β are zeros for the function ψ_n and in the second one that $\psi_n'(\alpha) > 0$ and $\psi_n'(\beta) < 0$. The RHS of (3.6.31) becomes instead:

$$\frac{2m}{\hbar^2}(E_{n+1} - E_n) \int_a^b dx \psi_n(x)\psi_{n+1}(x). \quad (3.6.33)$$

If, ab absurdo, ψ_{n+1} had not changed sign between α and β , then Eq. (3.6.32) would have an opposite sign with respect to the one of the function ψ_{n+1} ; Eq.(3.6.33) would have instead the same sign of ψ_{n+1} and this would give a contradiction. So ψ_{n+1} must change sign and, being continuous, it must have at least one zero between α and β . If ψ_n and ψ_{n+1} are eigenfunctions of the discrete spectrum they both vanish exponentially at the two boundaries of the interval $(-\infty, \infty)$. The $n - 1$ zeros of ψ_n divide the interval into n partial intervals, in each of which ψ_{n+1} has at least one zero; so, the function ψ_{n+1} has at least n nodes. Of course this does not prove that there is *only one* zero in each partial interval, but it is actually so and the function ψ_{n+1} has exactly n nodes.

Problem 13. Find the energy values for the bound states of a particle in the symmetric well potential plotted in Fig. 3.8.

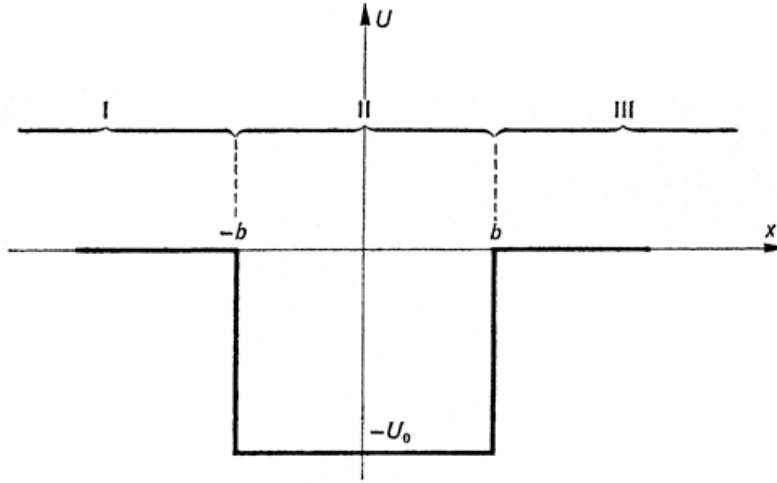


Fig. 3.8: Symmetric well potential.

Solution. As we have seen in **Sec. 2.4** the bound states can exist only in the range $-U_0 < E < 0$. Let us write the eigenvalue equations in the three regions *I, II, III* indicated above:

$$\begin{aligned} \psi_1''(x) - \frac{2m}{\hbar^2}|E|\psi_1(x) &= 0 & \text{for } x < -b, \\ \psi_2''(x) + \frac{2m}{\hbar^2}(U_0 - |E|)\psi_2(x) &= 0 & \text{for } |x| < b, \\ \psi_3''(x) - \frac{2m}{\hbar^2}|E|\psi_3(x) &= 0 & \text{for } x > b. \end{aligned} \quad (3.6.34)$$

Let us use the following compact notation:

$$\alpha \equiv \left(\frac{2m}{\hbar^2}|E| \right)^{1/2} > 0, \quad \beta \equiv \left(\frac{2m}{\hbar^2}(U_0 - |E|) \right)^{1/2} > 0. \quad (3.6.35)$$

In the three regions the wave function will have the form:

$$\psi(x) = \begin{cases} \psi_{III}(x) = Ae^{\alpha x} + Be^{-\alpha x} & x > b \\ \psi_{II}(x) = C \sin \beta x + D \cos \beta x & |x| < b \\ \psi_I(x) = Fe^{\alpha x} + Ge^{-\alpha x} & x < -b. \end{cases} \quad (3.6.36)$$

The wave function of a bound state must tend to zero at $\pm\infty$ and so $A = G = 0$. Note also that $|\psi_{III}|^2 = B^2 e^{-2\alpha x} \neq 0$ and $|\psi_I|^2 = F^2 e^{2\alpha x} \neq 0$, i.e. the particle can be found also outside the well potential differently than in the classical case where it should move within the well in order to have a positive kinetic energy.

Now we must impose that the wave function and its derivative be continuous also at $x = \pm b$, that is:

$$\begin{aligned} C \sin \beta b + D \cos \beta b &= B e^{-\alpha b} \\ -C \sin \beta b + D \cos \beta b &= F e^{-\alpha b} \\ \beta C \cos \beta b - \beta D \sin \beta b &= -\alpha B e^{-\alpha b} \\ \beta C \cos \beta b + \beta D \sin \beta b &= F \alpha e^{-\alpha b}. \end{aligned}$$

By making suitable combinations of the equations above we get:

$$2C \sin \beta b = (B - F) e^{-\alpha b} \quad (3.6.37)$$

$$2\beta C \cos \beta b = -\alpha(B - F) e^{-\alpha b} \quad (3.6.38)$$

$$2D \cos \beta b = (B + F) e^{-\alpha b} \quad (3.6.39)$$

$$2\beta D \sin \beta b = \alpha(B + F) e^{-\alpha b}. \quad (3.6.40)$$

If $C \neq 0$ then $B - F \neq 0$ and dividing (3.6.38) by (3.6.37) we get:

$$\beta \cotg \beta b = -\alpha. \quad (3.6.41)$$

If $D \neq 0$ then $B + F \neq 0$ and dividing (3.6.40) by (3.6.39) we get:

$$\beta \tg \beta b = \alpha. \quad (3.6.42)$$

Homework: exercise 6. Prove that Eqs. (3.6.41) and (3.6.42) cannot be satisfied at the same time.

