Quantum gravidynamics. I. Path integrals in curved space-time

1975 J. Phys. A: Math. Gen. 8 16

(http://iopscience.iop.org/0305-4470/8/1/007)

J. Phys. A: Math. Gen., Vol. 8, No. 1, 1975. Printed in Great Britain.

Quantum gravidynamics I. Path integrals in curved space–time

M Clutton-Brock†

School of Physics, The University, Newcastle upon Tyne, NE1 7RU, UK

Received 5 April 1974, in final form 24 May 1974

Abstract. The aim is to develop in as simple and elementary a fashion as possible the quantum dynamics of gravitating particles in a strongly curved classical background geometry. This type of theory is important when the radius of curvature of the universe is less than the Compton wavelength of the particles. Path integrals provide the simplest entry to quantum gravidynamics, and part I considers path integrals without spin. Spin will be introduced in part II, and perturbation expansions in part III.

Path integrals are convolutions of kernels for many small intervals: the curved space kernel for a small interval is found by suitably generalizing the free-particle flat space kernel. The structure of a relativistic path integral is very different from that of a non-relativistic path integral, and is illustrated by describing how to evaluate a path integral—at least in principle—by a crude Monte Carlo method.

Path integrals for several particles involve an action principle, which is expressed in terms of particle paths using an integral formulation of Mach's principle. When particles approach to within their gravitational radius, path integrals and even the space–time continuum become meaningless.

1. Quantum theories of gravitation in perspective

There is not just one quantum theory of gravitation, but at least three, of different levels of complexity and sophistication, and appropriate to different physical situations. This series of papers aims to develop the simplest level of theory, in as simple and elementary a fashion as possible. This level of theory, which I call quantum gravidynamics, describes the interactions of gravitating particles and gravitons in a classical background geometry. Because only the particles and gravitons are quantized, and not the background geometry, the interactions are essentially linear. There is a strong analogy with elementary quantum electrodynamics. Quantum gravidynamics becomes important during the Compton era, $t_u \lesssim 10^{-23}$ s, when the radius of curvature of the universe becomes less than the Compton wavelength of typical particles.

Most efforts to quantize gravitation are directed to the much more difficult problem of the non-linear quantization of the background geometry, which might be called quantum geometrodynamics (see, for example, reviews by de Witt 1972, and by Brill and Gowdy 1970). Quantum geometrodynamics becomes important during the Planck era, $t_u \lesssim 10^{-44}$ s, when the radius of curvature of the universe becomes less than the Planck length $(G\hbar/c^3)^{1/2}$. There are then violent quantum fluctuations in the geometry, leading to topological changes; wormholes, geons, and other exotic objects appear.

† Present address: CRESS, York University, Downsview, Ontario, Canada.

Quantum geometrodynamics still assumes that space–time is a continuum, but in the Schwarschild era, $t_u \lesssim 10^{-80}$ s, particles are squeezed together so closely that no test particle can make meaningful measurements of position. Then we need an even more profound theory, perhaps along the lines suggested by Penrose (1971, 1973), who tries 'to get rid of the continuum and build up physical theory from discreteness'.

2. Path integrals provide the most direct entry to quantum gravidynamics

The Compton era occupies about 20 decades of cosmic time. The background gravitational field is strong enough to create particle–antiparticle pairs (Audretsch 1973). CP violating reactions may produce a slight excess of particles over antiparticles (Kuzmin 1970), just sufficient to account for the present baryon–photon ratio. If the universe was homogenized during the Compton era, then statistical density fluctuations would have grown to about the right size to create galaxies (Clutton-Brock 1974). The initial conditions of the universe seem to be laid down during the Compton era.

The physics of the Compton era will be exceedingly complex and we need to use the simplest possible formalism in order to cope with the complex physics.

Now the central idea of Einstein's theory is one of beautiful simplicity. The gravitational field is identified with the curvature of space–time, and the inertial paths of particles are identified with geodesics. We might ask: how does a particle 'know' which path is a geodesic? Feynman (1948) gives an answer which is both astonishing and simple: the particle does *not* know. Instead, it travels over all possible paths. Each compatible path contributes to a process an amplitude e^{iS} where S is the action along the path. The probability amplitude for the process is the integral of e^{iS} over all compatible paths—a path integral.

Surely the simplest and most direct route to quantum gravidynamics is obtained by fusing these two beautiful ideas. So we shall in this series of papers study the idea of path integrals in curved space–time. To make things as clear as possible, we shall start with path integrals for spinless particles, and postpone the study of spin until part II.

The ultimate goal of quantum theory is to find transition amplitudes for processes, and so we wish to express transition amplitudes as path integrals. One can think of a path integral as a way of building up the amplitude for a complex process out of the amplitudes for many simple processes. The simplest process is for a particle to go from one point to another neighbouring point, and this is the kernel for a small interval which we find in § 3. We use this kernel in § 4 to build up transition amplitudes in the form of path integrals, and show in § 5 how they might be evaluated by Monte Carlo methods. Quantum gravidynamics proper begins when particles interact, and so we derive in § 6 an action integral for mutually gravitating particles.

One advantage of this approach is that it exposes with particular clarity the limits of its validity, as we see in § 7. The breakdown of the space–time continuum, which occurs in the Schwarzschild era, is reflected in the inability to form a path integral from the action for strongly gravitating particles.

