

Asymptotic Methods in Science and Engineering

S. G. Rajeev

Contents

Chapter 1. Convergent vs Asymptotic Expansions	6
Chapter 2. The Factorial	9
2.1. An Integral Representation	9
2.2. Stirling's Approximation	10
2.3. The Second Order	13
2.4. Some Integrals	15
Chapter 3. Laplace's Method	17
3.1. Expanding Around The Maximum	18
3.2. Higher Orders	19
3.3. Symbolic Computation	20
3.4. Graphical Interpretation	20
3.5. A Slight Generalization	22
Chapter 4. The Gamma Function	23
4.1. The Product Formula	23
4.2. Power Series in z	25
4.3. Series in Inverse Powers	26
Chapter 5. Generating Functions And Their Approximations	29
5.1. Euler Summation	32
5.2. Borel Summation	32
5.3. Pade' Approximants	33
Chapter 6. Continued Fractions	36
6.1. Continued fractions provide the best rational approximations to real numbers	36
6.2. General Continued Fractions	37
6.3. Second Order Linear Difference Equations	38
6.4. Rational Approximation of Analytic Functions Of A Complex Variable	39
6.5. An Example	40
6.6. Pade' Approximants as Continued Fractions	43
Chapter 7. Eigenvalues of a Tridiagonal Matrix	45
7.1. 2×2 Matrix	45
7.2. The Spectrum of a Tridiagonal Matrix	46
7.3. The Lanczos Algorithm Makes Any Matrix Tridiagonal	48
7.4. Appendix: The Discriminant of a Matrix	49
Chapter 8. Sturm-Liouville Problems	51

8.1. Variational principle	51
8.2. Picard Iteration	52
8.3. The Characteristic Function	52
8.4. The Ricatti Equation	54
8.5. The Resolvent (to be Completed)	55
Chapter 9. The Method of Steepest Descent	56
9.1. The Airy Function	58
Chapter 10. The WKB Method	63
10.1. The Eikonal	63
10.2. The Method of Characteristics	64
10.3. WKB for Sturm-Liouville Problems	65
Chapter 11. Symbol Calculus	67
11.1. Operators and their Kernels	67
11.2. From Symbols to Operators	69
11.3. The Star Product	70
11.4. Heisenberg Equation	72
11.5. The Star Exponential	72
11.6. The Trace	73
11.7. The Resolvent Symbol	74
11.8. Can you hear the shape of a drum?	75
Chapter 12. The Gutzwiller Trace Formula	76
12.1. The Poisson Summation Formula	77
12.2. The Propagator	79
Chapter 13. Singular Perturbation Theory for Differential Equations	82
13.1. Center Manifolds	82
13.2. Cyclotron Radiation	84
13.3. Center-Manifold of the Lorentz-Dirac Equation	87
Chapter 14. The Feynman-Kac Formula	90
14.1. The Heat Kernel of The Schrodinger Operator	90
14.2. Subdividing Time	91
14.3. Laplace's Method for Path Integrals	93
Chapter 15. The Gaussian Integral	94
15.1. The Normal Distribution	94
15.2. The Characteristic Function	95
15.3. The Expectation Values of a Multi-variate Gaussian	96
15.4. The Multi-Dimensional Gaussian Measure	97
Chapter 16. Laplace's Method For Multi-Dimensional Integrals	99
16.1. Example: $K = 4$	100
16.2. $K = 6$	100
Chapter 17. The Wiener Measure	104
17.1. Lebesgue Decomposition	104
17.2. The Space of Continuous Functions	105

CONTENTS

5

17.3. Probability of a Path	105
17.4. The Harmonic Oscillator	106

CHAPTER 1

Convergent vs Asymptotic Expansions

The prototype of a convergent infinite series is the geometric series

$$1 + z + z^2 + z^3 \dots$$

If $|z| < 1$ each successive term is smaller than the previous one and the sum tends to a finite value. To prove this, consider the partial sums

$$S_n(z) = 1 + z + \dots + z^n$$

Since

$$zS_n(z) = z + z^2 + \dots + z^{n+1} = S_n(z) - 1 + z^{n+1}$$

it follows that

$$S_n(z) = \frac{1 - z^{n+1}}{1 - z}.$$

The difference

$$|S_n(z) - \frac{1}{1-z}| = \frac{|z|^{n+1}}{|1-z|} < \frac{|z|^{n+1}}{(1-|z|)}.$$

As $n \rightarrow \infty$ this tends to zero. Thus the geometric series converges inside the unit circle

$$\lim_{n \rightarrow \infty} S_n(z) = \frac{1}{1-z}, \quad |z| < 1.$$

More generally consider the power series

$$\sum_n a_n z^n$$

If

$$|a_n| < \lambda^n$$

whenever $n > N$ for some λ , we can compare this series to the geometric series term by term

$$|a_n z^n| < |(\lambda z)^n|$$

So it will converge if

$$|z| < \lambda^{-1}.$$

In fact $f(z) = \sum_n a_n z^n$ is an analytic function of z within this circle of convergence. The most elementary situation is when $|a_n|$ tends to zero faster than an exponential;

i.e., $|a_n| < |\lambda|^n$ for **every** λ . In this case the series converges for any z and $f(z)$ is an entire function. Familiar examples are

$$\begin{aligned}\exp(z) &= \sum_{n=0}^{\infty} \frac{1}{n!} z^n \\ \sin(z) &= \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{(2n+1)!} z^{2n+1}\end{aligned}$$

Infinite power series with finite radius of convergence include other familiar examples

$$\begin{aligned}\log[1-z] &= \sum_{n=1}^{\infty} \frac{1}{n} z^n, \quad |z| < 1 \\ \operatorname{arctanh} z &= \sum_{n=0}^{\infty} \frac{z^{2n+1}}{2n+1}, \quad |z| < 1.\end{aligned}$$

A series like

$$\sum_{n=0}^{\infty} (-1)^n n! z^n$$

does not converge for **any** value of z (except the trivial case $z = 0$). Many of the standard approximation methods of physics give series of this type or worse. We will see why they work in spite of the lack of convergence. It will turn out that the usual definition of a convergent infinite series is too restrictive: there is a larger class of **asymptotic series** that include these cases of physical interest. Unlike for convergent series, some additional information (usually from the physics or geometry represented by the function) is necessary to make sense of the sum. The lack of convergence can be traced to certain essential singularities (“instantons”) at $z = 0$. Yet these singularities have a physical origin: e.g., tunneling in the application to quantum mechanics. By the technique of Borel summation, the divergent series above can be represented as an integral

$$\sum_{n=0}^{\infty} (-1)^n n! z^n \equiv f(z) = \int_0^{\infty} \frac{1}{1+zt} e^{-t} dt$$

which converges for $z > 0$ and has a branch cut along the negative real axis. This way we can find a meaning for the function represented by the power series. For example

$$f(2) \approx 0.461455$$

A plot of this interpretation of the function is given in the figure.

Even when a convergent series is available, an asymptotic series can provide a better alternative approximation. We will see an example of this in the Stirling approximation to the Gamma function.

A much more recent development (since the 1940s) is the theory of renormalization. Each term in the series is itself a divergent integral, and yet we are able to make sense of the expansion. The most accurate predictions in the history of science (e.g., the magnetic moment of the electron to fifteen decimal place accuracy)

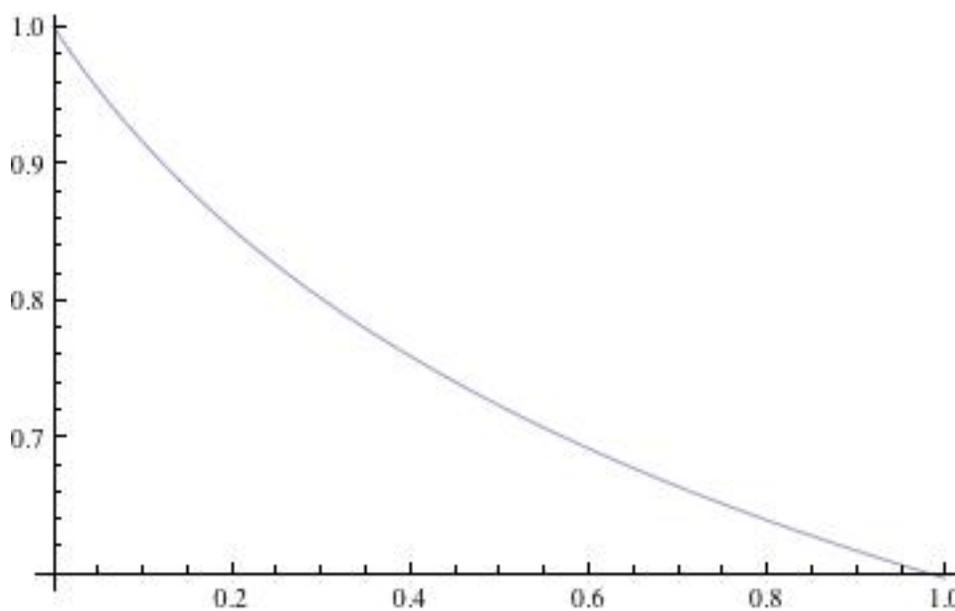


FIGURE 1.0.1.

are made this way. The mathematical basis of this is as yet unknown. 't Hooft has related this phenomenon also to certain essential singularities (“renormalons”). We will take a look at this mysterious theory towards the end of the course.

CHAPTER 2

The Factorial

Euler is the “Master of Us All” mathematical physicists. He was adept at manipulations that are, in the modern mathematical sense, illegal, yet gave physically correct answers. One of his favorite functions was the Gamma function, an extension to complex values of the factorial. It turns out that this function has an integral representation which is a prototype for the integrals that appear all over physics: quantum mechanics, wave optics, quantum field theory, statistical mechanics, all can be formulated in terms of integrals over many (sometimes infinite number of) variables. Learning how to calculate the factorial accurately is a good way to learn how to understand these theories.

2.1. An Integral Representation

Recall that

$$n! = n(n-1) \cdots 3 \times 2 \times 1.$$

for any positive integer. It is not hard to see that its value grows rapidly with n :

n	$n!$
1	1
2	2
3	6
4	24
5	120
6	720
7	5040
8	40320
9	362880
10	3628800

Many problems in probability require us to estimate the value of $n!$ for large n . How fast does it grow? It is useful to note that

PROPOSITION 1. *The factorial has an integral representation*

$$n! = \int_0^\infty t^n e^{-t} dt$$

PROOF. To prove this identity, start with the definition

$$I_n = \int_0^\infty t^n e^{-t} dt.$$

Note first of all that

$$I_0 = \int_0^{\infty} e^{-t} dt = 1$$

Moreover

$$\int_0^{\infty} t^n e^{-t} dt = - \int_0^{\infty} \frac{d}{dt} [t^n e^{-t}] dt + n \int_0^{\infty} t^{n-1} e^{-t} dt$$

Thus

$$I_n = nI_{n-1}.$$

Iterating this

$$I_n = n(n-1) \cdots 2 \times 1 \times I_0 = n!$$

□

Why bring in an integral to study a quantity that is defined using just the idea of multiplication? By looking at the integrand we can see where most of the contribution to the integral comes from; and this gives us an approximation method for large values of n which would otherwise take a large number of multiplications. A side benefit is that it allows us to extend the definition of the factorial to fractional values of n . We already have reason to believe that

$$0! = I_0 = 1.$$

We will see soon that

$$\left(\frac{1}{2}\right)! = \frac{\sqrt{\pi}}{2}$$

2.2. Stirling's Approximation

Let us plot the integrand for some moderately large value of n :

We see that the integrand is peaked around $t = n$; in fact the width of the peak decreases as n grows. So it should be possible to get a good approximation by expanding around this value.

If we write the integral as

$$\begin{aligned} \int_0^{\infty} t^n e^{-t} dt &= \int_0^{\infty} \exp[-t + n \log t] dt \\ &= \int_0^{\infty} \exp\left[n \left\{ \log t - \frac{t}{n} \right\}\right] dt \end{aligned}$$

Since we expect the region near $t = n$ to be important, let us make a change of variable that centers the graph around this value:

$$s = \frac{t}{n}$$

$$I_n = n \exp[n \log n] \int_0^{\infty} \exp[n \{\log s - s\}] ds$$

Or

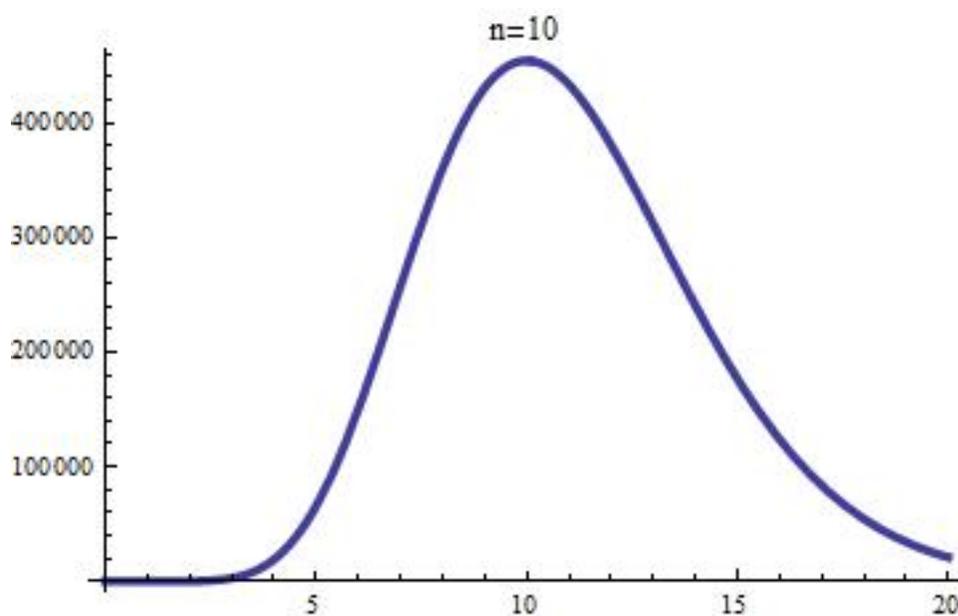


FIGURE 2.2.1.

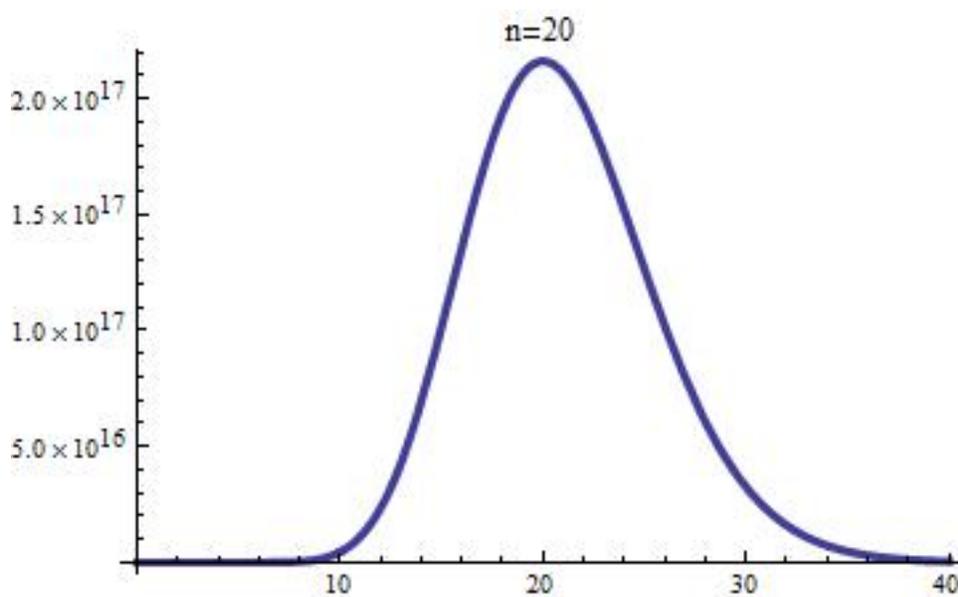


FIGURE 2.2.2.

$$I_n = n^{n+1} \int_0^{\infty} \exp [n \{\log s - s\}] ds$$

The maximum of $\log s - s$ occurs at $s = 1$, as we expected. If we expand around this point,

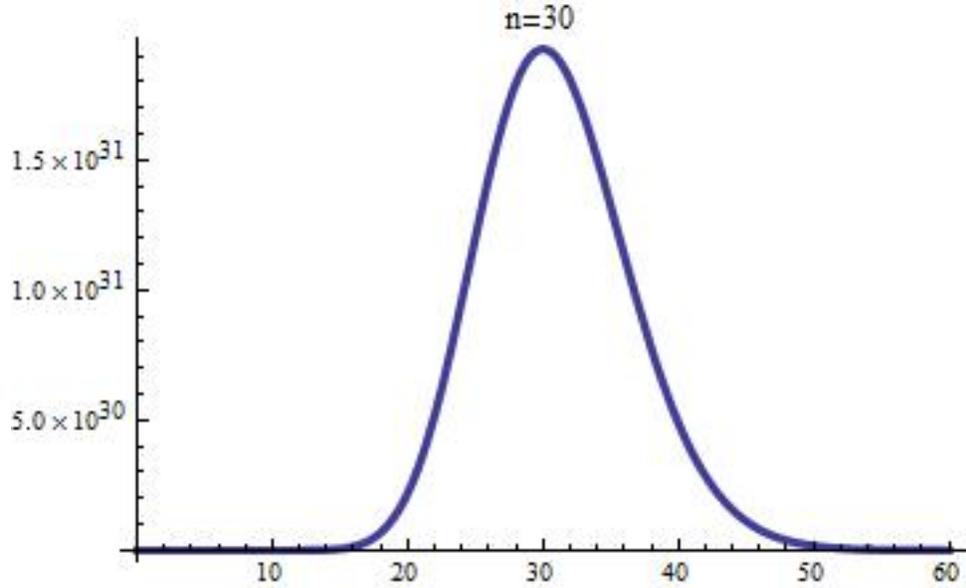


FIGURE 2.2.3.

$$\log s - s = -1 - \frac{1}{2}(s-1)^2 + \frac{1}{3}(s-1)^3 - \frac{1}{4}(s-1)^4 + O((s-1)^5)$$

$$I_n = n^{n+1} e^{-n} \int_0^\infty \exp \left[n \left\{ -\frac{1}{2}(s-1)^2 + \frac{1}{3}(s-1)^3 - \frac{1}{4}(s-1)^4 \dots \right\} \right] ds$$

This suggests yet another change of variable

$$n(s-1)^2 = u^2$$

$$s = 1 + \frac{u}{\sqrt{n}}$$

$$I_n = n^{n+\frac{1}{2}} e^{-n} \int_{-\sqrt{n}}^\infty \exp \left[-\frac{1}{2}u^2 + \frac{1}{3\sqrt{n}}u^3 - \frac{1}{4n}u^4 \dots \right] du$$

The point is that we now have n where we want it: the successive terms get smaller as s grows. If we boldly keep just the leading term we get an approximation

$$I_n \approx n^{n+\frac{1}{2}} e^{-n} \int_{-\infty}^\infty \exp \left[-\frac{1}{2}u^2 \right] du$$

The latter integral can be evaluated to be

$$\int_{-\infty}^\infty \exp \left[-\frac{1}{2}u^2 \right] du = \sqrt{2\pi}.$$

Thus

$$I_n \approx n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi}$$

How good is this approximation?

n	$n!$	Approx	%Error
1	1	0.922137	7.7863
2	2	1.919	4.04978
3	6	5.83621	2.72984
4	24	23.5062	2.0576
5	120	118.019	1.65069
6	720	710.078	1.37803
7	5040	4980.4	1.18262
8	40320	39902.4	1.03573
9	362880	359537.	0.921276
10	3628800	3.5987×10^6	0.829596

It is pretty good: the error is only about 1% for $n = 10$. Our method gets more accurate just when the brute force multiplication gets harder : when n gets larger.

n	Approx	%Error
2	1.919	4.04978
12	4.75687×10^8	0.691879
22	1.11975×10^{21}	0.378045
32	2.62447×10^{35}	0.260069
42	1.40222×10^{51}	0.198212
52	8.0529×10^{67}	0.160126
62	3.14277×10^{85}	0.134317
72	6.11636×10^{103}	0.115673
82	4.74881×10^{122}	0.101574
92	1.24272×10^{142}	0.0905383

EXERCISE 2. Show that the binomial distribution

$$\binom{n}{r} p^r (1-p)^{n-r}$$

tends to the Gaussian the limit $n \rightarrow \infty, p \rightarrow 0$ keeping $\frac{r}{n}$ and np fixed.

2.3. The Second Order

Can we improve the Stirling approximation?

$$I_n = n^{n+\frac{1}{2}} e^{-n} \int_{-\sqrt{n}}^{\infty} \exp \left[-\frac{1}{2} u^2 + \frac{1}{3\sqrt{n}} u^3 - \frac{1}{4n} u^4 \dots \right] du$$

Our approximation was to drop the higher order terms in the series, and also to replace the lower limit of the integration by infinity. It easy that the errors due to the latter are very small indeed:

$$\int_{-\sqrt{n}}^{\infty} \exp \left[-\frac{1}{2} u^2 \right] du = \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} u^2 \right] du - \int_{-\infty}^{-\sqrt{n}} \exp \left[-\frac{1}{2} u^2 \right] du$$

The relative error is

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\sqrt{n}} \exp \left[-\frac{1}{2} u^2 \right] du < \frac{1}{\sqrt{2\pi}} \int_1^{\infty} \exp \left[-\frac{n}{2} v \right] \frac{\sqrt{ndv}}{\sqrt{2v}} < \frac{\sqrt{n}}{\sqrt{2\pi}} e^{-\frac{n}{2}}$$

which decreases exponentially. The error due to ignoring higher order terms in the series decreases like a power of n instead:

$$\begin{aligned} I_n &= n^{n+\frac{1}{2}} e^{-n} \left(\int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} u^2 + \frac{1}{3\sqrt{n}} u^3 - \frac{1}{4n} u^4 \dots \right] du + O(e^{-n}) \right) \\ &= n^{n+\frac{1}{2}} e^{-n} \left(\int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} u^2 \right] \left\{ 1 + \frac{1}{3\sqrt{n}} u^3 + \frac{1}{2!} \left[\frac{1}{3\sqrt{n}} u^3 \right]^2 - \frac{1}{4n} u^4 \right\} du + O(e^{-n}) \right) \\ I_n &= n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi} \left(1 + \frac{1}{n} \left[\frac{1}{2!} \frac{1}{3^2} \langle u^6 \rangle - \frac{1}{4} \langle u^4 \rangle \right] + O(n^{-2}) + O(e^{-n}) \right) \end{aligned}$$

We encounter averages such as

$$\langle u^n \rangle \equiv \frac{\int_{-\infty}^{\infty} u^n \exp \left[-\frac{1}{2} u^2 \right] du}{\int_{-\infty}^{\infty} \exp \left[-\frac{1}{2} u^2 \right] du}$$

The odd powers average to zero by symmetry. The even powers can be evaluated to

$$\langle u^{2n} \rangle = (2n - 1)!!$$

For example,

$$\langle u^4 \rangle = 3, \quad \langle u^6 \rangle = 15$$

$$I_n = n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi} \left(1 + \frac{1}{n} \left[\frac{1}{2!} \frac{1}{3^2} 15 - \frac{1}{4} 3 \right] + O(n^{-2}) + O(e^{-n}) \right)$$

Simplifying

$$I_n = \sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}} \left(1 + \frac{1}{12n} + O(n^{-2}) + O(e^{-n}) \right)$$

We can see that the errors are substantially smaller:

n	$n!$	Approx1	%Error1	Approx2	%Error2
1	1	0.922137	7.7863	0.998982	0.101824
2	2	1.919	4.04978	1.99896	0.0518567
3	6	5.83621	2.72984	5.99833	0.0278913
4	24	23.5062	2.0576	23.9959	0.017137
5	120	118.019	1.65069	119.986	0.0115383
6	720	710.078	1.37803	719.94	0.00828033
7	5040	4980.4	1.18262	5039.69	0.00622504
8	40320	39902.4	1.03573	40318.	0.00484771
9	362880	359537.	0.921276	362866.	0.00388063
10	3628800	3.5987×10^6	0.829596	3.62868×10^6	0.00317601

Even in the worst case, $n = 1$, the error is about one part in a thousand. It looks as though we can do better and better by going to higher orders. What is to prevent us from calculating the answer to as much accuracy as we need?

It turns out that this Stirling expansion is divergent: we are expanding around an essentially singularity (the point at infinity). Although the answer gets better for the first few orders, it will eventually get worse. The larger n is, the smaller the higher orders are and moreover, the later the onset of the divergence. This is typical of all the approximation methods of physics.

Another important lesson is that a method that was supposed to work only for large n is actually giving excellent answers even when n is of order one. The validity of many of the expansions in physics has little to do with the actual value of the expansion parameter. It is a myth that the smallness of the coupling constant is what justifies perturbation theory in quantum mechanics, for example.

2.4. Some Integrals

We need often the constant

$$C = \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx$$

and averages such as

$$\langle x^n \rangle = \frac{\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} x^n dx}{\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx}.$$

Let us start with the evaluation of C . The trick is to look instead at C^2

$$C^2 = \left[\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx \right] \left[\int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2} dy \right]$$

Then we combine the two integrals into an integration on the plane

$$C^2 = \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x^2+y^2)} dx dy$$

and note that the integrand only depends on $r = \sqrt{x^2 + y^2}$. So we can pass to polar co-ordinates

$$C^2 = \int_0^{\infty} e^{-\frac{1}{2}r^2} 2\pi r dr$$

Setting $s = \frac{1}{2}r^2$

$$C^2 = 2\pi \int_0^{\infty} e^{-s} ds = 2\pi.$$

Thus

$$C = \sqrt{2\pi}$$

as promised.

To get the moments $\langle x^n \rangle$ it is useful to introduce the generating function

$$Z[J] = \langle e^{Jx} \rangle = \sum_{n=0}^{\infty} \frac{J^n}{n!} \langle x^n \rangle$$

Now,

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} e^{Jx} dx = e^{\frac{1}{2}J^2} \int e^{-\frac{1}{2}[x-J]^2} dx$$

By shifting the variable $x \rightarrow x + J$ we get

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} e^{Jx} dx = e^{\frac{1}{2}J^2} \int e^{-\frac{1}{2}x^2} dx = C e^{\frac{1}{2}J^2}$$

Thus

$$Z[J] = e^{\frac{1}{2}J^2}.$$

Expanding

$$Z[J] = \sum_{n=0}^{\infty} \frac{1}{2^n n!} J^{2n}$$

Thus

$$\langle x^n \rangle = 0, \quad \text{for } n \text{ odd}$$

which anyway is obvious by the symmetry $x \rightarrow -x$.

Comparing the even terms,

$$\sum_{n=0}^{\infty} \frac{J^{2n}}{(2n)!} \langle x^{2n} \rangle = \sum_{n=0}^{\infty} \frac{1}{2^n n!} J^{2n}$$

so that

$$\langle x^{2n} \rangle = \frac{(2n)!}{2^n n!} = \frac{(2n)(2n-1)(2n-2)(2n-3)\dots 3 \times 2 \times 1}{(2n)(2n-2)(2n-4)\dots 4 \times 2}$$

The denominator is the product of all the even numbers from 2 to $2n$. Thus

$$\langle x^{2n} \rangle = (2n-1)!!$$

which is the product of all the odd numbers from 1 to $2n-1$.

This number has an interesting combinatorial meaning. Put down a set of $2n$ points. Pair them up by drawing a line connecting any pair of points. $\langle x^{2n} \rangle$ is the number of such pairings. For, there are $2n-1$ ways of choosing the first pair, $2n-3$ ways to choose the next and so on.

There is no way to pair up all of them if the total number of points is odd. So the number of pairings is zero in that case, so that $\langle x^n \rangle$ is always equal to the number of pairings. Such graphical interpretations of Gaussian integrals have turned out to be very useful. Feynman turned this into a fine art.

CHAPTER 3

Laplace's Method

Many problems of physics can be reduced to the evaluation of integrals such as

$$Y(g) = \int e^{\frac{1}{g}S(\phi)} d\phi$$

In the simplest cases (such as the factorial) the integral is over a single real variable. More generally it may be a multiple integral. Even the deepest physical theories (quantum mechanics, quantum field theory, statistical mechanics) involve integrals such these, but over an infinite number of variables. The parameter g has different physical meanings in each subfield of physics: in combinatorics (such as the factorial) it is $g = \frac{1}{n}$. In quantum mechanics it is the Plank's constant, which measures the size of quantum effects. In wave optics, it is the wavelength of light, which is small when diffraction (interference) effects can be ignored. In statistical mechanics $g = kT$ is proportional to temperature: it is small when thermal fluctuations are small. Thus it is of great physical interest to understand an expansion of $f(g)$ around the point $g = 0$. This is despite the fact that this point is an essential singularity! We will generalize the ideas of the last chapter to develop a graphical method for calculating this expansion to as high an order as we need, in the case of a single variable of integration. The graphical method is a reinterpretation due to Feynman of this venerable method. As in many other cases, such a graphical user interface (gui) radically expanded the number of users. It made the method available to large numbers of people .

