Gauge and Gravitational Anomalies (Schwartz Ch.30)

Marcus Berg, January 24, 2025

1 Global anomaly

Global anomalies are not harmful. The measured decay $\pi^0 \to 2\gamma$ proceeds through an anomaly, Schwarz (30.11), and the Feynman diagram calculation gives the width (lifetime = $\hbar/\Gamma$, cf. p.105):

$$\Gamma(\pi^0 \to \gamma\gamma) = \frac{\alpha_e^2}{64\pi^3}\lambda^2 \frac{m_\pi^3}{m^2} = \frac{\alpha_e^2}{64\pi^3}\frac{m_\pi^3}{F_\pi^2} = 7.77 \text{ eV} \qquad \text{(experiment: } 7.73 \pm 0.16 \text{ eV)}$$

This process is nice for many reasons: there is a direct measurement of something that can be approximated by a relatively short calculation, and it is one of the few low-energy processes that it sensitive to the number of colors of quarks. So as Schwarz discusses, if there had not been three colors, the prediction would have been 1/3 of this, which is clearly ruled out with the given experimental precision!

Srednicki writes that *"such a low-energy process seems like a dusty relic to most of today's students"*, but it keeps being relevant. For example it has been used more recently as a sledgehammer to attack new approaches to quantum field theory; see "conceptual questions" below.

Another reason pion physics is instructive is in the recent work on alternative theories of *gravity*! In QCD, the pion π^0 is a Goldstone boson, generated by spontaneous symmetry breaking of a symmetry that QCD has when the quarks are taken to be massless (chiral symmetry). It is easier to work with scalars than the longitudinal components of massive gauge bosons, or of *gravitons*. In fact in recent work, π is the common notation for the longitudinal component of the graviton – if it had a mass. Experiments like LIGO constrain this mass, but to constrain it you have to first compute with it [18]. (Some of the recent developments were made in Stockholm!)

2 Gauge anomaly

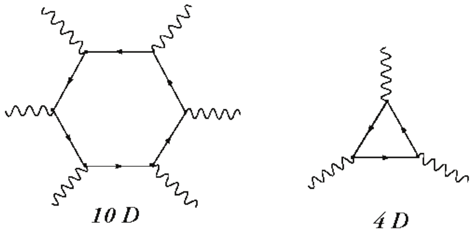
I find Harvey's 2005 lectures very nice. His statement what a gauge anomaly is is fairly standard, but particularly compact: *"gauge symmetries are not symmetries in the conventional sense of symmetries that act on the configuration space and lead to identical physics. Rather, they are redundancies in our description of the physics when we work in the space of gauge fields rather than its quotient by gauge transformations. Anomalies in a redundancy would not be a good thing. More concretely, when we work in the space of gauge field configurations, we need gauge invariance to remove negative norm states from the spectrum and a lack of gauge invariance due to anomalies would lead to fundamental inconsistencies."* [1]

One way to check for this is to take a loop diagram with external gauge bosons, and put one of the external polarizations longitudinal. This means the opposite of transverse, which is $p_\mu e^\mu = 0$, so longitudinal means $e_\mu \propto p_\mu$. Photons with that kind of polarization are the "negative-norm states" Harvey is talking about: the Hilbert-space norm is negative for timelike polarizations.

It is instructive to think about this for general integer dimension D. This is nicely described on p.148 of Green-Schwarz-Witten's string theory book[1]: you can insert a chiral projector P_L (in $D = 4$, $P_L = \frac{1}{2}(1 \pm \gamma^5)$) to implement a specific chirality of fermions running in the loop. If your Feynman diagram gives at least D gamma matrices, the γ^5 from the projector will give a D-dimensional ϵ tensor, that contracts into momenta p_i and polarizations e_i, where $i = 1, \ldots, n$ for an n-point correlation function. Because ϵ is totally antisymmetric, all p_i^μ and e_i^μ must be different to get something nonzero. You might think that therefore you need at least an $n = D/2$-point correlation function to get a nonzero contraction. But only $n-1$ momenta are independent due to momentum conservation,

[1] You might wonder why you need a string theory book to discuss field theory anomalies. Of course you don't, but most field theory books don't discuss general dimension D in any detail. There are a few exceptions, like Peskin & Schroeder occasionally venture out into general D, in fact Michael Peskin did some work in string theory and even wrote some nice lectures on the topic [8].

