

Heuristic Explanation of the Reality of Energy in \mathcal{PT} -Symmetric Quantum Field Theory

Carl M. Bender
Department of Physics
Washington University
St. Louis, MO 63130, USA

1 Introduction

A \mathcal{PT} -symmetric quantum theory is described by a (typically non-Hermitian) Hamiltonian that commutes with \mathcal{PT} , where \mathcal{P} represents space reflection and \mathcal{T} represents time reversal [1]. Hermitian Hamiltonians are boring because their energies are always real. In contrast, \mathcal{PT} -symmetric Hamiltonians are interesting because they usually have a parametric region of unbroken \mathcal{PT} symmetry in which the eigenvalues are all real and a region of broken \mathcal{PT} symmetry in which the some eigenvalues are complex [2, 3]. These regions are separated by a phase transition which has been repeatedly observed in laboratory experiments [4, 5, 6, 7, 8, 9, 10, 11, 12, 13].

Standard examples of \mathcal{PT} -symmetric quantum-mechanical Hamiltonians whose eigenvalues are all real and positive are

$$H = p^2 + ix^3 \tag{1}$$

and

$$H = p^2 - x^4. \tag{2}$$

The D -dimensional Euclidean-space field-theoretic equivalents of these quantum theories are described by the Lagrangians

$$L = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 + ig\phi^3 \tag{3}$$

and

$$L = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 - g\phi^4. \tag{4}$$

The purpose of this article is to explain in heuristic field-theoretic terms why the energy levels of these Hamiltonians are real. We will concentrate specifically on the ground-state energies (free energies) of these models.

2 Why a ϕ^3 theory is bad but an $i\phi^3$ theory is good

It is well-known that the quantum-mechanical Hamiltonian

$$H = p^2 + x^3 \tag{5}$$

is unacceptable because it does not possess a ground state. This is because the potential $V(x) = x^3$ is not bounded below. Let us try to understand how this problem comes to light in a field-theoretic context.

The Lagrangian for the corresponding ϕ^3 quantum field theory is

$$L = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 + g\phi^3. \tag{6}$$

The Feynman graphical rules for this quantum field theory are

$$\begin{aligned} \text{vertex amplitude :} & \quad -6g, \\ \text{line amplitude :} & \quad \frac{1}{p^2 + m^2}. \end{aligned} \tag{7}$$

The Feynman graphs contributing to the ground-state energy (the free energy) of this quantum field theory are the connected vacuum graphs. There are two graphs of order g^2 ; these graphs are shown in Fig. 1. In Fig. 2 two connected vacuum graphs of order g^4 are shown. Note that vacuum graphs can only have an even number of vertices.

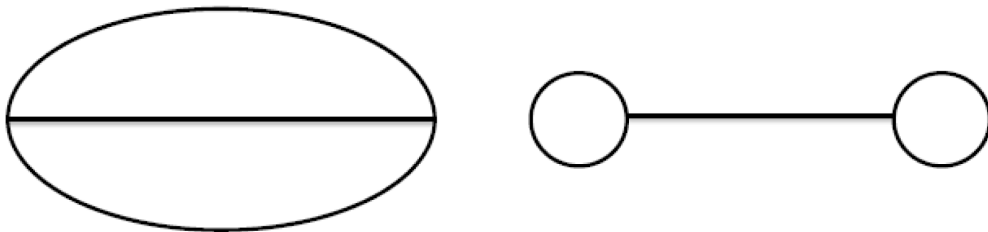


Figure 1: Connected vacuum graphs of order g^2 that contribute to the ground-state energy of a D -dimensional ϕ^3 quantum field theory.

