Journal of Uncertain Systems
Vol.5, No.x, pp.xx-xx, 2011
Online at: www.jus.org.uk

Why Feynman Path Integration?

Jaime Nava[1,*], Juan Ferret[2], Vladik Kreinovich[1],
Gloria Berumen[1], Sandra Griffin[1], and Edgar Padilla[1]
[1]*Department of Computer Science, University of Texas at El Paso, El Paso, TX 79968, USA*
[2]*Department of Philosophy, University of Texas at El Paso, El Paso, TX 79968, USA*

Received 19 December 2009; Revised 23 February 2010

Abstract

To describe physics properly, we need to take into account quantum effects. Thus, for every non-quantum physical theory, we must come up with an appropriate quantum theory. A traditional approach is to replace all the scalars in the classical description of this theory by the corresponding operators. The problem with the above approach is that due to non-commutativity of the quantum operators, two mathematically equivalent formulations of the classical theory can lead to different (non-equivalent) quantum theories. An alternative quantization approach that directly transforms the non-quantum action functional into the appropriate quantum theory, was indeed proposed by the Nobelist Richard Feynman, under the name of *path integration*. Feynman path integration is not just a foundational idea, it is actually an efficient computing tool (*Feynman diagrams*).

From the pragmatic viewpoint, Feynman path integral is a great success. However, from the foundational viewpoint, we still face an important question: why the Feynman's path integration formula? In this paper, we provide a natural explanation for Feynman's path integration formula.

Keywords: Feynman path integration, independence, foundations of quantum physics

1 Why Feynman Path Integration: Formulation of the Problem

Need for quantization. Since the early 1900s, we know that to describe physics properly, we need to take into account quantum effects. Thus, for every non-quantum physical theory describing a certain phenomenon, be it mechanics or electrodynamics or gravitation theory, we must come up with an appropriate quantum theory.

Traditional quantization methods. In quantum mechanics, a state of a physical system is described by a complex-valued *wave function* $\psi(x)$, i.e., a function that associates with each spatial location $x = (x_1, x_2, x_3)$ the value $\psi(x)$ whose meaning is that the square of the absolute value $|\psi(x)|^2$ is a probability density that the measured values of the spatial coordinates will be x. For a system consisting of several particles, a state can similarly described as a complex-valued function $\psi(x)$ where x is a tuple consisting of all the spatial coordinates of all the particles.

In this formulation, each physical quantity can be described as an *operator* that maps each state into a new complex-valued function. If we know this operator, then we are able to predict how the original state will change when we measure the corresponding physical quantity.

For example, in mechanics, each spatial coordinate x_i is described by an operator that transforms an arbitrary function $\psi(x)$ into a new function $x_i \cdot \psi(x)$, and each component p_i of the momentum vector $\vec{p} \stackrel{\text{def}}{=} m \cdot \vec{v}$ is described by the operator that transforms a state $\psi(x)$ into a new state $\mathrm{i} \cdot \dfrac{\partial \psi}{\partial x_i}$.

In general, these operators do not commute: e.g., when we first apply the operator x_i and then the operator p_i, then the resulting function will be, in general, different from the result of first applying p_i and then x_i. Since operators describe the process of measuring different physical quantity, this non-commutativity means that the result of measuring a momentum changes if we first measured the spatial coordinates. In other words, when we measure one of these quantities, this measurement process changes the state and thus, changes the result of measuring another quantity. On the example of spatial location and momentum, this phenomenon was first noticed already by W. Heisenberg – a physicist who realized that quantum analogues of physical

*Corresponding author. Email: jenava@miners.utep.edu (J. Nava).

quantities can be described by operators; this observation that measuring momentum changes the particle's location and vice versa is known as *Heisenberg's uncertainty principle*.

It is known that the equations of quantum mechanics can be written in terms of differential equations describing how the corresponding operators change in time. In the classical (non-quantum) limit, operators turn into the corresponding (commutative) scalar properties, and the corresponding equations turn into the classical Newton's equations. This result has led to the following natural idea of quantizing a theory: replace all the scalars in the classical description of this theory by the corresponding operators.

Limitations of the traditional quantization approach. The problem with the above approach is that due to non-commutativity of the quantum operators, two mathematically equivalent formulations of the classical theory can lead to different (non-equivalent) quantum theories. For example, if in the classical theory, if the expression for the force contains a term proportional to $x_i \cdot v_i$, i.e., proportional both to the spatial coordinates and to the velocities v_i (as, e.g., in the formula for the magnetic force), then we get two different theories depending on whether we replace the above expression with $x_i \cdot \frac{p_i}{m}$ or with $\frac{p_i}{m} \cdot x_i$.