So we must distinguish two different set of solutions of Eqs. (3.6.37)-(3.6.40):

$$C = 0, \quad B = F, \quad \beta \tg \beta b = \alpha \quad (3.6.43)$$

$$D = 0, \quad B = -F, \quad \beta \cotg \beta b = -\alpha. \quad (3.6.44)$$

In the first case given by Eq. (3.6.43) the solution within the well is $\psi(x) = D \cos \beta x$, which is an even function. In the second case given by (3.6.44) the solution is $\psi(x) = C \sin \beta x$, which is an odd function. The energies are determined by solving the equations $\beta \tg \beta b = \alpha$, $\beta \cotg \beta b = -\alpha$ (remember that, according to (3.6.35), α and β depend on the energy E). The equations can be solved only graphically. Putting $\xi = \beta b$ and $\eta = \alpha b$ we get:

$$E = -\frac{\hbar^2}{2mb^2} \eta^2. \quad (3.6.45)$$

Let us also notice that from the definition of ξ and η :

$$\eta^2 + \xi^2 = \frac{2mb^2}{\hbar^2} U_0 \quad (3.6.46)$$

while the equation $\beta \operatorname{tg} \beta b = \alpha$ becomes:

$$\xi \operatorname{tg} \xi = \eta. \quad (3.6.47)$$

Therefore in the first case we must solve the equations:

$$\xi \operatorname{tg} \xi = \eta, \quad \xi^2 + \eta^2 = \frac{2mb^2}{\hbar^2} U_0. \quad (3.6.48)$$

In the second one we have instead:

$$-\xi \operatorname{cotg} \xi = \eta, \quad \xi^2 + \eta^2 = \frac{2mb^2}{\hbar^2} U_0. \quad (3.6.49)$$

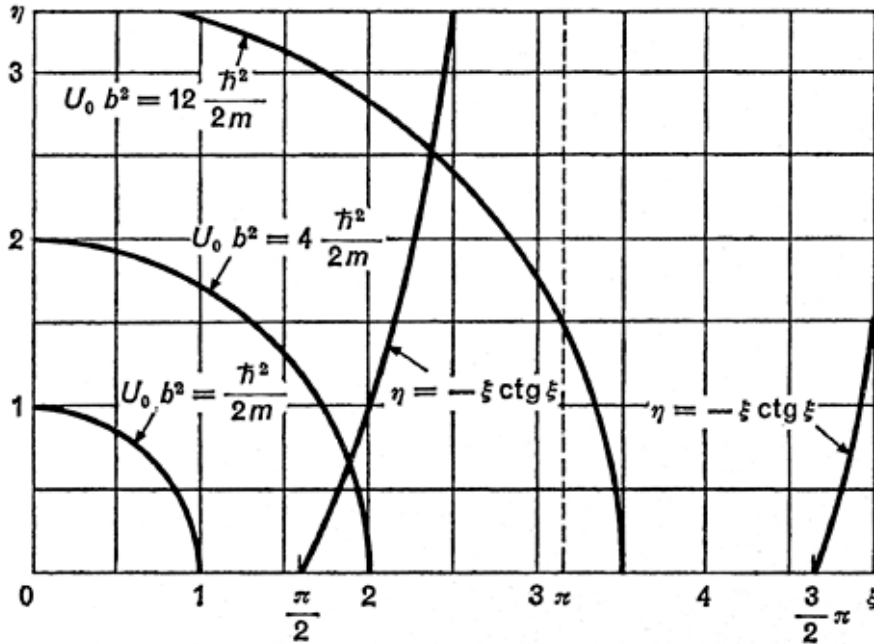


Fig. 3.9: Solution of the symmetric square well.

The solutions of (3.6.48) and (3.6.49) are given by the intersections of the circle $\xi^2 + \eta^2 = \frac{2mb^2}{\hbar^2} U_0$ with the curves $\eta = \xi \operatorname{tg} \xi$ and $\eta = -\xi \operatorname{cotg} \xi$ in the region $\xi, \eta > 0$. For example in Fig. 3.9 we plotted the second case in which the intersection is between the circle and $\eta = -\xi \operatorname{cotg} \xi$. It is clear that the number of the bound states increases with the radius of the circle, i.e. with $b^2 U_0$. This means that the number of bound states increases with the depth U_0 and the width b of the well. If the radius $R = \sqrt{\frac{2m}{\hbar^2} b^2 U_0}$ of the circle satisfies $N \frac{\pi}{2} < R < \frac{N+1}{2} \pi$ with $N = 0, 1, 2, \dots$, then the number of intersections between the circle and the curves, which gives the number of bound states, is $N + 1$. In the particular case $R < \pi/2$ there is only one bound state.