3. An approximate kernel for small intervals

The amplitude for a particle to go from x' to x'' is the kernel $K(x'', x')$. Now we can find the kernel for a free particle in flat space–time quite easily. Over a small interval, the effect of curvature is small, and so we can find an approximate kernel in curved space–time

by generalizing the flat-space kernel. The approximation will be good if the interval is small. The kernel for a large interval is found as a convolution of kernels for many small intervals: this is a path integral.

The kernel $K(x'', x')$ relates the wavefunction $\psi(x'')$ at x'' to the wavefunction $\psi(x')$ on a 3-surface V'. In relativistic quantum mechanics the 3-surface V' completely encloses the point x'':

$$\psi(x'') = \int K(x'', x')Q'\psi(x')\,\mathrm{d}V'. \tag{3.1}$$

We shall call Q the *path operator*, and it contains the specifically quantum part of the dynamics. For a Dirac particle, Q will be a spinor, while K will remain a scalar. For a free Klein–Gordon particle, Q contains Kirchoff's modification of Huygens' principle

$$Q = (\vec{\partial}_k - \overleftarrow{\partial}_k)N^k, \tag{3.2}$$

where $\vec{\partial}_k$ and $\overleftarrow{\partial}_k$ indicate differentiation to the right and left, and N^k is a vector orthogonal to the invariant element of 3-surface $\mathrm{d}V$ such that

$$N_k\,\mathrm{d}V = (-g)^{1/2}\epsilon_{klmn}\,\mathrm{d}x^l\,\mathrm{d}x^m\,\mathrm{d}x^n. \tag{3.3}$$

Because V' encloses x'', we can use Gauss' theorem to transform the 3-surface integral (3.1) into an integral over the 4-content C' enclosed by V':

$$\psi(x'') = \int D^{k'}[K(x'', x')(\vec{\partial}'_k - \overleftarrow{\partial}'_k)\psi(x')]\,\mathrm{d}C', \tag{3.4}$$

where D' indicates covariant differentiation with respect to x'. Now ψ obeys the free particle Klein–Gordon equation

$$(-D^k\partial_k + m^2)\psi = 0, \tag{3.5}$$

and so we may put (3.4) into the form

$$\psi(x'') = \int [(-D^{k'}\partial'_k + m^2)K(x'', x')]\psi(x')\,\mathrm{d}C'. \tag{3.6}$$

This yields immediately the differential equation that the kernel must obey

$$(-D^{k'}\partial'_k + m^2)K(x'', x') = \mathscr{I}(x'', x'), \tag{3.7}$$

where $\mathscr{I}(x'', x')$ is the invariant identity kernel.

In the presence of an electromagnetic potential ϕ, ψ must obey the equation

$$[-(D^k - \mathrm{i}e\phi^k)(\partial_k - \mathrm{i}e\phi_k) + m^2]\psi = 0, \tag{3.8}$$

and the path operator Q must be modified to

$$Q = [(\vec{\partial}_k - \mathrm{i}e\phi_k) - (\overleftarrow{\partial}_k + \mathrm{i}e\phi_k)]N^k \tag{3.9}$$

in order that the kernel should obey the differential equation

$$[-(D^{k'} + \mathrm{i}e\phi^{k'})(\partial'_k + \mathrm{i}e\phi'_k) + m^2]K(x'', x') = \mathscr{I}(x'', x'). \tag{3.10}$$

In flat space–time the free-particle differential equation (3.7) is easily solved as a Fourier integral

$$K(x'', x') = \frac{1}{(2\pi)^4}\int \exp[\mathrm{i}p_k(x'' - x')^k]\frac{\mathrm{d}^4p}{\eta^{ab}p_ap_b + m^2}. \tag{3.11}$$

The integration over p_0 is along the Feynman contour, which ensures that positive energies are propagated into the future, and negative energies into the past (Feynman 1949). The fact that only the appropriate pole of p_0 contributes means that the integration is over one sheet of the momentum surface

$$\Omega(p) \equiv \eta^{ab}p_a p_b + m^2 = 0. \tag{3.12}$$

When $x''-x'$ is future time-like or null, the positive energy sheet must be used; when $x''-x'$ is past time-like or null, the negative energy sheet must be used; when $x''-x'$ is space-like, either sheet may be used. Thus we can write the element of momentum surface in the alternative forms

$$\mathrm{d}\Omega(p) = \frac{1}{(2\pi)^4}\frac{\mathrm{d}^4 p}{\eta^{ab}p_a p_b + m^2}, \tag{3.13a}$$

or

$$\mathrm{d}\Omega(p) = \frac{1}{(2\pi)^3} U[-p_0(x''-x')^0]\delta[\Omega(p)]\,\mathrm{d}^4 p, \tag{3.13b}$$

where $U(\ldots)$ is the unit step function and $\delta(\ldots)$ the Dirac delta function.

While the flat-space kernel (3.11) is exact, the curved-space analogue can only be an approximation good for small intervals. The reason is connected with the impossibility of covering a region in curved space–time with a field of parallel momentum vectors. Thus, we cannot construct a momentum space over a region, but at most along a single curve—in our case, the geodesic joining x'' and x'. The kernel $K(x'', x')$ is, however, a sum over *all* paths joining x'' and x'. If we take the momentum vector p_k by parallel transport from x' to x'' along a non-geodesic path, it will not in general be parallel at x'' to the momentum vector transported along the geodesic. The deviation from parallelism is proportional to the area between the two paths, and so is of the second order in the interval. We may expect the error of the approximate kernel, which is based only on the geodesic path, to be also of the second order in the interval.