The corresponding mathematics is non-trivial even in the simplest case, when we only have two unknowns: the location and the momentum of each particle. The situation becomes much more complicated in quantum field theory, when each spatial value $A(x)$ of each field A is a new quantity, with lots of non-commutativity.

Towards Feynman's approach: the notion of the least action principle. Starting with Newton's mechanics, the laws of physics have been traditionally described in terms of differential equations, equations that explicitly describe how the rate of change of each quantity depends on the current values of this and other quantities.

For general systems, this is a very good description. However, for *fundamental* physical phenomena – like mechanics, electromagnetic field, etc. – not all differential equations make physical sense. For example, physics usually assumes that fundamental physical quantities such energy, momentum, and angular momentum, are conserved, so that it is impossible to build a perpetual motion machine which would go forever without source of fuel. However, many differential equations do not satisfy the conservation laws and are, therefore, physically meaningless. It is therefore desirable to restrict ourselves to mathematical models which are consistent with general physical principles such as conservation laws.

It turns out that all known fundamental physical equations can be described in terms of an appropriate optimization principle – and, in general, equations following from a minimization principle lead to conservation laws.

This optimization principle is usually formulated in the following form. Our goal is to find out how the state $\gamma(t)$ of a physical system changes with time t. For example, for a single particle, we want to know how its coordinates (and velocity) change with time. For several particle, we need to know how the coordinates of each particle change with time. For a physical field, we need to know how different field components change with time. Let us assume that we know the state $\underline{\gamma}$ at some initial moment of time $\underline{t}$ ($\gamma(\underline{t}) = \underline{\gamma}$), and we know the state $\overline{\gamma}$ at some future moment of time $\overline{t}$ ($\gamma(\overline{t}) = \overline{\gamma}$). In principle, we can have different *paths* (*trajectories*) $\gamma(t)$ which start with the state $\underline{\gamma}$ at the initial moment $\underline{t}$ and end up with the state $\overline{\gamma}$ at the future moment of time $\overline{t}$, i.e., trajectories $\gamma(t)$ for which $\gamma(\underline{t}) = \underline{\gamma}$ and $\gamma(\overline{t}) = \overline{\gamma}$.

For each fundamental physical theory, we can assign, to each trajectory $\gamma(t)$, we can assign a value $S(\gamma)$ such that among all possible trajectories, the actual one is the one for which the value $S(\gamma)$ is the smallest possible. This value $S(\gamma)$ is called *action*, and the principle that action is minimized along the actual trajectory is called the *minimal actio principle*.

For example, between every two spatial points, the light follows the path for which the travel time is the smallest. This fact explains why in a homogeneous medium, light always follows straight paths, and why in the border between two substances, light refracts according to Snell's law.

The motion $x(t)$ of a single particle in a potential field $V(x)$ can also described by the action $S = \int L\,dt$, where $L = \frac{1}{2} \cdot \left(\frac{dx}{dt}\right)^2 - V(x)$. In general, the action usually takes the form $S = \int L\,dx$, where dx means integration over (4-dimensional) space-time, and the function L is called the *Lagrange function*.

The language of action and Lagrange functions is the standard language of physics: new fundamental theories are very rarely formulated in terms of a differential equation, they are usually formulated in terms of an appropriate Lagrange function – and the corresponding differential equation can then be derived from

the resulting minimization principle. Since most classical (non-quantum) physical theories are formulated in terms of the corresponding action $S(\gamma)$, it is reasonable to look for a quantization procedure that would directly transform this action functional into the appropriate quantum theory. Such a procedure was indeed proposed by Richard Feynman, under the name of *path integration*.

Feynman's path integration: main formulas. In Feynman's approach (see, e.g., [4]), the probability to get from the state $\underline{\gamma}$ to the state $\overline{\gamma}$ is proportional to $\left|\psi\left(\underline{\gamma}\to\overline{\gamma}\right)\right|^2$, where

$$\psi=\sum_{\gamma:\underline{\gamma}\to\overline{\gamma}}\exp\left(\mathrm{i}\cdot\frac{S(\gamma)}{\hbar}\right),$$

the sum is taken over all trajectories γ going from $\underline{s}$ to $\overline{s}$, and $\hbar$ is Planck's constant.