Problem 14. Which is the parity of the energy eigenfunctions of the discrete spectrum for a time-independent Schrödinger equation with an even potential $V(x) = V(-x)$?

Solution. If $\psi(x)$ is a solution of the time-independent Schrödinger equation with energy E :

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x) = E\psi(x) \quad (3.6.50)$$

then it is easy to prove that also $\psi(-x)$ is a solution of the Schrödinger equation with the same value of energy E . In fact let us replace in (3.6.50) x with $-x$. What we obtain is:

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(-x) \right] \psi(-x) = E\psi(-x). \quad (3.6.51)$$

Since the potential is even, we have that $V(-x) = V(x)$, which means that we can rewrite (3.6.51) as:

$$\left[-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(-x) = E\psi(-x).$$

This proves that both $\psi(x)$ and $\psi(-x)$ are solutions of the Schrödinger equation associated with the same value of energy.

Now let us suppose the functions $\psi(x)$ and $\psi(-x)$ belong to the discrete spectrum. As we have seen in Problem 11, in this case there cannot be degeneracy. So the two wave functions must be linearly dependent: $\psi(x) = C\psi(-x)$. Changing $x \rightarrow -x$, we get $\psi(-x) = C\psi(x)$ which implies $\psi(x) = C\psi(-x) = C^2\psi(x) \Rightarrow C = \pm 1$. Therefore:

$$\psi(x) = \pm\psi(-x)$$

which means that the eigenfunctions are even or odd. In the discrete spectrum of a one-dimensional system the wave function ψ_{n+1} has one zero more than ψ_n . Since the even (odd) functions have an even (odd) number of zeros we can conclude that the eigenfunctions of the discrete spectrum of an even potential are alternatively even or odd.

In the continuous spectrum there can be degeneracy and so it is not necessary that $\psi(x) = C\psi(-x)$. This implies that the eigenfunctions can have a non-definite parity.

Problem 15. Study the time-independent Schrödinger equation for the step potential

$$V = \begin{cases} 0 & \text{for } x < 0 \\ V_0 > 0 & \text{for } x > 0 \end{cases}.$$

Solution. The Schrödinger equation in the region $x < 0$ is:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) = E\psi(x) \implies \frac{d^2}{dx^2} \psi + \frac{8\pi^2 m}{h^2} E \psi = 0.$$

In the region $x \geq 0$ the equation is instead:

$$-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + V_0 \psi(x) = E\psi(x) \implies \frac{d^2}{dx^2} \psi + \frac{8\pi^2 m}{h^2} (E - V_0) \psi = 0.$$

If we define $k_0^2 \equiv \frac{8\pi^2 m}{h^2} E$ and $k^2 = \frac{8\pi^2 m}{h^2} (E - V_0)$ then the two Schrödinger equations can be written as:

$$\begin{cases} \frac{d^2 \psi}{dx^2} + k_0^2 \psi = 0 & x < 0 \\ \frac{d^2 \psi}{dx^2} + k^2 \psi = 0 & x > 0. \end{cases}$$

The solutions can be written as:

$$\begin{cases} \psi_1(x) = A \exp(ik_0x) + B \exp(-ik_0x) & x < 0 \\ \psi_2(x) = C \exp(ikx) + D \exp(-ikx) & x > 0. \end{cases} \quad (3.6.52)$$

The boundary conditions are:

$$\begin{cases} \psi_1(0) = \psi_2(0) \\ \psi_1'(0) = \psi_2'(0) \end{cases}$$

and they give:

$$\begin{cases} A + B = C + D \\ k_0(A - B) = k(C - D). \end{cases} \quad (3.6.53)$$

Let us consider the two cases:

- 1) $E > V_0$
- 2) $0 < E < V_0$.

In *classical* mechanics in case 1) the particle slows down over the barrier because its kinetic energy becomes smaller in order to preserve the total energy. In case 2) instead the particle cannot go beyond the barrier because otherwise its kinetic energy would become negative.

Let us now analyze the situation in *quantum* mechanics and consider a wave which at the beginning is at $x = -\infty$ and propagates towards greater values of x . In the case 1) we have that both k_0 and k are real numbers, so the solution (3.6.52) is given by plane waves everywhere. Since in (3.6.53) we have two equations and four unknown quantities, we can put $D = 0$. With this choice Eq. (3.6.53) becomes:

$$\begin{cases} A + B = C \\ k_0(A - B) = kC \end{cases} \implies \frac{A + B}{A - B} = \frac{k_0}{k}. \quad (3.6.54)$$

Let us notice that we cannot choose also $B = 0$ because otherwise we would obtain $\frac{k_0}{k} = 1$ which is false since $k_0 > k$. Therefore we can say that, even if we consider a wave propagating towards greater values of x , for $x < 0$ we cannot throw away the wave propagating towards smaller values of x , since B must be different from zero. This means that, even if the energy of the particle is greater than the step potential, nevertheless the particle feels the step and is, at least in part, reflected by it.

