Relativistic Angular Momentum

1. User

consider the angular momentum of circular orbits with F=K/r^n in special relativity

2. Assistant

Take a particle of rest mass m in a central (inward) force field of magnitude

$$F(r) = \frac{K}{r^n}, \qquad K > 0,$$

moving on a circular orbit of radius r with constant speed v (so the acceleration is purely transverse to the velocity).

In special relativity the 3-momentum is $\mathbf{p} = \gamma m \mathbf{v}$, $\gamma = (1-\beta^2)^{-1/2}$, $\beta = v/c$. For uniform circular motion the acceleration is perpendicular to $\mathbf{v}$, so $d\gamma/dt = 0$ and

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} = \gamma m \mathbf{a}, \qquad a = \frac{v^2}{r}.$$

Hence the inward force balance for a circular orbit is

$$\frac{K}{r^n} = \gamma m \frac{v^2}{r}.$$

It is convenient to introduce the dimensionless parameter

$$\varepsilon \equiv \frac{K}{mc^2\, r^{\,n-1}}.$$

Then the force balance becomes

$$\varepsilon = \gamma \beta^2.$$

Using $\gamma = (1-\beta^2)^{-1/2}$, this can be solved explicitly. One clean way is to write $\beta^2 = \varepsilon/\gamma$, insert into γ, and solve the resulting quadratic; the physical (positive) solution is

$$\gamma = \frac{\varepsilon + \sqrt{\varepsilon^2 + 4}}{2}, \qquad \beta^2 = \frac{2\varepsilon}{\varepsilon + \sqrt{\varepsilon^2+4}}.$$

The (relativistic) orbital angular momentum magnitude is

$$L = |\mathbf{r} \times \mathbf{p}| = r\,(\gamma m v) = \gamma m v r = mcr\,\gamma\beta = mcr\,\sqrt{\gamma^2 - 1}.$$

Using $\varepsilon = \gamma\beta^2$ (equivalently $\gamma^2 - 1 = \varepsilon\gamma$), you can write it as

$$L^2 = Km\,r^{3-n}\,\gamma,$$

and therefore, in terms of r and the parameters,

$$\boxed{L(r) = \sqrt{Km}\,r^{\frac{3-n}{2}}\,\left[\frac{1}{2}\left(\varepsilon + \sqrt{\varepsilon^2 + 4}\right)\right]^{1/2}, \quad \varepsilon = \frac{K}{mc^2\,r^{n-1}}}$$

or, fully expanded,

$$L(r) = \sqrt{Km}\,r^{\frac{3-n}{2}}\left[\frac{1}{2}\left(\frac{K}{mc^2 r^{n-1}} + \sqrt{\frac{K^2}{m^2c^4r^{2n-2}} + 4}\right)\right]^{1/2}.$$

Two useful limits (as a consistency check) follow from $\varepsilon \ll 1$ and $\varepsilon \gg 1$.

If $\varepsilon \ll 1$ (weak field / large radius / Newtonian regime), $\gamma \simeq 1 + \varepsilon/2$, so

$$L \simeq \sqrt{Km}\,r^{\frac{3-n}{2}}\left(1 + \frac{\varepsilon}{4} + O(\varepsilon^2)\right),$$

recovering the Newtonian scaling $L_{\mathrm{Newt}} = \sqrt{Km}\,r^{(3-n)/2}$ at leading order.

If $\varepsilon \gg 1$ (ultra-relativistic), $\gamma \simeq \varepsilon$, giving

$$L^2 \simeq Kmr^{3-n}\,\varepsilon = \frac{K^2}{c^2}\,r^{4-2n}, \qquad \text{i.e.} \qquad L \simeq \frac{K}{c}\,r^{2-n},$$

so the mass drops out at leading order.

3. User

sldo consider the particular case n=2

4. Assistant

For $n = 2$ the central force is

$$F(r) = \frac{K}{r^2} \quad (K > 0),$$

and for a circular orbit of radius r and speed v, the acceleration is transverse to the velocity, so in special relativity

$$F = \frac{d\mathbf{p}}{dt} = \gamma m\,\frac{v^2}{r}, \qquad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

Force balance gives

$$\frac{K}{r^2} = \gamma m \frac{v^2}{r} \quad \Longrightarrow \quad K = \gamma m v^2 r.$$

The relativistic orbital angular momentum is

$$L = r\,(\gamma m v) = \gamma m v r.$$

Eliminating r between these two equations yields a special simplification valid only for $1/r^2$:

$$\frac{K}{L} = \frac{\gamma m v^2 r}{\gamma m v r} = v \quad \Longrightarrow \quad \boxed{v = \frac{K}{L}}.$$

So for $n = 2$, the speed of a circular orbit depends only on L, not on r. The relativistic constraint $v < c$ implies a strict lower bound on angular momentum:

$$\boxed{L > \frac{K}{c}}.$$

With $v = K/L$, the Lorentz factor becomes

$$\boxed{\gamma = \frac{1}{\sqrt{1 - \left(\frac{K}{Lc}\right)^2}}}.$$

Then the radius follows from $K = \gamma m v^2 r$ (or $L = \gamma m v r$):

$$r = \frac{K}{\gamma m v^2} = \frac{K}{\gamma m}\frac{L^2}{K^2} = \boxed{\frac{L^2}{mK}\sqrt{1 - \left(\frac{K}{Lc}\right)^2}} = \boxed{\frac{L}{mc}\sqrt{\left(\frac{Lc}{K}\right)^2 - 1}}.$$

Equivalently, if you want L as a function of r, introduce $\varepsilon = \frac{K}{mc^2 r}$. From the general circular-orbit relation $\varepsilon = \gamma\beta^2$ (with $\beta = v/c$) one finds

$$\gamma = \frac{\varepsilon + \sqrt{\varepsilon^2 + 4}}{2}, \qquad L^2 = Kmr\,\gamma,$$

so

$$\boxed{L(r) = \sqrt{Kmr}\left[\frac{1}{2}\left(\frac{K}{mc^2 r} + \sqrt{4 + \frac{K^2}{m^2c^4r^2}}\right)\right]^{1/2}}.$$

In the Newtonian limit $r \to \infty$ (so $v \ll c$), $\gamma \to 1$ and these reduce to $L \simeq \sqrt{Kmr}$ and $v \simeq \sqrt{K/(mr)}$. In the opposite limit $r \to 0$, one has $v \to c$ and $L \to K/c$ from above, so K/c is the asymptotic minimum angular momentum for circular motion under a $1/r^2$ force in special relativity.

5. User

and the particular case n=3

6. Assistant

For $n = 3$ the force magnitude is

$$F(r) = \frac{K}{r^3} \qquad (K > 0),$$

and for a circular orbit of radius r and speed v the relativistic transverse equation of motion gives

$$\frac{K}{r^3} = \gamma m \frac{v^2}{r} \quad\Longrightarrow\quad \frac{K}{r^2} = \gamma m v^2, \qquad \gamma = \frac{1}{\sqrt{1 - v^2/c^2}}.$$

The angular momentum is

$$L = r(\gamma m v) = \gamma m v r.$$

A useful simplification happens at $n = 3$. From $\frac{K}{r^2} = \gamma m v^2$ and $L = \gamma m v r$,

$$K = \gamma m v^2 r^2 = \frac{(\gamma m v r)^2}{\gamma m} = \frac{L^2}{\gamma m} \quad\Longrightarrow\quad \boxed{L^2 = Km\,\gamma}.$$

So the Lorentz factor is fixed by L:

$$\boxed{\gamma = \frac{L^2}{Km}}.$$

This immediately gives the speed as a function of L:

$$\beta^2 = 1 - \frac{1}{\gamma^2} = 1 - \left(\frac{Km}{L^2}\right)^2, \qquad \boxed{v(L) = c\sqrt{1 - \left(\frac{Km}{L^2}\right)^2}}.$$

Existence of a circular orbit requires $\gamma \geq 1$, hence

$$\boxed{L^2 \geq Km}.$$

At equality, $v = 0$ and the corresponding radius goes to infinity (see below), so any finite-radius circular orbit has $L > \sqrt{Km}$.

For the radius, use $r = L/(\gamma m v)$, or equivalently eliminate γ to get

$$r = \frac{K}{Lv},$$

so in terms of L,

$$\boxed{r(L) = \frac{K}{c\,L\sqrt{1 - \left(\frac{Km}{L^2}\right)^2}}}.$$

If you prefer L as a function of r, introduce

$$\varepsilon \equiv \frac{K}{mc^2r^2},$$

for which the circular-orbit condition is $\varepsilon = \gamma\beta^2$. Solving gives

$$\gamma = \frac{\varepsilon + \sqrt{\varepsilon^2 + 4}}{2},$$

and therefore

$$\boxed{L(r) = \sqrt{Km\,\gamma} = \sqrt{Km}\left[\frac{1}{2}\left(\frac{K}{mc^2r^2} + \sqrt{4 + \frac{K^2}{m^2c^4r^4}}\right)\right]^{1/2}}.$$

Checks in the two extremes: as $r \to \infty$, $\varepsilon \to 0$, $\gamma \to 1$, so $L(r) \to \sqrt{Km}$ (the Newtonian critical value). As $r \to 0$, $\varepsilon \to \infty$, $\gamma \sim \varepsilon$, and $L(r) \sim K/(cr)$, with $v \to c$.

7. User

confirm your results in the bibliography and now study the existence of non circular orbits and whan can we say of L on the general case and in the particulars n=2, 3

8. Assistant

I will treat this as the standard "special-relativistic particle in a time-independent central potential" model (flat spacetime, no radiation reaction, no GR), i.e. a particle of rest mass m moving under a central 3-force

$$\mathbf{F}(r) = -\frac{K}{r^n}\,\hat{\mathbf{r}} \qquad (K > 0),$$

coming from a scalar potential $V(r)$ with $F_r = -\frac{\mathrm{d}V}{\mathrm{d}r}$. For $n \neq 1$,

$$V(r) = -\frac{K}{n-1}\,r^{1-n} \qquad (\text{so } V \to 0 \text{ as } r \to \infty \text{ when } n > 1).$$

Web tool used: `web.run` (citations below).

1. What is always true (general, not necessarily circular)

Because the force is central, the (relativistic) angular momentum is conserved:

$$L = |\mathbf{r} \times \mathbf{p}| = \gamma m r^2 \dot{\varphi},$$

and the total energy is conserved:

$$E = \gamma m c^2 + V(r),$$

exactly as written in the circular-orbit paper's setup. arXiv

A convenient radial equation follows from

$$(E - V)^2 = m^2c^4 + c^2\Big(p_r^2 + \frac{L^2}{r^2}\Big),$$

so that

$$p_r^2 = \frac{(E-V)^2 - m^2c^4}{c^2} - \frac{L^2}{r^2},$$

and the allowed radii are those where $p_r^2 \geq 0$. Turning points satisfy $p_r = 0$. This is the cleanest way to discuss existence of non-circular orbits: for any fixed

(E, L), non-circular orbits exist whenever the inequality admits an interval of r and one chooses initial data with $p_r \neq 0$.

It is also useful to write the orbit equation with $u(\varphi) = 1/r$. Using $p_r = -L\, u'(\varphi)$, one gets the first integral

$$u'^2 + u^2 = \frac{(E - V(u))^2 - m^2c^4}{c^2L^2},$$

and differentiating gives a second-order equation valid for any central V:

$$u'' + u = -\frac{(E - V)}{c^2L^2}\,\frac{\mathrm{d}V}{\mathrm{d}u}.$$

For the power-law force ($n \neq 1$), i.e. $V(u) = -\frac{K}{n-1}u^{n-1}$, this becomes

$$u'' + u = \frac{KE}{c^2L^2}\,u^{n-2} + \frac{K^2}{(n-1)c^2L^2}\,u^{2n-3}.$$

Only $n = 2$ makes this linear; generically, power laws do not give closed-form conics.

1. Circular-orbit angular momentum: bibliographic confirmation (general n, and $n = 2, 3$)

For a circular orbit $r = r_0$ with speed v, the relativistic circular-motion condition is

$$F(r_0) = -\,m\gamma\,\frac{v^2}{r_0},$$

exactly as stated and used in “Circular Orbits under Central Forces in Special Relativity.” arXiv

Specialising that paper’s results to $F(r) = -K/r^n$, it explicitly classifies which angular momenta are possible for circular orbits, including:

- $n = 2$: there is a lower bound $L > K/c$. arXiv
- $n = 3$: there is a lower bound $L > \sqrt{Km}$. arXiv

It also gives the stability outcome: circular orbits are stable for $n \leq 2$, and unstable for $n \geq 3$ (with a mixed regime for $2 < n < 3$). arXiv

For $n = 2$ specifically, the existence of a limiting angular momentum (and the qualitative change in trajectories below it) is also emphasised in Boyer’s paper on relativistic Kepler/Coulomb trajectories. ar5iv+1

1. Existence of non-circular orbits, and what L does in general

There are two distinct questions about L:

(a) Conservation: In this central-force SR model, L is conserved for every orbit (circular or not). This is structural: $\dot{\mathbf{L}} = \mathbf{r} \times \mathbf{F} = 0$.

(b) Dynamical "barrier" role: Whether a nonzero L prevents reaching $r = 0$ depends on how singular the attractive potential is.

Use the radial inequality $p_r^2 \geq 0$ and examine $r \to 0$ for $n > 1$ (so $V \sim -\text{const}\, r^{1-n} \to -\infty$):

As $r \to 0$,

$$\frac{(E-V)^2}{c^2} \sim \frac{V^2}{c^2} \propto r^{\,2(1-n)} = r^{\,2-2n}, \qquad \frac{L^2}{r^2} \propto r^{-2}.$$

So:

• If $n < 2$: r^{2-2n} diverges more slowly than r^{-2}. The $-L^2/r^2$ term dominates and makes $p_r^2 < 0$ near 0. Thus any orbit with $L \neq 0$ is excluded from reaching the centre; only $L = 0$ can hit $r = 0$.

• If $n = 2$: both scale as r^{-2}, giving a genuine competition and a critical angular momentum (worked out explicitly below): $L_c = K/c$. This is exactly the special case highlighted in the literature. arXiv+1

• If $n > 2$: r^{2-2n} diverges faster than r^{-2}, so the potential-driven $(E-V)^2$ term dominates and $p_r^2 \to +\infty$. In this model, no matter how large L is, it does not kinematically forbid approaching $r = 0$; plunging ("collision") trajectories are always allowed by the inequality for suitable initial data.

This is the most robust general statement about "what can we say about L" beyond conservation: for $n \leq 2$, L can create a genuine centrifugal exclusion of the centre (always for $n < 2$, conditionally for $n = 2$); for $n > 2$, that exclusion disappears.

1. The special case $n = 2$ (Coulomb/Kepler force)

Here $V(r) = -K/r$ and the orbit equation becomes linear. Writing $u = 1/r$, one gets

$$u'' + \Big(1 - \frac{K^2}{L^2c^2}\Big)u = \frac{KE}{c^2L^2}.$$

Define

$$\alpha^2 := 1 - \frac{K^2}{L^2c^2}.$$

(i) Critical L and orbit types

- $L > K/c \iff \alpha^2 > 0$: $u(\varphi)$ is sinusoidal about a constant mean, so you get bound rosettes (for appropriate $E < mc^2$) and scattering trajectories (for $E \geq mc^2$), all with relativistic periapsis precession because the radial period in φ is $2\pi/\alpha \neq 2\pi$. A modern rigorous treatment states the corresponding rotation number as

$$\Theta = \sqrt{1 - \frac{K^2}{\ell^2 c^2}}$$

(where ℓ is the conserved angular momentum), and discusses periodic versus quasi-periodic behaviour depending on commensurability. ecuadif.uniovi.es

- $L = K/c \iff \alpha = 0$: the equation degenerates and u grows quadratically with φ, signalling marginal plunging behaviour.

- $L < K/c \iff \alpha^2 < 0$: $u(\varphi)$ grows/decays exponentially in φ, giving spirals that reach the force centre (collision) or spiral out, depending on time direction. This qualitative "spiral into the centre while conserving energy and angular momentum" behaviour is precisely the point of Boyer (and historically Darwin). ar5iv+1 A recent analysis even provides asymptotics near collision, including a logarithmic divergence in the polar angle (a spiral) as $t \to 0$. ecuadif.uniovi.es

(ii) Circular-orbit check inside $n = 2$

From the circular condition $K/r_0^2 = m\gamma v^2/r_0$ and $L = \gamma m r_0 v$, one finds $v = K/L$, hence $v < c \Rightarrow L > K/c$, matching the circular-orbit literature exactly. arXiv+1

1. The special case $n = 3$ (force K/r^3, potential $\propto -1/r^2$)

Here $V(r) = -\frac{K}{2r^2}$, i.e. $V(u) = -\frac{K}{2}u^2$, and the orbit equation becomes a nonlinear Duffing-type equation:

$$u'' + \Big(1 - \frac{KE}{c^2 L^2}\Big)u = \frac{K^2}{2c^2 L^2}\, u^3.$$

Two immediate consequences follow.

First, circular orbits require a lower bound on angular momentum and are unstable. The circular-orbit paper states $L > \sqrt{Km}$ for $n = 3$, and that for $n \geq 3$ circular orbits are unstable. arXiv So the only non-plunging motion at fixed radius is a measure-zero tuning, and any perturbation generically drives the motion away (either inward or outward).

Second, unlike $n = 2$, angular momentum does not protect the origin for noncircular motion. From the general small-r scaling, $n = 3$ lies in the $n > 2$ regime where $(E - V)^2$ dominates the $-L^2/r^2$ term, so collision trajectories remain allowed for any finite $L \neq 0$. In particular, there are no "$L > L_c$" noncollision families analogous to Coulomb/Kepler; the qualitative role of L is mainly (i) selecting whether a circular orbit exists at all (needs $L > \sqrt{Km}$) and (ii) fixing the timescale/geometry of the plunge or escape, rather than creating a strict centrifugal exclusion of $r = 0$.