so actually you need one more, i.e. at least an $n = D/2 + 1$-point function. So for $D = 4$ we have $n \geq 3$ (triangle anomaly), for $D = 6$ we have $n \geq 4$ (box anomaly) and for $D = 10$ we have $n \geq 6$ (hexagon anomaly). Higher-point functions can also be anomalous, for example by embedding the aforementioned diagrams as subdiagrams. These are the possibilities at lowest multiplicity (n, the number of external legs): (picture from [15])



So to check for an anomaly, we take one of these vector polarizations e_i^μ, say e_1^μ, and set it equal to p_1^μ to make it longitudinal, and see if the amplitude vanishes. The above argument shows that if you don't have enough external legs ($n < D/2 + 1$) it will vanish anyway, but the vanishing for low number of external legs is trivial and does not tell you anything about the anomaly.

Another point that is particularly clear from going to general dimensionality is that $D = 4k = 4, 8, 12, \ldots$ is quite different from $D = 4k + 2 = 2, 6, 10, \ldots$. For $D = 4k + 2$, the antiparticle of a left-handed particle is left-handed, so although in $D = 4$ it was possible to have different gauge group representations for left- and right-handed fermions, in $D = 10$ it is possible to have no right-handed fermions at all and still have no gauge anomaly (in contrast with Schwarz 30.2.4). There is now the possibility of a gravitational anomaly (see below). Again, this is nicely described in GSW, now p. 145.

3 Chiral anomaly

As discussed above, in four dimensions, a triangle diagram with a loop of chiral fermions can cause an gauge anomaly. In dimensional regularization this is somewhat tricky; following 't Hooft and Veltman, we have to assign some (reasonable, but at first ad-hoc-looking) rules about traces of gamma matrices. (Some more detail is given in Collins [19], but I'm not sure it really helps.) I will instead follow Schwarz section 30.2 and do this calculation using a formal argument about linear divergences. I will need this identity [5]

$$\mathrm{tr}(\gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \gamma^5) = -4i\epsilon^{\mu\nu\rho\sigma}$$

which is discussed e.g. on the Wikipedia page for gamma matrices. (It's good exercise to check this kind of identity, for example, if you didn't have the γ^5, can you see quickly that you wouldn't get $\epsilon^{\mu\nu\rho\sigma}$?) The result of the triangle loop diagram is basically

$$\partial_\mu J^{5\mu} \propto F_{\mu\nu} (*F)^{\mu\nu} \tag{3.1}$$

where $(*F)$ is the Hodge dual of F. This means that the axial current $J^{5\mu} = \bar{\psi}\gamma^5\gamma^\mu\psi$ is not conserved in the presence of certain gauge fields.

As a side remark, by "certain gauge fields", I meant that this does not affect all possible configurations of gauge fields. In particular, obviously the left-hand side is a total derivative of something, so the integral over all of space of the right-hand side vanishes unless there is some kind of boundary term. Perhaps surprisingly there often is a boundary term, I'll come back to this. (To be clear, I'm not saying it's OK if the theory has some problem only on some configurations of gauge fields, I'm just saying you may not notice for some configurations. The theory should be ideally be meaningful for all possible configurations of gauge fields, and more to the point, all such configurations should be included in the functional integral, as in Schwarz 30.3.)

4 Gauge vs. Chiral anomaly

So we will find an anomaly in the axial current $J^{5\mu}$, i.e. the axial current is not conserved. Even though sometimes the terms are used interchangeably, anomalous nonconservation of a chiral current does not automatically produce a gauge anomaly, i.e. production of longitudinal gauge bosons, in particular if there is no coupling $J^{5\mu}A_\mu$ in the Lagrangian! Indeed, in QED there is no such term, so in principle, the gauge boson A_μ won't notice that $\partial_\mu J^{5\mu} \neq 0$. But in the Standard Model there are Weyl fermions, and writing the above projector P_L there is such a term, from the γ^5 piece.

A nonconserved current cannot be consistently coupled to gauge bosons; there would be production of longitudinal polarization and loss of unitarity. There is a nice discussion in Schwarz 29.2 of loss of unitarity in WZ scattering already around 1 TeV without a Higgs boson. Note that this is more subtle than the situaton in QED: the photon is massless and doesn't have a third polarization component so producing such a component is clearly unphysical (in general, producing particles that are not on the menu of particles to begin with typically violates unitarity). But W and Z are massive and do have a physical longitudinal polarization, so there loss of unitarity is a little more subtle: a known bound on how quickly the amplitude can grow with energy is violated, the *partial wave bound* or the related *Froissart bound* [17].