In the case 2), $0 < E < V_0$, k is an imaginary number. In order to avoid a divergence of $\psi_2(x)$ we must put either $C = 0$ or $D = 0$. If we put $D = 0$ and we solve Eq. (3.6.53) we get:

$$A = \frac{k_0 + i|k|}{k_0} \frac{C}{2}, \quad B = \frac{k_0 - i|k|}{k_0} \frac{C}{2}. \quad (3.6.55)$$

Therefore there is a reflected wave but there is also a wave which goes beyond the step, i.e. we cannot put $C = 0$ because otherwise from (3.6.55) we would have $A = 0$ but this is not true because at the beginning we have the initial wave. So $C \neq 0$ and for $x > 0$ we get $\psi_2(x) = C \exp(-|k|x)$. This means that the wave can penetrate the step. Also the reflected wave cannot be thrown away: in fact we cannot choose $B = 0$ because otherwise we would have again that $k_0 = i|k|$ which is false.

Homework: exercise 7. Solve the Schrödinger equation for a potential like the one plotted in Fig. 3.10 and prove that the particles in this case can go from region I to region III even if they have less energy than the step (tunnel effect).

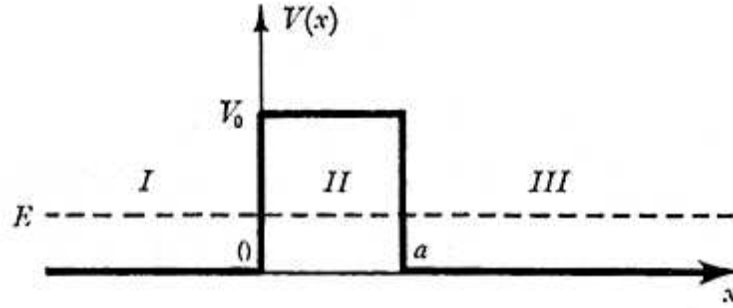


Fig. 3.10: Potential for the Tunnel Effect.

Transmission and Reflection Coefficients. Let us go back to the step potential. Remember that the probability current J in terms of the wave function ψ is given by: $J = \frac{\hbar}{4\pi im}(\psi^* \partial_x \psi - \psi \partial_x \psi^*)$. For $x < 0$ and in the case $E > V_0$ the wave function which propagates towards greater values of x gives the following probability current:

$$J^+(x < 0) = \frac{\hbar}{2\pi m} AA^* k_0. \quad (3.6.56)$$

The current associated with the wave function propagating towards smaller values of x is:

$$J^-(x < 0) = \frac{\hbar}{2\pi m} BB^* k_0. \quad (3.6.57)$$

Combining (3.6.54), (3.6.56) and (3.6.57) it is easy to prove that the reflection coefficient defined as $R \equiv \frac{J^-(x < 0)}{J^+(x < 0)}$ turns out to be:

$$R = \frac{BB^*}{AA^*} = \left(\frac{k_0 - k}{k_0 + k} \right)^2. \quad (3.6.58)$$

For $x > 0$ the current associated with the wave function $\psi_2 = Ce^{ikx}$ is:

$$J^+(x > 0) = \frac{\hbar}{2\pi m} CC^* k. \quad (3.6.59)$$

The transmission coefficient defined as: $T \equiv \frac{J^+(x > 0)}{J^+(x < 0)}$ becomes:

$$T = \frac{CC^* k}{AA^* k_0} = \frac{4kk_0}{(k_0 + k)^2}. \quad (3.6.60)$$

From (3.6.58) and (3.6.60) it is easy to prove that $R + T = 1$. If instead the energy E is smaller than the step potential V_0 , we get from (3.6.55) that:

$$R = \frac{BB^*}{AA^*} = 1, \quad (3.6.61)$$

i.e. it seems that the entire wave is reflected. This is not true. In fact, since $C \neq 0$, if we perform a measure, we have a certain probability of finding the particle also in the region $x > 0$.

Problem 16. Find the spectrum and the wave functions of the time independent Schrödinger equation for a harmonic oscillator $V = \frac{1}{2}kx^2$.

Solution. The Schrödinger equation is

$$-\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + \frac{1}{2}kx^2\psi(x) = E\psi(x). \quad (3.6.62)$$

With the following change of variables:

$$\xi \equiv \alpha x, \quad \alpha = \left(\frac{mk}{\hbar^2}\right)^{1/4}$$

Eq. (3.6.62) becomes:

$$\frac{d^2\psi}{d\xi^2} + (\lambda - \xi^2)\psi = 0 \quad (3.6.63)$$

where

$$\lambda \equiv \frac{2E}{\hbar} \left(\frac{m}{k}\right)^{1/2} = \frac{2E}{\hbar\omega_c}, \quad \text{with } \omega_c = \left(\frac{k}{m}\right)^{1/2}. \quad (3.6.64)$$

The function $\psi(\xi) = \xi^n \exp(-\frac{1}{2}\xi^2)$ is a normalizable solution of Eq. (3.6.63) at least for $\xi \rightarrow +\infty$. Let us calculate the derivatives:

$$\begin{cases} \frac{d\psi}{d\xi} = n\xi^{n-1}e^{-\xi^2/2} - \xi^{n+1}e^{-\xi^2/2} \\ \frac{d^2\psi}{d\xi^2} = \left(n(n-1)\xi^{n-2} - n\xi^n - (n+1)\xi^n + \xi^{n+2}\right)e^{-\xi^2/2}. \end{cases} \quad (3.6.65)$$

