the handy booklet of

Constantly Asked Physics Questions

Contents

Kinematics/General Physics 2

KGP1 - Why is $c/\hbar/k_B$/(other fundamental constant) the value it is? What would happen if it was different? 2

Classical Mechanics 4

Thermodynamics 4

Celestial/Orbital Mechanics 4

Fluid Mechanics 4

Classical Electrodynamics 4

CED1 - What is the Maxwell stress tensor σ and how does it work? 4

Analytical Mechanics 6

Special Relativity 6

SR1 - In what sense does $E = mc^2$, and what does it mean? . . . 6

SR2 - if photons are massless, how can $E = mc^2$? 9

General Relativity 11

Quantum Mechanics 11

General Quantum Field Theory/Many Body/Relativistic QM 11

Nuclear Physics 11

Quantum Electrodynamics 11

QED1 - How does $1 + 2 + 3 + \ldots = -1/12/\zeta$-regularization enter in the Casimir effect? 11

Particle Physics/High Energy Physics 16

HEP1 - Why are photons massless? 16

HEP2 - Why do photons not acquire a mass through the Higgs mechanism? 16

HEP3 - Why do photons not acquire a mass through quantum corrections/interaction with virtual particles? What is charge renormalization? 18

Quantum Gravity/String Theory 21

QG1 - If the Planck length/Planck time is the smallest measurable/possible length/time, then...? 21

Kinematics/General Physics

KGP1 - Why is $c/\hbar/k_B$/(other fundamental constant) the value it is? What would happen if it was different?

Why is c (the speed of light in a vacuum) exactly 299 792 458 m/s?

The metre and the second are arbitrary units that originally referred in their definition to natural phenomena that were relevant to the daily life of humans. A second would be 1/86400 of a day, the period of rotation of the Earth. The metre is one ten-millionth of the distance between the equator and the North Pole.

As soon as the necessary physics was consolidated, these definitions were replaced with the modern ones:

1) The second is 9 192 631 770 times the period of the radiation emitted in a specific atomic transition of ^{133}Cs.
2) The metre is the length travelled by light in vacuum in 1/299 792 458 of a second.

The first definition is still referring to a natural process, albeit much more exact than the rotation of the Earth. The bizzarre number involved is to make sure the new definition makes a modern second as similar as possible to the older second. Basically, the period of the transition radiation had been previously *measured* as being 1/9192631770 s (in older seconds).

The second one rests on the first and simultaneously fixes the value of c. The units are explicitly *designed* so that the value of c is 299 792 458. This value is similar to the previous, measured value of c in the older units.

Instead, the value of c in the new system is defined, not measured.

This means we can actually make c have any value we want by redefining the units. If we use lightyears and years for measuring lengths and times, we get

$$c = 1\ \text{ly/y}$$

But here it seems like we're just playing around with units. We aren't apparently actually *changing the speed of light*. Here's a little gedankenexperiment about that.

Assume there was a Universe where the speed of light was twice ours:

$$c' = 2c$$

Since the speed of light enters basically every relativistic phenomenon and many things about light and electromagnetism, you can bet the dynamics in this primed Universe will be very different from those in our unprimed one. For example, it's very likely it won't develop human life, at least in the same way as it happened for us.

However.

Consider the following change of variables in the primed Universe:

$$x' \to \tilde{x}' = x'/2, \quad t' \to \tilde{t}' = t'$$

The tilded coordinates are just the normal coordinates, but with space stretched by a factor of 2. In these coordinates, the value of the speed of light is:

$$\tilde{c} = \left(\frac{\tilde{x}'}{\tilde{t}'}\right)_{\text{computed for a light ray}} = \frac{1}{2}\frac{x'}{t'} = \frac{1}{2}c' = c$$

So we have to admit that, in the tilded coordinates, the speed of light is back to its original value in our Universe. In those coordinates, the primed Universe satisfies the same equations of motion as our own, and evolves identically. It has the same Big Bang, the same primordial nucleosynthesis, the same star formation, a Sun identical to ours, and an Earth and humans.

The humans are twice as tall as us, yes. But they don't know that. Since their Earth is twice as big, their metre (the tilded metre) is twice ours, and they measure their height to be normal in their tilded metres.

They finally measure the speed of light to be 299 792 458 metres per second.

So, nothing would happen. This is the manifestation of how fundamentally meaningless the value of the speed of light is. Mainly because there is no independent other speed with which to compare it, as all speeds in physics ultimately depend on it.

All of the above applies to the following set of independent fundamental physical constants:

c, $\hbar$, k_B, ϵ_0, G

and all those auxiliary constants formed by products of powers of those above. The above set is a maximal set of independent constants and is the largest set of constant for which you can simultaneously impose a fixed value by redefining the units. If you fix less than all of these, you get freedom in your system of units. For example, setting

$c = \hbar = k_B = \epsilon_0 = 1$

gives natural units (Lorentz-Heaviside variant), useful in high-energy physics. Since you omitted G, you still have a single arbitrary unit to fix. You can use the metre, for example, and the rest of the units follow. The time unit is cm, the energy unit is $\frac{c\hbar}{\mathrm{m}}$ and so on.

Adding $G = 1$ instead gives Planck units.

