Chapter 36

The Static Casimir Effect

Place a pair of parallel conducting plates sufficiently near each other and a force of attraction is generated. This is the Casimir effect (though possibly augmented by Van de Waals forces). The Casimir effect is often cited as evidence that zero point energy is real, because the force can be derived from the zero point energy of quantum fields in the vacuum state.

Last Update: 11/1/12

1. What Is Zero-Point Energy?

Consider the Hamiltonian of the quantum harmonic oscillator in one dimension, $\hat{H} = \frac{1}{2m}\hat{p}^2 + \frac{K}{2}x^2$. This Hamiltonian can be re-written in terms of the raising and lowering operators, $\hat{a}^+$ and $\hat{a}$, where $\hat{a} = \frac{\hat{p}}{\sqrt{2m}} - i\sqrt{\frac{K}{2}}\cdot x$, giving $\hat{H} = \hbar\omega\left(\hat{a}^+\hat{a} + \frac{1}{2}\right)$ where $\omega = \sqrt{K/m}$. In the vacuum state we have $\hat{a}|0\rangle = 0$, but the term ½ in the Hamiltonian implies that a quantum oscillator has non-zero energy even in the vacuum state, i.e., $\langle 0|\hat{H}|0\rangle = \frac{1}{2}\hbar\omega$.

The same situation is encountered for quantum fields. Consider a real scalar field with the usual Lagrangian $L = \frac{1}{2}\left[(\partial_\mu\phi)(\partial^\mu\phi) - m^2\phi^2\right]$. Expressing the free quantum field as a sum of plane waves in the usual way, $\phi(x) = \sum_{\bar{k}}\left(u_{\bar{k}}(x)a_{\bar{k}} + u^*_{\bar{k}}(x)a^+_{\bar{k}}\right)$, the Hamiltonian is found to be just the same as for the harmonic oscillator but with one oscillator for every mode. Thus we have,

$$\hat{H} = \sum_{\bar{k}}\hbar\omega_{\bar{k}}\left(\hat{a}^+_{\bar{k}}\hat{a}_{\bar{k}} + \frac{1}{2}\right) \qquad \text{hence,} \qquad \langle 0|\hat{H}|0\rangle = \frac{1}{2}\sum_{\bar{k}}\hbar\omega_{\bar{k}} \tag{1}$$

The sum is over an infinite number of modes, and the energy of the modes as $k \to \infty$ increases without limit in a spatial continuum. So every quantum field produces an infinite so-called 'zero point' energy (or vacuum energy) given formally by (1).

In most applications of quantum field theory, for example when calculating cross-sections in particle physics, this infinite zero point energy is simply shuffled away out of sight. This may be done in the form of 'normal ordering', for example, this being little more than a euphemism for removal of the unwanted zero point terms. But, assuming we could render the zero point energy finite in some way, could it be physically real? There is a hint that it might be. The spontaneous emission of radiation from a system in an excited state can be calculated as if it were stimulated emission, the stimulation being supplied by the assumed presence of one quantum in every mode of the field, see Schiff (1968).

Because the derivation of the Casimir force involves renormalisation it is desirable to perform the calculation in various independent ways to provide some confidence that the result does not depend upon the algorithm adopted for the renormalisation. The following sections present three different methods.

2. The Casimir Force – Derivation 1

Let us consider the Casimir set-up: a pair of parallel conducting plates. Suppose these are large square plates of side L placed a distance $d << L$ apart. Being conducting they impose boundary conditions on the electromagnetic field, so that the eigen-modes are changed compared to those in a vacuum with no plates present. This, it will be seen, can be regarded as the cause of the Casimir effect. For simplicity we will consider a scalar field as a surrogate for the electromagnetic field, and assume the boundary condition imposed by the plates is that the field vanishes, $\phi = 0$, on the plates, taken as $x = 0$ and $x = d$. The eigen-modes are thus,

$$\phi \propto \sin k_n x \cdot \sin k_y y \cdot \sin k_z z \qquad \text{where, } k_n = \frac{\pi n}{d} \text{ and, } n = 1,2,3... \qquad (2)$$