Bibliography used/confirmed via `web.run`: the circular-orbit classification and stability results are in "Circular Orbits under Central Forces in Special Relativity" (arXiv:physics/0407049). arXiv The existence of qualitatively new noncircular orbits (spirals/plunges) and the special role of the Coulomb/Kepler case are discussed by Boyer (arXiv:physics/0405090; Am. J. Phys. 72 (2004)). ar5iv+1 Modern dynamical/integrability statements for the relativistic Coulomb problem, including $\Theta = \sqrt{1 - K^2/(\ell^2 c^2)}$ and collision asymptotics, appear in Ortega & Rojas (2025 postprint). ecuadif.uniovi.es

9. User

now I want to address the gravity case, first we must settle if general relativity always approaches to K=GmM for any number of dimensions

10. Assistant

In ordinary $3 + 1$ general relativity, the weak-field/slow-motion ("Newtonian") limit gives

$$F(r) = \frac{G_4\, m\, M}{r^2},$$

so in your notation $F = K/r^n$ one has $n = 2$ and $K = G_4 mM$. This is the standard matching of GR to Newtonian gravity (Einstein equation ⇒ Poisson equation for the Newtonian potential, and geodesics ⇒ Newton's second law in the limit). phys.au.dk

In D-dimensional GR (meaning D spacetime dimensions, so $D - 1$ spatial), the Newtonian limit still exists, but two things change:

1. The radial falloff is different: Gauss's law/Poisson theory in $(D-1)$ spatial dimensions gives

$$\Phi(r) \propto -\frac{1}{r^{D-3}} \quad (D \geq 4), \qquad \Rightarrow \qquad F(r) \propto \frac{1}{r^{D-2}}.$$

1. The numerical coefficient depends on conventions for "G". A clean way to see the coefficient is to match the weak-field limit of the Schwarzschild–Tangherlini solution. In standard coordinates one has $g_{tt} = -(1-\mu/r^{D-3})$ and the mass parameter is related to the (D-dimensional) gravitational constant and the mass by

$$\mu = \frac{16\pi G_D\, M}{(D-2)\,\Omega_{D-2}\, c^2}, \qquad \Omega_d = \frac{2\pi^{(d+1)/2}}{\Gamma((d+1)/2)}.$$

arXiv

Using $g_{tt} \simeq -(1 + 2\Phi/c^2)$, this implies

$$\Phi(r) = -\frac{c^2\,\mu}{2\,r^{D-3}} = -\frac{8\pi G_D\,M}{(D-2)\Omega_{D-2}}\frac{1}{r^{D-3}},$$

and therefore the force on a test mass m is

$$F(r) = m\,|\nabla\Phi| = \frac{8\pi(D-3)}{(D-2)\Omega_{D-2}}\,\frac{G_D\,m\,M}{r^{D-2}}.$$

So, in the form $F = K/r^n$ (for $D \geq 4$):

$$n = D-2, \qquad K = \frac{8\pi(D-3)}{(D-2)\Omega_{D-2}}\,G_D\,m\,M.$$

This reduces to the familiar $K = G_4 mM$ in $D = 4$ because $\Omega_2 = 4\pi$, giving the prefactor $= 1$. arXiv It also makes explicit that the dependence on the product mM persists (linearity in M for the source and proportionality to m when you translate acceleration into "force"), but the identification "$K = GmM$" is only literally true in all dimensions if you define your dimension-dependent Newton constant by

$$F(r) \equiv \frac{G_{\mathrm{Newt},D}\,mM}{r^{D-2}}, \quad \text{so that} \quad G_{\mathrm{Newt},D} = \frac{8\pi(D-3)}{(D-2)\Omega_{D-2}}\,G_D.$$

Finally, the exceptional case is $D = 3$ (two spatial dimensions): the Green's function of the 2D Poisson operator is logarithmic, so $\Phi(r) \sim \ln r$ and $F(r) \sim 1/r$ rather than a power $1/r^{D-2}$ with $D-2 = 1$ coming from a $1/r^{D-3}$ potential. math.libretexts.org+1

11. User

do the results about a maximum angular momentum for bounded orbits extend to general relativity? summarize them. Also touch the pecular case of classical approxmations of gauge theories

12. Assistant

In 4D general relativity (test-particle motion in a stationary, spherically symmetric vacuum geometry), the role of angular momentum is structurally very similar to the special-relativistic "central potential" problem: you still have two constants of motion (energy and angular momentum), motion reduces to a 1D effective-potential problem, and "maximal angular momentum at fixed energy" still corresponds to the circular-orbit limit. What changes in GR is that the effective potential typically develops both a minimum and a maximum, which introduces a separatrix and an additional, genuinely relativistic lower threshold in angular momentum for long-lived bound motion.

For Schwarzschild, restricting to the equatorial plane, the conserved specific energy E and specific angular momentum $\ell = L/m$ (per unit rest mass) lead to a radial equation that can be written as an "energy balance" with an effective potential. Carroll's derivation exhibits the key new term responsible for strong-field behaviour: in GR the effective potential differs from the Newtonian one by an extra attractive contribution $\propto -GM\,\ell^2/r^3$, and the qualitative orbit types are read off by comparing E^2 to that potential. ned.ipac.caltech.edu

Now fix E (below the rest-energy-at-infinity value, so $E < 1$ in $c = G = 1$ units). As you increase ℓ, the centrifugal part raises the potential and shrinks the allowed radial interval $[r_{\text{min}}, r_{\text{max}}]$. The largest ℓ compatible with a bound orbit occurs when the two turning points coincide: the radial oscillation amplitude goes to zero and the orbit becomes circular. This is the direct GR analogue of the Newtonian (and special-relativistic) statement "for fixed energy, the circular orbit has maximal angular momentum". There is no single global upper bound on ℓ across all bound orbits, because as $E \to 1^-$ the bound orbit becomes very large and ℓ can grow without bound (just as in Newtonian gravity).

What GR adds is a lower threshold and a separatrix. For sufficiently large ℓ, Schwarzschild admits two circular orbits at the same ℓ: an outer stable one (a minimum of the effective potential) and an inner unstable one (a maximum). As ℓ is decreased, these two extrema merge and disappear at the innermost stable circular orbit (ISCO): Schwarzschild has no stable circular orbit below $r = 6GM/c^2$, and the effective barrier disappears entirely when $\ell < \sqrt{12}\,GM/c$ (equivalently $L < \sqrt{12}\,GMm/c$). ned.ipac.caltech.edu This is conceptually close to the special-relativistic Coulomb/Kepler phenomenon that circular orbits require $L > k/c$ for an inverse-square force, except that in GR the threshold is set by the spacetime geometry rather than the definition of relativistic angular momentum in a fixed potential. arXiv

For non-circular bound orbits in Schwarzschild, the clean way to describe the boundary between "bound" and "plunge" is via the separatrix in the (p, e) (semi-latus rectum, eccentricity) parametrisation: the separatrix occurs at

$$p_\star = 6 + 2e_\star$$

(in units where $M = 1$), and approaching it the orbit develops the familiar zoom–whirl behaviour before transitioning to plunge. arXiv In terms of the (E, ℓ) picture, this separatrix corresponds to the condition that E^2 just reaches the maximum of the effective potential (the unstable circular orbit); below that ℓ (at fixed E) there is no inner turning point and the particle falls in.

So, in Schwarzschild GR at fixed $E < 1$, bounded motion typically exists only for ℓ in an interval

$$\ell_{\min}(E) \;<\; \ell \;\le\; \ell_{\max}(E),$$

where $\ell_{\max}(E)$ is the circular-orbit (zero-eccentricity) limit (turning points coincide at the potential minimum), while $\ell_{\min}(E)$ is set by the unstable circular orbit / separatrix (turning point degeneracy at the potential maximum). The existence of two circular orbits for the same ℓ (stable outer, unstable inner) and the absolute ISCO bound are the most important departures from the "single circular orbit" picture. ned.ipac.caltech.edu

Two important qualifiers.

First, in higher-dimensional GR the extension can fail in the strongest possible way: for a single asymptotically flat higher-dimensional black hole (Schwarzschild–Tangherlini, Myers–Perry), it is "widely known" that in 5D there are no stable circular orbits, and in dimensions > 5 (for those backgrounds) stable circular orbits do not exist at all. arXiv In such cases the "maximum ℓ at fixed E" statement becomes largely moot because long-lived bound geodesic motion is absent (except in special multi-centre geometries or with additional structure).

Second, if you go beyond the conservative geodesic approximation, radiation breaks exact constancy of E and L. In electromagnetism, accelerated charges radiate (Larmor's formula), so classical "bound orbits in a fixed Coulomb potential" are not truly periodic: energy decreases and the orbit inspirals. pulsar.sternwarte.uni-erlangen.de The gravitational analogue (gravitational-wave emission) similarly causes slow inspiral of bound binaries in GR. In practice one often separates the problem into a conservative "geodesic/effective-potential" sector (where E, L are constants and the above bounds/separatrices apply) plus a dissipative sector that drives adiabatic evolution across those families.

On the "classical approximations of gauge theories" point: the conservative central-potential picture is itself a classical approximation to a gauge field theory. In QED, the Coulomb potential is the static-field limit, but once you include the dynamical gauge field you generically get radiation and therefore secular evolution of E and L. pulsar.sternwarte.uni-erlangen.de In non-Abelian Yang–Mills, the long-distance physics is not Coulombic at all: Tong emphasises that if you naïvely expect a $1/r$ potential you "would have been wrong", and in the confining phase the potential energy between a quark–antiquark pair grows linearly with separation, $V(r) \propto r$, with a string tension scale. damtp.cam.ac.uk This is why phenomenological "classical potential" models used for heavy quarkonia often take the Cornell form "Coulomb + linear", $V(r) = -\frac{4}{3}\,\alpha_s/r + \sigma r + \text{const}$, which is explicitly an effective (and partly quantum-informed) potential rather than a literal classical Yang–Mills field configuration. Wikipedia+1

Tools used: web.run (internet search; opening and reading web pages and PDFs for the cited GR, SR, and gauge-theory references).

13. User

how do we see all these results from the point of view of the action?

14. Assistant

From the action point of view, essentially everything you've been discussing (existence of circular/non-circular orbits, and upper/lower bounds on angular momentum for bounded motion at fixed energy) is a single mechanism: symmetry reduction + a 1D "turning-point" problem. The references below were located with the web tool; I cite them where they support the key structural steps.

Start with an action $S = \int L(q, \dot{q})\, d\lambda$. If the system is invariant under time translations then the corresponding Noether charge (energy) is conserved; if it is invariant under spatial rotations then the corresponding Noether charge (angular momentum) is conserved. In practice, for a central problem one chooses planar motion and polar coordinates (r, ϕ), and ϕ is cyclic, so

$$p_\phi \equiv \frac{\partial L}{\partial \dot{\phi}} = L = \text{const.}$$

Eliminating $\dot{\phi}$ in favour of the constant L reduces the dynamics to a one-degree-of-freedom radial problem. In the nonrelativistic setting this is often introduced explicitly as an "effective potential" reduction, with allowed motion determined by $U_{\text{eff}}(r) < E$, and circular orbits determined by $dU_{\text{eff}}/dr = 0$. users.physics.ox.ac.uk

In special relativity, the cleanest "action-first" entry point is to use a reparameterisation-invariant worldline action and then pick the parameter to be coordinate time. Tong writes the free relativistic particle action as the proper-time extremisation $S \propto \int ds$, and shows that choosing $\lambda = t$ gives the standard square-root Lagrangian (with the familiar $\sqrt{1 - v^2/c^2}$ factor). damtp.cam.ac.uk He then shows how adding the gauge-coupling term $q \int A_\mu dx^\mu$ produces, in the t-parametrised Lagrangian, an ordinary "potential term" (from A_0) plus velocity couplings (from $\mathbf{A}$). damtp.cam.ac.uk For a static central potential model one typically takes $\mathbf{A} = 0$ and $A_0 = \phi(r)$, i.e. an action-level origin for $U(r) = q\phi(r)$.

Once you have a time-independent central $U(r)$, the Hamiltonian form makes the reduction completely transparent:

$$H = \sqrt{m^2 c^4 + c^2 (p_r^2 + \tfrac{L^2}{r^2})} + U(r) = E,$$

equivalently

$$(E-U)^2 = m^2c^4 + c^2\Big(p_r^2 + \frac{L^2}{r^2}\Big).$$

This is already a 1D turning-point problem: the turning radii are the roots of $p_r^2(r) = 0$, and bounded non-circular motion exists iff there are two distinct positive roots $r_{\min} < r_{\max}$.

Now the "maximum angular momentum for bounded motion at fixed energy" is just the statement that, as you vary L with E held fixed, the two turning points can merge. The boundary of existence of (a family of) bounded orbits occurs exactly when the turning-point equation has a double root (a tangency), i.e.

$$p_r^2(r_\star) = 0, \qquad \frac{d}{dr}p_r^2(r)\Big|_{r_\star} = 0.$$

This double-root condition is precisely the circular-orbit condition in the reduced problem: it is the same "$dU_{\text{eff}}/dr = 0$" criterion, just written in a relativistic Hamiltonian form. users.physics.ox.ac.uk Stable vs unstable circular orbits correspond to whether the double root sits at a local minimum or maximum of the effective radial potential.

Applying this to your power-law force $F = K/r^n$ is straightforward at the action level: you choose $U(r)$ with $F = -U'(r)$, i.e. $U(r) = -\frac{K}{n-1}r^{1-n}$ for $n \neq 1$. Then the double-root conditions above determine the circular orbit curve (E, L) and therefore the extremal L values that bound non-circular bounded motion at fixed E. Two special features worth isolating are exactly the ones you asked for earlier:

- $n = 2$ (Coulomb/Kepler): $U \propto -1/r$ introduces a length scale once E is fixed, so the tangency picture gives a genuine upper bound $L \leq L_{\max}(E)$, saturated by the (stable) circular orbit; smaller L gives more eccentric bounded motion down to the separatrix.
- $n = 3$ ($U \propto -1/r^2$): the potential has the same radial scaling as the centrifugal term, so the reduced radial problem becomes scale-invariant in r. In practice this makes bounded non-circular motion non-generic: you do not get the "two turning points for an interval of L" structure unless the parameters are tuned (the same structural reason why $1/r^2$ is a classical borderline case).

In general relativity, the same logic is even more literally "action $\to$ constants $\to$ 1D reduction". The test-particle action is again the worldline length, but now in the curved metric: Tong's development makes explicit that the gravitational coupling is encoded by promoting the Minkowski metric in the action to a curved $g_{\mu\nu}(x)$. damtp.cam.ac.uk In a stationary, spherically symmetric spacetime (Schwarzschild), time translations and axial rotations correspond to Killing symmetries; the associated Noether charges are the conserved energy E and angular momentum L. Carroll's treatment phrases orbital dynamics in terms of an effective potential and notes that circular orbits occur when $dV_{\text{eff}}/dr = 0$, with stability determined by whether that extremum is a minimum or maximum. Physics LibreTexts

The GR extension of your "max/min L for bounded orbits" is: for fixed energy E, the boundaries of allowed bound motion again occur when the radial turning points merge (double root), i.e. when the orbit becomes circular (stable boundary) or sits on the unstable circular separatrix (plunge boundary). The Schwarzschild case has an extra qualitative feature absent in Newtonian mechanics: stable circular orbits cease to exist below an innermost stable circular orbit (ISCO). Carroll derives that in Schwarzschild the stable and unstable circular orbits merge when the angular momentum hits a critical value, and that the smallest stable circular orbit radius is $r = 6GM$ (in $c = 1$ units), with the associated critical angular momentum threshold $L = \sqrt{12}\,GM$ (again in $c = 1$ units). Physics LibreTexts In other words, GR introduces an absolute minimum angular momentum for the existence of any stable bounded orbit around a Schwarzschild black hole, and for each fixed E one typically gets an allowed interval in L bounded by (i) a stable circular tangency and (ii) an unstable circular separatrix, until those merge at the ISCO.

For null geodesics, the same action-reduction logic yields an impact-parameter criterion. A standard effective-potential treatment shows that the photon radial equation can be written as a "kinetic + potential" decomposition in terms of the impact parameter $b = L/E$, and that the unstable circular photon orbit sits at the maximum of the effective potential. Physics LibreTexts In Schwarzschild this gives the photon sphere at $r = 3M$ and a critical impact parameter $b = \sqrt{27}\,M$ separating scattering from capture. Physics LibreTexts This is exactly the null analogue of the double-root/tangency condition.

Finally, for "classical approximations of gauge theories", the action viewpoint clarifies what is—and is not—legitimate. The fundamental object is not a mechanical potential but the gauge-field action plus source/worldline couplings. A static inter-source "potential" emerges as an effective description after integrating out gauge fields (or, on the lattice/in Euclidean formulation, by looking at large-time behaviour of gauge-invariant correlators). A canonical statement is that the expectation value of a rectangular Wilson loop of temporal extent t and spatial separation r behaves at large t as $\langle W \rangle \propto e^{-tV(r)}$, thereby defining the static potential $V(r)$. TDLI-Indico (Indico) The same notes record the common phenomenological parameterisation $V(r) = A + B/r + \sigma r$: the B/r term corresponds to the Coulombic (short-distance / weak-coupling, "classical-like") behaviour, while the σr term reflects confinement (nonperturbative area-law physics), which is not obtainable from a naive classical field approximation of Yang–Mills in the same way Coulomb is from Maxwell. TDLI-Indico (Indico) In this sense, using a power-law K/r^n as a "gauge theory potential" is best understood as working with a particular effective worldline action meant to approximate some regime of the full gauge-field functional integral, not as a direct consequence of the microscopic classical Yang–Mills equations.

All of the “angular-momentum bounds” you’ve been discussing can therefore be read as discriminant statements about the reduced radial equation obtained after symmetry reduction of the action: bounded motion exists when there are two turning points; extremal values of L at fixed E occur when those turning points merge (double root), i.e. when the reduced action admits a circular solution; and GR’s distinctive input is that, in Schwarzschild, the stable/unstable circular extrema themselves annihilate at the ISCO, producing an absolute minimum L for stable bounded orbits. Physics LibreTexts

15. User

to complete the review we sould say something more about the long range term of the potential in generic Gauge Theories for D different of 4

16. Assistant

Let D be the spacetime dimension and $n = D - 1$ the number of spatial dimensions. For static external charges, the “long-range term” in the potential is most cleanly defined from the Wilson loop: for a large rectangle of spatial size r and Euclidean time extent T,

$$\langle W(\text{rect})\rangle \sim e^{-V(r)\,T}\,,$$

so $V(r)$ is what the gauge theory assigns to a static charge–anticharge pair. damtp.cam.ac.uk

In a Coulomb phase with a massless gauge boson, integrating out the non-dynamical A_0 produces an instantaneous kernel $j^0\,(\nabla^2)^{-1}\,j^0$. The spatial dependence is therefore the fundamental solution of the Laplacian in n dimensions:

$$\nabla^2\Phi_n(\mathbf{x}) = -\delta^{(n)}(\mathbf{x}), \qquad \Phi_n(r) \propto \begin{cases} -|r|, & n = 1,\\ -\ln r, & n = 2,\\ \;\; r^{2-n}, & n > 2, \end{cases}$$

so the Coulombic long-range potential behaves as

$$V(r) \propto g^2\,C\,\Phi_n(r) \;\sim\; \begin{cases} g^2\,r, & D = 2,\\ g^2\,\ln r, & D = 3,\\ g^2/r^{D-3}, & D \geq 4, \end{cases} \qquad \text{(Coulomb phase)}.$$

(The group/representation factor C is a quadratic Casimir-type coefficient, as in the standard non-Abelian Coulomb force computation. damtp.cam.ac.uk+1)

The Laplacian kernels themselves are the standard Green functions quoted above. mccuan.math.gatech.edu

Two special cases you asked for are particularly transparent in the literature. In $D = 3$ ($2 + 1$ dimensions), classical Maxwell theory gives a logarithmically growing potential ("log confinement"): $V(r) = \frac{Q^2}{2\pi} \log(r/r_0) + \text{const.}$ damtp.cam.ac.uk In $D = 2$ ($1 + 1$ dimensions), Gauss' law forces a constant electric field between separated opposite charges, and the field energy grows linearly with separation, $E = \frac{e^2}{2} L$, i.e. a linear potential. damtp.cam.ac.uk These are exactly the $n = 2$ (log) and $n = 1$ (linear) Green-function cases.

