Divergent series in quantum mechanics and quantum field theory: The issue of Borel summability

J. Zinn-Justin*

CEA-IRFU, Paris–Saclay campus,
F-91191 Gif-sur-Yvette cedex, FRANCE.

*email: jean.zinn-justin@cea.fr

Talk given at CERN on November 20, 2013.

Steepest descent method and asymptotic series

Let $S(x)$ be a real entire function and such that the integral

$$Z(g) = \frac{1}{\sqrt{2\pi g}} \int \mathrm{d}x \; \mathrm{e}^{-S(x)/g},$$

converges for $g > 0$. We also assume that $S(x)$ has a unique minimum at $x = 0$ where $S(x) = x^2/2 + O(x^3)$. For $g \to 0_+$, $Z(g)$ can be evaluated by the steepest descent method. Since the integral is dominated by the saddle point $x = 0$, this involves expanding $S(x)$ around $x = 0$. The integral is then formally given by an expansion of the form

$$Z(g) \approx \sum_{k=0}^{\infty} Z_k g^k .$$

Except for the Gaussian integral, the expansion cannot be convergent for any $g \neq 0$ because for $g < 0$ the integral diverges. The function $Z(g)$ thus has a cut starting from $g = 0$.

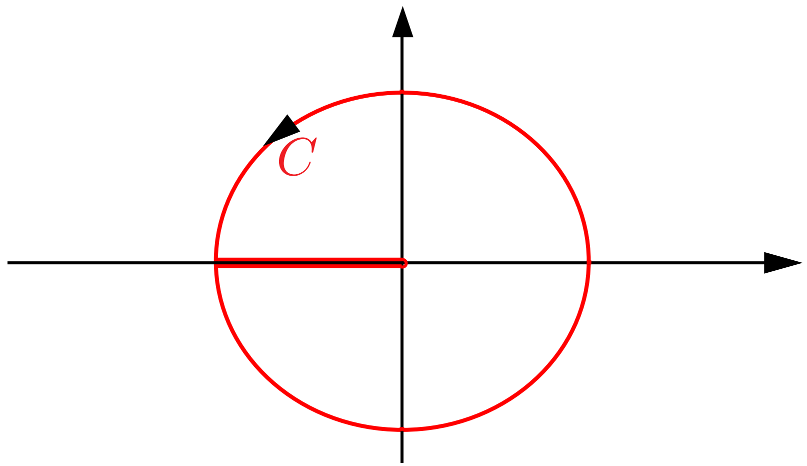


Fig. 1 Cauchy contour C_ε: one-cut example.

Large order behaviour and the steepest descent method

For $|g|$ small enough, one can use the Cauchy representation (see figure)

$$Z(g) = \frac{1}{2i\pi} \int_{C_\varepsilon} \mathrm{d}g' \, \frac{Z(g')}{g' - g} \,.$$

One shows that, for $g \to 0$, the discontinuity of $Z(g)$ on cuts is dominated by the contributions of other non-trivial (complex) saddle points.

One then finds

$$\text{disc.}Z(g) \underset{g\to 0}{\propto} \sum_{\text{leading saddle points } x_s} \mathrm{e}^{-S(x_s)/g} .$$

Due to the exponential decrease of disc. $Z(g)$ for $g \to 0$, near $g = 0$ one can replace the contour integral by the integral of disc. $Z(g)$ on the cut and expand the integral in powers of g. One infers

$$Z_k = \frac{1}{2i\pi} \int_{C_0} g^{-k-1} Z(g) \mathrm{d}g .$$

For $k \to \infty$, the integral is dominated by small values of g.

Using the behaviour of disc. $Z(g)$ for $g \to 0$, one thus finds

$$Z_k \underset{k\to\infty}{\sim} \frac{1}{2i\pi} \int^{0} g^{-k-1} \text{disc.}Z(g) \mathrm{d}g$$

$$\propto \sum_s \int^{0} g^{-k-1} \, \mathrm{e}^{-S(x_s)/g} \, \mathrm{d}g \propto k! \sum_s S^{-k}(x_s) .$$

Path and field integrals

The arguments (and the generic $k!$ behaviour) generalize at once to multiple integrals, path and field integrals in the context of loopwise or semi-classical expansions. The large order behaviour is then dominated by instantons, finite action solutions (in general complex), of euclidean classical field equations. However, there are two restrictions:

(i) Field theories with fermions and no boson self-interactions diverge with a smaller power of $k!$ due to the Pauli principle.

(ii) In renormalizable field theories (by contrast with super-renormalizable) additional contributions to the large order behaviour may be expected, generated by UV or, in massless theories, IR singularities.

Example: The $\phi^4_{d<4}$ field theory

In a $g\phi^4$ field theory, correlation functions are analytic in a cut-plane. For $g \to 0_-$, the saddle point solution $\phi(x) \equiv 0$ of the field equation dominates the real part while the non-trivial instanton solutions dominate the imaginary part on the cut. The euclidean field equation reads

$$(-\Delta_x + m^2)\phi(x) + gm^{4-d}\phi^3(x)/6 = 0 \,.$$

The leading instanton solution has the general form

$$\phi(x) = \frac{1}{\sqrt{-g}} m^{(d-2)/2} f(mr), \quad r = |x|.$$

The leading contribution yields

$$Z_k \sim \int_{-\infty}^{0} \mathrm{e}^{A/g}\, g^{-k-1} \mathrm{d}g \propto k!\, (-A)^{-k} \,,$$

where $-A$ is the instanton action. For $d = 3$, $A = 113.38350781527714(1)$.

Asymptotic series and Borel summability

For simplicity, we consider only the relevant example of $k!$ divergent series.

Let a function $F(z)$ a function analytic in a sector

$$S = \{0 < |z| \le \zeta\,, \quad |\mathrm{Arg}\, z| \le \alpha/2\},$$

where it satisfies

$$|F(z) - \sum_{k=0}^{n-1} F_k z^k| \le M n! \left(\frac{|z|}{A}\right)^n \quad \forall n > 0\,.$$

By contrast with a convergent series, the formal (asymptotic) series

$$[F](z) = \sum_{k=0} F_k z^k\,,$$

for $z \neq 0$ determines the function F only up to a finite error since the best estimate is obtained for a finite value of n. Indeed,

$$\min_n M\, A^{-n} |z|^n n! \underset{|z| \ll 1}{\propto} \mathrm{e}^{-A/|z|}\,.$$

The asymptotic series does not define, in general, a unique function, but only a class of functions that differ by terms of order $e^{-A/z}$.