The modes are not quantised in the *y* and *z* directions. The sum over modes in these directions is thus replaced by integration as follows,

$$\sum_{y,z \text{ modes}} \rightarrow \left[\frac{L}{2\pi}\right]^2 \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dk_y dk_z = \left[\frac{L}{\pi}\right]^2 \int_0^{\infty}\int_0^{\infty} dk_y dk_z \qquad (3)$$

Considering a zero mass field we have $\omega_{\bar{k}} = |\bar{k}|$ (using $c = 1$) so that the sum over the *y* and *z* modes only gives,

$$\sum_{y,z \text{ modes}} \omega_{\bar{k}} = \left(\frac{L}{\pi}\right)^2 \cdot \int_0^{\infty}\int_0^{\infty} \sqrt{\left(\frac{\pi n}{d}\right)^2 + k_y^2 + k_z^2} \cdot dk_y dk_z \qquad (4)$$

Putting $\kappa = \sqrt{k_y^2 + k_z^2}$ gives,

$$\sum_{y,z \text{ modes}} \omega_{\bar{k}} = \left(\frac{L}{\pi}\right)^2 \frac{\pi}{2} \int_0^{\infty} \sqrt{k_n^2 + \kappa^2} \cdot \kappa d\kappa \qquad (5)$$

This is, of course, divergent. It can be made finite (regularised) by inserting a factor of $e^{-\lambda s}$, where $s = \sqrt{k_n^2 + \kappa^2}$. We shall need to take the limit $\lambda \rightarrow 0$ eventually, to make the answer cut-off independent. Hence,

$$\sum_{y,z \text{ modes}} \omega_{\bar{k}} = \frac{L^2}{2\pi} \int_{k_n}^{\infty} s^2 e^{-\lambda s} ds = \frac{L^2}{2\pi} \frac{\partial^2}{\partial \lambda^2} \int_{k_n}^{\infty} e^{-\lambda s} ds = \frac{L^2}{2\pi} \frac{\partial^2}{\partial \lambda^2} \left(\frac{e^{-\lambda k_n}}{\lambda} \right) \qquad (6)$$

The discrete sum over the quantised modes in the *x*-direction is now simple, since,

$$\sum_{n=1}^{\infty} e^{-\lambda k_n} = \sum_{n=1}^{\infty} e^{-\tilde{\lambda} n} = \frac{e^{-\tilde{\lambda}}}{1 - e^{-\tilde{\lambda}}} = \frac{1}{e^{\tilde{\lambda}} - 1} \qquad \text{where, } \tilde{\lambda} = \frac{\pi\lambda}{d} \qquad (7)$$

Substituting into (6) we get the sum over all modes to be,

$$\sum_{\text{all modes}} \omega_{\bar{k}} = \sum_{n=1}^{\infty} \frac{L^2}{2\pi} \frac{\partial^2}{\partial \lambda^2} \left(\frac{e^{-\lambda k_n}}{\lambda} \right) = \frac{\pi^2 L^2}{2d^3} \frac{\partial^2}{\partial \tilde{\lambda}^2} \left(\frac{1}{\tilde{\lambda}\left(e^{\tilde{\lambda}} - 1\right)} \right) \qquad (8)$$

where it is understood that we will eventually need to take the limit $\tilde{\lambda} \rightarrow 0$. The term to be differentiated can be expanded in powers of $\tilde{\lambda}$ giving,

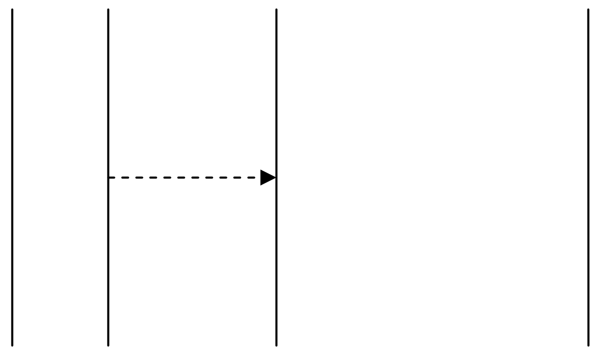
$$\frac{1}{\tilde{\lambda}\left(e^{\tilde{\lambda}}-1\right)}=\frac{1}{\tilde{\lambda}^2}-\frac{1}{2\tilde{\lambda}}+\frac{1}{12}-\frac{\tilde{\lambda}^2}{720}+O\left(\tilde{\lambda}^3\right) \qquad (9)$$