To be specific, we will consider the integral

$$Y(g) = \int_a^b e^{\frac{1}{g}S(\phi)} d\phi$$

where the function $S : R \rightarrow R$ has a unique maximum in the interval $[a, b]$ and g is a real positive variable.

$$Y(g) = \int_a^b e^{\frac{1}{g}S(\phi)} d\phi.$$

The basic idea is that for small positive g , most of the contribution to the integral comes from the neighborhood of the maximum of S , so that by expanding around it we get a power series in g . This is an asymptotic series which, although divergent, gives excellent approximations to the integral.

In the example we saw earlier,

$$n! = n^{n+1} \int_0^\infty \exp [n \{\log \phi - \phi\}] d\phi$$

so that we can identify

$$n! = n^{n+1}Y\left(\frac{1}{n}\right).$$

$$g = \frac{1}{n}, \quad S(\phi) = -\phi + \log \phi.$$

The case where the maximum is at one of the endpoints of integration introduces interesting new effects which we will study later. If the interval contains several minima of S , we should split it up into smaller intervals, each with one minimum. Examples where g is complex (purely imaginary) also appear in physics: the maxima of S also must be take into account in that case.

3.1. Expanding Around The Maximum

Let the maximum of S occur at some point $a < c < b$. Expand

$$S(\phi) = S_0 - \frac{1}{2G}(\phi - c)^2 + \sum_{k=3} S_k(\phi - c)^{rk}$$

where

$$S_0 = S(c), \quad -\frac{1}{G} = \left[\frac{d^2 S}{d\phi^2} \right]_{\phi=c},$$

and

$$S_3 = \frac{1}{3!} \left[\frac{d^3 S}{d\phi^3} \right]_{\phi=c}, \quad S_4 = \frac{1}{4!} \left[\frac{d^4 S}{d\phi^4} \right]_{\phi=c}$$

etc.

Here, the first order contribution is zero because a is an extremum. Since it is a maximum

$$G > 0.$$

Thus

$$Y(g) = \int e^{\frac{1}{g}S(\phi)} d\phi = e^{\frac{1}{g}S_0} \int e^{\frac{1}{g}[-\frac{1}{2G}(\phi-c)^2 + S_3(\phi-c)^3 + \dots]} d\phi$$

By changing variables to

$$\chi = \frac{1}{\sqrt{g}}[\phi - c]$$

$$Y(g) = \sqrt{g}e^{\frac{1}{g}S_0} \int e^{[-\frac{1}{2G}\chi^2 + \sqrt{g}S_3\chi^3 + gS_4\chi^4 + \dots]} d\chi$$

We now use the fact that

$$\int e^{-\frac{1}{2G}\chi^2} d\chi = \sqrt{2\pi G}$$

Thus

$$Y(g) = e^{\frac{1}{g}S_0} \sqrt{2\pi g G} \langle e^{\sqrt{g}S_3\chi^3 + gS_4\chi^4 + \dots} \rangle$$

where we define, for any function $f(\chi)$

$$\langle f(\chi) \rangle = \frac{\int f(\phi) e^{-\frac{1}{2G}\chi} d\chi}{\sqrt{2\pi G}}$$

It is convenient to set

$$Y(g) = e^{\frac{1}{g}S_0} \sqrt{2\pi g G} y(g), \quad y(g) = \langle e^{\sqrt{g}S_3\chi^3 + gS_4\chi^4 \dots} \rangle$$

We expand the exponential in a power series to get

$$y(g) = 1 + \langle \sqrt{g}S_3\chi^3 + gS_4\chi^4 + \frac{1}{2!} [\sqrt{g}S_3\chi^3]^2 \rangle + O(g^2)$$

It is obvious that the term of order \sqrt{g} is zero since $\langle \chi^3 \rangle = 0$. More generally, all the odd powers have zero expectation value, so that all the fractional powers of g will disappear. The even powers have expectation values

$$\langle \chi^{2n} \rangle = (2n-1)!! G^n$$

so that to first order

$$y(g) = 1 + g \left\{ 3!! S_4 G^2 + \frac{1}{2!} 5!! S_3^2 G^3 \right\} + O(g^2)$$

3.2. Higher Orders

In the second order term in g we will have contributions proportional to $S_3^4, S_3^2 S_4, S_3 S_5, S_4^2, S_6 \dots$

$$y_2(g) = g^2 \left\{ (11!!) G^6 \frac{S_3^4}{4!} + (9!!) G^5 \frac{S_3^2}{2!} S_4 + (7!!) G^4 S_3 S_5 + (7!!) G^4 \frac{S_4^2}{2!} + (5!!) S_6 G^3 \right\}$$

In general,

$$\begin{aligned} Y(g) &= \sqrt{g} e^{\frac{1}{g}S_0} \int e^{\left[-\frac{1}{2G}\chi^2 + \sum_{k=3} g^{\frac{k}{2}-1} S_k \chi^k \right]} d\chi \\ &= \sqrt{g} e^{\frac{1}{g}S_0} \int e^{-\frac{1}{2G}\chi^2} \prod_{k=3} e^{g^{\frac{k}{2}-1} S_k \chi^k} d\chi \end{aligned}$$

Expand each factor in a series, and calling the index of summation in the k th factor n_k ,

$$e^{g^{\frac{k}{2}-1} S_k \chi^k} = \sum_{n_k=0}^{\infty} g^{(\frac{k}{2}-1)n_k} \frac{S_k^{n_k}}{n_k!} \chi^{kn_k}$$

to get

$$Y(g) = \sqrt{g} e^{\frac{1}{g}S_0} \sum_{n_k=0}^{\infty} g^{\sum_{k=3} (\frac{k}{2}-1)n_k} \prod_{k=3} \frac{S_k^{n_k}}{n_k!} \int e^{-\frac{1}{2G}\chi^2} \chi^{\sum_{k=3} kn_k} d\chi$$

Evaluating the integral term by term,

$$Y(g) \equiv \int_a^b e^{\frac{1}{g}S(\phi)} d\phi \approx$$

$$(3.2.1) \quad e^{\frac{1}{g} S_0} \sqrt{2\pi g G} \sum_{n_3, n_4, \dots = 0}^{\infty} g^{\frac{1}{2} \sum_k k n_k - \sum_k n_k} \left(\sum_k k n_k - 1 \right)!! G^{\frac{1}{2} \sum_k k n_k} \prod_{k=3} \frac{S_k^{n_k}}{n_k!}$$

It is to be understood here that $m!! = 0$ for even values of m ; **beware** that this is an unusual use of the notation!!

Laplace's original method was "refined" by several mathematicians (e.g., Watson's Lemma), but they do not produce formulas that are any more useful in practical computations. Also, they do not generalize to higher dimensional integrals. So we have avoided them.

3.3. Symbolic Computation

This rather complicated formula has a graphical interpretation which helps us to visualize it (see below). This was useful in the days when finding combinatorial factors and adding fractions had to be done by the human brain. But with the easy availability of symbolic computation, it is a better strategy to translate the above formula directly into a program that can expand it out to any order desired. At order l , we have to find the non-negative integer solutions n_3, n_4, \dots to the equation

$$\sum_{k=3} \left(\frac{k}{2} - 1 \right) n_k = l$$

This is a variation on the ancient problem of finding partitions of l : the ways in which it can be written as a sum of natural numbers. The variation is that each n_k appears with a weight $\frac{k}{2} - 1$. Mathematica has a built-in command called `IntegerPartitions` that calculates these weighted partitions as well. In any case, they can be found recursively.

There are only a finite number of solutions because the weights are positive numbers. The largest value of k that can appear is given by setting $n_{k_{\max}} = 1$ and all the others equal to zero. Then

$$k_{\max} = 2l + 2$$

Next we would have a term involving $S_{k_{\max}-1} S_3$ and so on.

3.4. Graphical Interpretation

A graph is simply a set of points with lines connecting them. At any point (also called **vertex**) a certain number of lines will meet. This is the **co-ordination number** of that point. We allow for co-ordination numbers to be $2, 3, \dots$. We can translate each graph into an algebraic expression by the following **Feynman rules**:

- For each edge we put down a factor of G
- For each vertex of co-ordination number k we put a factor of S_k
- We multiply by the number of ways of contracting edges, $(\sum_k k n_k - 1)!!$
- We divide by the factor $\frac{1}{n_k!}$ which takes account of the permutations of vertices of the same co-ordination number.
- Multiply by g to the power of the number of edges minus the number of vertices.

Note that

- Graphs can be disconnected.

- Co-ordination numbers are at least 3; this rules out “external legs” in particular.

The last rules need some explanation. Since there are n_k vertices of co-ordination number k , there are $E = \frac{1}{2} \sum_k k n_k$ edges. (Each edge is shared by a pair of points. In particular, $\sum_k k n_k$ is always even.) And there are $V = \sum_k n_k$ vertices. Thus we have a factor of g^{E-V} . This quantity has an interesting topological meaning. By a theorem of Euler, it is related to the number of faces (or “loops”) of the graph.

EXAMPLE 3. Let us take the case of the Stirling expansion for the factorial.

$$n! = n^{n+1} Y\left(\frac{1}{n}\right), \quad g = \frac{1}{n}$$

$$S = -\phi + \log \phi$$

so that the maximum occurs at $\phi = 1$.

$$S_0 = -1, \quad G = 1$$

and

$$S_k = \frac{(-1)^k}{k}, \quad k = 3, 4, \dots$$

Then to second order,

$$n! = n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi} \left[1 + \frac{1}{n} \left\{ 3!! S_4 G^2 + \frac{1}{2!} 5!! S_3^2 G^3 \right\} + \frac{1}{n^2} \left\{ (11!!) G^6 \frac{S_3^4}{4!} + (9!!) G^5 \frac{S_3^2}{2!} S_4 + (7!!) G^4 S_3 S_5 + (7!!) G^4 \frac{S_4^2}{2!} + (5!!) S_6 G^3 \right\} + O(n^{-3}) \right]$$

which simplifies to

$$n! = n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi} \left[1 + \frac{1}{12n} + \frac{1}{288n^2} + O\left(\frac{1}{n^3}\right) \right]$$

Here is a table of the errors expressed as a percentage, for small values of n . You can see how the second order correction improves the answer as soon as n is moderately large.

n	%Error1	%Error2	%Error3
1	7.7863	0.101824	-0.218362
2	4.04978	0.0518567	-0.0314334
3	2.72984	0.0278913	-0.00963581
4	2.0576	0.017137	-0.00411784
5	1.65069	0.0115383	-0.00212137
6	1.37803	0.00828033	-0.00123182
7	1.18262	0.00622504	-0.000777329
8	1.03573	0.00484771	-0.00052145
9	0.921276	0.00388063	-0.000366569
10	0.829596	0.00317601	-0.000267405

EXERCISE 4. Calculate the next order term in the above expansion. Make a table of errors.

EXERCISE 5. Find the first few terms in the asymptotic expansion for large positive x of the Bessel function $K_\nu(x) = \int_0^\infty e^{-x \cosh t} \cosh \nu t \, dt$.

3.5. A Slight Generalization

Often we encounter a slight generalization of the integral in the previous section:

$$F(g) = \int_a^b f(\phi) e^{\frac{1}{g} S(\phi)} d\phi$$

where $f(\phi)$ is a function with finite derivatives at the maximum of S , which we again assume to be in the interior of the interval $a < c < b$.

$$f(\phi) \sim \sum_{r=0}^{\infty} f_r(\phi - c)^r$$

Earlier we had $f = 1$. Exactly the same argument allows us an expansion of this in powers of g as well. The graphs that appear will now have r external legs.

$$F(g) \equiv \int_a^b f(g) e^{\frac{1}{g} S(\phi)} d\phi \approx$$

$$e^{\frac{1}{g} S_0} \sqrt{2\pi g G} \sum_{n_1, n_3, n_4, \dots=0}^{\infty} g^{\frac{1}{2} n_1 + \frac{1}{2} \sum_{k=3} k n_k - \sum_k n_k} \binom{n_1 + \sum_{k=3} k n_k - 1}{n_1}!! G^{\frac{1}{2} \{n_1 + \sum_{k=3} k n_k\}} \prod_{k=3} \frac{S_k^{n_k}}{n_k!}$$

again with the understanding that $m!! = 0$ for even m .

EXERCISE 6. Find the first few terms in the asymptotic expansion for large positive x of the Bessel function $K_\nu(x) = \int_0^\infty e^{-x \cosh t} \cosh \nu t \, dt$.

EXERCISE 7. Write a symbolic program in Mathematica, Maple or Sage which implements Laplace's method, given functions S and f and the desired order of g .

The Gamma Function

The integral

$$n! = \int_0^{\infty} t^n e^{-t} dt$$

converges even for complex values of n , provided that $\operatorname{Re} n > -1$. So we can extend the definition of the factorial to complex values. This extension is called the Gamma function, apart from an awkward shift in the argument by one:

$$n! = \Gamma(n + 1)$$

The reason for this notation is obscure: many formulas would have been simpler if we had simply used $z!$ to represent the complex extension of the factorial. Thus

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$$

is an analytic function for $\operatorname{Re} z > 0$. Note that $\Gamma(1) = 0! = 1$. Using the recursion relation

$$\Gamma(z + 1) = z\Gamma(z)$$

obtained by integration by parts, we can extend the definition to the region with negative $\operatorname{Re} z$. Then $\Gamma(z)$ has poles at non-positive integers:

$$\Gamma(z) = \frac{\Gamma(z + 1)}{z} \sim \frac{1}{z}, \quad z \rightarrow 0$$

and more generally

$$\Gamma(z) = \frac{\Gamma(z + k + 1)}{z(z + 1)(z + 2) \cdots (z + k)} \sim \frac{(-1)^k}{k!(z + k)}$$

4.1. The Product Formula

The logarithmic derivative of the Gamma function is useful to establish many of its properties:

$$\psi(z) = \frac{d}{dz} \log \Gamma(z)$$

It satisfies the recursion

$$\psi(z + 1) = \frac{1}{z} + \psi(z)$$

By iterating

$$\psi(z) = -\frac{1}{z} - \frac{1}{z+1} - \cdots - \frac{1}{z+k} + \psi(z+k)$$

We can't take the limit $k \rightarrow \infty$ because the sum would diverge. But we can do that for the derivative of ψ :

$$\psi'(z) = \frac{1}{z^2} + \frac{1}{(z+1)^2} + \cdots + \psi'(z+k)$$

From the Stirling formula we can deduce that $\psi'(z) \sim \frac{1}{z}$ for large $|z|$ so we can take the limit

$$\psi'(z) = \frac{1}{z^2} + \sum_{k=1}^{\infty} \frac{1}{(z+k)^2}$$

By adding a constant to each term we can write this as the derivative of a convergent series:

$$\psi'(z) = \frac{d}{dz} \left\{ -\frac{1}{z} + \sum_{k=1}^{\infty} \left[\frac{1}{k} - \frac{1}{(z+k)} \right] \right\}$$

so there must be a constant such that

$$\psi(z) = -\gamma - \frac{1}{z} + \sum_{k=1}^{\infty} \left[\frac{1}{k} - \frac{1}{(z+k)} \right]$$

Here $\gamma \approx 0.577215$ is the **Euler constant**. Many useful results can be deduced from this series. Since

$$\psi(z) = \frac{d}{dz} \left\{ -\gamma z - \log z + \sum_{k=1}^{\infty} \left[\frac{z}{k} - \log \left(1 + \frac{z}{k} \right) \right] \right\}$$

we can conclude that

$$\log \Gamma(z) = -\gamma z - \log z + \sum_{k=1}^{\infty} \left[\frac{z}{k} - \log \left(1 + \frac{z}{k} \right) \right] + \text{constant}$$

By looking at the Laurent expansion around $z = 0$ we can conclude that the constant is zero.

This way we get Euler's product formula

$$\frac{1}{\Gamma(z)} = e^{\gamma z} z \prod_{k=1}^{\infty} \left[\left(1 + \frac{z}{k} \right) e^{-\frac{z}{k}} \right]$$

This remarkable result shows that $\frac{1}{\Gamma(z)}$ is an entire function: it is a uniformly convergent product of entire functions. So there must be a convergent expansion such that

$$\frac{1}{\Gamma(z)} = \sum_{n=1}^{\infty} a_n z^n$$

4.2. Power Series in z

The coefficients can be determined by going back to ψ and expanding it in a power series:

$$\begin{aligned}\frac{1}{z+k} &= \frac{1}{k} \frac{1}{1+\frac{z}{k}} = \frac{1}{k} \sum_{r=0}^{\infty} \frac{(-1)^r z^r}{k^r} \\ \psi(z) &= -\gamma - \frac{1}{z} + \sum_{k=1}^{\infty} \frac{1}{k} \left[1 - \sum_{r=0}^{\infty} \frac{(-1)^r z^r}{k^r} \right] \\ &= -\gamma - \frac{1}{z} + \sum_{k=1}^{\infty} \frac{1}{k} \left[- \sum_{r=1}^{\infty} \frac{(-1)^r z^r}{k^r} \right] \\ &= -\gamma - \frac{1}{z} - \sum_{r=1}^{\infty} (-1)^r z^r \sum_{k=1}^{\infty} \frac{1}{k^{r+1}}\end{aligned}$$

Recall the celebrated Riemann zeta:

$$\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}$$

so that

$$\psi(z) = -\gamma - \frac{1}{z} - \sum_{r=1}^{\infty} (-1)^r \zeta(r+1) z^r.$$

We can derive a recursion for the Taylor coefficients of the Gamma function in terms of those of ψ .

$$\psi(z) = -\frac{d}{dz} \log \frac{1}{\Gamma(z)} = -\frac{\left(\frac{1}{\Gamma(z)}\right)'}{\frac{1}{\Gamma(z)}}$$

so that

$$\frac{1}{\Gamma(z)} \psi(z) = -\left(\frac{1}{\Gamma(z)}\right)'$$

or

$$\begin{aligned}\sum_{n=1}^{\infty} a_n z^n \left[\gamma + \frac{1}{z} + \sum_{r=1}^{\infty} (-1)^r \zeta(r+1) z^r \right] &= \sum_{n=1}^{\infty} n a_n z^{n-1} \\ \sum_{n=1}^{\infty} a_n z^n \left[\gamma + \sum_{r=1}^{\infty} (-1)^r \zeta(r+1) z^r \right] &= \sum_{n=1}^{\infty} (n-1) a_n z^{n-1} \\ \sum_{m=1}^{\infty} a_m z^m \left[\gamma + \sum_{r=1}^{\infty} (-1)^r \zeta(r+1) z^r \right] &= \sum_{n=2}^{\infty} (n-1) a_n z^{n-1} \\ \sum_{n=1}^{\infty} z^n \left[\gamma a_n + \sum_{r=1}^{n-1} (-1)^r \zeta(r+1) a_{n-r} \right] &= \sum_{n=1}^{\infty} n a_{n+1} z^n\end{aligned}$$

Thus

$$a_{n+1} = \frac{\gamma}{n} a_n + \frac{1}{n} \sum_{r=1}^{n-1} (-1)^r \zeta(r+1) a_{n-r}$$

which determines each a_n in terms of lower elements in the sequence. Since $a_1 = 1$ we get

$$\begin{aligned} a_2 &= \gamma, & a_3 &= \frac{\gamma^2}{2} - \frac{\pi^2}{12}, \\ a_4 &= \frac{1}{3} \gamma \left(\frac{\gamma^2}{2} - \frac{\pi^2}{12} \right) \\ a_5 &= \frac{1}{3} \left(-\frac{\gamma\pi^2}{6} + \zeta(3) \right) \end{aligned}$$

etc.

4.3. Series in Inverse Powers

Suppose we expand

$$n! \sim n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi} \left[1 + \sum_{k=1}^{\infty} \frac{A_k}{n^k} \right]$$

We already know the first few terms

$$A_1 = \frac{1}{12}, \quad A_2 = \frac{1}{288}$$

etc. By taking the Laplace method to higher order we can calculate the higher order terms:

$$\begin{aligned} \frac{n!}{n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi}} &\sim 1 + \frac{0.0833333}{n} + \frac{0.00347222}{n^2} - \frac{0.00268133}{n^3} - \frac{0.000229472}{n^4} + \\ &\frac{0.000784039}{n^5} + \frac{0.0000697281}{n^6} - \frac{0.000592166}{n^7} - \frac{0.0000517179}{n^8} + \frac{0.000839499}{n^9} + \frac{0.000072049}{n^{10}} + O\left(\left(\frac{1}{n}\right)^{11}\right) \end{aligned}$$

Everything looks fine: the coefficients are getting smaller. But they will eventually grow to infinity. It is possible to relate these coefficients to Bernoulli numbers and use that to get their asymptotic behavior:

$$\begin{aligned} A_{2j} &\sim (-1)^{j+1} \frac{(2j-2)!}{6(2\pi)^{2j}}, \quad j \rightarrow \infty \\ A_{2j+1} &\sim (-1)^j \frac{(2j)!}{(2\pi)^{2(j+1)}}, \quad j \rightarrow \infty \end{aligned}$$

The Stirling formula itself shows that these diverge faster than any exponential of n . (See the book by Bender and Orszag for a proof). Or you can use Mathematica to calculate higher orders:

$$\begin{aligned} \frac{n!}{n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi}} &\sim \\ 1 + \frac{0.0833333}{n} + \frac{0.00347222}{n^2} - \frac{0.00268133}{n^3} - \frac{0.000229472}{n^4} + \frac{0.000784039}{n^5} + \frac{0.0000697281}{n^6} - \frac{0.000592166}{n^7} \end{aligned}$$

$$\begin{aligned}
& -\frac{0.0000517179}{n^8} + \frac{0.000839499}{n^9} + \frac{0.000072049}{n^{10}} - \frac{0.00191444}{n^{11}} - \frac{0.000162516}{n^{12}} + \frac{0.00640336}{n^{13}} + \frac{0.000540165}{n^{14}} \\
& -\frac{0.0295279}{n^{15}} - \frac{0.00248174}{n^{16}} + \frac{0.17954}{n^{17}} + \frac{0.0150561}{n^{18}} - \frac{1.3918}{n^{19}} - \frac{0.116546}{n^{20}} + \frac{13.398}{n^{21}} + \frac{1.1208}{n^{22}} - \frac{156.801}{n^{23}} - \frac{13.1079}{n^{24}} \\
& + \frac{2192.56}{n^{25}} + \frac{183.191}{n^{26}} - \frac{36101.1}{n^{27}} - \frac{3015.08}{n^{28}} + \frac{691346.}{n^{29}} + \frac{57721.3}{n^{30}} - \frac{1.52358 \times 10^7}{n^{31}} - \frac{1.27174 \times 10^6}{n^{32}} \\
& + \frac{3.82848 \times 10^8}{n^{33}} + \frac{3.19498 \times 10^7}{n^{34}} - \frac{1.08809 \times 10^{10}}{n^{35}} - \frac{9.07895 \times 10^8}{n^{36}} + \frac{3.47282 \times 10^{11}}{n^{37}} + \frac{2.89728 \times 10^{10}}{n^{38}} \\
& - \frac{1.23684 \times 10^{13}}{n^{39}} - \frac{1.03174 \times 10^{12}}{n^{40}} + O\left(\left(\frac{1}{n}\right)^{41}\right)
\end{aligned}$$

4.3.1. Comparison of Asymptotic to Convergent Expansions. Now we can compare the two schemes for approximating the factorial: the Stirling series we had before and the (reciprocal of the) convergent series we just obtained.

$$n! \approx n^{n+\frac{1}{2}} e^{-n} \sqrt{2\pi} \left[1 + \frac{1}{12n} + \frac{1}{288n^2} + \dots \right] \quad \text{vs} \quad \frac{1}{\sum_{r=1}^{\infty} a_r (n+1)^r}$$

n	convergent 2	convergent 7	convergent 12	convergent 17	convergent 22	$n!$
0.25	0.464706	0.93482	0.9064	0.906402	0.906402	0.906402
0.5	0.357304	1.00796	0.886231	0.886226	0.886227	0.886227
0.75	0.284275	1.54701	0.919265	0.919043	0.919062	0.919063
1.	0.23208	-4.68964	1.00239	0.999747	0.999994	1.
1.25	0.193343	-0.48646	1.15287	1.13038	1.13292	1.133
1.5	0.16373	-0.175665	1.47506	1.30629	1.3283	1.32934
1.75	0.140544	-0.0804508	3.11156	1.44182	1.59752	1.60836
2.	0.122027	-0.0412329	-1.81959	1.1982	1.90351	2.
2.25	0.106988	-0.0227033	-0.348258	0.56579	1.90909	2.54926
2.5	0.0945994	-0.0131819	-0.112141	0.185839	1.10643	3.32335
2.75	0.0842666	-0.00798623	-0.0420393	0.0576534	0.369141	4.42299
3.	0.0755547	-0.00501388	-0.0171793	0.0187194	0.10546	6.
3.25	0.0681387	-0.00324559	-0.00748784	0.00647034	0.0309338	8.28509
3.5	0.0617718	-0.0021578	-0.00344359	0.00237564	0.00964716	11.6317
3.75	0.0562638	-0.0014688	-0.00165916	0.00092149	0.00320174	16.5862
4.	0.0514658	-0.001021	-0.000832988	0.000375588	0.00112531	24.
4.25	0.0472601	-0.000723193	-0.000433861	0.000160086	0.00041663	35.2116
4.5	0.0435525	-0.000521003	-0.00023357	0.0000710563	0.000161708	52.3428
4.75	0.0402671	-0.000381146	-0.00012956	0.0000327248	0.0000655217	78.7845
5.	0.0373416	-0.000282755	-0.0000738467	0.0000155884	0.0000276124	120.

n	$n!$	asymptotic 0	asymptotic 1	asymptotic 2
0.25	0.906402	0.690194	0.920259	0.958603
0.5	0.886227	0.760173	0.886869	0.897427
0.75	0.919063	0.82641	0.918234	0.923335
1.	1.	0.922137	0.998982	1.00218
1.25	1.133	1.06124	1.13199	1.13435
1.5	1.32934	1.25843	1.32835	1.33029
1.75	1.60836	1.5343	1.60736	1.6091
2.	2.	1.919	1.99896	2.00063
2.25	2.54926	2.45714	2.54814	2.54983
2.5	3.32335	3.21495	3.32211	3.3239
2.75	4.42299	4.29152	4.42157	4.42354
3.	6.	5.83621	5.99833	6.00058
3.25	8.28509	8.07598	8.28306	8.28572
3.5	11.6317	11.3588	11.6292	11.6324
3.75	16.5862	16.2225	16.583	16.587
4.	24.	23.5062	23.9959	24.001
4.25	35.2116	34.5291	35.2062	35.2128
4.5	52.3428	51.3839	52.3355	52.3443
4.75	78.7845	77.4163	78.7745	78.7864
5.	120.	118.019	119.986	120.003

Although the series eventually converges to the right value, the number of terms to be kept grows very rapidly with n . The first few terms can be misleading: we even get negative values. The asymptotic series on the other hand, gives good answers for the first few terms but will diverge away if we keep too many terms. Convergent series are the way to prove theorems; but they are often lousy at calculating approximate numerical answers. So theoretical physics is mostly based on asymptotic expansions: convergent expansions are rarely available in interesting situations. Even when they are, they may not be the best approximation methods.

EXERCISE 8. Define $F(g) = \int e^{-\frac{1}{2}\phi^2 - g\phi^4} \frac{d\phi}{\sqrt{2\pi}}$. Find the coefficients in the expansion in powers of g . Does it converge? By a change of variables, show that $g^{\frac{1}{4}}F(g) = \int e^{-\frac{1}{2\sqrt{g}}\chi^2 - \chi^4} \frac{d\chi}{\sqrt{2\pi}}$. Obtain the coefficients in its expansion in powers of $z = g^{-\frac{1}{2}}$. Does this converge in some region in the z -plane? By numerically evaluating the integral as well as the leading terms of the two approximation methods, determine the region of validity of each.

Generating Functions And Their Approximations

A **sequence** is a map $a : \mathbb{N} \cup \{0\} \rightarrow \mathbb{C}$ that assigns to any non-negative integer a complex number. It is possible to add them termwise and multiply them by a complex number;

$$(a + b)_n = a_n + b_n, \quad [\lambda a]_n = \lambda a_n$$

i.e., the set of sequences form a **vector space**.

If we think of the terms of a sequence as the coefficients of a power series

$$a(z) = \sum_{n=0}^{\infty} a_n z^n$$

we say that $a(z)$ is the **generating function** of the sequence a .