However, for $\alpha > \pi$, a theorem states that any analytic function bounded by $e^{-A/|z|}$ vanishes. Then, the asymptotic series defines a unique function and is called Borel-summable. The function F can be recovered by introducing the auxiliary function

$$B_F(z) = \sum_k \frac{F_k}{k!} z^k$$

The function $B_F(z)$ is analytic is the union of a circle $|z| < A$, where the series converges, and a sector $|\mathrm{Arg}\, z| < \alpha/2 - \pi/2$.

F is then the Borel–Laplace transform of B_F:

$$F(z) = \int_0^\infty \mathrm{d}t\ e^{-t}\, B_F(zt)\mathrm{d}t\,.$$

Borel summation: application, the $(\phi^2)^2_{d<4}$ field theory

Number of interesting, continuous, phase transitions are described by the N-vector model, an $O(N)$ symmetric statistical field theory with an N-component field $\boldsymbol{\phi}(x)$. The partition function then reads

$$\mathcal{Z} = \int [\mathrm{d}\boldsymbol{\phi}(x)] \exp\left[-\mathcal{H}(\boldsymbol{\phi})\right],$$

where the Hamiltonian (or euclidean action) is given by

$$\mathcal{H}(\boldsymbol{\phi}) = \int \left\{ \frac{1}{2} \sum_{\mu} \left[\partial_\mu \boldsymbol{\phi}(x)\right]^2 + \frac{1}{2} r \boldsymbol{\phi}^2(x) + \frac{u}{4!} \left[\boldsymbol{\phi}^2(x)\right]^2 \right\} \mathrm{d}^d x \,.$$

The first values of N correspond to the transitions:

$N = 1$: liquid–vapour, binary mixtures, Ising-like ferromagnetic systems

$N = 2$: Helium superfluidity

$N = 3$: isotropic ferromagnetic systems

and $N = 0$ to statistical properties of long polymer chains or SAW.

Callan–Symanzik (CS) equations

Critical exponents can be derived from RG functions, calculated in the massless field theory, using the Wilson–Fisher (divergent) $\varepsilon = (4-d)$-expansion. However, as suggested by Parisi, one can also work at fixed dimension $d < 4$ in the massive theory (the critical domain), where, by contrast with the massless theory, the perturbative expansion is IR finite. Then, renormalized vertex functions can be defined by the renormalization conditions

$$\tilde{\Gamma}^{(2)}(p;m,g) = m^2 + p^2 + O(p^4), \quad \tilde{\Gamma}^{(4)}(0,0,0,0;m,g) = gm^{4-d}.$$

Vertex functions satisfy the CS equations

$$\left[m\frac{\partial}{\partial m} + \beta(g)\frac{\partial}{\partial g} - \frac{n}{2}\eta(g)\right]\tilde{\Gamma}^{(n)}(p_i;m,g) = m^2(2-\eta)\tilde{\Gamma}^{(n)}_{\phi^2}(p_i;m,g),$$

where $\Gamma^{(n)}_{\phi^2}$ is a vertex function with one $\int \mathrm{d}^d x\, \phi^2(x)$ insertion.

The fixed dimension scheme

RG functions are calculated directly in dimension 3 but, by contrast with the $\varepsilon = (4-d)$-expansion scheme, one no longer has a 'small' expansion parameter. However, Nickel has noticed that Feynman diagrams in dimension 3 can be more easily evaluated than in $\varepsilon = (4-d)$-dimensions. He has managed to calculate all diagrams contributing to RG functions η, η_2 (anomalous dimensions of ϕ and ϕ^2) up to seven loops, and the diagrams contributing to the β-function up to six loops.

For example, for $N = 1$, Nickel has obtained ($\tilde{g} = 3g/(16\pi)$)

$$\beta(\tilde{g}) = -\tilde{g} + \tilde{g}^2 - \tfrac{308}{729}\tilde{g}^3 + 0.3510695978\tilde{g}^4 \\ - 0.3765268283\tilde{g}^5 + 0.49554751\tilde{g}^6 - 0.749689\tilde{g}^7 + O\left(\tilde{g}^8\right).$$

At large orders, the coefficient of $\tilde{g}^k$ then behaves like

$$\beta_k \underset{k\to\infty}{\propto} (-a)^k k^b k!\,, \quad a = 0.147774232\ldots\,, \; b = 11/2\,.$$

The numerical determination of the zero of the β-function, which is a number of order 1, as well as the calculation of exponents, clearly requires a summation of the series.

In three dimensions, the perturbative expansion is proved to be Borel summable. It is thus natural to introduce the Borel–Laplace transformation (here, Borel–Leroy). For example, for the β-function one defines

$$B_\beta(\sigma, g) = \sum_k \frac{\beta_k}{\Gamma(k + \sigma + 1)} g^k ,$$

(σ is an adjustable parameter) where the function $B_\beta(\sigma, g)$ is analytic in a circle of radius $1/a$ and a neighbourhood of the positive real axis. Then,

$$\beta(g) = \int_0^{+\infty} t^\sigma \, \mathrm{e}^{-t} \, B_\beta(\sigma, gt) \mathrm{d}t \, .$$

The series defines the function B_β only in a circle. It is necessary to perform an analytic continuation, at least in the neighbourhood of the real positive axis. In practice, with a small number of terms, a precise continuation requires a domain of analyticity larger than what is proved. Le Guillou and Z.-J. (1977–1980) have assumed maximal analyticity compatible with existing knowledge, *i.e.*, analyticity in a cut-plane. A conformal mapping of the cut-plane onto a circle has then provided the required continuation.

Improved summation techniques and additional seven-loop contributions have later led to improved estimates of critical exponents.

Reference: R. Guida and J. Zinn-Justin, *J. Phys. A* 31 (1998) 8103, cond-mat/9803240, an improvement over the results first published in

J.C. Le Guillou and J. Zinn-Justin, *Phys. Rev. Lett.* 39 (1977) 95; *Phys. Rev.* B21 (1980) 3976.