Hence, substituting in (8) and carrying out the differentiations,

$$\sum_{\text{all modes}}\omega_{\bar{k}}=\frac{\pi^2L^2}{2d^3}\left[\frac{6}{\tilde{\lambda}^4}-\frac{1}{\tilde{\lambda}^3}-\frac{1}{360}+O\left(\tilde{\lambda}\right)\right]=\frac{\pi^2L^2}{2}\left[\frac{6d}{\pi^4\lambda^4}-\frac{1}{\pi^3\lambda^3}-\frac{1}{360d^3}+O(\lambda)\right] \qquad (10)$$

The first two terms of (10) are divergent as $\lambda\to 0$. To render the zero point energy finite we renormalise it by considering only the *change* in the total energy of the 'universe' caused by having the plates at a small spacing. Consider the situation represented by Figure 1.

Figure 1



Here we have introduced another plate at a great distance, $D >> d$, from the Casimir plates. Instead of considering only the energy associated with the field *between* the plates at $x=0$ and $x=d$, we also consider the energy associated with the pair of plates at $x=d$ and $x=D$. The latter is just the same as (10) but with d replaced by $(D-d)$. So, representing the function on the RHS of (10) as $f(d)$, the total energy of the 'universe' is $f(d)+f(D-d)$. This will still be divergent, of course, since both contributions are divergent and of the same sign. However we suggest that the physically meaningful quantity is the difference between this total energy and that which would result from moving the Casimir plate at $x=d$ to $x=\eta D$, where $0<\eta<1$ so that $\eta D >> d$. In this latter configuration, both pairs of plates are so far apart that we expect the physically meaningful energy to be zero. Consequently this reference configuration acts as an appropriate energy datum which we can subtract from the original configuration. Hence, the renormalised energy is suggested as being,

$$f(d)+f(D-d)-\left(f(\eta D)+f((1-\eta)D)\right) \qquad (11)$$

When this combination is taken we see that the first term on the RHS of (10), which is linear in d, produces zero net contribution, i.e., $d+(D-d)-[\eta D+(1-\eta)D]\equiv 0$. Similarly the second term also gives zero net contribution. This means we can now take the limit $\lambda\to 0$ and obtain a finite result. Moreover, the terms of order λ and above now vanish. Consequently only the third, λ-independent term, in (10) survives. From (1) the total zero point energy of a massless scalar field due to the parallel conducting' plates is therefore, (Scalar) Total Renormalised Energy = $-\frac{\pi^2 L^2 \hbar c}{1440 d^3}$ (12)

where we have now explicitly shown the c dependence to facilitate numerical evaluations. For the electromagnetic field the two polarisation (spin) states result in the energy being double (12). Note the negative sign.

To obtain the force on the plates we use $F = -\frac{\partial E}{\partial d}$, giving, for the electromagnetic field now,

(e/m field) Casmir Pressure = $\frac{F}{L^2} = -\frac{\pi^2 \hbar c}{240 d^4}$ (13)

The minus sign denotes that the plates are attracted towards each other. For example, at a spacing of 100nm the Casimir pressure is -13 Pa.

3. The Casimir Force – Derivation 2

For simplicity we illustrate this second method for deriving a renormalised Casimir force in one dimension only. The method used in §2 gives the following result for 1D,

$$E = \frac{\hbar c}{2}\sum_{n=1}^{\infty}\frac{\pi n}{d}e^{-\lambda n} = -\frac{\pi\hbar c}{2d}\frac{\partial}{\partial\lambda}\sum_{n=1}^{\infty}e^{-\lambda n} = -\frac{\pi\hbar c}{2d}\frac{\partial}{\partial\lambda}\left(\frac{1}{e^{\lambda}-1}\right) = \frac{\pi\hbar c}{2d}\left(\frac{1}{\lambda^2} - \frac{1}{12} + O\left(\lambda^2\right)\right) \qquad (14)$$