Classical Mechanics

Thermodynamics

Celestial/Orbital Mechanics

Fluid Mechanics

Classical Electrodynamics

CED1 - What is the Maxwell stress tensor σ and how does it work?

note: do not confuse this with the Maxwell tensor $F^{\mu\nu}$, which is the electromagnetic field.

Consider linear momentum:

$$P^i$$

it's the i component of the total linear momentum in the system. It will be an integral over space of some linear momentum density I'll call p^i:

$$P^i = \int d^3x \; p^i(x)$$

This quantity is conserved. It's also pretty reasonable that it's somewhat conserved *locally*, meaning that it doesn't just disappear somewhere and reappear magically somewhere else, it has to *flow*.

What I mean is that if the amount of linear momentum in a certain region of space changes, it must be because of some flux in/out the surface of that region:

$$\frac{d}{dt}P^i_V = \int_V d^3x \frac{d}{dt}p^i(x) = -\oint_{\partial V} d^2\vec{\Sigma} \; \cdot \vec{\Phi}^i$$

I just wrote that the variation of P^i_V (P^i restricted to the volume V), which is equal to the integral of the variation of the density, must be counterbalanced by some flux $\vec{\Phi}$ that crosses the boundary ∂V. If it's decreasing, then it must be leaking.

Now here's the thing: the stress tensor element σ_{ij} is precisely the *flux of P^i in the j direction.* It's how much P^i is flowing through a unit surface orthogonal to the j-direction. It is also symmetric (nontrivial, and in fact dependent on some choices) and it does transform like a 2-tensor, a matrix, which justifies the name.

So let's rewrite what we had:

$$\int_V d^3x \frac{d}{dt} p^i = -\oint_{\partial V} d^2\Sigma^j \ \sigma^{ij}$$

What I've done is:

- Rewritten the scalar product $\vec{\Sigma} \cdot \vec{\Phi}^i$ using indices as $\Sigma^j \Phi^{ij}$. Be careful about these indices: i means which component of the momentum we're talking about, j is the vector index of the flux itself (which is a vector).
- Recognized that $\Phi^{ij} = \sigma^{ij}$ from what we said earlier.

Now what you would like to do is to deduce a differential, infinitesimal form of the equation above (which is known as the integral continuity equation). You do this by integrating over a very small cube; I'll spare you the details, but it's an easy computation, and you end up with:

$$\frac{\partial p^i}{\partial t} + \partial_j \sigma^{ij} = 0$$

or, in vector form:

$$\frac{\partial \vec{p}}{\partial t} + \vec{\nabla} \cdot \sigma = 0$$

this is the continuity equation or local conservation (in differential form). (note that the density $\vec{p}$ is a vector, because it's the density of the vector $\vec{P}$.)

In an interacting theory of electromagnetic fields and matter, both contribute to total linear momentum. So, reasonably, both will have a stress tensor:

$$\sigma^{ij} = \sigma_f^{ij} + \sigma_m^{ij}$$

and they will **not** be separately conserved. Only their sum, total stress, obeys the continuity equation we just found. The physical interpretation is that momentum can be exchanged between fields and matter. When an electron produces radiation, for example, that radiation carries away momentum from the electron. We can substitute the decomposition in the continuity equation to obtain:

$$\frac{\partial \vec{p}_f}{\partial t} + \vec{\nabla} \cdot \sigma_f = -\frac{\partial \vec{p}_m}{\partial t} + \vec{\nabla} \cdot \sigma_m =: \vec{s}$$

Where I have defined the source term s.

This source term encapsulate the passage of momentum from charged matter to fields. So field momentum is not conserved separately, and s represents "generation" of momentum from charges. Fittingly, the equation is now called a continuity equation with sources.

So the Maxwell stress tensor is just σ_f, the stress tensor for only the electromagnetic field.

What I've detailed up to now is the physical interpretation in general of the stress tensor in any local theory; classical EM is just the first field theory one usually encounters. The actual form of the tensor for electromagnetism in terms of E & B is computed from the Poynting vector and a full derivation is presented in any decent CED textbook.

Analytical Mechanics

Special Relativity

SR1 - In what sense does $E = mc^2$, and what does it mean?

Ok, ok, ok. Let's take a deep breath.

When switching from Newtonian to relativistic physics, a couple of formulas have to be explicitly replaced. In particular, we can summarize this shift in the following substitutions for the mechanical energy and linear momentum of a body:

$$E_N = \frac{1}{2} m v^2 \longrightarrow E = \frac{1}{\sqrt{1 - (v/c)^2}} mc^2$$

$$p_N = mv \longrightarrow p = \frac{v/c}{\sqrt{1 - (v/c)^2}} mc$$

These might look a bit daunting, and additionally I just pulled them out of my ass. Just trust me that they can be derived rigorously. Let us concentrate on interpreting them and their consequences.