Table 2

Series summed by the method based on Borel transformation and mapping for the zero $\tilde{g}^*$ of the $\beta(g)$ function and the exponents γ and ν in the ϕ_3^4 field theory.

k	2	3	4	5	6	7
$\tilde{g}^*$	1.8774	1.5135	1.4149	1.4107	1.4103	1.4105
ν	0.6338	0.6328	0.62966	0.6302	0.6302	0.6302
γ	1.2257	1.2370	1.2386	1.2398	1.2398	1.2398

The parameter-free results obtained by this method, based on first principle field theory calculations, have survived more than 30 years of confrontation with experimental results, as well as lattice calculations. However, the progress in precision of the latter, as well as the superfluid experimental results, should be an encouragement to try to evaluate additional terms of the perturbative series.

Critical exponents from the $O(N)$ symmetric $(\phi^2)^2_3$ field theory

N	0	1	2	3
$\tilde{g}^*$	1.413 ± 0.006	1.411 ± 0.004	1.403 ± 0.003	1.390 ± 0.004
g^*	26.63 ± 0.11	23.64 ± 0.07	21.16 ± 0.05	19.06 ± 0.05
γ	1.1596 ± 0.0020	1.2396 ± 0.0013	1.3169 ± 0.0020	1.3895 ± 0.0050
ν	0.5882 ± 0.0011	0.6304 ± 0.0013	0.6703 ± 0.0015	0.7073 ± 0.0035
η	0.0284 ± 0.0025	0.0335 ± 0.0025	0.0354 ± 0.0025	0.0355 ± 0.0025
β	0.3024 ± 0.0008	0.3258 ± 0.0014	0.3470 ± 0.0016	0.3662 ± 0.0025
α	0.235 ± 0.003	0.109 ± 0.004	-0.011 ± 0.004	-0.122 ± 0.010
ω	0.812 ± 0.016	0.799 ± 0.011	0.789 ± 0.011	0.782 ± 0.0013
$\omega\nu$	0.478 ± 0.010	0.504 ± 0.008	0.529 ± 0.009	0.553 ± 0.012

A non-Borel summable example: the quartic double-well potential

While it is not always simple to prove Borel summability, by contrast, it is often easier to verify non-Borel summability. For instance, if a real function, analytic in a sector centred around the real positive axis, has an asymptotic expansion where, asymptotically, all terms have the same sign, the series cannot be Borel summable. Examples are provided by path (field) integrals with real instantons and, thus, physical barrier penetration.

This occurs in quantum mechanics for potentials with degenerate minima. Quantum tunnelling generates additional contributions (instanton contributions) to energy eigenvalues of order $\exp(-\text{const.}\ /\hbar)$, which have to be added to the expansion in powers of $\hbar$, and implies non-Borel summability.

In non-Borel summable situations, the determination of eigenvalues starting from their expansion for $\hbar$ small is a non-trivial problem.

The problem has been explicitly solved first in the case of the quartic double-well potential and the solution has later been generalized.

Generalized Bohr–Sommerfeld quantization formulae

The exact form of the small $\hbar$ expansion has been conjectured first in the case of the quartic double-well potential. We describe and motivate here the conjecture in this case, though the conjecture has been later generalized to potentials that are more general entire functions.

Note that, in what follows, the symbol g plays the role of $\hbar$ and the energy eigenvalues are measured in units of $\hbar$, a normalization adapted to perturbative expansions (by contrast with WKB expansions).

A few references:

J. Zinn-Justin, *Nucl. Phys.* B192 (1981) 125; B218 (1983) 333; *J. Math. Phys.* 25 (1984) 549.

U.D. Jentschura and J. Zinn-Justin, *Phys. Lett.* B 596 (2004) 138-144.

J. Zinn-Justin and U.D. Jentschura, *Ann. Phys. NY* 313 (2004) 197-267; *ibidem* 313 (2004) 269-325; U.D. Jentschura and J. Zinn-Justin, *Ann. Phys. NY* 326 (2011) 2186-2242.

The quartic double-well potential

The Hamiltonian corresponding to the quartic double-well potential can be written as

$$H = -\frac{g}{2}\left(\frac{\mathrm{d}}{\mathrm{d}q}\right)^2 + \frac{1}{g}V(q), \quad V(q) = \frac{1}{2}q^2(1-q)^2.$$

The Hamiltonian is symmetric in the exchange $q \leftrightarrow 1-q$ and thus commutes with the corresponding reflection operator P:

$$P\psi(q) = \psi(1-q) \ \Rightarrow [H, P] = 0\,.$$

The eigenfunctions $\psi_{\epsilon,N}(q)$ of H, where $\epsilon = \pm 1$, satisfy

$$H\psi_{\epsilon,N}(q) = E_{\epsilon,N}(g)\psi_{\epsilon,N}(q), \quad P\psi_{\epsilon,N}(q) = \epsilon\psi_{\epsilon,N}(q),$$

and $E_{\epsilon,N}(g) = N + 1/2 + O(g)$.

We have conjectured that the eigenvalues $E_{\epsilon,N}(g)$ have a complete (hyper-asymptotic) semi-classical expansion of the form

$$E_{\epsilon,N}(g) = \sum_{n=0} \epsilon^n E_N^{(n)}(g) \quad \text{with} \quad E_N^{(0)}(g) = \sum_k E_{Nk}^{(0)} g^k \text{ and}$$

$$E_N^{(n>0)}(g) = \left(\frac{2}{g}\right)^{Nn} \left(-\frac{\mathrm{e}^{-1/6g}}{\sqrt{\pi g}}\right)^n \sum_{l=0}^{n-1} [\ln(-2/g)]^l \, E_{Nl}^{(n)}(g),$$

where $E_N^{(0)}(g)$ and $E_{Nl}^{(n)}(g)$ are power series not Borel summable for $g > 0$.

In this expansion, they have to be first summed for g negative (where they are Borel-summable) and the values for g positive are then obtained by analytic continuation, consistently with the determination of $\ln(-g)$. In the analytic continuation from g negative to g positive, the Borel sums become complex with imaginary parts exponentially smaller by about a factor $\mathrm{e}^{-1/3g}$ than the real parts. These imaginary contributions are cancelled by the perturbative imaginary parts generated from the function $\ln(-2/g)$.

