6

The principle of least action in quantum mechanics

6.1 Introduction

Richard Feynman had to take the qualifying examination to become a candidate for the Ph.D. degree. At Princeton, the graduate students were not restricted to pursuing any special courses, but they had to take a very stiff qualifying examination; it was a long written examination for a day or two, followed by an oral examination. Three groups of professors would examine the candidate. The graduate students used to worry about this examination, and most of them dreaded it. They had a maximum of three chances in which to pass it, and some of them had failed it once or twice. To prepare for it was a serious matter.

In order to get ready for the qualifying examination, Feynman went to Cambridge, Massachusetts, where he studied for about six to eight weeks in the library of MIT during the summer of 1940. Nobody among the young people there knew him any longer, and he had the peace of mind to work in the library. He stayed at his old fraternity, where they gave him a room to stay. He was able to study without interruption or distraction. There he fully organized his knowledge of physics, things he knew and others which he did not know. He prepared a notebook, entitled 'Things I don't know about'. In it, he summarized all the subjects as best as he could, and he was rather proud of that notebook.

Feynman tried to reduce every part of physics to its fundamentals: to learn what were the essentials and the body of knowledge that was derived from them. In this way, he was able to see the pattern of each field: electrodynamics, relativity theory, quantum mechanics, statistical mechanics—everything was in his notebook. He would use this notebook many years later to prepare the lecture courses he would give at Cornell University and, later on, at the California Institute of Technology.[1] In this way, everything became organized in his mind, never to be forgotten. He made good use of this basic knowledge when, over two decades later, he gave his famous *Feynman lectures on physics* for freshmen and sophomore students at Caltech.[2] He was able to organize everything in a logical manner, so that the effort of memorization would be minimal. That required an arrangement by means of which all the logical interconnections were clear and self-contained. He completed a summary of all of fundamental physics in a concentrated and devoted study of about two months at MIT in preparation for his qualifying examination. Then he returned to Princeton, where he took the qualifying examination in the fall of 1940.

In the oral examination, one of the professors asked him about the order of colors in the rainbow: which color was at the top and which at the bottom? Feynman answered he did not know, and proceeded to work it out by drawing the curve for the index of refraction versus the wavelength.[1,3] In the curve, Feynman mixed up the indices of refraction for the blue and the red colors, but corrected himself just in time.[1,3] H. P. Robertson asked Feynman a question about stellar aberration and relativity, which he answered incorrectly.[1,3] When you look at a star through a telescope, it appears to move in a little circle due to aberration. How would the earth look in a telescope from a star? Feynman answered that the earth would also look as if it moved in a little circle. Feynman and Robertson had much argument about it, but Feynman succeeded in persuading him that he was right. When he thought more about it afterwards, he concluded that he had been wrong. The circles in the two cases would be of different sizes, and the two observations were not related. Special relativity would fail because acceleration is involved, and the motion would appear to be in a circle in which the velocity would change.[1,3]

After Feynman graduated from MIT, he wanted to get a summer job at the Bell Telephone Laboratories, and went out a couple of times to visit there. William Shockley, who knew him from the lab at MIT, would show Feynman around each time; he enjoyed these visits very much, but he did not get a job there. In the spring of 1941, Feynman went once again to the Bell Labs in New York to apply for a summer job. Shockley again showed him around and, afterwards, some of the young research scientists took him to a seafood restaurant for lunch. They were all pleased that they were going to have oysters but Feynman, who had grown up in a town near the ocean, could not even look at oysters; he forced himself to eat a couple of oysters, but felt awful and gave up. However, 'this time, which must have been my fourth or fifth time touring the Bell Labs, they accepted me. I was very happy. In those days it was hard to find a job where you could be with other scientists.'[4] But, instead of joining the Bell Labs for summer work, Feynman joined the army for wartime work.

Even before the United States of America declared war after Pearl Harbor in December 1941, there was a lot of propaganda about joining the war effort, and a lot of patriotic feelings about helping the country's business. Many businessmen were going to Plattsburg in upstate New York to join whatever service jobs were available. Feynman also thought that his technical ability to do physics might be of some use in the war effort. As a former member of the ROTC at MIT, he did not wish to join the Signal Corps, which he could have done. He talked to an army officer, who advised him to get basic training by signing up in Plattsburg, New York, and become a second lieutenant. The officer told him that that was the way the army was organized, and if he wanted to make a contribution that was the way to do it. Instead of joining Bell Labs, where he could also have done war-related research work, Feynman chose to work as a kind of engineer at the Frankfort Arsenal near Philadelphia. His job was to check everything for making a mechanical director for directing military artillery and shooting planes down. The fuses for the director were timed and powered by airburning some powder at high altitudes from airplanes, but there was not enough oxygen in the air to burn the powder.

Towards the end of summer Feynman's boss, an army colonel, realized that he was very useful and wanted to keep him in preference to the other engineers of the same rank. He offered Feynman his support and complete freedom to design the mechanical director, and tried to tempt him into staying on by promising him that he would become a big shot and could do whatever he wanted to do. In spite of all these promises, Feynman decided not to stay there and returned to Princeton. He felt sorry that he had given up the opportunity to work at the Bell Labs, where, through his scientific work, he could have made an effective contribution.

In Princeton, Feynman continued to work on his thesis from September to the end of November 1941. One morning, early in December, Robert R. Wilson, who was a young faculty member at Princeton, came into Feynman's office. He said: 'I have something to tell you. I am not supposed to say it, because it is an absolute secret. Since, however, after I tell you, you'll work on the project anyway, there's nothing to worry about.'[1,3] Wilson told Feynman that they were getting ready to build the atomic bomb in the United States, and they needed to separate uranium-235 from the ore in which it is found with the more abundant isotope uranium-238. Wilson had a process—different from the one that was eventually used—for separating the isotopes of uranium, which he wanted to develop. He told Feynman about his ideas and he wanted him to do the theoretical work to help them along. He asked Feynman to attend a meeting at 3:00 p.m., where these things would be discussed.

Feynman told Wilson that his secret was safe with him, but that he was seriously engaged on completing his thesis and did not wish to participate in the project. After Wilson left, Feynman went back to work on his thesis, but soon he began to pace the floor and think about Wilson's words. In Germany, under Hitler, there was a real possibility of developing an atomic bomb, and it gave him a fright to think that scientists under Hitler might develop the bomb first. So, Feynman decided to attend the meeting called by Wilson at 3:00 p.m., and by 4:00 p.m. the same afternoon he had a desk in an office, and was soon at work as a junior scientist at the OSRD (Office of Scientific Research and

Development) project on whether Wilson's method for the separation of isotopes 'was limited by the amount of current that you get in the ion beam'.[5] The experimentalists at Princeton were building the apparatus right there to verify the theoretical ideas.

What had happened was that all the young physicists to whom Wilson talked decided to stop their normal research and accepted to work on problems concerning the development of the atomic bomb. Feynman did the same. It was only after a while that he took a few weeks off to complete his thesis and take the final examination.