Proportionality means that the transition probability $P\left(\underline{\gamma}\to\overline{\gamma}\right)$ is equal to

$$P\left(\underline{\gamma}\to\overline{\gamma}\right)=C\left(\underline{\gamma}\right)\cdot\left|\psi\left(\underline{\gamma}\to\overline{\gamma}\right)\right|^2,$$

where the proportionality coefficient $C\left(\underline{\gamma}\right)$ should be selected in such a way that the total probability of going from a state $\underline{\gamma}$ to all possible states $\overline{\gamma}$ is equal to 1:

$$\sum_{\overline{\gamma}}P\left(\underline{\gamma}\to\overline{\gamma}\right)=C\left(\underline{\gamma}\right)\cdot\sum_{\overline{\gamma}}\left|\psi\left(\underline{\gamma}\to\overline{\gamma}\right)\right|^2=1.$$

Of course, in usual physical situations, there are infinitely many trajectories going from $\underline{\gamma}$ to $\overline{\gamma}$, so instead of the sum, we need to consider an appropriate infinite limit of the sum, i.e., an integral. Thus, in this approach, we integrate over all possible paths; because of this, the approach is called *Feynman path integration*.

Feynman's path integration: a successful tool. Feynman path integration is not just a foundational idea, it is actually an efficient computing tool. Specifically, as Feynman himself has shown, when we expand Feynman's expression for ψ into the Taylor series, each term in these series can be geometrically represented by a graph. These graphs have physical meaning – they can interpreted as representing a specific interaction between particle (e.g., that a neutron n^0 gets transformed into a proton p^+, an electron e^-, and an anti-neutrino $\widetilde{\nu}_e$: $n^0\to p^+ + e^- + \widetilde{\nu}_e$). These graphs are called *Feynman diagrams*. Their physical interpretation helps in computing the corresponding terms. As a result, by adding the values of sufficiently many terms from the Taylor series, we get an efficient technique for computing good approximations to the original value ψ.

This technique has been first used to accurately predict effects in quantum electrodynamics, a theory that has since become instrumental in describing practical quantum-related electromagnetic phenomena such as lasers.

This efficient technique is one of the main reasons why Richard Feynman received a Nobel prize in physics for his research in quantum field theory.

Comment. Feynman's path integration is not only an efficient computational tool: it also explains in explaining methodological questions. For example, in [7], Feynman's path integration is used to explain the common sense meaning of the seemingly counterintuitive Einstein-Podolsky-Rosen experiments – in which, in seeming contradiction to relativity theory, spatially separated particles instantaneously influence each other.

Feynman's path integration: remaining foundational question. From the pragmatic viewpoint, Feynman path integral is a great success. However, from the foundational viewpoint, we still face an important question: why the above formula?

What we do in this paper. In this paper, we provide a natural explanation for Feynman's path integration formula.

This explanation will be done on the physical level of rigor, we will not deal with subtleties of integration as opposed to a simple sum, all we do is justifying the general formula – without explaining how to tend to a limit.

2 Foundations of Feynman Path Integration: Main Ideas and the Resulting Derivation

Let us describe the main ideas that we will use in our derivation of path integration.

First idea: an alternative representation of the original theory. As a physical theory, we consider a functional S that assigns, to every possible path γ, the corresponding value of the action $S(\gamma)$.

From this viewpoint, a priori, all the paths are equivalent, they only differ by the corresponding values $S(\gamma)$. Thus, what is important for finding out the probability of different transitions is not which value is assigned to different paths, but how many paths are assigned to different values. In other words, what is important is the frequency with which we encounter different values $S(\gamma)$ when we select a path at random: if among N paths, only one has this value of the action, this frequency is $1/N$, if two, the frequency is $2/N$, etc.

In mathematical terms, this means that we consider the action $S(\gamma)$ as a *random variable* that takes different real values S with different probability.