The Newtonian expressions should be well-known; they are however incorrect when speed becomes relativistic ($(v/c) > 0.1$ is a good rule-of-thumb) and must be replaced by the formulas on the right.

m is the mass. Just mass. It's a constant *and* an invariant for the object. It does not depend on the frame of reference nor the *global* state of motion of the body (I'll clarify this adjective in a minute). Do not trust anyone talking about "relativistic mass", it's an old concept from when people were still trying to figure out this stuff and it makes everything *immensely* more complex (just a taste: there is a transverse and a longitudinal relativistic mass). Only ever discuss invariant mass, or just mass.

v is simply how much space the body travels over how much time, with space and time measured in a certain inertial frame. I really want to stress the simplicity of this definition, because people often get confused with time dilation and length contraction and other complications, while this definition is absolutely crystal clear:

$$v = \frac{dx}{dt}$$

where x and t are the space and time coordinates of the body in some coordinate system (reference frame), nothing to do with proper time or anything of that sort. No magic here.

And c, of course, is the speed of light in vacuum.

The v/c ratio and the $(1 - (v/c)^2)^{-1/2}$ thing are so ubiquitous in SR that we give them the following names:

$$\beta := v/c \qquad \gamma := \frac{1}{\sqrt{1-\beta^2}}$$

So the formulas simplify to $E = \gamma mc^2$ and $p = \beta\gamma mc$.

Now, the burden of proving that the relativistic expressions do actually reduce to the Newtonian expressions in the nonrelativistic limit is on us. The nonrelativistic limit is when $v \ll c$, or equivalently $\beta \ll 1$. To do so, let us recall the following Taylor expansion from calculus:

$$(1+\epsilon)^\alpha = 1 + \alpha\epsilon + \frac{\alpha(\alpha-1)}{2!}\epsilon^2 + \ldots$$

this is simply the Binomial expansion and it's a really useful one to keep in mind (at least the first order term). We expand the γ factor using $\epsilon = -\beta^2$ as such:

$$\gamma = \left(1-\beta^2\right)^{-1/2} = 1 + \frac{1}{2}\beta^2 + \ldots$$

Second-order in β is all we really need. So, finally, for the mechanical energy and momentum in the nonrelativistic limit we get:

$$E = mc^2 + \frac{1}{2}mv^2 + \dots$$
$$p = mv + \dots$$

In the expression for p I've stopped at the first term because the next is order β^3.

There's something seriously wrong. p looks like its Newtonian counterpart p_N, while E has an additional spurious term, mc^2. This is not a small term. In fact it's huge.

$E_0 := mc^2$ is the energy the body has when $v = 0$, so it's called the rest energy. Why does it not completely invalidate Newtonian mechanics?

Mostly, it's because it's impossible to tap into this energy. In nonrelativistic mechanics, mass is conserved (it is **not** conserved in special relativity). This means that in any physical process, E_0 is untouched. It decouples completely from the physics, and thus it's just an invisible energy shift.

Mechanics (and Physics in general) is insensible to global energy shifts. For example, if your mechanical energy is

$$E = \frac{1}{2}mv^2 + 49\,\mathrm{J}$$

nothing changes in your dynamics. You just added a constant, so what. Since mc^2 is effectively a constant in nonrelativistic physics, it does not affect dynamics and could not be derived even in principle by a nonrelativist not aware of special relativity. In fact, Newton just set that constant to zero for simplicity.

Now, we said that E_0 is the energy the body has when it's at rest. So we can conveniently divide our total energy in E_0 and a term we rightfully call kinetic energy:

$$E = \gamma mc^2 = mc^2 + (\gamma - 1)mc^2 =: E_0 + E_K$$

Since $\gamma - 1 \sim \frac{1}{2}\beta^2$ it's clear that the nonrelativistic limit is just $E_K \ll E_0$.

This is starting to make sense: E_0 is the energy the object has simply for existing, there is an energy cost associated with just having a mass. It is the difference between the energy of a state where the object exists (and is still) and one where it doesn't. It's the required energy to create it, or the yield if it's destroyed. Of course, this does not prove that it's possible to create or destroy mass, just that *if* there is a channel for that creation or destruction, that is the energy requirement.

Since in nonrelativistic mechanics the energies involved in processes (E_K) are much smaller than the E_0 for an object, creation and destruction of mass most certainly do not happen in nonrelativistic physics.

Since E_0 is the total energy of the object when it's still, it's reasonable that if the object was a composite system made of smaller units, it also includes the internal energy, not just the sum of the rest energies of the components.

Take for example a stationary box filled with a gas at temperature T. The overall, or average velocity of the gas is zero, but the single particle of the gas will have a nonzero velocity and consequently a kinetic energy E_K^i. The total energy is

$$\sum_i E^i = \sum_i mc^2 + \sum_i E_K^i$$

but we have said that this must be $E_0 = Mc^2$, with M the mass of the box, so

$$M = \sum_i m + \frac{1}{c^2}(\sum_i E_K^i)$$

So, the mass of the box is actually greater than the sum of the masses of the particles! Albeit, by a very, very small amount, that only gets relevant if E_K^i is at least of order mc^2. This shows that mass is not additive, and displays the so called "mass-energy equivalence" which is more correctly expressed as:

$$U = mc^2$$

that is, mass (as in, the inertia measured in Newtonian mechanics) is equivalent to the *total* **internal** energy, also including the energy to create the constituents.

This is why people will shout at you that $E = mc^2$ is not the full formula/is wrong. They're right. That E is supposed to be E_0.

SR2 - if photons are massless, how can $E = mc^2$?