The momentum vector can be defined along the whole geodesic by parallel transport, but its coordinates must be defined by reference to a single point. So what might seem the natural curved-space generalization of the momentum surface,

$$\Omega(p, x) \equiv g^{ab}(x)p_a p_b + m^2 = 0, \tag{3.14}$$

is in fact very inconvenient because of its dependence on position. So we shall work in terms of the 'mechanical momentum' q, which is defined as that function of the canonical momentum p and position x which makes the momentum surface take the simple form

$$\Omega(q) \equiv \eta^{\alpha\beta}q_\alpha q_\beta + m^2 = 0. \tag{3.15}$$

For a free particle in curved space–time, the components q_α are just the tetrad components of the canonical momentum p_k:

$$q_\alpha = \lambda^k_\alpha(x)p_k, \qquad p_k = \lambda^\alpha_k(x)q_\alpha. \tag{3.16}$$

For a charged particle in an electromagnetic potential ϕ, the mechanical momentum is given by

$$q_\alpha = \lambda^k_\alpha(p_k - e\phi_k), \qquad p_k = \lambda^\alpha_k q_\alpha + e\phi_k. \tag{3.17}$$

As an aside, we notice that Hamilton's principle

$$\delta S = \delta \int_{x'}^{x''} p_k \, \mathrm{d}x^k = 0 \qquad \text{subject to } \Omega(p, x) = 0, \tag{3.18}$$

can be expressed in the alternative form

$$\delta S = \delta \int_{x'}^{x''} p_k(q, x) \, \mathrm{d}x^k = 0 \qquad \text{subject to } \Omega(q) = 0, \tag{3.19}$$

which leads to equations of motion for q and x.

In curved space-time, the expression $p_k(x'' - x')^k$ has no invariant meaning, and must be replaced by the integral

$$S\begin{pmatrix} x'' \leftarrow x' \\ q \end{pmatrix} = \int_{x'}^{x''} p_k(q, x) \, \mathrm{d}x^k, \tag{3.20}$$

which is the action for the particle to go from x' to x'' with momentum q. The components of q at an arbitrary position x must be defined by parallel transport from some reference point x_q, so

$$p_k(q, x) = q_\alpha \lambda^\alpha_a(x_q) g^a(x_q, x)_k = q_\alpha g^\alpha(x_q, x)_k, \tag{3.21}$$

where $g^a(x_q, x)_k$ is the parallel propagator (Synge 1960). Provided x_q is on the geodesic, the momentum (3.21) and hence the action (3.20) is independent of where on the geodesic x_q lies. The integral (3.20) has then a simple geometrical interpretation: it is the scalar product of q_α with the vector $T^\alpha = T^\alpha(x_q; x'' \leftarrow x')$ tangent at x_q to the geodesic and of magnitude equal to the invariant measure along the geodesic from x' to x''. Thus, if u is an affine parameter

$$\begin{aligned} S\begin{pmatrix} x'' \leftarrow x' \\ q \end{pmatrix} &= q_\alpha T^\alpha(x_q; x'' \leftarrow x') \\ &= q_\alpha \lambda^\alpha_k(x_q)(u'' - u')\left(\frac{\mathrm{d}x^k}{\mathrm{d}u}\right)_q . \end{aligned} \tag{3.22}$$

When x_q is placed at x' or x'', T^α may be expressed in terms of the derivative of the world function $W(x'', x')$ defined by Synge (1960) as

$$W(x'', x') = \frac{1}{2}(u'' - u') \int_{u'}^{u''} g_{ab}(x) \frac{\mathrm{d}x^a}{\mathrm{d}u} \frac{\mathrm{d}x^b}{\mathrm{d}u} \, \mathrm{d}u. \tag{3.23}$$

Since

$$\begin{aligned} W^\alpha(x'', \alpha') &= \lambda^{\alpha k}(x'') \partial''_k W(x'', x') \\ &= \lambda^{\alpha k}(x'')(u'' - u') \frac{\mathrm{d}x''_k}{\mathrm{d}u} = T^\alpha(x''; x'' \leftarrow x'), \end{aligned} \tag{3.24}$$

and

$$\begin{aligned} W(x'', x')^\alpha &= \lambda^{\alpha k}(x') \partial'_k W(x'', x') \\ &= -\lambda^{\alpha k}(x')(u'' - u') \frac{\mathrm{d}x'_k}{\mathrm{d}u} = -T^\alpha(x'; x'' \leftarrow x'), \end{aligned} \tag{3.25}$$

we have the expressions

$$S\begin{pmatrix} x'' \leftarrow x' \\ q'' \end{pmatrix} = q''_\alpha W^\alpha(x'', x'), \qquad S\begin{pmatrix} x'' \leftarrow x' \\ q' \end{pmatrix} = -W(x'', x')^\alpha q'_\alpha, \tag{3.26}$$

which are useful when we wish to find the derivatives of the action.