Because of this analogy, we can use known representations of a random variable to represent the original physical theory. For random variables, there are many different representations. One of the most frequently used representations of a random variable α is the *characteristics function* $\xi_\alpha(\omega)$, a function that assigns, to each real number ω, the expected value $E[\cdot]$ of the corresponding exponential expression:

$$\xi(\omega) \stackrel{\text{def}}{=} E[\exp(\mathrm{i}\cdot\omega\cdot\alpha)].$$

In our case, the random variable is $\alpha = S(\gamma)$, where we have N paths γ with equal probability $1/N$. In this case, the corresponding expected value is equal to

$$\xi(\omega) = \frac{1}{N}\cdot\sum_\gamma \exp(\mathrm{i}\cdot S(\gamma)\cdot\omega).$$

Physical comment. One can notice that this expression is similar to Feynman's formula for ψ, with $\omega = \frac{1}{\hbar}$. This is not yet the desired derivation, because there are many different representations of a random variable, the characteristic function is just one of them. If we use, e.g., a cumulative distribution function or moments, we will get a completely different formulas. However, this similarity will be used in our derivation.

Mathematical comment. For a continuous 1-D random variable, with a probability density function (pdf) $\rho(x)$, the expected value corresponding to the characteristic function takes the form

$$\xi(\omega) \stackrel{\text{def}}{=} E[\exp(\mathrm{i}\cdot\omega\cdot\alpha)] = \int \exp(\mathrm{i}\cdot\omega\cdot x)\cdot\rho(x)\,dx.$$

One can see that, from the mathematical viewpoint, this is exactly the Fourier transform of the pdf (see, e.g., [2]). Thus, once we know the characteristic function $\xi(\omega)$, we can reconstruct the pdf by applying the inverse Fourier transform:

$$\rho(x) \propto \int \exp(\mathrm{i}\cdot\omega\cdot x)\cdot\chi(\omega)\,d\omega$$

(here $\propto$ means "proportional to").

For an n-dimensional random variable with a probability density $\rho(x) = \rho(x_1,\ldots,x_n)$, the characteristic function is similarly equal to the n-dimensional Fourier transform of the pdf and thus, the pdf can be uniquely reconstructed from the characteristic function by applying the inverse n-dimensional Fourier transform.

In general, the probability distribution can be uniquely determined once we know the characteristic function.

Second idea: appropriate behavior for independent physical systems. We want to derive a formula that transforms a functional $S(\gamma)$ into the corresponding transition probabilities. In many case, the physical system consists of two subsystems. In this case, each state γ of the composite system is a pair $\gamma = (\gamma_1, \gamma_2)$ consisting of a state γ_1 of the first subsystem and the state γ_2 of the second subsystem.

Often, these subsystems are *independent*. Due to this independence, the probability of going from a state $\gamma = (\gamma_1, \gamma_2)$ to a state $\gamma' = (\gamma_1', \gamma_2')$ is equal to the product of the probabilities P_1 and P_2 corresponding to the first and to the second subsystems:

$$P((\gamma_1, \gamma_2) \to (\gamma_1', \gamma_2')) = P_1(\gamma_1 \to \gamma_1') \cdot P_2(\gamma_2 \to \gamma_2').$$

In order to describe how to satisfy this property, let us recall that in physics, independence of two subsystems is usually described by assuming that the corresponding action is equal to the sum of the actions corresponding to two subsystems:

$$S((\gamma_1, \gamma_2)) = S_1(\gamma_1) + S_2(\gamma_2).$$

This relation is usually described in terms of the corresponding Lagrange functions: $L = L_1 + L_2$, which, after integration, leads to the above relation between the actions.

The reason for this sum representation is simple: the actual trajectory is the one that minimizes the total action, i.e., for which the corresponding (variational) derivatives are zeros: $\dfrac{\partial S}{\partial \gamma_1} = 0$ and $\dfrac{\partial S}{\partial \gamma_2} = 0$. When $S((\gamma_1, \gamma_2)) = S_1(\gamma_1) + S_2(\gamma_2)$, we have $\dfrac{\partial S_1(\gamma_2)}{\partial \gamma_1} = \dfrac{\partial S_2(\gamma_2)}{\partial \gamma_1} = 0$ and therefore,

$$\frac{\partial S}{\partial \gamma_1} = \frac{\partial S_1(\gamma_1)}{\partial \gamma_1} + \frac{\partial S_2(\gamma_2)}{\partial \gamma_1} = \frac{\partial S_1(\gamma_1)}{\partial \gamma_1} = 0$$

and

$$\frac{\partial S}{\partial \gamma_2} = \frac{\partial S_1(\gamma_1)}{\partial \gamma_2} + \frac{\partial S_2(\gamma_2)}{\partial \gamma_2} = \frac{\partial S_2(\gamma_2)}{\partial \gamma_2} = 0.$$

Thus, we have two independent systems of equations:

- the system $\dfrac{\partial S_1(\gamma_1)}{\partial \gamma_1} = 0$ that describes the motion of the first subsystem and which does not depend on the second subsystem at all, and
- the system $\dfrac{\partial S_2(\gamma_2)}{\partial \gamma_2} = 0$ that describes the motion of the second subsystem and which does not depend on the first subsystem at all.