To connect to pion decay above, a chiral anomaly could also produce a global anomaly without producing a gauge anomaly, so this is another example of the disconnection of the two concepts.

5 Functional integral

Fujikawa's functional integral calculation of the chiral anomaly is very nice, but it is well explained both in Schwarz 30.3 and in Peskin & Schroeder 19.2 and many other places, so there is no need to for me to repeat it. Just to give a little bit of flavor, a key identity for the gauge-covariant Dirac operator is Schwarz (10.106):

$$\not{D}^2 = D_\mu^2 + \frac{e}{2} F_{\mu\nu}\sigma^{\mu\nu}$$

where $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$, which produces $F_{\mu\nu}(*F)^{\mu\nu}$ very elegantly. This idea is developed in general in Peskin & Schroeder and later in Schwarz, and most powerfully by DeWitt [16]. Conceptually, with a functional integral it is easy to state how a symmetry of the classical Lagrangian becomes violated in the quantum theory: the measure of the functional integral is not invariant.

This could be a good place to at least mention the *Witten anomaly* [20], which is not the same anomaly as what I've been discussing so far. I will not have much to say about it.

6 Gravitational anomaly

In some sense this is an anomaly in diffeomorphism (general coordinate) invariance. This requires viewing gravity as a gauge theory, which takes a little bit more thinking than for ordinary gauge theory[2]. The way to understand it is to introduce Cartan's *repère mobile* (moving frame) from 1937[3], or *vielbein* in German ("multi-leg", so for $D = 4$ it's called "vierbein", "four-leg", not related to the polarization e_i^μ above): we "factorize" the metric into two frame parts with a new index a that is summed over:

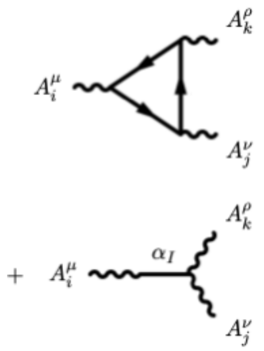
$$g_{\mu\nu} = e_\mu^a e_\nu^a = (\mathbf{e}_\mu)^T \mathbf{e}_\nu$$

where in the last equation I introduced vector notation for the a index. The point is that there is a new redundancy which is much like a gauge redundancy: the frame e_μ^a can be rotated by a local Lorentz transformation at each point and still give the same metric: $(\Lambda\mathbf{e}_\mu)^T \Lambda\mathbf{e}_\nu = \mathbf{e}_\mu \Lambda^T \Lambda \mathbf{e}_\nu = (\mathbf{e}_\mu)^T \mathbf{e}_\nu$. Using frame fields, the anomaly is checked for the same way as above: take a graviton amplitude and put polarizations on some leg longitudinal: $e_1^{\mu\nu} = p_1^\mu p_j^\nu$ for any j.

[2] The Becker-Becker-Schwarz string theory book has a nice review of this in Chapter 5.4.

[3] Hermann Weyl wrote in 1938 that "I found the book, like most of Cartan's papers, hard reading" [6].

The classic paper on gravitational anomalies is Alvarez-Gaume and Witten [2]. It contained a no-go theorem: in general dimension, it is very difficult to find anomaly-free Yang-Mills theories. In 1984, Green-Schwarz mechanism was discovered. In $D = 4$, an anomaly from the triangle diagram in the figure can be cancelled by exchange of the axion field α_I in the lower diagram (figure from [3]). This leads to a "generalized symmetry", where gauge transformations of the axion field cancel those of the gauge field, which was realized by Stückelberg before Higgs: $(A_\mu - m\partial_\mu\alpha)^2$ is a gauge invariant mass term for the gauge field, where the axion supplies the extra polarization (is "eaten" by A_μ).



This is typical of "string-inspired" quantum field theories. A field theorist would find the cancellation above weird: you have to include a coupling g in the $A_\mu\partial^\mu\alpha$ coefficient to be able to cancel a tree-level diagram against a loop diagram. But this follows from the low-energy limit of a single string diagram (Polchinski, Ch.7). So although it's a strange-looking field theory, it can be tested experimentally like any other field theory. The new gauge boson is called the "Z' (Z-prime) boson", and search strategies have been proposed, e.g. in [3]. It hasn't been discovered, but if it were, I like many people would call this evidence for extended objects, although it's clearly not a "proof".

This argument can also be turned around: you can "discover" a Green-Schwarz 2-form by "following the anomaly" [22].