Read SR1.

The formulas in SR1 are singular if $v = c$, which is certainly the case for a photon, a quantum of light. We need to rewrite them by getting rid of the velocity. (Or take a careful limit. That's a nice alternative. But we won't do that).

In classical mechanics, this is already done when switching from the Lagrangian to the Hamiltonian. You want to write $E(v)$ in function of $p(v)$. So you invert $p(v)$ as $v(p)$ and then substitute $E(v(p))$. Easier done than said:

$$p_N = mv \Rightarrow$$
$$v = \frac{p_N}{m}$$
$$E_N(v) = \frac{1}{2}mv^2 \Rightarrow E_N(p) = \frac{p_N^2}{2m}$$

Which you'll recognize as a standard kinetic Hamiltonian if you're into Hamiltonian mechanics and you'll ignore this sentence if you don't.

We can do the same with the relativistic case. But first, a neat fact about β and γ:

$$\gamma^2 - (\beta\gamma)^2 = 1$$

try it. It's very boring, but it's true. (It's cool because it implies that if $\gamma = \cosh(\eta)$, then $\sinh(\eta) = \beta\gamma$ and $\tanh(\eta) = \beta$ and if you don't think hyperbolic functions are the shit then I don't know what to tell you).

So, we write

$$\begin{aligned}
\gamma &= \sqrt{1 + (\beta\gamma)^2} \\
E = \gamma mc^2 &= \sqrt{1 + (\beta\gamma)^2} mc^2 \\
&= \sqrt{(mc^2)^2 + (\beta\gamma mc^2)^2} \\
&= \sqrt{(mc^2)^2 + (pc)^2}
\end{aligned}$$

So this is our $E(p)$, the so-called "full expression" for the mechanical energy of a relativistic body.

(If you Taylor-expand $E(p)$ around $p = 0$, you get $E = mc^2 + \frac{p^2}{2m} + \dots$. Go figure.)

Ok, so we've gotten rid of the gammas and betas. Just a last thing about them! The speed of an object is always recoverable from the energy and momentum:

$$\frac{p}{E}c^2 = \frac{\beta\gamma mc}{\gamma mc^2}c^2 = \beta c = v$$

And only now that we have built this architecture we plug in $m = 0$ to find about massless particles. We get

$$E = pc$$

and

$$v = c, \quad \gamma = \infty$$

So massless particles move always at the speed of light and have energy proportional to their momentum. In the limit where the momentum goes to zero, $p \to 0$, the energy also goes to zero. Instead, for massive objects the energy tends to the rest energy mc^2. Therefore it makes sense to extend the definition of the rest energy $E_0 = mc^2$ to photons, with $m = 0$, even if the cannot ever be brought to rest.

The energy of a photon is entirely kinetical.

Note that the previous expressions $E = \gamma(v)mc^2$ and $p = \beta(v)\gamma(v)mc$ when $m = 0$ are both indeterminate forms $(0 \cdot \infty)$. This makes sense: we have many photons, all with $v = c$, with different momentum and energy. We shouldn't be able therefore to determine the energy/momentum exactly just from the speed, so the math honourably breaks down.

General Relativity

Quantum Mechanics

General Quantum Field Theory/Many Body/Relativistic QM

Nuclear Physics

Quantum Electrodynamics

QED1 - How does $1 + 2 + 3 + \ldots = -1/12/\zeta$-regularization enter in the Casimir effect?

The Casimir effect is an attractive force between very close parallel conducting plates and it's a consequence of the quantum nature of the EM field. There are numerous ways to derive it, with various levels of rigour, but the one employing zeta-regularization holds a special place in my heart.

For simplicity, reduce from three to just one spatial dimension. This doesn't change the essential points of the computation. Then our parallel plates are actually two barriers, one at $x = 0$ and one at $x = a$, with a the spacing. We know conductors act in a way as to make the longitudinal component of the electric field vanish on them. So we can model these boundary conditions as

$$\vec{E}(t, 0) = 0 \qquad \vec{E}(t, a) = 0$$

The electric field satisfies the wave equation both inbetween and outside the plates:

$$\Box\vec{E}(t, x) = \left(\frac{1}{c^2}\partial_t^2 - \partial_x^2\right)\vec{E}(t, x) = 0$$

Now, since this equation is linear it's really tempting to write down a Fourier series for the electric field. We know that any function on the interval $[0, a]$ can be expanded in this basis of fundamental waves:

$$S_n(x) = \sin(k_n x)$$

with $k = \frac{\pi n}{a}$, $n = 1, 2, \ldots$. These are a bit different than the usual presentation of Fourier Series in that they use only sines (but halve the lowest period). If you are not familiar with Fourier sine series this is a good read.

We can decompose any function on the interval in this basis:

$$f(x) = \sum_{n=1}^{\infty} f_n \sin(k_n x)$$

So let's do that for the electric field at a given time t.

$$\vec{E}(t, x) = \sum_{n=-\infty}^{\infty} \sum_i A_{ni}(t) \sin(k_n x) \vec{e}_i$$

Since $\vec{E}$ is a vector, we have to introduce two basis polarization vectors $\vec{e}_y$, $\vec{e}_z$ pointing in the y and z directions, and sum over polarization ($i = y, z$). We have therefore two modes of oscillation for each value of n, corresponding to the two polarizations. The $\vec{e}_x$ polarization does not exist, because the electromagnetic field does not have a longitudinal component in free space.

