Orthodox quantization of Einstein's gravity: might its unrenormalizability be technically fathomable and physically innocuous?

Steven Kenneth Kauffmann
American Physical Society Senior Life Member

43 Bedok Road
01-11
Country Park Condominium
Singapore 469564
Tel & FAX: +65 6243 6334
Handphone: +65 9370 6583

and

Unit 802, Reflection on the Sea
120 Marine Parade
Coolangatta QLD 4225
Australia
Tel/FAX: +61 7 5536 7235
Mobile: +61 4 0567 9058

Email: SKKauffmann@gmail.com

Abstract

Many physical constants related to quantized gravity, e.g., the Planck length, mass, curvature, stress-energy, etc., are nonanalytic in G at $G = 0$, and thus have expansions in powers of G whose terms are progressively more divergent with increasing order. Since the gravity field's classical action is inversely proportional to G, the path integral for gravity-field quantum transition amplitudes shows that these depend on G only through the product $\hbar G$, and are nonanalytic in G at $G = 0$ for the same reason that all quantum transition amplitudes are nonanalytic in $\hbar$ at $\hbar = 0$, namely their standard oscillatory essential singularity at the classical 'limit'. Thus perturbation expansions in powers of G of gravity-field transition amplitudes are also progressively more divergent with increasing order, and hence unrenormalizable. While their perturbative treatment is impossible, the exceedingly small value of $\hbar G$ makes the semiclassical treatment of these amplitudes extraordinarily accurate, indeed to such an extent that purely classical treatment of the gravity field ought to always be entirely adequate. It should therefore be fruitful to couple classical gravity to other fields which actually need to be quantized: those fields' ubiquitous, annoying ultraviolet divergences would thereupon undergo drastic self-gravitational red shift, and thus be cut off.

Introduction

Gravity occupies a profoundly fundamental place in physical theory by virtue of its source being the energy-momentum tensor, whose presence accompanies *any* physical phenomenon whatsoever. This *universality* implies that the source of the gravitational field even includes a contribution from *itself*, which is the origin of its *nonlinearity*. Gravity's predominantly *negative* long-range field potential energy response to its sources can enable it to overwhelm and suppress those source components that happen to be sufficiently concentrated and strong, even plunging them into "black holes" in extreme cases, but its *fundamental* coupling strength is *smaller* than that of the other known forces to a mind-boggling degree: the mutual gravitational attraction between two even neutron-rich nuclei (e.g., tritons) is still weaker than their mutual Coulomb repulsion by a staggering factor of over 10^{35}.

As Einstein's gravitational theory and the quantum theory of dynamics are both crowning achievements of twentieth-century theoretical physics, it is entirely natural to try to combine them. The multitude of similarities between Einstein's gravity theory and Maxwell's electrodynamics, together with the qualifiedly successful (at least after 'renormalization') quantized treatment of the latter in the approximation context of a perturbation expansion in powers of the coupling strength, strongly suggests handling the quantization of gravity analogously, *especially* in view of the fact that its coupling strength is even *vastly weaker* than that of electrodynamics. Unfortunately, however, as has been realized for well over half a century [1], the perturbation expansions in powers of the gravitational coupling strength G of transition amplitudes of the canonically quantized gravitational field yield infinities whose severity increases progressively with order, which precludes 'renormalization'. Since 'renormalization' consists of a certain class of prescriptions that are imposed *atop* a perturbatively treated quantum field theory for the express purpose of *shunting aside* unwelcome infinities that it has produced ("sweeping them under the rug", in Feynman's blunt phrase), but *fails* to point to *any physical mechanism* which disallows their existence, its *inapplicability* to canonically quantized gravity is obviously *not of itself* adequate reason to conclude that that theory must be discarded—*especially* in light of the *robustness* of the twin pillars, namely Einstein's gravity and quantum dynamics, upon which it rests.