If we combine the vacuum graphs to obtain the ground-state energy $E_0(g)$ of the theory, we obtain a representation of $E_0(g)$ as a formal Taylor series in powers of g^2 :

$$E_0(g) = \sum_{n=0}^{\infty} A_n g^{2n}. \tag{8}$$

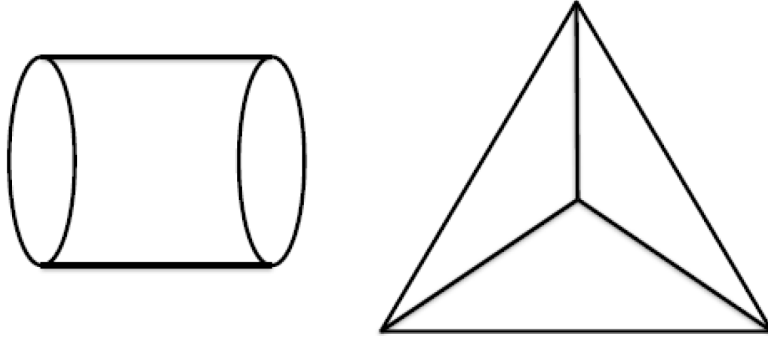


Figure 2: Two of the connected vacuum graphs of order g^4 that contribute to the ground-state energy of a ϕ^3 quantum field theory.

This series is divergent (it has a zero radius of convergence) because the coefficients A_n have factorial growth [14, 15, 16, 17]. In fact, A_n grows for large n like $n!$. The reason for the divergence of this series is simply that the graph amplitudes all have the same sign (they are all positive numbers) and the number of graphs having n vertices grows like $n!$.

If we try to use a summation technique such a Borel summation or Padé summation to sum the series in (8), we find that there is a cut in the g^2 plane on the positive axis. The discontinuity across this cut is the imaginary part of the ground-state energy. The fact that the energy is complex implies that the perturbative ground-state is unstable. The life-time of the ground state is just the reciprocal of the imaginary part of the ground-state energy.

There is a simple way to have a stable ground state in a cubic potential: We merely replace g by ig and obtain the \mathcal{PT} -symmetric Lagrangian in (3). The Feynman rules for this Lagrangian are

$$\begin{aligned} \text{vertex amplitude :} & \quad -6ig, \\ \text{line amplitude :} & \quad \frac{1}{p^2 + m^2}. \end{aligned} \tag{9}$$

Using these Feynman rules, we find that the perturbation series for the ground-state energy has the form

$$E_0(g) = \sum_{n=0}^{\infty} (-1)^n A_n g^{2n}. \tag{10}$$

Now the perturbation series alternates in sign and is Borel summable [18]. There is no discontinuity across the cut in the g^2 plane, and we conclude that the ground-state energy is real! Of course, this conclusion is not a surprise for the case $D = 1$ (quantum mechanics) because it was proved rigorously by Dorey, Dunning, and Tateo that the eigenvalues of the Hamiltonian in (1) are all real [19, 20].

This argument illustrates that the mechanism by which the eigenvalues of a cubic potential with an imaginary coupling become real has a very simple heuristic field-theoretic explanation. However, the argument is much more complicated and interesting for the case of a quartic potential, as we now show.

3 Ground-state energy for a quantum field theory with quartic coupling

Let us now consider a conventional quantum field theory with quartic coupling:

$$L = \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}m^2\phi^2 + g\phi^4. \quad (11)$$

The Feynman graphical rules for this quantum field theory are

$$\begin{aligned} \text{vertex amplitude :} & \quad -24g, \\ \text{line amplitude :} & \quad \frac{1}{p^2 + m^2}. \end{aligned} \quad (12)$$

Let us use these Feynman rules to calculate ground-state energy perturbatively. In Fig. 3 we show the unique vacuum graph of order one that contributes to the free energy. In Fig. 4 we display all three vacuum graphs of order two, but only the first two graphs contribute to the free energy because the third vacuum graph is disconnected.



Figure 3: The only vacuum graph to first order in g that contributes to the free energy of the Lagrangian in (11).