$A_{ni}(t)$ is the coefficient of the expansion of $\vec{E}(t)$ in the mode given by n and i, and obviously depends on time.

The expression obtained for the electric field automatically satisfies the boundary conditions *if* the coefficient A_{ni} are not too crazy (it's still an infinite sum of functions, it's no joke).

Let us study the time evolution of A_{ni}. Since the equation is linear, we just plug into the equation one single term from our expansion of the electric field:

$$\Box(A_{ni}(t) \sin(k_n x)) = \frac{1}{c^2} \left(\partial_t^2 A_{ni}(t)\right) \sin(k_n x) - k_n^2 A_{ni}(t) \sin(k_n x) = 0 \Rightarrow$$

$$\partial_t^2 A_{ni}(t) + \omega_n^2 A_{ni}(t) = 0$$

with $\omega_n = c k_n$. This is the equation for a harmonic oscillator! Of course this is no surprise, it's well-known that

$$\vec{E}_{ni}(t,x) = A_{ni}(t)\sin(k_n x)\vec{e}_i$$

is a standing-wave solution for the wave equation. The interesting bit is that this can be a way to get insight about the quantization without actually studying quantum field theory. This is because anyone who has studied a bit of basic quantum mechanics knows about the quantum harmonic oscillator (QHO) and how it's energy spectrum is given by

$$\mathcal{E} = \hbar\omega\left(m + \frac{1}{2}\right) \qquad m = 0, 1, 2, \ldots$$

So we hope that substituting our classical harmonic oscillator with its quantum counterpart we would be doing something presumably similar to what actually quantizing the EM field is. Let's try that. Each mode (n, i) has a QHO associated with energy

$$\mathcal{E}^{ni} = \hbar\omega_n\left(m^{ni} + \frac{1}{2}\right)$$

and the total energy of the quantum EM field is given by $\mathcal{E} = \sum_n \sum_i \mathcal{E}^{ni}$.

We have unknowingly discovered photons.

The number m^{ni} is actually *the number of photons in the mode* (n, i), and takes the name of occupation number. Since photons are indistinguishable, you can describe the total state just by specifying how many photons have momentum $\hbar k$ and polarization $\vec{\varepsilon}$, how many $\hbar k'$ and $\vec{\varepsilon}'$, and so on; so just the full set of occupation numbers. The state with lowest energy possible is the one with 0 photons in each single mode, and we call that the vacuum. All the m^{ni} vanish and we are left with the energy per mode:

$$\mathcal{E}_0^{ni} = \frac{\hbar\omega_n}{2}$$

this is called the zero-point energy of the QHO, and is well-known to students of quantum mechanics. Even with zero photons in a specific mode, that mode has a certain energy. We can write down the total energy of the vacuum as the sum of the zero-point energies over all modes:

$$\mathcal{E}_0 = \sum_{n=1}^{\infty}\sum_i \frac{\hbar\omega_n}{2} = \frac{\hbar c\pi}{a}\sum_{n=1}^{\infty} n = \frac{\hbar c\pi}{a}(1 + 2 + 3 + 4 + \ldots) = \infty$$

Ugh. It's divergent. The total vacuum energy inbetween the plates is infinite. Physically, we can interpret it as being due to the ever-increasing contribution of high-momentum modes (it's a "UV" thing in QFT-speak). This divergence is most annoying and we must get rid of it someway. The reason I'm making you compute the total vacuum energy is more or less this: if the vacuum energy depends on the value of the separation a, then it acts as a sort of potential energy, that induces a force on the plates. I mean

$$U(a) = \mathcal{E}_0(a) \Rightarrow F(a) = -\frac{d}{da}U(a)$$

but if $U(a)$ is infinity that's kind of difficult to work with. There must be another infinity we should subtract to cancel the divergence. An interesting candidate is the *vacuum energy in the region without the plates if the plates were not there*. We can subtract that:

$$U(a) = \mathcal{E}_0(a) - \mathcal{E}_0^{\text{no plates}}$$

Makes sense, $U(a)$ it's the energy associated with putting the plates there. Now $\mathcal{E}_0^{\text{no plates}}$ is computed in a way entirely analogous to how we calculated $\mathcal{E}_0(a)$, but with a significant difference: the wavenumber k is a continuous variable. Inbetween the plates the electric field was constrained by the boundary conditions and therefore only multiples of the fundamental wavenumber $k_1 = \frac{\pi n}{a}$ were allowed. No such restriction exists without them and long story short our potential ends up being the difference of a divergent sum and a divergent integral:

$$U(a) = \frac{\hbar c \pi}{a}\left(\sum_{n=1}^{\infty} n - \int_0^{\infty} n dn\right) = \infty - \infty$$

