The Nature of the Casimir Effect

Sami Ortoleva[1]

[1]*Department of Physics, University of California at San Diego, La Jolla, CA 92093-0319*

The Casimir effect has been of steadily increasing interest over the past six and a half decades. Here we discuss its origins from the zero point energy of the modes of the system. The importance and non-negligibility of the zero point energy is first motivated using ordinary quantum mechanics. The effect is discussed in the context of real scalar quantum field theory for parallel plates with Dirchlet boundary conditions. Finally, different geometries are discussed in the context of repulsive forces.

INTRODUCTION

The Casimir effect, first introduced in its modern form as a manifestation of zero point energy in 1948, has been of increasing interest. Indeed, citations of Casimir's original paper on the topic have been increasing exponentially with a half-life of only 12 years [1]. Much of this can be attributed to the confluence of work on the vacuum catastrophe in fundamental physics with the continuing ascension of MEMS/NEMS in the technological sector, bringing the problem under the purview of both pure theorists as well as applied physicists and engineers.

While zero point energy is a relatively uncontroversial aspect of ordinary quantum mechanics—likely on account of its finite magnitude in that regime—its discussion in the context of field theory remains tinged with discomfort in much the same way that renormalization seems to have been until the work of Wilson in the 1970s. Various authors have gone through considerable effort to recast the problem in terms of retarded interactions between the objects themselves [1]. Here we will focus on the mode expansion technique, directly addressing the zero point energy in the fields themselves. But first, let's see why the Casimir effect cannot be disposed of by a redefinition of the Hamiltonian.

ZERO POINT ENERGY IN ORDINARY QM

One of the classic problems in ordinary quantum mechanics is that of the one dimensional infinite square well. Consider such a well of width L with the minimum of the potential defined to be zero. Assuming the system is in its ground state, then the energy of the particle is $E = \hbar^2\pi^2/2mL^2$.

Consider now an extension of this problem. A (massive, thin) plate is placed within the well at some position x_p. Two distinguishable particles of equal mass are placed in the well, one on each side of the impenetrable plate. Then the ground state energy of the system is the sum of the ground state energies of each of the particles:

$$E = \frac{\hbar^2\pi^2}{2m}\left[\frac{1}{x_p^2} + \frac{1}{(L-x_p)^2}\right]$$

If we assume that the plate moves slowly enough for the adiabatic theorem to hold, then a (conservative) force can be calculated. The energy of these particles, which is an even function about $x_p = L/2$, gives rise to a force $F_p = -\partial_{x_p} E$ on the plate of the form:

$$F_p = \frac{\hbar^2\pi^2}{m}\left[\frac{1}{x_p^3} - \frac{1}{(L-x_p)^3}\right]$$

This force is directed towards $x_p/2$, and is a natural consequence of the zero point energy of the ground state of the infinite square well. While the Hamiltonian can be redefined to eliminate the zero point energy for a particular set of boundary conditions, allowing the boundaries themselves to be dynamical forces us to recognize that variations of the zero point energy with the system configuration gives rise to measurable consequences.

REAL SCALAR FIELDS

With the previous example from ordinary quantum mechanics as a guide, we now turn to the simplest quantum field theory. Define a real scalar field $\phi(\vec{x}, t)$ on $\vec{x} \in \Omega$ with Lagrangian density $\mathcal{L} = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2$. Generally speaking, we will be interested in embedding some objects—call them $\{O_i\}$—into Ω, enforcing physically appropriate boundary conditions for ϕ on the various spatial boundaries $\{\partial\Omega, \partial\Omega_{O_i}\}$, and computing the resulting variation in the zero point energy as a function of the location and/or orientation of O_i. The general procedure is as follows [2]:

- Determine the classical modes of the system and their associated temporal frequencies ω_j, where j is a stand-in for any suitable indexing of the modes.
- Compute the zero point energy as $E_0 = \sum_j \hbar\omega_j/2$ using an appropriate regulator.
- Extract measurable quantities from E_0, then remove the regulator.