When the coupling constant g is positive (the case of the conventional ϕ^4 theory), the vertex amplitude in (12) is negative. Thus, the perturbative expansion

$$E_0(g) = \sum_{n=0}^{\infty} B_n g^{2n} \quad (13)$$

for the ground-state energy is an *alternating* series. As in the case of a cubic interaction, the coefficients B_n for this series grow like $n!$ for large n and the series is divergent. However, because the terms in the series alternate in sign, the Borel sum of

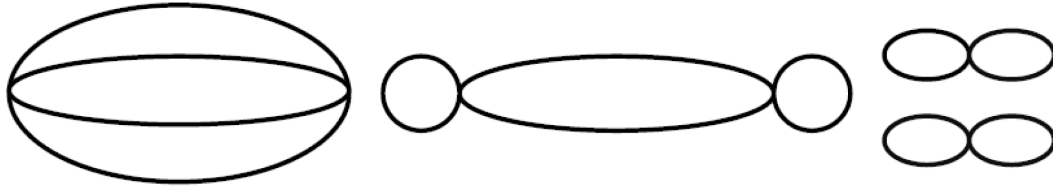


Figure 4: The three vacuum graphs of order two for the Lagrangian in (11). Only the first two graphs contribute to the free energy; the third is not connected, and therefore does not contribute.

the series does not have a cut on the positive axis in the complex- g plane. Hence, the ground-state energy is real when g is positive. This result is not surprising because it is well known that the spectrum for a quantum field theory with a positive quartic self-interaction is bounded below.

Let us follow the procedure used in Sec. 2 for the case of a cubic selfinteraction and replace g by $-g$. By this formal procedure, we replace the Lagrangian in (11) by that in (4). Now, the perturbation series in (13) no longer alternates in sign. Thus, we seem to reach the conclusion that the Borel sum of this perturbation series now has a cut on the positive- g axis and that the ground-state energy is complex. However, if this simple and convincing argument were correct, it would contradict the proof by Dorey, Dunning, and Tateo [19, 20] that the energy levels of the quantum-mechanical theory in (2) are real. It is hard indeed to see the flaw in this argument and the point of this paper is to explain this subtle aspect of \mathcal{PT} quantum theory.

In brief, we show in Sec. 4 that while the ground-state energy of an $ig\phi^3$ is entirely perturbative in character, the ground-state energy of a $-g\phi^4$ theory has both perturbative and nonperturbative contributions. We demonstrate that when the nonperturbative contributions are included, they exactly cancel the discontinuity across the cut in the Borel sum of the divergent perturbation series (13) and, as a result, the ground-state energy is purely real!

4 Perturbative and nonperturbative contributions to the free energy of a $-g\phi^4$ field theory

As a warm-up calculation, we begin by looking at the cubic $ig\phi^3$ theory. Then, we examine a $-g\phi^4$ theory.

4.1 The Cubic Theory

We begin by showing that there are *no* nonperturbative contributions to the ground-state energy of an $ig\phi^3$ theory. We construct this theory as a functional integral and for simplicity in this paper we work only in zero dimensions. When $D = 0$, the functional-integral representation becomes an ordinary integral. In order for a Euclidean-space functional integral of the general form $Z = \int \mathcal{D}\phi \exp(-\int d^D x L)$ to converge, it is necessary for the path of integration to terminate in the appropriate Stokes wedges in the complex- ϕ plane. For the cubic theory in zero dimensions this functional integral takes the form

$$Z = \int_C d\phi \exp\left(-\frac{1}{2}m^2\phi^2 - ig\phi^3\right), \quad (14)$$

where the integration contour C is shown in Fig. 5. Note that the Stokes wedges lie below the real axis in the complex- ϕ plane and are left-right symmetric (\mathcal{PT} symmetric). These wedges have an angular opening of 60° and are centered about -30° and -150° , as shown in Fig. 5.