Moreover, we have conjectured that all the series $E_N^{(0)}(g), E_{Nl}^{(n)}(g)$ can be obtained by expanding, for $g \to 0$, a generalized Bohr–Sommerfeld quantization formula, which in the case of the double-well potential reads

$$\frac{1}{2\pi}\Gamma^2(\tfrac{1}{2}-B)\left(-\frac{2}{g}\right)^{2B(E,g)} \mathrm{e}^{-A(E,g)} + 1 = 0\,.$$

The functions A and B are given by asymptotic series of the form

$$B(E,g) = -B(-E,-g) = E + \sum_{k=1}^{\infty} g^k b_{k+1}(E),$$

$$A(E,g) = -A(-E,-g) = \frac{1}{3g} + \sum_{k=1}^{\infty} g^k a_{k+1}(E).$$

The coefficients $b_k(E)$ and $a_k(E)$ are polynomials in E, alternatively odd or even, of degree k. The three first orders, for example, are

$$B(E,g) = E + g\left(3E^2 + \tfrac{1}{4}\right) + g^2\left(35E^3 + \tfrac{25}{4}E\right) + O\left(g^3\right),$$

$$A(E,g) = \tfrac{1}{3}g^{-1} + g\left(17E^2 + \tfrac{19}{12}\right) + g^2\left(227E^3 + \tfrac{187}{4}E\right) + O\left(g^3\right).$$

Multi-instanton contributions at leading order

When the functions A and B are approximated by their leading order, resp., the spectral equation reduces to ($\epsilon = \pm 1$)

$$\frac{e^{-1/6g}}{\sqrt{2\pi}}\left(-\frac{2}{g}\right)^{E} = -\frac{\epsilon i}{\Gamma(\frac{1}{2}-E)} \Leftrightarrow \frac{\cos \pi E}{\pi} = \epsilon i \frac{e^{-1/6g}}{\sqrt{2\pi}}\left(-\frac{2}{g}\right)^{E}\frac{1}{\Gamma(\frac{1}{2}+E)} \,.$$

From the viewpoint of the path integral representation, this approximated spectral equation then corresponds to a summation of the leading order multi-instanton contributions.

For example, the term of order $e^{-1/6g}$ in the expansion of E_N,

$$E_N^{(1)}(g) = -\frac{1}{N!}\left(\frac{2}{g}\right)^{N+1/2}\frac{e^{-1/6g}}{\sqrt{2\pi}}\left(1+O(g)\right),$$

is the one-instanton contribution.

The term of order $\mathrm{e}^{-2/6g}$,

$$E_N^{(2)}(g) = \frac{1}{(N!)^2}\left(\frac{2}{g}\right)^{2N+1}\frac{\mathrm{e}^{-1/3g}}{2\pi}\left[\ln(-2/g) - \psi(N+1) + O\left(g\ln g\right)\right],$$

$(\psi = (\ln\Gamma)')$ can be identified with the two-instanton contribution.

More generally, the term of order $\mathrm{e}^{-n/g}$, the n-instanton contribution, has at leading order the form

$$E_N^{(n)}(g) = \left(\frac{2}{g}\right)^{n(N+1/2)}\left(-\frac{\mathrm{e}^{-1/6g}}{\sqrt{2\pi}}\right)^n\left[P_N^{(n)}\left(\ln(-g/2)\right) + O\left(g\left(\ln g\right)^{n-1}\right)\right],$$

in which $P_N^{(n)}(\sigma)$ is a polynomial of degree $(n-1)$.

Verifications: large order behaviour of perturbation series. After an analytic continuation from g negative to g positive, the Borel sums become complex with an imaginary part exponentially smaller by about a factor $\mathrm{e}^{-1/3g}$ than the real part. Simultaneously, the function $\ln(-2/g)$ also becomes complex with an imaginary part $\pm i\pi$. Since the sum of all contributions is real, the imaginary parts must cancel. For example, the non-perturbative imaginary part of the Borel sum of the perturbation series $E_0^{(0)}(g)$ cancels the perturbative imaginary part of the two-instanton contribution:

$$\mathrm{Im}\, E_0^{(0)}(g) \underset{g\to 0_+}{\sim} \frac{1}{\pi g}\,\mathrm{e}^{-1/3g}\,\mathrm{Im}\left[P_0^{(2)}\bigl(\ln(-g/2)\bigr)\right] = -\frac{1}{g}\,\mathrm{e}^{-1/3g}\,.$$

The coefficients $E_{0k}^{(0)}$ of the perturbative expansion have the integral representation $(k>1)$:

$$E_{0k}^{(0)} = \frac{1}{\pi}\int_0^\infty \mathrm{Im}\left[E_0^{(0)}(g)\right]\frac{\mathrm{d}g}{g^{k+1}}.$$

From the asymptotic estimate of $\operatorname{Im} E_0^{(0)}$ for $g \to 0$, on then derives the large order behaviour of the perturbative expansion (checked numerically):

$$E_{0k}^{(0)} \underset{k\to\infty}{=} -\frac{1}{\pi} 3^{k+1} k!\,(1+O(1/k))\ .$$

Similarly, from $\operatorname{Im} P_0^{(3)}$ one derives the large order behaviour of the expansion of the one-instanton contribution (also checked numerically):

$$E_{0k}^{(1)} \underset{k\to\infty}{=} -\frac{3^{k+2}}{\pi} k! \left[\ln 6k + \gamma + O\left(\frac{\ln k}{k}\right)\right].$$

Other verification: the real part of the two-instanton contribution. Considering two lowest eigenvalues $E_\pm$ of H, we have calculated the ratio