Here we will discuss a parallel plate geometry embedded in a finite-volume cubic universe of side-length L, as done by Pluniun et al. [2]. Consider two (infinitely thin) plates, parallel to the x-y plane, that enforce $\phi = 0$ on their boundaries. We will use periodic boundary conditions and assume the plates are separated by a distance $d \ll L$. Then the modes *between* the plates have the spectrum

$$\omega_{n_x,n_y,n_z} = \sqrt{m^2 + \frac{\hbar^2\pi^2}{L^2}(n_x^2+n_y^2) + \frac{\hbar^2\pi^2}{d^2}n_z^2},$$

and those *outside* the plates have the spectrum

$$\omega_{n_x,n_y,n_z} = \sqrt{m^2 + \frac{\hbar^2\pi^2}{L^2}(n_x^2+n_y^2) + \frac{\hbar^2\pi^2}{(L-d)^2}n_z^2}.$$

Here $n_x, n_y \in \{... -1, 0, 1...\}$, and $n_z \in \{1, 2, 3...\}$ due to the loss of translational symmetry along $\hat{z}$.

Clearly, the zero point energy itself, $E_0 = \sum \hbar\omega/2$, will diverge. This is due to the presence of an infinite number of modes, since there are an infinite number of spatial points. However, differences in the zero point energy will turn out to be finite. To this end, define

$$\Delta E_0(d) \equiv E_0(d) - E_0^{single\ plate}.$$

Here, $E_0^{single\ plate}$ is the zero point energy of a single plate being in the volume, which is the finite-spatial-volume equivalent of taking $d \to \infty$. Then, converting the sums over n_x and n_y into integrals since L is taken to be large, we find:

$$\Delta E_0(d) = \frac{\hbar L^2}{2(2\pi)^2}\int d^2k_{\shortparallel}\left[\sum_{n_z=1}^{\infty}\sqrt{m^2+\vec{k}_{\shortparallel}^2+\left(\frac{2\pi}{d}\right)^2 n_z^2} - \int_0^{\infty} dn_z\sqrt{m^2+\vec{k}_{\shortparallel}^2+\left(\frac{2\pi}{d}\right)^2 n_z^2}\right] \tag{1}$$

Both the sum and the integral over n_z are divergent. To make sense out of this difference of infinite quantities mathematically, we must regularize each term in ΔE_0 first, then take the difference, and only afterwards remove the regulator [3]. But is there a physical justification for such a procedure?

In certain cases, there is. Take the example of the electromagnetic field, and assume the plates are infinitely conductive. This enforces the boundary condition $\hat{n} \times \vec{E} = 0$ on the electric field. However, any real conductor will have its conductivity $\sigma(\omega) \to 0$ as $\omega \to \infty$, if for no other reason than the inertia of its electrons. Thus, when computing ΔE_0, modes of sufficiently high frequencies do not see the plates at all, and therefore do not contribute to ΔE_0. So we may anticipate this by discarding those high frequency modes before carrying out the sums. Then, if ΔE_0 is found to be well behaved in the limit of no cutoff, we may remove the regulator. [3]

We might expect that such an argument can be made quite generically when dealing with material-induced boundary conditions. This is because the effect a material system has on the field must be due to the response of its constituent particles to the field's perturbations. All materials have a finite response time to external stimuli due to their constituent particles' inertia, and thus will have a linear response coefficient that tends to zero as $\omega \to \infty$, as before. Even materials with massless quasi-particle excitations are bound by this discussion, as the range of ω that is affinely related to $\vec{k}$ is limited.

Following this previous discussion, Pluniun et al. [2] calculated the Casimir energy, ΔE_0, in the low mass limit:

$$\Delta E_0(d) = \frac{-L^2\pi^2}{1440}\frac{1}{d^3}\left[1 + O(d^2m^2)\right],$$

and in the high mass limit:

$$\Delta E_0(d) = -\frac{L^2 m^{\frac{3}{2}}}{16\pi^{\frac{3}{2}}}\frac{1}{d^{\frac{3}{2}}}e^{-2md}.$$

Both of these cases show that $\Delta E_0(d)$ is minimized for $d \to 0$, and thus there exists an attractive force between the plates.