Retaining only the finite term as $\lambda \to 0$ gives the renormalised Casimir energy for a massless scalar field in 1D as,

(1D Scalar) Total Renormalised Energy = $-\frac{\pi\hbar c}{24d}$ (15)

A slick way to this result is to note that, although $\sum_{n=1}^{\infty} n$ is obviously divergent, the Riemann zeta function defined by,

$$\varsigma(s) = \sum_{n=1}^{\infty} n^{-s} \qquad (16)$$

is, nevertheless, well defined at $s = -1$, namely $\varsigma(-1) = -1/12$ [see, for example, Abramowitz and Stegun (1965), or Edwards (1974)]. Using this bit of mathematical magic, the result for the 1D Casimir energy can be written straight down,

$$E = \frac{\hbar c}{2}\sum_{n=1}^{\infty}\frac{\pi n}{d} = \frac{\pi\hbar c}{2d} \times -\frac{1}{12} = -\frac{\pi\hbar c}{24d} \qquad (17)$$

in agreement with the previous method, (15). The series in (16) which defines the zeta function does not apply at $s = -1$, of course. However it does apply for $\Re(s) > 1$, and the resulting function can then be analytically continued to other parts of the Argand plane, including (it so turns out) to $s = -1$. This is how a finite value for $\varsigma(-1)$ is found. The mathematics of the zeta function is perfectly legitimate. It is, however, rather an act of faith that this will yield the correct renormalised energy, the connection being only the formal algebraic appearance of (16). However similar objections can be made for any regularisation/ renormalisation methodology.

Essentially the same method can be employed for 3D cavities, using generalised, or Epstein, zeta functions – see for example Lim and Teo (2007) which also addresses temperature effects.

4. The Casimir Force – Derivation 3

The third method to be discussed can be used to derive all components of the energy-momentum tensor of the zero-point field, though we shall concentrate on the energy density alone here. For the ordinary massless scalar field the energy density is given in terms of the field by,

$$T_{tt} = \frac{1}{2}\left[(\partial_t \phi)^2 + (\partial_x \phi)^2 + (\partial_y \phi)^2 + (\partial_z \phi)^2\right] \tag{18}$$

Define Hadamard's elementary function as the vacuum expectation value of the symmetrical product of the field at two different spacetime positions,

$$D^{(1)}(x, x') = \langle 0 | [\phi(x)\phi(x') + \phi(x')\phi(x)] | 0 \rangle \tag{19}$$

This is one of the several variants of propagator or Green's function for the scalar field. For a free, massless scalar field it can be evaluated explicitly to be,

$$D^{(1)}(x, x') = -\frac{1}{2\pi^2 (x_\mu - x'_\mu)(x^\mu - x'^\mu)} \tag{20}$$

The relevance of this function is that it is readily seen that the energy density, (18), can be expressed in terms of it, thus,

$$T_{tt} = \frac{1}{4} LIM_{x' \to x}\left[(\partial_t \partial_{t'} + \partial_x \partial_{x'} + \partial_y \partial_{y'} + \partial_z \partial_{z'}) D^{(1)}(x, x')\right] \tag{21}$$

Similar expressions apply for the other components of the stress energy tensor, $T_{\mu\nu}$, so knowledge of $D^{(1)}$ suffices to find them all.

In the case of a pair of parallel plates the Hadamard function can be found using the method of images. Thus each of the infinite number of images obtained by successive reflections in the two 'mirrors' gives rise to a term in the sum,

$$D_B^{(1)}(x, x') = \frac{1}{2\pi^2} \sum_{n=-\infty}^{\infty} \left\{ \begin{array}{l} \dfrac{1}{(x - x' - nd)^2 + (y - y')^2 + (z - z')^2 - (t - t')^2} \\ -\dfrac{1}{(x + x' - nd)^2 + (y - y')^2 + (z - z')^2 - (t - t')^2} \end{array} \right\} \tag{22}$$