To summarize, the approximate kernel for a small interval in curved space–time is

$$\hat{K}(x'', x') = \int \exp\left[\mathrm{i}S\begin{pmatrix} x'' \leftarrow x' \\ q \end{pmatrix}\right] \mathrm{d}\Omega(q), \tag{3.27}$$

where $\mathrm{d}\Omega(q)$ is the element of momentum surface:

$$\begin{aligned} \mathrm{d}\Omega(q) &= \frac{1}{(2\pi)^4} \frac{\mathrm{d}^4 q}{\eta^{\alpha\beta} q_\alpha q_\beta + m^2} \\ &= \frac{1}{(2\pi)^3} U(-q_0 T^0) \delta[\Omega(q)]\, \mathrm{d}^4 q, \end{aligned} \tag{3.28}$$

which lies on the positive energy sheet when $x'' \leftarrow x'$ is future time-like or null, on the negative energy sheet when $x'' \leftarrow x'$ is past time-like or null, and on either sheet when $x'' \leftarrow x'$ is space-like.

The curved-space kernel (3.27) seems to me to be the simplest covariant generalization of the flat-space kernel, but it is not the only possible generalization. The reader may feel that the particular choice (3.27) is in need of further justification.

The real justification for an approximation lies in an error estimate. It is more economical to postpone the discussion of errors to part III, where we can treat the Klein–Gordon and Dirac kernels at the same time. We shall show that the error is of the second order in the interval, which is sufficiently small for the kernel to be used in a path integral.

4. Transition amplitudes as path integrals

The transition amplitude to go from a state ψ_I defined on a 3-surface V_I to a state ψ_F defined on a 3-surface V_F is

$$T_\mathrm{FI} = \int\int \tilde{\psi}_\mathrm{F} Q_\mathrm{F} K(x_\mathrm{F}, x_\mathrm{I}) Q_\mathrm{I} \psi_\mathrm{I}\, \mathrm{d}V_\mathrm{F}\, \mathrm{d}V_\mathrm{I}. \tag{4.1}$$

The kernel (3.27) for a small interval is just $\mathrm{e}^{\mathrm{i}S}$ integrated over the momentum surface. The kernel for a large interval is a convolution of kernels for many small intervals, and should therefore be $\mathrm{e}^{\mathrm{i}S}$ integrated over all paths in both momentum and coordinate space. We aim therefore to express the transition amplitude as a path integral

$$T_\mathrm{FI} = \int \tilde{\psi}_\mathrm{F} Q(\mathrm{path}) \psi_\mathrm{I} \exp[\mathrm{i}S(V_\mathrm{F} \leftarrow \mathrm{path} \leftarrow V_\mathrm{I})]\, \mathrm{d}(\mathrm{path}). \tag{4.2}$$

To see how to do this, consider the simple path

$$\{\mathrm{path}\} = \begin{Bmatrix} x_1 \leftarrow x_3 \leftarrow x_5 \\ q_2 \leftarrow q_4 \end{Bmatrix} \tag{4.3}$$

where the particle goes from x_5 to x_3 with momentum q_4 and from x_3 to x_1 with

momentum q_2. The kernel for this path is

$$K(x_1, x_5) = \int K(x_1, x_3) Q_3 K(x_3, x_5)\, \mathrm{d}V_3$$

$$= \iiint \exp\left[\mathrm{i}S\begin{pmatrix} x_1 \leftarrow x_3 \\ q_2 \end{pmatrix}\right] Q_3 \exp\left[\mathrm{i}S\begin{pmatrix} x_3 \leftarrow x_5 \\ q_4 \end{pmatrix}\right] \mathrm{d}\Omega_2\, \mathrm{d}V_3\, \mathrm{d}\Omega_4. \tag{4.4}$$

Now identify x_F with x_1 and x_I with x_5, and substitute (4.4) into (4.1) to obtain

$$T_{\mathrm{FI}} = \int \cdots \int \mathrm{d}V_1\, \mathrm{d}\Omega_2\, \mathrm{d}V_3\, \mathrm{d}\Omega_4\, \mathrm{d}V_5 \tilde{\psi}_\mathrm{F}(x_1) Q_1 \exp\left[\mathrm{i}S\begin{pmatrix} x_1 \leftarrow x_3 \\ q_2 \end{pmatrix}\right]$$

$$\times Q_3 \exp\left[\mathrm{i}S\begin{pmatrix} x_3 \leftarrow x_5 \\ q_4 \end{pmatrix}\right] Q_5 \psi_\mathrm{I}(x_5). \tag{4.5}$$

Now allow the path operators to operate on the exponentials, and obtain

$$Q_1 = N_1^k \left\{ -\overleftarrow{\partial}_{k1} + \mathrm{i}\partial_{k1} S\begin{pmatrix} x_1 \leftarrow x_3 \\ q_2 \end{pmatrix} \right\}, \tag{4.6a}$$

$$Q_3 = \mathrm{i}N_3^k \partial_{k3} \left\{ -S\begin{pmatrix} x_1 \leftarrow x_3 \\ q_2 \end{pmatrix} + S\begin{pmatrix} x_3 \leftarrow x_5 \\ q_4 \end{pmatrix} \right\}, \tag{4.6b}$$

$$Q_5 = N_5^k \left\{ -\mathrm{i}\partial_{k5} S\begin{pmatrix} x_3 \leftarrow x_5 \\ q_4 \end{pmatrix} + \vec{\partial}_{k5} \right\}. \tag{4.6c}$$