That notwithstanding, it was promulgated as "pragmatically motivated" dogma in the late 1960's that *all* unrenormalizable quantum field theories were henceforth to be regarded as being "beyond the pale" on the grounds that if the *perturbation expansion* for a quantum field theory was *unviable*, it simply was *not conceivable* that any *practicable alternate means* of satisfactorily extracting its results could exist or be developed. This "pragmatic", but obviously *not physically based*, dogma, which *also* directly flouts Einstein's admonition not to cravenly limit research efforts to "the thinnest part of the board", then drove a single-minded effort to create renormalizable theories at essentially any cost: Occam's razor and its companion guidelines of conservatism, continuity, and known empirical support for undertaking modification of physical theory were upended as fields were abruptly swapped for strings, the four dimensions of space-time were simply increased to "whatever it takes", and fermionic Noether currents which anticommute rather than causally commute at spacelike separations were ascribed with physical existence—all done *not* to accommodate existing *physical knowledge* but simply in brainstorming support of a headlong tunnel-vision effort to create theoretical structures whose perturbative infinities are "under control" in a particular preconceived sense.

It was not *inevitable*, however, that the catastrophic failure of the perturbation expansion in powers of G for canonically quantized gravity need have so abjectly contributed to the above-described departures from the guidelines which have well served theoretical physics research for centuries. The keen sense of frustration which arose from that failure *ought* to have been tempered by the realization that theories can be qualitatively probed with simple, if somewhat blunt tools that at least have the virtue of being entirely *nonperturbative*. One such rough tool is the time-honored one of dimensional analysis—this seems almost made to order for the canonically quantized gravitational field with its highly suggestive "Planck trio" of applicable constants, namely G, c, and $\hbar$. It is well-known that from these universal constants Planck entities having the dimensions of mass, length, and time are readily constructed, namely the Planck mass $\sqrt{\hbar c/G}$, the Planck length $\sqrt{\hbar G/c^3}$, and the Planck time $\sqrt{\hbar G/c^5}$, which, in turn, are basic building blocks of further Planck entities that have any dimension one might wish to nominate (with the exception of dimensionlessness). Such entities are of particular interest to us for extremely *small* values of G, which accord with both G's actual *physical* value *and* with the point $G = 0$ about which the disastrous perturbation expansion of canonically quantized gravity is made. We are immediately struck by the fact that *all three* of the basic Planck entities given above, notwithstanding that they are obviously *perfectly well-defined* (indeed elementary), nevertheless have perturbation expansions in powers of G whose terms become progressively more severely divergent with increasing order—which is precisely the *same* "perturbatively catastrophic" behavior that has for so many decades been regarded as the death knell of gravity's straightforward canonical quantization! It now becomes apparent that canonically quantized gravity *may* have been *prematurely* written off all those decades ago on entirely inadequate grounds—any otherwise well-defined quantity that, like the above three basic Planck entities, is *nonanalytic* in G at the point $G = 0$, will normally have just such a "catastrophic" perturbation expansion in powers of G. There are, indeed, many *other* physically interesting (and perfectly well-defined) Planck entities that *also* fall precisely into this category, including, inter alia, the Planck curvature $c^3/(\hbar G)$, Planck energy density $c^7/(\hbar G^2)$, Planck acceleration $\sqrt{c^7/(\hbar G)}$, and Planck wave number $\sqrt{c^3/(\hbar G)}$. We now need to inquire into why $G = 0$ is a point of nonanalyticity of so many physical entities which flow from canonically quantized gravity (conceivably including, it now seems not implausible, its transition amplitudes).