where the sum in (22) is over even *n* only. The $n = 0$ term in (22) is discarded to render it finite when the limit $x' \to x$ is taken (this is the renormalisation process for this method). Substitution in (21) yields the energy density. However if the reader works this through he will not reproduce the results found in §2 and §3. The reason is that the energy density turns out to retain an *x*-dependence. This is due to the fact that the ordinary scalar field is not conformally invariant and the energy-momentum tensor has a non-zero trace, $T^\mu{}_\mu \neq 0$. As such it is a poor surrogate for the electromagnetic field. Instead we can employ the conformal scalar field for which the expression for the energy density is more complicated than (18), and leads to a correspondingly different relationship with the Hadamard function, i.e., different from (21). Nevertheless the above equations indicate the nature of the method and the details can be found in Birrell and Davies (1982), or in de Witt (1975). When applied to the conformal scalar field this method does produce the same result for the renormalised energy as 2 and 3. In fact, the whole energy-momentum tensor for the zero-point field is,

(Conformal Scalar) $$\left(T_{\mu\nu}\right) = -\frac{\pi^2 \hbar c}{1440 d^4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -1 & \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad (23)$$

The energy density, T_{tt}, is consistent with (12) when the latter is divided by the volume between the plates, $L^2 d$, to convert it to a density. The zero-point field has a positive pressure in the transverse (*y* and *z*) directions, but a negative pressure (or tension) in the direction separating the plates. This latter produces the attractive Casimir force between the plates.

5. Is the Casimir Force Always Attractive?

A Casimir force is not only produced between parallel plates. Other geometries of conducting materials will also give rise to a Casimir force. Moreover, the force is not confined to conducting materials but dielectrics will also produce a force. Finally, the nature of the dielectric fluid between the solid bodies will influence the Casimir force. The first demonstration that geometry could reverse the sign of the Casimir energy for conducting bodies, resulting in repulsion, was due to Boyer (1968). Boyer was interested in the force on a conducting spherical shell. The motivation was to attempt a model for the electron in which its electrostatic self-repulsion was stabilised by (he hoped) an attractive Casimir force. Lamentably Boyer found that the force was repulsive for a spherical geometry, so the model failed.

Since then it has been found that the Casimir force may be repulsive in a variety of cases, including cases involving dielectric materials. In particular, a repulsive Casimir force may be generated by a suitable choice for the dielectric properties of the fluid between the two bodies, compared with that of the bodies themselves. However, most experimental measurements of Casimir forces are in the attractive regime. Only recently has a measurement of a repulsive Casimir force been claimed, see Munday, Capasso and Parsegian (2009).

The measurement of a repulsive Casimir force is a major triumph, confirming a long standing prediction, though it presently requires independent replication. It is potentially of more than academic interest. Casimir forces generally have become of greater practical concern recently due to the rise of nanotechnology at a scale where Casimir forces become significant, even dominant. The usual attractive Casimir force can lead to unwanted stiction in nanoscale objects whose parts are intended to move. On the other hand, repulsive Casimir forces hold out the hope that such stiction effects could be avoided. They could even lead to levitation on these small size scales, and hence elimination of friction. A further potential use is for switching nanoscale devices, see Torricelli et al (2010).

6. Experimental Measurements of the Casimir Force

Historically there was a suspicion that the Casimir force might be explained as a Van der Waals force between the two surfaces. In fact, in all experiments both Van der Waals forces and Casimir forces operate. They can be distinguished, though, by their different dependence upon the spacing. All experiments involve investigating the variation of the force with changing separation. Typically Van der Waals forces may dominate for spacings less than a few tens of nm, whereas Casimir forces might dominate at ~100nm and above (though these figures will depend upon the situation).