After Feynman had been working on the atomic bomb project for a few months, he got a little tired, and thought he would take six weeks off and finish his thesis. He quit the project to take his degree. John Wheeler, who had in the meantime joined the Metallurgical Laboratory at the University of Chicago, also urged Feynman to take a little time off to complete the thesis. When Feynman returned to his thesis, he found himself entirely unable to work on it. First, he was tired from the work on the atomic bomb project; second, he could not turn from one thing he had been working on intensively to something else. 'So I lazed around and I felt very guilty. Then all of a sudden ideas began to come and I wrote it all down. I solved the [thesis] problem, whatever it was, and wrote it all up. Wheeler also kept on insisting that I had enough stuff even if I did not solve any of these problems, if I never applied my theory to electrodynamics, which was the purpose of it.'[1]

6.2 Formulation of the problem

We have noted that already as an undergraduate student at MIT in the 1930s (1935–39) Richard Feynman had realized from his reading of the books of Walter Heitler (1936) and Paul Dirac (1935) that 'the fundamental problem of the day was that the quantum theory of electricity and magnetism was not completely satisfactory' and he looked at it 'as a challenge and an inspiration'.[6] His personal feeling was, that since they didn't get a satisfactory answer to the problem he had to do it himself. In a long series of conversations a couple of weeks before he died, Feynman remarked: 'All I could remember in Heitler and Dirac, and I remember distinctly well, was that they could not solve the problems. They were getting infinities, and the last sentence in Dirac's book was that some new ideas were needed. That's the main thing I got from Dirac: that new ideas were needed. This, to me, meant that I did not have to study the old ones; I didn't have to read exactly what Dirac was doing. I didn't have to find out what Heitler was telling me, what Pauli or Weisskopf, or whoever it was, was doing with the field theory, because they were all getting infinities. So I didn't have to understand what they were doing. All that was wrong; that was going to give me trouble. So what I would learn from them, if I studied them, was *what not to learn*, and I didn't learn quantum electrodynamics from those books. I learned it from an article by Fermi in the *Reviews of Modern Physics*[7] later, that there are nothing but a bunch of oscillators, the world is made up of harmonic oscillators, and each one of them is a quantum oscillator. What could be simpler? Nothing to it! Why all this hokey-pokey?'[1]

At that time Feynman understood that there existed two kinds of difficulties in the quantum electrodynamical theories. The first one, which was the classical one, still existed, namely, an infinite energy of the electron with itself. The other difficulty, as we have also noted earlier, arose from the infinite number of degrees of freedom in the electromagnetic field. Each such degree of freedom or 'mode' has a ground state energy of $\frac{1}{2}\hbar\omega$, where $\omega = 2\pi\nu$ is the circular frequency, ν being the frequency of the mode and $\hbar$ being Planck's constant. Because of the infinite number of modes associated with the electromagnetic field in a given finite volume, an infinite energy will be confined in this volume. Actually, the latter difficulty may be removed simply by changing the zero point, from which the energy is measured, but Feynman took this point very seriously.

The solution to these difficulties, which Feynman himself suggested, was as follows: (1) to eliminate the self-interaction of the particles—'hence, no self-energy!'; (2) to eliminate the field completely—there will be 'no infinite degrees of freedom of the field'.[6] These two steps are actually connected, because if one wishes to avoid the self-interaction of the particles, one has to avoid the field itself, otherwise each charged particle will interact with the common field, to the creation of which this particle has made its own contribution; hence, the self-interaction of the particles will still exist.

This general idea seemed to be so appealing that Feynman 'fell deeply in love with it'.[6] It seemed to be a possible solution, since the electromagnetic field of the particles is completely determined by the position and the motion of the charged particles. Then one could expect that the electromagnetic field might be expressed in terms of the particle variables and, therefore, it will not have any independent degrees of freedom, and the infinities due to the latter will be removed. Feynman's general plan was 'first to solve the classical problem, to get rid of the infinite self-energy in the classical theory, and to hope that when I made a quantum theory of it, everything would be just fine.'[8]

The resulting classical action-at-a-distance theory had its own problems, which we have discussed in the previous chapter. This theory was based on the principle of least action, with a new action but without the electromagnetic field. The main unsolved problem of this theory, that is, how to quantize it, still remained. When quantization was attempted, the electromagnetic field brought its own problems. Therefore, the idea of excluding the electromagnetic field entirely from the theory of charged material particles seemed to be quite appealing.

At the level of the classical theory, the new principle of least action that had been employed in the action-at-a-distance theory leads to an essential agreement with the usual form of electrodynamics. Besides being consistent with the description of point charges, it gives a unique law of radiation damping which could be checked experimentally. But if the electromagnetic field is a derived concept, there arise some physical questions which are in conflict with experiments at the quantum level, and which Feynman discussed in the introduction to his dissertation.[9]

After the concept of the electromagnetic field has been completely excluded, one has still to explain well-known processes such as the photoelectric effect and the Compton effect without light-quanta. Feynman's opinion was that 'since these phenomena deal with the interaction of light with matter their explanation may lie in the quantum aspects of matter, rather than requiring photons of light. This supposition is aided by the fact that if one solves the problem of an atom being perturbed by a potential varying sinusoidally with the time, which would be the situation if matter were quantum mechanical and light classical, one finds indeed that it will in all probability eject an electron whose energy shows an increase of $h\nu$, where ν is the frequency of variation of the potential. In a similar way an electron perturbed by the potential of two beams of light of different frequencies and different directions will make transitions to a state in which its momentum and energy is changed by an amount just equal to that given by the formulas for the Compton effect, with one beam corresponding in direction and wavelength to the incoming photon and the other to the outgoing one. In fact, one may correctly calculate in this way the probabilities of absorption and induced emission of light by an atom.'[10]

When, however, we come to spontaneous emission and the mechanism of the production of light, we come much closer to the real reason for the apparent necessity of photons. The fact that an atom emits spontaneously at all is impossible to explain by the simple picture given above. In empty space an atom emits light and yet there is no potential to perturb the system and thus force it to make a transition. The explanation of modern quantum mechanical electrodynamics is that the atom is perturbed by the zero-point fluctuations of the quantized radiation field. 'It is here that that the theory of action-at-a-distance gives us a different viewpoint. It says that an atom in empty space would, in fact, *not* radiate. Radiation is a consequence of the interaction with other atoms (namely, those in the matter which absorb the radiation). We are then led to the possibility that the spontaneous radiation of an atom in quantum mechanics also may not be spontaneous at all, but may be induced by the interaction with other atoms, so that all of the apparent quantum properties of the light and the existence of photons may be nothing more than the result of matter interacting with matter directly, and according to quantum mechanical laws.'[10]

Thus we see that Feynman was looking for the new physical concept of describing electromagnetic interactions between charged particles to resolve not only technical problems, but the fundamental difficulties in our understanding of the physical world. He pointed out two difficulties concerning the quantum analog of the action-at-a-distance theory of electromagnetic interactions. The first was that it might not be correct to represent fields as a set of harmonic oscillators with their own degrees of freedom, since in the action-at-a-distance theory the field is entirely determined by the particles. On the other hand, an attempt to deal quantum mechanically directly with the particles, which would seem to be the most satisfactory way to proceed, is faced with the circumstance that the equations of motion of the particles are expressed classically as a consequence of a principle of least action, and cannot, it appears, be expressed in Hamiltonian form.

'For this reason a method of formulating a quantum analog of systems for which no Hamiltonian, but rather a principle of least action, exists has been worked out. It is a description of this method which constitutes this thesis. Although the method was worked out with the express purpose of applying it to the theory of action-at-a-distance, it is in fact independent of that theory, and complete in itself.'[11]

The second difficulty mentioned by Feynman, to avoid which the Lagrangian method should be used is, strictly speaking, not quite correct. Actually, the action-at-a-distance theories do allow the Hamiltonian prescription if one uses the generalization of the canonical formulation of classical mechanics as given by M. Ostrogradski.[12] As far as we know, no attempt in this direction has been made up to now, but one can show from Ostrogradski's formulation of the principle of least action that this argument concerning Feynman's motivation to invent the path integral method is groundless.†

6.3 Least action in classical mechanics

Feynman started his dissertation by describing the concept of the mathematical notion of a functional, which would play a predominant role in his investigations. 'To say F is a functional of the function $q(\tau)$ means that F is a number whose value depends on the *form* of the function $q(\tau)$ (where τ is just a parameter used to specify the form of $q(\tau)$). Thus, $F=\int_{-\infty}^{\infty} q(\tau)\exp(-\tau^2)\,d\tau$ is a functional of $q(\tau)$ since it associates with every choice of the function $q(\tau)$ a