We can simplify (4.6) somewhat if we use

$$\partial_{k1} S\begin{pmatrix} x_1 \leftarrow x_3 \\ q_2 \end{pmatrix} = q_{\alpha 2} \partial_{k1} \int_{x_3}^{x_1} g^\alpha(x_2, x)_a\, \mathrm{d}x^a$$

$$= +g_k(x_1, x_2)^\alpha q_{\alpha 2} + \mathrm{O}(\Delta^2), \tag{4.7a}$$

and

$$\partial_{k3} S\begin{pmatrix} x_1 \leftarrow x_3 \\ q_2 \end{pmatrix} = -g_k(x_3, x_2)^\alpha q_{\alpha 2} + \mathrm{O}(\Delta^2), \tag{4.7b}$$

to obtain

$$Q_1 = N_1^k \{ -\overleftarrow{\partial}_{k1} + \mathrm{i}g_k(x_1, x_2)^\alpha q_{\alpha 2} \} + \mathrm{O}(\Delta^2), \tag{4.8a}$$

$$Q_3 = \mathrm{i}N_3^k \{ g_k(x_3, x_2)^\alpha q_{\alpha 2} + g_k(x_3, x_4)^\alpha q_{\alpha 4} \} + \mathrm{O}(\Delta^2), \tag{4.8b}$$

$$Q_5 = N_5^k \{ \mathrm{i}g_k(x_5, x_4)^\alpha q_{\alpha 4} + \vec{\partial}_{k5} \} + \mathrm{O}(\Delta^2). \tag{4.8c}$$

The approximation (4.8) is of the same order as that of the kernel, and so it is just as good to use (4.8) as to use the exact expression (4.6). Moreover, in an electromagnetic potential (4.6) must be modified, but the approximation (4.8) remains valid without modification.

When the path operators are in the form (4.6) or (4.8), we can separate them from the exponentials, and write the transition amplitude in the form

$$T_{\mathrm{FI}} = \int \cdots \int \tilde{\psi}_\mathrm{F}(x_1) Q_1 Q_3 Q_5 \psi_\mathrm{I}(x_5)$$

$$\times \exp\left[\mathrm{i}S\begin{pmatrix} x_1 \leftarrow x_3 \leftarrow x_5 \\ q_2 \leftarrow q_4 \end{pmatrix}\right] \mathrm{d}V_1\, \mathrm{d}\Omega_2\, \mathrm{d}V_3\, \mathrm{d}\Omega_4\, \mathrm{d}V_5. \tag{4.9}$$

This can be generalized to an N-step path integral in the form (4.2) provided we identify

$$Q(\text{path}) = Q_1 Q_3 \dots Q_{2N+1}, \tag{4.10}$$

$$\mathrm{d}(\text{path}) = \mathrm{d}V_1\, \mathrm{d}\Omega_2\, \mathrm{d}V_3\, \mathrm{d}\Omega_4 \dots \mathrm{d}\Omega_{2N}\, \mathrm{d}V_{2N+1}, \tag{4.11}$$

where $Q_3 \dots Q_{2N-1}$ take the form (4.6*b*) or (4.8*b*), and Q_1 and Q_{2N+1} take the form (4.6*a*) and (4.6*c*) or (4.8*a*) and (4.8*c*).

In order to use this recipe for transition amplitudes, we have to know how to construct wavefunctions ψ_1 and ψ_F corresponding to physical states of definite energy ($\pm$) and momentum. This is complicated by our inability to construct fields of parallel momentum vectors, so that there is no curved space analogue of the momentum eigenfunction $\psi_p = \exp(\mathrm{i}p_k x^k)$. Fortunately, we do not need to define the wavefunction over all space–time: it is sufficient to define the wavefunction, together with its derivative, on a 3-surface (V_I for ψ_I and V_F for ψ_F).

One way of constructing a wavefunction ψ on a 3-surface V is to start with a momentum $\bar{p}$ at a point $\bar{x}$ in V, and to use the world function $W(x, \bar{x})$ to define that part of the momentum field $p(x)$ which lies in the 3-surface V. Thus we construct a phase function

$$S(x) = -W(x, \bar{x})_l \bar{p}^l, \tag{4.12}$$

and put the wavefunction on the 3-surface equal to

$$\psi(x; \bar{p}) = \mathrm{e}^{\mathrm{i}S} = \exp[-\mathrm{i}W(x, \bar{x})_l \bar{p}^l]. \tag{4.13}$$

The motivation is that

$$\partial_k S(x) = -W_k(x, \bar{x})_l \bar{p}^l = g_k(x, \bar{x})_l \bar{p}^l + \mathrm{O}(\Delta^2), \tag{4.14}$$

so the gradient of $S(x)$ is nearly equal to the momentum $\bar{p}$ taken by parallel transport from $\bar{x}$ to x. Nearly, but not exactly, for

$$\partial_k S \partial^k S + m^2 = \mathrm{O}(\Delta^2), \tag{4.15}$$

so that $\partial_k S$ does not correspond to a momentum vector of constant mass. However, while $\psi = \mathrm{e}^{\mathrm{i}S}$ gives the value of the wavefunction lying on the 3-surface V, its derivative normal to the 3-surface can be assigned independently. So we take the part of the momentum lying in the 3-surface V to be

$$p^k_{\|}(x) = (g^{kl} + N^k N^l)\partial_l S, \tag{4.16}$$

where N^k is the unit normal to V, and the part of the momentum normal to V to be

$$p_{\perp}(x) = \pm(m^2 + g_{kl} p^k_{\|} p^l_{\|})^{1/2}. \tag{4.17}$$

The derivative of the wavefunction normal to V is taken to be

$$N^k \partial_k \psi(x, \bar{p}) = \mathrm{i}p_{\perp}(x)\psi(x, \bar{p}). \tag{4.18}$$