EXAMPLE 9. Suppose P_n is the number of ways in which a positive integer can be written as a sum of other positive integers; i.e., the number of solutions to the equation

$$n = \sum_{k=1}^{\infty} k m_k$$

where $m_k = 0, 1, \dots$ is the number of times k appears in that partition of n .

For example, $3 = 1 + 1 + 1 = 1 + 2$ so that $P_3 = 2$. Similarly $P_4 = 5$, $P_5 = 7$, $P_6 = 11$, $P_{98} = 50198136$ etc. The generating function $P(q) = \sum_{n=1}^{\infty} P_n q^n = \sum_{n_1, n_2, \dots=0}^{\infty} q^{n_1} q^{2n_2} q^{3n_3} \dots$

$$\begin{aligned} P(q) &= \sum_{n=1}^{\infty} P_n q^n = \sum_{n_1, n_2, \dots=0}^{\infty} q^{n_1 + 2n_2 + 3n_3 + \dots} \\ &= \sum_{n_1=0}^{\infty} q^{n_1} \sum_{n_2=0}^{\infty} q^{2n_2} \sum_{n_3=0}^{\infty} q^{3n_3} \dots \\ P(q) &= \prod_{k=1}^{\infty} \frac{1}{1 - q^k}. \end{aligned}$$

The fact that this product converges when $|q| < 1$ shows that the power series $\sum_n P_n q^n$ converges as well, when $|q| < 1$. This by itself shows that P_n has to grow for large n slower than an exponential. More precisely, the radius of convergence of the power series

$$\limsup_n |P_n|^{\frac{1}{n}} < 1.$$

Hardy and Ramanujan developed a deep theory of the asymptotic behavior of P_n for large n by studying the singularities of this function $P(q)$. In the leading order

$$P_n \sim e^{\pi\sqrt{\frac{2n}{3}}}$$

in agreement with the sub-exponential growth we noted earlier.

EXAMPLE 10. The generating function for partitions into odd numbers is $\prod_{r=1}^{\infty} \frac{1}{1-q^{2r+1}}$. That for distinct parts is $\prod_{k=1}^{\infty} (1+q^k)$. From the identity

$$\prod_{r=1}^{\infty} (1-q^{2r+1}) \prod_{k=1}^{\infty} (1+q^k) = 1$$

(Prove it!) it follows that the number of ways of writing n as a sum of distinct parts and the number of ways of splitting it into odd parts are the same. This means that the partition function of a system of fermions with energies $1, 2, 3, \dots$ is the same as that of a system of bosons of energies $1, 3, 5, 7, \dots$

There is a natural notion of multiplication for sequences that is induced by the pointwise product of the functions $a(z)$:

$$a(z)b(z) = \sum_{p=0}^{\infty} a_p z^p \sum_{q=0}^{\infty} b_q z^q$$

$$(ab)_n = \sum_{p+q=n} a_p b_q.$$

Since p and q are ≥ 0 this is always a finite sum. Thus it makes sense to multiply power series term by term even if the series itself does not converge! We say in this case that $\sum_n a_n z^n$

is a **formal power series**: it is merely a convenient notation for the sequence a . Thus we can do algebra with formal power series: they form a ring, which contains the ring of polynomials.

Using the notation of a power series, even when z cannot be assigned a complex value, allows us to manipulate them using the familiar rules for functions. Thus, it makes sense to say that derivative of a sequence is

$$[a']_n = (n+1)a_{n+1}$$

since

$$\frac{d}{dz} \sum_{n=0}^{\infty} a_n z^n = \sum_{n=0}^{\infty} (n+1)a_{n+1} z^n$$

For example,

$$\frac{d}{dz} \sum_{n=0}^{\infty} n! z^n = \sum_{n=0}^{\infty} (n+1) \times (n+1)! z^n$$

makes sense for formal power series even if the series converges only at the one point $z = 0$. When is it possible to divide a formal power series by another?

$$a(z)b(z) = 1 \implies$$

$$a_0 b_0 = 1, \quad a_0 b_1 + a_1 b_0 = 0, \quad a_0 b_2 + a_1 b_1 + a_2 b_0 = 0 \cdots$$

Thus, as long as $a_0 \neq 0$, we can solve these equations recursively:

$$b_0 = \frac{1}{a_0}, \quad b_n = -\frac{1}{a_0} \sum_{p=1}^{n-1} a_p b_{n-p}, \quad n > 0$$

More generally, we can form a ratio $\frac{f(z)}{g(z)}$ provided that $g_0 \neq 0$. The coefficients of $h(z) = \frac{f(z)}{g(z)}$ are determined by solving

$$f_n = \sum_{p+q=n} h_p g_q$$

i.e.,

$$f_n = g_0 h_n + \sum_{p=1}^{n-1} h_p g_{n-p}$$

from which we get the recursion relation

$$h_0 = \frac{f_0}{g_0}$$

$$h_n = \frac{1}{g_0} f_n - \frac{1}{g_0} \sum_{p=1}^{n-1} h_p g_{n-p}$$

This is how we found the coefficients of the expansion of $\frac{1}{\Gamma(z)}$ in the last chapter. Without convergence it also makes sense to compose formal power series

$$h(z) \equiv f \circ g(z) = f(g(z)) = \sum_{r=0}^{\infty} f_r [g(z)]^r$$

If $g_0 = 0$, (this is the **opposite** of the condition needed to divide by $g(z)$) we can write

$$h_0 = f_0$$

$$h_n = \sum_r f_r \sum_{p_1+p_2+\cdots+p_r=n} g_{p_1} g_{p_2} \cdots g_{p_r}, \quad n > 0$$

For any given n there are only a finite number of terms in this sum: each term corresponds to a partition of n . Sometimes we may need to find the compositional inverse of $f(z)$. It exists as long as $f_0 = 0$ and $f_1 \neq 0$:

$$f(g(z)) = z$$

Setting $g_0 = 0$ we get

$$f_1 g_1 z + [f_2 g_1^2 + f_1 g_2] z^2 + [f_1 g_1^3 + f_2 g_1 +] z^3 + \cdots = 0$$

$$g_1 = \frac{1}{f_1}, \quad g_2 = -\frac{f_2}{f_1} g_1^2,$$

etc.

There is a formula of Lagrange that expresses the general term of the inversion of a formal power series.

5.1. Euler Summation

Can we actually calculate the value of the generating function for some value z , even if the sum diverges? We might be able to give it a meaning if we have additional physical information about the quantity it represents.

For example, if

$$f(z) = \sum_n f_n z^n$$

converges for $|z| < 1$, and the function $f(z)$ has a limit as $z \rightarrow 1$ we can assign that as the value of the series. This is **Euler summation**. For example

$$\sum_{n=0}^{\infty} (-1)^n := \lim_{z \rightarrow 1} \frac{1}{1+z} = \frac{1}{2}.$$

Such cheap tricks have to be justified independently based on the physical application, however.

5.2. Borel Summation

Suppose f_n grows faster than a power of n so that

$$f(z) = \sum_n f_n z^n$$

is not a convergent series. This is the case for perturbation theory in quantum mechanics, for example. But it may be that the **Borel transform**

$$B(z) = \sum_{n=0}^{\infty} \frac{a_n}{n!} z^n$$

converges for some k . Then we can recover $f(z)$ by inverting the Borel transform

$$f(z) = \int_0^{\infty} B(tz) e^{-t} dt.$$

Sometimes we may need to iterate this several times

$$B_k(z) = \sum_{n=0}^{\infty} \frac{a_n}{(n!)^k} z^n$$

to get a convergent answer. The inversion is

$$f(z) = \int_0^{\infty} B_k(t_1 \cdots t_k z) e^{-t_1 - t_2 \cdots - t_k} dt_1 \cdots dt_k$$

Often such series arise as asymptotic expansion of some integral by Laplace's method. Then we know that the physical quantity it represents is meaningful, the series being divergent is only the fault of the approximation method. This is the case for perturbation theory in quantum mechanics, for example. The integral is hard to calculate, but makes perfect sense; the expansion can be calculated term

by term, but diverges. By combining the two pieces of information we can calculate the value needed.

EXAMPLE 11. Suppose $b(z) = \sum_n \frac{a_n}{n!} z^n$ converges. Since

$$n! = \int_0^\infty t^n e^{-t} dt$$

$$\sum_n a_n z^n = \sum_{n=0}^\infty n! \frac{a_n}{n!} z^n = \int_0^\infty b(tz) e^{-t} dt$$

If you can calculate the integral, you can sum the series.

EXAMPLE 12. Suppose $\mu_n = \int_{-\infty}^\infty \phi^n p(\phi) d\phi$ are the moments of a random variable with probability density $p(\phi) d\phi$. Then the generating function of moments

$$\sum_n \mu_n J^n := \int \frac{1}{1 - \phi J} p(\phi) d\phi$$

might make sense as an integral. This can be used to give a meaning to the sum. For the exponential distribution,

$$\sum_{n=0}^\infty n! (-1)^n := \lim_{J \rightarrow -1} \int_0^\infty \frac{1}{1 - \phi J} e^{-\phi} d\phi \approx 0.596347$$

5.3. Pade' Approximants

But what if we only know a few terms of a divergent series? Can we still find a good approximation for its value? This is the situation confronting us in physical calculations. The best answer is not necessarily just to add the first few terms of the series. For example,

$$\log[1 - z] = - \sum_n \frac{z^n}{n}$$

has a branch cut starting at $z = 1$. The polynomial $-\sum_{n=1}^M \frac{z^n}{n}$ has no such singularity. So it will be a bad approximation. We can do better by approximating the function by a ratio of polynomials

$$f(z) \approx \frac{\sum_{n=1}^M A_n z^n}{\sum_{n=1}^N B_n z^n}$$

The particular case $N = 0$ is the power series terminated at order M . When $N > 0$, this has poles along the zeros of the denominator. Many such zeros packed close together can approximate a branch cut. With $M = N = 5$,

$$\log[1 - z] \approx \frac{-\frac{137z^5}{7560} + \frac{11z^4}{36} - \frac{47z^3}{36} + 2z^2 - z}{-\frac{z^5}{252} + \frac{5z^4}{42} - \frac{5z^3}{6} + \frac{20z^2}{9} - \frac{5z}{2} + 1}$$

with poles at

$$1.04922, 1.29999, 2., 4.33341, 21.3174$$

and residues

0.130412, 0.404437, 1.13778, 4.49394, 53.8334

Even when the function being approximated is entire, a ratio of polynomials can be a better approximation than a polynomial.

The coefficients A_n, B_n can be determined from the power series

$$f(z) = \sum_n f_n z^n$$

matching the two:

$$\sum_{n=0}^M A_n z^n = \left[\sum_{p=0}^{M+N} f_p z^p \right] \left[\sum_{q=0}^N B_q z^q \right] + O(z^{M+N+1}).$$

to the desired order. We can write these as a system of linear equations (setting $B_0 = 1$ with no loss of generality) which determine the numerator and denominator. The B_q are determined by solving the $N \times N$ system

$$\begin{pmatrix} f_{M+N-1} & f_{M+N-2} & f_{M+N-3} & \cdot & \cdot & \cdot & f_M \\ f_{M+N-2} & f_{M+N-3} & f_{M+N-4} & \cdot & \cdot & \cdot & f_{M-1} \\ f_{M+N-3} & f_{M+N-4} & f_{M+N-5} & \cdot & \cdot & \cdot & f_{M-2} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ f_M & f_{M-1} & f_{M-2} & \cdot & \cdot & \cdot & f_1 \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \\ B_3 \\ \cdot \\ \cdot \\ \cdot \\ B_N \end{pmatrix} = \begin{pmatrix} -f_{M+N} \\ -f_{M+N-1} \\ -f_{M+N-2} \\ \cdot \\ \cdot \\ \cdot \\ -f_{M+1} \end{pmatrix}$$

and then the A coefficients are simply given by

$$A_n = \sum_{q=0}^n f_{n-q} B_q$$

If the series $\sum_n f_n z^n$ diverges, or converges only slowly, this trick can give quite good answers. Even for convergent series, it can improve the rate of convergence.

EXAMPLE 13. $f(J) = \int_0^\infty \frac{1}{1+\phi^J} e^{-\phi} d\phi$

As we saw before, this function has the asymptotic series

$$f(J) = \sum_{n=0}^\infty (-1)^n n! J^n$$

Suppose we terminate this series a few terms and calculate the values. We get good answers for small J but the answers do not get better by including higher order terms as J grows. But using the same information repackaged as Pade' approximants does get better answers, as we can see in the figure.

The blue line is exact (from the integral); the solid green and solid red are sums of 2 and 4 terms respectively; and the dashed green and dashed red are the Pade' diagonal approximants of order 1 and 2, which use the same coefficients as the corresponding power series.

We can see how well the combination of asymptotic series with Pade' approximants works for large values of J , by extending the range of the plot. As J becomes large, the error in the low Pade' approximants of order 1 and 2 grow; but by going to a higher order we again get good approximations. How do we choose the

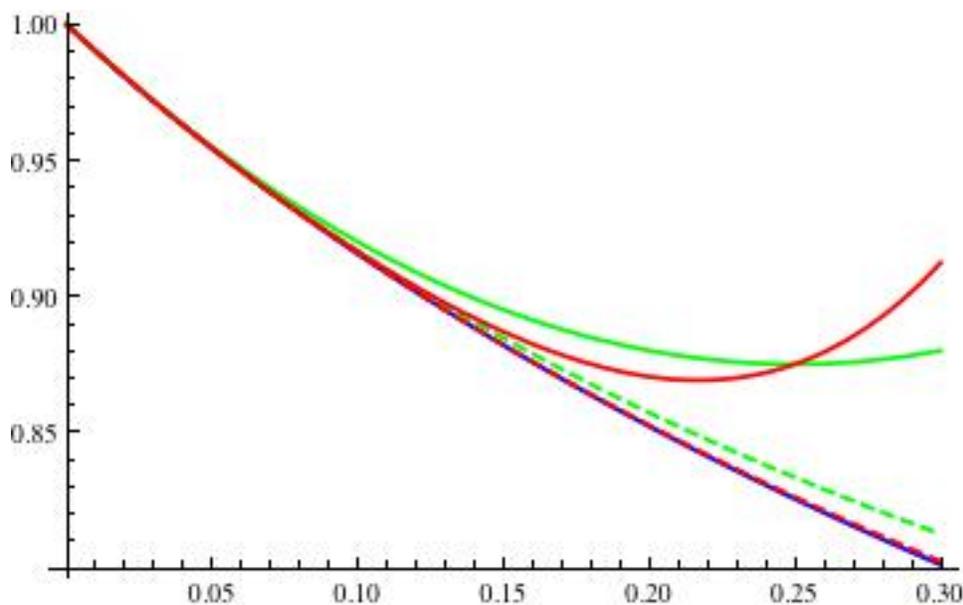


FIGURE 5.3.1.

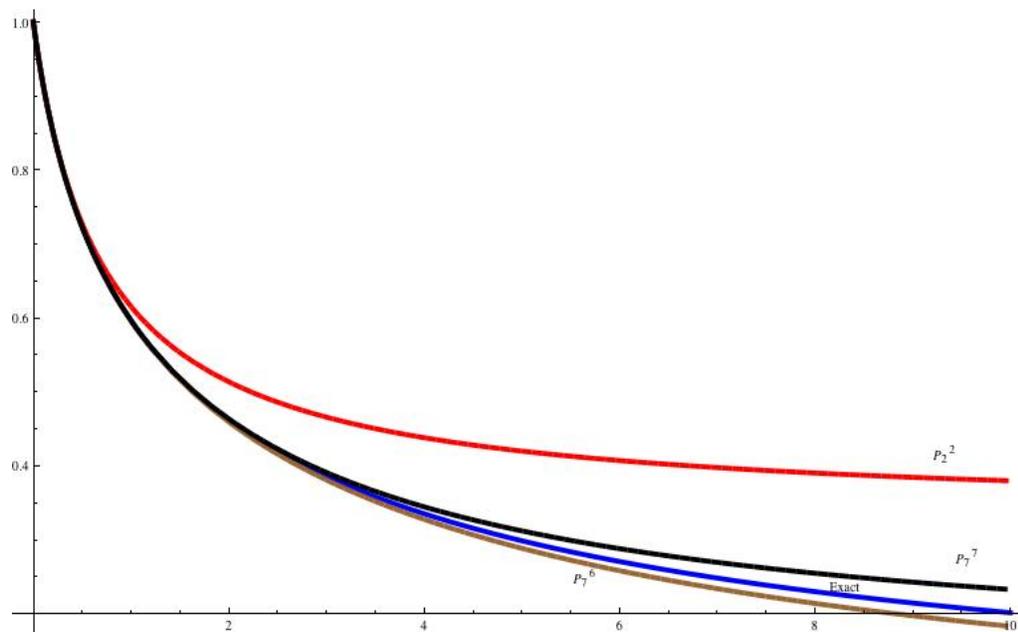


FIGURE 5.3.2.

orders of the numerator and the denominator in the Pade' fraction? By looking at the behavior of the integral, we see that $\int_0^\infty \frac{1}{1+\phi J} e^{-\phi} d\phi \sim \frac{J}{\log J}$ for large J . (Prove this!). So we should expect the answer to be somewhere between $P_N^N(J)$ (which goes as J^0 for large J) and $P_N^{N-1}(J)$, which goes as J^{-1}). See the plot to see that this is indeed true. The error in the power series itself is intolerable in this range.

Continued Fractions

6.1. Continued fractions provide the best rational approximations to real numbers

The decimal expansions of a positive real number is a kind of power series: $x = \sum_n x_n 10^{-n}$ where the coefficients x_n are integers between 0 and 9. If we terminate this series at some point (keep some finite number of decimal places) we get a rational approximation to the number. But this is not necessarily the best approximation to that number with that denominator. The best approximations are given by continued fractions.

Given a real number x , define $a_0 = \lfloor x \rfloor$ to be the largest integer smaller than x . For example, $\lfloor 2.1 \rfloor = 2 = \lfloor 2.8 \rfloor$. Then $\{x\} = x - \lfloor x \rfloor$ is a positive number smaller than one, its fractional part. If it is non-zero, the reciprocal $x_1 = \frac{1}{\{x\}}$ is greater than one. In the same spirit, define the reciprocal of its fractional part: $x_2 = \frac{1}{\{x_1\}}$, $a_2 = \lfloor x_2 \rfloor$ and so on

$$x_{n+1} = \frac{1}{\{x_n\}}, \quad a_{n+1} = \lfloor x_{n+1} \rfloor$$

unless any of the fractional parts vanish, in which the the sequence terminates. This expresses the real number as a **simple continued fraction**

$$x = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 \dots}}}$$

This sequence $a_0, a_1 a_2 \dots$ terminates if and only if x is rational. Otherwise, it gives the most rapidly converging rational approximation to x . A more convenient notation is $x = [a_0, a_1, a_2, \dots]$

6.1.1. Celebrated examples are.

$$\sqrt{2} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 \dots}}}$$

$$\pi = [3, 7, 15, 1, 292, 1, 1, 1, 2, 1, 3, 1, 14, 2, 1, 1, 2, 2, 2, 2, \dots]$$

6.1.2. Quadratic Equations. It is possible to solve a quadratic equation in terms of a simple continued fraction, giving a series of approximations to the solution in terms of rational functions of the coefficient. Euler improved even this elementary fact, although it was known centuries before (for example to Indian mathematicians).

Any quadratic equation

$$ax^2 + bx + c = 0$$

can be brought to the form

$$y^2 = \Delta$$

by rational transformations

$$x = \frac{-b + y}{2a}, \quad \Delta = b^2 - 4ac$$

The combination $\Delta = b^2 - 4ac$ is the discriminant of the quadratic: if it is zero, the solutions coincide and the equation is trivial to solve. If it is negative, the solution is complex. If it is positive, find a rational number y_0 such that

$$y_0^2 < \Delta$$

This is the starting point of an approximation method to solving the quadratic:

$$y^2 - y_0^2 = \Delta - y_0^2$$

$$(y + y_0)(y - y_0) = \Delta - y_0^2$$

Thus any quadratic equation can be brought to the form

$$y = y_0 + \frac{\Delta - y_0^2}{y_0 + y}$$

This leads to the continued fraction

$$y = y_0 + \frac{\Delta - y_0^2}{2y_0 + \frac{\Delta - y_0^2}{2y_0 + \frac{\Delta - y_0^2}{y_0 + \dots}}}$$

As long as $\Delta - y_0^2 > 0$, this converges to a positive real number which solves the quadratic. With

$$\Delta = 2, \quad y_0 = 1$$

we get the continued fraction we mentioned earlier

$$\sqrt{2} = 1 + \frac{1}{2 + \frac{1}{2 + \frac{1}{2 + \dots}}}$$

6.2. General Continued Fractions

More generally, we encounter fractions which may not have unity as numerator:

$$x = b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \frac{a_3}{b_3} \dots}}$$

which we denote by something easier to type:

$$b_0 + \frac{a_1}{b_1 +} \frac{a_2}{b_2 +} \frac{a_3}{b_3 +} \dots$$

or by

$$x = b_0 + \mathbb{K}_{n=1}^{\infty} \frac{a_n}{b_n}$$

A moment's thought will show that the transformation

$$a_n \rightarrow r_{n-1}r_n a_n, \quad b_n \rightarrow r_n b_n$$

for $r_n \neq 0, r_0 = 1$ leaves the value of the fraction unchanged:

$$b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \frac{a_3}{b_3} \dots}} = b_0 + \frac{r_1 a_1}{r_1 b_1 + \frac{r_1 r_2 a_2}{r_2 b_2 + \frac{r_2 r_3 a_3}{r_3 b_3} \dots}}$$

Such equivalence transformations are sometimes used to bring a continued fraction to some convenient form; e.g., $a_n = 1$ or $b_n = 1$.

6.3. Second Order Linear Difference Equations

The recurrence relation (linear difference equation)

$$u_n = b_n u_{n-1} + a_n u_{n-2}$$

plays an important role in the theory of c.f. Such equations arise often as discrete approximations to differential equations.

EXAMPLE 14. Suppose we have a Schrodinger equation in one dimension

$$(6.3.1) \quad -\psi'' + [V(x) - E]\psi = 0, \quad x \in [a, b].$$

A standard approach to solving it numerically is to approximate the real line by a discrete set of points

$$x_n = a + \frac{(b-a)}{N}n, \quad n = 0, 1, \dots, N$$

$$u_n = \psi(x_n)$$

$$\psi''(x_n) \approx \frac{u_{n+1} - 2u_n + u_{n-1}}{N^{-2}(b-a)^2}$$

EXERCISE 15. Find a_n, b_n for the discrete version of the Schrodinger equation above.

(6.3.1) has two linearly independent solutions. Let A_n, B_n be the solutions with initial conditions

$$A_0 = b_0, \quad A_{-1} = 1$$

$$B_0 = 1, \quad B_{-1} = 0$$

respectively. They turn out to be the numerator and denominator of the continuant

$$b_0 + \frac{a_1}{b_1 + \frac{a_2}{b_2 + \dots + \frac{a_n}{b_n}}} = \frac{A_n}{B_n}$$

Thus, solving the recursion relations converts the c.f. into an ordinary fraction. To prove this, note first the following result:

LEMMA 16. For any ξ ,

$$b_0 + \frac{a_1}{b_1 +} \frac{a_2}{b_2 +} \cdots \frac{a_n}{+ \xi} = \frac{\xi A_{n-1} + a_n A_{n-2}}{\xi B_{n-1} + a_n B_{n-2}}$$

PROOF. Proceed by induction. The case $n = 1$ is obvious:

$$b_0 + \frac{a_1}{\xi} = \frac{b_0 \xi + a_1}{\xi} = \frac{\xi A_0 + a_1 A_{-1}}{\xi B_0 + a_1 B_{-1}}.$$

Assume that the identity holds for some m . Define $\eta = b_m + \frac{a_{m+1}}{\xi}$. Then,

$$b_0 + \frac{a_1}{b_1 +} \frac{a_2}{b_2 +} \cdots \frac{a_{m+1}}{+ \xi} = b_0 + \frac{a_1}{b_1 +} \frac{a_2}{b_2 +} \cdots \frac{a_m}{+ \eta}$$

By the inductive hypothesis

$$\begin{aligned} b_0 + \frac{a_1}{b_1 +} \frac{a_2}{b_2 +} \cdots \frac{a_m}{+ \eta} &= \frac{\eta A_{m-1} + a_m A_{m-2}}{\eta B_{m-1} + a_m B_{m-2}} \\ &= \frac{\left[b_m + \frac{a_{m+1}}{\xi} \right] A_{m-1} + a_m A_{m-2}}{\left[b_m + \frac{a_{m+1}}{\xi} \right] B_{m-1} + a_m B_{m-2}} \\ &= \frac{\xi [b_m A_{m-1} + a_m A_{m-2}] + a_{m+1} A_{m-1}}{\xi [b_m B_{m-1} + a_m B_{m-2}] + a_{m+1} B_{m-1}} \\ &= \frac{\xi A_m + a_{m+1} A_{m-1}}{\xi B_m + a_{m+1} B_{m-1}} \end{aligned}$$

which proves the identity for $n = m + 1$. □

If we now put in $\xi = b_n$ and use the recursion relation

$$b_0 + \frac{a_1}{b_1 +} \frac{a_2}{b_2 +} \cdots \frac{a_n}{+ b_n} = \frac{A_n}{B_n}$$

as promised.

6.4. Rational Approximation of Analytic Functions Of A Complex Variable

Polynomials form a ring analogous to the ring of integers. If $f(z)$ is a complex function which is meromorphic in some neighborhood of the point at infinity, we can define $a_0(z) = [f(z)]$ to be the unique polynomial such that $f(z) - a_0(z)$ vanishes at infinity. Also, $[f] = f - [f]$ be the "fractional part" of f . Then $f_1 = \frac{1}{[f]}$ has a pole at infinity (unless $f = a_0$ identically in which case we stop) and we can define $a_1(z) = [f_1(z)]$, and continue with $f_2 = \frac{1}{[f_1]}$, $a_2 = [f_2]$ and so on unless $f_r - a_r$ vanishes identically. Thus we can express such a function as a continued fraction

$$f = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}$$

If the function is analytic in some neighborhood of the origin, (it could have a pole of finite order at the origin) it can be represented by a power series

$$f(z) = \sum f_n z^n$$

The power series representation of an analytic function is like the decimal expansion of a number. Knowing the first N coefficients will give an accuracy of $|z|^{N+1}$ in the calculation of $f(z)$. But the same coefficients can be used to construct a fraction of polynomials that gives a more accurate estimate of $f(z)$, if we know something about the asymptotic behavior of f . This can work well even when $f(z)$ has an essential singularity at the origin, for example from the Laplace method applied to an integral.

EXAMPLE 17. An ancient example: rational approximation of \sin

Converting to modern units, Bhaskara's formula is

$$\sin \frac{\pi}{n} \approx \frac{16(n-1)}{5n^2 + 4n - 4}$$

which works well in the useful range $0 < \frac{\pi}{n} < \frac{\pi}{2}$: the error is less than a percent or so.

EXAMPLE 18. Another ancient example: $\tan(x)$

Being the ratio of two entire functions, it is intuitive that \tan should have a nice continued fraction expansion. The power series expansion has coefficients for which there is no simple formula (Bernoulli numbers) and moreover diverges at $|x| = \pi$. A continued fraction is

$$\tan x = -\frac{1}{x - \frac{1}{x - \frac{3}{x - \frac{5}{x - \dots}}}}$$

with negative odd integer coefficients.