That's looking undoubtedly better, but it's still bad. We are still summing and integrating to infinity *separately*, then subtracting. Let's instead trying summing/integrating *simultaneously*, and see if something cancels before we go to infinity. For this we introduce a regulator Λ, a maximum value for n. Then we hope to find that in the limit $\Lambda \to \infty$ we get a finite result not depending on Λ. This is also physically reasonable: the conductor blocks the electric field by having electrons oscillate with the same frequency in a way to cancel out the field of an incoming wave; above a certain frequency, we expect the conductor to not be able to catch up. The conductor *must* be permeable to EM waves above some frequency. Anyways, so we do

$$U(a) = \frac{\hbar c \pi}{a}\left(\sum_{n=1}^{\Lambda} n - \int_0^{\Lambda} n dn\right)$$

Which is better, because the sum of the first Λ integers goes as $\sim \frac{1}{2}\Lambda^2$ and the integral from 0 to Λ goes as $\sim \frac{1}{2}\Lambda^2$, so the leading contributions cancel! This does not prove that the difference then is finite as $\Lambda \to \infty$, but it's a good start.

There is an explicit formula for calculating this kind of divergent series minus divergent integral things and it's the Euler-MacLaurin formula, but it's really convoluted computationally. I just want to say that what we would get it's exactly the result if we had, naively, made the substitution

$$1 + 2 + 3 + \ldots \to -\frac{1}{12}$$

which is what is dictated by ζ-regularisation. This is, very handwavingly, because ζ-regularisation makes some kind of statement that could be depicted schematically as:

$$1 + 2 + 3 + \ldots = \infty - \frac{1}{12}$$

and we then subtract exactly that infinity when removing $\mathcal{E}_0^{\text{no plates}}$. Now, this is not a rigorous statement, but it's undoubtedly much easier to compute this way, and we immediately get our result:

$$\begin{aligned} U(a) &= \frac{\hbar c\pi}{a}\frac{-1}{12} = -\frac{\hbar c\pi}{12a} \\ \Rightarrow F(a) &= -\frac{d}{da}U(a) = -\frac{\hbar c\pi}{12a^2} \end{aligned}$$

That's cool, and it's attractive, but that's not the expression for the Casimir force on Wikipedia. Remember we've done this in one dimension; in three dimensions you get

$$\begin{aligned} \frac{U(a)}{A} &= -\frac{\pi^2}{720}\frac{\hbar c}{a^3} \\ \Rightarrow \frac{F(a)}{A} &= -\frac{\pi^2}{240}\frac{\hbar c}{a^4} \end{aligned}$$

which is the well-known expression - everything is per unit of area of the plates. (In this case, the divergent sum-integral pair is not the same as $1 + 2 + 3 + \ldots$, but it's equally nasty, and the same arguments apply).

AFAIK the Casimir force with the expression above was experimentally observed in at least two great experiments, first by Lamoreaux in 1997 at the U of Washington and then by Bressi, Carugno, Onofrio, Ruoso in 2002 at the University of Padua. There is now an entire field of Casimir nanophysics.

Good reads:

- For more on the justification and meaning of the use of ζ-regularisation in the Casimir effect and vacuum energy: https://en.wikiversity.org/wiki/Quantum_mechanics/Casimir_effect_in_one_dimension

- Terrence Tao on ζ-reg: https://terrytao.wordpress.com/2010/04/10/the-euler-maclaurin-formula-bernoulli-numbers-the-zeta-function-and-real-variable-analytic-continuation/
- For a concise derivation in three dimensions using ζ-reg: https://en.wikipedia.org/wiki/Casimir_effect#Derivation_of_Casimir_effect_assuming_zeta-regularization
- Some people are strongly critical or at least skeptic of vacuum fluctuations as the correct way to see the Casimir effect. A sensible argument is http://arxiv.org/abs/hep-th/0503158

Particle Physics/High Energy Physics

HEP1 - Why are photons massless?

Photons, being gauge bosons of a gauge theory, are a priori massless. At least classically, a mass term for the photon in the Lagrangian:

$$\frac{m^2}{8\pi} A_\mu A^\mu$$

is not invariant under the gauge transformation

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda$$

and thus gauge invariance disallows a photon mass. The massive photon Lagrangian one would instead obtain with the above term is the Proca theory of the massive vector boson.

However, photons could still in principle acquire a mass through at least a couple of mechanisms. The first is also present at the classical level and is the Higgs mechanism and variants, where gauge invariance is "broken", in some careful sense (gauge symmetry breaking is more involved than global symmetry breaking). Why the photon is not affected by the Higgs mechanism is treated in HEP2. The second happens upon quantization and is the quantum correction to the mass. Particles in general acquire quantum corrections to their physical parameters that can be investigated as being due with interaction with virtual particles as the original particle travels from point A to point B. The photon is again protected from this phenomenon and this is explained in HEP3.

HEP2 - Why do photons not acquire a mass through the Higgs mechanism?

By definition!

Photons are by definition the single component of the original $SU(2) \times U(1)$ electroweak gauge field that leaves the Higgs vacuum expectation value invariant.

This means that the VEV is uncharged for the photon, and the photon aquires no mass.