For generic (especially non-Abelian) gauge theories, the true asymptotic long-range term depends on the phase. A useful minimal trichotomy is: Coulomb $V(r) \sim 1/r^{D-3}$ (or $\log r$, or r in low D); confining $V(r) \sim \sigma r$ (flux tube/string tension); Higgs/screened $V(r) \to$ const at large r. damtp.cam.ac.uk+1 With dynamical fundamental matter, even a confining flux tube can break by pair creation, so the large-r potential ultimately crosses over to a constant (string breaking). damtp.cam.ac.uk

"Classical approximation" issues become sharper when $D \neq 4$ because the gauge coupling is typically dimensionful in the usual normalisation where it multiplies interaction vertices. From canonical dimensions in d spacetime dimensions, $[A_\mu] \sim E^{(d-2)/2}$ and $[\psi] \sim E^{(d-1)/2}$, so the vertex $g\,\bar{\psi}\gamma^\mu A_\mu \psi$ forces $[g] \sim E^{(4-d)/2}$. Department of Physics Thus $d < 4$ theories are super-renormalisable and tend to become strongly coupled in the IR on dimensional grounds (making purely "Coulombic" classical reasoning unreliable at large distances unless you are demonstrably in a Coulomb phase), while $d > 4$ gauge theories are non-renormalisable EFTs, where the long-distance behaviour can remain Coulombic but the UV completion controls what "generic" ultimately means. Department of Physics+1

From the action viewpoint, all of these long-range statements are different ways of reading the same object: the effective action for conserved sources after integrating out gauge fields (giving the Green-function kernel), or equivalently the area/perimeter law behaviour of Wilson loops (giving linear vs Coulomb vs screened asymptotics for $V(r)$). damtp.cam.ac.uk+2damtp.cam.ac.uk+2

17. User

We want to address in the future the existence of a maximal angular momentum for columb and how it relates to the existence of a planck length in relativity, and how both provide interesting pivots to sustitute Planck constant, either adimensionaly or with a geometrical dimension. First of all, we want to understand quantisation, what is a different beast.

Read and transcribe the contents of https://a.rivero.nom.es/research/9803035.pdf and aim for a complete formalisation of the paper, first for the static case, then with quantum mechanics (action in a time line), then for field theory on D=4, and in each case the interpretation of the "halved expresion" with links to the literature.

You could use as fundamentals the known results on path integral but also the formalism ot the tangent groupoid. I had some sparse notes in https://a.rivero.nom.es/research/ode.pd and https://a.rivero.nom.es/research/9710026.pdf

It could be useful if for the formalisation and proof of the static case you provide code for lean to be input into some helper. In any case, work all the details and expound the results with wide detail.e

18. Assistant

Tools used: (1) `file_search` (reading the three PDFs you uploaded: 9803035, ode, 9710026); (2) `web.run` (to retrieve and cite the background literature you requested: Dirac 1933; Feynman 1948; Connes tangent groupoid; Carinena et al; Wilson–Kogut; Schwinger–Keldysh; delta function/oscillatory integrals; Born rule).

Transcription of "A short derivation of Feynman formula" (quant-ph/9803035)

arXiv:quant-ph/9803035 14 Mar 1998

A short derivation of Feynman formula

Alejandro Rivero*

Dep. Física Teórica, Universidad de Zaragoza, 50009 Zaragoza, Spain

February 22, 2002

Abstract

The complex exponential weighting of Feynman formalism is seen to happen at the classical level. (Finiteness of) Feynman path integral formula is suspected then to appear as a consistency condition for the existence of certain Dirac measures over functional spaces.

1. The simplest variational problems can be easily formulated in terms of distribution theory. For instance, recall the static problem, to find the minimum of a function and evaluate some quantity in that minimum. We can reformulate it as:

Given a function $f(x)$, find a Dirac measure δ_f concentrated in the critical points of f.

The answer is obviously $\delta(f'(x))$. Its exponential form,

$$\langle \delta(f'(x)) \mid g(x) \rangle = \iint e^{izf'(x)} g(x)\, dz\, dx = \iint \lim_{y\to x} e^{iz\frac{f(y)-f(x)}{y-x}} g(x)\, dz\, dx \quad (1)$$

is interesting for itself, but we can carry it to a more amazing shape by making the substitution $\varepsilon = (y-x)/z$. Then the previous limit is asymptotically equivalent to

$$\langle \delta_f \mid g \rangle = \iint \lim_{\varepsilon \to 0} e^{\frac{i}{\varepsilon}(f(y)-f(x))} g(x)\, dx\, dy\, \frac{1}{\varepsilon}. \tag{2}$$

This is done controlling $x-y$ and the oscillating character of the exponential.

Now, if f has only an extremal point, we can choose to work with the "halved" expression,

$$\langle \delta_f^{1/2} \mid O \rangle = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon^{1/2}} \int e^{\frac{i}{\varepsilon} f(x)} O(x)\, dx \tag{3}$$

from which we can recover (2) by taking modulus square,

$$\langle \delta_f \mid g \rangle = \langle \delta_f^{1/2} \mid O \rangle \, \langle \delta_f^{1/2} \mid O \rangle^*, \tag{4}$$

with $g(x) = O(x)O^*(x)$. Of course, other games are possible, changing the regularization method — i.e. the role of ε — but all of them are equivalent in the vicinity of $\varepsilon = 0$.

The whole point here is that the basic structure of path integral, the complex exponential weighting, is already present in the integral presentation of the most elementary variational problem, the principle of virtual work, which we teach in general physics courses.

It is intriguing to notice that Dirac guessed the exponential weighting directly from quantum mechanics, trying to build contact transformations [3], instead of uplifting it from its classical version, as we are doing here.

*rivero@sol.unizar.es

1. To go from the zero-dimensional problem (static) to the one dimensional (classical dynamics) we need to generalize (3) to spaces of functions of one variable, time. There we can not directly assert the convergence of the regularization, and we need to follow an indirect route, inspired in Wilson-Kogut triangles [8, ch. 12]. Feynman formula will appear as a convergence condition: the regularizated measure has a limit if and only if the Feynman measure over paths has a finite one.

First lets restate the question: We are given a functional $L[\phi]$ and associated contour conditions $(\phi_0, t_0), (\phi_1, t_1)$ determining a space F of functions. The problem, again, is to find a Dirac measure over this space F concentrated in the critical points of the functional L. Inspired in (3), we propose as answer the limit of the discretized functional:

$$\langle \delta^L_{\varepsilon,\varepsilon_0} \mid O[\phi] \rangle = \int \cdots \int \frac{1}{\varepsilon^{n/2}} \exp\left(\frac{i}{\varepsilon}\, \varepsilon_0 \sum_{i=0}^{n} L_{\varepsilon_0}\left[x_i, \frac{x_{i+1}-x_i}{\varepsilon_0} \right] \right) O[\phi] \,\prod dx_i. \tag{5}$$

Where $\varepsilon_0 = (t_1 - t_0)/(n+1)$, $x_0 = \phi_0$, $x_{n+1} = \phi_1$, and we must take both limits $\varepsilon, \varepsilon_0 \to 0$. Each integration in dx_i is a mirror of formula (3), concentrating in the values of x_i where the function $L[x_1, \dots, x_n]$ takes its extremal value keeping the rest of $\{x_j\}$ fixed.

At this point, we could directly define a proportionality constant between ε and ε_0, say $\varepsilon = h\varepsilon_0$, to join both limits and then claim (8) directly.

1. But better we would like to try a more sophisticated process, and ask ourselves about the convergence of the $\varepsilon, \varepsilon_0 \to 0$ limit. To do it, we introduce an arbitrary scale h which controls the limit process: each term would be moved back according a transformation

$$\varepsilon \to \tau\, \varepsilon_h(\varepsilon), \qquad \varepsilon_0 \to \tau\, \varepsilon_h(\varepsilon_0), \tag{6}$$

whose exact form still elude us. Surely we must to take in account that, in addition to $\varepsilon, \varepsilon_0$, also the quantities $\xi_i \equiv x_{i+1} - x_i$ go to zero. In fact, Feynman proof of Schrödinger equation [4] relies heavily in an approximation for small ξ. In our case, it is implied an adjustment in ξ which we can hide under the carpet of the Lagrangian.

On the other side, transformation (6) reduces the number n of points where we fix the classical path. So, alternatively, we could try to build $\varepsilon \to 0$ as a discrete series of bipartitions $\varepsilon_{n+1} = \varepsilon_n/2$, and then the control as a block summation back to the expression of level $\varepsilon_0 = h$.

1. In any case, the limit would be convergent if and only if the controlled series

$$\langle \delta^L_{h,\varepsilon_0} \mid O[\phi] \rangle = \int \cdots \int \frac{1}{(h\varepsilon_0)^{n/2}} \exp\left(\frac{i}{h} L^n_{\varepsilon_0}[\phi_0, x_1, \dots, x_n, \phi_1] \right) O[\phi] \,\prod dx_i \tag{7}$$

converges (compare with the control of a Wiener measure through the normalized Brownian bridge, see e.g. [7, sec. 3.1.2]).

Notice that the limit of this second series is Feynman path integral formula,

$$\langle \delta_{h,0} \mid O \rangle = \int e^{\frac{i}{h} S[\phi]} O[\phi]\, (d\phi), \tag{8}$$

as searched. Note also that our indirect travel gives the normalization factor $(h\varepsilon)^{1/2}$ almost directly from (3). Feynman [4] prefers to get it in the course of its approximation for small ξ.

Finally, note that the constant h we have introduced to control the series is arbitrary, and we can repeat the construction for any other value $h = \varepsilon_{00} > 0$. In this form we get a third series

$$\langle \delta_{\varepsilon_{00}} \mid O \rangle = \int e^{\frac{i}{\varepsilon_{00}} S[\phi]} O[\phi] \, (d\phi) \tag{9}$$

which is also a solution of the proposed regularization problem, and additionally fulfills that it is invariant under the control transformation:

$$\tau_\mu \langle \delta_h \mid = \langle \delta_{\mu h} \mid . \tag{10}$$

In the spirit of Wilson-Kogut transformations, we would like to call (5),(7),(9) respectively a bare series, a renormalized series, and the dressed series associated to the measure being defined.

The transformation which lets invariant the dressed series "corresponds to" Gell-Mann old RG transformation, its invariants being associated to some mean value equations. For instance, we could "formally" manipulate eq (8) for $O = \delta L / \delta \phi$ and then we get

$$\left\langle \frac{\delta L}{\delta \phi} \right\rangle = \int e^{\frac{i}{h} L[\phi]} \frac{\delta L}{\delta \phi} \, d\phi = \int e^{\frac{i}{h} L[\phi]} \, dL = \delta \left(\frac{1}{h} \right) , \tag{11}$$

which is obviously invariant under the transformation. RG invariance in this context relates to Ehrenfest theorem.

1. As a final remark, let us to note that the derivation here exposed could be nicely formulated in the framework of the tangent groupoid of the configuration space as defined by Connes [2] (see also [1,5] and references therein). Elements of tangent groupoid can be "chained" as arrays of vectors $(x, x_1)(x_1, x_2) \cdots (x_n, y)$, functions over them are operator kernels such that $(ab)(x_i, x_k) = \int a(x_i, x_k) b(x_k, x_j) \, dx_k$. More, Connes groupoid relates to the groupoid of paths only if we make a scaling $(x, y, \varepsilon) \to (x, y, 2\varepsilon)$ after composition of arrows.

A deeper understanding of this framework would be useful before to try to generalize the construction to dimensions higher than one. For dim > 1 we would expect more exotic fixed points, and additional entities (fermions?) are surely needed in order to keep the properties of line integration, and stokes-like theorems, that usually are codifyed inside the wedge product of differentials. Also, known issues related to the kind of metric (Euclidean, Minkowski...) should become more relevant.

This work has been inspired from discussions with the theoretical teams of Zaragoza University and Costa Rica university, whose patience the author wants to thank. Partial support from project MEC.xxx.yyy must be acknowledged.

This document is a working draft. Comments are welcome, but please check the database [6] for more recent work in the subject.

References

[1] J.F. Cariñena et al. Connes's Tangent Groupoid and Deformation Quantization, preprint math/9802102

[2] A. Connes, Non Commutative Geometry, Academic Press 1994

[3] P.A.M. Dirac, The Lagrangian in Quantum Mechanics, Phisik. Zeitschr. der Sowjetunion, 3, p. 64, 1933

[4] R.P. Feynman, Space-Time Approach to Non-relativistic Quantum Mechanics, Rev. Mod. Phys. 20, 2, p. 357, 1948

[5] A. Rivero, Introduction to the tangent grupoid, preprint dg-ga/9710026

[6] http://xxx.lanl.gov/find/math,physics/1/Rivero/0/1/0/1998/3/0

[7] G. Roepstorff, Path Integral Approach to Quantum Physics, Springer TMP.

[8] K.G. Wilson and J. Kogut, Phys Rep 12, 2, p. 75 (1974)

(End of transcription.) 9803035 (1)

Formalisation, static case (0-dimensional variational problem)

1. What object is being constructed?

Given a smooth $f : \mathbb{R} \to \mathbb{R}$, define its critical set

$$\mathrm{Crit}(f) = \{x \in \mathbb{R} : f'(x) = 0\}.$$

The paper's "Dirac measure concentrated in the critical points" is

$$\delta_f := \delta(f'(x))$$

viewed as a distribution (and, under nondegeneracy conditions, a finite positive measure supported on $\mathrm{Crit}(f)$).

The standard distribution identity behind this is: if $\varphi(x)$ has isolated simple zeros $\{x_i\}$ with $\varphi'(x_i) \neq 0$, then

$$\delta(\varphi(x)) = \sum_i \frac{\delta(x - x_i)}{|\varphi'(x_i)|},$$

hence with $\varphi = f'$,

$$\delta(f'(x)) = \sum_{x_i \in \mathrm{Crit}(f)} \frac{\delta(x - x_i)}{|f''(x_i)|}.$$

This is the rigorous meaning of “concentrated on critical points” in the generic Morse (nondegenerate) case.

1. Exponential (Fourier) representation and the move to the “difference quotient”

The first step is the Fourier representation of the delta distribution:

$$\delta(u) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{izu} \, dz,$$

in the distributional sense (i.e. after pairing with a test function). This is a classical representation of δ as an oscillatory integral. dlmf.nist.gov

Plugging $u = f'(x)$ gives the paper’s (1) (up to the 2π convention):

$$\langle \delta(f'(x)), g \rangle = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} e^{izf'(x)} g(x) \, dx \, dz.$$

Now the paper rewrites $f'(x)$ as the limit of a difference quotient:

$$f'(x) = \lim_{y \to x} \frac{f(y) - f(x)}{y - x},$$

and then changes variables so that “$y - x$” scales with the regularisation parameter.

A clean way to formalise exactly what is going on is to write the “amazing shape” (2) using the scaling $y = x + \varepsilon z$. Define, for $\varepsilon > 0$,

$$I_\varepsilon(g) := \int_{\mathbb{R}} \int_{\mathbb{R}} e^{\frac{i}{\varepsilon}(f(x+\varepsilon z) - f(x))} \, g(x) \, dz \, dx.$$

Then as $\varepsilon \to 0$,

$$\frac{f(x + \varepsilon z) - f(x)}{\varepsilon} \to z f'(x),$$

so formally

$$I_\varepsilon(g) \to \iint e^{izf'(x)} g(x) \, dz \, dx = 2\pi \langle \delta(f'(x)), g \rangle.$$

If one re-expresses dz as $(dy)/\varepsilon$ via $y = x + \varepsilon z$, this becomes exactly the paper's (2):

$$I_\varepsilon(g) = \iint e^{\frac{i}{\varepsilon}(f(y)-f(x))} g(x) \, \frac{dy}{\varepsilon} \, dx,$$

with the crucial (often implicit) point that y is not an independent far-away variable: the scaling means only the near-diagonal region $y \approx x$ contributes as $\varepsilon \to 0$. That "near diagonal" idea is precisely what the tangent groupoid packages geometrically (pairs (x, y) with a scale parameter ε and a limiting tangent vector).

1. 9710026 n" as a half-density / amplitude

The paper then defines the "halved" object

$$A_\varepsilon(O) := \varepsilon^{-1/2} \int e^{\frac{i}{\varepsilon} f(x)} \, O(x) \, dx,$$

and claims that (in the unique critical point case) $|A_\varepsilon(O)|^2$ recovers the full delta-object.

The algebra is:

$$|A_\varepsilon(O)|^2 = \frac{1}{\varepsilon} \iint e^{\frac{i}{\varepsilon}(f(x)-f(y))} \, O(x) \, \overline{O(y)} \, dx \, dy.$$

Now apply the same "near-diagonal scaling" $y = x + \varepsilon z$. Then $dy = \varepsilon dz$ and

$$|A_\varepsilon(O)|^2 = \iint e^{-\frac{i}{\varepsilon}(f(x+\varepsilon z)-f(x))} \, O(x) \, \overline{O(x + \varepsilon z)} \, dz \, dx.$$

As $\varepsilon \to 0$, $\overline{O(x + \varepsilon z)} \to \overline{O(x)}$, so the limit becomes

$$\iint e^{-izf'(x)} \, |O(x)|^2 \, dz \, dx = 2\pi \int \delta(f'(x)) \, |O(x)|^2 \, dx.$$

This matches the paper's schematic (4) with $g(x) = |O(x)|^2$, provided one is in a regime where the off-diagonal contributions are negligible/controlled (unique stationary point, or a limiting procedure that suppresses interference terms).

Conceptually, this is the standard quantum-mechanical pattern "probability = modulus squared of amplitude", i.e. the Born rule in its simplest functional-analytic form. math.ru.nl+1

The other (more geometric) interpretation of "halved" is: on a Lie groupoid, convolution algebras are canonically formulated using half-densities to make the convolution coordinate-free. This is stated explicitly in tangent-groupoid quantisation literature (they often work with a Riemannian metric to identify densities, but the conceptual reason remains: kernels live naturally as half-densities). matwbn.icm.edu.pl+1

That provides a precise “why a square root?”: amplitudes are naturally half-densities; probabilities are densities.