6.5. An Example

The example $f(z) = z + \sqrt{1+z^2}$ is useful to understand how fractions approximate a function with a branch cut. There is a branch cut which we can choose to be the line connecting $-i$ to i . The two branches

$$f_+(z) = z + \sqrt{1+z^2}, \quad f_-(z) = z - \sqrt{1+z^2}$$

are related by

$$f_-(z) = -\frac{1}{f_+(z)}.$$

Also,

$$f_+(z) = 2z + \frac{1}{f_+(z)}$$

leading to the continued fractions

$$f_+(z) = 2z + \frac{1}{2z + \frac{1}{2z + \dots}}$$

$$f_-(z) = -\frac{1}{2z + \frac{1}{2z + \dots}}.$$

The even order continuants of the first cf are rational functions with a zero at the origin and poles placed along the line connecting $-i$ to $+i$. The odd order

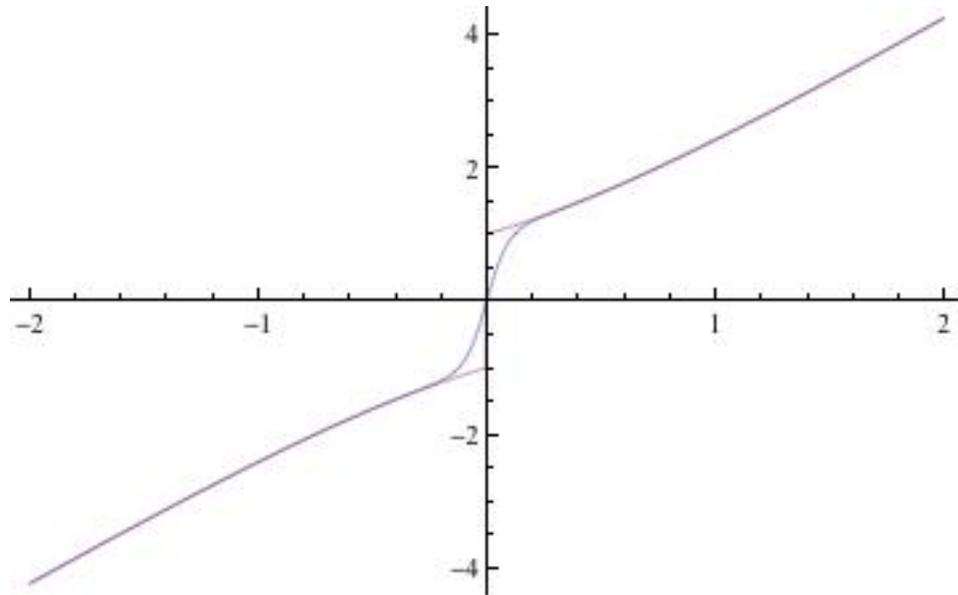
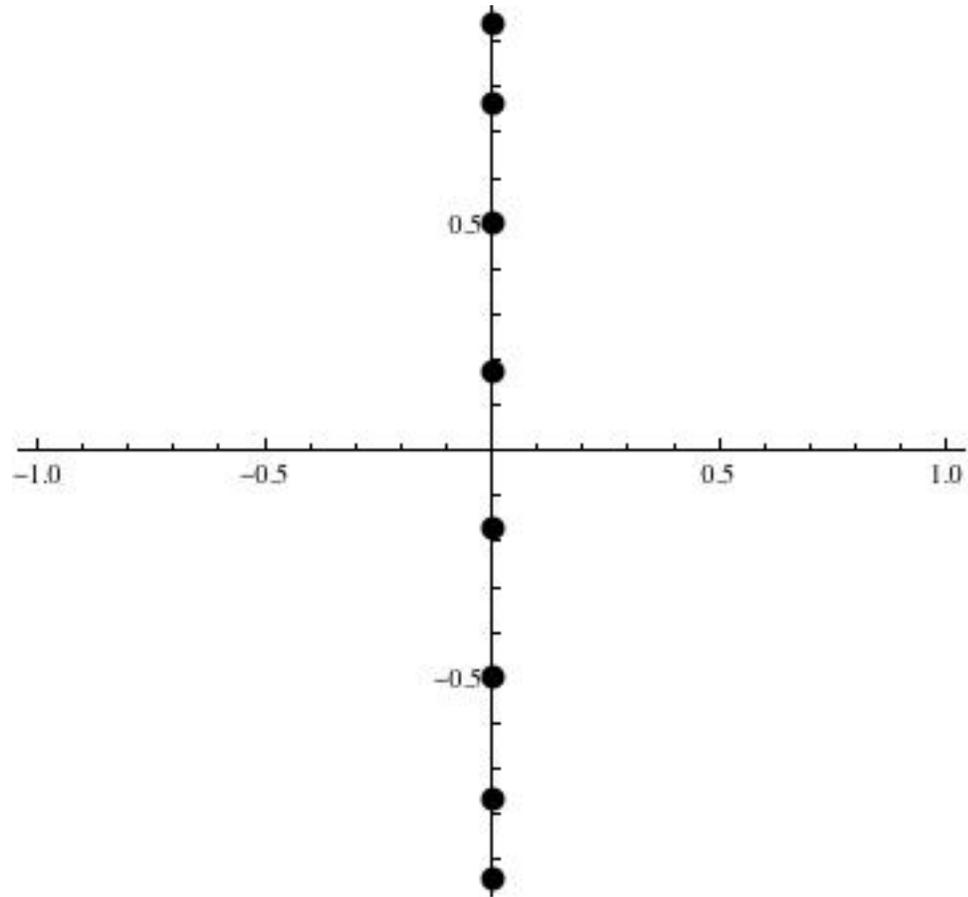


FIGURE 6.5.1.

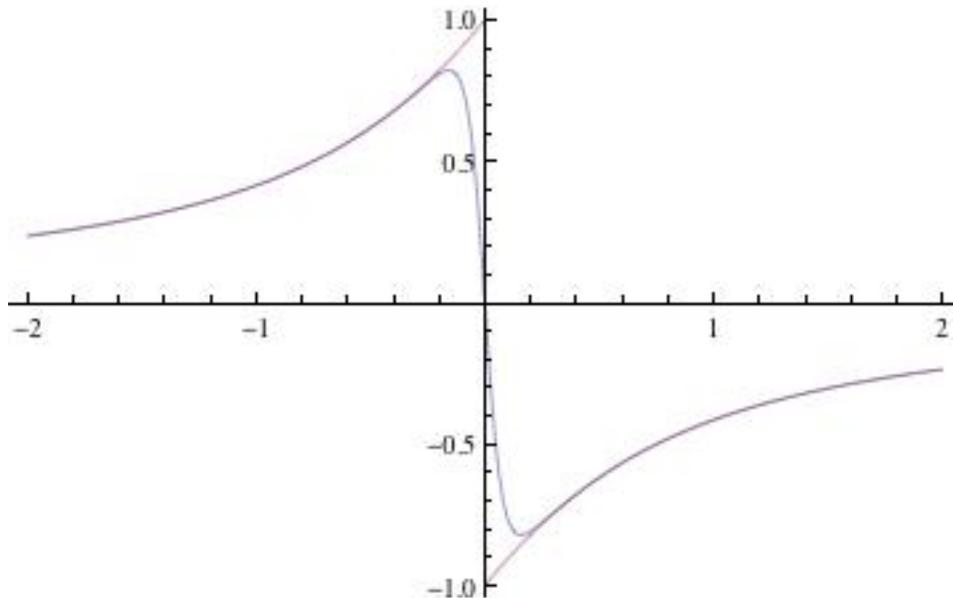
contiguants of f_+ also have poles along this line, but have a pole at the origin. Away from the origin, either of these give a good fit to $f_+(x)$ for real x . But those with a zero at the origin provide a an approximation that is smooth on the real axis. At order 8, we have

$$R_8(z) = 2z + \frac{8(z + 10z^3 + 24z^5 + 16z^7)}{1 + 40z^2 + 240z^4 + 448z^6 + 256z^8}$$

The poles along the imaginary axis can be found.



Similarly, the negative of the reciprocals of the odd continuants of f_+ provide smooth approximations to $f_-(x)$.



6.6. Pade' Approximants as Continued Fractions

Suppose the Pade' approximant $P_M^M(z)$ is expressed as a continued fraction

$$P_M^M(z) = \frac{c_0}{1+} \frac{c_1 z}{1+} \frac{c_2 z}{1+} \cdots \frac{c_{2M-1} z}{1+ c_{2M} z},$$

and that c_n are all nonzero. (It is said to be **normal** c.f. in this case.) Then the next Pade' approximant is the same, except that a new term is added at the end:

$$P_{M+1}^M(z) = \frac{c_0}{1+} \frac{c_1 z}{1+} \frac{c_2 z}{1+} \cdots \frac{c_{2M-1} z}{1+} \frac{c_{2M} z}{1+ c_{2M+1} z}$$

Again,

$$P_{M+1}^{M+1}(z) = \frac{c_0}{1+} \frac{c_1 z}{1+} \frac{c_2 z}{1+} \cdots \frac{c_{2M-1} z}{1+} \frac{c_{2M} z}{1+} \frac{c_{2M+1} z}{1+ c_{2M+2} z}$$

So if we arrange them as a ste-ladder $P_0^0, P_1^0, P_1^1, P_2^1, P_2^2, \dots$ only one new coefficient needs to be computed at each step (as long as it is non-zero, so that the c.f. is normal). That is, they are the even and odd terms of the partial sums of the c.f. So

$$P_M^M(z) = \frac{A_{2M}(z)}{B_{2M}(z)}, \quad P_{M+1}^M(z) = \frac{A_{2M+1}(z)}{B_{2M+1}(z)}$$

satisfy the recursion

$$A_n(z) = A_{n-1}(z) + c_n z A_{n-2}(z), \quad B_n(z) = B_{n-1}(z) + c_n z B_{n-2}(z)$$

$$A_0 = c_0, \quad A_{-1} = 0, \quad B_0 = 1, \quad B_{-1} = 1$$

This can be proved inductively as before (See Bender-Orszag). In most situations, the hard work is in calculating the coefficients of a formal power series (e.g., by Laplace's method on multiple or infinite dimensional integrals). Converting it into Pade' fractions is relatively easy, as it is a problem in a single variable. So we will not elaborate on this point.

Eigenvalues of a Tridiagonal Matrix

Finding the eigenvalue of a square matrix (more generally, singular values of a rectangular matrix) is a problem that arises in every kind of applied mathematics, from quantum mechanics to graph theory. A particular case is a tridiagonal matrix, for which only the diagonal terms and the next-to-diagonal terms are non-zero. Continued fractions can be used to solve these problems. Moreover, many matrices can be reduced to tridiagonal form efficiently by a succession of linear transformations.

7.1. 2×2 Matrix

Let us start with the simplest case of a two by two matrix. This is always tridiagonal

$$T = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

The eigenvalues are obtained by solving a quadratic equation

$$\lambda_{\pm} = \frac{(\alpha + \delta) \pm \sqrt{\Delta}}{2}, \quad \Delta = (\alpha - \delta)^2 + 4\gamma\beta$$

The quantity Δ is the **discriminant** of the matrix. If it is zero, the eigenvalues coincide. In fact

$$\Delta = (\lambda_+ - \lambda_-)^2.$$

EXERCISE 19. Prove that a hermitean 2×2 matrix has coincident eigenvalues only if it is a multiple of the identity.

Solution: $\gamma = \beta^*, \alpha = \alpha^*, \delta = \delta^* \implies \Delta = |\alpha - \delta|^2 + 4|\beta|^2$. Thus if $\Delta = 0$, we get $\alpha = \delta, \beta = 0$.

Since the solution only involves a square root, we can find a continued fraction approximation for it. There is no loss of generality in choosing $\alpha = -\delta$, as it only involves adding a multiple of the identity to the matrix, shifting both eigenvalues by the same amount. Thus the characteristic equation

$$\det \begin{pmatrix} \alpha - \lambda & \beta \\ \gamma & -\alpha - \lambda \end{pmatrix} = 0$$

which can be turned into

$$\lambda = \alpha + \frac{\gamma\beta}{\alpha + \lambda}$$

giving the continued fraction for one of the eigenvalues:

$$\lambda_+ = \alpha + \frac{\gamma\beta}{\alpha} \frac{\gamma\beta}{\alpha} \dots$$

Since the trace of the matrix is zero, the other eigenvalue is just the negative of this one.

EXAMPLE 20. The matrix $\begin{pmatrix} 2z & -i \\ i & 0 \end{pmatrix}$ has eigenvalues $z \pm \sqrt{z^2 + 1}$, for which we constructed the rational approximations in the last chapter.

7.2. The Spectrum of a Tridiagonal Matrix

Consider the tridiagonal matrix

$$T = \begin{pmatrix} \alpha_0 & \beta_1 & 0 & 0 & \cdot \\ \gamma_1 & \alpha_1 & \beta_2 & 0 & \cdot \\ 0 & \gamma_2 & \alpha_2 & \alpha_3 & \cdot \\ 0 & 0 & \gamma_3 & \alpha_3 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

Let $A_k = \det T_k$ be the determinant of the $(k+1) \times (k+1)$ submatrix T_k of T starting with α_0 and ending with α_k . Thus

$$A_0 = \alpha_0, \quad A_1 = \alpha_0\alpha_1 - \beta_1\gamma_1$$

Applying Kramer's rule to the last row and column, we get

$$A_k = \alpha_k A_{k-1} - \beta_k \gamma_k A_{k-2}.$$

To agree with the above special case, we impose initial conditions

$$A_0 = \alpha_0, \quad A_{-1} = 1.$$

Now we see the close relation between tridiagonal matrices and continued fractions: these are exactly the numerators of the c.f.

$$F_k = \alpha_0 + \frac{-\beta_1\gamma_1}{\alpha_1 +} \frac{-\beta_2\gamma_2}{\alpha_2 +} \dots \frac{-\beta_k\gamma_k}{\alpha_k} = \frac{A_k}{B_k}$$

Just compare with the earlier section with $b_k = \alpha_0, a_k = -\beta_k\gamma_k$. The denominators satisfy the same recursion but with the initial conditions

$$B_0 = 1, \quad B_{-1} = 0$$

Another fraction with the same numerator is useful in finding the eigenvalues of T . By replacing $\alpha_k \rightarrow \alpha_k - \lambda$ we find that the characteristic function of the submatrix

$$A_k(\lambda) = \det[T_k - \lambda]$$

satisfies

$$A_k(\lambda) = [\alpha_k - \lambda]A_{k-1}(\lambda) - \beta_k\gamma_k A_{k-2}.$$

$$A_0(\lambda) = \alpha_0 - \lambda, \quad A_{-1} = 1.$$

Then the ratio

$$R_k(\lambda) = \frac{A_k(\lambda)}{A_{k-1}(\lambda)}$$

satisfies

$$R_k(\lambda) = (\alpha_k - \lambda) - \frac{\beta_k \gamma_k}{R_{k-1}(\lambda)}$$

which gives a kind of “transpose” of the earlier c.f. The rational functions $R_k(\lambda)$ are better to use than the polynomials $A_k(\lambda)$ because they grow only like λ for large λ while $A_k(\lambda) \sim \lambda^k$. Thus for large k , the evaluation of $A_k(\lambda)$ will be prone to numerical overflow even for modest values of λ and k . (e.g., $\lambda \sim 2, k = 100, \lambda^k \sim 10^{30}$).

7.2.1. Numerical Approximation of the Spectrum of a Hermitean Tridiagonal Matrix. In many applications (e.g., quantum mechanics) T is hermitean (i.e., α_k are real and $\beta_k = \gamma_k^*$) so that the eigenvalues of T_k are real. We can find an approximation to the spectrum of T by finding the roots of $R_k(\lambda)$ for successive values of k . Assume that all the β_k are non-zero, as otherwise the problem breaks up into matrices which can be solved separately.

We start with the observation that α_0 is the root of $R_0(\lambda)$. Next we find the roots of the degree two rational function $R_1(\lambda) = (\alpha_1 - \lambda) - \frac{|\beta_1|^2}{\alpha_0 - \lambda}$. For $\lambda = \alpha_0 + \epsilon$, for small positive ϵ , this function has a large positive value; yet as $\lambda \rightarrow \infty$ it has a large negative value. So for some $\lambda > \alpha_0$ it must vanish. Since the function is monotonically decreasing in this range, it is easy to find this root by bisecting the interval.

EXAMPLE 21. Suppose $f : [a, b] \rightarrow R$ has positive derivative and $f(a) < 0$ and $f(b) > 0$. So there is a unique solution for the equation $f(x) = 0$ in this interval. To find it within some tolerance ϵ , we use the iteration

$$\text{If } f\left(\frac{a_k + b_k}{2}\right) > 0, \text{ set } b_{k+1} = \frac{a_k + b_k}{2} \text{ else } a_{k+1} = \frac{a_k + b_k}{2}.$$

starting with

$$a_0 = a, \quad b_0 = b$$

and stopping when $|a_k - b_k| < \epsilon$.

Similarly, for a root $< \alpha_0$ as well.

More generally, in between two roots of $R_{k-1}(\lambda)$, there will lie a root of $R_k(\lambda)$; which can be found as above. In addition, there could be a new eigenvalue added at the upper or lower end of the spectrum. Thus, we get a controlled sequence of approximations to the eigenvalues of a tridiagonal matrix.

EXERCISE 22. Write a program to find numerically the eigenvalues of the tridiagonal matrix with $\alpha_n = n + s(-1)^n, \beta_n = \gamma_n = g\sqrt{n}$. Plot the first few eigenvalues as a function of g for the case $s \approx 0$ and then also close to resonance $s \approx \frac{1}{2}$. This is the parity even sector of the celebrated Rabi model of an atom interacting with a laser.

7.2.2. Rational Approximation to Eigenvalues. This is a new idea and how useful it is remains to be seen.

Sometimes it is useful to analytically continue the spectrum to complex values of some parameter on which T depends. In such cases we would like to have an analytic rather than numerical approximation method. The problem is, of course, that polynomials of degree larger than four do not have a solution in terms of radicals. But this does not prevent us from finding roots by the iteration of rational functions. MacMullen has shown that a quintic can be solved by iterating rational functions of one variable; McMullen and Doyle have shown that a sextic can be solved by iterating rational functions of two variables, and so on. So we should be able to find the spectrum of T_k as the iteration of functions such as ϕ_k , which depends on α 's and β 's. For example, we can find the eigenvalues of T_k by iterating the function $\phi_k : C^2 \rightarrow C^2$

$$\phi_k : \{y_1, y_2\} \mapsto \left\{ y_1 - R_k(y_1) \frac{y_1 - y_2}{R_k(y_1) - R_k(y_2)}, y_1 \right\}$$

We choose the initial condition close to an eigenvalue of T_{k-1} (which we assume have already been found). Once the desired accuracy has been achieved ($|y_1 - y_2|$ is small enough) we move to the next interval eigenvalue of T_{k-1} and so on. This iteration of ϕ_k amounts to finding roots by the section method, a variant of Newton-Raphson. It should converge because the eigenvalues of T_{k-1} are close to those of T_k .

As the memory available for symbolic computation grows exponentially with time, more interesting problems will become accessible by this method.

7.3. The Lanczos Algorithm Makes Any Matrix Tridiagonal

Apply Gram-Schmidt orthogonalization to the sequence of vectors $A^k|\psi_0\rangle$ for any starting vector: it produces a basis (for the **Krylov subspace**, which is the subspace spanned by this sequence) in which A is tridiagonal. More precisely,

Start with any state $|\psi_0\rangle$ of unit norm and let $b_0 = 0$. Define recursively

$$|\psi_{n+1}\rangle = A|\psi_n\rangle - \alpha_n|\psi_n\rangle - \beta_n|\psi_{n-1}\rangle, \quad n = 0, 1, 2, \dots$$

where

$$\alpha_n = \frac{\langle \psi_n | A | \psi_n \rangle}{\langle \psi_n | \psi_n \rangle}, \quad \beta_n = \frac{\langle \psi_{n-1} | A | \psi_n \rangle}{\langle \psi_{n-1} | \psi_{n-1} \rangle}$$

We continue this recursion until $|\psi_{K+1}\rangle = 0$ for some K ; or continue indefinitely, in which case we set $K = \infty$. Then it can be proved by induction that $|\psi_n\rangle, n = 0, 1, \dots, K$ is an orthogonal sequence. But they are not necessarily of length one. In the subspace spanned by the $|\psi_n\rangle$ for $n = 0, 1, \dots, K$ the operator A is represented by a tridiagonal matrix:

$$\frac{\langle \psi_m | A | \psi_n \rangle}{\langle \psi_m | \psi_m \rangle} = \begin{pmatrix} \alpha_0 & \beta_1 & 0 & 0 & \cdot \\ 1 & \alpha_1 & \beta_2 & 0 & \cdot \\ 0 & 1 & \alpha_2 & \beta_3 & \cdot \\ 0 & 0 & 1 & \alpha_3 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

Now we apply the earlier methods for tridiagonal matrix. This method works best when the eigenvalues of A have a large gap (so that the vectors $A^k|\psi\rangle$ converges to the ground state quickly) and when A is a sparse matrix (so that it is easy to calculate $H|\psi\rangle$ repeatedly). There could be an alternative to perturbation theory in terms of continued fractions.

7.4. Appendix: The Discriminant of a Matrix

This section is put in for completeness and is not required for the rest of the course.

Let $f(\lambda)$ and $g(\lambda)$ be two polynomials in some complex variable λ of degrees m, n respectively. What is the condition that they have a common zero (equivalently, a common factor h)? In this case there would exist polynomials a, b of degree $n-1$ and $m-1$ such that

$$a(\lambda)f(\lambda) + b(\lambda)g(\lambda) = 0$$

(simply set $a = \frac{g}{h}, b = -\frac{f}{h}$). This equation can be recast as a homogenous linear equation for the coefficients of a, b ,

$$S \begin{pmatrix} a \\ b \end{pmatrix} = 0$$

where S is an $(m+n) \times (m+n)$ matrix

$$S = \begin{pmatrix} f_m & f_{m-1} & \cdots & f_0 & 0 & \cdots & 0 \\ 0 & f_m & \cdots & f_1 & f_0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \cdot & f_1 & f_0 \\ g_n & g_{n-1} & \cdot & \cdot & g_0 & 0 & \cdots \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & g_1 & g_0 \end{pmatrix}$$

The determinant of this **Sylvester** matrix is a polynomial of degree mn in the coefficients of f, g , called the **resultant**. Its vanishing is the condition for the two polynomials to have a common root. Of particular interest is the case where $g = f'$. The resultant of f and its derivative is called its **discriminant**. The discriminant vanishes precisely when two roots of f coincide.

The **discriminant of a matrix** is the discriminant of its characteristic polynomial. It is a polynomial of degree $n(n-1)$ in its matrix elements. Its vanishing is the condition that one of its characteristic values is degenerate. The discriminant of a real symmetric matrix is non-negative. A moment's thought shows that (up to a constant multiple) it is equal to

$$\Delta(A) \propto \prod_{i < j} (\lambda_i - \lambda_j)^2$$

in terms of the eigenvalues.

7.4.1. van der Monde Determinant. There is another way to think of the discriminant as a determinant. Note first that

$$\prod_{i < j} (\lambda_i - \lambda_j) \propto \det \begin{pmatrix} 1 & \lambda_1 & \lambda_1^2 & \cdots & \lambda_1^{N-1} \\ 1 & \lambda_2 & \lambda_2^2 & \cdots & \lambda_2^{N-1} \\ 1 & \lambda_3 & \lambda_3^2 & \cdots & \lambda_3^{N-1} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 1 & \lambda_N & \lambda_N^2 & \cdots & \lambda_N^{N-1} \end{pmatrix}$$

The point is that if any pair of λ 's are equal, two columns will become identical and the determinant vanishes. Moreover, by counting powers we see that the r.h.s. is a homogenous polynomial of degree $1 + 2 + \cdots + N - 1 = \frac{N(N-1)}{2}$ in the λ 's. So it cannot have any other zeros. (There is a constant multiple that is undetermined). If we now take the matrix on the rhs and multiply with its transpose we get

$$\Delta(A) \propto \det \begin{pmatrix} 1 & G_1 & G_2 & \cdots & G_{N-1} \\ G_1 & G_2 & G_3 & \cdots & G_N \\ G_2 & G_3 & G_4 & \cdots & G_{N+2} \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ G_{N-1} & G_N & G_{N+1} & \cdots & G_{2N} \end{pmatrix}$$

where the **moments** of A are defined by

$$G_n = \frac{1}{N} \sum_k \lambda_k^n = \frac{1}{N} \text{tr } A^n$$

REMARK 23. It would be interesting to derive a recursion relation for the discriminants of tridiagonal matrices, analogous to that for the determinants.

REMARK 24. The discriminant of a hermitean matrix is positive. It can be written as the sum of squares of polynomials of degree $\binom{n}{2}$. It is not known how many such terms are needed, but $n!$ is sufficient. There are interesting papers by P.D. Lax and by N. V. Ilyushechkin on this topic. Each of these terms would have to vanish separately if a hermitean matrix is to have degeneracies. An example of such a quantity is the determinant of the linear transformation $u \rightarrow [A, u]$ on the space of anti-symmetric matrices.

CHAPTER 8

Sturm-Liouville Problems

It is worthwhile to solve eigenvalue problems in one dimension, as they are models for more complicated situations. To be specific, let us consider the problem of solving the differential equation

$$(8.0.1) \quad \psi'' + [\lambda - v(x)]\psi = 0$$

with the **Dirichlet**

$$\psi(a) = 0 = \psi(b)$$

or **Neumann** boundary conditions

$$\psi'(a) = 0 = \psi'(b)$$

More generally we have some linear combinations

$$\alpha_1\psi(a) + \alpha_2\psi'(a) = 0$$

$$\beta_1\psi(b) + \beta_2\psi'(b) = 0$$

These are called **Sturm-Liouville problems**; many interesting problems can be reduced to this form. They are the continuous analogues of tridiagonal eigenvalue problems.

8.1. Variational principle

The SL equation is the condition for an extremum of $\int_a^b [\psi'^2 + v(x)\psi^2] dx$ subject to the condition that $\int_a^b \psi^2(x) dx = 1$; in fact, λ is the Lagrange multiplier enforcing this constraint. Equivalently we can extremize

$$\frac{\int [\psi'^2 + v(x)\psi^2(x)] dx}{\int \psi^2(x) dx}$$

This fact, first studied in acoustics, is called the **Rayleigh-Ritz variational principle**.

8.2. Picard Iteration

There is a unique solution to the ODE (8.0.1) with the initial condition $\psi'(a) = -\alpha_1, \psi(a) = \alpha_2$ and any complex value of λ .

$$\frac{d}{dx} \begin{pmatrix} \psi(x) \\ \psi'(x) \end{pmatrix} = Q(\lambda, x) \begin{pmatrix} \psi(x) \\ \psi'(x) \end{pmatrix}, \quad Q(\lambda, x) = \begin{pmatrix} 0 & 1 \\ v(x) - \lambda & 0 \end{pmatrix}.$$

We can write this as an integral equation (with Neumann b.c.. The others are similar)

$$\begin{pmatrix} \psi(x) \\ \psi'(x) \end{pmatrix} = \begin{pmatrix} \alpha_2 \\ -\alpha_1 \end{pmatrix} + \int_a^x Q(\lambda, x_1) \begin{pmatrix} \psi(x_1) \\ \psi'(x_1) \end{pmatrix} dx_1$$

By iterating this, the solution can be expressed as an infinite series (called the **Dyson series** by physicists, or **Picard iteration** by mathematicians)

$$\begin{pmatrix} \psi(\lambda, x) \\ \psi'(\lambda, x) \end{pmatrix} = \sum_{r=1}^{\infty} \int_{x > x_r > x_{r-1} \cdots x_1 > b} dx_1 \cdots dx_r Q(\lambda, x_r) \cdots Q(\lambda, x_1) \begin{pmatrix} \alpha_2 \\ -\alpha_1 \end{pmatrix}$$

This series converges as long as $\nu(x)$ is continuous: the integral is on a bounded region. In particular, it is an analytic function of λ .

8.3. The Characteristic Function

Thus evaluating it at the other boundary point gives an entire function of λ :

$$\Delta(\lambda) = \beta_1 \psi(\lambda, b) + \beta_2 \psi'(\lambda, b)$$

The zeros of this function are the eigenvalues of (8.0.1). Being the zeros of an entire function, they are isolated points although some of them might be multiple zeros of finite order. Thus $\Delta(\lambda)$ can be thought of as proportional to the **characteristic function** of the operator $\frac{d^2}{dx^2} + \lambda - v(x)$.

$$\det \left[\frac{d^2}{dx^2} + \lambda - v \right] = C \Delta(\lambda)$$

Here C might depend on the boundary conditions, but not on λ .

EXAMPLE 25. The most elementary example is $v(x) = 10, a = 0, b = \pi$ with Dirichlet boundary conditions

$$\psi'' = \lambda \psi, \quad \psi(0) = \psi(\pi) = 0$$

The solution is

$$\psi(x) = \frac{\sin \sqrt{\lambda} x}{\sqrt{\lambda}},$$

Despite the square root, this is an entire function of λ . Thus,

$$\Delta(\lambda) = \frac{\sin \sqrt{\lambda} \pi}{\sqrt{\lambda}}$$

which has roots at

$$\lambda = n^2.$$

To understand this better, recall the product formula

$$\sin z = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2\pi^2}\right)$$

so that

$$\Delta(\lambda) = \pi \prod_{n=1}^{\infty} \left(1 - \frac{\lambda}{n^2}\right)$$

In order to make sense of

$$\det \left[\frac{d^2}{dx^2} + \lambda \right] = \prod_{n=1}^{\infty} (n^2 - \lambda) \equiv C\Delta(\lambda)$$

we have to make sense of the infinite constant

$$C = \frac{1}{\pi} \prod_{n=1}^{\infty} n^2$$

Recall the Riemann zeta function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

so that

$$\zeta'(s) = - \sum_{n=1}^{\infty} \frac{\log n}{n^s}$$

Although the series converges only for $\text{Re } s > 1$, we can make sense of it by analytic continuation. The only singularity of $\zeta(s)$ is a simple pole at $s = 1$. In particular it is regular at $s = 0$. Thus, $\zeta'(0)$ makes perfect sense:

$$\zeta'(0) = -\frac{1}{2} \log(2\pi)$$

In this sense,

$$\sum_{n=1}^{\infty} \log n \equiv -\zeta'(0) = \frac{1}{2} \log(2\pi)$$

and

$$\prod_{n=1}^{\infty} n^2 \equiv e^{-2\zeta(0)} = 2\pi.$$

So we conclude that

$$\det \left[\frac{d^2}{dx^2} + \lambda \right] = 2 \frac{\sin \sqrt{\lambda}\pi}{\sqrt{\lambda}}$$

Of course, this answer is specific to the boundary conditions chosen.