† Action-at-a-distance theories do allow Hamiltonian formulation in the framework of a proper infinite-dimensional generalization of the canonical formalism, given by Ostrogradski.[12] The formal difference between the usual theories and the action-at-a-distance theory is that in the action-at-a-distance theory the time variable in some coordinates is shifted by a time interval which describes the retarded or the advanced interactions. By using a Taylor series expansion with respect to the time of the shifted variables in the interaction terms of the action-at-a-distance theory, one arrives at a Lagrangian with infinitely higher-order derivatives. Then, by employing all derivatives of the coordinates under consideration as new generalized coordinates, one can build canonical momenta and the corresponding classical Hamiltonian as standard formulas. In this way, Feynman could have obtained a Hamiltonian formulation of the action-at-a-distance theories, and such an approach could have been useful in overcoming some of the difficulties in formulating these theories.

number, namely the integral. Also, the area under a curve is a functional of the function representing the curve, since to each such function a number, the area, is associated. The expected value of the energy in quantum mechanics is a functional of the wave function. Again, $F=q(0)$ is a functional, which is especially simple because its value depends on the (value) of the function $q(\tau)$ at the point $\tau=0$.'[13]

For a long time the important role of the functionals in mechanics, optics, geometry, and other mathematical and physical domains, had been realized. The concept of a functional first occurred in the papers of Johann Bernoulli, and was further developed by Leonhard Euler, Joseph Louis Lagrange, and several other mathematicians, up to Vito Volterra and David Hilbert, all of whom built the foundations of variational calculus and functional analysis. It is important to emphasize that the notion of the functional appeared in connection with the so-called variational principles in different domains, the most important of which is the principle of least action.

The first principle of such a kind in science was Pierre de Fermat's principle of least time, which explained the laws of geometrical optics. Fermat (1601–65) discovered that in every medium light rays travel from one point to another in such a way as to make the time taken a minimum. A similar law had been discovered by Hero of Alexandria (c. 125 BC) many centuries before Fermat for the special case in which a light ray is reflected by a mirror.[14] Fermat's principle of least time permits one to explain all the laws of propagation, diffraction, reflection, and refraction of rays in geometrical optics and was the first successful variational principle in science. As we know today, it is remarkable that this principle actually reflects the properties of the wave phenomenon, namely the propagation of light waves. But Fermat did not introduce the notion of a functional or give a modern formulation of his principle.

The next step toward the principle of least action was taken by Johann Bernoulli in 1696 when he had to formulate the problem of the brachistochrone: Given two points at different heights, not lying in the same vertical line, to find the curve, called brachistochrone, which connects these points and has the property that the body which goes down on this curve will take the least amount of time to reach the lower point starting from the upper point and moving under the force due to its weight alone. In other words, one is looking for the fastest curve of descent. The solution of this 'wonderful and unheard of' problem,' as Leibniz called it, was the cycloid and was given by Leibniz himself, by Isaac Newton, Johann Bernoulli, and the Marquis de L'Hôpital.

The most important consequences for the later development of science were that, owing to Bernoulli's problem, mechanical problems were considered in the context of variational principles and the extremely useful idea of the internal connection between mechanics and optics was established.

The successful solution of Johann Bernoulli's problem led to a new question: Is it possible to derive all mechanical laws from some variational principle? The principal question was what kind of functional could one use for this purpose? This question first appeared in the seventeenth century and, after a large number of investigations, was solved at the end of the eighteenth and the beginning of the nineteenth centuries. The new mechanical quantity needed for the new variational principle was called a 'mechanical action' and the corresponding functional an 'action functional'.

The notion of mechanical action was first formulated in an incomplete form by Leibniz, which was published only in 1890 by C. J. Gerchard: Leibniz's *Dynamica de potentia et leqibus nature corporeae*, was written in 1669 during his journey to Italy. Leibniz called this quantity '*actio formalis*'. For a body with a mass m and velocity v, which travels a distance s in time t, Leibniz defined the action as mvs, or mv^2t if one employs modern conventional notation. Although he was not able to formulate the corresponding variational principle, Leibniz nevertheless tried to make some progress in this direction.

The principle of least action was first formulated by Maupertuis on 15 April 1744 in his scientific essay 'Accord de différentes lois de la Nature qui avaient jusqu'ici paru incompatibles', where he considered mainly the spreading of light. Earlier in 1740 Maupertuis had insisted that there exists some function which has an extremum in the equilibrium states of mechanical bodies. And, finally, in 1746 Maupertuis proclaimed the principle of least action as the most general principle of Nature. He declared: 'Lorsqu'il arrive quelque changement dans la Nature, la quantité d'action, nécessaire pour ce changement, est la plus petite qu'il soit possible.' In his investigations Maupertuis was looking not only for the simplest rational principles but also for the *theological* foundation of mechanics. In his opinion, 'perfection of the Supreme Being in His divine wisdom would be incompatible with anything other than utter simplicity and minimum expenditure of action'. After he had examined several simple concrete problems, namely the direct impact of two perfectly elastic bodies and the direct impact of two perfectly inelastic bodies, as well as the refraction of light, Maupertuis almost empirically derived the expression mvs for the action. His conclusion was completely independent of the work of Leibniz, which Maupertuis did not know.

The principle of least action was not strictly formulated by Maupertuis in a truly mathematical sense. His main accomplishment was that he had formulated this principle as a *general principle of nature*, but his theolological interpretation of the principle of least action has been the subject of more controversy than any other physical principle and has met great resistance from many scientists and philosophers. Among those who, in heated discussions, had different opinions and took an active part were the following: Euler, D'Arcy, König, Courtivron, Kraft, Clemm, D'Alembert, Voltaire, and even the Prussian king Frederick II, who issued the command to burn the pamphlet entitled *Histoire du docteur Akakia et du natif de Saint-Malo*, written by Voltaire against Maupertuis's theological interpretation of the principle of least action. As D'Alembert commented, the heated discussions were similar to religious discussions with respect to their fervor, and the number of people who took part in these discussions without any actual understanding of the problem was larger than any!

The first mathematician who considered the principle of least action as an exact dynamical statement in a strictly mathematical sense was Leonhard Euler. He had invented this principle independently of, and several months before, Maupertuis, but Euler's first paper on that principle was published on 12 June 1744, or two months after the paper of Maupertuis. Much earlier, in dealing with the problem of the motion of a particle with mass m between two fixed points in a gravitational field, Euler had proved that the real trajectories are those for which the integral $\int mv\, ds$ has an extremum, i.e. a minimum or a maximum. Euler was able to give strict meaning to these ideas not only because he was a great mathematician but also because he had great experience in dealing with isoperimetric problems, which he had obtained in working with his teacher Johann Bernoulli. Euler proved the general case of this problem in 1728. The problem was to find the shortest curve between two points on a certain surface, the so-called geodesic lines. Euler proved that one has to consider the path length integral, $S=\int_A^B ds$, where ds is the elementary length on the curve. The curve with a minimal or maximal length, which connects the points A and B on a given surface, is that curve for which this integral has a stationary value, i.e. an extremum.

For the mathematical formulation of the extremum conditions, one has to consider—in addition to the trajectory from the point A to the point B—another trajectory between the same points or, in other words, one has to consider the variation of the trajectory. The two trajectories pass through the fixed points A and B. Now one may examine the values of the path length integral of S for these trajectories. The difference δS between the lengths of the two trajectories, if calculated to the first-order terms with respect to the deviation between these trajectories, is called the variation of the functional S. Then the path with extremal length from point A to point B is that for which the variation of the functional S equals zero, i.e. $\delta S=0$.

Thus we have got a variational principle which defines in general the extremal path from a point A to a point B. Euler derived a differential equation whose solution gave the extremal path; this was an ordinary differential equation of the second order which solved the isoperimetric problem completely.