Small G and the classical limit

We see that as G tends toward zero, the Planck wave number just mentioned increases without bound, which at least *suggests* that $G \to 0$ drives the canonical quantization of gravity toward its *classical limit.* Furthermore, in line with what would be expected of a wave number marker for the classical limit, the Planck wave number *also* increases without bound as $\hbar$ tends toward zero—in fact, the Planck wave number depends on G through the *product* $(\hbar G)$. It is, of course, well-known that quantum theories behave in an extremely nonsmooth asymptotic fashion as they are driven to their classical limit (e.g., when $\hbar \to 0$), so we now glimpse fragments of an argument as to why many physical entities which flow from canonically quantized gravity might be expected to be nonanalytic in both $\hbar$ at $\hbar = 0$ and in G at $G = 0$. In order to present that argument in a clear, systematic way for this theory's transition amplitudes in particular, we shall first review the reasons why quantum transition amplitudes *in general* are normally nonanalytic in $\hbar$ at $\hbar = 0$. We shall also briefly discuss a path-integral-based stationary phase asymptotic semiclassical expansion approach to quantum transition amplitudes which is valid as $\hbar \to 0$.

General quantum transition amplitudes such as $\langle\psi_f|\exp(-i\widehat{H}(t_2 - t_1)/\hbar)|\psi_i\rangle$ can be rewritten in terms of the eigenspectrum of $\widehat{H}$ as $\sum_E \langle\psi_f|E\rangle \exp(-iE(t_2 - t_1)/\hbar)\langle E|\psi_i\rangle$, and therefore will obviously almost *always* be *nonanalytic* in $\hbar$ at $\hbar = 0$. A systematic treatment of such transition amplitudes as $\hbar$ approaches their point of nonanalyticity at $\hbar = 0$ can be developed from their *path integral expression*,

$$
\begin{aligned}
&\langle\psi_f|\exp(-i\widehat{H}(t_2 - t_1)/\hbar)|\psi_i\rangle = \\
&\quad \int d^n\mathbf{q}_2\, \langle\psi_f|\mathbf{q}_2\rangle \int d^n\mathbf{q}_1\, \langle\mathbf{q}_1|\psi_i\rangle \times \\
&\quad \int \mathcal{D}\{\, (\mathbf{q}(t), \mathbf{p}(t)) \;\mid\; t\,\epsilon\,[t_1, t_2], \mathbf{q}(t_1) = \mathbf{q}_1, \mathbf{q}(t_2) = \mathbf{q}_2\} \times \\
&\qquad \exp(i\textstyle\int_{t_1}^{t_2} dt\, (\dot{\mathbf{q}}(t)\cdot\mathbf{p}(t) - H(\mathbf{q}(t), \mathbf{p}(t)))/\hbar),
\end{aligned}
$$

which again makes their *nonanalyticity* in $\hbar$ at $\hbar = 0$ manifest. As $\hbar \to 0$, however, the pure phase integrand of the path integration oscillates increasingly rapidly except in an increasingly small neighborhood of the path which renders it *stationary.* This stationary path is readily seen to satisfy $\dot{\mathbf{q}}(t) = \nabla_{\mathbf{p}(t)} H(\mathbf{q}(t), \mathbf{p}(t))$ and $\dot{\mathbf{p}}(t) = -\nabla_{\mathbf{q}(t)} H(\mathbf{q}(t), \mathbf{p}(t))$, which are, of course, Hamilton's *classical* equations of motion, subject to the end constraints imposed on the paths integrated over, namely that $\mathbf{q}(t_1) = \mathbf{q}_1$ and $\mathbf{q}(t_2) = \mathbf{q}_2$. The *full* mathematical development of this sort of stationary phase approximation to an integral over a pure phase integrand which is driven by a parameter to oscillate increasingly rapidly is well known to produce an *asymptotic expansion* of the integral in that parameter. To be sure, this stationary phase asymptotic expansion technology is not normally presented in the context of *path* or *functional* integrals, but all the needed concepts and theorems, including polynomials, derivatives, the Taylor expansion, Gaussians, and analytic integration of the products of polynomials with Gaussians, are readily extended from the mathematics of multivariate functions to that of functionals. This stationary phase asymptotic expansion technique applied to the quantum path integral, with $\hbar \to 0$ being the driver of the increasingly rapid pure phase integrand oscillations, provides the systematic *semiclassical* expansion of *quantum transition amplitudes*—a methodology of considerable promise for strong interactions, which *has not yet* been tapped. It is to be cautioned, however, that the approach may be a formidable consumer of computational resources, as it, in principle, requires all the classical paths generated by *every possible pair* of end constraints.