EXERCISE 26. Find $\det \left[\frac{d^2}{dx^2} + \lambda \right]$ on the interval $[0, \pi]$ subject to the mixed boundary condition $\psi(0) = \psi'(\pi) = 0$. It is useful to know that $\zeta(s, a) = \sum_{n=1}^{\infty} \frac{1}{(n+a)^s}$ has the expansion $\zeta(s, \frac{1}{2}) \sim \left[-\frac{1}{2} \log 2 \right] s + O(s^2)$.

8.4. The Ricatti Equation

It is possible to rewrite the SL equation as a nonlinear, but first order differential equation. Let us consider Neumann b.c., so that $\psi(a) \neq 0$.

$$R = \frac{\psi'}{\psi}, \quad \psi(x) = \psi(a) e^{\int_a^x R(x_1) dx_1}$$

Then we get the **Ricatti equation**

$$R' = v(x) - \lambda - R^2$$

$$[R + \sqrt{v(x) - \lambda}][R - \sqrt{v(x) - \lambda}] = -R'$$

Thus

$$R = \sqrt{v(x) - \lambda} + \frac{-R'}{R + \sqrt{v(x) - \lambda}}$$

It is interesting to compare with the WKB method, which is an asymptotic power series in $\frac{1}{\lambda}$.

EXAMPLE 27. The Airy function vanishes as $x \rightarrow \infty$ and satisfies the differential equation

$$\text{Ai}'' - x\text{Ai} = 0$$

Thus

$$R(x) = \frac{\text{Ai}'(x)}{\text{Ai}(x)}$$

satisfies

$$R(x) = -\sqrt{x} + \frac{-R'}{R - \sqrt{x}}.$$

We take the negative root as we expect Ai to decrease for positive x . We can turn this into a recursion relation to get a sort of continued fraction in \sqrt{x} (i.e., rational functions of \sqrt{x} rather than x):

$$R_0 = -\sqrt{x}, \quad R_n = -\sqrt{x} + \frac{-R'_{n-1}}{R_{n-1} - \sqrt{x}}$$

This works well for large x , but has spurious singularities for small x .

8.5. The Resolvent (to be Completed)

A typical problem in physics is to find the spectrum of some operator, for example the hamiltonian of some quantum system. This information is contained in its resolvent

$$R(\lambda) = \frac{1}{\lambda - A}$$

If A is a matrix, this is a rational function of λ . That is, each matrix element of $R(\lambda)$ is a rational function of λ . Explicitly,

$$R_{ij}(\lambda) = \frac{(-1)^{i+j} \det[\lambda - A]^{ij}}{\det[\lambda - A]}$$

where $[\lambda - A]^{ij}$ is the matrix obtained by omitting the i th row and j th column. If the characteristic polynomial has no double roots, we can resolve into partial fractions

$$R(\lambda) = \sum_n \frac{1}{\lambda - a_n} P_n$$

The poles are at the characteristic values of A . The operators P_n are the projections to the eigenspaces of each eigenvalue.

CHAPTER 9

The Method of Steepest Descent

Consider again

$$\int_a^b e^{\frac{1}{g}S(\phi)} d\phi$$

where $S(\phi)$ is complex valued. Suppose we can find a contour in the complex so that $S(\phi)$ has constant imaginary part. Then that factor can be pulled out of the integral and it will reduce to a Laplace-type integral we have already studied: it will be dominated by the maxima of the real part of $S(\phi)$ along this contour. Now, along this curve where the imaginary part of $S(\phi)$ is constant, this real part will change most rapidly. So this is called the path of **steepest descent**.

Being an analytic function, wherever the real part has vanishing derivative, the imaginary will also have an extremum. So we must find the points in the complex ϕ -plane where $S'(\phi) = 0$; then deform our contour of integration so that it passes through these points along curves with constant imaginary part for $S(\phi)$. Applying the Laplace method to this deformed integral will give us an asymptotic expansion for the integral.

It is best to look at examples.

EXAMPLE 28. The Fresnel integral $\int_{-\infty}^{\infty} e^{iz^2} dz = \sqrt{\pi} e^{i\frac{\pi}{4}}$

The integral on the lhs is not absolutely convergent: the integrand has magnitude one. It can be given a meaning if the contour of integral is deformed a bit so that $\text{Im } z^2 > 0$ as $|z| \rightarrow \infty$. That is, at infinity the contour must lie within two wedges

$$I : 0 < \arg z < \frac{\pi}{2}$$

$$II : \pi < \arg z < \frac{3\pi}{2}.$$

So if we lower the contour a bit below the negative real axis and a bit above the positive real axis, the integral will converge. The stationary point of the exponent is at $z = 0$.

$$\int_{C_1} e^{iz^2} dz = \int_{-\infty}^0 e^{i(x-ia)^2} dx + \int_{-ia}^{ia} e^{iz^2} dz + \int_0^{\infty} e^{i(x+ia)^2} dx, \quad a > 0$$

The rhs is independent of a as long as it is positive.

The curve passing through this saddle point, on which the imaginary part of the exponent is constant is given

$$\text{Im } iz^2 = 0, \quad x^2 - y^2 = 0, \quad z = x + iy$$

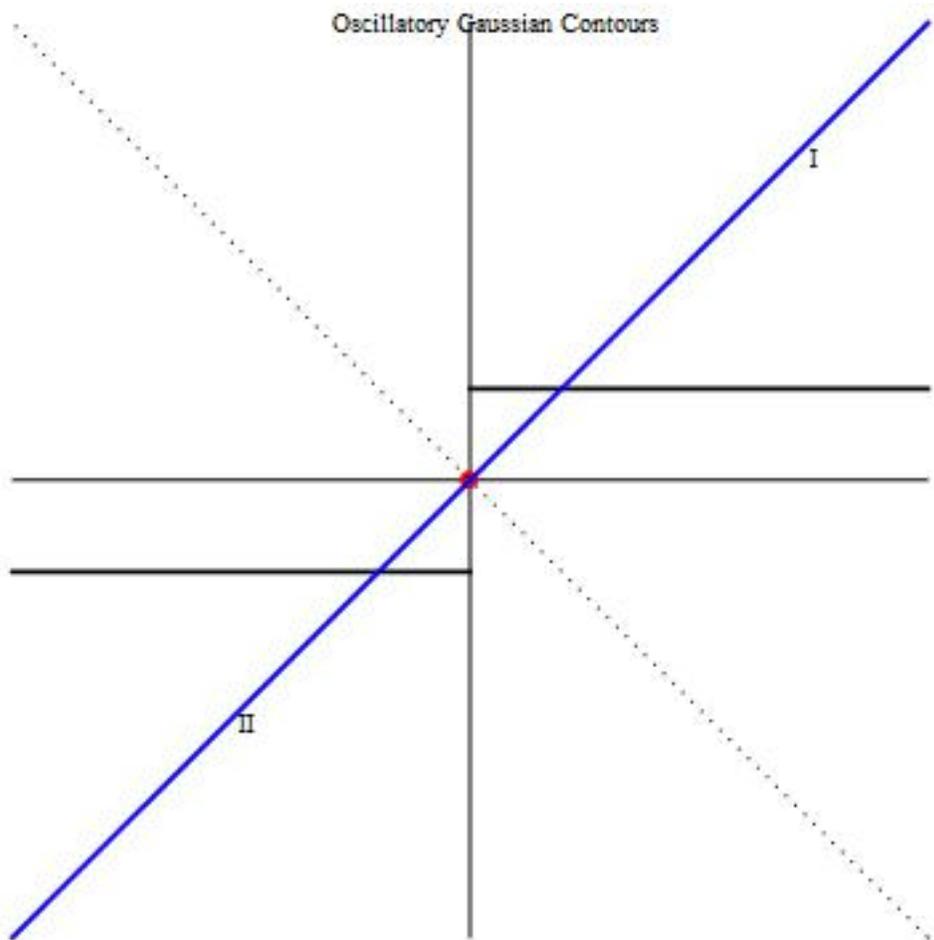


FIGURE 9.0.1.

This is a pair of straightlines. One of these connects the region II to region I .

$$C_2 : x = y$$

So we can deform our contour into this one:

$$z = (1 + i)x, \quad -\infty < x < \infty$$

$$\int_{C_1} e^{iz^2} dz = \int_{C_2} e^{iz^2} dz = (1 + i) \int_{-\infty}^{\infty} e^{-2x^2} dx$$

since

$$iz^2 = -2x^2, \quad dz = (1 + i)dx$$

We can easily evaluate the Gaussian integral

$$\int_{-\infty}^{\infty} e^{-2x^2} dx = \sqrt{\frac{\pi}{2}}$$

In summary

$$\int_{-\infty}^{\infty} e^{iz^2} dz = \sqrt{\pi} \frac{1+i}{\sqrt{2}} = \sqrt{\pi} e^{i\frac{\pi}{4}}.$$

See M. Born and E. Wolfe *Principles of Optics* for the optical applications of this integrals.

9.1. The Airy Function

Airy's differential equation is

$$\psi'' = x\psi$$

The Fourier transform

$$\tilde{\psi}(k) = \int \psi(x) e^{-ikx} dx$$

satisfies the first order equation

$$k^2 \tilde{\psi} = -i \frac{d}{dk} \tilde{\psi}$$

Thus a solution is

$$\psi(x) = \int e^{i\frac{1}{3}k^3 + ikx} \frac{dk}{2\pi}$$

More generally,

$$\int_C e^{i\frac{1}{3}k^3 + ikx} \frac{dk}{2\pi}$$

over some contour C in the complex k -plane (for which it converges) is a solution. To converge, the contour C must tend to infinity along an angle θ such that $\text{Im } e^{3i\theta} > 0$. That is, along one of the three wedges,

$$I : 0 < \theta < \frac{\pi}{3}, \quad II : \frac{2\pi}{3} < \theta < \pi, \quad III : \frac{4\pi}{3} < \theta < \frac{5\pi}{3}$$

We define the Airy function by the integral that lies asymptotically within the first two wedges, above the positive real axis:

$$(9.1.1) \quad \text{Ai}(x) = \int_{-\infty+ia}^{\infty+ia} e^{i[kx + \frac{1}{3}k^3]} \frac{dk}{2\pi}, \quad a > 0$$

Because the integrand has no singularities, the integral does not depend on a (as long as it is positive). A little sloppily, this can also be written as,

$$\text{Ai}(x) = \int_0^{\infty} \cos[kx + \frac{1}{3}k^3] \frac{dk}{\pi}.$$

The extrema of

$$S[k, x] = i[kx + \frac{1}{3}k^3]$$

are at

$$k = \pm\sqrt{-x}.$$

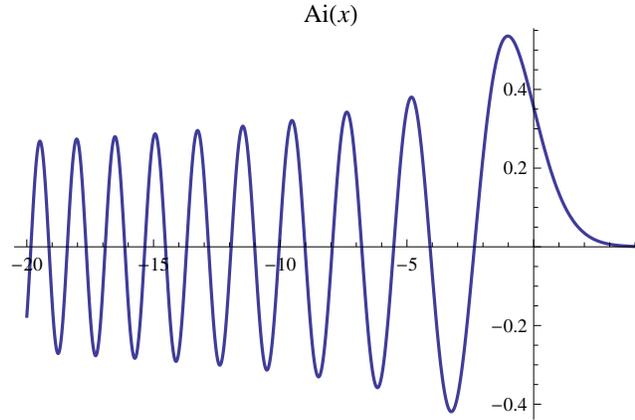


FIGURE 9.1.1.

9.1.1. When x is positive. Deform the contour of integration (9.1.1) so that it passes through one or more of these points along curves of constant imaginary part for $S(\phi)$.

Set

$$k = k_1 + ik_2$$

so that (assuming x is real)

$$\operatorname{Re}(S) = -\frac{1}{3}k_2(3k_1^2 - k_2^2 + 3x)$$

$$\operatorname{Im}(S) = \left[\frac{k_1^2}{3} - k_2^2 + x \right] k_1$$

If $x > 0$, the saddle points are on the imaginary axis and $\operatorname{Im}(S) = 0$ there. So the curve of constant $\operatorname{Im}(S)$ passing through the saddles degenerates into the union of the imaginary axis

$$k_1 = 0$$

and the hyperbola

$$\frac{k_1^2}{3} - k_2^2 + x = 0.$$

Thus we deform the original contour C_1 to this hyperbola C_2 .

We can eliminate k_2 in favor of k_1 along the contour C_2 so that

$$S = -\frac{2}{3}x^{\frac{3}{2}}[1 + \phi^2]^{\frac{3}{2}}, \quad \phi = \frac{1}{\sqrt{3x}}k_1$$

$$Ai(x) = \frac{\sqrt{3x}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{2}{3}x^{\frac{3}{2}}[1+\phi^2]^{\frac{3}{2}}} d\phi$$

We can apply the Laplace method to this integral by expanding around the maximum at $\phi = 0$.

The contribution of this saddle point is

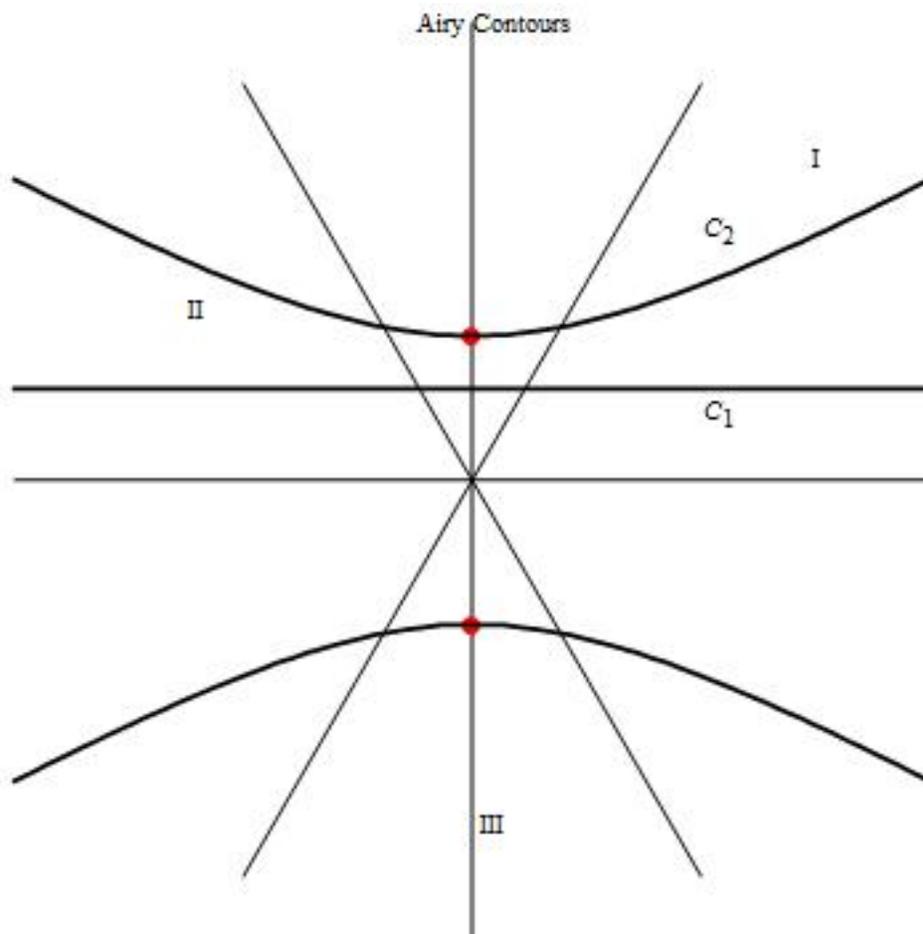


FIGURE 9.1.2.

$$\text{Ai}(x) \sim \frac{1}{2\sqrt{\pi x^{\frac{1}{4}}}} e^{-\frac{2}{3}x^{\frac{3}{2}}}$$

Note that the integral along the curve of constant phase through the other saddle point at $-i\sqrt{x}$ diverges: its endpoints are outside the wedge of convergence.

9.1.2. When x is negative. Now the saddle points are on the real axis $k = \pm\sqrt{|x|}$. The imaginary parts of the exponent at these points are $\mp\frac{2}{3}|x|^{\frac{3}{2}}$. The curves on which the imaginary part is constant are the curves

$$\frac{1}{3}k_1(k_1^2 - 3k_2^2 - 3|x|) = \mp\frac{2}{3}|x|^{\frac{3}{2}}$$

There is a branch C_{2+} of these curves that pass connect the asymptotic regions *I* and *III*

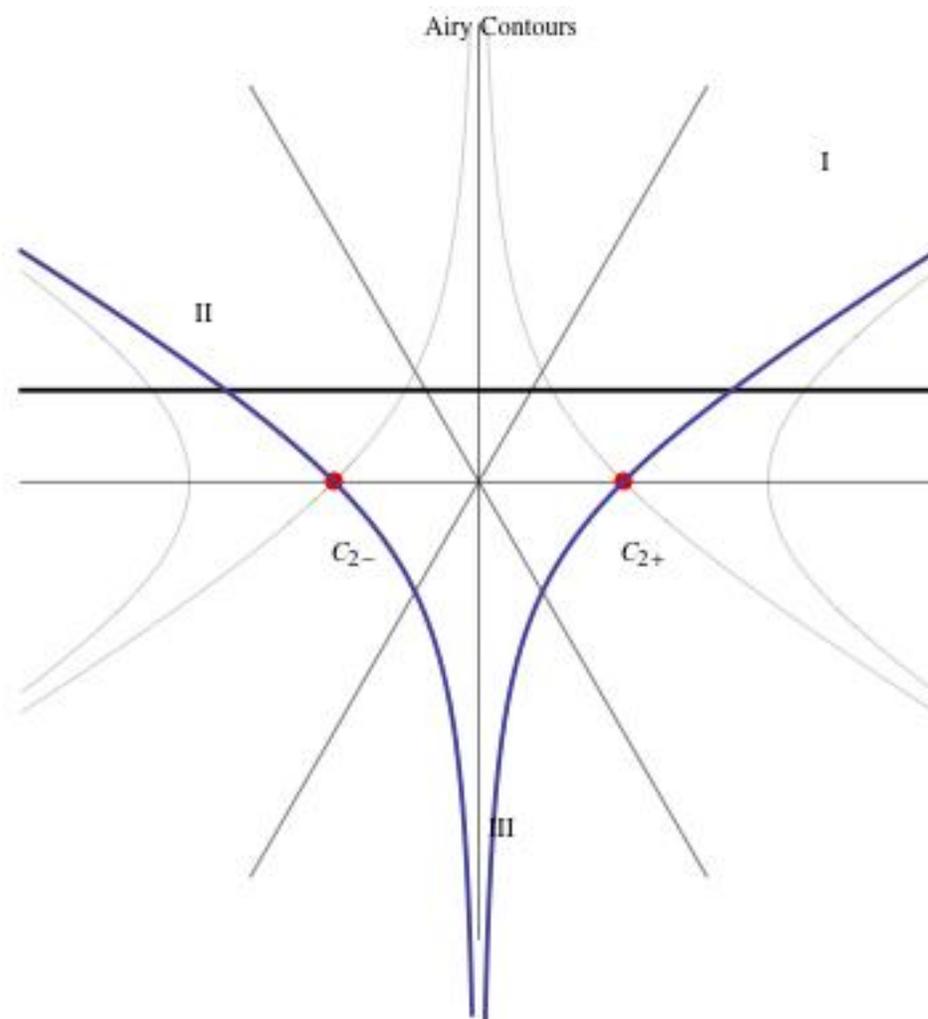


FIGURE 9.1.3.

$$k_2 = \left\{ \begin{array}{ll} -\sqrt{\frac{1}{3}k_1^2 - |x| + \frac{2}{3k_1}|x|^{\frac{3}{2}}} & \text{for } 0 < k_1 < \sqrt{|x|} \\ \sqrt{\frac{1}{3}k_1^2 - |x| + \frac{2}{3k_1}|x|^{\frac{3}{2}}} & \text{for } k_1 > \sqrt{|x|} \end{array} \right\}$$

and its mirror image C_{2-} that connect regions *II* and *III*. Together they form a contour C_2 of constant phase that can be deformed into C_1 .

The real part of the exponent along C_{2+} is

$$-\frac{1}{3}k_2(3k_1^2 - k_2^2 + 3x) = -\frac{2}{9\sqrt{3}}|x|^{\frac{3}{2}}(1 - \phi)^2(1 + 2\phi)^2\phi^{-\frac{3}{2}}\sqrt{2 + \phi}, \quad k_1 = \sqrt{|x|\phi},$$

Thus we get

$$\text{Ai}(x) = \frac{1}{2\pi} e^{-i\frac{2}{3}|x|^{\frac{3}{2}}} \int_0^\infty e^{-\frac{2}{9\sqrt{3}}|x|^{\frac{3}{2}}(1-\phi)^2(1+2\phi)^2\phi^{-\frac{3}{2}}\sqrt{2+\phi}} \left[\frac{dk_1}{d\phi} + i\frac{dk_2}{d\phi} \right] d\phi + c.c.$$

Again, the integral can be evaluated by the Laplace method, by expanding around the maximum at $\phi = 1$. An extra phase $e^{i\frac{\pi}{4}}$ comes from the change in the measure $dk = \left[\frac{dk_1}{d\phi} + i\frac{dk_2}{d\phi} \right] d\phi$: note that the angle that the curve C_2 makes at the saddle point $\sqrt{|x|}$ is $\frac{\pi}{4}$.

$$\text{Ai}(x) \sim \frac{1}{\sqrt{\pi}} |x|^{-\frac{1}{4}} \sin \left[\frac{2}{3}|x|^{\frac{3}{2}} + \frac{\pi}{4} \right], \quad x \rightarrow -\infty$$

For more on these matters see *Asymptotic Expansions* by A. Erdelyi. (Dover 1956).

CHAPTER 10

The WKB Method

Wave equations are at the heart of many problems of physics. In optics,

$$(10.0.2) \quad \nabla^2 \psi + k^2 \nu^2(x) \psi = 0$$

where $\nu(x)$ is the refractive index. In quantum mechanics,

$$\nabla^2 \psi + [\lambda - v(x)] \psi = 0$$

where $v(x)$ is the potential. In both cases $k = \sqrt{\lambda}$ is the wavenumber, inversely proportional to the wavelength. In the limit of $k \rightarrow \infty$ (small wavelength) we have a simplification: the wave moves mostly along a well-defined path, instead of spreading out. Interference effects become small as well. This path is determined as the solution of a system of ordinary differential equations, Hamilton's equations. It is possible to go further and solve the wave equation in a power series in $\frac{1}{k}$. This is an asymptotic series, which is usually not convergent.

This method has its origins in the nineteenth century, when it was applied to sound and light waves. (Airy, Stokes, Rayleigh..) It had a resurgence in the 1920s when W, K and B (who were fortunate enough to be graduate students just when the Schrodinger equation was discovered) applied it to quantum mechanics. Kramer went on to apply it to solve many problems in chemical physics. Brillouin later did seminal work on scattering by crystals.

We can bring the equation being studied in the convenient form

$$\nabla^2 \psi = k^2 u^2(x) \psi$$

10.1. The Eikonal

The difficulty with letting $k \rightarrow \infty$ directly is that it is a singular limit. Consider the simplest case when $u(x) = -1$. The solution

$$e^{\pm ikx}$$

oscillates wildly as $k \rightarrow \infty$. This suggests that we seek a solution of the form

$$\psi(x) = e^{kS(x)}$$

Although ψ itself will not have a good limit as $k \rightarrow \infty$, perhaps $S(x)$ behaves better. Indeed,

$$\frac{1}{k} \nabla^2 S + (\nabla S)^2 = u^2$$

allowing for a solution in power series

$$S = \sum_{r=0} k^{-r} S_r$$

In leading we get the **eikonal equation**

$$(\nabla S_0)^2 = u^2$$

In the next order

$$\nabla S_0 \nabla S_1 + \frac{1}{2} \nabla^2 S_0 = 0$$

In higher orders,

$$\nabla S_0 \cdot \nabla S_r = -\frac{1}{2} \left[\nabla^2 S_{r-1} + \sum_{n=1}^{r-1} \nabla S_n \cdot \nabla S_{r-n} \right]$$

The point is that we are now solving first order rather than second order equations. If the number of independent co-ordinates is one, these equations can be solved by integration

$$S_0(x) = \pm \int_{x_0}^x u(q) dq$$

$$S_r = - \int_{x_0}^x \frac{1}{2u(q)} \left[\sum_{n=1}^{r-1} S'_n(q) S'_{r-n}(q) + S''_{r-1}(q) \right] dq$$

etc.

10.2. The Method of Characteristics

Even when the number of co-ordinates is greater than one, first order equations (even if non-linear) are simpler to solve than second order ones. For more on this see the **second** volume of Courant-Hilbert.

But, instead of just evaluating an integral, we have to now solve a system of ordinary differential equations. Just think of $H = p^2 - u^2$ as the hamiltonian and realize that the eikonal equation is just the Hamilton-Jacobi equation, with $p = \nabla S$. Pass now to Hamilton's equations of mechanics instead.

$$\frac{dq^i}{dt} = 2p_i, \quad \frac{dp_i}{dt} = -\frac{\partial u^2}{\partial q^i}$$

Given some reference point x_0 , there is a unique solution starting there and ending at x

$$q(0) = x_0, \quad q(1) = x$$

Then the action of this trajectory

$$S_0(x) = \int_0^1 p \cdot \dot{q}(t) dt$$

solves the eikonal equation. The higher orders become ODEs as well:

$$\frac{dq^i}{dt} \frac{\partial S_r}{\partial q^i} \equiv \frac{dS_r}{dt} = -\frac{1}{4} \left[\sum_{n=1}^{r-1} \nabla S_n \cdot \nabla S_{r-n} + \nabla^2 S_{r-1} \right]$$

which can also be solved

$$S_r(x) = \frac{1}{4} \int_0^1 \left[\sum_{n=1}^{r-1} \nabla S_n \cdot \nabla S_{r-n} + \nabla^2 S_{r-1} \right] dt$$

where the integral is taken along the path that satisfies Hamilton's equations.

10.3. WKB for Sturm-Liouville Problems

Recall

$$\psi'' + [\lambda - v(x)]\psi = 0, \quad \alpha_1\psi(a) + \alpha_2\psi'(a) = 0 = \beta_1\psi(b) + \beta_2\psi'(b)$$

With $\lambda = k^2$, $u^2(x) = \frac{v(x)}{k^2} - 1$ the Sturm-Liouville problem (8.0.1) becomes

$$\psi'' = k^2 u^2(x) \psi, \quad ()$$

The leading solution can be either of two signs

$$S_0(x) = \int_a^x u(q) dq$$

but the next order

$$S_1 = -\frac{1}{2} \log u$$

and higher order terms

$$S_r = \int_a^x \frac{1}{2u(q)} \left[-\sum_{n=1}^{r-1} S'_n(q) S'_{r-n}(q) + S''_{r-1}(q) \right] dq$$

are independent of this choice of sign. The solution satisfying the b.c. at a is some linear combination.

$$\psi(x) = \frac{1}{\sqrt{u}} \left[A e^{k \int_a^x u(q) dq} + B e^{-k \int_a^x u(q) dq} \right] e^{\sum_{r=2} k^{1-r} S_r}$$

10.3.1. Turning Points. At the **turning points** at which $k^2 = v(x)$ (so that $u = 0$) the above expansion breaks down: the higher order terms get too big. The regions where $k^2 > v(x)$ are very different from those where $k^2 < v(x)$: in the former, S_0 is imaginary and the solution is oscillatory. In the latter, S_0 is real and it is a linear combination of a term that grows exponentially and another that decreases exponentially. In quantum mechanics, $k^2 - v(x) > 0$ is the classically allowed region where the kinetic energy is positive. When $k^2 < v(x)$ the solution decays exponentially: the probability of finding the particle must decrease as we move away from the turning point into the classically forbidden region.