We take g to be small ($g \ll 1$) and rescale the integral in (14) to obtain

$$Z = \frac{1}{g} \int_C d\phi \exp\left[-\frac{1}{g^2} \left(\frac{1}{2}m^2\phi^2 - i\phi^3\right)\right]. \quad (15)$$

We then perform a steepest-descent evaluation of the integral. There are two saddle points, a perturbative saddle point at the origin $\phi = 0$ and a nonperturbative saddle point on the positive-imaginary axis at $\phi = \frac{1}{3}im^2$. (We refer to the saddle point at $\phi = 0$ as *perturbative* because it gives rise to the Feynman perturbation series.) Note that the nonperturbative saddle point plays no role in the asymptotic evaluation of the integral because, as we can see from Fig. 5, the steepest-descent contour that terminates in the Stokes wedges is the hyperbola

$$y = \frac{1}{6} \left(m^2 - \sqrt{m^4 + 12x^2}\right), \quad (16)$$

where $\phi = x + iy$. This hyperbola only passes only through the saddle point at the origin and the path terminates in the centers of the two Stokes wedges. The steepest path through the nonperturbative saddle point is the hyperbola

$$y = \frac{1}{6} \left(m^2 + \sqrt{m^4 + 12x^2}\right). \quad (17)$$

This path is actually a steepest-ascent path, and it cannot terminate in the Stokes wedges. Thus, it plays no role in the calculation of the behavior of Z for small g . Evidently, the behavior of Z for small g is entirely perturbative in character and it is completely determined by the Feynman-diagram expansion in (10).

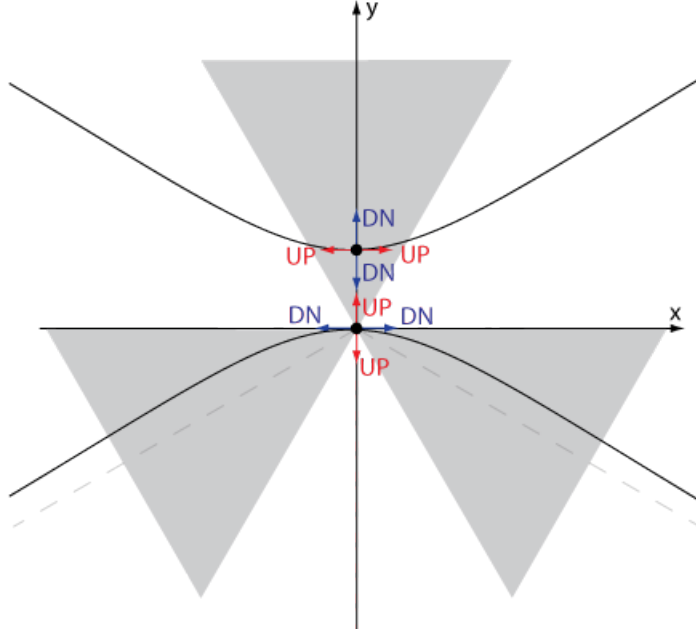


Figure 5: The complex- ϕ plane for the integral in (14). The perturbative and non-perturbative saddle points, the Stokes wedges, and steepest paths through the saddle points are shown. The relevant steepest-descent path passes through the perturbative saddle point at the origin. The path through the other saddle point on the positive-imaginary axis plays no role in the small- g asymptotic behavior of Z .

4.2 The Quartic Theory

Now let us consider the \mathcal{PT} -symmetric quartic quantum field theory described by

$$Z = \int_C d\phi \exp\left(-\frac{1}{2}m^2\phi^2 + g\phi^4\right). \quad (18)$$

The contour of integration C terminates in two Stokes wedges of angular opening 45° , one centered about -45° and the other centered about -135° . These Stokes wedges incorporate the \mathcal{PT} symmetry of the boundary conditions. The Stokes wedges are shown in Fig. 6. Treating g as small, we expand Z as a formal series in powers of g :

$$Z = \int_C d\phi e^{-m^2\phi^2/2} \sum_{n=0}^{\infty} \frac{g^n \phi^{4n}}{n!}. \quad (19)$$

Using Watson's Lemma [18], we now interchange orders of integration and summation and obtain the (divergent) weak-coupling asymptotic series expansion for Z :

$$Z \sim \frac{\sqrt{2\pi}}{m} \sum_{n=0}^{\infty} \frac{g^n (4n-1)!}{m^{4n} n!}. \quad (20)$$