For particle dynamics, as it is treated above, the classical action functional is,

$$\textstyle\int_{t_1}^{t_2} dt\, (\dot{\mathbf{q}}(t)\cdot\mathbf{p}(t) - H(\mathbf{q}(t), \mathbf{p}(t))).$$

For the gravitational field, however, it turns out that the classical action functional is *inversely* proportional to G, i.e., it equals the constant factor $(-c^4/(16\pi G))$ times the curvature scalar integrated over generally invariant space-time [2]. Therefore for the canonically quantized gravitational field, $(16\pi\hbar G)/c^4$ may be expected to play a role analogous to that played by $\hbar$ *alone* in quantized dynamics generally. Thus we may *indeed* expect canonically quantized gravitational field theory to be *nonanalytic* in G at $G = 0$, as well as in $\hbar$ at $\hbar = 0$, and its perturbation expansion in powers of G to be a disaster. However, since

$(16\pi\hbar G)/c^4$ is notable for its extreme smallness, we may *also* expect the stationary phase *semiclassical* asymptotic expansion of canonically quantized gravitational field theory to produce extraordinarily accurate results indeed. In fact, this approach may be expected to yield *such* good results that simply resorting to the purely *classical* gravitational field ought to be entirely adequate.

Another way to appreciate the predominantly *classical* character of the canonically quantized gravitational field is to consider the detectability of individual gravitons. The extreme weakness of the gravitational coupling strength G makes individual gravitons essentially undetectable *unless* they have extraordinarily high energy. Any process capable of emitting such gravitons would almost certainly involve extremely strong gravitational fields in the immediate vicinity of their region of emission, fields which would tend to gravitationally red shift those very gravitons to lower energy. It thus might be problematic for gravitons energetic enough to be individually detectable to actually be available. Furthermore, the total phase space for a graviton to decay into two gravitons that both travel in its original direction is nonvanishing (albeit for three or more gravitons it does vanish). This decay is suppressed both by the weakness of G and by a d-wave orbital angular momentum barrier, but its rate should rise strongly with energy, thus also depleting the availability of gravitons energetic enough to be individually detectable. (Interestingly, two-photon decay of a photon is ruled out in spin $\frac{1}{2}$ quantum electrodynamics by Furry's theorem.) Finally, the distinctly *macroscopic* magnitudes of the Planck mass (which at nearly 22 mcg is comparable to that of a small puntuation mark cut out of a glossy page), the Planck momentum (which at over 23,000 g km/hr is comparable to that of a bullet), and the Planck energy (which at over 540 kWh would supply a household for many days) hardly suggest a significant need to take *quantum* corrections to the gravitational field into account.

It is quite interesting that in the course of pondering the *quantization* of the gravitational field, one is driven to the conclusion that this endeavor is largely *unnecessary.* Einstein even more fiercely opposed the quantization of gravity than he opposed the quantum theory generally—as fate would have it, quantized gravity theory *itself* turns out to be disinclined to disagree with him to any *significant* extent. The dominantly *classical* character of *universal* gravitation turns out to provide a deep validation of the Copenhagen interpretation of quantum theory: in principle any 'pure' quantum state is necessarily (albeit usually *extremely* weakly!) coupled to the *universal* gravitational field, which is effectively a "classical observer" that must in due course bring about the "collapse" of its quantum coherence. What a pity that Einstein never confronted the reverberating ironies implicit in this line of thought! (In actual practice, of course, the vastly more strongly coupled electromagnetic field is *very* much more likely to play this "quantum coherence collapsing" role, but electromagnetism is in principle neither universally coupled nor necessarily dominantly classical.)