We have so far ignored the possibility that geodesics may start to cross, so that there is no longer a unique geodesic joining two points, and the world function is no longer well defined. We can avoid this trouble by forming a wave packet of limited spatial extent as an integral over the part of $\bar{p}$ lying in the 3-surface V:

$$\begin{aligned} \varphi(x) &= \int \psi(x; \bar{p}) f(\bar{p})\, \mathrm{d}^3 \bar{p}_{\|}, \\ N^k \partial_k \varphi(x) &= \int \mathrm{i}p_{\perp}(x)\psi(x; \bar{p}) f(\bar{p})\, \mathrm{d}^3 \bar{p}_{\|}. \end{aligned} \tag{4.19}$$

This is a perfectly well defined procedure for constructing wavefunctions, but it is not obvious what the physical interpretation should be. One might be tempted to think that $\psi(x;\bar{p})$ represents the state of a particle which, if it crossed V at x, would have momentum $p(x)$. But a little reflection will convince one that no measurement could verify—or falsify—such an interpretation.

In principle the way to interpret wavefunctions was suggested by Feynman in his 1948 paper: one uses a path integral to calculate the amplitudes of the results of well defined measurements. Thus one can measure the momentum of a charged particle by examining the curvature of its track in a magnetic field. To predict the results of such a measurement, one calculates a path integral with 'gates' which define the track.

To work out that idea in any detail would be a major research project in its own right.

5. Monte Carlo evaluation of path integrals

Each 3-surface V_{2n+1} in the path completely encloses the previous point x_{2n-1}. Each step $x_{2n+1} \leftarrow x_{2n-1}$ must be small, so that the kernel $K(x_{2n+1}, x_{2n-1})$ can be a good approximation; therefore each 3-surface V_{2n+1} must be small. This gives a relativistic path integral a very different structure from a non-relativistic path integral, where the 3-surfaces are infinite space-like hypersurfaces, and successive intervals may be arbitrarily large. The easiest way to visualize the structure of a relativistic path integral is to imagine the way in which it might be evaluated by a crude Monte Carlo method.

The basic idea of the Monte Carlo method is to replace an integral by a sample average. We generate a pseudo-random sequence of paths of the type

$$\{\text{path}\} = \begin{Bmatrix} x_1 \leftarrow x_3 \leftarrow \ldots \leftarrow x_{2N+1} \\ q_2 \leftarrow q_4 \leftarrow \ldots \leftarrow q_{2N} \end{Bmatrix}, \tag{5.1}$$

where as we shall see even the number of steps N is a random variate. The generating process picks points x_{2n+1} to lie in $\mathrm{d}V_{2n+1}$ with probability $P(\mathrm{d}V_{2n+1})$, and momenta q_{2n} to lie in $\mathrm{d}\Omega_{2n}$ with probability $P(\mathrm{d}\Omega_{2n})$. Each path is assigned a weight

$$W(\text{path}) = \frac{\mathrm{d}V_1}{P(\mathrm{d}V_1)} \frac{\mathrm{d}\Omega_2}{P(\mathrm{d}\Omega_2)} \frac{\mathrm{d}V_3}{P(\mathrm{d}V_3)} \frac{\mathrm{d}\Omega_4}{P(\mathrm{d}\Omega_4)} \cdots \frac{\mathrm{d}\Omega_{2N}}{P(\mathrm{d}\Omega_{2N})} \frac{\mathrm{d}V_{2N+1}}{P(\mathrm{d}V_{2N+1})}. \tag{5.2}$$

The transition amplitude is then estimated by the sample average

$$\hat{T}_{\mathrm{FI}} = N_{\text{path}}^{-1} \sum_{\text{paths}} \{W(\text{path})\tilde{\psi}_{\mathrm{F}} Q(\text{path})\psi_{\mathrm{I}} \exp[\mathrm{i}S(V_{\mathrm{F}} \leftarrow \text{path} \leftarrow V_{\mathrm{I}})]\}. \tag{5.3}$$

To generate the paths, we start with the two 3-surfaces V_{F} and V_{I} with V_{I} completely enclosing V_{F}, as in figure 1.

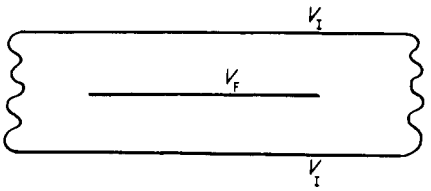

Figure 1.