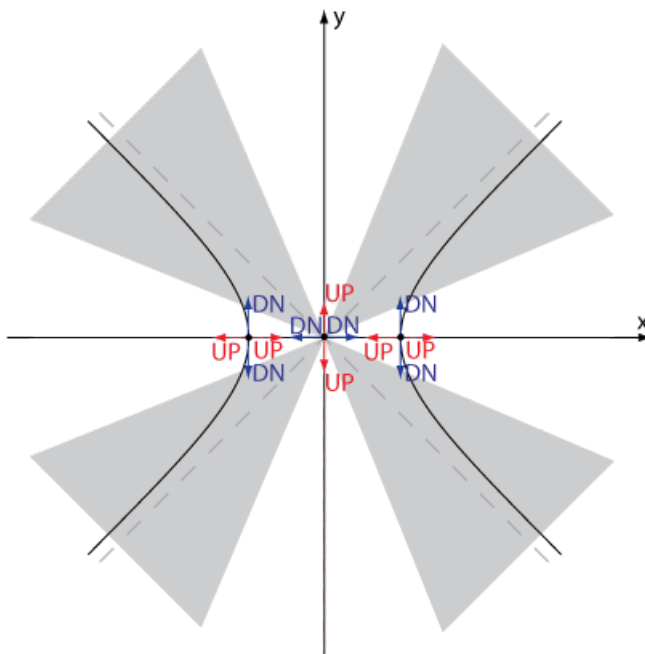


Figure 6: The complex- ϕ plane for the integral in (18). The perturbative and non-perturbative saddle points, the Stokes wedges, and steepest paths through the three saddle points are shown. The steepest-descent path passes through all three saddle points, the perturbative saddle point at the origin as well as the other two non-perturbative saddle points on the real axis. All three saddle points are needed to demonstrate that the small- g asymptotic behavior of Z is purely real.

Using the Stirling approximation $n! \sim n^n e^{-n} \sqrt{2\pi n}$ for large n , we can replace the series in (20) by the series

$$Z \sim \frac{1}{m\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{16^n g^n \Gamma(n-1/2)}{m^{4n}}. \quad (21)$$

We now perform a Borel sum of this series by introducing the identity

$$\Gamma(n-1/2) = \int_0^{\infty} dt e^{-t} t^{n-3/2} \quad (n \geq 1) \quad (22)$$

and interchanging orders of summation and integration. The result is

$$Z = \frac{1}{m\sqrt{\pi}} \int_0^\infty dt \frac{e^{-t}}{t^{3/2}(1-16gt/m^4)}. \quad (23)$$

(This integral is divergent at $t = 0$, but we can ignore this because the divergence is an artifact. It arises here because we have not bothered to exclude the $n = 0$ term in the sum. What is at issue is the imaginary part of this integral, and the imaginary part is not affected by the behavior at $t = 0$.)

Note that there is a discontinuity across the cut on the positive- g axis in the g plane. To determine this discontinuity, we use the identity

$$\frac{1}{x} = \mathbf{P} \left(\frac{1}{x} \right) + i\pi\delta(x), \quad (24)$$

where \mathbf{P} indicates *principal part*. We find that

$$\begin{aligned} \text{Im } Z &= \frac{\sqrt{\pi}}{m} \int_0^\infty dt e^{-t} t^{-3/2} \delta(1-16gt/m^4) \\ &= \frac{\sqrt{\pi}}{m} (16g/m^4)^{3/2} \exp\left(-\frac{m^4}{16g}\right). \end{aligned} \quad (25)$$

Now let us perform a steepest-descent analysis of the integral in (18) for Z . We begin by scaling out \sqrt{g} and obtain

$$Z = \frac{1}{\sqrt{g}} \int_C d\phi \exp \left[-\frac{1}{g} \left(\frac{1}{2} m^2 \phi^2 - \phi^4 \right) \right]. \quad (26)$$

There are three saddle points, a perturbative saddle point at the origin $\phi = 0$ and two nonperturbative saddle points on the real- ϕ axis at $\phi = \pm \frac{1}{2}m$. As is shown in Fig. 6, the steepest-descent path emerges from the perturbative saddle point at the origin and runs along the real axis until it hits the two nonperturbative saddle points. Then it turns downward and runs along the hyperbolas

$$y = -\sqrt{x^2 - m^2/4}, \quad (27)$$

which terminate as $x \rightarrow \pm\infty$ in the centers of the Stokes wedges.