1. Stationary phase: what the halved expression actually “converges to”

Strictly speaking, $A_\varepsilon(O)$ does not converge as a number for $\varepsilon \to 0$ without additional prescriptions, because the phase oscillates like $e^{if(x_0)/\varepsilon}$ at the critical point x_0. What does converge, in the usual semiclassical sense, is an asymptotic expansion in powers of ε obtained by stationary phase. This is precisely the kind of situation where oscillatory integrals define distributions but not pointwise limits; the delta function itself is the archetypal example. en.wikipedia.org+1

In the nondegenerate single critical point case (Morse), stationary phase gives

$$\int e^{\frac{i}{\varepsilon}f(x)}O(x)\,dx \sim e^{\frac{i}{\varepsilon}f(x_0)}e^{i\frac{\pi}{4}\operatorname{sgn}(f''(x_0))}\sqrt{\frac{2\pi\varepsilon}{|f''(x_0)|}}\,O(x_0) + O(\varepsilon^{3/2}),$$

hence

$$A_\varepsilon(O) \sim e^{\frac{i}{\varepsilon}f(x_0)}(\text{phase})\sqrt{\frac{2\pi}{|f''(x_0)|}}\,O(x_0) + O(\varepsilon).$$

Taking the modulus squared removes the wild phase and yields a well-defined leading term:

$$|A_\varepsilon(O)|^2 \to \frac{2\pi}{|f''(x_0)|}\,|O(x_0)|^2,$$

which is exactly $\langle \delta(f'), |O|^2 \rangle$ (up to the 2π convention), because $\delta(f') = \delta(x - x_0)/|f''(x_0)|$ in the single nondegenerate root case.

This calculation is the precise static analogue of how the semiclassical limit of the path integral is dominated by classical extrema, with the Van Vleck / Hessian determinant supplying the prefactor.

Lean skeleton for the static “Dirac measure on critical points” (measure-level)

This does not attempt to formalise stationary phase or oscillatory integrals inside Lean (that is substantial analysis infrastructure); instead it formalises the measure that corresponds to $\delta(f')$ in the generic “finite nondegenerate critical points” case: a weighted sum of Dirac measures supported on critical points.

```
/-
  Rivero (quant-ph/9803035), static case:  (f') is a (weighted) Dirac measure
  supported on critical points when f' has isolated simple zeros.

  This file formalizes the *measure-theoretic* object:
     _f :=  _{x Crit} (1 / ‖f''(x)‖) •  _x
  and basic evaluation properties.
-/

import Mathlib.Analysis.Calculus.Deriv
import Mathlib.MeasureTheory.Measure.Dirac
import Mathlib.MeasureTheory.Integral.Bochner
import Mathlib.Data.Real.Basic

open scoped BigOperators
open MeasureTheory

namespace Rivero9803035

noncomputable section

-- Critical point predicate
def IsCritical (f :   →  ) (x :  ) : Prop := deriv f x = 0

-- Weight used in  (f') = Σ  (x-xi)/|f''(xi)|
def critWeight (f :   →  ) (x :  ) :   :=
(Real.norm (deriv (deriv f) x)) ¹

-- A "Dirac measure concentrated on a finite set Crit"
def critDiracMeasure (f :   →  ) (Crit : Finset  ) : Measure   :=
  x in Crit, (critWeight f x) • (Measure.dirac x)

-- Evaluation of integrals against the finite weighted Dirac measure
theorem integral_critDiracMeasure
    (f :   →  ) (Crit : Finset  ) (g :   →  ) :
    (  x, g x  (critDiracMeasure f Crit))
      =   x in Crit, (critWeight f x) * g x := by
  classical
  -- Expand the measure as a Finset sum and use simp rules for integrals
  -- against dirac and scaling.
  simp [critDiracMeasure, critWeight, Finset.integral_sum_measure]

/-
  Support is contained in Crit (as a set), hence in Crit(f) if Crit enumerates it.
  (One can strengthen this with additional lemmas equating the support to Crit
   when all weights are nonzero.)
```

```
-/
theorem support_critDiracMeasure_subset
    (f :   → ) (Crit : Finset  ) :
    (critDiracMeasure f Crit).support   (Crit : Set  ) := by
  classical
  -- Uses `support_sum` + `support_smul` + `support_dirac`.
  intro x hx
  -- Proof details omitted; fill using lemmas in MeasureTheory.
  sorry

end Rivero9803035
```

The key point is: once you accept (from distribution theory) the identity $\delta(f') = \sum \delta_{x_i}/|f''(x_i)|$ in the nondegenerate case, the "Dirac measure concentrated at critical points" becomes a straightforward finite measure; Lean can handle that structure even if it does not yet handle the oscillatory-limit construction that motivates it.

Formalisation, quantum mechanics (1D time): from (3) to the path integral

1. What is the analogue of "$\delta(f')$"?

Now the configuration is a path $\phi : [t_0, t_1] \to M$ (for simplicity $M = \mathbb{R}^d$), and the functional is the action

$$S[\phi] = \int_{t_0}^{t_1} L(\phi(t), \dot{\phi}(t))\, dt.$$

The "critical points" are classical paths satisfying Euler–Lagrange:

$$\frac{\delta S}{\delta \phi} = 0.$$

So the direct analogue of the static object is the functional delta

$$\delta\Big(\frac{\delta S}{\delta \phi}\Big),$$

a "Dirac measure on path space concentrated on classical solutions".

The paper's (5) is exactly the time-sliced attempt to build that delta-functional by repeating the static construction at each intermediate point x_i.

1. The time-slicing is the action-level version of the tangent groupoid "near diagonal" scaling

Each time slice 9803035 (1) x_i and increments $\xi_i = x_{i+1} - x_i$. The small-time limit requires control of $\xi_i \to 0$ jointly with $\varepsilon_0 \to 0$, exactly as the paper stresses (and notes that Feynman's Schrödinger derivation relies on the small-ξ expansion).

Geometrically, the collection $(x_i, x_{i+1}, \varepsilon_0)$ is a chain of arrows in the "secant" part of the tangent groupoid; the limit $\varepsilon_0 \to 0$ gives tangent vectors (x_i, v_i), i.e. velocities, which is exactly the "vector as scaled pair of points" viewpoint emphasised in the tangent groupoid construction.

1. Why does the exponential weight appear, and why is the normalisation $(h\varepsilon_0)^{-n/2}$?

In Feynman's time-slicing derivation the key kernel is the short-time propagator; the measure factor is fixed so that the composition law holds and the Schrödinger equation emerges. link.aps.org+1

Rivero's paper reframes this: the same normalisation already appears in the static "half" expression (3), and when you iterate the "half" construction n times you get $\varepsilon^{-n/2}$, which becomes $(h\varepsilon_0)^{-n/2}$ once you identify the control parameter h (to be interpreted as $\hbar$) tying together the two limits.

That matches the standard physics story: the factor $(2\pi i\hbar\Delta t/m)^{-d/2}$ per time slice is fixed by consistency of composition and th 9803035 (1) l; Rivero's point is that the structure is already present as "square-root delta" normalisation in the static setting.

1. The "RG control" step and the Wilson–Kogut analogy

The paper's step (6)–(10) is a Wilsonian idea: define a refinement flow (more slices, smaller ε_0), and control it by mapping back to a fixed scale h. It explicitly cites Wilson–Kogut, and the analogy is conceptually accurate: Wilson–Kogut's central idea is a flow under rescaling/coarse-graining; here rescaling is tied to time-slicing and to the scale parameter in the exponential. max2.physics.sunysb.edu+1

This RG viewpoint is also developed in your "ode" notes in the language of scaling between the "natural" path-composition groupoid $(x, y, t)(y, z, t') = (x, z, t+t')$ and Connes' tangent-groupoid composition $(x, y, \varepsilon)(y, z, \varepsilon) = (x, z, \varepsilon)$, with "halving scales" as the key operation.

Interpretation of the halved expression in quantum mechanics

The halved expression becomes the propagator amplitude:

$$K(\phi_1, t_1; \phi_0, t_0) \sim \int_{\phi(t_0)=\phi_0}^{\phi(t_1)=\phi_1} e^{\frac{i}{\hbar}S[\phi]}\,\mathcal{D}\phi.$$

Probabilities are $|K|^2$, which at the path-integral level is exactly a doubling of path variables (a "forward" path and a "backward" path), the same structural reason why $\langle\delta^{1/2}|\cdot\rangle\langle\delta^{1/2}|\cdot\rangle^*$ appears in the static case. This "doubling" becomes a formalism in its own right in QFT/nonequilibrium physics: the Schwinger–Keldysh closed-time-path integral explicitly doubles fields (bra/ket) to compute real expectation values, and $|Z|^2$-type objects are naturally represented on such a contour. link.springer.com+1

Formalisation, field theory in D=4 (configuration is a field on spacetime)

1. The field-theory analogue of the variational "Dirac measure"

In classical field theory on $D = 4$, one has a field configuration Φ (e.g. scalar, gauge field, spinor), action functional

$$S[\Phi] = \int_{\mathbb{R}^{1,3}} \mathcal{L}(\Phi, \partial\Phi)\, d^4x,$$

and Euler–Lagrange equations

$$\frac{\delta S}{\delta \Phi} = 0.$$

The direct analogue of the static $\delta(f')$ is the functional delta

$$\delta\left(\frac{\delta S}{\delta \Phi}\right),$$

a "Dirac measure on field space concentrated on classical solutions".

Formally, the same Fourier/oscillatory-integral logic suggests that such a functional delta can be represented (after regularisation) by an oscillatory functional integral, and that the "halved" expression should again be an amplitude whose modulus squared yields the full delta-like concentrating object.

1. Why D=4 is where renormalisation becomes the real obstruction (and why Rivero's "convergence condition" becomes "renormalisability/continuum limit")

In quantum mechanics (0+1D), time slicing is usually sufficient to define the propagator (modulo standard measure subtleties). In QFT (4D), the continuum limit requires both a UV regulator (e.g. lattice spacing a or momentum cutoff Λ) and a renormalisation procedure to keep predictions finite as the regulator is removed.

This is exactly the content of the Wilson–Kogut viewpoint: physics at long distances is obtained by integrating out short-distance degrees of freedom and rescaling, producing an RG flow of couplings; the continuum limit exists at (or flows into) an RG fixed point. max2.physics.sunysb.edu+1

So in the language of 9803035, the statement "the regularizated measure has a limit iff the Feynman measure has a finite one" becomes, in 4D field theory, "the functional integral exists as a continuum object only after renormalisation, and only for theories for which the RG flow admits a controlled continuum limit".

1. The tangent groupoid / Connes–Kreimer pivot

The paper already hints that higher-dimensional generalisation will have "more exotic fixed points" and possibly require additional entities (fermions, wedge products) for Stokes-like properties.

In the noncommutative-geometry/QFT interface, Connes–Kreimer show that perturbative renormalisation is controlled by a Hopf algebra of rooted trees (and related structures), explicitly tying renormalisation to algebraic/combinatorial groupoid-like structures. arxiv.org+2alainconnes.org+2

Your "ode" notes connect this rooted-tree Hopf algebra to Butcher's rooted-tree calculus for Runge–Kutta composition ("composition proceeds by halving scales"), and explicitly frames it as "quantization implied by renormalization" within the tangent groupoid mindset.

This provides a coherent D=4 storyline consistent with the 9803035 program:

- A regularised functional integral is defined at a finite scale (lattice spacing a, cutoff Λ, or groupoid scale ε).
- A coarse-graining/composition step naturally changes the scale (e.g. two steps ε compose to 2ε).
- A "renormalisation" map rescales back to a reference scale (your h) and updates couplings/parameters.
- The existence of a consistent limit as the regulator is removed is equivalent to existence of a renormalised continuum theory.

Interpretation of the "halved expression" across the three levels, with literature pointers

Static (0D):

The halved object is a square root of a concentrating distribution. Algebraically, it matches "amplitude squared gives density"; geometrically, it matches the half-density formalism for groupoid convolution (kernels are naturally half-densities; products yield densities). matwbn.icm.edu.pl+2math.ru.nl+2

Quantum mechanics (0+1D):

The halved object is the propagator amplitude $K \sim \int e^{iS/\hbar}\mathcal{D}\phi$. Its modulus squared naturally requires a doubled path description (bra/ket), which becomes systematic in Schwinger–Keldysh. link.aps.org+2jontalle.web.engr.illinois.edu+2

Field theory (D=4):

The halved object is the generating functional/amplitude (in-out), whereas modulus squared and physical probabilities/real-time expectation values are naturally represented by a doubled-field formalism on a closed time contour (in-in / Schwinger–Keldysh). link.springer.com+2max2.physics.sunysb.edu+2

Tangent groupoid link (all levels):

The consistent appearance of (i) "near-diagonal scaling", (ii) composition changing the scale, and (iii) half-density normalisation is precisely what Connes' tangent groupoid packages: it interpolates between "finite differences / kernels" ($\varepsilon > 0$) and "tangent/cotangent objects" ($\varepsilon = 0$), and continuity of the product yields a strict deformation/quantisation picture. alainconnes.org+1

The paper 9803035 can therefore be read as proposing a unified slogan: "the complex exponential weighting is already the oscillatory integral representation of a Dirac(-type) object localising on extrema; quantisation is what you get when you insist this localisation procedure be compatible with composition across scales," with the halved expression identifying the amplitude/half-density whose modulus square produces the concentrating measure.

19. User

Now lets consider our discussions in the start of the conversation, that is also associated to two old notes of mine, https://a.rivero.nom.es/research/0404086.pdf https://a.rivero.nom.es/research/Areal.pdf transcribe them and see if they help to consider a quantisation method based on Planck length and if we can join the whole corpus into a new approach to explain path integrals and how they are forced to exist as a correction to classical mechanics. Organize all of the corpora as an article on fundations of quantum mechanics, but that could be useful as complementary lecture in the first month of a QFT course.

20. Assistant

Tools used in this answer: (1) file_search on the uploaded PDFs for transcription and internal citations; (2) Python + container image rendering to visually verify the equations in the PDFs; (3) web.run to confirm key bibliographic anchors (Dirac/Feynman, tangent groupoid quantization, minimal-length/GUP reviews, Connes–Kreimer renormalization).

ARTICLE (lecture-note style): Minimal length, areal speed, and why path integrals "want to exist"

1. What is being unified across the notes

Across the five notes you provided, there are two complementary lines:

A. "Path integrals from classical variational calculus" (9803035 + tangent-groupoid notes)

You recast “the classical equations of motion” as the support of a Dirac measure on a space of configurations/paths, then use Fourier/oscillatory representations of that Dirac measure to obtain a complex exponential of the action. This already gives, in finite dimension, a clean origin for the phase factor exp(iS/ħ), and in 1D (time) it suggests the need for an RG-style control of the regulator that forces a parameter h to survive as a finite constant. 9803035 (1)

ode

9710026

This dovetails with the historically standard Dirac → Feynman action-amplitude idea. informationphilosopher.com+1

B. “Planck length/area as a geometrical seed for quantum scales” (0404086 + Areal)

You propose that if there is a minimal area/time scale (Planck), then classical gravity in 3+1 has a special “cancellation” property: demanding one Planck area per Planck time for a circular gravitational orbit gives the Compton scale R = ħ/(mc). In the “areal speed” language, the Planck areal speed ȦP = cLP acts as a baseline, and inverse-square gravity makes Newton’s constant drop out when expressing the circular-orbit radius in Planck units.

This belongs to the wider minimal-lengt 0404086 APS Link+2Springer Nature Link+2 Areal (4) arch2 turn1search3

Putting A + B together suggests a coherent storyline: if Nature enforces a nonzero minimal geometric scale, then the “regulators” that classical calculus implicitly sends to 0 cannot be taken to 0 physically; the correct replacement is an oscillatory/phase-weighted integral. That is exactly the mathematical shape of the path integral, and the deformation parameter can be “geometrized” via Planck length/area.

What follows is a structured synthesis that is meant to be usable as complementary reading early in a QFT course, while staying faithful to what your notes actually do.

1. Static variational problems as measures; why an exponential of the action appears

Start from the minimal (finite-dimensional) variational problem:

Given a smooth function f(x), find its critical points f (x)=0 (or minima under extra conditions), and evaluate observables g(x) at those points.

A distribution-theoretic encoding is: “the solution set” is supported by the Dirac distribution (f (x)). In pairing form,

 (f (x)) | g(x) =   (f (x)) g(x) dx.

Your note 9803035 rewrites this by using a Fourier representation of :

(u) = e^{i z u} dz/(2),

so

(f) | g = e^{i z f (x)} g(x) dz dx.

Then it uses a difference quotient f (x) = lim_{y→x} (f(y)−f(x))/(y−x), substitutes = (y−x)/z, and arrives at an oscillatory integral with phase (f(y)−f(x))/ . In your notation one gets an asymptotic form like

_f | g = lim_{ →0} e^{ i (f(y)−f(x))/ } g(x) dx (dy/).

This is the key point: an exponential of an "action difference divided by a small pa 9803035 (1) re any quantum postulate, purely from representing (f) via Fourier/oscillatory integrals.

1.1 The "halved expression" and why it looks like an amplitude

In 9803035, once you are in the regime "f has only one extremal point," you introduce a "halved" object:

_f^{1/2} | O = lim_{ →0} (1/ ^{1/2}) e^{ i f(x)/ } O(x) dx,

and then you recover _f | g by taking a modulus-square-like product:

_f | g = _f^{1/2} | O _f^{1/2} | O *,

with g(x)=O(x)O*(x).

Interpretation (mathematically clean, physically standard):

- The "full" object (f 9803035 (1)

e) distribution concentrating on classical solutions.

- The "half" object is naturally complex and behaves like an amplitude; the probability-like object is its modulus square.
- In semiclassical stationary-phase approximations, this is exactly the pattern: a complex phase e^{i f/ } times a prefactor whose square gives a real density on critical points.

This "half-density" viewpoint is also natural in geometric quantization and in groupoid/operator-kernel formulations, where one often quantizes half-densities so that composition laws work cleanly. Your later tangent-groupoid notes are consistent with this reading.

2. From 9710026 why a regulator + an RG-style control naturally produces Feynman's kernel

2.1 Why the infinite-dimensional generalization is nontrivial

For classical mechanics, the object is not f(x) but an action functional S[ ] on a path space (or field configuration space). "Classical solutions" are the stationary points S/ = 0.

The direct analogue of (f (x)) is a (formal) Dirac measure concentrated on the critical points of S, i.e. something like (S/). But in function spaces you cannot just assume the oscillatory integral regulator converges; you need a controlled limiting procedure.

This is where 9803035 introduces a Wilson–Kogut inspired viewpoint: the time-slicing regulator has to be controlled under refinement/coarse-graining to have a limit.

2.2 Time slicing and the emergence of an h scale

In your 9803035, the discretized f 9803035 (1) troduces two small parameters (schematically and , one tied to the "Fourier" regularization and one to time slicing). You then propose that convergence requires introducing a scale h that controls how the parameters transform under refinement:

→ _h( ), → _h( ),

with the idea that repeated bisection (n→2n, step halving) and subsequent rescaling should leave the "dressed" series invariant.