Modern, accurate, measurements of the Casimir force between conducting surfaces arguably started with those of Lamoreaux (1997), and there has been significant experimental activity ever since. The magnitude of the force of attraction has been confirmed, though not with great precision in the case of flat parallel plates. This is probably because of the difficulty of ensuring precise parallel alignment of flat plates at separations of the very small order required (Bressi, et al, 2002, report confirmation of the theoretical force to an accuracy of ~15% in this configuration). However, experiments based on more favourable geometries, such as the force between a plane and a sphere, or between two cylinders, have reported confirming the theoretical force to an accuracy of ~1%, see for example, Mohideen and Roy (1998), Harris et al (2000), Ederth (2000) and Decca et al (2003). The claimed precision of the experiments over the last decade appears to be impressive confirmation of the Casimir effect. However, comparisons between theory and experiment are not easy. In particular temperature effects can be problematical. So the claim of ~1% agreement may be an overstatement. The reader is referred to recent reviews such as Klimchitskaya (2009) for a more complete account.

7. Is Zero-Point Energy Real?

Assuming that we are content to take the existence of the Casimir force as having been demonstrated by experiment, does this mean that the zero-point field is physically real? This does not necessarily follow, though the more precise the agreement between theory and experiment the more convincing we may find it. An interesting theoretical problem arises if we take the zero-point energy at face value. For an attractive Casimir force the absolute (renormalised) zero-point energy is negative. What are we to make of negative energy? In particular, how does it gravitate?

References

L.I.Schiff (1968), “Quantum Mechanics”, 3rd edition, McGaw-Hill (see §57).

S.C.Lim and L.P.Teo (2007), “Finite temperature Casimir energy in closed rectangular cavities: a rigorous derivation based on a zeta function technique”, J.Phys.A: Math.Theor. 40, 11645-11674.

M.Abramowitz and I.A.Stegun, “Handbook of Mathematical Functions”, Dover Publications, NY, 1965.

H.M.Edwards, “Riemann’s Zeta Fnction”, Dover Publications, NY, 1974.

N.D.Birrell and P.C.W.Davies, “Quantum Fields in Curved Space”, Cambridge University Press, 1982.

B.S.DeWitt, “Quantum Field Theory in Curved Spacetime”, Physics Reports, Vol.19C, No.6, August 1975.

J. N. Munday, F. Capasso, and V. A. Parsegian, Nature 457, 170 (2009).

T. H. Boyer, Phys. Rev. 174, 1764 (1968).

G. L. Klimchitskaya, U. Mohideen, V. M. Mostepanenko, The Casimir force between real materials: experiment and theory, Rev. Mod. Phys. v.81, N4, pp.1827-1885 (2009) arXiv:0902.4022

Bressi, G., Carugno, G., Onofrio, R., Ruoso, G., (2002), "Measurement of the Casimir Force Between Parallel Metallic Surfaces", quant-ph/0203002, 1 March 2002.

Decca, R.S., Lopez, D., Fischbach, E., Krause, D.E., (2003), "Measurement of the Casimir Force Between Dissimilar Metals", Phys.Rev.Lett. **91**, 050402, (arXiv:quant-ph/0306136)

Ederth, Thomas, (2000), "Template-Stripped Gold Surfaces with 0.4nm rms Roughness Suitable for Force Measurements: Application to the Casimir Force in the 20-100nm Range", Phys.Rev.A **62**, 062104 (arXiv:quant-ph/0008009).

Harris, B.W., Chen, F., Mohideen, U., (2000), "Improved Precision Measurement of the Casimir Force Using Gold Surfaces", quant-ph/0005088.

Lamoreaux, S.K., (1997), "Demonstration of the Casimir Force in the 0.6 to 6 Micron Range", Phys.Rev.Lett., **78**, 5-8, 6 Januray 1997.

Mohideen, U, and Roy, A., (1998), "A Precision Measurement of the Casimir Force from 0.1 to 0.9 Microns", Phys.Rev.Lett. **81**, 4549 (arXiv: physics/9805038)

G. Torricelli, P. J. van Zwol, O. Shpak, C. Binns, G. Palasantzas, B. J. Kooi, V. B. Svetovoy, and M. Wuttig, "Switching Casimir forces with phase-change materials", Phys. Rev. A 82, 010101 (2010)– Published July 9, 2010.

Move to Chapter on dynamic Casimir or drop,
Bressi, G., Carugno, G., Onofrio, R., Ruoso, G., (2008), Phys.A:Math.Theor. **41**, 312002. THIS ONE IS EXPERIMENT FOR DYNAMIC CASIMIR