Suppose x_0 is a turning point. Assume for that $v'(x_0) \neq 0$. (This case of **simple turning point** is the typical one.) To be specific that $v'(x_0) > 0$. Then for $x > x_0$ the solution must decrease exponentially:

$$\psi(x) \approx C \frac{1}{\sqrt{\frac{v(x)}{k^2} - 1}} e^{-\int_{x_0}^x \sqrt{v(x) - k^2} dq}, \quad x > x_0$$

In the oscillatory region

$$\psi(x) \approx C_1 \frac{1}{\sqrt{1 - \frac{v(x)}{k^2}}} \sin \left[\int_{x_0}^x \sqrt{k^2 - v(q)} dq + \eta \right] \quad x < x_0$$

for constants C_1, η . To find the a relation between the constants, we have to match these solutions across the turning point x_0 . It turns out that $\eta = \frac{\pi}{4}, C_1 = 2C$.

The WKB approximation breaks down near the turning point, as $u(x_0) = 0$. In the immediately vicinity of a turning point, we can use another approximation: replace the potential by a linear function

$$v(x) \approx v(x_0) + v'(x_0)[x - x_0]$$

The Schrodinger equation for such a potential can be solved exactly in terms of the Airy function. Then we match the solutions in the three regions $x < x_0, x \approx x_0, x > x_0$. See Landau and Lifshitz for more details.

10.3.2. WKB approximation for eigenvalues. In this leading order, we can get an approximate formula for the spectrum of the Sturm-Liouville problem. Just consider the oscillatory regions of the wavefunctions (it is exponentially small in the forbidden regions) and impose the boundary condition on the phases.

$$\int_a^b \sqrt{\lambda - v(x)} dx = \left[n + \frac{1}{2} \mu(\lambda) \right] \pi + \arctan \frac{\alpha_2}{\alpha_1} - \arctan \frac{\beta_2}{\beta_1} \quad n = 0, 1, 2 \dots$$

where $\mu(\lambda)$ is the number of turning points (assumed to be simple) in the interval $[a, b]$ for energy λ .

This idea can be generalized to systems with several degrees of freedom as long the H-J equation is separable: we can reduce it to a collection of one-dimensional problems. Gutzwiller has derived a remarkable trace formula that gives the partition function $\text{tr} e^{-tH}$ as a sum over classical trajectories, in the case of chaotic systems. This is in the early stages of development.

Symbol Calculus

Observables of classical mechanics are functions on phase space. It is possible to think of the observables of quantum mechanics as functions on phase space as well, by translating the usual representation of them as operators in Hilbert space. The induced rule for multiplying functions is more complicated: it is not commutative. This **star product** is still associative, however. This point of view allows a systematic method-an asymptotic expansion- to calculate quantum corrections to classical theories. The function on phase space R^{2n} associated to a differential operator on $L^2(R^n)$ is called its **symbol**. There is an extensive mathematical literature on it, with applications ranging from acoustics to radar.

11.1. Operators and their Kernels

To be specific, consider a classical system whose configuration space is R^n : it has n degrees of freedom represented by real variables. In the Schrodinger picture the wave functions are complex-valued function

$$\psi : R^n \rightarrow C$$

with the inner product

$$\langle \psi, \phi \rangle = \int \psi^*(q)\phi(q) dq$$

A linear operator can be represented in terms of its integral kernel

$$\hat{A}\psi(q) = \int A(q, q')\psi(q') dq'$$

In general $A(q, q')$ is a distribution. In Dirac's notation

$$A(q, q') = \langle q | \hat{A} | q' \rangle$$

EXAMPLE 29. The hamiltonian is usually a differential operator

$$\hat{H} = \frac{1}{2}p^2 + V(q), \quad H(q, q') = -\frac{\hbar^2}{2} \delta''(q - q') + V(q)\delta(q - q')$$

EXAMPLE 30. The heat kernel is the solution of the differential equation

$$\frac{\partial}{\partial t} h_t(q, q') = \frac{\partial^2}{\partial q^2} h_t(q, q')$$

subject to the initial condition

$$\lim_{t \rightarrow 0^+} h_t(q, q') = \delta(q, q').$$

We can think of it as an exponential of the laplacian

$$\hat{h}_t = e^{t\partial^2}$$

The solution of the above differential equation is a gaussian

$$h_t(q, q') = \frac{e^{-\frac{(q-q')^2}{2t}}}{[2\pi t]^{\frac{n}{2}}}$$

A function of position alone (such as potential energy) has an integral kernel that is the continuous analogue of a diagonal matrix

$$V(q)\psi(q) = \int V(q, q')\psi(q')dq', \quad V(q, q') = V(q)\delta(q - q')$$

We can similarly represent a function of momentum alone as a multiplication operator on the Fourier transform of the wave function

$$T\left(-i\hbar\frac{\partial}{\partial q}\right)\psi(q) = \int e^{i\frac{p\cdot q}{\hbar}}\frac{dp}{(2\pi\hbar)^n}T(p)\tilde{\psi}(p), \quad \tilde{\psi}(p) = \int \psi(q)e^{-i\frac{p\cdot q}{\hbar}}dq$$

If we invert the Fourier transform we get

$$T\left(-i\hbar\frac{\partial}{\partial q}\right)\psi(q) = \int T(q, q')\psi(q')dq'$$

with

$$T(q, q') = \int T(p)e^{\frac{1}{\hbar}p\cdot(q-q')}\frac{dp}{(2\pi\hbar)^n}$$

Sometimes it is convenient to choose units such that $\hbar = 1$.

EXAMPLE 31. This is one way to get the above heat kernel

$$\begin{aligned} e^{t\partial^2}\psi(q) &= \int \frac{dp}{(2\pi)^n}e^{-tp^2+ip\cdot q}\tilde{\psi}(p) \\ &= \int dq' \int \frac{dp}{(2\pi)^n}e^{-tp^2+ip\cdot q-ip\cdot q'}\psi(q') \end{aligned}$$

Thus

$$h_t(q, q') = \int \frac{dp}{(2\pi)^n}e^{-tp^2+ip\cdot q-ip\cdot q'}$$

which is a Gaussian integral that can be evaluated by ‘completing the squares’.

EXAMPLE 32. Recall the identity

$$\psi(q + b) = e^{b\frac{\partial}{\partial q}}\psi(q)$$

which is just the Taylor series in other words:

$$\psi(q + b) = \sum_{n=0}^{\infty} b^n \frac{1}{n!} \frac{\partial^n \psi}{\partial q^n}$$

It follows that

$$e^{ib\hat{p}}\psi(q) = \psi(q + \hbar b)$$

The position operator is just a multiplication

$$e^{ia\hat{q}}\psi(q) = e^{iaq}\psi(q)$$

Thus

$$e^{ib\hat{p}}e^{ia\hat{q}}\psi(q) = e^{ia(q+\hbar^{-1}b)}\psi(q + \hbar b)$$

while

$$e^{ia\hat{q}}e^{ib\hat{p}}\psi(q) = e^{iaq}\psi(q + \hbar b)$$

The Weyl relation follows from this:

$$e^{ia\hat{q}}e^{ib\hat{p}} = e^{-i\hbar a \cdot b} e^{ib\hat{p}}e^{ia\hat{q}}$$

Written more symmetrically this is

$$e^{\frac{1}{2}i\hbar a \cdot b} e^{ia\hat{q}}e^{ib\hat{p}} = e^{-\frac{1}{2}i\hbar a \cdot b} e^{ib\hat{p}}e^{ia\hat{q}}$$

The unitary operator

$$\hat{U}(a, b) = e^{\frac{1}{2}i\hbar a \cdot b} e^{ia\hat{q}}e^{ib\hat{p}}$$

satisfies the multiplication law

$$\hat{U}(a, b)\hat{U}(a', b') = e^{\frac{1}{2}i\hbar[a' \cdot b - a \cdot b']}\hat{U}(a + a', b + b')$$

It is clear that we can associate a function on the classical phase space

$$T(p) + V(q)$$

with any operator of the form

$$T(\hat{p}) + V(q)$$

Many interesting hamiltonians of quantum mechanics are of this type. Is there a way to set up a one-one correspondence between operators (or their integral kernels) and functions on the phase space? If we could do this, we would be able to study the relation between quantum and classical mechanics much more conveniently. We will also be able to develop a semi-classical asymptotic expansion systematically.

11.2. From Symbols to Operators

How to convert a function on the phase space into an operator on the quantum Hilbert space? If it is a function of position alone, or of momentum alone, or even a sum of two such operators, we just saw how to set up such a one-one correspondence. If we have a product of position and momentum there is an ambiguity: the precise meaning depends on the order in which these operators act on the wavefunction. Each interpretation has its supporters, but I favor the **symmetric ordering**, because it preserves the hermitian character of observables. We should think of the operator corresponding to pq as $\frac{1}{2}(\hat{p}q + q\hat{p})$ because it is hermitean while $\hat{p}q$ or $q\hat{p}$ are not. Using the Fourier transform,

$$\tilde{A}(q, p) = \int A'(a, b) e^{i[a \cdot q + b \cdot p]} da db$$

we can find the operator corresponding to any function if we know how to do it for the basic functions

$$\tilde{U}(a, b|q, p) = e^{i[a \cdot q + b \cdot p]}$$

It can be interpreted to mean as $e^{ia\hat{q}}e^{ib\hat{p}}$, or $e^{ib\hat{p}}e^{ia\hat{q}}$ depending on whether \hat{p} or \hat{q} acts first. A Solomonic resolution of this would be to go half way, essentially treating momentum and position on an equal footing. That amounts to the rule

$$e^{i[a \cdot q + b \cdot p]} \rightarrow \hat{U}(a, b) = e^{\frac{1}{2}i\hbar a \cdot b} e^{ia\hat{q}} e^{ib\hat{p}}$$

Now,

$$\hat{U}(a, b) = e^{i\hbar a \cdot [q + \frac{b}{2}]} \psi(q + \hbar b)$$

$$\begin{aligned} \hat{A}\psi(q) &= \int da db A'(a, b) e^{i\hbar a \cdot [q + \frac{b}{2}]} \psi(q + \hbar b) \\ &= \int \tilde{A}\left(q + \frac{\hbar b}{2}, p\right) e^{-ip \cdot b} \psi(q + \hbar b) \frac{dp}{(2\pi\hbar)^n} db \\ &= \int \left[\int \tilde{A}\left(\frac{q+q'}{2}, p\right) e^{\frac{i}{\hbar} p \cdot (q-q')} \frac{dp}{(2\pi\hbar)^n} \right] \psi(q') dq' \end{aligned}$$

Thus the integral kernel of the operator corresponding to the symbol $\tilde{A}(q, p)$ is

$$A(q, q') = \int \tilde{A}\left(\frac{q+q'}{2}, p\right) e^{\frac{i}{\hbar} p \cdot (q-q')} \frac{dp}{(2\pi\hbar)^n}$$

We can think of this as an integral transform, analogous to the Fourier transform. It is actually half-way between a Fourier transform and a multiplication: If $\tilde{A}(q, p)$ happens to be independent of p it is a multiplication; if it is independent of q , it is a Fourier transform. It is often called the **Weyl transform**. It is not difficult to work out its inverse

$$\tilde{A}(q, p) = \int A\left(q + \frac{u}{2}, q - \frac{u}{2}\right) e^{-\frac{i}{\hbar} p \cdot u} du$$

Thus we can set up a dictionary between functions on phase space and operators (integral kernels) on wavefunctions.

11.3. The Star Product

It is important to note that the Weyl transform captures all the properties of the operator in its symbol: it is not just a classical approximation. So it must be possible to translate the multiplication of operators into some operation on symbols. As $\hbar \rightarrow 0$ it must reduce to the usual pointwise multiplication; we should be able to expand it in powers of \hbar to calculate quantum corrections.

The elementary case is the multiplication of the unitary operators

$$\hat{U}(a, b)\hat{U}(a', b') = e^{\frac{1}{2}i\hbar[a' \cdot b - a \cdot b']}\hat{U}(a + a', b + b')$$

which leads to

$$\tilde{U}(a, b) * \tilde{U}(a', b') = e^{\frac{1}{2}i\hbar[a' \cdot b - a \cdot b']} \tilde{U}(a + a', b + b')$$

The general case follows by applying a Fourier transform to this basic formula. A moment's thought will show that the above formula can be rewritten as

$$\tilde{U}(a, b) * \tilde{U}(a', b') = \tilde{U}(a, b) e^{-\frac{i\hbar}{2} \left(\overleftarrow{\frac{\partial}{\partial p}} \frac{\partial}{\partial q} - \overleftarrow{\frac{\partial}{\partial q}} \frac{\partial}{\partial p} \right)} \tilde{U}(a', b')$$

The left arrow is to denote that the differentiation acts on the quantity to the left. Each differentiation by $\frac{\partial}{\partial p}$ on the left pulls out a factor of b and on the right pulls out b' and similarly for $\frac{\partial}{\partial q}$. By applying a Fourier transform we see that this formula continues to be true for all symbols:

$$\tilde{A} * \tilde{B} = \tilde{A} e^{-\frac{i\hbar}{2} \left(\overleftarrow{\frac{\partial}{\partial p}} \frac{\partial}{\partial q} - \overleftarrow{\frac{\partial}{\partial q}} \frac{\partial}{\partial p} \right)} \tilde{B}$$

We can also write it explicitly by expanding the exponential

$$\tilde{A} * \tilde{B}(q, p) = \sum_{n=0}^{\infty} \left(-\frac{i\hbar}{2} \right)^n \frac{1}{n!} \tilde{A} \left(\overleftarrow{\frac{\partial}{\partial p}} \frac{\partial}{\partial q} - \overleftarrow{\frac{\partial}{\partial q}} \frac{\partial}{\partial p} \right)^n \tilde{B}$$

Expanding the power in a binomial series and rearranging,

$$\tilde{A} * \tilde{B}(q, p) = \sum_{r,s=0}^{\infty} \left(-\frac{i\hbar}{2} \right)^{r+s} \frac{(-1)^s}{r!s!} \tilde{A}_{j_1 \dots j_r}^{i_1 \dots i_r} \tilde{B}_{i_1 \dots i_s}^{j_1 \dots j_s}$$

where

$$\tilde{A}_{i_1 \dots i_r}^{j_1 \dots j_r} = \frac{\partial^{r+s} \tilde{A}}{\partial p_{j_1} \dots \partial p_{j_r} \partial q^{i_1} \dots \partial q^{i_s}}$$

If you did not know where formula for the star product comes from, you would be amazed that it is associative:

$$\tilde{A} * (\tilde{B} * \tilde{C}) = (\tilde{A} * \tilde{B}) * \tilde{C}.$$

To zeroth order in \hbar this is the usual product. The next order term in the Poisson bracket. After that is a second order derivative term and so on.

$$\begin{aligned} \tilde{A} * \tilde{B}(q, p) &= \tilde{A}(q, p) \tilde{B}(q, p) - \frac{i\hbar}{2} \left\{ \frac{\partial \tilde{A}}{\partial p_i} \frac{\partial \tilde{B}}{\partial q^i} - \frac{\partial \tilde{A}}{\partial q^i} \frac{\partial \tilde{B}}{\partial p_i} \right\} \\ &+ \left(-\frac{i\hbar}{2} \right)^2 \frac{1}{2} \left\{ \frac{\partial^2 \tilde{A}}{\partial p_i \partial p_j} \frac{\partial^2 \tilde{B}}{\partial q^i \partial q^j} + \frac{\partial^2 \tilde{A}}{\partial q^i \partial q^j} \frac{\partial^2 \tilde{B}}{\partial p_i \partial p_j} - 2 \frac{\partial^2 \tilde{A}}{\partial p_i \partial q^j} \frac{\partial^2 \tilde{B}}{\partial p_j \partial q^i} \right\} + \dots \end{aligned}$$

The essential example, from which all else follows, is

$$p * q = pq - \frac{i\hbar}{2}$$

The canonical commutation relations

$$p * q - q * p = -i\hbar$$

can be verified..

Real functions over to hermitean operators and vice versa. Of course the product of hermitean operators is not hermitean always; that is why the star product of two real functions might have an imaginary part.

11.4. Heisenberg Equation

In this point of view the focus is on observables rather than states. Given the symbol of the hamiltonian, we can get the equation for the time evolution of any observable

$$-i\hbar \frac{\partial \tilde{A}}{\partial t} = \tilde{H} * \tilde{A} - \tilde{A} * \tilde{H}$$

The most important are the equations for position and momentum. Because the second and higher order derivatives vanish, they just reduce to Hamilton's equations:

$$\frac{dq^i}{dt} = \frac{\partial \tilde{H}}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial \tilde{H}}{\partial q^i}$$

More complicate observables will have quantum correctins to their time evolution.

11.5. The Star Exponential

We can define the exponential of a symbol with respect to the star product by

$$e_*^{\tilde{A}} = \sum_{n=0}^{\infty} \frac{1}{n!} \tilde{A} * \tilde{A} * \dots * \tilde{A} \text{ (n times)}$$

Or we can define it by the differential equation

$$\frac{d}{dt} e_*^{t\tilde{A}} = \tilde{A} * e_*^{t\tilde{A}}$$

EXAMPLE 33. Consider the hamiltonian of the harmonic oscillator

$$\tilde{H}(q, p) = \frac{p^2 + q^2}{2}$$

Suppose

$$\tilde{h}_t = e_*^{-t\tilde{H}}$$

To find it we have to solve the equation

$$\frac{\partial}{\partial t} \tilde{h}_t(q, p) = -\frac{p^2 + q^2}{2} \tilde{h}_t(q, p) - \frac{i\hbar}{2} \left\{ -p \frac{\partial \tilde{h}_t}{\partial q} + q \frac{\partial \tilde{h}_t}{\partial p} \right\} - \left(-\frac{i\hbar}{2} \right)^2 \frac{1}{2} \left\{ \frac{\partial^2 \tilde{h}_t}{\partial q^2} + \frac{\partial^2 \tilde{h}_t}{\partial p^2} \right\}$$

There are no higher order terms because the hamiltonian is a quadratic function. It is not hard to guess that the solution must be a gaussian. Based on the rotation symmetry of the system in the $p - q$ plane, we make the ansatz

$$\tilde{h}_t(q, p) = e^{a_t[p^2 + q^2] + b_t}$$

The second term in the PDE above is then a Poisson bracket of two terms that depend only on $p^2 + q^2$; so it vanishes. Thus¹

$$\begin{aligned}\dot{a}_t &= -\frac{1}{2} - \left(-\frac{i}{2}\right)^2 \frac{1}{2} [(2a_t)^2] \implies a_t = -\tanh \frac{t}{2} \\ \dot{b}_t &= -\left(-\frac{i}{2}\right)^2 \frac{1}{2} \times 2a_t, \implies b_t = -\log \cosh \frac{t}{2} \\ \tilde{h}_t(q, p) &= \frac{1}{\cosh \frac{t}{2}} e^{-[p^2+q^2] \tanh \frac{t}{2}}\end{aligned}$$

In particular, the limit $t \rightarrow \infty$ gives the ground state energy $\frac{1}{2}$:

$$\tilde{h}_t(q, p) \rightarrow e^{-\frac{t}{2}} \tilde{P}_0(q, p)$$

and the symbol of the projection operator to the ground state

$$\tilde{P}_0(q, p) = 2e^{-[p^2+q^2]}.$$

11.6. The Trace

The trace of an operator is the sum over all its diagonal elements

$$\text{tr } A = \int A(q, q) dq$$

It is easy to check that in terms of symbols

$$\text{tr } A = \int \tilde{A}(q, p) \frac{dpdq}{[2\pi\hbar]^n}$$

This is the integral on phase space with respect to the Liouville measure, normalized such that it counts the number of quantum states.

EXERCISE 34. Find the partition function $Z(t) = \text{tr } e_*^{-t\tilde{H}}$ for the harmonic oscillator. Using its expression in terms of the spectrum $\text{tr } e_*^{-t\tilde{H}} = \sum_n e^{-tE_n}$ find the spectrum of the harmonic oscillator.

Solution. Using the above solution,

$$\begin{aligned}\text{tr } e^{-t\tilde{H}} &= \int \frac{dpdq}{[2\pi]} \frac{1}{\cosh \frac{t}{2}} e^{-[p^2+q^2] \tanh \frac{t}{2}} \\ &= \frac{1}{2 \sinh \frac{t}{2}} \\ &= \sum_{n=0}^{\infty} e^{-(n+\frac{1}{2})t}\end{aligned}$$

which gives the usual spectrum. Note also that $2e^{-[p^2+q^2]}$ has trace one, which is the degeneracy of the ground state:

$$\int \frac{dqdp}{2\pi} 2e^{-[p^2+q^2]} = 1.$$

¹We set $\hbar = 1$ for simplicity.

11.7. The Resolvent Symbol

The resolvent of an operator is

$$R(\lambda) = \frac{1}{\lambda - A}$$

When A is a differential operator, the resolvent is its Green's function. It contains all the information about eigenvalues and eigenvectors of A . If A is diagonalizable with purely discrete spectrum,

$$R(\lambda) = \sum_n \frac{1}{\lambda - a_n} P_n$$

where P_n is the projection operator to the eigenspace with eigenvalue a_n .

$$AP_n = a_n P_n, \quad P_n^2 = P_n$$

The symbol of the resolvent satisfies

$$\tilde{R}(\lambda) * (\lambda - \tilde{A}) = 1$$

If we expand

$$\tilde{R}(\lambda) = \sum_{n=0}^{\infty} \hbar^n r_n(\lambda)$$

we can get recursion relations for the coefficients

$$r_0(\lambda)[\lambda - \tilde{A}] = 1$$

$$r_1(\lambda) - \frac{i}{2} \left\{ r_0, \tilde{A} \right\} = 0$$

$$r_2(\lambda)[\lambda - \tilde{A}] - \left(-\frac{i}{2} \right)^2 \frac{1}{2} \left\{ \frac{\partial^2 \tilde{r}_0}{\partial p_i \partial p_j} \frac{\partial^2 \tilde{A}}{\partial q^i \partial q^j} + \frac{\partial^2 \tilde{r}_0}{\partial q^i \partial q^j} \frac{\partial^2 \tilde{A}}{\partial p_i \partial p_j} - 2 \frac{\partial^2 \tilde{r}_0}{\partial p_i \partial q^j} \frac{\partial^2 \tilde{A}}{\partial p_j \partial q^i} \right\} = 0$$

etc.

The solutions are

$$r_0 = \frac{1}{\lambda - \tilde{A}}$$

$$r_1 = 0$$

$$r_2 = -\frac{1}{8} r_0 \left\{ \frac{\partial^2 \tilde{r}_0}{\partial p_i \partial p_j} \frac{\partial^2 \tilde{A}}{\partial q^i \partial q^j} + \frac{\partial^2 \tilde{r}_0}{\partial q^i \partial q^j} \frac{\partial^2 \tilde{A}}{\partial p_i \partial p_j} - 2 \frac{\partial^2 \tilde{r}_0}{\partial p_i \partial q^j} \frac{\partial^2 \tilde{A}}{\partial p_j \partial q^i} \right\}$$

etc.

This gives a way to solve a quantum system by an asymptotic expansion even when exact solutions are not possible.

11.8. Can you hear the shape of a drum?

An interesting application of this procedure is to the solution of inverse problems. Knowing the sound produced by a drum, can you deduce its shape (the boundary conditions on the Laplace operator)? This is a famous problem posed by M. Kac. More generally, is it possible to deduce the potential (or refractive index in optics) of a Schrodinger operator knowing the spectrum? Such problems are important in applications such as oil exploration and medical imaging. Exact solutions are impossible in practical situations. Semi-classical asymptotics is usually feasible and gives a good first approximation when the wavelength is small compared to the features explored. To leading order, this amounts to “ray tracing”. The expansion we described above allows an expansion in powers of the wavelength.

We can deduce almost immediately the area of the drum from the partition function of the spectrum. Suppose we have the Laplace operator with some b.c. The configuration space is some bounded domain of R^n . The Laplace operator has symbol

$$\tilde{A} = p^2.$$

Consider the partition function

$$Z(t) = \text{tr} e^{-t\tilde{A}}$$

Since p appears only in the combination tp^2 , in the limit of small t , we can keep just the leading terms of the expansion above

$$\begin{aligned} Z(t) &\sim \int dq \frac{dp}{[2\pi]^n} e^{-tp^2} \\ &\sim V \frac{1}{(2t)^{\frac{n}{2}}} \end{aligned}$$

as $t \rightarrow 0$. Here, $V = \int dx$ is the size of the domain in R^n . When $n = 2$, (the case of the drum) it is the area.

CHAPTER 12

The Gutzwiller Trace Formula

The WKB method, which leads to the Bohr-Sommerfeld quantization rules, was originally devised for a system with one degree of freedom. It can be extended to systems that can be decomposed into independent degrees of freedom by a canonical transformation. That is, when the Schrodinger equation is separable. But most systems that occur in nature are not of this type. Their classical limits are chaotic. There must still be a semi-classical approximation for such systems.

An example of such a chaotic quantum system is a Sinai billiard. Particles move freely on the plane until reflected by a boundary. Equivalently, we can think of waves reflected by mirrors at the boundary. If the boundary is a rectangle (as in the actual billiard table) the Schrodinger equation can be separated in Cartesian coordinates. If the boundary is a union of convex curves (i.e., they bulge inwards) the classical motion (or the path of rays) is chaotic. At each reflection the divergence between two paths is increased. Yet, the paths have to stay within a finite area. These two somewhat incompatible conditions lead to chaos. What are the quantum energy levels (natural frequencies) of such a system?

The first step forward in this direction was taken by A. Selberg, one of the great number theorists of the twentieth century. He was studying geodesics on a Riemann surface of genus g greater than two. Such a surface can be thought of as a hyperboloid, with some additional boundary conditions. The boundary is a union of $4g$ geodesics with opposite pairs of equal length. A particle that hits a point on the boundary is instantly transported to the corresponding point on the opposite edge, with its velocity now pointed inwards. Here again, any small divergence between trajectories will grow with time. Not because of the boundary, but because of the negative curvature of the interior.

Chaotic classical systems (such as the billiard or the Riemann surface) have only a countably infinite number of closed trajectories. This is very different from the Kepler problem for example, where there is a closed trajectory for any value of energy and angular momentum. In an unstable system, most initial conditions will not lead to closed trajectories. You have to aim the billiard ball just right if it is to return to the original point with the same velocity after a number of reflections. Thus, chaotic classical systems have a kind of spectrum: the actions of the closed trajectories. For geodesics, the action is the same as the length.

The quantization of the problem of geodesics is the solution of the eigenvalue equation for the Laplace operator. Selberg derived a remarkable formula giving the eigenvalues of the laplacian in terms of the lengths of the classical geodesics. The catch is that this is not a one-one correspondence: a sum over a function of eigenvalues is equal to a sum over another function (a kind of “dual”) of the lengths. An elementary version of this is the Poisson summation formula for sums over integers.

Selberg's result was an exact result in a specific system with a high degree of symmetry. (The surface has constant curvature). In general we would expect at least an approximate relation between classical spectrum of actions and the quantum spectrum of energy eigenvalues. Gutzwiller derived such a formula, which has become the foundational result on quantum chaos. His original derivation used path integrals, a kind of Laplace formula for infinite dimensional integrals. It can also be derived by using symbol calculus. Each derivation gives a different physical insight into this remarkable formula, a gem of twentieth century theoretical physics.

Let us begin with an elementary example, the Poisson summation formula for functions of the circle.

12.1. The Poisson Summation Formula

PROPOSITION 35. *For a positive number L (the period),*

$$\sum_{a \in LZ} \frac{e^{-\frac{a^2}{2t}}}{\sqrt{2\pi t}} = \frac{1}{L} \sum_{k \in \frac{2\pi}{L}Z} e^{-k^2 t}.$$

This formula is remarkable because the l.h.s. has terms that get smaller for $t \rightarrow 0$ while the r.h.s. has terms that decrease for large t . It is relating the small time behavior to the large time behavior.

The idea of the proof is to construct the solution of the heat equation (the Schrodinger equation with imaginary time) on the circle of perimeter L , in two different ways. Since the solution is unique this will give us a useful identity.

$$\frac{\partial}{\partial t} h_t(x) = \frac{\partial^2}{\partial x^2} h_t(x)$$

$$h_t(x+a) = h_t(x), \quad a = \dots - 2L, -L, 0, L, 2L \dots$$

$$\lim_{t \rightarrow 0} h_t(x) = \sum_{a \in LZ} \delta(x+a)$$

If we look at the problem on the real (i.e, without periodicity)

$$\frac{\partial}{\partial t} g_t(x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} g_t(x)$$

$$\lim_{t \rightarrow 0} g_t(x) = \delta(x)$$

the solution is a Gaussian:

$$g_t(x) = \frac{e^{-\frac{x^2}{2t}}}{\sqrt{2\pi t}}$$

One way to get the solution with periodic boundary conditions is to sum over all points separated by one period:

$$h_t(x) = \sum_{a \in LZ} g_t(x+a).$$

This is the sum on the l.h.s. of the Poisson formula.