A Gaussian approximation to the contribution of Z from the nonperturbative saddle points is imaginary and it *cancels exactly* the imaginary part of the perturbative contribution to Z in (25), which comes from the discontinuity across the cut in the Borel sum. This explains at a quantum-field-theoretic level the mechanism behind the remarkable property that non-Hermitian \mathcal{PT} -symmetric quantum theories have real eigenvalues.

The author is supported by the U.K. Leverhulme Foundation and by the U.S. Department of Energy.

References

- [1] C. M. Bender, Contemp. Phys. **46**, 277 (2005) and Repts. Prog. Phys. **70**, 947 (2007).
- [2] C. M. Bender and S. Boettcher, Phys. Rev. Lett. **80**, 5243 (1998).
- [3] C. M. Bender, S. Boettcher, and P. N. Meisinger, J. Math. Phys. **40**, 2201 (1999).
- [4] A. Guo, G. J. Salamo, D. Duchesne, R. Morandotti, M. Volatier-Ravat, V. Aimez, G. A. Siviloglou, and D. N. Christodoulides, Phys. Rev. Lett. **103**, 093902 (2009).
- [5] C. E. Rüter, K. G. Makris, R. El-Ganainy, D. N. Christodoulides, M. Segev, and D. Kip, Nat. Phys. **6**, 192-195 (2010).
- [6] J. Rubinstein, P. Sternberg, and Q. Ma, Phys. Rev. Lett. **99**, 167003 (2007).
- [7] K. F. Zhao, M. Schaden, and Z. Wu, Phys. Rev. A **81**, 042903 (2010).
- [8] Y. D. Chong, L. Ge, and A. D. Stone, Phys. Rev. Lett. **106**, 093902 (2011).
- [9] Z. Lin, H. Ramezani, T. Eichelkraut, T. Kottos, H. Cao, and D. N. Christodoulides, Phys. Rev. Lett. **106**, 213901 (2011).
- [10] N. M. Chtchelkatchev, A. A. Golubov, T. I. Baturina, and V. M. Vinokur, arXiv:1008.3590 [cond-mat.supr-con].
- [11] C. Zheng, L. Hao, and G. L. Long, arXiv:1105.6157 [quant-ph].
- [12] S. Bittner, B. Dietz, U. Guenther, H. L. Harney, M. Miski-Oglu, A. Richter, and F. Schaefer, arXiv:1107.4256 [quant-ph].
- [13] J. Schindler, A. Li, M. C. Zheng, F. M. Ellis, T. Kottos, arXiv:1109.2913 [cond-mat].
- [14] C. M. Bender and T. T. Wu, Phys. Rev. Lett. **21**, 406 (1968). “Analytic Structure of Energy Levels in a Field-Theory Model”
- [15] C. M. Bender and T. T. Wu, Phys. Rev. **184**, 1231 (1969).
- [16] C. M. Bender and T. T. Wu, Phys. Rev. Lett. **27**, 461 (1971).
- [17] C. M. Bender and T. T. Wu, Phys. Rev. D **7**, 1620 (1973), Reprinted in J. C. Le Guillou and J. Zinn-Justin (eds.), *Large-order behaviour of perturbation theory* (North-Holland, 1990), p. 41-57.

- [18] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw Hill, New York, 1978).
- [19] P. Dorey, C. Dunning, and R. Tateo, J. Phys. A: Math. Gen. **34**, L391 (2001) and **34**, 5679 (2001).
- [20] P. Dorey, C. Dunning, and R. Tateo, J. Phys. A: Math. Gen. **40**, R205 (2007).