Consequences of the second (independence) idea. In probabilistic terms, the fact that $S(\gamma) = S((\gamma_1, \gamma_2)) = S_1(\gamma_1) + S_2(\gamma_2)$ means that the random variable $S(\gamma)$ is the sum of two independent random variables $S_1(\gamma_1)$ and $S_2(\gamma_2)$. To simplify our analysis of this situation, we will use the fact that the the characteristic function $\xi_\alpha(\omega)$ of the sum $\alpha = \alpha_1 + \alpha_2$ of two independent random variables α_1 and α_2 is equal to the product of the characteristic functions $\xi_{\alpha_1}(\omega)$ and $\xi_{\alpha_2}(\omega)$ corresponding to these variables: $\xi_\alpha(\omega) = \xi_{\alpha_1}(\omega) \cdot \xi_{\alpha_2}(\omega)$. Indeed, by definition,

$$\xi_\alpha(\omega) = E[\exp(\mathrm{i} \cdot \omega \cdot \alpha)] = E[\exp(\mathrm{i} \cdot \omega \cdot (\alpha_1 + \alpha_2))] = E[\exp((\mathrm{i} \cdot \omega \cdot \alpha_1) + (\mathrm{i} \cdot \omega \cdot \alpha_2)].$$

Since $\exp(x + y) = \exp(x) \cdot \exp(y)$, we get

$$\xi_\alpha(\omega) = E[\exp(\mathrm{i} \cdot \omega \cdot \alpha)] = E[\exp(\mathrm{i} \cdot \omega \cdot \alpha_1) \cdot \exp(\mathrm{i} \cdot \omega \cdot \alpha_2)].$$

Due to the fact that the variables α_1 and α_2 are independent, the expected value of the product is equal to the product of expected values:

$$\xi_\alpha(\omega) = E[\exp(\mathrm{i} \cdot \omega \cdot \alpha)] = E[\exp(\mathrm{i} \cdot \omega \cdot \alpha_1)] \cdot E[\exp(\mathrm{i} \cdot \omega \cdot \alpha_2)] = \xi_{\alpha_1}(\omega) \cdot \xi_{\alpha_2}(\omega).$$

Because of this relation, instead of the original dependence of the transition probability p on the functional $S(\gamma)$, we will look for the dependence of p on $\xi(\omega)$. We have already argued that all we can use about $S(\gamma)$ is

the corresponding probability distribution, and this probability distribution can be uniquely determined by the corresponding characteristic function.

To describe a function, we must describe its values for different inputs. In these terms, we are interested in the dependence p on the variables $\xi(\omega)$ corresponding to different values ω: $p(\xi) = p(\xi(\omega_1), \ldots, \xi(\omega_n), \ldots)$. The fact that for a combination of two independent systems, the transition probability should be equal to the product of the corresponding probabilities means that

$$p(\xi_1 \cdot \xi_2) = p(\xi_1(\omega_1) \cdot \xi_2(\omega_1), \ldots, \xi_1(\omega_n) \cdot \xi_2(\omega_n), \ldots) =$$

$$p(\xi_1(\omega_1), \ldots, \xi_1(\omega_n), \ldots) \cdot p(\xi_2(\omega_1), \ldots, \xi_2(\omega_n), \ldots).$$

This equation can be further simplified if we move to a log-log scale, i.e., we use $P \stackrel{\text{def}}{=} \ln(p)$ as the new unknown and the values $Z_i = X_i + \mathrm{i} \cdot Y_i \stackrel{\text{def}}{=} \ln(\xi(\omega_i))$ as the new variables, i.e., if instead of the original function $p(\xi(\omega_1), \ldots, \xi(\omega_n), \ldots)$, we consider a new function

$$P(Z_1, \ldots, Z_n, \ldots) = \ln p(\exp(Z_1), \ldots, \exp(Z_n), \ldots)$$

Vice versa, once we know the dependence $P(Z_1, \ldots, Z_n, \ldots)$, we can reconstruct the original function

$$p(\xi(\omega_1), \ldots, \xi(\omega_n), \ldots)$$

as

$$p(\xi(\omega_1), \ldots, \xi(\omega_n), \ldots) = \exp(P(\ln(\xi(\omega_1)), \ldots, \ln(\xi(\omega_n)), \ldots)$$

Since $\ln(x \cdot y) = \ln(x) + \ln(y)$, in terms of P and Z_i, the independence requirement takes the form

$$P(Z_{11} + Z_{21}, \ldots, Z_{1n} + Z_{2n}, \ldots) = P(Z_{11}, \ldots, Z_{1n}, \ldots) + P(Z_{21}, \ldots, Z_{2n}, \ldots),$$

where we denoted $Z_{ij} \stackrel{\text{def}}{=} \ln(\xi_i(\omega_j))$.