This is an RG statement in spirit: define a "bare" discretized expression, a "renorm 9803035 (1) ontrol map, and the "dressed" (invariant) one.

The limiting expression you reach is the familiar Feynman-type formula:

^{h,0} | O = e^{ i S[ ]/h } O[ ] (d),

and you explicitly note this is "Feynman path integral formula" and that the normalization factor appears in the same way as in the static "halved" expression.

This matches the historical record: Dirac already emphasized the Lagrangian/action s 9803035 (1) eynman explicitly postulated that each path contributes an amplitude whose phase is the classical action in units of ħ, with a sum/integral over paths yielding the wavefunction/propagator. informationphilosopher.com+1

Key synthesis point:

- In your framework, the complex exponential is not an extra axiom; it is the Fourier/oscillatory representation of a "Dirac measure on solutions."
- The appearance of a finite parameter h is tied to the requirement that the limiting procedure be consistent under refinement, i.e. an RG/continuum-limit consistency requirement.

This is one precise sense in which "path integrals are forced to exist as a correction to classical mechanics": the formal classical object (S/) needs a regulator, and the consistent continuum limit wants a nontrivial scale to survive.

3. Tangent groupoid: a clean geometric home for "finite differences → derivatives" and for deformation quantization

3.1 What the tangent groupoid does conceptually

Your "Introduction to the tangent groupoid" (9710026) and "Revisiting the tangent groupoid" (ode.pdf) stress a central geometric idea:

- A derivative is a limit of secants (finite differences).
- The tangent groupoid packages, in a single object, both:
- the pair groupoid M×M (finite displacements), and
- the tangent bundle TM (infinitesimal displacements),

with a deformation parameter (often called or t) that interpolates between them.

This is closely aligned with the standard Connes tangent groupoid approach to strict deformation quantization; see e.g. work on Connes' tangent gro 9710026 t ode arXiv

3.2 Why this matters for "why quantum"

In ode.pdf you explicitly frame classical mechanics as having a hidden ambiguity when it mixes differential and integral calculus:

- a → 0 is the "differential" scale,
- b → ∞ is the "integral/thermodynamic" scale,
- classical mechanics effectively chooses the product a · b = 0,
- but Nature can choose a finite value a · b = h, producing a "nontrivial correction."

In groupoid language, "a→0" is pushing to the tangent bundle boundary (purely classical). Keeping a finite deformation parameter means you are not exactly on the boundary; you live in the interpolati ode natural algebraic structure becomes noncommutative (or, in semiclassical terms, deforms Poisson to commutator).

This supplies a geometric interpretation of the "forced correction" idea:

- Classical mechanics corresponds to taking the tangent limit strictly.
- A minimal length/area principle makes it physically impossible to take that limit all the way.
- The "finite " regime is mathematically a deformation quantization regime.

3.3 Rooted trees / renormalization link

Your 9803035 already points to Wilson–Kogut as the right language for controlling the continuum limit. The Connes–Kreimer Hopf-algebra viewpoint makes explicit how renormalization is governed by combinatorics of rooted trees and recursive subtraction. arXiv

Even if one does not adopt all of that machinery here, the consistent picture is:

- "Summing over discretizations" + "requiring invariance under refinement" is an RG idea.
- RG fixed-point structure is exactly where a universal scale (like ħ, or a geometric substitute) is expected to appear.

4. The Planck-length/area side: areal speed, inverse-square gravity, and the Compton scale

This is where 0404086 and Areal add a second pillar: they propose a specific geometrical quantity (areal speed) and a specific cancellation property in 3+1 dimensions.

4.1 Planck areal speed and the "areal" rewriting of angular momentum quantization

Areal speed Ȧ is the rate at which area is swept by a radius vector. For central forces it is constant (Kepler's second law); in Newtonian mechanics it is directly tied to angular momentum:

Ȧ = L/(2m).

Your "On Planck Areal Speed" defines for a particle of mass m a natural areal speed scale

Ȧ_m = ħ/m = c L_m,

where L_m is the reduced Compton length. It then defines the Planck areal speed

Ȧ_P = L_P^2/T_P = c L_P = ħ/M_P.

If angular momentum is quantized as J = nħ, then for circular orbits

Ȧ = nħ/(2m) = (1/2) n (M_P/m) Ȧ_P.

So the usual "L = nħ" becomes, equivalently, "areal speed is quantized in units set by ħ/m," and can Areal (4) anck areal speed.

4.2 The "quantum gravity haiku": Planck area per Planck time   Compton radius

In 0404086 Areal (4) cle of mass m, for which radius will a circular gravitational orbit around the particle sweep one Planck area in exactly one Planck time?

The answer given is

R = ħ/(m c),

the Compton radius (reduced Compton length).

This is the striking “cancellation”: Planck length disappears from the final answer in 3+1-dimensional gravity (inverse-square).

4.3 The same cancellation in the areal-speed formula for inverse-square gravity

In Areal, for an inv 0404086 orce with coupling K, a circular orbit has radius

r = (4m/K) Ȧ^2.

For Newtonian gravity, K = G M m, so

r = (4/(G M)) Ȧ^2.

You then rewrite it in Planck units to show that Newton’s constant cancels the Planck area, yielding

r = 4 L_M (Ȧ^2 / Ȧ_P^2),

where L_M = ħ/(cM) is the reduced Compton Areal (4) mass. In particular, if Ȧ = (1/2) Ȧ_P, then r = L_M.

So the “haiku” is embedded in a more general identity: inverse-square gravity ties the orbit radius (in Planck units) to the Compton length, independent of G, once you express things in areal-speed units.

4.4 Dimensionality: why q Areal (4) th notes stress that the cancellation is special:

If one considers a generic force G m m / r^q, then demanding “area swept in a Planck time is an integer multiple of Planck area” leads to a dependence on G unless q=2 (inverse-square).

You then connect this to the idea that if q is fixed by the Fourier transform of a wave propagator (i.e. the massless field Green’s function), then inverse-square is tied to 3 spatial dimensions (3+1 spacetime).

This is not a proof that spacetime must be 3+1, but it is a coherent “selection mechanism” within your logic: the Planck-to-Compton cancellation is a 3+1 inverse-square feature.

4.5 Relation to minimal length/GUP literature

Your Areal note explicitly situates the discussion in the long-running “minimal length” literature (Mead; Padmanabhan; later reviews and GUP frameworks).

This theme is widely reviewed in the minimal-length/GUP literature. APS Link+2Springer Nature Link+2

5. A Planck-length-based quantisation Areal (4) oposal consistent with the corpus

Here is a way to assemble the corpus into a quantization method that uses Planck length/area as the pivot (while remaining honest about what is established vs conjectural).

5.1 What is already mathematically established in the corpus

(1) “Classical solutions as Dirac measures”

The idea that the classical equations can be encoded as a distribution supported on critical points, and that a Fourier representation introduces a complex exponential, is explicit in 9803035 and is mathematically standard at the level of distributions.

(2) “Square-root/half-density structure”

The “halved expression” in 9803035 gives an amplitude whose modulus square recovers the (regularized) measure concentrated on solutions.

(3) “Consistency under refinement requires a contr 9803035 (1) dynamic extension suggests that the existence of a continuum limit for the regulated functional integral is tied to an RG-style control, leaving a finite scale h. 9803035 (1) poid packages the classical limit and its deformations”

Your tangent-groupoid notes provide the geometric framework in which a finite deformation parameter has a clean meaning: it is not yet at the tangent boundary. 9803035 (1) is consonant with the broader strict deformation quantization literature around Connes' tangent groupoid. arXiv

(5) “In inverse-square gravity in 3+1, Planck-area/time considerations map to Compton length”

This is the core result of 0404086 + Areal: the Planck scale can cancel to yield a quantum (Compton) scale.

5.2 The quantization method (as a structured set of postulates + constructions)

Postulate P1 (geometric cutoff): There exists a minimal effective length (or area) scale _* (or a minimal areal speed), plausibly of order Planck. This is the “no →0 physically” input.

Construction C1 (deformed extremum selection):

Instead of the strict classical selection functional (S/), define a regulated oscillatory functional:

Z_{ }(O) =   exp(i S[ ]/) O[ ] (d),

and interpret classical mechanics as the →0 stationary-phase boundary. This is exactly the construction pattern in 9803035, but now with the philosophical shift: is not “just a regulator,” it is tied to a physical minimal scale.

Construction C2 (RG/consistency fixes the meaningful):

To have consistency under refinement/coarse graining (time slicing), require a control map   that makes the family {Z_ } stable under discretization changes. Your 9803035 introduces precisely this, with h as the scale that r 9803035 (1) g."

Identification step I1 (geometrizing h with Planck data):

In standard QM you identify h with ħ. In the "Planck-length pivot" stance, one can equivalently regard ħ as a derived scale from (c, G, L_P) by ħ = (c^3/G) L_P^2, i.e. the phase S/ħ can be rewritten as a phase proportional 9803035 (1) vided by L_P^2. This is not new physics; it is just algebraic inversion of the Planck-unit relation, but it realizes the "substitute ħ by a geometric scale" idea in a concrete way.

Interpretation I2 (why inverse-square gravity is special in this scheme):

Your areal-speed cancellation says: if you impose a Planck-geometric discreteness condition on central gravitational motion in 3+1, the resulting characteristic scale is Compton, i.e. a quantum scale. In other words, Planck-scale discreteness does not stay hidden at ultra-high energies; in inverse-square gravity it reappears as the Compton scale.

So, within this approach:

- The "path integral exists" because the correct way to encode "solutions of variational equations" in the presence of a nonzero minimal scale is by an oscillatory (Fourier) representation rather than a sharp .
- The "Planck length pivot" exists because ħ can be traded for L_P (together with c and G), and because gravitational central motion provides a mechanism by which Planck discreteness feeds into Compton scales.

5.3 Where the approach is conjectural (and what would count as completion)

The corpus is strongest on motivating the oscillatory form and on the Planck Compton cancellation in inverse-square gravity. What is not yet a theorem is:

- A rigorous derivation that a specific physical minimal length implies a specific value of =h=ħ in the deformation quantization of all systems (not just gravitational or central ones).
- A fully covariant field-theoretic version where the measure and regulator are controlled in a mathematically clean way (this is where tangent groupoids, renormalization, and half-densities likely matter).

But the direction is coherent: "quantization as a deformation forced by nonzero cutoff," rather than "quantization as an independent axiom."

6. Linking back to the relativistic angular-momentum bounds for central forces

Even though the prompt here is quantization, your Areal note already includes a key bridge to the earlier "relativistic central forces" discussion:

- In special relativity, circular orbits for an inverse-square gravitational coupling have a minimum possible angular momentum as r→0, namely GMm/c. This is consistent with known SR central-force analyses: there is a critical angular momentum scale K/c for a 1/r potential.

Written in areal-speed form, since L = 2m Ȧ in the Newtonian central-force kinematics, the SR bound is stated in Areal as

GMm/c < 2m Ȧ.

This is conceptually aligned with the "minimal scale forbids arbitrary localization" theme: when angular momentum is too Areal (4) too low), the orbit structure collapses, pushing one to either GR corrections or to a cutoff interpretation.

The same section then remarks Areal (4) eintroduce Planck-scale dependence into the orbit radius, but a GUP-induced correction to the Compton wavelength can compensate at first order for a particular coefficient choice.

That is exactly the type of "interplay" one expects in minimal-length scenarios. Springer Nature Link+1

So the corpus already contains the seed of the Coulomb/gravity + maximal/minimal angular momentum theme, connected to Planck length via the deformation-of-classical-or Areal (4)

7. How to present this as a first-month QFT complement

A workable 2–3 lecture arc (conceptual, not computational-heavy) is:

Lecture 1: Classical mechanics as measures on solution sets

- Variational principle → Euler–Lagrange.
- Encode "solution set" by (S/).
- Fourier representation of → oscillatory integrals.
- The "half" object as amplitude; modulus square recovers concentrated measure.

Lecture 2: Time slicing, composition, and why exp(iS/ħ) is the unique stable form

- Discretize time; show composition properties imply normalization factors (compare free-particle kernel normalization).
- Introduce RG control argument: consistent refinement needs a scale h.

- Connect to Dirac and F 9803035 (1) cally standard endpoint. informationphilosopher.com+1

Lecture 3: UV cutoffs, Planck length, and why 3+1 inverse-square is special

- Minimal length motivation (Mead/Padmanabhan; modern reviews). APS Link+1
- Areal speed and Planck areal speed; gravita 9803035 (1) o Compton.
- Dimensionality: long-range potentials depend on spatial dimension; inverse-square corresponds to 3 spatial dimensions; connect to why the cancellation is special.
- (Optional QFT bridge) Renormalization is also a "consistency under changing resolution" principle; rooted-tree Hopf algebra viewpoint can be mentioned at a high level. arXiv

Where "classical approximations of gauge theories" enter (as in 0404086):

- You note that constant forces can arise as limits (you mention QCD), and that inverse quartic forces connect to contact-interaction approximations (Fermi).

Pedagogically: this illustrates that "effective force laws" (and their scaling with distance) can shift with the regime and the integrated-out degrees of freedom—exactly the mindset QFT needs.

Appendix A: Transcription of "Some minor examples on discrete geometry" (0404086)

```
Some minor examples on discrete geometr:contentReference[oaicite:76]{index=76}UPT (Universic

By assuming a minimum value for area measurement, the emergence of quantum mechanics can be
easily motivated from simple consideration of gravitational forces. Here we provide some exa
and extensions that can be used for pedagogical purposes.
At the same time, the role of Planck units is shown to be of some theoretical influence ever
low energies.

A QUANTUM GRAVITY HAIKU

Given a particle of mass m, for which radius will a circular gravitational orbit around the
have the property of sweeping one Planck area in exactly one Planck time?

With this question we began a thread at PhysicsForums website [1] and the user Marcus, alway
optimistic, praised as a "QG haiku". As a catchy motivation, I argued that below that radius
it should be possible to use Planck time beats to divide area into regions smaller than prev
And so, a fundamental break of physics will happen at quantum Kepler length of the particle
```

Of course this is exaggerated, but really the question is about emerging quantum mechanics
i.e. about to accept the discretization of area as a first principle, then deriving from it
of quantum mechanics. Markopoulou and Smolin [3] have showed that QM emerges if some stochas
principle is incorporated to the description. Here this small puzzle shows that such princip
a requirement, as the answer is

$$R = \hbar/(m c)$$

i.e., Compton radius of the particle m.

We are saved from contradiction because relativistic quantum mechanics comes to help us: no body can be localized beyond its Compton radius.

In 1949, Osborne [8], in an unnoticed (except by [9]) paper, studied the possibility of meas
the curvature of a Schwarzschild solution of mass m using geodesical triangles from a test p
He applied quantum uncertainty to the test particle and then he derived sequentially Planck
Compton length and Planck length as successive bounds barring the measurement of curvature.
Our example reverses the path, taking Planck length as the fundamental principle.

The results of [8] show that, while we have invoked classical gravity for simplicity,
general relativity also contains the same argument. Also, similar results could surely be de
from other standpoints, for example the Veneziano [7], formula 4.2 string.

DEPENDENCE ON NUMBER OF DIMENSIONS

Note that Planck length has disappeared above, giving place to the usual QM relationship. This is a peculiarity of gravity on 3+1 dimensional space.

Consider a generic force $G m m / r^q$ so that the units of G will depend on q. In general, asking $A(t_P)$ to be a multiple n of Planck Area A_P, we have

$$2n = G^{1/2 - 1/q} m^{1/2} r^{3/2 - q/2} c^{3/q - 1} \hbar^{-1/q}$$

And only for the usual inverse-square law, q = 2, we get to cancel Newton constant.

If we assume that the value of q comes from Fourier transformation of a wave propagator, then we are forced to fix space time to be 3+1.

A DEPENDENCE ON INDETERMINACY

But there is also a dimension-independent way to inverse square forces.
Time before Feynman, the mathematician J.L. Synge proposed [10] to link the potential energy
to the total energy of the photons exchanged to generate V. It did not work very well,
but it suggests the following argument:

Assume that the preferent wavelength for exchanged photons is of the order of the distance r
between particles, and that this exchange happens under the cloak of indeterminacy principle
Then we have a momentum of order $p \quad \hbar/r$. On the other hand the photon has an energy
$E = p c = \hbar c / r$ with the associated time $t \quad \hbar/E = r/c$.
Thus

$$F = \Delta p/\Delta t \quad \hbar c / r^2$$

(To justify this Δt, imagine for instance a stable circular orbit. In this situation the par
changes momentum but keeps the energy constant. Thus the photon is virtual and it can only e
during the indeterminacy time. This is the key use of virtual, off-shell, particles)

It is possible to do the same trick for massive mediators if we start from
$\Delta p = (1/c)\sqrt{E^2 - m^2 c^4}$ and, ad hoc, $\Delta t = r/c$. Then one gets a short distance approximati
of yukawian force.

Thus from quantum indeterminacy it seems that forces should be always inverse square.
And again if we want them to come from a wave propagator we are forced to fix space time to

CANCELLATION, OR INDEPENDENCE

It could be worth to research the mutual cancellation of the two previous arguments.
We could define gravitational force as the result of a virtual exchange in the above way.
Then the units of G should be independent of space time, and we would always be able to canc
and Newton constant to obtain purely the Compton radius.

In exchange, the Fourier transform of the potential would be a "standard wave propagator"
only for three spatial dimensions.

Regretly this mechanism imposes upon us the need to invoke quantum mechanics,
thus it is muddier than the first, QG only, procedure.

EMERGENCE OF QUANTUM MECHANICS

On other hand, if we have got Compton Length, can we get quantum mechanics from it?
It is tricky. Compton Length is not exactly a quantum condition, but the result of pair crea
via indeterminacy principle, for extreme localisation of energy.
We could think that consistency of quantum gravity implies pair creation and Zitterbewegung,
but not the whole quantum mechanics.

Still, we can try in an antique way: the Bohr-Sommerfeld quantum condition can be formulated
at least for circles, via a Newton-Kepler principle: any bound particle sweeps a multiple
of Compton Area in a unit of Compton Time. This principle does not need gravity; it works
for any central force. Note that we have shifted the point of view: instead of considering t
of the central particle, here we have a fixed force field and we consider the mass of the or
particle.

The usual way to get BS quantisation is to invoke the De Broglie wavelength to check for destructive interference. And then, also, a bound particle sweeps a multiple of De Broglie A in a unit of De Broglie Time. Really, if we use any speed v to define area and time, the sam apply. While in the first example Planck Length was cancelled out, here speed simplifies and left only with the quantisation constant.

A historically minded reader could here enjoy the setup of the area principle in Newton [2], book 1, sect 2, Prop 1]; it is defined first for discrete areas and impulses.