For $\xi \rightarrow \infty$ we get that $\frac{d^2\psi}{d\xi^2} \sim \xi^{n+2}e^{-\xi^2/2}$. In the limit $\xi \rightarrow +\infty$ we can disregard λ with respect to ξ^2 and Eq. (3.6.63) is satisfied:

$$\frac{d^2\psi}{d\xi^2} - \xi^2\psi \sim \xi^{n+2}e^{-\xi^2/2} - \xi^2\xi^n e^{-\xi^2/2} = 0. \quad (3.6.66)$$

Now we let us try to find a complete (and not only asymptotic) solution of (3.6.63) of the form:

$$\psi(\xi) = H(\xi)e^{-\xi^2/2} \quad (3.6.67)$$

with $H(\xi)$ polynomial in ξ . If we insert (3.6.67) into (3.6.63) we obtain the following equation for $H(\xi)$:

$$H''(\xi) - 2\xi H'(\xi) + (\lambda - 1)H(\xi) = 0. \quad (3.6.68)$$

Let us make this ansatz for $H(\xi)$:

$$H(\xi) = \xi^s [a_0 + a_1\xi + a_2\xi^2 + \dots] \quad (3.6.69)$$

with $a_0 \neq 0$ and $s \geq 0$. If we use (3.6.69) in (3.6.68) and we equate the coefficients of the same powers of ξ we get:

$$\begin{aligned}
 s(s-1)a_0 &= 0 \\
 (s+1)sa_1 &= 0 \\
 (s+2)(s+1)a_2 - (2s+1-\lambda)a_0 &= 0 \\
 (s+3)(s+2)a_3 - (2s+3-\lambda)a_1 &= 0 \\
 \dots\dots\dots \\
 (s+\nu+2)(s+\nu+1)a_{\nu+2} - (2s+2\nu+1-\lambda)a_\nu &= 0.
 \end{aligned} \tag{3.6.70}$$

Since we have chosen $a_0 \neq 0$, we get from the first equation that $s = 0$ or $s = 1$. The second equation gives $s = 0$ or $a_1 = 0$ or both of them. The third equation gives a_2 as a function of a_0 ; the fourth equation gives a_3 as a function of a_1 ; in general the ν -th equation gives $a_{\nu+2}$ as a function of a_ν . According to what we have seen in Problems 11 and 14, the functions $\psi(x)$ cannot be degenerate in the discrete part of the spectrum and, since the potential is even, they will be either even or odd: $\psi(x) = \pm\psi(-x)$. This implies that $H(\xi)$ must be either even or odd. From (3.6.69) we get the even or odd $H(\xi)$ by choosing $a_{2\nu+1} = 0$ and $s = 0$ or $a_{2\nu+1} = 0$ and $s = 1$ respectively.

In (3.6.70) there will be a finite or an infinite number of terms depending on how we choose s and λ . If the series does not stop, then the ratio of the coefficients $\frac{a_{\nu+2}}{a_\nu}$, for large ν , goes as:

$$\frac{a_{\nu+2}}{a_\nu} \xrightarrow{\nu \rightarrow \infty} \frac{2s+2\nu+1-\lambda}{(s+\nu+2)(s+\nu+1)} \sim \frac{2}{\nu}. \tag{3.6.71}$$

For large values of ν the a_ν behave just like the coefficients of the series expansion of

$$\xi^n e^{\xi^2} \tag{3.6.72}$$

with n integer¹. This means that the power series (3.6.69) becomes an exponential which, combined with the factor $e^{-\xi^2/2}$ of (3.6.67), gives $\xi^n e^{\xi^2/2}$. Such a function is not normalizable and so it cannot be a solution for the Schrödinger equation. In order to have an acceptable solution of the Schrödinger equation the series (3.6.69) must stop. This means that in the last term of (3.6.70) λ must satisfy the equation:

$$\lambda = 2s + 2\nu + 1. \tag{3.6.73}$$

If the previous relation holds then $a_{\nu+2} = 0$ and the series stops. Now if we call $n = s + \nu$ we obtain that $\lambda = 2n + 1$. By replacing λ with E , see Eq. (3.6.64), we get

$$E_n = \left(n + \frac{1}{2} \right) \hbar\omega_c \tag{3.6.74}$$

which implies that the spectrum of a harmonic oscillator is discrete. The zero-point energy is given by the $n = 0$ term:

$$E_0 = \frac{1}{2} \hbar\omega_c. \tag{3.6.75}$$

¹In fact two consecutive terms in the series expansion of (3.6.72) are given by $\xi^n (\xi^2)^m \frac{1}{m!}$ and $\xi^n (\xi^2)^{m-1} \frac{1}{(m-1)!}$ and their ratio goes just as $1/m$.

Also in this case the zero-point energy is due to the uncertainty principle. In fact the eigenfunction associated with E_0 is $\psi_0(\xi) = Ne^{-\xi^2/2}$, which means that the particle is characterized by a finite uncertainty Δx in the position. From Heisenberg's relations the associated uncertainty in the momentum must be finite $\Delta p \neq 0$.

From Eq. (3.6.68) we get that the equation to solve in order to get the eigenfunctions of a harmonic oscillator is:

$$H_n'' - 2\xi H_n' + 2nH_n = 0 \quad (3.6.76)$$

whose solutions are the Hermite polynomials:

$$H_0(\xi) = 1, \quad H_1(\xi) = 2\xi, \quad H_2(\xi) = 4\xi^2 - 2$$

which satisfy the following recurrence relations:

$$H_n' = 2nH_{n-1}, \quad H_{n+1} = 2\xi H_n - 2nH_{n-1}.$$

We want to stress the fact that also in the case of the harmonic oscillator the discreteness of the spectrum comes from the renormalizability requirement of the wave function ψ_n .