Let us further simplify this formula by expressing it in terms of real X_{ij} and imaginary parts Y_{ij} of the real numbers $Z_{ij} = X_{ij} + \mathrm{i} \cdot Y_{ij}$. Since

$$\xi_i(\omega_j) = \exp(Z_{ij}) = \exp(X_{ij} + \mathrm{i} \cdot Y_{ij}) = \exp(X_{ij}) \cdot \exp(\mathrm{i} \cdot Y_{ij},$$

the real part X_{ij} is equal to the logarithm of the absolute value $|\xi_i(\omega_j)|$ of the complex number $\xi_i(\omega_j)$, and Y_{ij} is equal to the phase of this complex number (recall that every complex number z can be represented as $\rho \cdot \exp(\mathrm{i} \cdot \theta)$, where ρ is its absolute value and θ its phase.)

Specifically, we introduce a new function

$$\mathcal{P}(X_1, Y_1, \ldots, X_n, Y_n, \ldots) \stackrel{\text{def}}{=} P(X_1 + \mathrm{i} \cdot Y_1, \ldots, X_n + \mathrm{i} \cdot Y_n, \ldots).$$

In terms of this new function, the above relation takes the form

$$P(X_{11} + X_{21}, Y_{11} + Y_{21}, \ldots, X_{1n} + X_{2n}, Y_{1n} + X_{2n}, \ldots) =$$

$$P(X_{11}, Y_{11}, \ldots, X_{1n}, Y_{1n}, \ldots) + P(X_{21}, Y_{21}, \ldots, X_{2n}, Y_{2n}, \ldots).$$

In other words, for every two tuples $z_1 = (X_{11}, Y_{11}, \ldots, X_{1n}, Y_{1n}, \ldots)$ and $z_2 = (X_{21}, Y_{21}, \ldots, X_{2n}, Y_{2n}, \ldots)$, once we define their sum component-wise

$$z_1 + z_2 \stackrel{\text{def}}{=} (X_{11} + X_{21}, Y_{11} + Y_{21}, \ldots, X_{1n} + X_{2n}, Y_{1n} + X_{2n}, \ldots),$$

we get

$$\mathcal{P}(z_1 + z_2) = \mathcal{P}(z_1) + \mathcal{P}(z_2).$$

It is known that every continuous (or even measurable) function satisfying this equation is linear, i.e., has the form

$$\mathcal{P}(X_1, Y_1, \ldots, X_n, Y_n, \ldots) = a_1 \cdot X_1 + b_1 \cdot Y_1 + \ldots + a_n \cdot X_n + b_n \cdot Y_n + \ldots$$

(Let us recall that we operate on the physical level of rigor, where we ignore the difference between the finite sum and the limit (infinite) sum such as an integral. In general, since there are infinitely many possible values ω and thus, infinitely many variables $\xi(\omega)$, we will need an integral.)

In terms of the function $P(Z_1, \ldots, Z_n, \ldots)$ this formula takes the form

$$P(Z_1, \ldots, Z_n, \ldots) = a_1 \cdot X_1 + b_1 \cdot Y_1 + \ldots + a_n \cdot X_n + b_n \cdot Y_n + \ldots,$$

where $X_i \stackrel{\text{def}}{=} \mathrm{Re}(Z_i)$ and $Y_i \stackrel{\text{def}}{=} \mathrm{Im}(Z_i)$.