A little simpler: basically, $SU(2) \times U(1)$ is a four dimensional group of transformation. The Higgs is a field which takes value in a four-dimensional (two-complex dimensional) vector space and which is transformed ("rotated") by these transformations. Now after electroweak symmetry breaking the Higgs aquires a VEV, which just mean that in all of space it assumes the value of a specific vector in that 4-dimensional space. This vector is not invariant under the original gauge group, this means that it breaks the symmetry. There is however a 1-dimensional subgroup of the gauge group that still leaves the VEV invariant and thus that symmetry remains unbroken. That group's generator is defined to be the photon and the preserved gauge symmetry assures the photon has no mass. The other three generators instead do interact with the VEV and acquire mass. They are decomposed in three orthogonal generators by electric charge: W^+, W^-, Z^0.

And now, more detailed: assume WLOG that the Higgs has VEV as such:

$$\phi = \begin{pmatrix} 0 \\ \phi_0 \end{pmatrix}$$

with ϕ_0 real, and the generic gauge group element acts on ϕ as

$$\phi \to \exp\left(i\left(\frac{g'}{2}B \cdot \not{\mathbb{1}} + gW_1T^1 + gW_2T^2 + gW_3T^3\right)\right)\phi$$

Where B and W_i are respectively the gauge field for weak hypercharge ($U(1)$) and weak isospin ($SU(2)$), $T^i = \frac{\sigma^i}{2}$ are generators for $SU(2)$, and g, g' are the corresponding coupling constants.

We impose $\phi' = \phi$ to find the little group of the VEV (the isotropy group). For infinitesimal generators, the above reduces to:

$$\begin{pmatrix} g'B + gW_3 & g(W_1 + iW_2) \\ g(W_1 - iW_2) & g'B - gW_3 \end{pmatrix}\begin{pmatrix} 0 \\ \phi_0 \end{pmatrix} = 0$$

This gives immediately $W_1 = W_2 = 0$ (as they are real generators) and $g'B - gW_3 = 0$; this means that the generator $A = \frac{1}{\sqrt{g'^2+g^2}}(g'W_3 + gB)$ solution to the latter two equations generates the one-dimensional isotropy group. This generator is the photon.

The other three generators do modify the value of the Higgs and are orthogonalized into the $W^{\pm} = \frac{1}{\sqrt{2}}(W_1 \pm iW_2)$ and $Z^0 = \frac{1}{g^2+g'^2}(gW_3 - g'B)$.

The mass generation is evidenced by expanding the kinetic term of the Higgs Lagrangian around the VEV. This gives a mass term for the gauge boson which is precisely the norm squared of $\left(\frac{g'}{2}B \cdot \not{\mathbb{1}} + gW_iT^i\right)\begin{pmatrix}0\\ \phi_0\end{pmatrix}$. Therefore the orthonormal states A, $W^{\pm}$, Z^0 above diagonalize the mass matrix, and A is the single one with eigenvalue 0, as was shown before.

HEP3 - Why do photons not acquire a mass through quantum corrections/interaction with virtual particles? What is charge renormalization?

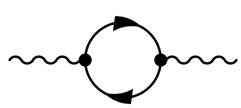
The short answer is that quantum corrections to the photon propagator do not give it a mass because of the Ward identity, which is a consequence of gauge invariance.

The probability amplitude for a photon to go from point A to point B is given by what is called the propagator. The “bare” propagator is given by the following expression:

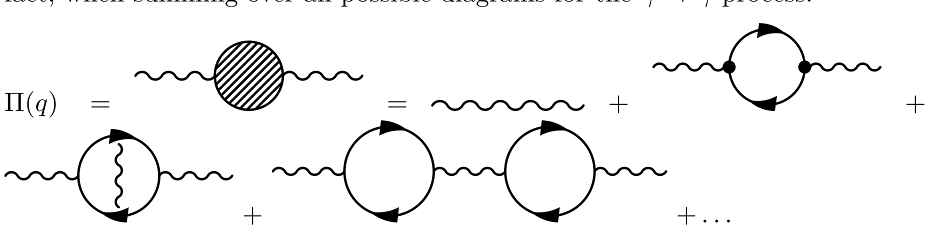
 $= \pi^{\mu\nu}(q) = \frac{-ig^{\mu\nu}}{q^2+i\epsilon}$

where q is the four-momentum of the photon and ϵ is a funny thing you shouldn’t worry about. In fact, put it to 0. This is how a massless propagator should look. The mass of the particle is given by where the propagator is singular (has a pole), so in this case $q^2 = m^2 = 0$. (If the photon was massive, in the denominator we would have $q^2 - m_\gamma^2 + i\epsilon$). The propagator has two Lorentz indices μ and ν because the photon has a polarization. To get the probability amplitude, you actually have to contract the indices of the propagator with your desired polarization ε_μ

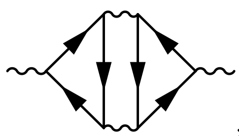
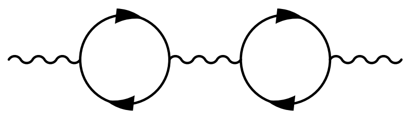
But this is just the bare propagator, this is all tree-level. We want to compute what happens to the propagator when we include higher-order corrections. More practically: the fact that a photon can produce a virtual electron-positron pair, which then reannihilate into a photon, as such:



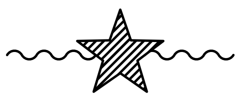
should affect the probability of the photon going from point A to point B. In fact, when summing over all possible diagrams for the $\gamma \to \gamma$ process:



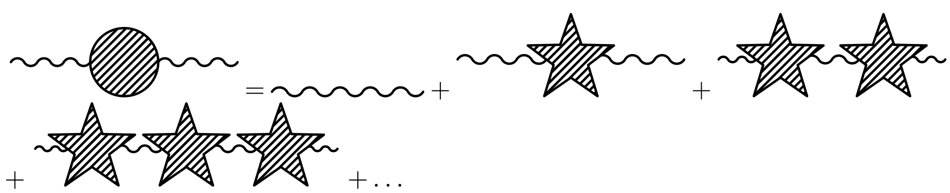
we should get the “dressed”, or physical propagator $\Pi^{\mu\nu}$. (A quick review of Feynman diagrams: time goes from left to right, wavy lines are photons, solid lines going right are electrons, left are positrons). Now we want to rearrange the terms of the previous series. Call a diagram one-particle-irreducible (1PI) if you cannot split it in two by cutting one of the internal lines (the two wavy lines at the beginning and end don't count). For example, the following diagrams are respectively 1PI and not 1PI.


,


Now consider the sum of all 1PI diagrams in the original series, and denote it as:

$S =$ 

My claim is that the full series of all diagrams for the photon propagator is equal to:



This is actually really simple. Think about it.

So, remembering that concatenating Feynman diagrams means we have to multiply them, we get the geometric series:

$$\Pi = \pi + \pi S\pi + \pi S\pi S\pi + \pi S\pi S\pi S\pi + \ldots = \pi\left(1 + S\pi + (S\pi)^2 + (S\pi)^3 + \ldots\right)$$

except these are matrices (with μ and ν indices), so we have to fix that before summing the geometric series. Here we introduce the fundamental Ward identity. If you have any process in which a photon is one of the external lines (incoming or outgoing particles), then the probability amplitude $\mathcal{M}$ satisfies

$$\mathcal{M}^\mu q_\mu = 0$$

where q_μ is the momentum of the photon.

Our 1PI propagator $S^{\mu\nu}(q)$ should satisfy the Ward identity. It must therefore be proportional to the projector on the subspace orthogonal to q_μ. It's not hard to convince oneself that this is given by

$$S^{\mu\nu}(q) = (q^2 g^{\mu\nu} - q^\mu q^\nu)\Phi(q^2) = \Delta^{\mu\nu}(q)\Phi(q^2)$$

Where $\Phi(q^2)$ is some scalar function (we also have used that the propagator must be Lorentz-invariant). Since Δ is a projector, $\Delta^2 = \Delta$ (in the sense of matrices). Moreover, also the product $\Delta' = \Delta\pi$ is a projector, as you can readily compute. We return to our series:

$$\Pi(q) = \pi\left(1 + \Delta'\Phi + \Delta'\Phi^2 + \Delta'\Phi^3 + \ldots\right) = \pi + \pi\left(\Delta'\left(\frac{1}{1-\Phi(q)} - 1\right)\right)$$

So, finally

$$\Pi^{\mu\nu}(q) = \frac{-i}{q^2(1-\Phi(q^2))}\left(g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}\right) + \frac{-i}{q^2}\left(\frac{q^\mu q^\nu}{q^2}\right)$$

Now, we can just drop all terms with $q_\mu q_\nu$, because the Ward identity tells us they will not contribute to the scattering amplitude (the Ward identity is just telling us that the longitudinal polarization of the photon is unphysical). Our final expression for the dressed propagator is

$$\Pi^{\mu\nu} = \frac{-ig^{\mu\nu}}{q^2(1-\Phi(q^2))}$$

which is identical to our original bare propagator π, just multiplied by the function $(1-\Phi(q^2))^{-1}$. Remembering what we said about poles and masses, for the photon mass to be preserved we would need that this function does not cancel the pole at $q^2 = 0$ of π. The cancellation would need $\Phi(q^2)$ to have itself a second-order pole at $q^2 = 0$, but this would seem impossible. However, it is absolutely not obvious, and in fact it is false in 2 spacetime dimensions. In the Schwinger model, in 2D, $\Phi(0)$ has a pole canceling q^2 and the photon mass is shifted to a finite value.

But our world is luckily higher-dimensional and the photon is safe from mass renormalization.

However, the propagator still acquires a global multiplicative factor

$$Z_3 := \frac{1}{1-\Pi(q^2)}$$

Since the photon propagator connects two interaction vertices, and these bring each a power of the fundamental charge e, the quantum correction to the photon propagators is making the electromagnetic interaction stronger. In particular, for each photon line there are two vertices and one factor of Z_3, so

$$e^2 \to Z_3 e^2$$

which basically mean we could understand this shift as $e \to \sqrt{Z_3}e$. This is why the procedure we performed above is called charge renormalization. We just derived the well-known fact that the strength of the EM interaction depends on the four-momenta involved.

To derive exactly that Φ has no unwanted pole and how exactly $e(q^2)$ runs with energy we would need to compute Φ explicitly (with some clever regularization, since it's divergent) and I don't want to do that. For that, refer to any good QFT/QED textbook.

Quantum Gravity/String Theory

QG1 - If the Planck length/Planck time is the smallest measurable/possible length/time, then...?

to answer.