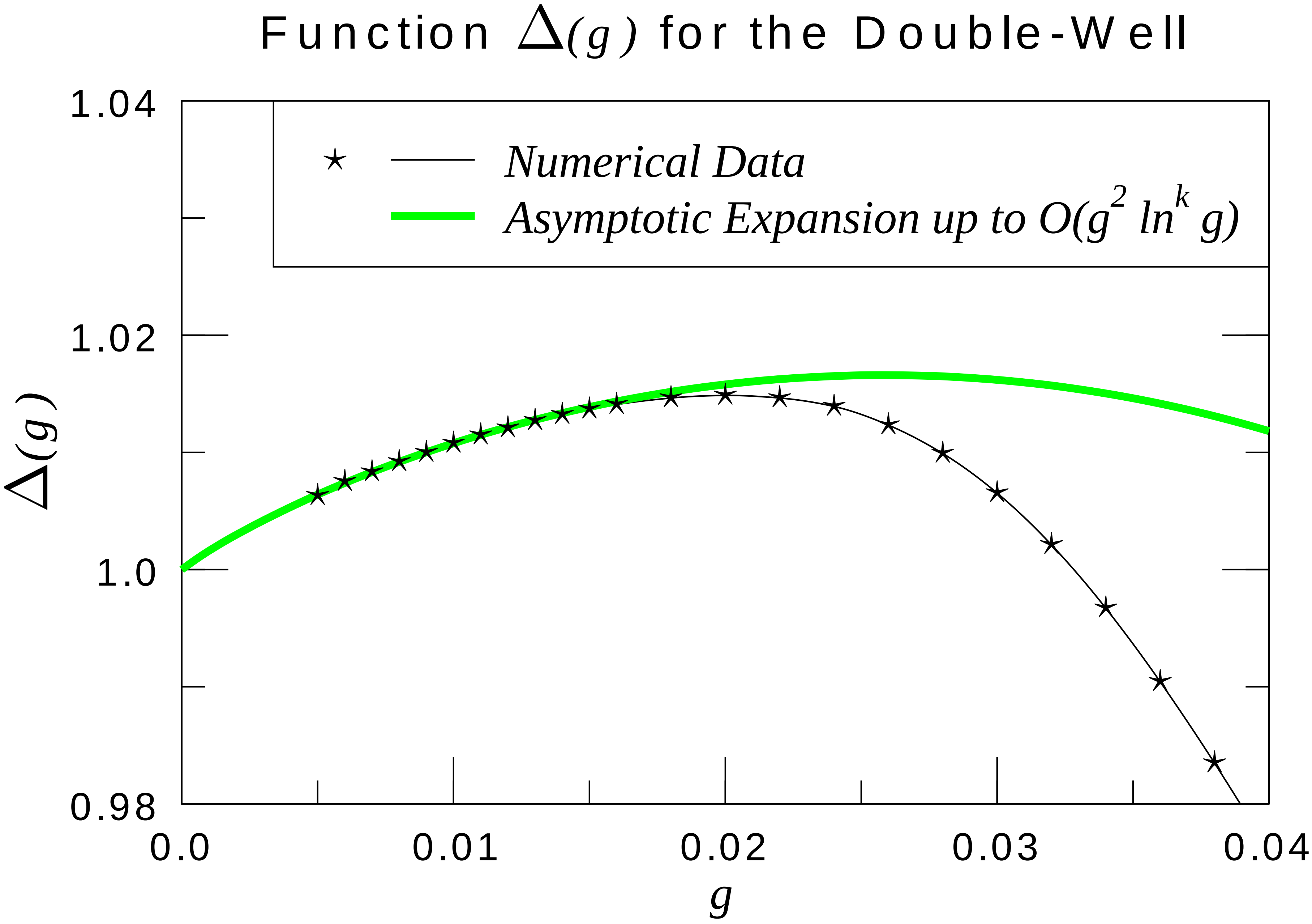
$$\Delta(g) = 4\frac{\left\{\frac{1}{2}\left(E_+ + E_-\right) - \operatorname{Re}\left[\text{Borel sum of } E_0^{(0)}(g)\right]\right\}}{\left(E_+ - E_-\right)^2 \left(\ln 2g^{-1} + \gamma\right)}.$$

In the sum $(E_+ + E_-)$, contributions corresponding to an odd number of instantons cancel. For $g \to 0$, the numerator thus is dominated by the real part of the two-instanton contribution proportional to $\mathrm{Re}\, P^{(2)}$. The difference $(E_+ - E_-)$ is dominated by the one-instanton contribution.

Using explicit expressions, and performing an expansion in powers of g and inverse powers of $\ln(2/g)$ and keeping only the first few terms in $\{1/\ln(2/g)\}$ in each term in the g-expansion, one finds

$$\Delta(g) = 1 + 3g - \tfrac{23}{2} \tfrac{g}{\ln(2/g)} \left[1 - \tfrac{\gamma}{\ln(2/g)} + \tfrac{\gamma^2}{\ln^2(2/g)} + O\left(\tfrac{1}{\ln^3(2/g)}\right)\right]$$
$$+ \tfrac{53}{2} g^2 - 135 \tfrac{g^2}{\ln(2/g)} \left[1 - \tfrac{\gamma}{\ln(2/g)} + \tfrac{\gamma^2}{\ln^2(2/g)} + O\left(\tfrac{1}{\ln^3(2/g)}\right)\right] + O\left(g^3\right) .$$

The higher-order corrections, which are only logarithmically suppressed with respect to the leading terms $1 + 3g$, change the numerical values quite significantly, even for small g.



Function Δ(g) for the Double-Well
Numerical Data
Asymptotic Expansion up to O(g^2 ln^k g)
1.04
1.02
1.0
0.98
0.0
0.01
0.02
0.03
0.04
g
Δ(g)

Another example: The periodic cosine potential

For illustration purpose, we mention another example that has been investigated thoroughly, the cosine potential $\frac{1}{16}(1-\cos 4q)$, which differs from the preceding one because the potential is still an entire function but no longer a polynomial.The periodicity of the potential allows classifying eigenfunctions according to their behaviour under a translation of one period:

$$\psi_\varphi(q+\pi/2) = \mathrm{e}^{i\varphi}\,\psi_\varphi(q).$$

The conjectured spectral equation then takes the form

$$\left(\frac{2}{g}\right)^{-B} \frac{\mathrm{e}^{A(E,g)/2}}{\Gamma(\frac{1}{2}-B)} + \left(\frac{-2}{g}\right)^{B} \frac{\mathrm{e}^{-A(g,E)/2}}{\Gamma(\frac{1}{2}+B)} = \frac{2\cos\varphi}{\sqrt{2\pi}}.$$

The first few terms of the perturbative expansions of the functions A and B are

$$B = E + g\left(E^2+\tfrac{1}{4}\right) + g^2\left(3E^3+\tfrac{5}{4}E\right) + O\left(g^3\right),$$

$$A = g^{-1} + g\left(3E^2+\tfrac{3}{4}\right) + g^2\left(11E^3+\tfrac{23}{4}E\right) + O\left(g^3\right).$$

Perturbative and WKB expansions from Schrödinger equations

In the simplest examples these conjectures, motivated by semi-classical evaluations of path integrals (instanton calculus), have obtained independent confirmation by considerations based on the Schrödinger equation.