3.7 Dirac Delta

A *function* $y = f(x)$ is a map, e.g. between \mathbb{R} and \mathbb{R} , between \mathbb{C} and \mathbb{C} or between \mathbb{R} and \mathbb{C} . A *functional* is a map between a space of functions f and the space of real or complex numbers and it is indicated as $y = \mathcal{F}[f]$. In other words a functional associates a number to every function f of its domain. In particular a *Dirac delta* $\delta(x - x_0)$ associates to every function $f(x)$ the number $f(x_0)$:

$$\int_{-\infty}^{\infty} dx \delta(x - x_0) f(x) = f(x_0). \quad (3.7.1)$$

So the Dirac delta $\delta(x - x_0)$ is like a very peaked function, different from zero only for $x \neq x_0$. From the definition itself of the Dirac delta (3.7.1) we get:

$$\int_{-\infty}^{\infty} dx \delta(x - x_0) = 1.$$

This condition suggests that it is possible to “represent” the Dirac delta distribution like a limit of a suitable sequence of functions. For example, if we choose $x_0 = 0$, we have that the action of the Dirac delta $\delta(x)$

$$\int_{-\infty}^{+\infty} dx \delta(x) f(x) = f(0) \quad (3.7.2)$$

is the same as the one achieved by the following limit:

$$\delta(x) \sim \lim_{l \rightarrow \infty} \frac{l}{\sqrt{\pi}} e^{-l^2 x^2} \quad (3.7.3)$$

where the limit must be done within the integral. In fact, by performing the following change of variables $\tilde{x} \equiv lx$ we get:

$$\frac{l}{\sqrt{\pi}} \int_{-\infty}^{\infty} dx e^{-l^2 x^2} f(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} d\tilde{x} e^{-\tilde{x}^2} f(\tilde{x}/l).$$

If we send $l \rightarrow \infty$ we obtain:

$$\lim_{l \rightarrow \infty} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} d\tilde{x} e^{-\tilde{x}^2} f(\tilde{x}/l) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} d\tilde{x} e^{-\tilde{x}^2} f(0) = f(0).$$

So in the limit $l \rightarrow \infty$ the sequence of functions $\frac{l}{\sqrt{\pi}} e^{-l^2 x^2}$ is a good approximation of the Dirac delta distribution. Such a sequence of functions is plotted in Fig. 3.11.

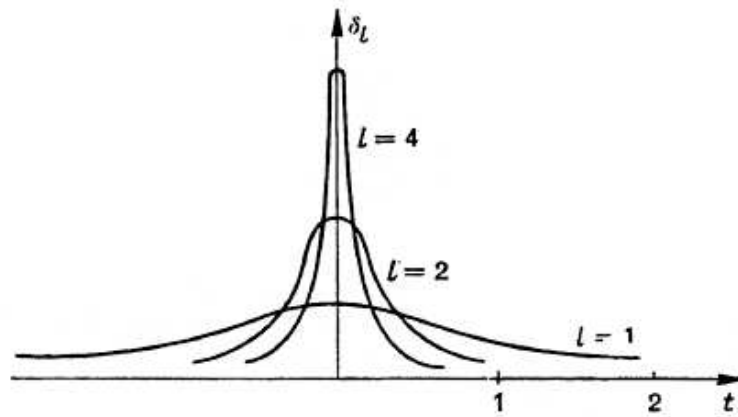


Fig. 3.11: Representation of the Dirac delta.

Other representations of the Dirac delta $\delta(x)$ are given by:

$$\frac{1}{l\pi} \frac{\sin^2 lx}{x^2}, \quad \frac{1}{l\pi} \frac{1}{x^2 + \frac{1}{l^2}}, \quad \frac{1}{\pi} \frac{\sin lx}{x}.$$

The Dirac delta $\delta(x - x_0)$ is the generalization to the continuous case of the Kronecker delta δ_j^i defined as:

$$\delta_j^i = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad i, j \in 1, 2, \dots, N. \quad (3.7.4)$$

From the definition itself of Kronecker delta every vector f^i can be rewritten as $f^i = \sum_j \delta_j^i f^j$. In fact if “ i ” is a continuous index then f^i becomes a function of x , $f^x = f(x)$, and the sum becomes an integral:

$$f(x) = \int dy \delta_y^x f^y = \int dy \delta(x - y) f(y) \quad (3.7.5)$$

which is just the action of the Dirac delta $\delta(x - y)$. In order to get some practice with the Dirac delta distributions we are going to perform some problems on them.