7 Anomalies in the worldline formalism

As Schwarz writes: it was crucial for Schwinger to develop a gauge-invariant proper-time formalism to solve the seeming gauge dependence of earlier calculations of $\pi^0 \to 2\gamma$.

It's not a coincidence that a book that is entirely about anomalies [9] is almost entirely expressed in the worldline formalism. See conceptual questions below.

8 Anomalies in string theory: p-forms and submanifolds

In string theory there are basically two new ingredients: new gravitational-like fields (p-index antisymmetric tensors), and submanifolds embedded in higher dimensions (e.g. Dirichlet-branes, where open strings end). Antisymmetric tensors and embedded submanifolds can both studied in field theory, but some connections only become evident in string theory. For example, the *descent method* (discussed nicely in Harvey) makes it natural to express a D-dimensional anomaly in terms of a differential form of rank $D + 2$.[4] This is useful because the original anomaly can be written as the exterior derivative of some $D + 1$-form (again, this is differential form language), and then there is an issue about gauge invariance, which is resolved by going up one more step in rank to represent it. It is surprising that a $D + 2$-form should be relevant, since it cannot even be defined in D dimensions! Again, this is discussed in Becker-Becker-Schwarz. Harvey points out that the potentially anomalous groups ($SU(N)$ for $N \geq 3$, $SO(4N + 2)$ and E_6) all have nonvanishing fifth homotopy group $\pi_5(G) \neq 0$, so there is some interesting topology here; this is also discussed in Weinberg Vol.2.

With submanifolds, one might want to argue (like e.g. Uranga does) that not only the D dimensional theory should be non-anomalous, but also every probe of dimensionality $d < D$ that can be inserted into the theory should have a non-anomalous d-dimensional theory on its world-surface. This leads to so-called "K-theory constraints". In field theory it is not so clear why one should require this (if there are only point particles, who cares if a membrane theory is anomalous?) but in string theory there are *dualities* that relate theories with objects of various dimensionalities, so we

[4] Differential forms are nicely explained e.g. in Polchinski's or Naharaka's books, or on Wikipedia. Incidentally, the descent method is also discussed on the Wikipedia page for "Gauge anomaly", but currently I don't find that page very useful.

would not want an anomalous theory to show up when pursuing a relation from a non-anomalous theory.

This also means that anomalies on various submanifolds can "communicate". This is called *anomaly inflow*, i.e. a higher-dimensional anomaly can "leak into" a submanifold and cancel its anomaly. One simple model that Harvey nicely discusses in his 2005 review (section 4.1) where some of these points can be introduced is *axion electrodynamics*, which is QED coupled to an additional pseudoscalar Φ (the axion),

$$\Phi F_{\mu\nu} \tilde{F}^{\mu\nu} \; .$$

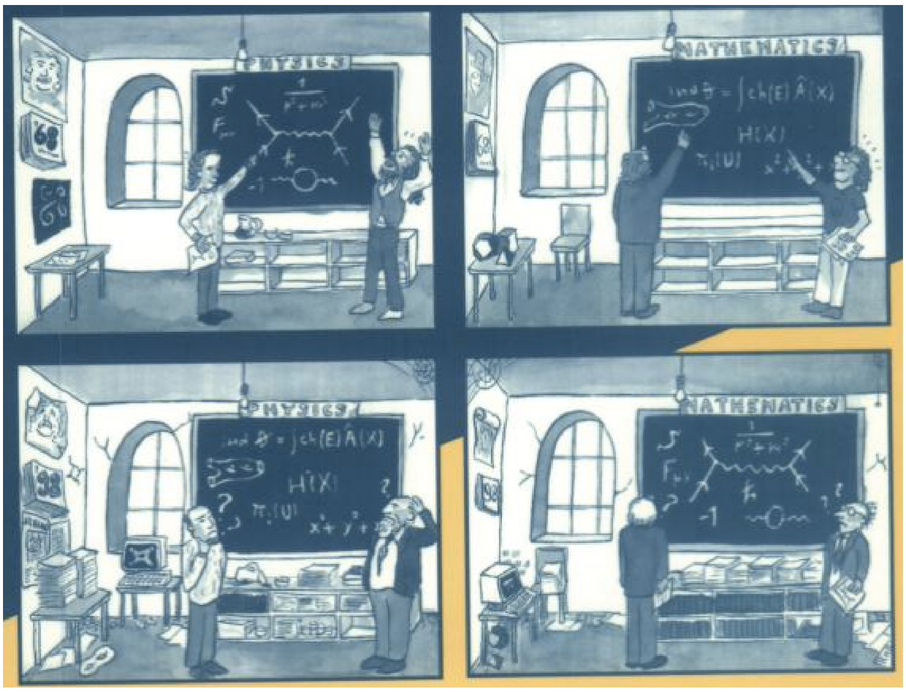
Notice that even though the $F\tilde{F}$ term is a total derivative by itself, the above term is not a total derivative; the derivative from partial integration hits $\Phi(x)$. A related topic is the *Freed-Witten anomaly*, related to the above-mentioned Witten anomaly.