Another way is to expand the periodic function in a Fourier series

$$h_t(x) = \sum_{k \in \frac{2\pi}{L}Z} \tilde{h}_t(k) e^{ikx}$$

to get

$$\frac{\partial \tilde{h}_t(k)}{\partial t} = -k^2 \tilde{h}_t(k)$$

and

$$\lim_{t \rightarrow 0} \tilde{h}_t(k) = \frac{1}{L}.$$

so that

$$\tilde{h}_t(k) = \frac{1}{L} e^{-k^2 t}.$$

This solution gives the r.h.s. of the Poisson formula.

12.1.1. Physical Interpretation. This result has a nice physical interpretation in terms of quantum mechanics (except that the time is imaginary). Consider a particle moving on the circle with perimeter L . Its classical trajectories are solutions of

$$\ddot{x} = 0$$

that is, straight lines in space-time. If we require the positions at time zero and at time T to be the same modulo a translation a by a multiple of L we will get

$$x(t) = a \frac{t}{T} + x_0$$

The action of this trajectory is

$$\int_0^T \dot{x}^2(t) dt = \frac{a^2}{2T}$$

The l.h.s. is the sum over the exponentials of the actions of all the closed trajectories, except for the factor of $\frac{1}{\sqrt{2\pi T}}$. The r.h.s. is the sum over the exponentials of the energies multiplied by T ; i.e., $\text{tr} e^{-\hat{H}t}$ (again, except for the factor $\frac{1}{L}$). So what we have here is a formula of the type

$$\text{tr} e^{-\hat{H}t} \sim \sum_{\text{closed}} e^{-S}$$

that relates a quantum mechanical quantity to its classical analogue.

Gutzwiller's formula is a generalization of this idea to more general systems for which also the closed trajectories are countable. The factor $\frac{1}{\sqrt{2\pi t}}$ will be explained as a kind of Jacobian: the determinant of the derivative of momentum w.r.t. to position.

EXERCISE 36. Suppose f is a smooth function rapidly decreasing at infinity and $\tilde{f}(k) = \int f(x) e^{-ikx} dx$ its Fourier transform. Then there is a generalization of the Poisson formula

$$\sum_{a \in LZ} f(a) = \sum_{k \in \frac{2\pi}{L}Z} \tilde{f}(k)$$

EXERCISE 37. Recall the **Riemann zeta function** $\zeta(s) = \sum_{k=1}^{\infty} \frac{1}{k^s}$. Using the Poisson formula, we can derive the **functional relation**

$$\xi(s) = \xi(1-s)$$

where

$$\xi(s) = \pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s).$$

REMARK 38. It is believed that all the zeros of $\xi(s)$ lie on the axis of symmetry $\text{Re } s = \frac{1}{2}$. Proving this **Riemann hypothesis** is the most famous celebrated unsolved problem in all of mathematics. It is not part of the above exercise. Selberg's motivation in deriving his trace formula was to find a generalization of this zeta function where the sum over integers is replaced by the sum over a non-abelian group (the fundamental group of the Riemann surface).

12.2. The Propagator

The kernel of the unitary operator $e^{-\frac{i}{\hbar} \hat{H}t}$ is called the propagator:

$$U(t, q, q') = \langle q | e^{-\frac{i}{\hbar} \hat{H}t} | q' \rangle.$$

It can be expanded in terms of the eigenvectors of the hamiltonian (when the spectrum is pure point; i.e., bound states only)

$$U(t, q, q') = \sum_n e^{-\frac{i}{\hbar} E_n t} \psi_n(q) \psi_n(q'), \quad \hat{H} \psi_n = E_n \psi_n$$

Each eigenvector is normalized to have length one. The trace

$$\text{tr} e^{-\frac{i}{\hbar} \hat{H}t} = \sum_n e^{-\frac{i}{\hbar} E_n t} \rho(E_n)$$

contains the information about the energy eigenvalues E_n and their degeneracies $\rho(E_n)$. More generally, if we have a continuous spectrum as well

$$\text{tr} e^{-\frac{i}{\hbar} \hat{H}t} = \int e^{-\frac{i}{\hbar} Et} \rho(E) dE$$

where $\rho(E) dE$ is the **spectral density** (or **density of states** in physicist's jargon). It contains a δ -function for every discrete eigenvalue and a continuous distribution for the scattering states.

Often it is useful to analytically continue the above quantities into imaginary time to get the **heat kernel**

$$h_t(q, q') = \langle q | e^{-\frac{1}{\hbar} \hat{H}t} | q' \rangle$$

which satisfies the differential equation

$$\hbar \frac{\partial h_t(q, q')}{\partial t} = -\hat{H} h_t(q, q')$$

with the initial condition

$$h_t(q, q') \rightarrow \delta(q, q'), \quad t \rightarrow 0.$$

A closely related operator is the resolvent

$$\hat{R}(\lambda) = \frac{1}{\hat{H} - \lambda} = \frac{1}{\hbar} \int_0^\infty e^{\frac{1}{\hbar}\lambda t} e^{-\frac{1}{\hbar}\hat{H}t} dt$$

which is well-defined (is a bounded operator) if λ is **not** in the spectrum of \hat{H} . Its kernel is called the Green's function

$$R(\lambda; q, q') = \langle q | \frac{1}{\hat{H} - \lambda} | q' \rangle = \sum_n \frac{1}{E_n - \lambda} \psi_n(q)^* \psi_n(q')$$

Again the traces contain the information about eigenvalues and degeneracies. The trace

$$Z_t = \text{tr} e^{-\hat{H}t}$$

is the **partition function**. This contains all the information of statistical mechanics where t has the meaning of the *inverse* of temperature, for the hamiltonian of some quantum system. It is equal to

$$Z_t = \int e^{-Et} \rho(E) dE$$

in terms of the density of states. Our aim is to derive a semi-classical approximation for this partition function. A Schrodinger operator is a second order differential operator on functions $\psi : R^n \rightarrow C$

$$\hat{H}\psi = \frac{1}{2} (i\hbar\partial)^2 \psi + V(x)\psi$$

12.2.1. Small Time Approximation. If we ignore all the \hbar dependent terms, the star product reduces to the ordinary product and

$$\tilde{h}_t = e^{-\hat{H}t} + O(\hbar)$$

so that the partition just becomes its classical analogue

$$Z_{t,0} = \int e^{-\tilde{H}(x,p)t} \frac{dx dp}{[2\pi\hbar]^n}$$

The only place where \hbar appears is in the normalization of the volume of phase space.

If the hamiltonian has the form

$$\tilde{H}(x, p) = \frac{1}{2} p^2 + V(x)$$

we can evaluate this integral to get

$$Z_{t,0} = \int e^{-V(x)t} \frac{dx}{[\sqrt{2\pi t\hbar}]^n}$$

For the isotropic harmonic oscillator for example, $V(x) = \frac{1}{2}\omega^2 x^2$

$$Z_{t,0} = \frac{1}{[t\hbar\omega]^n}$$

¹This is actually an important point: this constant is undetermined in the classical theory. It can equivalently be thought of as an additive constant in the entropy. Quantum theory determines this constant in terms of \hbar : in terms of entropy it is related to the third law of thermodynamics.

while the exact answer is

$$Z_t = \frac{1}{[2 \sinh \frac{t\hbar\omega}{2}]^n}.$$

Clearly, the small time behavior is captured correctly, but not for large time.

12.2.2. Semiclassical Approximation. We can make the ansatz

$$h_t(q, q') = e^{-\frac{1}{\hbar} S_t(q, q')}$$

and an expansion

$$S_t(q, q') = \sum_{n=0}^{\infty} \hbar^n S_n(q, q')$$

in the heat equation. To leading order we get a Hamilton-Jacobi equation

$$\frac{\partial S_{t,0}}{\partial t} = \frac{1}{2} \left[\frac{\partial S_{t,0}}{\partial q} \right]^2 - V(q)$$

with the b.c. that $S_t(q, q') \rightarrow \frac{(q-q')^2}{2t}$ as $t \rightarrow 0$. (Then $h_t(q, q')$ will tend to the delta-function). The potential energy has the “wrong” sign because this is the imaginary time version of the Hamilton-Jacobi equation. The solution is the action of the (imaginary time) trajectory that connects the two points in time t .

$$\ddot{q} = \frac{\partial V}{\partial q}.$$

Again, the sign of the force is the opposite of that in Newton’s equations.

If we only want the energy levels, we will need the trace of h_t ; that is we will put $q = q'$ and integrate over all values of q . This means that we will need a sum over all closed orbits of period t .

The next order term will determine the term S_1 in terms of derivatives of $S_{t,0}$:

$$h_t(q, q') = e^{-\frac{1}{\hbar} S_{t,0}(q, q')}$$

This will work out to involve the determinant of the second derivative of the action with respect to initial conditions. Thus we get a formula

$$\text{tr} e^{-\frac{1}{\hbar} \hat{H}t} = \sum \int dq e^{-\frac{1}{\hbar} S_t(q, q)} \sqrt{\left[\det \frac{1}{2\pi} \frac{\partial^2 S}{\partial q \partial q'} \right]_{q=q'}}$$

the sum being over all orbits of period t starting and ending at q . This is the Gutzwiller trace formula.

EXERCISE 39. Work out the first order correction to S_t in the above expansion to get the $\sqrt{\det}$ factor in the Gutzwiller formula.

EXERCISE 40. Verify this formula for the particle on a circle and the simple harmonic oscillator.

Singular Perturbation Theory for Differential Equations

The semi-classical expansion of the Schrodinger equation

$$-\frac{\hbar^2}{2}\partial^2\psi + V\psi = E\psi$$

is an example of singular perturbation theory: if we simply set $\hbar = 0$ in the equation, we change its order. In this case, we change from a second order equation to a zeroth order equation. Physically, it is clear that we should not interpret the classical limit as simply setting $\hbar = 0$ in the above equation: we would then lose the kinetic energy and not just quantum effects. It is wrong to expand ψ itself in powers of \hbar . The point is that the solution has an essential singularity as $\hbar \rightarrow 0$, so the correct this is to put

$$\psi = e^{\frac{1}{\hbar}S}$$

and then to expand $S = \sum_{r=0}^{\infty} S_r \hbar^r$. In the leading order we get the Hamilton-Jacobi equation.

Such singular perturbations occur in many branches of physics. Perhaps the deepest is in the Navier-Stokes equation of fluid mechanics

$$\frac{\partial \mathbf{v}}{\partial t} = \nu \partial^2 \mathbf{v} + \mathbf{v} \cdot \partial \mathbf{v} - \nabla p, \quad \operatorname{div} \mathbf{v} = 0$$

along the with the boundary condition that $\mathbf{v} = 0$ at a spatial boundary.

This is a second order nonlinear system of PDEs. The first term on the rhs is proportional to viscosity. The limit of an ideal fluid $\nu \rightarrow 0$ is singular because the order of the equation changes from two to one. It turns out viscosity has the biggest effect near the boundary: there is a **boundary layer** in which the viscosity cannot be ignored no matter how small it is. Non-linear PDEs such as this are among the most difficult problems in all of mathematics and physics. So we will mostly look at ODEs in this course.

13.1. Center Manifolds

A basic reference is NEIL FENICHEL, JOURNAL OF DIFFERENTIAL EQUATIONS 31, 53-98 (1979) based on earlier work of Anosov.

We can think of the independent variable as time. Suppose there are two dependent variables (each of which can be a vector) $x \in R^m$ and $y \in R^n$:

$$\dot{x} = f(x, y, \epsilon), \quad \dot{y} = g(x, y, \epsilon)$$

The set of initial conditions (“the phase space”) of the system is R^{m+n} . When $\epsilon = 0$ the system degenerates to

$$\dot{x} = f(x, y, 0), \quad 0 = g(x, y, 0).$$

Thus, in this limit initial conditions cannot be chosen independently: the y variables are determined by the x variables through the second equation, in terms of some function $h(x)$:

$$y = h(x)$$

and there is a limiting equation for the x variables

$$\dot{x} = f(x, h(x), 0).$$

How do these solutions change when ϵ is small but not zero? The tricky part is that now there could be other solutions, whose y co-ordinate might grow with time, invalidating the starting approximation above. It turns out that under reasonable conditions, it is still possible to find an m -dimensional submanifold of R^{m+n} on which the solution lies.

$$y = h(x, \epsilon)$$

The equation of time evolution can be projected on this submanifold (called the **center manifold**) and effectively replace the original equation.

$$\dot{x} = f(x, h(x, \epsilon), \epsilon)$$

For this to happen, $h(x, \epsilon)$ must satisfy the partial differential equation

$$\epsilon f^a(x, h(x, \epsilon), \epsilon) \frac{\partial h^j(x, \epsilon)}{\partial x^a} = g^j(x, h(x, \epsilon), \epsilon).$$

This equation can be solved as an asymptotic expansion in powers of ϵ , determining the center manifold. Determining the center manifold is itself part of the solution of the dynamics of such problems.

There is then an expansion for solutions $x(t, \epsilon)$ that lie on this submanifold as well, although not for those outside of it. We illustrate the idea with a linear example.

13.1.1. Expansion Method for the Center Manifold. Determining the center manifold (the function $h(x, \epsilon)$) is essential to understanding systems with vastly different time scales. We can approach this by an expansion in powers of ϵ .

$$h(x, \epsilon) = \sum_n h_n(x) \epsilon^n$$

with $h_0(x) = h(x)$ as above. Then

$$f^a(x, h(x), 0) \frac{\partial h^j(x)}{\partial x^a} = \frac{\partial g^j}{\partial \epsilon}(x, h(x), 0) + h_1^k(x) \frac{\partial g^j}{\partial y^k}(x, h(x), 0)$$

If the matrix $\frac{\partial g^j}{\partial y^k}(x, h(x), 0)$ is invertible, this can be solved for h_1 . And so on for higher orders.

13.2. Cyclotron Radiation

Consider a charged particle moving in a constant magnetic field. Its velocity satisfies the Lorentz equation

$$\frac{d\mathbf{v}}{dt} = \frac{e}{m} \mathbf{v} \times \mathbf{B}$$

Consider the special case where the velocity is normal to the magnetic field, which we can assume to be along the third axis:

$$\frac{dv_1}{dt} = \omega_c v_2, \quad \frac{dv_2}{dt} = -\omega_c v_1$$

where the **cyclotron frequency** is

$$\omega_c = \frac{e}{m} B$$

Introducing the complex variable $v = v_1 + iv_2$ our equation becomes

$$\frac{dv}{dt} = -i\omega_c v$$

The solution is obvious

$$v = v(0)e^{-i\omega_c t}$$

This describes a particle in uniform circular motion with angular frequency ω_c . Devices with electrons in such orbits in a magnetic field are called cyclotrons. Typically the frequencies are in the MegaHertz range. These devices started as tools for nuclear physics research, but now have found applications in medicine as a source of radiation for cancer treatment.¹

Now, every charged particle with an acceleration will radiate. The electrons in a cyclotron lose energy as well: although the magnitude of velocity is constant, its direction is changing, so that the particle is accelerating. The only reason cyclotrons work is that energy is replenished by pumping in microwave radiation. In the absence of that, the electrons would lose energy slowly and come to rest. The radiation emitted by the electron is mostly in the forward direction, so it exerts a recoil force on it which will tend to slow the electron down. This force is notoriously difficult to calculate, being the force exerted on the electron by itself. If we naively use the formula of electrodynamics, the force is infinite in the limit of a point particle. Dirac showed that this infinity can be absorbed into a redefinition of the mass of the electron. The residual effects of this **renormalization** is a self-force proportional to the derivative of acceleration. We get the Lorentz-Dirac equation

$$\frac{d\mathbf{v}}{dt} = \frac{e}{m} \mathbf{v} \times \mathbf{B} + \epsilon \ddot{\mathbf{v}}, \quad \epsilon = \frac{2}{3} \frac{e^2}{mc^3}.$$

The parameter ϵ has units of time and is equal (in the case of the electron) to about

$$\epsilon \approx 2 \times 10^{-23} \text{s}$$

¹If we include relativistic effects, the frequency has a dependence on the energy (not just the mass). Such devices are called synchrotrons. We will consider only the non-relativistic case here.

This is a fantastically small time compared to the other parameter in the problem, the inverse of the cyclotron frequency:

$$\epsilon\omega_c \sim 10^{-17}$$

But simply setting ϵ equal to zero is not correct as it changes the order of the equation: the solution is not analytic in ϵ . Again, restricting the motion to the plane normal to B we get

$$\dot{v} = -i\omega_c v + \epsilon \ddot{v}$$

13.2.1. Runaway Solutions. Being linear ODE with constant coefficients we can solve it by the ansatz

$$v = Ae^{\lambda t}$$

We get a quadratic equation for λ

$$\epsilon\lambda^2 - \lambda - i\omega_c = 0$$

with two solutions

$$\lambda_{\pm} = \frac{1 \pm \sqrt{1 + 4i\omega_c\epsilon}}{2\epsilon}$$

So the solution is

$$v = A_+ e^{\lambda_+ t} + A_- e^{\lambda_- t}$$

In the limit of small ϵ

$$\lambda_- \approx -i\omega_c - \epsilon\omega_c^2 + O(\epsilon^2)$$

$$\lambda_+ \approx \frac{1}{\epsilon} + i\omega_c + O(\epsilon)$$

The term $A_- e^{-\lambda_- t} \approx A_- e^{-i\omega_c t - \epsilon\omega_c^2 t}$ describes a slowly decreasing, periodic function. This is the physically correct solution for the velocity of the electron. But there is another solution

$$A_+ e^{\frac{1}{\epsilon} + i\omega_c t + \dots}$$

which has the non-analytic dependence on ϵ . This is unphysical, as it grows with time! The energy of the electron (proportional to the absolute square of v) increases with time, and very rapidly at that. Such runaway solutions are a source of confusion even now in the physics literature (including in some popular text books on electrodynamics). If this term is present in the solution, it will eventually dominate over the decreasing solution, no matter how small the coefficient A_+ . So the phase space must be identified with the subspace on which this term is set exactly to zero.

Conceptually, the equation of motion should not have been second order in the velocities: in mechanics, we should not have the freedom to choose the initial acceleration at will. It must somehow be determined by the initial position and velocity: the physically correct choice must be the one that kills off the runaway term in the solution.

13.2.2. The Landau-Lifshitz Equation. A clue is that the radiation term is actually small. So we can approximate

$$\ddot{v} \approx \frac{d}{dt} [-i\omega_c v]$$

so that Lorentz-Dirac equation is replaced by the Landau-Lifshitz equation, which is first order in the velocities:

$$\dot{v} = -i\omega_c v - i\epsilon\omega_c \dot{v}$$

The solution is easy

$$\dot{v} = \frac{-i\omega_c}{1 + i\epsilon\omega_c} v, \implies v = v_0 e^{\frac{-i\omega_c}{1+i\epsilon\omega_c} t}$$

which is a decreasing function. Does this approximation hold in higher orders? That is, are the higher order terms smaller, or is there lurking yet another runaway solution?

13.2.3. The Center Manifold for Cyclotron Motion. The application of center manifold theory to resolve completely the runaway solutions of the Lorentz-Dirac equation is the work of Herbert Spohn. We only work out the particular case of constant magnetic field.

By identifying $x = (v_1, v_2) \in R^2$ and

$$y_a = \dot{v}_a$$

we can write the equation as

$$\dot{x}_a = f_a(x, y, \epsilon), \quad f_a(x, y, \epsilon) = y_a$$

$$\epsilon \dot{y}_a = g_a(x, y, \epsilon), \quad g_a(x, y, \epsilon) = y_a - \omega_{ab} x_b$$

The matrix ω_{ab} has components

$$\omega_{12} = -\omega_{21} = \omega_c, \quad \omega_{11} = \omega_{22} = 0.$$

In the limit $\epsilon = 0$

$$y_a = h_a(x), \quad h_a(x) = \omega_{ab} x_b$$

The center manifold is given by the function $h(x, \epsilon)$ satisfying the PDE

$$\epsilon h_b(x, \epsilon) \partial_b h_a(x, \epsilon) = h_a(x, \epsilon) - \omega_{ab} x_b$$

If we set

$$h_a(x, \epsilon) = h_{ab} x_b$$

this reduces to an algebraic equation for the 2×2 matrix h

$$\epsilon h_{bc} h_{ab} = h_{ac} - \omega_{ac}$$

or

$$\epsilon h^2 = h - \omega$$

This can be solved as a power series in $\epsilon\omega$. Or, we can write it as

$$h = \frac{1 - \sqrt{1 - 4\epsilon\omega}}{2\epsilon}$$

where the square root of a matrix is defined as the power series

$$\sqrt{1 + A} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} A^n$$

etc. Notice that this is the branch of the square root that tends to ω as $\epsilon \rightarrow 0$. If the initial condition lies on the submanifold

$$y_a = h_{ab}(\epsilon)x_b$$

the solutions stay on it for ever. On this submanifold, the equation reduces to

$$\dot{x}_a = h_{ab}(\epsilon)x_b$$

The solution is (in complex notation $x = x_1 + ix_2$)

$$x(t) = e^{\lambda t}x(0)$$

where

$$\lambda = \omega_c \{ -i [1 - 2\epsilon^2\omega_c^2] + \epsilon\omega_c (\epsilon^2\omega_c^2 - 1) \}$$

We can see that this is the decreasing periodic solution, except that the cyclotron frequency and damping constant have been corrected slightly:

$$\text{Re}\lambda = -\epsilon\omega_c^2 + \epsilon^3\omega_c^4 + \dots$$

$$\text{Im}\lambda = -\omega_c [1 - 2\epsilon^2\omega_c^2 + \dots]$$

In view of the smallness of $\epsilon\omega_c \approx 10^{-17}$ there is no practical reason to go to higher orders. But there is nothing to prevent us from going to as high an order as we want. Note that although the equations we are solving are linear, the determination of the center manifold requires us to solve a quadratic equation for the matrix h .

We conclude that the Landau-Lifshitz equation is the first order in an expansion in powers of ϵ for the orbits of the radiating electron. That is the important conceptual point.

13.3. Center-Manifold of the Lorentz-Dirac Equation

We start with

$$\ddot{x}^\mu = \frac{e}{m} F_\nu^\mu \dot{x}^\nu + \epsilon [\delta_\nu^\mu - \dot{x}^\mu \dot{x}_\nu] \ddot{x}^\nu$$

Here t is proper time and $\epsilon = \frac{2}{3} \frac{e^2}{mc^3}$.

Define

$$v^\mu = \dot{x}^\mu, \implies v^\mu \ddot{x}_\mu = 0$$

$$y^\mu = \ddot{x}^\mu = [\delta_\nu^\mu - \dot{x}^\mu \dot{x}_\nu] \ddot{x}^\nu, \implies v^\mu y^\mu = -y^2$$

to get the system

$$\dot{x}^\mu = v^\mu,$$

$$\dot{v}^\mu = y^\mu$$

$$\epsilon \dot{y}^\mu = -\epsilon v^\mu y^2 + y^\mu - \frac{e}{m} F_\nu^\mu v^\nu$$

We seek a function $h(x, v, \epsilon)$ such that

$$\epsilon \left[v^\nu \frac{\partial h^\mu(x, v, \epsilon)}{\partial x^\nu} + h^\nu \frac{\partial h^\mu(x, v, \epsilon)}{\partial v^\nu} + v^\mu h^2(x, v, \epsilon) \right] = h^\mu(x, v, \epsilon) - \frac{e}{m} F_\nu^\mu v^\nu.$$

Then the equation of motion is

$$\dot{x}^\mu = v^\mu$$

$$\dot{v}^\mu = h^\mu(x, v, \epsilon)$$

Expanding

$$h(x, v, \epsilon) = \sum_{n=0}^{\infty} \epsilon^n h_n(x, v)$$

we have

$$h_0^\mu(x, v) = \frac{e}{m} F_\nu^\mu v^\nu$$

$$h_1^\mu = v^\nu \frac{\partial h_0^\mu(x, v, \epsilon)}{\partial x^\nu} + h_0^\nu \frac{\partial h_0^\mu(x, v, \epsilon)}{\partial v^\nu} + v^\mu h_0^2$$

$$h_1^\mu = v^\sigma \partial_\sigma \left[\frac{e}{m} F_\nu^\mu \right] v^\nu + \left(\frac{e}{m} \right)^2 [\delta_\nu^\mu - v^\mu v_\nu] (F^2)_\rho^\nu v^\rho$$

etc. To this order, we get the Landau-Lifshitz equation, proposed (without proof) in their classic text:

$$\ddot{x}^\mu = \frac{e}{m} F_\nu^\mu \dot{x}^\nu + \epsilon \left\{ \dot{x}^\sigma \partial_\sigma \left[\frac{e}{m} F_\nu^\mu \right] \dot{x}^\nu + \left(\frac{e}{m} \right)^2 [\delta_\nu^\mu - \dot{x}^\mu \dot{x}_\nu] (F^2)_\rho^\nu \dot{x}^\rho \right\}$$

We can calculate higher order corrections, but they are very small. It is worth doing for conceptual reasons, however. As we carry this out to higher orders, we will get more and more derivatives and powers of the field on the rhs. The important point is that the rhs will only involve velocities \dot{x}^μ and not higher derivatives wrt time such as \ddot{x}^μ , \dddot{x}^μ .

After many years of confusion, the problem of radiation reaction was finally settled by Herbert Spohn when he derived this equation using center manifold theory.

REMARK 41. Idea for Research: Is there a version of center manifold theory for PDEs? For example, the Skyrme model, or quantum corrections to GR all produce differential equations of higher order. Can we remove the runaway solutions of those theories this way? The equations obtained this way are dissipative, so don't have a canonical formalism. Is there a modified version (complex hamiltonian?) of canonical formalism that works for them?

The Feynman-Kac Formula

Feynman, following some ruminations of Dirac, discovered a formulation of quantum mechanics in terms of integrals over the paths of particles. This has turned out to be the most convenient way to think of relativistic quantum theories. The mathematics needed is the integral calculus on function spaces. At present we have a rigorous mathematical theory only for the case of functions of a single variable. The main obstruction to a rigorous theory of integration over functions of several variables is the appearance of spurious divergences, which have to be removed by a mysterious process known as **renormalization**. Precursors to this idea are present as early as in Euler's theory of divergent series. Finding a formalism for quantum field theory free of these divergences remains one of the grand challenges of mathematics and of physics. We will start down this path by starting with quantum mechanics. Instead of Feynman's approach we will follow later work of Kac which is a bit easier: it deals with real rather than complex integrands. It relates quantum mechanics to the Wiener integral, the only well-understood integration theory on function spaces.

14.1. The Heat Kernel of The Schrodinger Operator

We have already seen that solving the spectral problem for a Schrodinger operator

$$\hat{H} = -\frac{1}{2}\partial^2 + V$$

is equivalent to finding the heat kernel. That is, the solution to the PDE

$$\frac{\partial h_t(x, y)}{\partial t} = \frac{1}{2}\partial^2 h_t - V(x)h_t, \quad \lim_{t \rightarrow 0} h_t(x, y) = \delta(x, y)$$

If the spectrum is pure-point (i.e., just eigenvalues of finite multiplicity)

$$h_t(x, y) = \sum_n e^{-E_n t} \sum_a \psi_{na}^*(x) \psi_{na}(y)$$

where

$$\hat{H}\psi_{na} = E_n \psi_{na}$$

and a labels the degeneracy of the eigenvalue E_n . Also we know that this is the kernel of the operator $\hat{h}_t = e^{-\hat{H}t}$

$$h_t(x, y) = \langle x | \hat{h}_t | y \rangle$$

In the simplest case $V(x) = 0$ we have

$$\langle x | e^{t\frac{1}{2}\partial^2} | y \rangle = \frac{e^{-\frac{(x-y)^2}{2t}}}{[2\pi t]^{\frac{n}{2}}}$$

On the other hand,

$$\langle x | e^{-tV} | y \rangle = e^{-tV(x)} \delta(x, y)$$

Thus, each term in the hamiltonian can be easily exponentiated. For operators,

$$e^{\hat{A}+\hat{B}} \neq e^{\hat{A}}e^{\hat{B}}$$

so we cannot directly use these formulas for the exponential of each term to deduce that of the hamiltonian. If the commutators are small,

$$e^{\hat{A}+\hat{B}} = e^{\hat{A}}e^{\hat{B}}e^{-\frac{1}{2}[\hat{A},\hat{B}]}+\dots$$

This can be verified by expanding both sides to the required order in the commutator: the ignored terms involve repeated commutators of higher order. Thus for small t we can hope that

$$e^{t\frac{1}{2}\partial^2-tV} \approx e^{\frac{t}{2}\partial^2}e^{-tV} [1 + O(t^2)]$$

In terms of kernels

$$h_t(x, y) = \frac{e^{-\frac{(x-y)^2}{2t}}}{[2\pi t]} e^{-tV(y)} [1 + O(t^2)]$$

The idea of the Feynman-Kac formula is to break up the time interval into small pieces, for each of which the above approximation can be used.