Returning to the problem of the principle of least action, one can reduce Euler's discovery to the mathematical theorem

$$\delta A=0, \tag{6.1}$$

where δA is the variation of the action functional $A=\int mv\, ds$ in the problem of the motion of a mechanical particle between two fixed point centers in the plane in a gravitational field. Euler investigated this principle in several mechanical problems: the particle in a uniform gravitational field, in a central field, and under the action of constant horizontal and vertical forces; he was able to verify the validity of his principle in all these problems. But Euler's philosophy was basically different from that of Maupertuis. In the introduction of his paper, Euler wrote: 'Since all processes in Nature obey certain maximum or minimum laws, there is no doubt that the curves which bodies describe under the influence of arbitrary forces also possess some maximum or minimum property. It does not seem so easy, however, to define this property *a priori* from metaphysical principles. But, as it is possible to determine the curves themselves with the aid of a direct method, one should be able, upon thorough examination of these curves, to conclude what quantity in them must be a maximum or minimum.'[15]

Euler's result was that the quantity with extremal property is an action integral A, i.e. an integral of the momentum mv of the body on the path length ds. It is necessary to stress a very important justification which Euler had used implicitly in his principle. He had assumed that one can obtain the velocity v of a body as a function of its position in space from the law of conservation of energy, i.e.

$$\tfrac{1}{2}mv^2 + U(x, y, z) = E = \text{constant}, \tag{6.2}$$

where E is the energy constant and $U(x, y, z)$ is the potential energy of the body with the position given by the coordinates x, y, z. Euler assumed that both the real path of the body in space and the virtual path, which is needed to calculate the variation of the action functional, obey the law of conservation of energy. The corresponding variations of the paths are called *isoenergetic* and one has to regard the variation of action, denoted as $\delta_E A$, to be strictly valid only in a mathematical sense.

Euler's interpretation of the principle of least action was expressed at the end of his paper: '. . . Since the bodies resist every change of their state by reason of their inertia, they yield to the accelerating forces as little as possible, at least if they are free. It therefore follows that in the actual motion the effect arising from the forces should be less than if the body or bodies were caused to move in any other manner. Although the force of this conclusion does not convince one as satisfactory, I do not doubt that it will be possible to justify it with the aid of metaphysics. I leave this task, however, to others who are proficient in metaphysical studies.'[15] Although Euler had formulated the principle of least action as a rigorous mathematical theorem in the cases he had considered, the physical meaning of this fundamental statement was not clear to him; he had not derived dynamical equations with the help of this principle.

The next step toward the justification of the principle of least action was made by Lagrange. In his famous *La mécanique analytique*, he had considered the general case of the motion of several particles interacting by means of conservative forces, derived from certain potentials. Lagrange stated that the system moves in such a way that the total action, which is equal to the sum of the actions of all the particles, is stationary on a fixed value of the total energy: $\delta_E A = \delta_E A_1 + \delta_E A_2 + \delta_E A_3 + \cdots = 0$. This first correct formulation of the principle of least action in the general case made it possible for Lagrange to derive Newton's equations of motion for each particle.

Thus Lagrange had proved that the principle of least action in his formulation is completely equivalent to Newton's dynamical laws and may be considered as a new foundation of dynamics in general. In spite of this important observation, the principle of least action was treated only as an interesting statement for the next half century, and even in 1837 Poisson spoke about it as 'a useless rule'.[16]

Lagrange had another very important success in analytical mechanics, which was also of great importance for the subsequent development of the principle of least action. He had introduced the so-called *generalized coordinates* q_1, q_2, q_3, and so on (as many as one needs to describe unambiguously the configuration of the system at any moment t). These parameters may have different physical or geometrical meaning, like distances, angles, etc., and they possess the property of being able to express all Cartesian coordinates as functions of the q's: $x = x(q)$, $y = y(q)$, $z = z(q)$. For a given mechanical system, the number f of generalized coordinates is called the number of degrees of freedom, and one may say that any generalized coordinate q_α ($\alpha = 1, \ldots, f$) describes the corresponding degrees of freedom of the system. Using D'Alembert's principle of virtual work, Lagrange had derived new equations for such systems. The new Lagrange equations read:

$$\frac{\partial A}{\partial q_\alpha} = \frac{\partial L}{\partial q_\alpha} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_\alpha}\right) = 0 \qquad (\alpha = 1, \ldots, f). \tag{6.3}$$

Here $\dot{q}_\alpha$ are the generalized velocities, i.e. the derivatives of the generalized coordinates with respect to the time t ($\dot{q}_\alpha = dq_\alpha/dt$), and L is the so-called Lagrangian of the system. It is a function of the generalized velocities, generalized coordinates, and possibly of the time; and is given by $L = L(\dot{q}, q, t) = T - U$, the difference between the kinetic and potential energy of the system. The expression $\partial A/\partial q_\alpha$ in equation (6.3) is called the variational derivative of the action functional.

Lagrange's equations have the same form as Euler's and, therefore, they are often referred to as the Euler–Lagrange equations. But their meaning is completely different. These are dynamical equations, completely equivalent to those of Newton in the case of a system without constraints, and they allow one to solve any mechanical problem, if the forces acting upon the system are conservative, in the most economical way. The other significant property of these equations is their invariance under arbitrary coordinate transformations. Therefore they have got an extremely large number of applications in different mechanical problems and also, after corresponding generalization, in field theory.

Lagrange himself did not derive his equations from a variational principle. This was done by William Rowan Hamilton, who gave us the modern form of the principle of least action, the so-called Hamilton's principle of least action. Hamilton's main achievement in this field was that he was able to avoid the *isoenergetic* restriction on the paths in the variational principle. He proved that Euler–Lagrange equations may be derived from the principle of least action (equation (6.1)) without any constraints on the virtual paths in the configuration space of the system, where the action functional may be written as

$$A[q(t)] = \int_{t_1}^{t_2} L(\dot{q}, q, t)\, dt. \tag{6.4}$$

Here t_1 and t_2 are the initial and the final instants of time and L is the Lagrangian of the given system.

Hamilton's variational principle gives us a new viewpoint on dynamics. One may consider the initial point in configuration space of the system, given by the coordinates $q_\alpha(t_1)$ $(\alpha = 1, \ldots, f)$, and the final point, given by

$$q_\alpha(t_2)\ (\alpha = 1, \ldots, f).$$

One may connect these two points with all possible paths in this space. These are just the so-called virtual paths. Then there arises the question: Which one of these paths will be the real path of actual motion of the given system? The answer is: The path which makes Hamilton's action stationary or, in other words, the path which obeys the Hamiltonian variational principle. This path is the solution of the Euler–Lagrange equations.

The final form of the principle of least action is extremely general and finds a fundamental place in a large number of physical theories. Moreover, Hamilton's principle changed the style of doing calculations in theoretical physics. All modern field theories, like classical and quantum electrodynamics, theory of gravitation, quantum chromodynamics, theory of electroweak interactions, and theories of supergravity and superstrings, were derived from the corresponding straightforward generalization of Hamilton's principle for systems with infinitely many degrees of freedom. Today the principal problem of every new fundamental theory is how to find the corresponding Lagrangian. This principle is very useful also for calculations in many practical problems.