We consider the Schrödinger equation written as

$$[H\psi](q) \equiv -\frac{g}{2}\psi''(q) + \frac{1}{g}V(q)\psi(q) = E\psi(q).$$

The restriction to potentials $V(q)$ that are entire functions allows extending the Schrödinger equation and its solutions to the q complex plane.

To generate semi-classical expansions, it is convenient to derive a Riccati equation from the Schrödinger equation by setting

$$S(q) = -g\psi'/\psi\,,\ \Rightarrow\ gS'(q) - S^2(q) + 2V(q) - 2gE = 0\,.$$

We further set $S = S_+ + S_-$, where in the sense of a series expansion

$$S_\pm(q,-g,-E) = \pm S_\pm(q,g,E).$$

Then, the Riccati equation decomposes into

$$gS'_- - S_+^2 - S_-^2 + 2V(q) - 2gE = 0\,, \quad gS'_+ - 2S_+S_- = 0\,.$$

The Fredholm determinant of $H - E$ is related to S_+ by

$$\ln\det(H-E) = \mathrm{tr}\,\ln(H-E) = \frac{1}{g}\int \mathrm{d}q\, S_+(q,E).$$

On infers that the spectral equation, or quantization condition, can then be written as

$$\frac{1}{2i\pi g}\lim_{\varepsilon\to 0_+}\int \mathrm{d}q\,[S_+(q,E_N - i\varepsilon) - S_+(q,E_N + i\varepsilon)] = N + \tfrac{1}{2}\,,\ N \geq 0\,.$$

In the case of analytic potentials, the domain of integration can be locally deformed in the q complex plane into a contour C that encloses the N zeros of the eigenfunction. The spectral equation then becomes

$$B(E_N, g) \equiv -\frac{1}{2i\pi g}\oint_C \mathrm{d}z\, S_+(z, E_N) = N + \tfrac{1}{2}\,.$$

WKB and perturbative expansions

This elegant formulation, restricted, however, to one dimension and analytic potentials, bypasses the difficulties generally associated with turning points and thus allows connecting perturbative and WKB expansions.

Potential with unique minimum. We first assume that the potential V has a unique minimum at $q = 0$ with $V(0) = 0$.

A systematic expansion for $g \to 0$, at Eg fixed, of the solution of the Riccati equation in the complex q-plane, inserted into the expression

$$B(E, g) = -\frac{1}{2i\pi g} \oint_C \mathrm{d}z\, S_+(z, E),$$

leads to the 'exact' WKB expansion.

At leading order, one finds

$$S(q) \sim S_+(q) = S_0(q), \quad S_0(q) = \sqrt{2V(q) - 2gE}\,.$$

For $E > 0$, the function S_0 has two branch points $q_1 < q_2$ on the real axis. We place the cut between q_1 and q_2 and choose the determination of S_0 to be positive for $q > q_2$.

In the semi-classical limit, the contour C in the spectral equation encloses the cut of $S_0(q)$. One recovers the Bohr–Sommerfeld quantization condition in the form

$$B(E, g) = -\frac{1}{2i\pi g} \oint_C \mathrm{d}z \sqrt{2V(z) - 2Eg} = N + \tfrac{1}{2}\,.$$

In this form, the WKB expansion can be further expanded in powers of g at E fixed, to recover the perturbative expansion.

Potentials with degenerate minima

We now consider potentials of double-well type, with two degenerate minima as displayed in the figure. For E small enough, the function $S_0(q)$ has now four branch points $q_1, \ldots, q_4$ on the real axis.

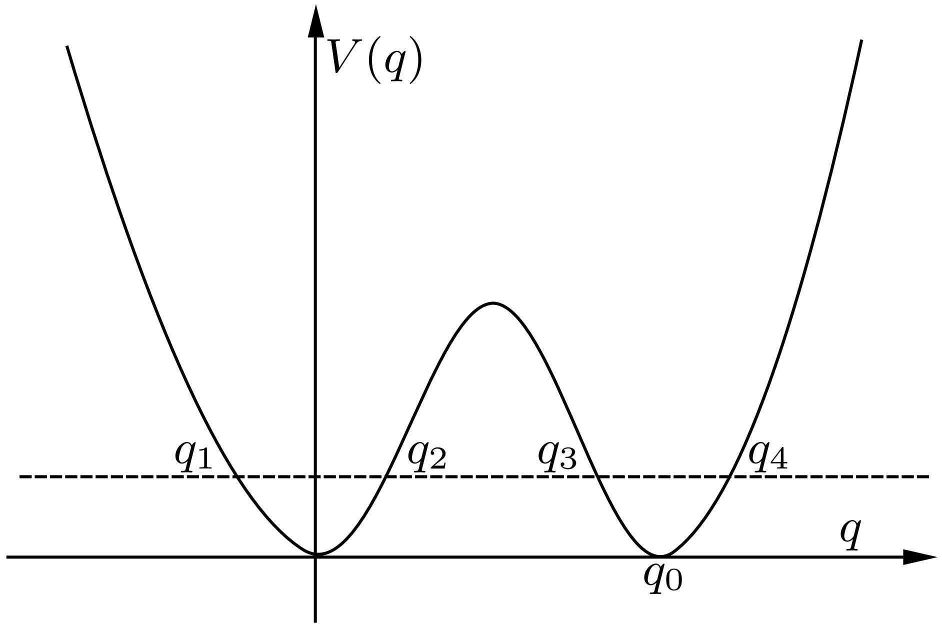


Fig. 3 Potentials with degenerate minima: The four turning points.

Intuitively, one expects the spectral equation to now involve two functions $B_1(E,g)$ and $B_2(E,g)$ obtained by integrating S_0 around the two cuts $[q_1,q_2]$ and $[q_3,q_4]$. However, starting from E large, a careful analysis shows that the contour C' that surrounds the cut along $[q_2,q_3]$, when the cuts of S_0 are placed differently, is also involved.

Comparing with conjectured form of the spectral equation, one then infers

$$\frac{1}{g}\oint_{C'} \mathrm{d}z\, S_+(z) = A(E,g) + \ln(2\pi) - \sum_{i=1}^{2} \ln\Gamma\big(\tfrac{1}{2} - B_i(E,g)\big) + B_i(E,g)\ln(-g/2C_i),$$

where the C_i's are normalization constants and $A(E,g)$ is the function that appears in the generalized Bohr–Sommerfeld formula.