Problem 17. *The potential energy of a system is $U(x) = \tilde{U}(x) + \alpha\delta(x - x_0)$ where $\tilde{U}(x)$ is a bounded function. Which is the behavior of the solution of the Schrödinger equation $\psi(x)$ and of its derivative near the point x_0 ?*

Solution. The equation is

$$-\frac{\hbar^2}{2m}\psi'' + [\tilde{U}(x) + \alpha\delta(x - x_0)]\psi = E\psi. \quad (3.7.6)$$

The wave function must be continuous because of its probabilistic meaning. Let us study the behavior of the derivative by integrating Eq. (3.7.6) between $x_0 - \epsilon$ and $x_0 + \epsilon$ with $\epsilon > 0$ and let us send $\epsilon \rightarrow 0$:

$$\begin{aligned} -\frac{\hbar^2}{2m} \int_{x_0-\epsilon}^{x_0+\epsilon} dx \psi'' &= - \int_{x_0-\epsilon}^{x_0+\epsilon} dx [\tilde{U}(x) + \alpha\delta(x - x_0) - E]\psi(x) \\ &\Downarrow \\ -\frac{\hbar^2}{2m} [\psi'(x_0 + \epsilon) - \psi'(x_0 - \epsilon)] &= -\alpha\psi(x_0) - \int_{x_0-\epsilon}^{x_0+\epsilon} [\tilde{U}(x) - E]\psi(x). \end{aligned} \quad (3.7.7)$$

If $\psi(x)$ is continuous and $\tilde{U}(x)$ is limited we can use the mean value theorem. In the limit $\epsilon \rightarrow 0$ the integral on the RHS of (3.7.7) is zero, so we get:

$$\lim_{\epsilon \rightarrow 0} -\frac{\hbar^2}{2m} [\psi'(x_0 + \epsilon) - \psi'(x_0 - \epsilon)] = -\alpha\psi(x_0) \quad (3.7.8)$$

which means

$$\psi'(x_0 + 0) - \psi'(x_0 - 0) = \frac{2m\alpha}{\hbar^2}\psi(x_0). \quad (3.7.9)$$

This equation tells us that the derivative is not continuous at the point x_0 . Remember what we have proved in Problem 12: for potentials with a finite gap the derivative of the wave function must be continuous. In this case instead, since the Dirac delta at the point x_0 implies an infinite gap of the potential in x_0 , the results of Problem 12 cannot hold and the derivative is discontinuous.

Problem 18. Find the energy levels and the wave functions of the discrete spectrum of a particle in a potential $U(x) = -\alpha\delta(x)$ with $\alpha > 0$.

Solution. The Schrödinger equation

$$-\frac{\hbar^2}{2m}\psi'' - \alpha\delta(x)\psi = E\psi$$

for $x \neq 0$ becomes the usual equation of a free particle:

$$-\frac{\hbar^2}{2m}\psi'' = E\psi. \quad (3.7.10)$$

In order to have normalizable solutions of (3.7.10) the wave functions must be:

$$\begin{aligned} \psi(x) &= A \exp(-kx) & x > 0 \\ \psi(x) &= B \exp(kx) & x < 0. \end{aligned} \quad (3.7.11)$$

For the continuity of the wave function the two coefficients A and B must be equal. The relation between E and k can be obtained by inserting (3.7.11) into (3.7.10):

$$-\frac{\hbar^2}{2m} A k^2 \exp(-kx) = E A \exp(-kx) \implies E = -\frac{\hbar^2 k^2}{2m} \quad (3.7.12)$$

where k must be determined. Using the results of Problem 17, the gap in the derivative of the wave function is:

$$\psi'(x_0 + 0) - \psi'(x_0 - 0) = \frac{-2m\alpha}{\hbar^2} \psi(x_0) \quad (3.7.13)$$

which in our case becomes:

$$-kA \exp(-k \cdot 0) - kA \exp(k \cdot 0) = \frac{-2m\alpha}{\hbar^2} A \Rightarrow k = \frac{m\alpha}{\hbar^2}. \quad (3.7.14)$$

This implies that there is *only one* value of k which corresponds to a solution of our problem. The associated energy level is given by:

$$E = -\frac{\hbar^2}{2m} k^2 = -\frac{\hbar^2}{2m} \frac{m^2 \alpha^2}{\hbar^4} = -\frac{m\alpha^2}{2\hbar^2}. \quad (3.7.15)$$

Homework: exercise 8. Prove that the normalized eigenfunction associated with the energy level (3.7.15) is: $\psi = \sqrt{k} \exp[-k|x|]$.

Problem 19. Prove that the mean value of a conservative force calculated in an energy eigenstate of the discrete spectrum is zero.

Solution. Since the force is conservative it can be written in terms of a potential U as $F = -\frac{dU}{dx}$. The mean value of F is:

$$\begin{aligned} \bar{F} &= \int dx F(x) P(x) = \int dx F(x) \psi_n^* \psi_n = - \int dx \psi_n^* \frac{dU}{dx} \psi_n = \\ &= - \int_{-\infty}^{\infty} dx \frac{d}{dx} (U |\psi_n|^2) + \int_{-\infty}^{\infty} dx (\psi_n^* U \psi_n' + \psi_n^{*'} U \psi_n). \end{aligned} \quad (3.7.16)$$

The first integral in (3.7.16) gives $-U |\psi_n|^2 \Big|_{-\infty}^{\infty} = 0$ since the wave function ψ_n tends to zero for $x \rightarrow \pm\infty$. In order to evaluate the second integral of (3.7.16) we must remember that ψ_n is a solution of the time-independent Schrödinger equation:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \psi_n}{\partial x^2} + U \psi_n = E_n \psi_n, \quad -\frac{\hbar^2}{2m} \frac{\partial^2 \psi_n^*}{\partial x^2} + U \psi_n^* = E_n \psi_n^*. \quad (3.7.17)$$