Thus we see that Hamilton's principle of least action looks like a fundamental law of nature and there exists the great fundamental question as to what lies behind this principle, what its meaning is, and what makes its large number of applications possible. The physicists before P.A.M. Dirac and Richard Feynman were not able to answer this question, because the principle of least action is perhaps the most basic principle of classical physics, *but its full significance lies in quantum physics; one can find it only in Feynman's path-integral formulation of quantum mechanics.*

6.4 Feynman's simple 'toy' models

After the first step of Feynman's program for the resolution of the difficulties of quantum electrodynamics at MIT had been realized in collaboration with John Archibald Wheeler, the main problem was the quantization of the Wheeler–Feynman action-at-a-distance theory. In the fall of 1941, John Wheeler was still trying to solve this problem by making use of a trick of Dirac's. His idea was to replace the 'difficult' term, $(-da_\mu da^\mu)^{1/2}$, in the action functional of the action-at-a-distance theory with the linear term $\gamma_\mu da^\mu$, where γ_μ are the Dirac matrices. Therefore Wheeler advised Feynman not to bother to work on the problem of the quantization of the action-at-a-distance theory, since Wheeler thought that he already knew the solution. But since Feynman did not know what Wheeler had in mind, he still had to find it out for himself.[1] So, although Feynman had discovered the new extremely productive general method for quantization of classical systems, because of the complexities of the Wheeler–Feynman classical theory of electromagnetic interactions he first considered simple 'toy' models and sought to work out the quantum mechanics of these toy models. Feynman's idea was to 'put the essential in, but keep everything else simple'.[1]

Since the fall of 1939, when Richard Feynman went to Princeton as a graduate student, he had been assigned to work as John Wheeler's assistant. However, with the beginning of the new fall semester in 1941, Feynman received the award of a Charlotte Elizabeth Procter Fellowship open 'only to unmarried men who are in their terminal year' of graduate studies at Princeton University. 'Then I started to work on my thesis very seriously. The problem was, of course, to make a quantum theory of the classical action-at-a-distance electrodynamics, and the form I preferred to express it in was the principle of least action, involving particles only, no field, and in which the interaction occurred at two different times. It wasn't just the velocities that were involved, but coordinates at a certain time and velocity. There was no Hamiltonian for the system, and none of the standard things that you have to convert by the standard method to quantum mechanics—the Hamiltonian, the momentum operators, etc., were not there. There was no momentum operator, because the action was of a new form. It was a direct attempt to get quantum electrodynamics. I had a classical theory with an action, but not with the Hamiltonian, and the problem was how to go to a reasonable quantum analog? The standard method of going to quantum mechanics from classical mechanics assumed that there was a Hamiltonian, but in this form there wasn't. If I had expressed it in terms of fields there might have been, but I was most reluctant to do that, and I insisted upon representing only the particles. In fact, it was because I wished to get rid of the infinite degrees of freedom of the field. I had a principle of least action involving only the particles, with delayed interaction, just the way I wanted it. This was my MIT program all over again, and it was completely satisfying: to make a classical theory with action at a distance, with delays in the interaction. The only thing that had changed since MIT was that now I had delayed as well as advanced interactions, but only the coordinates of the particles were mentioned in the fundamental law, which was the principle of least action. My first step from MIT, to fix the classical theory, had been done; now I had to go on to the quantum theory.

'Since Wheeler was fiddling around to write the paper and had told me that I should not worry about the quantum part because he "already had solved it", I had a free hand. All through that year he had been having troubles with it, and I decided that I would try to find out how the quantum theory of action-at-a-distance works by myself, since he (Wheeler) had never shown me how it worked.'[1,3]

The first thing Feynman did was to consider a toy model with a particle of mass m and one degree of freedom, described by the coordinate x, which moves in a potential $V(x)$, and interacts with itself in a distant mirror by means of retarded and advanced waves. The time it takes for light to reach the mirror from the particle is assumed to be constant, and equal to $\frac{1}{2}T_0$. The action functional of this system can be written as

$$A[x(t)] = \int_{-\infty}^{\infty} \{\tfrac{1}{2}m[\dot{x}(t)]^2 - V[x(t)] + k^2\dot{x}(t)\dot{x}(t+T_0)\}\, dt, \qquad (6.5)$$

where k depends on the charge of the particle and its distance from the mirror. One can easily check in this problem that the quantity

$$E(t) = \tfrac{1}{2}m[\dot{x}(t)]^2 + V[x(t)] + k^2\dot{x}(t)\dot{x}(t+T_0) - \int_{t}^{t+T_0} \ddot{x}(t-T_0)\ddot{x}(t+T_0)\, dt \quad (6.6)$$

is conserved. The first two terms on the right-hand side represent the kinetic plus potential energy, $V(x)$, of the particle, and depend only on the state of the particle at a given moment t. But the additional terms, representing the energy of the self-interaction of the particle via the mirror, require one to know the motion of the particle for each moment from time $t - T_0$ to $t + T_0$.

'Can we really talk about conservation, when the quantity conserved depends on the path of the particles over considerable ranges of time?' asked Feynman.[17] Many quantities, which are constants may be devised, but we should not be inclined to say that they actually represent quantities of interest, in spite of their constancy. 'The conservation of a physical quantity is of considerable interest because in solving problems it permits us to forget a great number of details. The conservation of energy can be derived from the laws of motion, but its value lies in the fact that the use of it in certain broad aspects of a problem may be discussed, without going into the great detail that is often required by direct use of laws of motion.'[17]

This remark shows how Feynman had to master the mathematical treatment of physical problems. The problem why the same conserved quantity occurs in different domains has not been resolved completely up to now. In his doctoral thesis Feynman sketched the original solution of this problem, which seems to have been ignored later. He proposed that one should 'require two things if a quantity $I(t)$ is to attract our attention as being dynamically important. The first is that it be conserved, $I(t_1)=I(t_2)$. The second is that $I(t)$ should depend only locally on the path. This is to say, if one changes the path at some time t' in a certain (arbitrary) way, the change which is made in $I(t)$ should decrease to zero as t' gets further and further from t.'[18] The expression for the energy $E(t)$, written above, satisfies this condition.

Feynman considered a completely general procedure for deriving first integrals from the action functional of a given system. This procedure was based on a deep relationship between symmetries and conservation laws, which had been first realized by Emmy Noether.[19] This fundamental result showed that a conservation law corresponds to any continuous group of symmetries of the action functional of the system under consideration. For example, the law of the conservation of energy corresponds to the invariance of the action functional under time translations, conservation of the full linear momentum of the system is a consequence of the invariance of the action functional under space translations, and the conservation of the angular momentum is a consequence of the invariance under space rotations. Thus, the so-called Noether's theorem relates the fundamental conservation laws with the properties of space and time, and gives us a method of calculating the corresponding conserved quantities from the action functional.

Feynman used a similar approach to conservation laws, but he derived his own formulas for the conserved quantities, which differed from Emmy Noether's. His approach to this problem was a more general one, because he did not use the special property of the classical Lagrangian to be a function of the generalized coordinates and generalized velocities only. Feynman's approach could be used in the theories with advanced and retarded interactions, where Noether's theorem does not work without proper modifications. Feynman 'proved a thing called Noether's theorem, not knowing that it was known.'[1]

The second toy model in Feynman's dissertation described two one-dimensional particles interacting with a one-dimensional harmonic oscillator. This example is quite instructive for action-at-a-distance electrodynamics. First, Feynman supposed that he had 'two particles A and B which do not interact directly with each other, but there is the harmonic oscillator, with which both particles A and B interact. The harmonic oscillator, therefore serves as an intermediary by means of which particle A is influenced by the motion of particle B and vice versa. In what way is this interaction through the intermediate oscillator equivalent to a direct interaction between the particles A and B, and can the motion of these particles be expressed by means of a principle of least action, not involving the oscillator? (In the theory of electrodynamics this is the problem as to whether the interaction of particles through the intermediary of the field oscillators can also be expressed as a direct action at a distance.)' [20] The corresponding action functional is

$$A[x(t), y(t), z(t)] = \int [L_y + L_z + \tfrac{1}{2}(m\dot{x}^2 - m\omega^2 x^2) + (I_y + I_z)x] \, dt, \quad (6.7)$$

where $\frac{1}{2}(m\dot{x}^2 - m\omega^2 x^2)$ is the Lagrangian of the harmonic oscillator with a coordinate x, mass m, and frequency ω; L_y and L_z are the Lagrangians of the particles (their concrete form is not essential for the following consideration); and the last term $(I_y + I_z)x$ is the Lagrangian of the interaction between the oscillator and the particles. The coordinates y and z refer to the particles A and B, respectively. The function I_y depends on the coordinates y of atom A alone, and the function I_z on the coordinate z of atom B. Then the Lagrangian equation of motion reads

$$m\ddot{x} + m\omega^2 x = I_y(t) + I_z(t) = Y(t). \quad (6.8)$$

In what way is this interaction through the intermediate oscillator equivalent to direct action between the particles A and B, and how can the motion of these particles be expressed by means of a principle of least action, not involving the oscillator? Because of its linearity, one can solve the equation of the oscillator, equation (6.8), with the arbitrary right-hand side $Y(t)$. After the oscillator equation has been solved, one can substitute the solution $x(t)$ in the equation of motion of the atoms A and B and thus eliminate the intermediate oscillator. But there exist infinitely many solutions of this type of inhomogeneous equation. Feynman showed that the proper choice of the initial conditions of the oscillator is needed to ensure energy conservation for the rest of the system after the elimination of the oscillator. For this purpose, Feynman's generalization of Noether's theorem was needed.