The expansion for Eg small of its WKB expansion yields the perturbative expansion of the function $A(g,E)$. To identify with the WKB expansion, the function $\Gamma(\frac{1}{2} - B)$ has to be replaced by its asymptotic expansion for B large:

$$-\ln\Gamma(\tfrac{1}{2} - B) \underset{E\to\infty}{\sim} B\ln(-B) - B \underset{g\to 0}{\sim} B\ln(-E) + \cdots \ .$$

More recently, further insight into the algebraic properties of these expansions and the cancellation of imaginary contributions has been reported in

G.V. Dunne and M. Ünsal, *Generating energy eigenvalue trans-series from perturbation series*, [hep-th] arxiv:1306.4405.

In particular, a direct relation between the functions A and B has been derived:

$$\frac{\partial E_{\text{pert.}}}{\partial B} = -6Bg - 3g^2 \left.\frac{\partial A}{\partial g}\right|_B .$$

An identical relation, up to the coupling normalization, holds for the cosine potential:

$$\frac{\partial E_{\text{pert.}}}{\partial B} = -2Bg - g^2 \left.\frac{\partial A}{\partial g}\right|_B ,$$

or for the Fokker–Planck quartic potential. It remains to generalize these relations for asymmetric wells, multiple wells and to discover an instanton interpretation.

Instantons and multi-instantons: the quartic double-well

The initial motivation for our conjectures came from a summation of leading order multi-instanton contributions to path integrals.

Partition function and spectrum

The path integral formalism allows calculating directly the quantum partition function, which for Hamiltonians with a discrete spectrum has the expansion

$$\mathcal{Z}(\beta) \equiv \mathrm{tr}\ \mathrm{e}^{-\beta H} = \sum_{N\geq 0} \mathrm{e}^{-\beta E_N} .$$

The spectrum can be inferred from the Fredholm determinant $\mathcal{D}(E) = \det(H - E)$ and thus from the trace $G(E)$ of the resolvent of H, which is related to the partition function by

$$G(E) = \mathrm{tr}\, \frac{1}{H - E} = \int_0^\infty \mathrm{d}\beta\ \mathrm{e}^{\beta E}\, \mathcal{Z}(\beta)\, ,\ \ \frac{\mathrm{d}}{\mathrm{d}E} \ln \mathcal{D}(E) = G(E).$$

For the double-well potential $V(q) = \frac{1}{2}q^2(1-q)^2$, one can use the reflection operator $P\psi(q) = \psi(1-q)$ to separate eigenvalues according to the symmetry properties of eigenfunctions in the exchange $q \leftrightarrow (1-q)$. This leads to consider the two functions

$$\mathcal{Z}_{\pm}(\beta) = \operatorname{tr}\left[\tfrac{1}{2}(1 \pm P)\,\mathrm{e}^{-\beta H}\right] = \sum_{N=0} \mathrm{e}^{-\beta E_{\pm,N}} = \tfrac{1}{2}\big(\mathcal{Z}(\beta) \pm \mathcal{Z}_{\mathrm{a}}(\beta)\big).$$

In the path integral formulation,

$$\mathcal{Z}(\beta) \equiv \operatorname{tr}\,\mathrm{e}^{-\beta H} = \int_{q(-\beta/2)=q(\beta/2)} [\mathrm{d}q(t)]\,\mathrm{e}^{-\mathcal{S}(q)/g},$$

$$\mathcal{Z}_{\mathrm{a}}(\beta) \equiv \operatorname{tr}\left(P\,\mathrm{e}^{-\beta H}\right) = \int_{q(-\beta/2)+q(\beta/2)=1} [\mathrm{d}q(t)]\,\mathrm{e}^{-\mathcal{S}(q)/g},$$

where $\mathcal{S}(q)$ is the euclidean action,

$$\mathcal{S}(q) = \int_{-\beta/2}^{\beta/2} \left[\tfrac{1}{2}\dot{q}^2(t) + V\big(q(t)\big)\right] \mathrm{d}t\,.$$

Perturbation theory. For $g \to 0$, $\mathcal{Z}(\beta)$ can be evaluated by the steepest descent method applied to the path integral. The two saddle points $q(t) \equiv 0$ and $q(t) \equiv 1$ give the same contribution. Thus, the eigenvalues are twice degenerate to all orders in a perturbative expansion in powers of g:

$$E_{\pm,N}(g) \approx E_N^{(0)}(g) \equiv \sum_{k=0}^{\infty} E_{N,k}^{(0)} g^k .$$

Instantons: the two lowest eigenvalues ($\beta \to \infty$). In the case of the path integral representation of $\mathcal{Z}_{\mathrm{a}}(\beta)$, constant solutions of the euclidean equation of motion do not satisfy the boundary conditions.

The lowest eigenvalues are obtained in the infinite β limit. Solutions with an action that then has a finite limit, necessarily correspond to paths which connect two minima of the potential. Non-constant solutions with finite action are called instanton solutions.

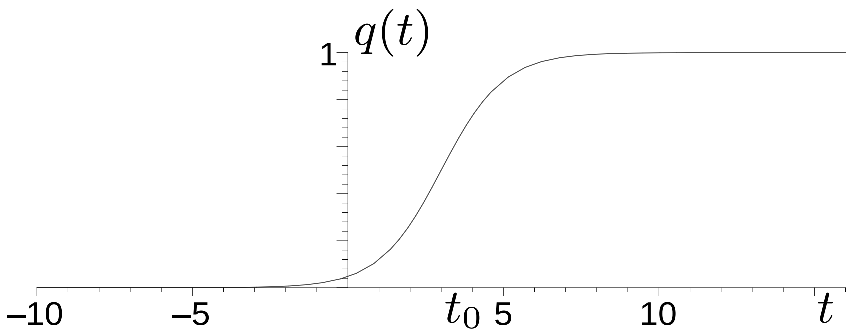


Fig. 4 The one-instanton configuration.