9 Anomalies in mathematics: index theorems

A big generalization of both the Chern-Gauss-Bonnet theorem and the Riemann-Roch theorem is the *Atyiah-Singer index theorem* [4] from the 1960s, that captures all these things and expresses them nicely in terms of characteristic classes (see e.g. [1] section 3.6):

$$\text{ind}\, \not{D} = \int \text{ch}(F) \widehat{A}(R) \tag{9.1}$$

I won't attempt to explain this beautiful formula here, but the basics are contained in the Becker-Becker-Schwarz chapter mentioned above, and it is nicely explained in more detail in the mathematical physics book by Nakahara [21], Ch. 11 onward. Instead, here's a cartoon by Robbert Dijkgraaf that shows the development from 1968 (top two panels) to 1998 (bottom two panels) [7] where this formula appears:



10 Some conceptual questions

Here are some questions that I had myself at some point. They may not be very illuminating to you.

Question: Why did we need to compute with three gauge bosons just to see that the single current $J^{5\mu}$ is not conserved? (I gave the counting $n \geq D/2 + 1$ above so technically that is "why", but is there a conceptual reason for this, and is there another way to compute it?)

Answer: Feynman diagrams in the usual sense are not optimized for this kind of calculation. In Peskin & Schroeder, they insist on first computing it two other ways to make this point clearer. In section 33 in Schwarz, he uses the worldline formalism to compute the one-point function of a single current in a gauge field background A:

$$\langle A|\partial_\mu J^{5\mu}|A\rangle = -\frac{\alpha}{2\pi}F_{\mu\nu}\tilde{F}^{\mu\nu} \ .$$

This is conceptually the best way to understand the anomaly in my opinion; the extra gauge bosons in the Feynman diagram are just a perturbative representation of this background gauge field.

Question: *Why did we worry at all whether QED has a chiral anomaly, it only comes from massless chiral fermions, and the electron isn't massless*? (Well, the end result is that QED doesn't have a chiral anomaly! But the question is, why did we even worry about that? It's the same for all other fermions in the Standard Model: it seems like there are no massless fermions in nature.)

But for high-energy processes, the electron mass is completely negligible. From this point of view, anomaly freedom could perhaps be thought of as a technical trick to be able to work without masses even if the actual theory with mass is anomaly-free. But it's deeper than that: all fermion masses in the Standard Model come from the Higgs mechanism, and the theory before spontaneous symmetry breaking should make sense, so it should better not be anomalous. In other words, we are right to worry about anomalies!

In fact not even all massless fermions can be anomalous even in principle: *"only massless particles for which no mass term is allowed that is consistent with the potentially anomalous symmetry can contribute to the anomaly".* Otherwise, we could use Pauli-Villars (add fictitious heavy particle) and renormalize while preserving the symmetry, at least if it is not broken by masses.

Question: *Since there are no chiral anomalies at all in odd dimensions, couldn't the deeper explanation for anomaly freedom just be that the world is five-dimensional? Or eleven-dimensional?*

Well, yes, many people have thought about this. But for example string theory seems to have a preference for $D = 10$, and ten is even worse than four, since there can be gravitational anomalies that you cannot have in $D = 4$. There is the idea that string theory descends from a theory in $D = 11$ called M-theory and there is a lot of discussion about this, perhaps this does help explain something. But in M-theory there are lower-dimensional objects of dimension 2 and 5, so it's not for free there either.

By the way, while thinking about higher dimensions, I find this connection useful: a gauge anomaly in $2k$ dimensions is related to a chiral anomaly in $2k + 2$ dimensions. So a gauge anomaly in $D = 4$ is related to a chiral anomaly in $D = 6$ ([1] section 3).

Question: *I have a great idea how to quantize without producing anomalies! How do I check it?*

The "$\pi^0 \to 2\gamma$ sledgehammer" is that the above width Γ is well-measured, so if your nice new theory has the great feature that it automatically gets rid of anomalies, it will predict $\Gamma = 0$, which is ruled out. [5]

References

[1] J. A. Harvey, "TASI 2003 lectures on anomalies," hep-th/0509097.