The next step was extremely important for all subsequent investigations of action-at-a-distance theories. We wish to derive the new equations of motion of atoms A and B from a principle of least action with a new functional which does not include the oscillator degree of freedom. The new action functional must depend only on the coordinates y and z of each of the atoms A and B. Feynman showed that it is possible to find such a new action functional only if one chooses a definitely determined solution of the oscillator equation, a symmetric one which included one-half advanced and one-half retarded interaction between the atoms A and B. This new functional reads

$$A[y(t), z(t)] = \int_{-\infty}^{\infty} (L_y + L_z) \, dt + \frac{1}{2m\omega} \int_{-\infty}^{\infty} \int_{-\infty}^{t} \sin \omega(t-s) Y(t) Y(s) \, ds \, dt. \quad (6.9)$$

Feynman's main result was that one could actually completely eliminate the oscillator degree of freedom of the initial system, and thus reduce all effects of the presence of the oscillator in an action-at-a-distance between the atoms *A* and *B* of quite specific nature, namely including in equal proportion advanced and retarded interactions. Exactly the same result had been shown to take place in the Wheeler–Feynman classical time-symmetric electrodynamics.

6.5 Feynman's invention of the new method for the quantization of classical systems

Bearing in mind that in the action-at-a-distance theories 'we are faced with the circumstance that the equations of motion of the particles are expressed classically as a consequence of a principle of least action and cannot, it appears, be expressed in Hamiltonian form',[11] Feynman concluded that we need 'a formulation of quantum mechanics . . . which does not require the idea of a Hamiltonian or momentum operator for its expression. It has, as a central mathematical idea, the analog of the action integral of classical mechanics.'[21]

In the spring of 1941, Feynman was already actively looking for a new approach to quantum mechanics directly from the classical Lagrangian. One day, when he was struggling with this problem, he went to a beer party in the Nassau tavern in Princeton. There he met Professor Herbert Jehle, who had recently arrived from Europe. He sat down near Feynman and together they began to talk about various scientific problems. Feynman asked him if he knew any way of doing quantum mechanics starting with the action, or where the action integral came into quantum mechanics. Jehle said that he did not, but told him of a paper by Dirac in which the Lagrangian, at least, came into quantum mechanics.[22]

Next day Jehle and Feynman went to the Princeton Library, and Jehle showed him Dirac's paper,[23] and they studied it together. The paper began with the words: 'Quantum mechanics was built upon a foundation of analogy with the Hamiltonian theory of classical mechanics. This is because the classical notion of canonical coordinates and momenta was found to be one with a very simple quantum analogue, . . .

'Now there is an alternative formulation for classical dynamics, provided by the Lagrangian. This requires one to work in terms of coordinates and velocities instead of coordinates and momenta. The two formulations are, of course, closely related, but there are reasons for believing that the Lagrangian one is the more fundamental.

'In the first place the Lagrangian method allows one to collect together all the equations of motion and express them as the stationary property of a certain action function. (This action function is just the time integral of the Lagrangian.) There is no corresponding action principle in terms of the coordinates and momenta of the Hamiltonian theory. [This last statement is not true, but this mistake has no bearing upon what follows.] Secondly, the

Lagrangian method can easily be expressed relativistically, on account of the action function being a relativistic invariant; while the Hamiltonian method is essentially nonrelativistic in form, since it marks out a particular time variable as the canonical conjugate of the Hamiltonian function.

'For these reasons it would seem desirable to take up the question of what corresponds in the quantum theory to the Lagrangian method in classical theory.'[23]

This was exactly what Feynman was looking for. Dirac then said that in quantum mechanics there exists a quantity which describes the time evolution of the wave function $\psi(x)$. This quantity carries the wave function $\psi(x_1)$ at a time t_1 to the wave function $\psi(x_2)$ at time t_2. Dirac called this quantity $(x_{t_2}|x_{t_1})$, but later Feynman preferred to denote it as $K(x_2, t_2; x_1, t_1)$ and we will follow Feynman's notation. Dirac pointed out that this function K is '*analogous*' to $\exp[(i/\hbar)S]$, where $S=S(x_2, t_2; x_1, t_1)$ is the classical action, as a function of the intial coordinate x_1, at the initial instant of time t_1, and the corresponding final coordinate x_2 and instant t_2. This function S was first introduced by Hamilton and carries the coordinate x_1 of the classical particle at the instant t_1 to the coordinate x_2 at the instant t_2, describing the classical evolution as a canonical transformation developing in time. The function S is called Hamilton's principal function and may be obtained by calculating the classical action functional on the real classical path from the point with the coordinate x_1 at instant t_1 to the point with the coordinate x_2 at instant t_2. Hence,

$$S=S(x_2, t_2; x_1, t_1)=\int_{t_1}^{t_2} L[x(t), x(t), t]\, dt, \tag{6.10}$$

where $x(t)$ describes the classical path of the mechanical system. Dirac also said that for an infinitesimally small time difference ε, when the initial time instant is t and the final time instant is $t+\varepsilon$, the quantum quantity $K(X, t+\varepsilon; x, t)$, x being the initial coordinate and X being the final one, is *analogous* to $\exp[(i/\hbar)\varepsilon L(X, t+\varepsilon; x, t)]$. Here one must consider the classical Lagrangian not as a function of the coordinate x and velocity v at time t but as a function of the initial coordinate x at initial time t and final corresponding coordinate X and time $t+\varepsilon$.

As Feynman read this in Dirac's paper, he asked Herbert Jehle what Dirac meant by 'analogous': did he mean that they were equal? Jehle thought not, but Feynman decided to see what happened if they were made equal.

So Feynman tried to put them equal to each other in the simplest example with the Lagrangian $L=\frac{1}{2}m\dot{x}^2-V(x)$, and he very soon found out that they were not equal but proportional, if one chose a suitable constant of proportionality, $A=(2\pi i\hbar\varepsilon/m)^{1/2}$. By doing this, Feynman obtained the result

$$\psi(X, t+\varepsilon)=\int \exp[(i/\hbar)\varepsilon L(X, t+\varepsilon; x, t)]\psi(x, t)\,\frac{dx}{A}, \tag{6.11}$$

and just calculated the integral by means of the Taylor series expansion, thus working out the Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\frac{\partial^2}{dx^2}+V(x)\right)\psi(x,t)=i\hbar\frac{\partial}{\partial t}\psi(x,t). \tag{6.12}$$

Feynman turned to Jehle, who did not quite follow, and told him that Dirac meant that they were proportional. Herbert Jehle had taken out a little notebook and was rapidly copying it down from the blackboard, and said, 'No, no, this is an important discovery. You Americans are always trying to find out how something can be used. That's a good way to discover things!' [22]