We choose the first point x_1 on $\mathrm{d}V_1$ in $V_1 = V_{\mathrm{F}}$ with probability $P(dV_1)$. Then we

construct a small 3-surface V_3 completely enclosing x_1. Then we choose a momentum q_2 on $\mathrm{d}\Omega_2$ with probability $P(\mathrm{d}\Omega_2)$, and a point x_3 on $\mathrm{d}V_3$ in V_3 with probability $P(\mathrm{d}V_3)$. We arrive at the situation in figure 2.

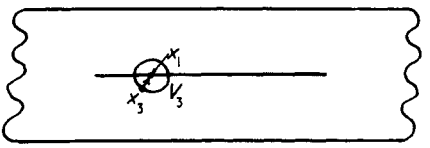

Figure 2.

At each step, we construct a small 3-surface V_{2n+1} enclosing the previous point x_{2n-1}. We choose a momentum q_{2n} on $\mathrm{d}\Omega_{2n}$ with probability $P(\mathrm{d}\Omega_{2n})$ and a point α_{2n+1} on $\mathrm{d}V_{2n+1}$ with probability $P(\mathrm{d}V_{2n+1})$. Then we ask whether x_{2n+1} lies outside the 3-surface V_1: if it does, the path is terminated, as in figure 3.

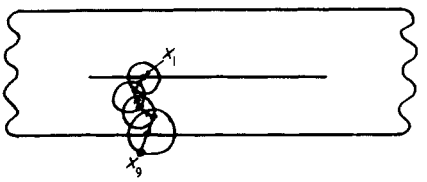

Figure 3.

We can take x_1 as the point at which the geodesic joining the two last points crosses the 3-surface V_1, and we can then have to find $P(\mathrm{d}V_1)/\mathrm{d}V_1$. Alternatively, we can identify x_1 with the last point selected, which amounts to a slight redefinition of V_1.

The crude Monte Carlo method we have described is purely illustrative. To make a reasonably efficient computational procedure, we would have to give great attention to methods of variance reduction, such as importance sampling, correlated paths, Russian roulette and splitting, etc.

6. Mutually gravitating particles

The path operator Q contains the quantum part of the dynamics. If there are no specifically quantum forces—spin dependent forces, for example—then the path operator for two particles is just the outer product of the two path operators for the individual particles:

$$Q(^{\text{path of}}_{\text{A and B}}) = Q(^{\text{path}}_{\text{of A}}) \otimes Q(^{\text{path}}_{\text{of B}}). \tag{6.1}$$

It is quite obvious that, since baryon interactions are strongly spin dependent, (6.1) would not be valid for a path integral of two baryons. For the present we are concerned with gravitational theory. There should be no quantum contributions to the gravitational interaction, so (6.1) should be valid for particles interacting primarily via their mutual gravitational field.

The action for two independent particles is just the sum of the actions for the two separate particles:

$$S(^{\text{path of}}_{\text{A and B}}) = S(^{\text{path}}_{\text{of A}}) + S(^{\text{path}}_{\text{of B}}) = \int_{\text{A}} q_a \lambda^a_{a\,\text{A}}(x)\, \mathrm{d}x^a_{\text{A}} + \int_{\text{B}} q_\mu \lambda^\mu_{k\,\text{B}}(x)\, \mathrm{d}x^k_{\text{B}}. \tag{6.2}$$

The tetrad at A will be altered by the gravitational field at A due to particle B, and vice versa. We want an expression for the change in the tetrad in terms of the particle paths. We can find such an expression using the formulation of Mach's principle pioneered by Lynden-Bell (1967), and elaborated by Sciama *et al* (1969). This form of Mach's principle expresses the metric tensor as an integral over the energy momentum tensor:

$$g_{ab}(x) = \int G_{ab}(x, y)_{kl} T^{kl}(y)\, \mathrm{d}C(y). \tag{6.3}$$

The bitensor $G_{ab}(x, y)_{kl}$ is the Green function of the differential equation

$$\Box\phi_{ab} - 2R_{a\cdot b}^{\cdot c\cdot d}\phi_{cd} = 0, \tag{6.4}$$

subject to the gauge condition

$$\phi^b_{a|b} - \tfrac{1}{2}\phi^b_{b|a} = 0. \tag{6.5}$$

The Green function depends implicitly on the metric, and so (6.3) is a non-linear integral equation. It is important to notice that a small change in the metric causes a change in the Green function which is of the second order of smallness, as Lynden-Bell has stressed. Thus it will be a good approximation to use in the Green function only the background metric which ignores the presence of A and B. The change in the energy momentum tensor due to B is

$$\delta T^k_{l}(y) = \mathscr{I}(y, x)\lambda^\mu_l(x) q_\mu \dot{x}^k, \tag{6.6}$$
$$\quad\ \mathrm{B} \qquad\quad \mathrm{B}\ \mathrm{B}\ \mathrm{B}$$

and so the change in the metric at A due to B is

$$\delta g_{ab}(x) = \int G_{ab}(x, x)^m_k \lambda^\mu_m(x) q_\mu \, \mathrm{d}x^k. \tag{6.7}$$
$$\mathrm{A}\quad \mathrm{A}\ \mathrm{B} \qquad\quad \mathrm{B}\ \mathrm{B} \qquad\quad \mathrm{B}$$

The change in the metric corresponding to a change $\delta\lambda^\alpha_a$ in the tetrad is

$$\delta g_{ab} = \eta_{\alpha\beta}(\delta\lambda^\alpha_a \lambda^\beta_b + \lambda^\alpha_a \delta\lambda^\beta_b). \tag{6.8}$$