UGLY DIMENSIONAL ANALYSIS

Naive dimensional analysis can be used also to justify inverse-square forces.
In natural units, force has a dimension $[L]^{-2}$. In absence of masses, the only scale avail is the separation r between particles. Thus either the force will have a dimensional couplin (spoiling renormalizability: try naive power counting) or it must use this unique scale avai then inducing a dependence on inverse square of distance.

Also, we can use a regulator mass M in adimensional way:

$$f = (K/r^2)\ (1 - (M\ r)^p)^q$$

and then we get a sort of approximation to short range, yukawian, forces (note q=1/2, p=2 fo

Even if trivial, it is perhaps worth to remark that, when we add some masses, naive dimensio offers the possibility of justifying constant and inverse quartic forces.
The corresponding equations, with a K still adimensional, are

$$f = K\ m\ m\ , \quad \text{and} \quad f = K / (m\ m\ x^4)$$

This is of some value because a constant force appears as a limit of QCD, while inverse quartic is a way to approach Fermi theory of contact interactions.
The mass in this later case is known to come from the electroweak bosons M_W and M_Z and not from the fermions involved.

SCALES OF MASS

Another user of PF, nicknamed Orion1, suggested to try the two body problem instead of the K
I am a bit slow to follow Orion1 calculations, so I have redone them, basically confirming t
Now, it is interesting to look also to the intermediate steps, so let me play them here.

We have two bodies 1,2 circling around the center of mass, with a common angular velocity such that $^2 R_i = G m_j / R^2$. Here R is the sum of both radius.
The equation is consistent with the center of mass condition

$$R_1 m_1 = R_2 m_2$$

The sum of cases 1 and 2 let us to solve for ,

$$= \sqrt{(G M / R^3)}$$

Now we impose that the area $A_i(t_P)$ must be a multiple n_i of Planck Area.
This translates to

$$n_i = (1/2) \sqrt{ (c M/\hbar) (R_i^2 / R^{3/2}) }$$

Or, using the C.M. condition to substitute R,

$$R_i = (4\hbar/c) (M^2 / m_j^3) n_i^2$$

Note now that using again this condition over the already solved radiouses,
we get a condition on the multiples of area, namely $(m_1/m_2)^2 = n_2/n_1$.
Or, say,

$$m_1^2 n_1 = m_2^2 n_2$$

Now lets go for the total angular momentum $L = m_1 \; R_1^2 + m_2 \; R_2^2$.
Substituting and after a little algebra we get

$$L = (2\hbar/m_P) [M^{3/2} / (m_1 m_2)^{3/2}] [(m_1^7 n_1^4 + m_2^7 n_2^4) / (m_1^3 n_1^2$$

Which, using the relationship between n and m, simplifies to

$$L = (2\hbar/m_P) [(m_1 + m_2)/(m_1 m_2)] m_1^2 n_1$$

Orion' case $L = \hbar$, $m_1 = m_2 \; m$ gives us, accordingly,

$$m = m_P / (4 n)$$

Also, if we take m_1 a lot greater than m_2, we recover the initial Compton for R_2 and als

$$L_{m_1>>m_2} \; 2 n_1 \hbar (m_1^2 / (m_P m_2)) = 2 n_2 \hbar (m_2 / m_P)$$

which shows that Planck mass keeps its role as a bound.

Last, a interesting mistake happens if we try to impose simultaneously low quantum numbers
(n_1 and n_2 small) and big mass differences (m_1 a lot greater than m_2).
Then we are driven to write

$$L^{WRONG}_{m_1>>m_2} \; 2 n_1 \hbar (m_1/m_P) (m_1/m_2)^{3/2}$$

that is not completely out of physical ranges, if for instance we put m_1 = 175 GeV and m_2

The first section of this note has taught us that QFT scales can be implied by cancelling pl
This last equation, even unjustified, tell us that planckian scales can be adequate to study

* Electronic address: arivero@unizar.es

[1] www.physicsforums.com/showthread.php?threadid=14007
see also
www.physicsforums.com/showthread.php?threadid=14992

[2] Sir Isaac Newton, Philosophiae Naturalis Principia Mathematica 1729
Online html with images at
http://members.tripod.com/~gravitee/booki2.htm
TEX source and pdf at
http://vorlon.cwru.edu/~ames/principia/

[3] Fotini Markopoulou, Lee Smolin Quantum Theory from Quantum Gravity
http://arxiv.org/abs/gr-qc/0311059

[4] Lee Smolin Matrix models as non-local hidden variables theories
http://arxiv.org/abs/hep-th/0201031

[5] Abhay Ashtekar, Stephen Fairhurst, Joshua L. Willis
Quantum gravity, shadow states, and quantum mechanics
http://arxiv.org/abs/gr-qc/0207106

[6] Ashtekar Abhay Ashtekar, Jerzy Lewandowski, Hanno Sahlmann
Polymer and Fock representations for a Scalar field
http://arxiv.org/abs/gr-qc/0211012

[7] M. J. Duff, L. B. Okun, G. Veneziano Trialogue on the number of fundamental constants
http://arxiv.org/abs/physics/0110060

[8] M.F.M. Osborne, Quantum-Theory Restrictions on the General Theory of Relativity,
Phys. Rev. 75, 1579-1584 (1949)
http://link.aps.org/abstract/PR/v75/p1579

[9] Misner 1957 review. Charles W. Misner Feynman Quantization of General Relativity
Rev. Mod. Phys. 29, 497-509 (1957)
http://link.aps.org/abstract/RMP/v29/p497

[10] J. L. Synge, Angular Momentum, Mass-Center and the Inverse Square Law in Special Relati
Phys. Rev. 47, 760-767 (1935)
http://link.aps.org/abstract/PR/v47/p760

[11] Alejandro Rivero Standard Model Masses and Models of Nuclei,
http://arxiv.org/abs/nucl-th/0312003

[12] for the next step in the ladder, check the plots at [11];
the masses of the standard model, in turn, could generate the extra coupling that guara

Source:

Appendix B: Transcription of “On Planck Areal Speed” (Areal.pdf)

Qeios

LETTER

On Planck Areal Speed

Alejandro Rivero

Instituto de Biocomputacion
y Fisica de Sistemas Complejos,
Universidad de Zaragoza,
50002 Zaragoza, Spain

E-mail: arivero@unizar.es

Abstract. We explore the concept of areal speed concerning mi:contentReference[oaicite:78]{i
We also note that the general relativity correction is compensated by a Generalized Uncertai

https://doi.org/10.32388/Q7SKTZ

1 Areal speed and minimal length

The question of the existence of a minimal scale of length associated with Planck units was

Usually, a minimal scale of length determines a minimal distance, a minimal area or a minima

Particularly we can build the Planck Areal Speed A_P from the Planck length L_P and the Plan

$$A_P = L_P^2 / T_P = c L_P = \hbar / M_P \quad (1)$$

And here an operational quantisation of the area and time involved in an areal speed does no

$$A = (p/q) A_P = k A_P, \quad (2)$$

perhaps only of topological interest. Still, let’s take some time to consider it.

In principle A_P is always smaller than most areal speeds happening in quantum mechanics. If

$$A = n\hbar/(2m) = (1/2) n (M_P/m) A_P \quad (3)$$

and as long as the test mass m is well smaller than Planck mass, we can guarantee that the c

2 Kepler length

For a circular orbit with an inverse square force of coupling K, the radius of the orbit is

$$r = (4m/K) A^2 = (n^2/(K m)) M_P^2 A_P^2 \quad (1)$$

Now, something peculiar happens with gravity, where $K = G M m$. The radius has no dependency

$$r = (4/(G M)) A^2 = 4 L_M (A^2/A_P^2) = 4 L_M (p^2/q^2) \qquad (2)$$

so we get explicitly $L_M = \hbar/(c M)$, the reduced Compton wavelength of the particle creating

Of course this is a re-statement of

$$A = (1/2) \sqrt{G M r} \qquad (3)$$

And anyway this regime is outside of the expected range of areal velocities. It can be studi

3 Gravity from lengths

A more interesting note, to us, is that we can reverse the argument to produce an emergent c

A Newtonian gravity is a mutual force law with a single universal coupling constant between

Let's try to prove the consistency of this definition, i.e., that the analytical form of a N

$$F(M,m,r) = G (M m)/r^2 \qquad (1)$$

for some constant G.

Let A_0 be the universal areal speed at Compton length radius. We also define the constants

$$2mA_0 = r_M m v, \qquad (2)$$

where v is the tangential speed of the particle and $r_M = \hbar/(c M)$ is the Compton length radi

$$F(M,m,r_M) = (m/r_M) v^2 = (c m M 4A_0^2/\hbar) (1/r_M^2) [c M r_M/\hbar]^n \qquad (3)$$

For the force to be mutual, i.e symmetrical under the exchange m   M, we need n = 0. Now let

$$G_0 = 4 c A_0^2/\hbar = 4 \hbar c / M_0^2 \qquad (4)$$

and we see that, as expected, the functional form of the force law is, for any m and M-and t

$$F(M,m,r_M) = G_0 (M m)/r_M^2 \qquad (5)$$

To see that this is the final closed form of the force equation, note that if a second force

$$0 = \Delta(M,m,r_M) = \Delta(m,M,r_M) = \Delta(m,M,r_m) \qquad (6)$$

Thus it has two different zeros in the radius coordinate for each arbitrary pair (M,m) of ma

This proves the assertion, but depending in the condition of a single coupling constant. Not

$$(G_0 M m)/L_0^{(2-n)} * (M^M/m^m)^{(2-n)/(M-m)} * (M/m)^{-((2-n) M m)/(I_0 (M-m)) r} * r$$

that have a secondary coupling I_0, a multiple of $\hbar/c$. A different attempt could be to ask t

4 Gravity, the strongest classical force

At Compton length from the center of force, all the forces would generate stable orbits a lo

Gravity is the strongest force that allows stable orbits with an areal speed as low as Planc

5 Extra dimensions

Which carries us to consider how the dimensionality of space affects the cancellation.

In general, if we allow force to be any power r^{-q} of the radius, the areal speed will be

$$G^{1/2 - 1/q} m^{1/2} r^{3/2 - q/2} c^{3/q - 1} \hbar^{-1/q} \qquad (1)$$

and only for $q = 2$ Newton's constant cancels out. Thus here we have a non-relativistic varia

In fact, our proof above fixes gravity to be ~ $1/r^2$, similarly to [12]. But it could be obj

6 GUP and relativity

Finally, let us consider a first approximation to relativistic effects.

If we enable special relativity, the circular orbits of a given M,m system have a minimum po

$$(G M m)/c < 2mA \qquad (1)$$

and $2A/c$ is bounded by the Schwarzschild radius, $r_M = G M / c^2$.

If we consider general relativity, the Schwarzschild effective potential amounts to a correc

$$r = 2 k^2 L_M [1 \pm \sqrt{(1 - 3 L_P^2/(k^2 L_M^2))}] \qquad (2)$$

and so

$$r_\pm \quad 2 k^2 L_M [1 \pm (1 - 3 L_P^2/(2 k^2 L_M^2))] \qquad (3)$$

which introduces our solution of interest,

$$r_+ \quad 4 k^2 L_M (1 - 3 L_P^2/(4 k^2 L_M^2)) \qquad (4)$$

as well as the branch of unstable solutions
    r_-   + (3/2) r_M

So we get an extra correction that seems to reintroduce Planck length L_P and thus Newton's

    L _M = L_M (1 +   L_P^2/L_M^2)   (5)

We can combine multiplicatively both corrections so that at first order they cancel, if   =

7 Acknowledgments

While his work was supported in part by Grant No. PID2022-136374NB-C22, funded by the Spanis

References

[1] C.A. Mead, Possible Connection Between Gravitation and Fundamental Length, Phys. Rev. 13
[2] T. Padmanabhan, Planck Length as the Lower Bound to All Physical Length Scales, Gen. Rel
[3] M. Maggiore, A Generalized uncertainty principle in quantum gravity, Phys. Lett. B 304 (
[4] L.J. Garay, Quantum gravity and minimum length, Int. J. Mod. Phys. A 10 (1995) 145 [gr-q
[5] S. Doplicher, K. Fredenhagen and J.E. Roberts, The Quantum structure of space-time at th
[6] C. Rovelli and L. Smolin, Discreteness of area and volume in quantum gravity, Nucl. Phys
[7] S. Hossenfelder, Minimal Length Scale Scenarios for Quantum Gravity, Living Rev. Rel. 16
[8] P. Bosso, G.G. Luciano, L. Petruzziello and F. Wagner, 30 years in: Quo vadis generalize
[9] B. Carr, L. Modesto and I. Premont-Schwarz, Generalized Uncertainty Principle and Self-d
[10] M.J. Lake and B. Carr, The Compton-Schwarzschild correspondence from extended de Brogli
[11] F.R. Tangherlini, Schwarzschild field in n dimensions and the dimensionality of space p
[12] J.L. Synge, Angular momentum, mass-center and the inverse square law in special relativ
[13] B.J. Carr, The Black Hole Uncertainty Principle Correspondence, Springer Proc. Phys. 17
[14] A. Rivero, Some minor examples on discrete geometry, 2004.
[15] A. Rivero, Some bounds extracted from a quantum of area, 2006.

Source:

Bibliographic confirmations used explicitly in the synthesis

- Feynman's "Space-Time Approach…" and the core exp(iS/ħ) + sum-over-paths postulate: APS Link
- Dirac's "The Lagrangian in Quantum Mechanics": informationphilosopher.com
- Tangent groupoid and strict quantization (Connes tangent groupoid principle): arXiv Areal (4) enormalization + rooted trees (Connes–Kreimer): arXiv

- Minimal length and GUP modern reviews / critiques: Springer Nature Link+1
- Early minimal-length arguments (Mead; Padmanabhan): APS Link+1

Final synthesis paragraph (no calls to action): The corpus is internally consistent in the following sense: (i) representing "classical solutions" as a Dirac measure and then using Fourier/oscillatory representations naturally produces complex exponentials of the action, and the "halved expression" is naturally an amplitude whose modulus square yields a concentrated (classical) distribution; (ii) demanding a consistent continuum limit under refinement motivates an RG-style control that leaves a finite scale h, matching the role of ħ in Dirac/Feynman; (iii) independently, imposing a Planck-geometric discreteness condition on inverse-square gravity in 3+1 yields the Compton scale and can be expressed cleanly via Planck areal speed, highlighting why inverse-square in 3 spatial dimensions is special; together these support a coherent viewpoint in which path integrals are not an arbitrary reformulation but a structurally stable "finite-scale correction" to the idealized classical limit, with Planck length/area providing a geometrical pivot for the deformation parameter.

21. User

It is nice you in your thinking look for Osborne's 1949, because Tong started in Cambridge as teacher assistant to Osborne. Anyway, our next goal is a mathematically correct narrative that argues that quantum mechanics was unavoidable since the first calculation of a uncuttable volume by Democritus (cone and pyramid, see also Hilbert 3rd Problem). It drives straight to newton calculus but then for trajectories and the way to properly define the classical solutions is as limits of the quantum solutions. I even heard a history about newton delaying the principle because two friends had told him that the naive limit of areas in the first chapter could be problematic.

that narrative you shoud write as considering to be a chapter 1 of a book on quantum field theory, or at least on relativistic quantum mechanics of elementary particles.

You could find some help leveraging, via bibliography search, my historical articles here uploaded and you can find more online. Also https://a.rivero.nom.es/research/0603123.pdf contains some history notes, plus more ruminations on the quantum of area. No need to quote any of them in the elaboration, in fact I prefer if you do not.

while it is a chapter 1, it is not just a narrative, it should be mathematically formal, and doing consistent steps that view quantum mechanics and then quantum fields as the inescapable conclusion, so that we can read the chapter also as another way for fundation of quantum mechanics.

22. Assistant

Tools used for this answer: web.run (bibliography confirmation and primary sources).

Chapter 1

From indivisibles to amplitudes: why quantum mechanics and fields are hard to avoid

1.1 The Democritean dilemma and what "uncuttable" really means

The earliest sharp place where "continuum physics" appears as a mathematical necessity is not in dynamics, but in geometry: the attempt to define volume from cutting.

Plutarch reports an episode (ascribed to Democritus and discussed by Chrysippus) about a cone cut by a plane parallel to its base: are the two new faces "equal" or "unequal"? If each successive slice is unequal to the previous, the cone looks like a staircase with jumps; if equal, the cone becomes a cylinder. The point is not a childish paradox: it is the recognition that "a solid is made of slices" does not become a definition unless one specifies a limit procedure that makes "infinitely near" meaningful. loebclassics.com+2people.math.harvard.edu+2

That is already a physics-grade lesson: a quantity defined by cutting must come with a stability axiom under refinement of the cut. Modern language calls this continuity (or countable additivity, depending on the framework). Without such an axiom, "finite cutting" can support multiple additive invariants at once.

Hilbert's 3rd problem makes the point brutally precise. In 2D, equal area implies scissors-congruence (Wallace–Bolyai–Gerwien), so "area by cutting-and-pasting" is essentially unique. In 3D, Hilbert asked if equal-volume polyhedra must be scissors-congruent. Dehn showed the answer is no by constructing another additive dissection invariant (the Dehn invariant) which obstructs equidecomposability even at equal volume. arxiv.org+2ncatlab.org+2

So "uncuttable volume" can be formalized as follows: if you insist that volume should be derivable purely by finite dissection and reassembly, you do not get a unique notion unless you add extra structure (continuity under limits, invariance under isometries, etc.). Hilbert 3 shows that additivity under finite cuts is too weak in 3D. arxiv.org+1

This will be our template: whenever a physical quantity is obtained by "summing local pieces", the limit is not optional; it is the definition. In mechanics, the corresponding "summing local pieces" is not slicing solids, but composing short-time motions.

1.2 From exhaustion to calculus: "limit" as a structural axiom

The method of exhaustion (Eudoxus/Euclid/Archimedes) is, in modern terms, a stability principle: approximate a curved region by inscribed and circumscribed polygonal regions whose difference tends to zero. This is exactly the kind of principle Newton later made explicit as “first and last ratios” in the Principia: quantities that converge to equality become “ultimately equal.” en.wikisource.org+1

The common feature is that the continuum object is not built out of indivisibles by summing them; it is built as a limit of refinements where the error is controlled. That is the real content behind Democritus’ cone dilemma: not that geometry is inconsistent, but that “summing slices” needs a rule for the vanishing thickness limit. loebclassics.com+1

This becomes a foundational pattern:

- In geometry, “volume” wants additivity plus stability under refinement.
- In kinematics/dynamics, “trajectory” wants composition of short-time steps plus stability under time refinement.

The second bullet is the bridge to quantum mechanics.