So the second integral of (3.7.16) becomes:

$$\begin{aligned} &\int_{-\infty}^{\infty} dx \left[\left(E_n + \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right) \psi_n^* \psi_n' + \psi_n^{*'} \left(E_n + \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \right) \psi_n \right] = \\ &= \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \left[\frac{\partial^2 \psi_n^*}{\partial x^2} \psi_n' + \psi_n^{*'} \frac{\partial^2 \psi_n}{\partial x^2} \right] + E_n \int_{-\infty}^{\infty} dx \frac{d}{dx} (\psi_n^* \psi_n) = \\ &= \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} dx \frac{d}{dx} (\psi_n^{*'} \psi_n') = \frac{\hbar^2}{2m} \psi_n^{*'} \psi_n' \Big|_{-\infty}^{\infty} = 0. \end{aligned} \quad (3.7.18)$$

In the proof above it is crucial that the wave function and its derivative tend to zero at $\pm\infty$.

Homework: exercise 9. Prove that it is not true that the mean value of the force is zero if the wave function is not an energy eigenstate.

Problem 20. Find the energy values for which a particle does not reflect on the following potential barrier: $U(x) = \alpha[\delta(x) + \delta(x - a)]$.

Solution. Let us suppose the particle moves from left to right. If we want to avoid reflection, the wave function must have the form:

$$\psi_k(x) = \begin{cases} \exp ikx & x < 0 \\ A \sin kx + B \cos kx & 0 < x < a \\ C \exp[ik(x - a)] & x > a. \end{cases} \quad (3.7.19)$$

In the sector $x < 0$ we do not have the reflected part $R \exp -ikx$. In the sector $x > a$, for symmetry, we must have only the transmitted wave because if we had a wave propagating towards smaller values of x , we would have such a kind of wave also for $x < 0$. Now we can impose the boundary conditions on the wave function and on its derivative both in $x = 0$ and in $x = a$. We can use the results already obtained in the case of Dirac delta potentials; the wave function will be continuous while the derivative will have a finite gap:

$$\begin{cases} \psi'(x_0 + 0) - \psi'(x_0 - 0) = \frac{2m\alpha}{\hbar^2} \psi_{x_0} \\ \psi(x_0 + 0) = \psi(x_0 - 0). \end{cases} \quad (3.7.20)$$

Imposing (3.7.20) both on x_0 and x_a we get

$$B = 1 \quad (3.7.21)$$

$$kA - ik = 2m\alpha/\hbar^2 \quad (3.7.22)$$

$$A \sin ka + B \cos ka = C \quad (3.7.23)$$

$$ikC - kA \cos ka + kB \sin ka = 2m\alpha C/\hbar^2. \quad (3.7.24)$$

We have four equations and A, B, C, k as unknown variables. From (3.7.22) we get

$$kA = ik + 2m\alpha/\hbar^2 \Rightarrow A = i + 2m\alpha/k\hbar^2. \quad (3.7.25)$$

Using (3.7.23) and (3.7.25), Eq. (3.7.22) gives:

$$C = \left(i + \frac{2m\alpha}{k\hbar^2} \right) \sin ka + \cos ka$$

while (3.7.24) becomes:

$$\begin{aligned} ik \left[\left(i + \frac{2m\alpha}{k\hbar^2} \right) \sin ka + \cos ka \right] - k \left(i + \frac{2m\alpha}{k\hbar^2} \right) \cos ka + k \sin ka = \\ = \frac{2m\alpha}{\hbar^2} \left[\left(i + \frac{2m\alpha}{k\hbar^2} \right) \sin ka + \cos ka \right] \end{aligned}$$

which implies:

$$\operatorname{tg} ka = -k \frac{\hbar^2}{\alpha m}. \quad (3.7.26)$$

Eq. (3.7.26) can be solved graphically by looking for the intersections of the line $y = -\frac{k\hbar^2}{\alpha m}$ with the curve $y = \operatorname{tg} ka$; the points of intersection give the values of k and, consequently, of the energies $E = \frac{\hbar^2 k^2}{2m}$ for which there is no reflection.

Bibliography

- [1] L.D. Landau and E.M. Lifshitz, “*Course of Theoretical Physics, vol.1: Mechanics*”, (London, Pergamon Press).
- [2] H. Goldstein, “*Classical Mechanics*”, (Reading, MA, Addison-Wesley. Pub.Co. 1980).
- [3] P. Caldirola, R. Cirelli and R. Prosperi, “*Introduzione alla Fisica Teorica*” (UTET, Torino, 1982).
- [4] E. Fermi, “*Notes on Quantum Mechanics*” (University of Chicago Press, Chicago, 1961).
- [5] C. Cohen-Tannoudji, B. Diu and F. Laloe, “*Quantum Mechanics*” (Wiley, New York, 1977).
- [6] V. I. Kogan and V. M. Galitskij, “*Problems in Quantum Mechanics*” (Englewood Cliffs, Prentice-Hall, NJ, 1963).
- [7] F. Constantinescu and E. Magyari, “*Problems in Quantum Mechanics*” (Pergamon Press, Oxford, 1971).
- [8] R. M. Eisberg and R. Resnick, “*Quantum physics of atoms, molecules, solids, nuclei, and particles*” (Wiley, New York, 1974).
- [9] R. P. Feynman and A. R. Hibbs, “*Quantum Mechanics and Path Integrals*” (McGraw-Hill Book Co., New York, 1965).