Dominantly classical gravity in a quantum world

This dominantly classical character of gravitation sorely needs, however, for the purposes of theoretical physics, to be appropriately conjoined with the markedly *quantum* characteristics which so many other physical phenomena, such as electromagnetism, can manifest. A straightforward formal approximation technology for accomplishing this has been proposed by Boucher and Traschen [3], wherein a hybrid partially quantized Hermitian density operator is taken to be merely a *function* of those phase-space variables (e.g., the gravitational ones) which are to be left as unquantized c-numbers. The same hybridization applies to the Hamiltonian and other dynamical variables of interest. The equation of motion of the hybrid density then is taken to involve a natural hybrid commutator-cum-Poisson bracket of that density with the hybrid Hamiltonian (the factors of the Poisson bracket part of this hybrid bracket must, of course, be ordered so as to ensure the hybrid bracket's Hermiticity). Albeit straightforward and natural, this approach *cannot* be regarded as the realization of some manner of quantum/classical dynamical 'subtheory', because there is no *guarantee* that its hybrid density remains *positive* as time evolves, as pointed out by Boucher and Traschen [3]. Therefore this approach definitely falls in the category of being an *approximation* technology with the characteristic property of being subject to manifest failure *if* applied well *beyond* its appropriate scope. It is in fact impossible for classical degrees of freedom to *strictly* maintain their quintessential *inherent determinism* once they are permitted to *interact* with *quantum* degrees of freedom that are *not* bound by such determinism.

In a much less methodical vein, it is instructive to try to tease out in a rough, qualitative manner some of the salient implications of the predominantly classical gravitational field for quantum phenomena. A driving goal of high energy particle physics is to resolve natural phenomena at ever smaller spatial scales. To resolve an object of extremely short length l, we need quanta of of momenta around $\hbar/l$ or larger to be absorbed and then reemitted (i.e., scattered) by that object. As we suppose l to have ever smaller values, we may safely assume such quanta to be ultrarelativistic, i.e. photon-like. Upon the quantum's absorption, the object of length l will have an energy of at least $\hbar c/l$, which will generate a dimensionless gravitational potential of around $-\hbar G/(c^3 l^2)$ at its extremities, and this, in turn, would tend to reduce the momentum of the *reemitted* quantum by the factor $(1 - (\hbar G)/(c^3 l^2))$ because of gravitational redshift. (We have used very crude Newtonian-like gravitational guesstimates here—these do not include nonstatic corrections nor take account of the needed self-consistency iterations.) The thrust of this crude exercise is clearly that a short *enough* length will be very difficult to resolve, as the requisitely high momentum quantum will, after absorption into the object of this length, tend to redshift its *reemitted* counterpart toward extinction. So the very *means* of resolving a small enough region has the side effect of redshifting *itself* (upon reemission) toward nonexistence—the *necessarily* energetic probe drives its tiny target in the direction of becoming an invisible 'black hole'. The above expressions strongly suggest that this effect will indeed 'bite' when the target length l is significantly less than $\sqrt{\hbar G/c^3}$, the Planck length. Furthermore, if it is not possible to resolve length intervals significantly smaller than the Planck length, it is rather clear that 'stopwatches' which reliably record time intervals significantly shorter than the Planck time $\sqrt{\hbar G/c^5}$ cannot be constructed either. Thus we would expect space-time below the Planck scale to be not so much a 'quantum foam' as *intractably opaque*. A less crude, more detailed exposition of this argument is to be found in Ng and van Dam [4].