1.3 Newton’s polygonal motion and the problem of taking the limit

Newton’s Proposition 1 (Book I) is the key structural move in the Principia: replace a smooth orbit by a polygonal chain traced by inertial motion punctuated by instantaneous impulses toward a center. Equal times correspond to equal inertial steps; the areas of the triangles swept by the radius are equal; in the limit of infinitely many steps, the polygon becomes a curve and the equal-area law becomes exact. humanities.mcmaster.ca+2Mathematics+2

Mathematically, this is again an exhaustion argument, but now in time rather than space. The catch is that “take the limit” must be specified, or different refinements can lead to different limiting objects. Modern historical analysis of Prop. 1 emphasizes that Newton’s justification is concise and implicitly relies on the lemmas about limiting polygons/rectangles in the opening of Book I; and that understanding the continuum limit requires reading those lemmas as the controlling device. arxiv.org+1

Two things matter here for our purposes:

(1) Motion is being defined by a refinement limit of compositions. The primitive operation is composition of small steps (a discrete groupoid-like composition), and the continuum trajectory is its limit.

(2) The invariant that survives refinement is not a pointwise property of the curve; it is an area swept (a 2-dimensional quantity). Newton’s proof is already telling you that “area” is more stable than “velocity at an instant.”

This point is where your earlier “quantum of area” theme naturally attaches: classical mechanics is quietly written in terms of invariant areas (symplectic areas, action integrals) even before quantum mechanics makes them explicit.

1.4 Action as the additive invariant for composed motion

Now move from Newton’s geometric argument to modern variational mechanics.

For a configuration manifold Q and a Lagrangian $L(q, \dot{q}, t)$, the action functional on a path $q : [t_i, t_f] \to Q$ is

$$S[q] \;=\; \int_{t_i}^{t_f} L(q, \dot{q}, t)\, dt.$$

Its fundamental property is additivity under concatenation of paths:

if $q = q_1 \star q_2$ with a matching endpoint at time t, then

$$S[q] = S[q_1] + S[q_2].$$

This is the temporal analogue of volume additivity under cutting.

In Hamiltonian form on phase space T^*Q, one writes

$$S[q, p] = \int_{t_i}^{t_f} (p \cdot \dot{q} - H(q, p, t))\, dt,$$

and the geometric object behind it is the canonical 1-form $\theta = p\, dq$ with symplectic 2-form $\omega = d\theta$. For one degree of freedom, the "action variable"

$$J = \oint p\, dq$$

is literally the signed area enclosed by the orbit in the (q, p)-plane; for many degrees of freedom it generalizes as a symplectic area. This is the precise sense in which classical mechanics is organized around area-like invariants.

So, at the purely classical level, the correct "indivisible" is not a slice of space, but a segment of action. If you insist on defining motion by refining time-slices, the additive quantity that is stable under refinement is S, not the instantaneous velocity.

Up to now there is still no quantum mechanics. The unavoidable step comes when we ask: what does it mean to put a probability theory on paths that composes correctly under time-slicing?

1.5 From additivity to amplitudes: why $\exp(iS/\hbar)$ is forced

Suppose we want a dynamical rule that is:

- Local in time (built by composing short-time evolution),
- Compatible with superposition (intermediate alternatives add),
- Compatible with a probability interpretation.

The object that composes by integration over intermediate states is the propagator (kernel) K:

$$K(q_f, t_f; q_i, t_i) \;=\; \int_Q dq\, K(q_f, t_f; q, t)\, K(q, t; q_i, t_i).$$

This is the Chapman–Kolmogorov composition law (for Markov processes) and also the composition law for quantum propagators (unitary evolution kernels).

If you demand (i) composition by inserting intermediate configurations and (ii) additivity of the "weight" under concatenation of paths, you are driven to multiplicative weights on path segments, hence exponential of an additive functional. Concretely, if the weight of a path is $W[q]$ and you want

$$W[q_1 \star q_2] = W[q_1]\, W[q_2],$$

then writing $W[q] = \exp(A[q])$ forces

$$A[q_1 \star q_2] = A[q_1] + A[q_2],$$

so A is an additive functional. The natural additive functional already present is the action S. Therefore the only stable possibility is

$$W[q] = \exp\!\left(\frac{i}{\hbar} S[q]\right)$$

(up to normalization and possible boundary terms), where $\hbar$ is the constant setting the scale of phase oscillations.

This is not merely Feynman's aesthetic choice; it is the unique way to reconcile (a) additivity of classical action with (b) multiplicativity of composed evolution.

Historically, Dirac stated the key idea that the quantum amplitude is related to the classical action, and Feynman made the "sum over histories" formulation explicit and systematic. ncatlab.org+4informationphilosopher.com+4ifi.unicamp.br+4

At this point one can already see the "halving" phenomenon in a clean mathematical way: probabilities add, but amplitudes compose linearly before squaring. In the simplest reading, the wavefunction is a square root (a "half-density") of a probability density; passing from amplitude to probability doubles the exponent in Euclidean weights and removes the phase in Minkowski weights. This is the structural origin of many "half" factors: the quantum object is not a probability on paths, but an amplitude whose modulus-square is a probability density. (In semiclassics, the Van Vleck–Morette determinant is also literally a square root of a Jacobian, another "half-density" appearance.)

1.6 Why classical trajectories are better treated as limits of quantum ones

The slogan "classical mechanics is the $\hbar \to 0$ limit of quantum mechanics" is often stated loosely. A mathematically correct version must specify what kind of limit and in what topology.

There are two complementary rigorous routes:

A. Stationary phase (semiclassical) as $\hbar \to 0$

Consider an oscillatory integral

$$I(\hbar) = \int e^{\frac{i}{\hbar} S(x)} a(x)\, dx.$$

As $\hbar \to 0$, contributions away from critical points of S cancel by destructive interference; dominant contributions come from stationary points $dS = 0$. The stationary phase theorem makes this precise, yielding an asymptotic series whose leading term depends on the Hessian determinant at the critical points. In the path integral, the same mechanism selects classical paths (solutions of Euler–Lagrange) as the $\hbar \to 0$ saddle points, with a prefactor given by a functional determinant (the infinite-dimensional analogue of the Hessian). courses.physics.illinois.edu+2APS Link+2

So classical trajectories are not "the only paths that matter"; they are the stationary points that dominate the small-$\hbar$ asymptotics. This is already a mathematically precise sense in which classical solutions are limits of quantum solutions.

B. Euclideanization: Wiener measure and small-noise limits

For nonrelativistic quantum mechanics, after Wick rotation $t \mapsto -i\tau$, the propagator becomes a heat kernel and the path integral becomes a genuine measure-theoretic object: a Wiener integral (Feynman–Kac). en.wikipedia.org+1

In that setting, the "typical path" is not differentiable. Brownian motion sample paths are almost surely nowhere differentiable, and they have unbounded total variation but finite quadratic variation. dspace.mit.edu+2math.uchicago.edu+2

This is the geometric mirror of the Democritus cone dilemma: if you define a global object by refining partitions, the typical limiting object is rough, not smooth. Smoothness requires an additional limiting process.

In Euclidean quantum mechanics, the classical path re-emerges as a small-noise or small-$\hbar$ limit: the measure concentrates near minimizers of the Euclidean action. This is the rigorous probabilistic version of stationary phase (now Laplace's method rather than oscillatory cancellation).

Either way, the mathematically correct statement is: classical solutions are limiting objects extracted from a family of quantum (or Euclideanized) objects parameterized by $\hbar$, not primitive objects defined first.

This is exactly the pattern suggested by the history: volume is defined by limit of refinements, not by indivisible slices; likewise, motion is defined by a limit of composed amplitudes, and classical trajectories are extracted as limits/saddles.

1.7 A precise “delta-functional” intuition and why it leads to oscillatory integrals

There is a compact way to summarize why “classical first” is unstable.

The set of classical paths is defined by the Euler–Lagrange equation $\delta S/\delta q = 0$. One might wish to define a “classical measure” on path space supported only on solutions—formally something like

$$\mu_{\mathrm{cl}} \sim \delta\left(\frac{\delta S}{\delta q}\right) Dq.$$

But even in finite dimensions, a delta distribution is most naturally represented as a Fourier transform:

$$\delta(x) = \frac{1}{2\pi\hbar}\int e^{\frac{i}{\hbar}px}\, dp.$$

In infinite dimensions, this heuristic becomes exactly the oscillatory integral structure behind the path integral: enforcing constraints by delta functionals tends to introduce auxiliary “Fourier” variables, and the universal way to get localization to solutions in a limit is through rapid oscillation (or exponential decay after Wick rotation).

So “sum over histories with phase $S/\hbar$” is not a decorative reformulation; it is what you obtain when you represent “sharp classical constraint” as a limit of smoother objects. This is the same mathematical move as replacing “volume as a sum of slices” by “volume as a limit of approximations”: you trade an ill-defined “indivisible” picture for a controlled limit of well-defined approximants.

1.8 Why quantum fields appear once relativity is taken seriously

So far the discussion is general enough to be nonrelativistic. QFT enters when you add relativistic locality and the physical meaning of “a particle”.

A clean way to state the issue (as Tong does early in his QFT notes) is that there is a scale—the Compton wavelength—below which the very concept of a single pointlike particle breaks down: attempting to localize a particle too sharply forces energies high enough to create particle–antiparticle pairs, so “one particle in a box” is not a stable notion. damtp.cam.ac.uk+1

The structural consequence is that the correct configuration space is not “positions of finitely many particles”, but field configurations, with a Hilbert space that naturally accommodates variable particle number (Fock space). The dynamics is again built from an action principle, now on fields $\phi(x)$:

$$S[\phi] = \int d^4x \; \mathcal{L}(\phi, \partial\phi).$$

The same inevitability argument then repeats at the next level:

- Locality wants the action to be an integral of a local Lagrangian density.
- Composition wants evolution built by slicing time and integrating over intermediate field configurations.
- Superposition wants amplitudes, not probabilities, to be the linear composing objects.

Therefore the functional integral

$$Z = \int \mathcal{D}\phi \; e^{\frac{i}{\hbar} S[\phi]}$$

is the natural continuation of the path-integral logic, now on an infinite-dimensional configuration space.

The "equations of motion" survive, but in QFT the exact analog of Euler–Lagrange is an identity among correlation functions (Schwinger–Dyson), obtained by functional integration by parts. This makes explicit, in an equation form, the earlier claim that classical equations are not replaced but lifted into a distributional/functional framework. arxiv.org+1

1.9 The deformation viewpoint: tangent groupoids and why $\hbar$ behaves like a geometric parameter

There is a further conceptual consolidation that matches your stated direction: interpret quantization as a deformation, not as a discretization.

In strict deformation quantization, one constructs a family of algebras $A_\hbar$ such that as $\hbar \to 0$ the commutator reproduces the Poisson bracket:

$$\frac{1}{i\hbar}[a_\hbar, b_\hbar] \;\longrightarrow\; \{a_0, b_0\}.$$

Connes' tangent groupoid provides a geometric mechanism for such a deformation: it interpolates between the pair groupoid $M \times M$ (operators/kernels) and the tangent bundle TM (classical phase space via Fourier transform). The associated C^*-algebra yields a strict quantization scheme, and Landsman and others made the Weyl/Moyal rules appear naturally within this construction. arxiv.org+1

This matters for foundations because it reframes "$\hbar \to 0$" as a geometrical degeneration: quantum kinematics is not a different universe; it is a deformation of the classical one, with $\hbar$ acting like the parameter that collapses a groupoid convolution algebra to a commutative algebra of functions on phase space.

From this viewpoint, replacing $\hbar$ by a geometrical constant (Planck length/area) becomes conceptually coherent: any fundamental length introduces a preferred scale for the refinement limit and can serve as the deformation parameter in a suitably rescaled action. What must remain is a dimensionless phase $S/\hbar$ (or its analog), i.e. a scale that compares action-like quantities to a fundamental unit. Whether that unit is taken as $\hbar$ directly or derived from a geometric length with c and G is a secondary choice once you accept the deformation logic.

1.10 A single thread tying Democritus, Newton, path integrals, and QFT

The narrative can now be stated as a sequence of mathematically disciplined steps.

(1) "Additivity under cutting" is insufficient without a limiting axiom. Democritus' cone dilemma is the early recognition; Hilbert 3 proves in modern terms that finite dissection additivity does not uniquely fix volume in 3D. en.wikipedia.org+3loebclassics.com+3jstor.org+3

(2) Newton's Principia uses the same structure in time: motion is composed from short inertial segments; the continuum orbit is a limit; the robust invariant under refinement is an area law. The mathematical subtlety is the passage to the limit, which Newton supports by his "first and last ratios" lemmas and which has been analyzed carefully by historians of mathematics. humanities.mcmaster.ca+2arxiv.org+2

(3) In modern mechanics, the additive invariant for refinement under time-slicing is the action S. Requiring local composition pushes you toward multiplicative weights, hence exponential of an additive functional.

(4) Superposition and probability demand amplitudes that compose linearly before taking absolute squares. This is the structural origin of the "half" phenomenon: amplitudes are square roots (in the measure-theoretic sense) of probabilities.

(5) Once you accept composition + superposition + locality, the only stable family of weights is $e^{iS/\hbar}$, as anticipated by Dirac and formulated as "sum over histories" by Feynman. informationphilosopher.com+2APS Link+2

(6) Classical solutions are then recovered, not by declaring only one path exists, but as the stationary-phase (or Laplace) limit as $\hbar \to 0$. In the Euclideanized (Feynman–Kac) setting, the "typical path" is nowhere differentiable, and smooth classical trajectories emerge only as concentration limits—an exact analogue of how smooth volumes emerge from refined approximations rather than from indivisible slices. en.wikipedia.org+2dspace.mit.edu+2

(7) Relativity destabilizes the single-particle configuration space; localization below Compton scale produces particle number fluctuations, forcing field configurations and hence QFT. damtp.cam.ac.uk+1

(8) Deformation quantization (e.g., tangent groupoids) makes the $\hbar \to 0$ passage a geometric degeneration, reinforcing the view that "classical" is a limit point in a family rather than a self-standing starting point. arxiv.org+1

This chapter-level conclusion is the precise sense in which quantum mechanics can be presented as "unavoidable": once physical quantities are defined as refinement-stable limits of composed local pieces, and once superposition is treated as a structural requirement (not an experimental add-on), the path-integral/amplitude formalism is the mathematically natural completion of the classical limit procedure.

23. User

I think we are missing what uncuttable means here and for Democritus: the only way to calculate the volume of a pyramid is to make it indefinitely cuttable. So what we need is some way to refer our indefinite cutting back to an scale where we can measure. All the historical primary sources hint this way, that Democritus always considered the duality between atoms and vacuum, one separating the other, and how the vacuum is cut by inserting atoms. This is what comes to our days in the way of inserting more points in a runge kutta calculation or in a quantum Feynman path, and it even surfaces in the renormalisation insertions.

24. Assistant

In this context, "uncuttable" is best read as "not decomposable into smaller pieces without introducing extra structure," rather than "a smallest geometric chunk exists." The Democritean cone dilemma arises precisely because if you model a "slice" as a single plane section, then in the limit of an "infinitesimal" slice you lose the information that distinguishes cone from cylinder: consecutive sections become indistinguishable as mere circles.

Your 9904021 note makes the operational point sharply: a real cut produces two faces, and the "zero-thickness" idealisation is not "one plane" but "two planes approaching each other," i.e. a pair of parallel planes (a covector, in the Schouten–Golab sense) with a modulus (separation) sent to zero. 9904021 That is the missing ingredient in the naive paradox. The distinction between cone and cylinder is then not in "the area of the limiting circle" but in how you glue successive bifacial slices: continuity is imposed between one face of a slice and the adjacent face of the next slice, and that "internal structure" survives the modulus$\to$0 limit. 9904021

Once you phrase "slice" as an object that must be composable (gluable), "uncuttable" shifts meaning: the irreducible thing is not a little 3D chunk with thickness, but an oriented, ordered, weighted geometric datum that you can sum/compose. That is exactly how your 0006065 note motivates Democritus' three atom-words (rhythmos/diathige/trope): an "atom" in this mathematical sense has no extension (no “size”) in the cutting direction, yet carries finite measure in the transverse directions, and to add such things you must preserve ordering and orientation and assign a weight. 0006065v1 This is also where the atoms/vacuum duality becomes precise in modern terms: “size” is not a property of the atoms (forms/covectors) but of the vacuum/ambient structure in which they sit; the continuum supplies the metric scale, while the “atoms” supply the composable differential data. 0006065v1

Your “we need to refer indefinite cutting back to a scale where we can measure” is, in modern language, the move from a raw limit to a controlled limit.

In numerical dynamics and RK, “inserting more points” is literally refining a partition; but the core issue is not that you refine, it is that you need a rule that compares refinements at different step sizes without changing the physical content. In your ode note, this is stated as the RG-like exchange between “infinitely many infinitesimal steps” and “infinitely many intermediate points,” controlled respectively by a cutoff and an arbitrary reference scale h. ode The same note then spells out the renormalisation logic: introduce a cutoff, subtract divergent parts, and define the finite part relative to a chosen scale h; changing h without changing physics is the RG group property. ode

That same “controlled refinement” is exactly what shows up in your path-integral formalisation. In 9803035, the passage from a discretised variational problem to the Feynman formula is presented as: the double limit (, →0) is meaningful only after introducing a control scale h and an explicit rescaling map that “moves each term back” to that scale; the Feynman weight exp(iS/h) is then the convergence condition for the controlled (renormalised/dressed) series. 9803035 (1) In other words: the “indefinite cutting” (refining the time-slicing / inserting more intermediate points) becomes well-defined only after you impose a renormalisation-style identification between different discretisations.

Two further links make your atoms/vacuum picture line up with “Feynman paths” and “renormalisation insertions.”

First, your 0302285 note explicitly identifies the path integral as the limit of an equal-time discretisation “just as Newton Proposition 1,” and simultaneously explains that the tangent-groupoid interpolation (>0 operator product versus =0 commutative product) is a (de)quantisation continuity condition; choosing a limiting procedure corresponds to choosing a star product, whose star-exponential gives (after Fourier transform) a Feynman path integral. 0302285

0302285

Second, your tangent-groupoid introduction (9710026) phrases “reference back to a measurable scale” in exactly RG terms: a “bare line” of finite differences [x ,y ,] approaching the critical surface is rescaled by a renormalisation flow _{ / } to a fixed “physical” scale , and only then does one speak of a meaningful limit object. 9710026 This is essentially the abstract version of “keep cutting, but keep comparing cuts at a fixed ruler.”