For the quartic double-well potential, such solutions for $\beta \to \infty$ are

$$q_{\mathrm{c}}(t) = \left(1 + \mathrm{e}^{\pm(t-t_0)}\right)^{-1} \ \Rightarrow \ \mathcal{S}(q_{\mathrm{c}}) = 1/6\,,$$

where for β finite, $t_0 \in [-\beta/2, \beta/2]$. Thus, $\mathcal{Z}_{\mathrm{a}}(\beta) = O(\mathrm{e}^{-1/6g})$. Taking the time t_0 as a collective coordinate, and integrating over the remaining fluctuations in the Gaussian approximation, one infers for the two lowest eigenvalues ($\epsilon = \pm 1$)

$$E_{\epsilon,0}(g) = \lim_{\beta\to\infty} -\frac{1}{\beta} \ln \mathcal{Z}_\epsilon(\beta) \underset{g\to 0}{\approx} E_0^{(0)}(g) - \epsilon E_0^{(1)}(g),\ E_0^{(1)}(g) \sim \frac{1}{\sqrt{\pi g}}\,\mathrm{e}^{-1/6g}\,.$$

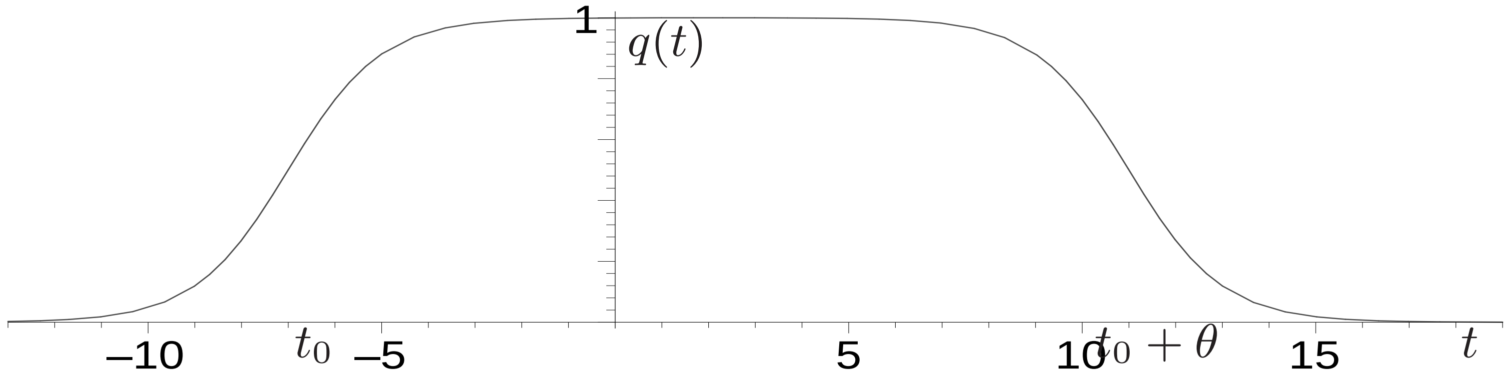


Fig. 5 A two-instanton configuration.

The two-instanton configurations

Two-instanton configurations depend on one additional time parameter θ, the separation between instantons, in the limit $\theta \to \infty$ decompose into two instantons (see figure) and for θ large must minimize the variation of the action. A configuration that satisfies these conditions can be obtained in the following way: we parametrize the one-instanton solutions as

$$q_{\pm}(t) = f\bigl(\mp(t - t_0)\bigr), \quad f(t) = 1/\left(1 + \mathrm{e}^{t}\right) = 1 - f(-t).$$

Then, the two-instanton configuration, up to a global translation, can be chosen as

$$q_{\mathrm{c}}(t) = f(t-\theta/2) + f(-t-\theta/2) - 1 = f(t-\theta/2) - f(t+\theta/2).$$

The corresponding classical action is

$$\mathcal{S}(q_c) = \tfrac{1}{3} - 2\,\mathrm{e}^{-\theta} + O\left(\mathrm{e}^{-2\theta}\right).$$

In analogy with the partition function of a classical gas (instantons being identified with particles), one calls the quantity $-2\,\mathrm{e}^{-\theta}$ interaction potential between instantons.

It is simple to extend the result to β large but finite. Symmetry between θ and $\beta-\theta$ indeed implies

$$\mathcal{S}(q_c) = \tfrac{1}{3} - 2\,\mathrm{e}^{-\theta} - 2\,\mathrm{e}^{-(\beta-\theta)} + \text{exponentially smaller contributions.}$$

The n-instanton configuration

More generally, the n-instanton configuration, which is constructed from n instantons separated by times θ_i with

$$\sum_{i=1}^{n} \theta_i = \beta\,,$$

yields the classical action

$$\mathcal{S}_{\mathrm{c}}(\theta_i) = \frac{n}{6} - 2\sum_{i=1}^{n} \mathrm{e}^{-\theta_i} + O\left(\mathrm{e}^{-(\theta_i+\theta_j)}\right).$$

For n even, the n-instanton configurations contribute to $\mathrm{tr}\,\mathrm{e}^{-\beta H}$, while for n odd they contribute to $\mathrm{tr}\,\left(P\,\mathrm{e}^{-\beta H}\right)$.

The integration over the Gaussian fluctuations around the multi-instantons at fixed collective coordinates θ_i can then be performed explicitly.

The n-instanton contribution: discussion

To go beyond the one-instanton approximation, one must take into account the interaction between instantons. However, one notices that the interaction between instantons, $-2\,e^{-\theta}/g$, for $g > 0$ is attractive. Therefore, for $g \to 0$, the path integral is dominated by configurations in which the instantons are close. However, for θ_i finite the concept of instanton is no longer meaningful, since the configurations cannot be distinguished from fluctuations around the constant or the one-instanton solution.

Such a difficulty could have been anticipated. Indeed, the large order behaviour analysis shows that the perturbative expansion in the case of potentials with degenerate minima is not Borel summable and ambiguous at the two-instanton order. But then the notion of contributions of the two-instanton order $e^{-(2/6)g}$, or even smaller, is not meaningful.

To proceed any further, one must provide a definition for the sum of the perturbative expansion.

In the example of the double-well potential, one can prove that the perturbation series is Borel summable for g negative. Simultaneously, for g negative, the interaction between instantons becomes repulsive and the instanton configurations become meaningful. Therefore, we define the sum of the complete perturbative expansion, including multi-instanton contributions, as the analytic continuation from $g < 0$ negative to $|g| \pm i0$ of all quantities consistently.

With this definition, summing explicitly the leading order multi-instanton contributions, one obtains the spectral conditions

$$\Delta_\epsilon(E) = \frac{1}{\Gamma(\frac{1}{2} - E)} + \epsilon i \left(-\frac{2}{g}\right)^E \frac{\mathrm{e}^{-1/6g}}{\sqrt{2\pi}} = 0,$$

which, initially, have motivated our conjectures.