If we impose the gauge condition

$$\eta_{\alpha\beta}(\delta\lambda^\alpha_a \lambda^\beta_b - \lambda^\alpha_a \delta\lambda^\beta_b) = 0 \tag{6.9}$$

corresponding to no rotation of the tetrad, then we obtain

$$\begin{aligned}
\delta\lambda^\alpha_a(x) &= \tfrac{1}{2}\lambda^\alpha_b \delta g^b_a(x) \\
&\quad\ \mathrm{A} \\
&= \tfrac{1}{2}\lambda^\alpha_b \int G^b_a(x, x)^m_k \lambda^\mu_m(x) q_\mu \, \mathrm{d}x^k \\
&\qquad\quad \mathrm{A}\ \mathrm{B} \qquad\quad \mathrm{B}\ \mathrm{B} \qquad\quad \mathrm{B} \\
&= \tfrac{1}{2}\int G^\alpha_a(x, x)^\mu_k q_\mu \, \mathrm{d}x. \\
&\qquad\quad \mathrm{A}\ \mathrm{B} \qquad\quad \mathrm{B}\ \mathrm{B} \qquad\quad \mathrm{B}
\end{aligned} \tag{6.10}$$

The change in the action integral (6.2) due to the change (6.10) in the tetrad at A caused by B is

$$\begin{aligned}
\delta S &= \int \delta\lambda^\alpha_a q_\alpha \, \mathrm{d}x^a \\
&\qquad\ \mathrm{A} \qquad\ \mathrm{A} \\
&= \tfrac{1}{2}\int\int q_\alpha \, \mathrm{d}x^a G^\alpha_a(x, x)^\mu_k q_\mu \, \mathrm{d}x^k. \\
&\qquad\ \mathrm{A} \qquad\ \mathrm{A} \qquad\quad \mathrm{A}\ \mathrm{B} \qquad\quad \mathrm{B}\ \mathrm{B} \qquad\quad \mathrm{B}
\end{aligned} \tag{6.11}$$

This integral is symmetric in A and B, and so gives also the change in action due to the change in the tetrad at B caused by A. The total action for A and B is therefore

$$S(^{\text{path of}}_{\text{A and B}}) = \int q_\alpha \lambda^\alpha_a(x) \, \mathrm{d}x^a + \int q_\mu \lambda^\mu_k(x) \, \mathrm{d}x^k + \tfrac{1}{2} \int\int q_\alpha \, \mathrm{d}x^a G^\alpha_a(x, x)^\mu_k q_\mu \, \mathrm{d}x^k. \tag{6.12}$$

The classical paths of mutually gravitating particles are found by requiring that the integral (6.12) be stationary subject to the constraints

$$\Omega(q) = 0, \qquad \Omega(q) = 0. \tag{6.13}$$

7. Strongly interacting particles: the breakdown of the space–time continuum?

It is legitimate to form the Green function from the background metric provided that the distances between the particles remain much larger than their gravitational radii Gm/c^2. In the Schwarzschild era particles approach to within their gravitational radii, and the Green function is strongly affected by their mutual gravitational interaction. In classical theory, this simply means that the Machian integral equation is implicitly non-linear. The quantum theory, however, meets severe difficulties.

Path integrals must be taken over all paths, both geodesic and non-geodesic. However, the geodesic nature of paths is a consequence of Einstein's field equations; when the paths are not geodesic, the field equations have no solution, there is no metric, no Green function, and no action integral.

Perhaps we could tinker with the field equations so as to force a solution, but I doubt whether this is the right approach. What is happening is that a single elementary particle causes such a large change in the geometry experienced by neighbouring particles that the background geometry no longer has any meaning. Then test particles can no longer make meaningful measurements of local geometry, and the space–time continuum becomes meaningless.

Acknowledgments

I began this work at the University of Washington, where I was greatly helped by discussions with Dr P C Peters. I am particularly grateful to Dr Peters for his suggestion that the use of the mechanical momentum eliminates position from the momentum surface.

This work emerged into a coherent whole while I was at Newcastle. I am grateful to the University of Newcastle-upon-Tyne for its hospitality and for the award of the Wilfred Hall Fellowship.

References

Audretsch J 1973 *Nuovo Cim.* **17B** 284–98
Brill D R and Gowdy R H 1970 *Rep. Prog. Phys.* **33** 413–88
Clutton-Brock M 1974 *Astrophys. Space Sci.* in the press
De Witt B S 1972 *Magic Without Magic* ed J R Klauder (San Francisco: Freeman)

Feynman R P 1948 *Rev. Mod. Phys.* **20** 367–87
—— 1949 *Phys. Rev.* **76** 749–59
Kuzmin V A 1970 *JETP Lett.* **12** 228–30
Lynden-Bell D 1967 *Mon. Not. R. Astron. Soc.* **135** 413–28
Penrose R 1971 *Quantum Theory and Beyond* ed T Bastin (Cambridge: Cambridge University Press)
Penrose R and MacCallum M A H 1973 *Phys. Rep.* **6** 241–315
Sciama D W, Waylen P and Gilman R 1969 *Phys. Rev.* **187** 1762–6
Synge J L 1960 *Relativity, The General Theory* (Amsterdam: North-Holland)