With that framing, "inserting atoms into the vacuum" has a clean modern translation:

- In classical integration/ODE solving, you insert partition points into a continuum; the continuum is "vacuum," the partition points are "atoms," and the meaningful object is not the individual point but the composition law that lets refinements be compared (Butcher composition / scaling / controlled limit). ode
- In the Feynman path integral, you insert intermediate configuration points; the amplitude is only well-defined after a rule that makes the refinement limit controlled by a scale parameter h (your "dressed series" viewpoint). 9803035 (1)
- In perturbative QFT renormalisation, subdivergences are handled by "insertions" (counterterms / local corrections). Conceptually, these are the analogue of refining the description at shorter scales and then re-expressing the result at a fixed reference scale so that observables become scale-independent, exactly the logic you highlight in the RG passages. ode

So the missing piece in "uncuttable" is not that Democritus needed a literal minimal slice, but that the act of cutting/refining forces you to choose (and then control) a gluing/scale-comparison structure. Without it, "the infinitesimal slice" collapses to a featureless object (a circle) and the paradox fires; with it, you retain the ordering/orientation/weight (your rhythmos/diathige/trope package) that makes the limiting reconstruction nontrivial. 0006065v1

9904021

25. User

In my papers I note that two authors, possibly Barrow and some of the editors of the principia, stopped Newton to publish some notes; this is narrated later by Newton that only refers two "two friends" or two coleagues or something so, but now I can not find the exact reference. Look around.

26. Assistant

Tooling used: `web.run` (search + open).

The closest (and, I think, exact) primary-source match is Newton's long letter to Henry Oldenburg dated 24 October 1676 (the "Epistola posterior" sent for Leibniz). In the normalised Newton Project transcription, Newton explicitly mentions Barrow as the intermediary who sent Collins a "compendium" of Newton's methods, and then says that Collins repeatedly urged him to publish. Newton then adds that, about five years earlier (1671), "with friends persuading" he had decided to publish his optical treatise and (re)wrote a mathematical tract to publish alongside it; but the controversy triggered by his earlier telescope/light letter produced "frequent interruptions" and deterred him—famously, he says he reproached himself for “catching at a shadow” and thereby losing his peace. newtonproject.ox.ac.uk

That same episode is later used by Newton himself as a retrospective explanation for delayed publication. In Newton’s drafts to Des Maizeaux (post-Leibniz), he writes (in effect): if asked why he did not publish the “Book of Quadratures” sooner, it was for the same reason he did not publish the theory of colours sooner, and that he already gave the reason in the 24 Oct 1676 letter; he also reiterates that he was “upon a design of publishing” in 1671 but “forbore” until 1704. newtonproject.ox.ac.uk+1 A separate Newton Project fragment on the dispute notes (third-person summary) that the quadratures book had been “some years before in the hands of some of his friends & thereby he was induced to publish it,” again linking “friends” with the publication timing. newtonproject.ox.ac.uk

If what you remembered was specifically “two friends/colleagues” in Newton’s own narration: in the 24 Oct 1676 letter Newton names Barrow and Collins in immediate succession, in exactly the “manuscript transmitted by Barrow → publication urged by Collins” storyline, and then attributes the actual deterrent not to them but to the objections/controversy that followed his optical correspondence. newtonproject.ox.ac.uk

27. User

i” “De Analysi”, y su tratado sobre las fluxiones. Barrow y Collins jugaron un papel en disuadirlo de publicarlo” but I need a primary source or at least enough authority. I think the point is that Newton recall of the episode is that it first delayed the publication and was induced to publosh by the few people that knew the treatises. The piece that lacks is that if it was not publoshed from first sublision is because the friends feedback induced newton to do some rework. And a lot of known rework is in the part of the angular momentum under central forces.

28. Assistant

Herramienta usada: `web.run` (búsquedas y apertura de transcripciones/ediciones digitales).

En las fuentes primarias que he podido verificar en abierto, la frase “Barrow y Collins jugaron un papel en disuadirlo de publicarlo” sólo es sostenible (en sentido estricto) para Collins y sólo en un momento concreto; para Barrow, lo que aparece es más bien lo contrario: actúa como intermediario para difundir el manuscrito, y su correspondencia con Collins (tal como se conserva/edita) presenta el envío y la valoración entusiasta del texto, no una disuasión.

La pieza más “dura” (primaria) que sí respalda una disuasión/dilación explícita es una carta de John Collins a Newton del 5 de julio de 1671. Collins le dice que no ve “gran razón” para precipitar la publicación (“no great reason to overhasten the publication of your thoughts”), y le recomienda usar su propio criterio. newtonproject.ox.ac.uk Esto encaja con tu necesidad: no es sólo que Newton fuese reservado; aquí hay un corresponsal de primera línea diciéndole que no corra.

La fuente primaria más importante para reconstruir el episodio "publicar / re-trabajar / retraerse" es Newton mismo, en la carta a Oldenburg del 24 de octubre de 1676 (la "Epistola posterior", escrita para ser comunicada a Leibniz). Newton afirma (i) que el compendio de sus series se comunicó a Collins a través de Barrow, (ii) que Collins "no dejó de sugerir" que lo hiciera público ("non destitit suggerere ut hæc publici juris facerem"), y (iii) que "hace cinco años" (1671) "por persuasión de amigos" decidió publicar un tratado de óptica y colores y, en ese contexto, retomó las series y escribió un tratado para publicarlo conjuntamente; pero las controversias y objeciones sobre su carta óptica le "disuadieron" del plan ("me prorsus a consilio deterruerunt"). newtonproject.ox.ac.uk+1 En otras palabras: el freno más explícito que Newton menciona no es Barrow/Collins, sino el coste social-institucional de las disputas (y su aversión a ellas), que le empuja a replegarse y, a la vez, a reorientar el trabajo.

Sobre Barrow: lo que se cita una y otra vez (y está anclado en las cartas de 1669 tal como aparecen reproducidas en la tradición editorial del Commercium Epistolicum y en historias críticas) es que Barrow anuncia a Collins que un "amigo" le ha traído papeles muy generales, luego le envía el manuscrito, y más tarde revela el nombre de Newton y autoriza a compartirlo con Brouncker. Un compendio moderno muy usado reproduce literalmente dos pasajes: el de 20 julio 1669 ("A friend of mine here…") y el de 20 agosto 1669 ("I am glad my friend's paper giveth you so much satisfaction… his name is Mr Newton… you may impart the papers… to my Ld Brouncker"). nzdr.ru+2en.wikisource.org+2 Eso muestra difusión controlada, no disuasión.

La "pieza que falta" que describes (no se publica porque el feedback de amigos obliga a re-trabajar) se puede sostener con un encadenamiento prudente de evidencias, pero con matices. Lo verificable es: (a) Collins aconseja no precipitarse en 1671. newtonproject.ox.ac.uk (b) Newton declara que, al planear publicar (por persuasión de amigos), reabrió el tema y escribió un tratado más amplio (lo que sugiere re-elaboración respecto del compendio de 1669). newtonproject.ox.ac.uk+1 (c) Newton atribuye el abandono del plan editorial, de forma explícita, a la experiencia de controversia tras su carta óptica, no a una crítica matemática concreta de "dos amigos". newtonproject.ox.ac.uk Con eso se obtiene una narrativa sólida: hubo consejo de cautela desde Collins; hubo intención de publicar inducida por "amigos" (plural) que detonó re-trabajo; y el factor decisivo de retirada fue el contexto polémico.

Sobre tu frase concreta "Barrow y Collins lo disuadieron": tal como queda en las fuentes consultadas, lo más ajustado sería algo como "Collins le recomendó no apresurar la publicación (1671), mientras que Barrow actuó como canal de comunicación del manuscrito (1669); Newton más tarde atribuyó su retirada del plan editorial al coste de las controversias tras su carta óptica (1676)". Eso preserva lo que necesitas (hay 'disuasión' documentable) sin forzar a Barrow a un papel que, en lo que aparece editado y citable, no desempeña.

No he localizado, con búsqueda en abierto, una cita primaria donde Newton diga literalmente "two friends" en este contexto (De Analysi/fluxiones/publicación). Lo que sí aparece en Newton (1676) es la formulación latina "suadentibus amicis" (amigos persuadiendo), sin cardinalidad. newtonproject.ox.ac.uk

29. User

review newton project and Commertium Epistolicum for specifics about the part on the movement under central forces and how the areas need to be reduced to infinitesimals. We have discussed elsewhere how these areas, for central potentials, are essentially proportional to dE dt, and we want to know concrete references of Newton and colleagues about the process of making such quantities negligible or why not. We know that Nature choose not to allow them to zero :-)

30. Assistant

In Newton's Principia the reduction of "areas" (and the other geometric magnitudes used in orbital dynamics) to something like infinitesimals is explicitly a limit procedure, not a commitment to indivisible geometric atoms.

Newton lays this out in Book I, Section I (the "method of first and last ratios"). In the Scholium following Lemma XI, he says he avoided the "method of indivisibles" because its hypothesis "seems somewhat harsh", and instead reduced the demonstrations to "the first and last sums and ratios of nascent and evanescent quantities", i.e. "the limits of those sums and ratios"; and if he speaks as if quantities were "made up of particles", he means "evanescent divisible quantities … always the limits of sums and ratios." Project Gutenberg+1 He then makes the key meta-point for your "uncuttable vs indefinitely cuttable" concern: when he says "least, evanescent, ultimate" he does not mean a determinate minimal magnitude; he means quantities "always diminished without end." Project Gutenberg

Where this enters central-force motion is Book I, Section II, Proposition I (the area law). Newton approximates the curved orbit by a polygon whose vertices are reached in equal time steps, with a centripetal "impulse" at each vertex; in each equal time interval one gets an equal triangle of swept area, so the summed swept areas are proportional to elapsed time. Then the "continuously acting" centripetal force is obtained by the explicit limiting step: "let the number of those triangles be augmented, and their breadth diminished in infinitum … their ultimate perimeter … will be a curve line", so the force "will act continually", while the areas remain proportional to times. Project Gutenberg+1 This is Newton's concrete mechanism for "reducing areas": define the areal velocity as a limit of polygonal area increments as the time-step goes to zero, justified by the preceding limit-lemmas and the "not indivisibles, but limits" Scholium.

Project Gutenberg+1

The Newton Project material you pointed to (and the Commercium Epistolicum context) is useful because it records Newton being very explicit about when he permits "negligible" terms and what that means operationally. In the Newton Project's "An account of the Commercium Epistolicum" (a text known to be Newton's), he distinguishes (i) demonstrations, where the time-moment o is treated as "indefinitely (not infinitely) small", the work is done in finite figures "without any approximation", and only after the computation is complete one "supposes the moment o to become infinitely small & vanish into nothing" to read off "last ratios"; versus (ii) investigations, where he may "suppose the moment o & the figures to be infinitely little" and use approximations (he gives the example "putting the arch & its chord sine & tangent equal") and even omit writing the o. newtonproject.ox.ac.uk+1 This speaks directly to your "why make such quantities negligible or why not": "negligible" is not an ontological claim about nature containing literal zero-sized sectors; it is a statement about a controlled limiting operation (in proofs), and about heuristic small-quantity approximations (in discovery work), with Newton openly separating the two modes. newtonproject.ox.ac.uk+1

On the specific "movement under central forces" aspect: Proposition I is the forward direction (centripetal $\Rightarrow$ equal areas in equal times), and Proposition II is the converse (equal areas in equal times $\Rightarrow$ acceleration directed to the centre). Proposition I is also where Newton ties the polygonal "impulsive" model to the continuous force by an explicit passage to the limit using Corollary 4 and the "ultimate perimeter is a curve" remark. Project Gutenberg+1 That limiting move is the exact historical locus where "indefinite cutting" (finer time partitions; thinner triangles) becomes the mathematically defined object Newton needs (areal velocity as a limit).

About your remark that, "elsewhere", these areas are "essentially proportional to $dE\,dt$": Newton's swept-area triangles are configuration-space areas, and in modern Hamiltonian language they are tied most directly to the cyclic angle θ and its conjugate momentum $p_\theta = L$. In action terms, the relevant piece is $p_\theta\,d\theta$ (or, equivalently, the rotational Noether charge), whereas the extended-phase-space term involving energy is $-H\,dt$; for an autonomous central potential, $H = E$ is conserved along solutions, so the (E, t) piece is structurally different from the configuration-space areal law even though both live naturally inside the same symplectic/action framework. The clean Newton-to-modern dictionary is: "areas proportional to times" $\Leftrightarrow$ dA/dt constant $\Leftrightarrow$ $L = 2m\,dA/dt$ conserved, and the point at which "areas become negligible" is precisely the $\Delta t \to 0$ limit Newton formalises as "ultimate ratios," not an assertion of a smallest geometric grain. Project Gutenberg+2Project Gutenberg+2

31. User

so can we fit all of this in our Narrative “from newton to the path integral”? Do a separate note, self contained

32. Assistant

Note: From Newton’s “first and last ratios” to the path integral

Newton’s core move, in the Principia, is to make “indefinite cutting” into a mathematically controlled operation: one refines a partition, proves statements at the finite (polygonal) level, and then takes a limit that is explicitly not a claim about indivisible ultimate magnitudes. In the Scholium to Lemma XI he insists that “ultimate ratios” are “limits towards which the ratios of quantities decreasing without limit do always converge,” and that when he speaks of “least” or “evanescent” quantities one must not imagine “quantities of any determinate magnitude,” but ones “diminished without end.” en.wikisource.org+1

This philosophy becomes concrete in the first central-force theorem: Proposition I (area law). Newton divides time into equal parts and replaces a continuous centripetal force by a succession of instantaneous “impulses” at the vertices of a polygonal path. On each time-step the radius vector sweeps out a triangle; because the impulses are directed to the centre, successive triangles have equal area, hence equal areas in equal times. The continuous force is then obtained by a limiting refinement: “let the number of those triangles be augmented, and their breadth diminished in infinitum; … their ultimate perimeter … will be a curve line,” so the force “will act continually,” and the swept areas remain proportional to times. Project Gutenberg This is the precise “reduce areas to infinitesimals” step: not “areas go to zero in Nature,” but “area increments are defined as the limit of a refinement procedure,” justified by the limit-lemmas.

A second, even more explicit, Newtonian datum is the methodological separation between (i) demonstrations with a time-moment treated as “indefinitely (not infinitely) small,” carried out with finite figures and only at the end “supposing the moment … to vanish,” and (ii) investigations where one treats the moment and the figures as “infinitely little,” uses approximations (e.g. arc/chord/sine/tangent taken equal), and may even omit writing the small parameter. newtonproject.ox.ac.uk This is structurally the same distinction modern physics makes between controlled limits (definitions/theorems) and “small-parameter heuristics” (approximations, asymptotics, perturbation theory).

At this point the bridge to an action-based formulation is straightforward. In modern notation, the area law is equivalent to conservation of angular momentum L: for a particle of mass m, the areal velocity is $\dot{A} = \frac{1}{2}r^2\dot{\theta}$, and $L = mr^2\dot{\theta} = 2m\,\dot{A}$. Newton’s triangles are a discrete-time definition of $\dot{A}$, and the limit “triangles augmented, breadth diminished” is the passage $\Delta t \to 0$ giving a well-defined conserved quantity for the continuous dynamics. Project

Gutenberg In an action formalism, the same conservation is Noether's theorem applied to rotational symmetry of the Lagrangian $L(q, \dot{q}) = T(\dot{q}) - V(r)$: the cyclic coordinate θ implies $p_\theta = L$ constant, and "equal areas in equal times" is the geometric shadow of that symmetry.

Now replace Newton's polygonal orbit by a time-sliced trajectory in configuration space. For any path $q(t)$ over $[t_i, t_f]$, the classical action is

$$S[q] = \int_{t_i}^{t_f} \mathcal{L}(q, \dot{q})\, dt,$$

and on a partition $t_i = t_0 < t_1 < \cdots < t_N = t_f$ one has the additivity

$$S[q] \approx \sum_{k=0}^{N-1} \mathcal{L}\Big(q_k, \frac{q_{k+1} - q_k}{\Delta t}\Big)\, \Delta t,$$

exactly analogous to Newton's additivity of swept areas as sums of triangles. Newton's refinement "more triangles, thinner triangles" corresponds to "more time-slices, smaller Δt." The central question becomes: what mathematical object remains stable under arbitrary refinement?

Dirac's 1933 observation is that the Lagrangian/action is the natural invariant controlling time-evolution in quantum mechanics (he motivates this by the action principle and the relativistic invariance of the action). informationphilosopher.com Feynman's 1948 formulation then posits the refinement-stable rule: the contribution of a single path is an exponential of the action "in units of $\hbar$," and amplitudes are sums over alternatives (paths). journals.aps.org The key structural reason the exponential appears is the same "composition under cutting" that already drives Newton's polygon proof: if you cut a time interval into two pieces and demand that the propagator compose (a Markov/semigroup property in amplitude form), then the weight assigned to a history must be multiplicative under concatenation, hence of the form exp(additive functional); the action is precisely the additive functional singled out by the classical variational principle, so $\exp(iS/\hbar)$ is the natural multiplicative refinement-stable weight. journals.aps.org+1

This also clarifies your "we need to refer indefinite cutting back to a scale where we can measure." Newton's "moment o" plays the role of an explicit small parameter whose disappearance is controlled (and whose omission in heuristic work is a conscious approximation). newtonproject.ox.ac.uk In the path integral, $\hbar$ is exactly such a control scale: it renders the phase dimensionless and governs the stationary-phase limit $\hbar \to 0$, in which contributions concentrate near classical stationary points of S. In that precise sense, classical mechanics is recovered not as "the truth before cutting," but as the asymptotic regime of the quantum refinement-stable construction.

Two modern mathematical frameworks package this "refinement plus control scale" idea.

First, strict deformation quantization via the tangent groupoid (Connes' construction, developed further by Landsman and others) formalises quantisation as a continuous deformation from a commutative algebra of classical observables to a noncommutative quantum algebra, with the deformation parameter playing the role of $\hbar$. Landsman explicitly treats Lie groupoids and recovers Connes' tangent groupoid in this context. Springer Nature+1 Here, "inserting more points" becomes "comparing products at different deformation scales" in a single geometric object.

Second, renormalisation theory is what happens when refinement produces divergences rather than a well-defined limit: you introduce a cutoff/scale, define finite renormalised quantities relative to that scale, and demand that changing the scale does not change observables. Connes–Kreimer's Hopf-algebra/Riemann–Hilbert formulation is a precise algebraic expression of "insertions" (counterterms) as the mechanism that restores consistency under refinement. archives.ihes.fr Structurally, it is again Newton's theme: you cut more finely, but you must also transport results back to a fixed "measuring" framework so comparisons make sense.

In this narrative, "Nature chooses not to allow them to zero" can be read as: the mathematical limit $\Delta t \to 0$ is a definition/idealisation (Newton is explicit about this), but the physically meaningful theory may come with a regulator (quantum $\hbar$, field-theoretic renormalisation scales, or potentially a minimal length in quantum gravity scenarios) that prevents treating refinement as a literal physical process beyond some regime. Newton's own discipline is still the right template: do finite constructions, state precisely what is meant by the limiting step, and keep separate the controlled limit (definition/theorem) from the heuristic neglect of small quantities. en.wikisource.org+1