In the fall of 1946, Princeton University was celebrating its bicentennial, on the occasion of which numerous festivities, including various series of lectures were organized. In one of these sessions, devoted to science and organized by Eugene Wigner, Feynman was invited to introduce Dirac and, after his lecture, comment upon it. 'It was like the ward-heeler of the 54th district (in New York City) introducing the president of the United States. Dirac sent me his paper, in his own handwriting, to read and I had to comment on it. After Dirac's lecture, I made my comments; I tried to simplify Dirac's very technical talk for the benefit of high school teachers and others who were not familiar with the things that Dirac had talked about. But the other physicists, like Bohr and Weisskopf, who were there did not give a damn about these other people, and they criticized my attempt to "explain Dirac" in my simplified way. After I had made my criticism, people were standing around and discussing Dirac's paper, and I looked through the window and saw that Dirac was lying on the lawn outside looking up in the sky. I had never really sat and talked to him before then. But there was this question which I very much wanted to ask him, so I walked up to him and said: "Professor Dirac, you wrote in a paper[23] in which you talk about the analogy between exp $(i\varepsilon L)$ and the difference between two points." He said, "yes." I said, "Did you know that they are not just analogous, they are equal or rather proportional." He said, "Are they?" I said, "Yes." "Oh, that's interesting," was his comment. I wanted to know whether I had discovered something or not, but he had never sat down to find out whether they were equal or proportional. He just said, "No, I didn't know, are they? That's interesting!" That was the first time I talked to him personally.' [1]

In his paper Dirac was not able to complete this line of his investigations on quantum mechanics because his point of view was based on the opinion that the correspondence between the function K and the exponent of the classical action function is only an approximate semiclassical relation. From the very beginning of relativistic quantum mechanics it had been recognized that the expression $\exp[(i/\hbar)S]$ gave the semiclassical approximation to the exact quantum wave function. Therefore Dirac was looking for a proper and exact quantum analog of Hamilton's principal function S, and he found relations between the corresponding exact quantum Hamiltonian wave function and other quantum operators. Another step in this direction was taken by Edmund Whittaker.[24] Up to then this approach seems to have been quite formal and did not lead to any essentially new results. Hence, the crucial formal step to Feynman's new method was to look at the limit when ε goes to zero. In this limit one reaches an exact result for infinitesimal times.

Thus Feynman found the relation between the Lagrangian and quantum mechanics, which was an important result of his dissertation, but still for infinitesimal times. Several days later, when he was lying in bed, he worked out the next fundamental step. Feynman described it as follows: '. . . I'm lying in bed—I can still see the bed. And I can't sleep too well. And the bed was next to the wall. I got my feet up against the wall, leaning my head off on one side of the bed. You know that kind of stuff. And I'm picturing this thing and I'm putting more and more lengthy times, I have to do this again and again, and so I've got this exponential iL times again, times again, integrate it, integrate it. But the product of all the exponentials is the exponential of the sum of the L's, which is the action. So I go, AAAAAHHHHH, and I jumped, "That's the action!" That was a moment of discovery!'[1]

Now Feynman was able, by using N times the formula (6.11), to obtain exactly the right result for the function $K(X, T; x, t)$. He had to construct the expression

$$\int \cdots \int \exp\left((i/\hbar) \sum_{i=0}^{N-1} L[(x_{i+1}-x_i)/(t_{i+1}-t_i),\, x_{i+1}]\,(t_{i+1}-t_i) \right) \frac{dx_N}{A_N} \cdots \frac{dx_1}{A_1}, \tag{6.13}$$

where $t=0, t_1, t_2, \ldots, t_{N-1}, t_N=T$ are certain instants of time, which divide the time interval from the initial instant t to the final instant T into a large number of small intervals from t_1 to t_{i+1} of duration ε ($i=1, 2, \ldots, N$), such that $t_i=t+i\varepsilon$. Then, in the limit when ε goes to zero, we reach the exact quantum function K. In this limit, the expression in the exponent in equation (6.13) resembles Riemann's integral for the classical action functional:

$$A = \lim_{\varepsilon \to 0} \left(\sum_{i=0}^{N-1} L[(x_{i+1}-x_i)/(t_{i+1}-t_i),\, x_{i+1}]\,(t_{i+1}-t_i) \right). \tag{6.14}$$

Feynman's conclusion was that equation (6.11) 'is equivalent to Schrödinger's differential equation for the wave function ψ. Thus, given a classical system described by a Lagrangian, which is a function of velocities and coordinates only, a quantum mechanical description of an analogous system may be written down directly, without working out a Hamiltonian.'[25]

This approach thus promised to solve the main problem, which Feynman was trying to attack in his thesis: that is, the quantization of a classical system without knowing its Hamiltonian. In addition, it turned out that he obtained a new general procedure of quantization for classical systems.† The physical meaning of expression (6.13) and the meaning of the underlying limiting procedure was treated six years later in Feynman's paper on the 'Space–time approach to non-relativistic quantum mechanics' in the *Reviews of Modern Physics*.[28]

6.6 Further development of Feynman's results in his thesis

In his Ph.D. thesis Feynman did not develop further the physical interpretation of his new method; rather he greatly developed its mathematical formalism.

First of all, he derived a new method of calculation of quantum averages. From this, Feynman derived the generalization of Ehrenfest's theorem, which says that in quantum mechanics the classical equations of motion are fulfilled by the average values of the quantum quantities. In particular, Feynman arrived at the classical Euler–Lagrange equations (6.13) and the classical Newtonian equations in the form of quantum averages. But the most unexpected result was the fact that Feynman's relation for quantum averages is fundamental, in that 'when compared to corresponding expressions in the usual form of quantum mechanics, it contains, . . ., in one equation, both the equations of motion and commutation rules for **p** and **q**.'[29] Thus Feynman actually invented a completely new formulation of quantum mechanics in terms of classical notions.

Then, keeping in mind his main point, i.e. the quantization of action-at-a-distance theories, Feynman investigated a proper generalization of the new formalism for the case when classical action is a functional of a more general type than equation (6.4). He assumed the action to be, for example, like the action of the particle in an external potential U which interacts with itself through a distant mirror, as in equation (6.5), or like the action functional in

† The quantization procedure is the rule which gives the quantum Hamiltonian for a given classical system. In the so-called canonical quantization procedure, one takes the classical Hamiltonian $H(p, q, t)$ and substitutes the classical coordinates and momenta in this function by means of the corresponding quantum operators. This leads to the quantum Hamiltonian $\mathbf{H} = H(\mathbf{p}, \mathbf{q}, t)$, which one can use in the Schrödinger equation. But this procedure is quite ambiguous, as Schrödinger himself mentioned in his early work on wave mechanics, because the result depends on the ordering of the noncommuting quantum operators **p** and **q** in the classical Hamiltonian, and there exist, in addition, certain ambiguities of other kind.[26] Feynman's new quantization procedure looks very attractive, because 'the form of Schrödinger equation which will be arrived at will be definite and will not suffer from the type of ambiguity one finds if one tries to substitute $(\hbar/i)\partial/\partial q$ for p in the classical Hamiltonian.'[25] Unfortunately, this belief of Feynman's was not justified, and the same ambiguity presents itself in his method, as Feynman himself showed later.[27]

equation (6.9) for two particles interacting through an intermediate oscillator after the latter's degree of freedom has been eliminated.

However, there also appeared some additional difficulties of principle, a few of which Feynman overcame in his thesis, but some others remain unsolved even to this day. One such problem was the nonexistence of the wave function for action-at-a-distance theories, which Feynman succeeded in establishing. He proposed to take the viewpoint that 'the wave function is just a mathematical construction, useful under certain particular conditions to analyze the problem presented by the more generalized quantum mechanical equations . . . but not more generally applicable. It is not unreasonable that it should be impossible to find a quantity like a wave function, which has the property of describing the state of the system at one moment, and from which the state at other moments may be derived. In more complicated mechanical systems . . . the state of motion of a system at a particular time is not enough to determine in a simple manner the way that the system will change in time. It is also necessary to know the behavior of the system at other times; information which a wave function is not designed to furnish. An interesting, and at present unsolved, question whether there exists a quantity analogous to a wave function for these more general systems. . ..