[5] There could be other ways to get $\Gamma = 7.8$ eV than the standard way, I just don't know any frameworks that really do that, short of replacing quantum mechanics. An example of this would be the (at times absurd) blog discussion moderated by my string theory teacher Jacques Distler [10] about a "new quantization method" used by some workers in *loop quantum gravity*. This is summarized in a paper by Robert Helling [11]. (Also see [12] by Robert's former supervisor, which provides more context but goes far beyond the point I was trying to illustrate.) I don't understand the details, but to me it does not seem very attractive to try to change quantum mechanics. Unfortunately, Weinberg now disagrees with me on this, and recently Distler has been writing papers on this too! [13, 14] The modification they are trying was introduced by Göran Lindblad at KTH. Lindblad graciously gave a talk about it in my seminar series in Stockholm, but I understood very little.

[2] L. Alvarez-Gaume and E. Witten, "Gravitational Anomalies," Nucl. Phys. B **234** (1984) 269. doi:10.1016/0550-3213(84)90066-X

[3] P. Anastasopoulos, M. Bianchi, E. Dudas and E. Kiritsis, "Anomalies, anomalous U(1)'s and generalized Chern-Simons terms," JHEP **11** (2006), 057 doi:10.1088/1126-6708/2006/11/057 [arXiv:hep-th/0605225 [hep-th]].

[4] "Atiyah–Singer index theorem", Wikipedia page
https://en.wikipedia.org/wiki/Atiyah-Singer_index_theorem
The original paper is:
M.F. Atiyah, I. M. Singer, "The Index of Elliptic Operators on Compact Manifolds" (1963), Bull. Amer. Math. Soc., 69 (3): 422–433

[5] "Gamma matrices", Wikipedia entry, https://en.wikipedia.org/wiki/Gamma_matrices

[6] H.Weyl, "Cartan on Groups and Differential Geometry" (1938),
http://projecteuclid.org/euclid.bams/1183500670

[7] P. Deligne, P. Etingof, D. S. Freed, L. C. Jeffrey, D. Kazhdan, J. W. Morgan, D. R. Morrison and E. Witten, "Quantum fields and strings: A course for mathematicians. Vol. 1, 2," Providence, USA: AMS (1999) 1-1501

[8] M. E. Peskin, "An Introduction To The Theory Of Strings," SLAC-PUB-3821
http://inspirehep.net/record/218654/files/slac-pub-3821.pdf

[9] F. Bastianelli and P. van Nieuwenhuizen, "Path integrals and anomalies in curved space," (2006), Cambridge.

[10] "The LQG Landscape", bloggdiskussion
https://golem.ph.utexas.edu/~distler/blog/archives/000855.html

[11] R. C. Helling, "Lessons from the LQG string," hep-th/0610193.

[12] H. Nicolai, K. Peeters and M. Zamaklar, "Loop quantum gravity: An Outside view," Class. Quant. Grav. **22** (2005) R193 [hep-th/0501114].

[13] S. Weinberg, "Lindblad Decoherence in Atomic Clocks," Phys. Rev. A **94** (2016) no.4, 042117 [arXiv:1610.02537 [quant-ph]].

[14] J. Distler and S. Paban, "von Neumann's Formula, Measurements and the Lindblad Equation," arXiv:1702.01724 [quant-ph].

[15] CEA Saclay webpage,
http://ipht.cea.fr/en/Phocea-SPhT/ast_visu_spht.php?id_ast=231

[16] B. S. DeWitt, "The global approach to quantum field theory. Vol. 1, 2," Int. Ser. Monogr. Phys. **114** (2003) 1.

[17] "Froissart Bound", Scholarpedia, by Mr. Froissart!
http://www.scholarpedia.org/article/Froissart_bound

[18] C. de Rham, J. T. Deskins, A. J. Tolley and S. Y. Zhou, "Graviton Mass Bounds," arXiv:1606.08462 [astro-ph.CO].

[19] J. C. Collins, "Renormalization : An Introduction to Renormalization, The Renormalization Group, and the Operator Product Expansion" (1984), Cambridge University Press.

[20] https://en.wikipedia.org/wiki/Anomaly_(physics)#Witten_anomaly

[21] Nakahara, "Geometry, Topology and Physics" (2003), CRC Press.

[22] W. M. Chen, Y. t. Huang and D. A. McGady, "Anomalies without an action," [arXiv:1402.7062 [hep-th]].