Thus, the dependence $p(\xi(\omega_1), \ldots, \xi(\omega_n), \ldots)$ takes the form

$$p(\xi(\omega_1), \ldots, \xi(\omega_n), \ldots) = \exp(P(\ln(\xi(\omega_1)), \ldots, \ln(\xi(\omega_n)), \ldots) =$$

$$\exp(a_1 \cdot X_1 + b_1 \cdot Y_1 + \ldots + a_n \cdot X_n + b_n \cdot Y_n + \ldots).$$

If we represent each complex number $\xi(\omega_i)$ by using its absolute value $\rho(\omega_i) = |\xi(\omega_i)|$ and the phase Y_i, as $\xi(\omega_i) = \rho(\omega_i) \cdot \exp(\mathrm{i} \cdot Y_i)$, then the above formula takes the form

$$p(\xi(\omega_1), \ldots, \xi(\omega_n), \ldots) = \exp(P(\ln(\xi(\omega_1)), \ldots, \ln(\xi(\omega_n)), \ldots) =$$

$$\exp(a_1 \cdot \ln(\rho(\omega_1)) + b_1 \cdot Y_1 + \ldots + a_n \cdot \ln(\rho(\omega_n)) + b_n \cdot Y_n + \ldots).$$

Since $\exp(x + y) = \exp(x) \cdot \exp(y)$, $\exp(a \cdot x) = (\exp(x))^a$, and $\exp(\ln(x)) = x$, we get

$$p(\xi(\omega_1), \ldots, \xi(\omega_n), \ldots) = \exp(P(\ln(\xi(\omega_1)), \ldots, \ln(\xi(\omega_n)), \ldots) =$$

$$\exp(a_1 \cdot \ln(\rho(\omega_1))) \cdot \exp(b_1 \cdot Y_1) \cdot \ldots \exp(a_n \cdot \ln(\rho(\omega_n)) \cdot \exp(b_n \cdot Y_n) \cdot \ldots =$$

$$(\rho(\omega_1))^{a_1} \cdot \exp(b_1 \cdot Y_1 + \cdot \ldots \cdot (\rho(\omega_n))^{a_n} \cdot \exp(b_n \cdot Y_n) \cdot \ldots$$

The phase value are determined modulo $2 \cdot \pi$, because $\exp(\mathrm{i} \cdot 2 \cdot \pi) = 1$ and hence,

$$\exp(\mathrm{i} \cdot (Y_i + 2 \cdot \pi)) = \exp(\mathrm{i} \cdot Y_i) \cdot \exp(\mathrm{i} \cdot 2 \cdot \pi) = \exp(\mathrm{i} \cdot Y_i).$$

If we continually change the phase from Y_i to $Y_i + 2 \cdot \pi$, we thus get the same state. The transition probability p_i should therefore not change if we simply add $2 \cdot \pi$ to one of the phases Y_i. When we add $2 \cdot \pi$, the corresponding factor $\exp(b_i \cdot Y_i)$ in the formula for p changes to

$$\exp(b_i \cdot (Y_i + 2 \cdot \pi)) = \exp(b_i \cdot Y_i + b_i \cdot 2 \cdot \pi) = \exp(b_i \cdot Y_i) \cdot \exp(b_i \cdot 2 \cdot \pi).$$

Thus, the only case when the expression for p does not change under this addition is when $\exp(b_i \cdot 2 \cdot \pi) = 1$, i.e., when $b_i \cdot 2 \cdot \pi = 0$ and $b_i = 0$. So, the resulting expression for p takes the form

$$p(\xi(\omega_1), \ldots, \xi(\omega_n), \ldots) = |\xi(\omega_1)|^{a_1} \cdot \ldots \cdot |\xi(\omega_n)|^{a_n} \cdot \ldots$$

We have already mentioned that each value $\xi(\omega_i)$ is equal to the Feynman sum, with an appropriate value of the parameter corresponding to Planck's constant $\hbar_i = 1/\omega_i$. So, we are close to the justification of Feynman's path integration: we just need to explain why in the actual Feynman's formula (the formula which is well justified by the experiments), there is only one factor $|\xi(\omega_i)|$, corresponding to only one value $\hbar_i = 1/\omega_i$. For this, we will need the third idea.

Third idea: maximal set of possible future states. To explain our third (and final) idea, let us recall one of the main differences between classical and quantum physics:

- in classical (non-quantum) physics, once we know the initial state γ, we can uniquely predict all future states of the system and all future measurement results γ' – of course, modulo the accuracy of the corresponding measurements;
- in quantum physics, even when we have a complete knowledge of the current state γ, we can, at best, predict *probabilities* $p(\gamma \to \gamma')$ of different future measurement results γ'.

In quantum physics, in general, many different measurement results γ' are possible. Of course, it is possible to have $p(\gamma \to \gamma') = 0$: for example, in the traditional quantum mechanics, the wave function may take 0 values at some spatial locations.