'Quantum mechanics can be worked out entirely without a wave function, by speaking of matrices and expectation values only. In practice, however, the wave function is a great convenience, and dominates most of our thought in quantum mechanics.'[30]

Feynman developed the new procedure to calculate expectation values of the physical quantities, which gives the known results in the usual cases. But he was not able to prove that in theories of the action-at-a-distance type his formulas give real values for physical quantities. Later on, in a letter to C. Kelber, Feynman wrote: 'In the thesis I was trying to generalize the ideas to apply to any action function at all—not just the integral of a function of velocities and position . . . [Feynman means the Lagrangian] I met with a difficulty. An arbitrary action functional S produces results which do not conserve probability; for example, the energy values come out complex. I do not know what this means nor was I able to find that class of functionals which would be guaranteed to give real values for the energies.'[31] This was the reason why Feynman never published this part of his thesis.

In the last section of his Ph.D. thesis Feynman discussed the quantum problem of two atoms A and B, each of which interacts with an oscillator, which we have discussed earlier (see Section 6.4). '. . . To what extent can the motion of the oscillator be disregarded and atoms be considered as interacting directly?' Feynman noted that 'this problem has been solved in a special case by Fermi,[32] who has shown that the oscillators of the electromagnetic field which represent longitudinal waves could be eliminated from the Hamiltonian, provided an additional term be added representing instantaneous Coulomb interactions between particles. . . .

'... Drawing on the classical analogue we shall expect that the system with the oscillator is not equivalent to the system without the oscillator for all possible motions of the oscillator, but only for those for which some property (i.e. the initial and final position) of the oscillator is fixed. These properties, in the cases discussed, are not properties of the system at just one time, so we will not expect to find the equivalence simply by specifying the state of the oscillator at a certain time, by means of a particular wave function. *It is just for this reason that the ordinary methods of quantum mechanics do not suffice to solve this problem.*'[33]

By direct calculation of the expectation values of the system of particles Feynman derived the same effective active functional (equation (6.9)) as the action functional when one eliminates the oscillator at the classical level. The final result of this investigation of Feynman's, i.e. the particles interacting through an intermediate oscillator, was that at the quantum level, just as for the classical level, one can eliminate the oscillator and describe the particles as interacting at a distance, by using an action of the type given in equation (6.9).

In his concluding remarks Feynman mentioned that 'the problem of the form that relativistic quantum mechanics, and the Dirac equation, take from this point of view, remains unsolved. Attempts to substitute, for the action, the classical relativistic form (integral of proper time) have met with difficulties associated with the fact that the square root involved becomes imaginary for certain values of the coordinates over which the action is integrated.'

Feynman reiterated the fact that 'the final test of any theory lies, of course, in experiment. No comparison to experiment has been made in [this] paper. The author hopes to apply these methods in quantum electrodynamics. It is only out of some such direct application that an experimental comparison can be made.'[34]

As Feynman found later, 'it turned out there still remained difficulties, and the thing was not satisfactory, but I didn't realize it at the time in my excitement though, of course, I had solved the difficulties I had previously seen. But it was only a temporary error, for I thought that everything was all right and I wrote up the thesis. The parts that are right, of course, are still a representation of quantum mechanics without delay and so on. But the generalizations that are contained in the thesis are probably erroneous as written.'[1]

This was the situation in the spring of 1942 when Feynman wrote up his dissertation at the strong urging of John Wheeler. At that time Wheeler was working on the first atomic pile with Fermi in Chicago, and he wrote a letter to Feynman in which he told him that he (Feynman) 'had done enough for a thesis', and Wheeler urged him 'very strongly to write up what you have in the remaining weeks before you get into the situation in which I now find myself.'[35]

Feynman had worked on the OSRD atomic bomb project at Princeton from December 1941 to March 1942. Then, in April 1942 the Manhattan

Project was established, and Feynman continued working for it, and upon graduation after completing his thesis he went back to doing war-related research. The commencement took place around the middle of June and Feynman received his Ph.D. degree. It was a regular commencement, with academic gowns and all the hoopla of graduation from Princeton. His parents came to attend the commencement from Far Rockaway, New York, and were very proud of him.

Also, soon after commencement, on 29 June 1942, Richard Feynman and Arline Greenbaum were married.

Notes and References

1. R.P. Feynman, Interviews and conversations with Jagdish Mehra, in Austin, Texas, April 1970, and Pasadena, California, January 1988.

2. R.P. Feynman, M. Sands, and R.B. Leighton, *The Feynman lectures on physics*. Addison-Wesley, Reading, Massachusetts, 1963.

3. R.P. Feynman, Interviews with Charles Weiner (American Institute of Physics), in Pasadena, California, 1966.

4. R.P. Feynman, *SYJMF*, p. 100.

5. R.P. Feynman, Los Alamos from below, *Engineering and Science*, January/February 1976, p. 12.

6. R.P. Feynman, The development of the space-time view of quantum electrodynamics, Nobel Lecture, 11 December 1965. *Science* **153**, 699 (1966), p.1.

7. E. Fermi, *Rev. Mod. Phys.* **4**, 87 (1932).

8. R.P. Feynman, Ref. 6, p. 2.

9. R.P. Feynman, The principle of least action in quantum mechanics. A dissertation presented to the faculty of Princeton University in candidacy for the degree of Doctor of Philosophy, May 1942, p. 1.

10. R.P. Feynman, Ref. 9, pp. 4–5.

11. R.P. Feynman, Ref. 9, p. 5.

12. E. Whittaker, *A treatise on analytical dynamics of particles and rigid bodies*. Cambridge University Press, 1927, 1944, p. 245.

13. R.P. Feynman, Ref. 9, p. 8.

14. W. Yourgrau and S. Mandelstam, *Variational principles in dynamics and quantum theory*. Dover, New York, 1968, p. 5.

15. L. Euler, *Methodus inveniendi lineas curves maximi minimive proprietate*

quadentes, additamentum II, *Opera Omnia*, Series I, Vol. 24 (ed. C. Carathéodory) 1952, LII–LV, 298–308.

16. S.D. Poisson, *Traité de mécanique*, Vol. 1, Paris, 1811.

17. R.P. Feynman, Ref. 9, pp. 13–14.

18. R.P. Feynman, Ref. 9, p. 14.

19. E. Noether, Invariante Variationsproblem, *Nach. König. Gesellsch. Göttingen, math.-phys. K*1. **2**, 235 (1918).

20. R.P. Feynman, Ref. 9, pp. 18–19.

21. Ref. 9, p. 28.

22. R.P. Feynman, Ref. 6, p. 5.

23. P.A.M. Dirac, The Lagrangian in quantum mechanics. *Phys. Z. Sowjetunion* **3**(1), 64 (1933).

24. E. Whittaker, On Hamilton's principal function in quantum mechanics, *Proc. R. Soc. Edinb. A* **LXI**, 1940/41.

25. R.P. Feynman, Ref. 9, p. 34.

26. E. Schrödinger, *Ann. Phys.* (*Leipzig*) **79**, 734 (1926).

27. R.P. Feynman, Ref. 26, p. 377.

28. R.P. Feynman, Space–time approach to non-relativistic quantum mechanics, *Rev. Mod. Phys.* **20**, 376 (1948).

29. R.P. Feynman, Ref. 9, p. 38.

30. R.P. Feynman, Ref. 9, p. 49.

31. R.P. Feynman, Letter to C. Kelber, 21 February 1949, Feynman Archive, Caltech.

32. E. Fermi, *Rev. Mod. Phys.* **4**, 131 (1932).

33. R.P. Feynman, Ref. 9, p. 66.

34. R.P. Feynman, Ref. 9, p. 74.

35. J.A. Wheeler, Letter to R.P. Feynman, 26 March 1942, Feynman Archive, Caltech.