In many quantum field theories, such as quantum electrodynamics, the presence of virtual particles of *arbitrarily large* energy can be a source of mathematical divergences known as the 'ultraviolet catastrophe'. A virtual particle of very high energy E, however, can only exist for a very short time of order $\hbar/E$ before it must be reabsorbed. Hence its evanescent presence will have been confined to a region whose length is around $\hbar c/E$. Again resorting to very crude Newtonian-like gravitational guesstimation, we obtain that it will have given rise to an average dimensionless gravitational potential of around $-GE^2/(3\hbar c^5)$ in that region, which roughly reduces its energy from E to $E(1-GE^2/(3\hbar c^5))$. We would therefore expect virtual particle energies E to be limited to being not greatly larger than the Planck energy $\sqrt{\hbar c^5/G}$, or else the virtual particle tends to disappear entirely into a black hole of its own making. This gravitational limit to virtual particle energies in turn yields a gratifying natural cutoff for the 'ultraviolet catastrophe' divergences—the correctness of the basic thrust of this universal natural cutoff idea for the 'ultraviolet catastrophe' divergences has very recently received some detailed support in the case of quantized scalar fields via a model wherein each virtual scalar particle is subjected to the gravitational fields produced by its virtual companions [5]. Crudely inserting such a Planck-scale cutoff into the divergent electromagnetic correction to the electron's bare mass in QED yields a result roughly comparable to the bare mass itself. For a charged spin 0 particle, however, the electromagnetic mass contribution would be within about an order of magnitude of the Planck mass, independent of the particle's bare mass. Perhaps not so coincidentally, the known charged spin 0 particles are all believed to be *composed* of charged spin $\frac{1}{2}$ particles (i.e., quarks). In contrast, the divergent apparent corrections to the electron's *charge* must be purely *artifacts* of the unphysical 'ultraviolet catastrophe', since even *virtual* processes are *just* as formally constrained to *conserve charge* in QED as they are to respect gauge invariance. Gauge noninvariant infinities in QED are recognized as unphysical artifacts that need to be *subtracted out*, so charge nonconserving infinities must be handled *likewise*, but unfortunately they were *enshrined* very early on as infinite 'charge renormalizations' in flawed analogy with "effective charge reductions" *within* polarizable media.[1] Once classical gravitation and its consequent virtual particle energy cutoff have been properly incorporated into QED such that charge conservation continues to be respected, there can be no alternative but for infinite 'charge renormalizations' to thereupon *vanish identically*, which is precisely what calculations confirm [6].

[1]This very useful way to view the electrostatic interaction of two probe charges *within* a polarizable medium obviously *falls away* once the probe charge separation becomes *much larger* than the dimensions of the polarizable specimen being considered—as their separation $\to\infty$, that separation squared times the electrostatic force on at least one of the probe charges approaches the *same* limit as in the *absence* of the polarizable specimen. In QED an infinite 'charge renormalization' artifact *already* occurs for a *single* virtual electron-positron pair that has a vanishingly small spacelike total four-momentum. Because this virtual pair is "off mass shell" by at least *two electron masses*, and therefore must annihilate *within the corresponding time* $\hbar/(2m_e c^2)$ to recreate the virtual photon whose dissociation gave it birth, its effective "polarizable specimen dimension" can encompass no more than about half an electron Compton wavelength. The minuscule extent of this "virtual pair polarizable specimen"—no larger than an electron's *own* quantum relativistic size—makes it obvious that this speck *cannot*, even during its fleeting existence, have effected *any* 'charge renormalization', let alone an *infinite* amount, so that phenomenon clearly *must* be an ultraviolet divergence *artifact*.

References

[1] S. Weinberg, *Gravitation and Cosmology* (Wiley, New York, 1972), p. 289.

[2] Ibid., p. 364.

[3] W. Boucher and J. Traschen, Phys. Rev. D **37**, 3522 (1988).

[4] Y. Jack Ng and H. van Dam, Mod. Phys. Lett. A **9**, 335 (1994).

[5] R. Casadio, www.arXiv.org, 0806.0501v2 [hep-th] (2008).

[6] S. K. Kauffmann, www.arXiv.org, hep-th/9505191 (1995).