From this viewpoint, if we have several candidates for a true quantum theory, we should select a one for which the set of all inaccessible states γ' should be the smallest possible – in the sense that no other candidate theory has a smaller set. Let us look at the consequences of this seemingly reasonable requirement.

As we have mentioned, in general, the transition probability is the product of several expressions $|\xi(\omega_i)|^{a_i}$ corresponding to different values $\hbar_i = 1/\omega_i$. The product is equal to 0 if and only if at least one of these expressions is equal to 0. Thus, the set of all the states γ' for which the product is equal to 0 is equal to the union of the sets corresponding to individual expressions. Thus, the smallest possible set occurs if the whole product consists of only one term, i.e., if

$$p(\xi) = |\xi(\omega)|^a$$

for some constants ω and a.

The only difference between this terms and Feynman's formula is that here, we may have an arbitrary value a, while in Feynman's formula, we have $a = 2$.

Fourth idea: analyticity and simplicity. In physics, most dependencies are real-analytical, i.e., expandable in convergent Taylor series. For a complex number $z = x + \mathrm{i} \cdot y$, the function $|z|^a = \left(\sqrt{x^2 + y^2}\right)^a = (x^2 + y^2)^{a/2}$ is analytical at $z = 0$ if and only if a is non-negative even number, i.e., $a = 0$, $a = 2$, $a = 4$, ...

The choice $a = 0$ corresponds to a degenerate physical theory $p =$ constant, in which the probability of the transition from a given state γ to every other state γ' is exactly the same. Thus, the simplest non-trivial case is the case of $a = 2$, i.e., the case of Feynman integration.

Physical comments.

- The simplicity arguments are, of course, much weaker than the independence arguments given above. Instead of these simplicity arguments, we can simply say that the empirical observation confirms that $p = |\psi|^2$ as in Feynman path integration but not that $p = |\psi|^4$ or $p = |\psi|^6$ etc., as in possible alternative theories.
- The alternatives $a \neq 2$ have a precise mathematical meaning which can be illustrated on the example of a system consisting of a single particle. For this particle, we can talk about the probability of this particle to be at a certain spatial location x. For $a = 2$, this is proportional to $|\psi(x)|^2$, so the fact that the total probability of finding a particle at one of the locations is equivalent to requiring that $\int |\psi(x)|^2\, dx = 1$, i.e., to the fact that the wave function belongs to the space L^2 of all square integrable functions.

  For $a \neq 2$, the corresponding probability density is proportional to $|\psi(x)|^a$, so we arrive at a similar requirement that $\int |\psi(x)|^a\, dx = 1$. In mathematical terms, this means that the corresponding wave function ψ belongs to the space L^p of all the functions whose p-th power is integrable, for $p = a \neq 2$. It is worth mentioning that the idea of using L^p space has been previously proposed in the foundations of quantum physics; see, e.g., [1, 3, 5, 6].

Acknowledgments

This work was supported in part by the National Science Foundation grants HRD-0734825 and DUE-0926721, and by Grant 1 T36 GM078000-01 from the National Institutes of Health.

The authors are thankful to all the participants of the 6th Joint UTEP/NMSU Workshop on Mathematics, Computer Science, and Computational Sciences (El Paso, Texas, November 7, 2009) for valuable discussions.

References

[1] Bugajski, S., Delinearization of quantum logic, *International Journal of Theoretical Physics*, vol.32, no.3, pp.389–398, 1993.

[2] Cormen, T. H., C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, Cambridge, Massachusetts, 2009.

[3] Davies, E. B., and J. T. Lewis, An operational approach to quantum probability, *Communications in Mathematical Physics*, vol.17, no.3, pp.239–260, 1970.

[4] Feynman, R., R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

[5] Foulis, D. J., and M. K. Bennett, Effect algebras and unsharp quantum logics, *Foundations of Physics*, vol.24, no.10, pp.1331–1352, 1994.

[6] Hamhalter, J., Quantum structures and operator algebras, In: K. Engesser, D. M. Gabbay, and D. Lehmann (eds.), *Handbook of Quantum Logic and Quantum Structures: Quantum Structures*, Elsevier, Amsterdam, pp.285–333, 2007.

[7] Sinha, S., and R. D. Sorkin, A sum-over-histories account of an EPR(B) experiment, *Foundations of Physics Letters*, vol.4, pp.303–335, 1991.