Plunien et al. describe the exponential damping of ΔE_0 in the high-mass case as "...reflecting the fact that the Casimir energy vanishes in the classical limit of particles with large mass." [2] It is a curious classical limit, as the zero point energy diverges regardless of the mass of the scalar particle, and fact that the field is massive is exactly what makes it have no clean classical limit as a field theory (as opposed to electromagnetism and gravitation). What perhaps can be said more precisely is that, in the large mass limit, the summand in (1) varies less from integer increments in n_z, and so the sum approaches the integral exponentially rapidly as $m \gg 2\pi/d$. In more physical terms, the mass dominates the energy of the modes at low $|\vec{k}|$, and modes with nearly $\vec{k}$-independent frequencies $\omega_{\vec{k}}$ do not contribute to the Casimir energy ΔE_0.

REPULSIVE CASIMIR FORCES

It is tempting to think that the Casimir force, which in the case of parallel plates is attractive, generally acts to compress an object O procided ϕ satisfies Dirchlet boundary conditions on $\partial\Omega_O$. The (*incorrect*) reasoning goes as follows: within the enclosed volume Ω_O, the spectrum of the Hamiltonian is discretized. Decreasing the size of Ω_O increases the spacing between the mode frequencies rapidly enough that, despite the smaller volume of Ω_O, the zero total point energy of the fields drops. Therefore the Casimir force is attractive, and tends to shrink enclosed volumes.

Indeed, Casimir himself seemed to think so, and used this to attempt to patch the classical Abraham-Lorentz model of the electron as a spherical shell of charge. In his augmented model, the supposedly attractive Casimir force would balance the electrostatic repulsion of the charge distribution, and the balance point would allow a purely geometric calculation of the fine structure constant by requiring the electrostatic energy to be equal to that of the rest mass of the electron.[4]

The heuristic argument for the generic attractive nature of the Casimir force does not hold, however, as Boyer showed in 1968 [4]. The failure in this thinking is that it neglects the evanescent surface modes on the exterior of $\partial\Omega_O$. The contribution of these modes is evidently large enough that the lower energy configuration is one where the radius $r \to \infty$ [1]. For a spherical conductor, the relevant Casimir energy *increases* as the radius $r \to 0$. Thus, spherical conductors tend to be expanded by the Casimir effect, with

$$\Delta E_0 = +\frac{0.04618...}{r}.$$

While the calculation of this result is rather involved, it seems to be rather general for spherical geometries for both vector and scalar fields in various dimensions [5]. Unfortunately, there does not seem to be any simple criteria for whether a given geometry experiences a compressive or expansive force.

It should be noted that a sphere is not the only known geometry for which repulsive Casimir forces have been shown to exist. Cylindrical-type geometries exhibit repulsive Casimir forces as well, though only if the length of the cylinder is long enough. For short prisms of various cross sections, including circular, the Casimir force is attractive. [5]

Acknowledgements

I'd like to thank Donald Knuth for writing TEX, and Leslie Lamport for building LATEX atop it. Without them I'd be wrestling with an ugly Word document.

[1] S. K. Lamoreaux, "The Casimir force: Background, experiments, and applications," Rept. Prog. Phys. **68**, 201 (2005).

[2] G. Plunien, B. Muller and W. Greiner, "The Casimir Effect," Phys. Rept. **134**, 87 (1986).

[3] H. B. G. Casimir, Indag. Math. **10**, 261 (1948) [Kon. Ned. Akad. Wetensch. Proc. **51**, 793 (1948)] [Front. Phys. **65**, 342 (1987)] [Kon. Ned. Akad. Wetensch. Proc. **100N3-4**, 61 (1997)].

[4] T. H. Boyer, "Quantum electromagnetic zero point energy of a conducting spherical shell and the Casimir model for a charged particle," Phys. Rev. **174**, 1764 (1968).

[5] K. A. Milton, E. K. Abalo, P. Parashar, N. Pourtolami, I. Brevik and S. A. Ellingsen, "Repulsive Casimir and Casimir-Polder Forces," J. Phys. A **45**, 374006 (2012) [arXiv:1202.6415 [hep-th]].