14.2. Subdividing Time

Since

$$\hat{h}_t = e^{-t\hat{H}}$$

we have

$$\hat{h}_t = \hat{h}_{t-t_1}\hat{h}_{t_1}, \quad t > t_1 > 0$$

In terms of Kernels,

$$h_t(x, y) = \int h_{t-t_1}(x, x_1)h_{t_1}(x_1, y)dx_1$$

We can divide the interval $[0, t]$ any number of times

$$\hat{h}_t = \hat{h}_{t-t_N} \cdots \hat{h}_{t_2-t_1}\hat{h}_{t_1}, \quad t > t_N \cdots t_2 > t_1 > 0$$

to get

$$h_t(x, y) = \int h_{t-t_N}(x, x_N) \cdots h_{t_2-t_1}(x_2, x_1)h_{t-t_1}(x, x_1)h_{t_1}(x_1, y)dx_1 \cdots dx_N$$

This looks more natural if we change notation slightly and call $y = x_0$

$$h_t(x, x_0) = \int h_{t-t_N}(x, x_N) \cdots h_{t_2-t_1}(x_2, x_1) h_{t-t_1}(x, x_1) h_{t_1}(x_1, x_0) dx_1 \cdots dx_N$$

Then x_0 can be thought of as the position of the particle at time 0, x_1 that at time t_1 and so on, and finally x is the position at time t .

$$h_t(x, x_0) = \int \prod_{k=1}^{N+1} h_{t_k-t_{k-1}}(x_k, x_{k-1}) dx_k$$

with the understanding that

$$t_{N+1} = t, \quad t_0 = 0.$$

For N big enough, the intervals $t_k - t_{k-1}$ will become small enough that we can approximate

$$h_{t_k-t_{k-1}}(x_k, x_{k-1}) \approx e^{-\frac{1}{2} \frac{(x_k - x_{k-1})^2}{t_k - t_{k-1}} - [t_k - t_{k-1}] V(x_{k-1})} \frac{1}{[2\pi(t_k - t_{k-1})]^{\frac{N}{2}}} [1 + O([t_k - t_{k-1}]^2)]$$

Thus we get

$$h_t(x, x_0) = \int e^{-\left[\frac{1}{2} \sum_{k=1}^N \frac{(x_k - x_{k-1})^2}{t_k - t_{k-1}} + \sum_{k=1}^N [t_k - t_{k-1}] V(x_{k-1})\right]} [1 + O([t_k - t_{k-1}]^2)] \prod_{k=1}^N \frac{dx_k}{[2\pi(t_k - t_{k-1})]^{\frac{N}{2}}}$$

In the limit as $N \rightarrow \infty$, the quantity in the exponential tends to something simple:

$$-\left[\frac{1}{2} \sum_{k=1}^N \frac{(x_k - x_{k-1})^2}{t_k - t_{k-1}} + \sum_{k=1}^N [t_k - t_{k-1}] V(x_{k-1})\right] \rightarrow S[x] = -\int_0^t \left[\frac{1}{2} \dot{x}^2(s) + V(x(s))\right] ds$$

This quantity $S[x]$ depends on a path $x(s)$ connecting x_0 at time zero to x at time t . It is called the **action**. (The sign of the kinetic energy is not what you are used to in mechanics. That is because we are solving Schrodinger's equation in imaginary time. This is easier technically than the more physical case of real time.)

In this limit we are integrating over all possible paths that start at x_0 at time 0 and end at x at time t . Just as the limit of a sum can be thought of as an integral,

$$\sum_{k=1}^N f(t_k) [t_k - t_{k-1}] \rightarrow \int_0^t f(s) ds$$

a limit of such integrals can be thought of as a new kind of integral on paths: a **path integral**.

$$h_t(x, x_0) = \int_{\substack{x(0) = x_0 \\ x(t) = x}} e^{S[x]} \mathcal{D}[x]$$

The symbol $\mathcal{D}[x]$ is to be thought as the limit of $\prod_{k=1}^N \frac{dx_k}{[2\pi(t_k - t_{k-1})]^{\frac{N}{2}}}$ as $N \rightarrow \infty$. But it is not to be taken literally, just as the derivative $\frac{df}{dx}$ of calculus is not literally the ratio of two infinitesimal quantities. The proper definition of the path integral,

just like the $\epsilon - \delta$ definition of the derivative, is a subtle piece of analysis. We will take that up later. For now we are just developing the physical intuition needed for that construction.

This is Feynman's variational principle: if we integrate over all paths starting at x_0 at time zero and ending at x at time t , the weight of each path being given by its action as $e^{-S[x]}$, we get the heat kernel of the Schrodinger equation. A moments thought will show that if we had not set $\hbar = 1$ we would have obtained

$$h_t(x, x_0) = \int_{\substack{x(0) = x_0 \\ x(t) = x}} e^{\frac{1}{\hbar} S[x]} \mathcal{D}[x]$$

14.3. Laplace's Method for Path Integrals

For small \hbar we should expect that the path integral is dominated by the path of largest action: the variational principle of classical mechanics follows as an approximation from Feynman's variational principle of quantum mechanics. (Again we get the path of largest rather than least action because our sign conventions differ from that in mechanics. The idea is the same, though.) Moreover by expanding around this maximum of the integrand, we should be able to get an asymptotic expansion in powers of \hbar . This is the most physically clear way of deriving the semi-classical approximation of quantum mechanics. To put this through in practice, we will have to learn how to evaluate (and define) Gaussian integrals over the space of paths. The techniques needed were developed by Wiener, as part of his mathematical formulation of Einstein's theory of Brownian motion.

The Gaussian Integral

We are interested in developing a theory of integration over an infinite number of variables. The first instinct might be to define some generalization of the Lebesgue measure $d\phi$ to the infinite dimensional case. But that turns out to be impossible: there is no translation invariant measure except in finite dimensions. However, the Gaussian measure given by a positive matrix K

$$e^{-\frac{1}{2}\phi^T K \phi} d\phi$$

does have a generalization to the infinite dimensional case. In the limit, K can be a positive operator, such as the laplacian or a Schrodinger operator. To understand this, we start by reviewing the finite dimensional case. The most natural language to use is that of probability distributions.

15.1. The Normal Distribution

By far the common probability distribution in science is the normal distribution. Given a mean μ and variance G , the probability that such a random variable will take values in some interval $[a, b]$ is

$$P(a \leq \phi < b) = \int_a^b e^{-\frac{1}{2}G^{-1}(\phi-\mu)^2} \frac{d\phi}{\sqrt{2\pi G}}$$

That this is a positive quantity is obvious; the identity

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx = \sqrt{2\pi}$$

proves that as $a \rightarrow -\infty$ and $b \rightarrow \infty$ this probability tends to one, as it should.

A basic result of statistics is that the average of a large number of independent random variables (each with finite variance) tends to such a distribution. For example, if the measurement of a physical quantity is subject to a large number of independent errors, the measured value will be normally distributed. This is also called the Gaussian distribution, after Karl Frederick Gauss, one of the most respected mathematicians of all time. Much of asymptotic analysis is centered on Gaussian integrals, which are the average values with respect to this distribution. It is obvious by symmetry around the point μ that

$$\langle \phi \rangle = \int_{-\infty}^{\infty} \phi e^{-\frac{1}{2}\frac{(\phi-\mu)^2}{\sigma^2}} \frac{dx}{\sqrt{2\pi G}} = \mu$$

By adding a constant and multiplying by another constant, we can bring any normal random variable to a standard normal variable with zero mean and unit standard deviation. Using the identity

$$\frac{\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} x^2 dx}{\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} dx} = 1$$

we can see that

$$\text{Var}(\phi) = \langle \phi^2 \rangle - \langle \phi \rangle^2 = G.$$

Thus μ and G are indeed the mean and variance of the random variable. The probability density has a familiar bell-shape.

15.2. The Characteristic Function

For any probability distribution, expectation values such as $\langle \phi^n \rangle$ are called its **moments**. The moments are determined by their generating function

$$Z_\phi(J) = \sum_{n=0}^{\infty} \frac{J^n}{n!} \langle \phi^n \rangle$$

The idea is that if we know $Z(J)$ we can find the moments by expanding it around the origin in a Taylor series and reading off the coefficients. A moments thought will show that $Z(0) = 1$. And that $Z[iJ]$ is the Fourier transform of the probability density:

$$Z_\phi(iJ) = \langle e^{iJ\phi} \rangle = \int e^{iJ\phi} p(\phi) dx\phi.$$

By inverting the Fourier transform, we can recover the probability density function:

$$p(\phi) = \int Z(iJ) e^{-iJ\phi} \frac{dJ}{2\pi}$$

Thus $Z(J)$ determines the p.d.f., which is why it is also called the characteristic function. The effect of adding a constant to the random variable is to shift $Z(J)$ by a multiplicative factor:

$$Z_{\phi+a}(J) = e^J Z_\phi(J)$$

The effect of multiplying the random variable by a constant is to multiply J by a constant as well:

$$Z_{a\phi}(J) = Z_\phi(aJ)$$

By evaluating the integral (“completing the square”), the characteristic function of a Gaussian of mean μ and variance G is found to be

$$Z_\phi(J) = e^{\mu J} e^{-\frac{1}{2} G J^2}.$$

15.3. The Expectation Values of a Multi-variate Gaussian

Often we will need to deal with vector-valued random variables: the position of a molecule of a gas for example. The Gaussian has a generalization to this multi-dimensional case: the component $\phi \cdot J$ of the random variable ϕ in some direction J must be a Gaussian. Since $\langle \phi \cdot (J + J') \rangle = \langle \phi \cdot J \rangle + \langle \phi \cdot J' \rangle$, there must be a vector μ such that

$$\langle \phi \cdot J \rangle = \mu \cdot J,$$

This vector μ is the mean of ϕ . Similarly, the variance

$$\text{Var}(\phi \cdot J)$$

is a quadratic function of J . So there must be a positive matrix G such that

$$\langle (\phi \cdot J)^2 \rangle - \langle \phi \cdot J \rangle^2 = J^T G J.$$

This matrix G is the variance of ϕ . Using the result for a single Gaussian random variable,

$$\langle e^{\phi \cdot J} \rangle = e^{\mu \cdot J} e^{-\frac{1}{2} J^T G J}$$

so that the mean and variance determine all the other moments: they are determined by expanding in powers of J .

It is useful to think of $\phi = (\phi^1, \dots, \phi^n)$ in terms of its components in some Cartesian co-ordinate system

$$\langle \phi^i \rangle = \mu^i, \quad \langle \phi^i \phi^j \rangle - \langle \phi^i \rangle \langle \phi^j \rangle = G^{ij}.$$

Then the variance matrix G^{ij} is symmetric and positive in the sense that

$$J^T G J = G^{ij} J_i J_j \geq 0$$

with equality only in the case $J = 0$.

By expanding the characteristic function in a power series, we can get the expectation values of products of the components, such as $\langle \phi^i \phi^j \phi^k \phi^l \rangle$. The explicit formula is quite useful and is called the Wick expansion in quantum field theory.

PROPOSITION 42. (The Wick Expansion) For a Gaussian random variable of mean μ and variance G , the expectation value

$$\langle \phi^{i_1} \dots \phi^{i_n} \rangle = \sum_{\text{pairings}} \prod_{\text{unpaired } j} \mu^j \prod_{\text{pairs } a,b} G^{i_a i_b}$$

is given by a sum over pairings: the contribution of each pair is the corresponding matrix element of G , and each unpaired component contributes the corresponding component of the mean.

The idea of a pairing is best described graphically. Put down points ("vertices") labelled by the indices i_1, \dots, i_n . For each graph we can get by connecting pairs of these vertices by lines, we get a contribution to the above sum.

In the simplest case of zero pairs, we get just the product of the components of the mean. The next set of terms correspond to one factor of G and $n - 2$ factors

of μ . After that we would get the graphs with two pairs of vertices connected by lines and $n - 4$ singletons and so on:

$$\begin{aligned} \langle \phi^{i_1} \dots \phi^{i_n} \rangle &= \mu^{i_1} \mu^{i_2} \dots \mu^{i_n} + G^{i_1 i_2} \mu^{i_3} \mu^{i_4} \dots \mu^{i_n} + \text{permutations} \\ &+ G^{i_1 i_2} G^{i_3 i_4} \mu^{i_5} \dots \mu^{i_n} + \text{permutations} \end{aligned}$$

For example,

$$\langle \phi^i \rangle = \mu^i$$

There is just one graph with one vertex. Then

$$\langle \phi^i \phi^j \rangle = G^{ij} + \mu^i \mu^j$$

which agrees with the definition of the Gaussian. Then

$$\begin{aligned} \langle \phi^i \phi^j \phi^k \rangle &= \mu^i \mu^j \mu^k + G^{ij} \mu^k + G^{ik} \mu^j + G^{jk} \mu^i \\ \langle \phi^i \phi^j \phi^k \phi^l \rangle &= \mu^i \mu^j \mu^k \mu^l + \\ &G^{ij} \mu^k \mu^l + G^{ik} \mu^j \mu^l + G^{il} \mu^j \mu^k + G^{jk} \mu^i \mu^l + G^{jl} \mu^i \mu^k + G^{kl} \mu^i \mu^j \\ &+ G^{ij} G^{kl} + G^{ik} G^{jl} + G^{il} G^{jk} \end{aligned}$$

and so on. The proof is simply to expand the characteristic function above in a series and collect coefficients of J 's.

15.4. The Multi-Dimensional Gaussian Measure

We can calculate the expectation values of any function (that can be approximated by polynomials) using the moments obtained by the Wick expansion.¹ Still there is interest in understanding the measure that generates these expectation values:

$$\langle f(\phi) \rangle = \int f(\phi) p(\phi) d\phi.$$

We can get it by inverting the Fourier transform. The answer should be

$$p(\phi) d\phi = e^{-\frac{1}{2} \phi^T K \phi} \frac{d\phi}{Z}$$

where Z is a normalization factor determined the condition

$$\begin{aligned} \int p(\phi) d\phi &= 1 \\ Z &= \int e^{-\frac{1}{2} \phi^T K \phi} d\phi. \end{aligned}$$

In the special case where K is diagonal, the probability density splits as a product

¹There is no loss of generality in setting the mean $\mu = 0$: we can always put it back by a translation of the variable by a constant vector.

$$p(\phi)d\phi = \prod_{j=1}^n e^{-\frac{1}{2}\phi_j k_j \phi_j} \frac{d\phi_j}{Z_j}$$

In this case

$$\langle \phi^i \phi^j \rangle = \delta^{ij} k_j^{-1}$$

and

$$Z_j = \sqrt{2\pi k_j^{-1}}.$$

But any positive matrix can be diagonalized by an orthogonal transformation of determinant one. Thus in general

$$\langle \phi^i \phi^j \rangle = [K^{-1}]^{ij}$$

$$Z = \sqrt{\det [2\pi K^{-1}]}.$$

We see that the variance G is the inverse of K .

Laplace's Method For Multi-Dimensional Integrals

We return to a theme from the beginning of the course. The evaluation of an integral

$$Z(g) = \int e^{\frac{1}{g}S(\phi)} d^N \phi$$

over a finite number N of variables. Suppose the integrand has a unique maximum at some point φ in the interior of the domain of integration. In the limit of $g \rightarrow 0$ we would expect the integral to be dominated by the contribution of the neighborhood of this maximum. We expand

$$S(\phi) = S(\varphi) - \frac{1}{2}(\phi - \varphi)^T G^{-1}(\phi - \varphi) + \sum_{k=3}^{\infty} S_{i_1 \dots i_k}(\phi - \varphi)^{i_1}(\phi - \varphi)^{i_2} \dots (\phi - \varphi)^{i_k}$$

where

$$-[G^{-1}]_{ij} = \left[\frac{\partial^2 S}{\partial \phi^i \partial \phi^j} \right]_{\phi=\varphi}$$

$$S_{i_1 \dots i_k} = \frac{1}{k!} \left[\frac{\partial^k S}{\partial \phi^{i_1} \dots \partial \phi^{i_k}} \right]_{\phi=\varphi}$$

Since we are at a maximum, the second derivative is a negative matrix; so G is a positive matrix. Now we change variables to

$$\chi = \frac{1}{\sqrt{g}}[\phi - \varphi]$$

and substitute the expansion in to the integral to get

$$Z(g) = g^{\frac{N}{2}} e^{\frac{1}{g}S(\varphi)} \int e^{\left[-\frac{1}{2}\chi^T G \chi + \sum_{k=3}^{\infty} g^{\frac{k}{2}-1} S_{i_1 \dots i_k} \chi^{i_1} \dots \chi^{i_k} \right]} d^N \chi$$

Expanding the exponential

$$Z(g) = e^{\frac{1}{g}S(\varphi)} \sqrt{\det[2\pi g G]} \sum_{n_k=0}^{\infty} \frac{g^{\sum_{k=3}^{\infty} (\frac{k}{2}-1)n_k}}{n_k!} \frac{\int e^{-\frac{1}{2}\chi^T G^{-1}\chi} [S_{i_1 \dots i_k} \chi^{i_1} \dots \chi^{i_k}]^{n_k} d^N \chi}{\int e^{-\frac{1}{2}\chi^T G^{-1}\chi} d^N \chi}$$

This can be evaluated by a combinatorial method, which has a graphical interpretation as well.

- (1) Choose some even integer $K > 3$ (twice the number of edges of the graph)

- (2) For each partition of $K = \sum_{k=3} kn_k$ with $n_k > 0$, there are n_k vertices, each with k edges coming out. Label these edges with indices taking values $1 \dots N$.
- (3) Connect edges to each other in all possible ways; it is possible to connect an edge to another from the same vertex or to one from a different vertex. There should be no unpaired edge (recall that K is even).
- (4) To each edge labelled by i and j associate a factor G^{ij} . To each vertex of order k associate a factor $S_{i_1 \dots i_k}$. These indices will be summed over the range $1 \dots N$. Divide by a factor $\frac{1}{n_k!}$ where n_k is the number of times a vertex of a given order k appears.
- (5) Sum over all such partitions.

These diagrams are called Feynman diagrams.

16.1. Example: $K = 4$

Let us consider the lowest order terms in this expansion. The only partition is $4 = 4 \times 1$. The only graph is the "figure eight" graph.

$$Z(g) = e^{\frac{1}{g}S(\varphi)} \sqrt{\det[2\pi gG]} [1 + 3gS_{i_1 i_2 i_3 i_4} G^{i_1 i_2} G^{i_3 i_4} + \dots]$$

16.2. $K = 6$

The only partition is $6 = 3 \times 2$ which gives rise to two graphs:

$$Z(g) = e^{\frac{1}{g}S(\varphi)} \sqrt{\det[2\pi gG]} [$$

$$1 + 3gS_{i_1 i_2 i_3 i_4} G^{i_1 i_2} G^{i_3 i_4} + \frac{g}{2!} S_{i_1 i_2 i_3} S_{i_4 j_1 j_2} [3 \times 2G^{i_1 j_1} G^{i_2 j_2} G^{i_3 j_3} + 3^2 G^{i_1 i_2} G^{i_3 j_3} G^{j_1 j_2}] + O(g^2)]$$

In the last term there are 3^2 ways of choosing an index from the first vertex to pair with one from the second vertex: the remaining contraction is determined once this is chosen. In the first, there are three ways of choosing the index to pair with the first vertex, 2 ways to choose a mate for the second and then the last is determined. Note that $3^2 + 3 \times 2 = 15 = (6 - 1)!!$ is the number of terms in the six point function of the Gaussian (Wick's theorem).

EXERCISE 43. Enumerate all the partitions and Feynman diagrams for $K = 8$. Find the corresponding terms in the Laplace expansion.

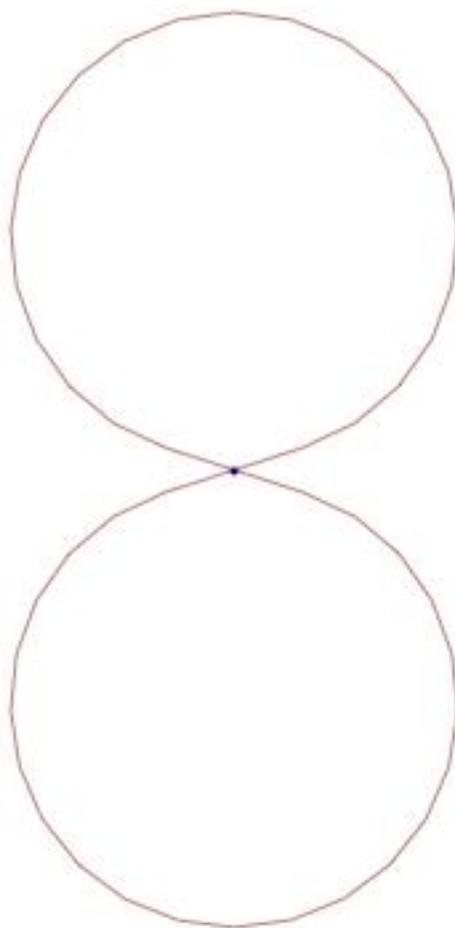


FIGURE 16.1.1.



FIGURE 16.2.1.



FIGURE 16.2.2.

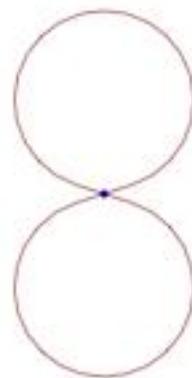
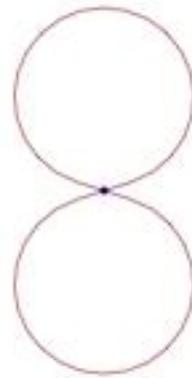


FIGURE 16.2.3.

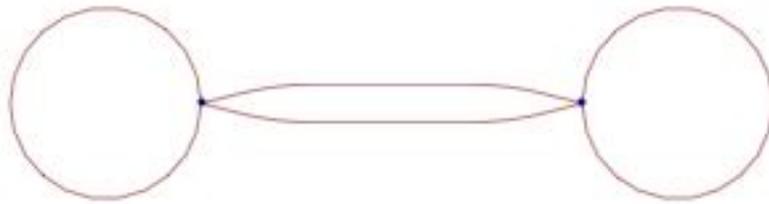


FIGURE 16.2.4.



FIGURE 16.2.5.

The Wiener Measure

There is no translation invariant measure in infinite dimensions. The closest we can come to it is the Gaussian measure. A Gaussian measure on the space of continuous functions of one variable is of great interest. It was constructed by Wiener as a mathematical realization of Einstein's theory of Brownian motion.

17.1. Lebesgue Decomposition

Even for a continuous real random variable, the probability of a particular value is zero; the correct question to ask is, what is the probability that the variable takes values within some semi-open interval $[a, b)$. For such a variable, it will be given by

$$P(\xi \in (a, b]) = \Phi(b) - \Phi(a)$$

where $\Phi(x) = \text{Prob}(\xi \leq x)$ is the cumulative distribution function. The expectation value of any function of the random variable can now be calculated as an integral

$$\langle f \rangle = \int f(x) d\Phi(x)$$

This function $\Phi(x)$ is non-decreasing; it has jump discontinuities at the discrete values of the random variable, and is the integral of some continuous function (the probability density function) for the continuous range of values of ξ .

If Φ has a derivative $p(x) = \frac{d\Phi}{dx}$ we can write this as

$$\langle f \rangle = \int f(x) p(x) dx$$

But in many physically interesting cases we may get a sum as well

$$\langle f \rangle = \sum_n f(x_k) p_k + \int f(x) p(x) dx$$

Typically, $\Phi(x)$ does not have a derivative at all: in addition to jump discontinuities, it may have a piece that is continuous but not differentiable. (Singular continuous spectrum). An example is the Cantor function, which maps the middle-third Cantor set onto the unit interval. Lebesgue showed that there is a unique decomposition of an expectation value

$$\langle f \rangle = \sum_n f(x_k) p_k + \int f(x) p(x) dx + \int f(x) d\Phi_{\text{sing}}(x)$$

into a discrete part, a continuous part and a singular continuous part. Here, p_k are positive numbers with $\sum_k p_k \leq 1$ and $p(x)$ is a continuous positive function,

and $\Phi_{\text{sing}}(x)$ is a continuous function that is non-decreasing, but which is nowhere differentiable.

17.2. The Space of Continuous Functions

The typical continuous function $\phi : [0, T] \rightarrow R$ is not differentiable. That there are functions that are continuous everywhere but differentiable nowhere was a surprising discovery of nineteenth century analysis. They arose as Fourier sums of the type

$$\phi(t) = \sum_n \phi_n e^{int}$$

If $\sum_n |\phi_n|$ is convergent, this sum is uniformly convergent and hence defines a continuous function. But typically $\sum_n n|\phi_n|$ will diverge, so the derivative $\phi'(t)$ is not continuous. (e.g., $\phi_n = \frac{1}{n^2}$). The set of continuous functions admits a norm (a length)

$$|\phi| = \sum_{t \in [0, T]} |\phi(t)|$$

By completing with this norm, we get a Banach space.

17.3. Probability of a Path

When a particle is executing Brownian motion, its direction changes all the time from hitting other particles in the medium. So its path is not differentiable. But the path is continuous. What is the probability of a path? Of course, this is an ill-define question.

More precisely, what is the probability that at times

$$t_1 < t_2 \cdots < t_n$$

the path will fall within windows

$$\phi(t_k) \in [a_k, b_k)$$

For Brownian motion, physical considerations suggest this is a Gaussian: the position of the particle is determined as the sum of a large number of independent increments. There are so many collisions that the increments $\phi(t_k) - \phi(t_{k-1})$ are independent random variables. The process is reflection invariant, meaning that the mean increment is zero: there is no drift in any direction. It is stationary, meaning that the variance of the increment only depends on the difference $t_k - t_{k-1}$. with all this in mind we get

$$\text{Prob}(\phi(t_k) \in [a_k, b_k)) = \int_{a_k}^{b_k} \prod_{k=1}^n d\phi_k \prod_{k=2}^n \frac{e^{-\frac{(\phi_k - \phi_{k-1})^2}{2\sigma(t_k - t_{k-1})}}}{[2\pi\sigma(t_k - t_{k-1})]}$$

If we were to omit any subset of the conditions on $\phi(t_k)$ this formula would remain unchanged: it is invariant under “integrating out” variables. This formula defines the Wiener measure.

17.4. The Harmonic Oscillator

We saw by the Feynman-Kac formula that

$$\text{tr} e^{-\hat{H}t} = \int_{\phi(0)=\phi(t)} e^{-\frac{1}{2} \int \dot{\phi}^2(s) ds - \int V(\phi(s)) ds} \mathcal{D}\phi$$

In the particular case $V(\phi) = \frac{1}{2}\phi^2$ this is a Gaussian integral. Recall that in finite dimensions

$$\int d\phi e^{-\frac{1}{2}(\phi, A\phi)} = \sqrt{\det [2\pi A^{-1}]}$$

In our case, we can interpret the domain of integration as the vector space of periodic functions of period t . Then

$$A = -\frac{d^2}{ds^2} + 1$$

On the above vector space the eigenfunctions of this operator are

$$1, \sin \left[\frac{2\pi s}{t} k \right], \cos \left[\frac{2\pi s}{t} k \right], \quad k = 1, 2, \dots$$

with eigenvalues

$$1, 1 + \left[\frac{2\pi}{t} \right]^2 k^2, 1 + \left[\frac{2\pi}{t} \right]^2 k^2, \quad k = 1, 2, \dots$$

Thus

$$\begin{aligned} \sqrt{\det A} &= \prod_{k=1}^{\infty} \left[1 + \left[\frac{2\pi}{t} \right]^2 k^2 \right] \\ &= \prod_{k=1}^{\infty} \left[\left[\frac{2\pi}{t} \right]^2 k^2 \right] \prod_{k=1}^{\infty} \left[1 + \left[\frac{t}{2\pi k} \right]^2 \right] \end{aligned}$$

The first product diverges. But it can be given a meaning using zeta function regularization. The second can be deduced from

$$\sinh z = z \prod_{k=1}^{\infty} \left[1 + \frac{z^2}{\pi^2 k^2} \right]$$

so that

$$\prod_{k=1}^{\infty} \left[1 + \left[\frac{t}{2\pi k} \right]^2 \right] = \frac{\sinh \frac{t}{2}}{\frac{t}{2}}$$

Now,

$$\prod_{k=1}^{\infty} k = e^{\sum_{k=1}^{\infty} \log k} = e^{-\zeta'(0)} = e^{\frac{1}{2} \log[2\pi]} = \sqrt{2\pi}$$

$$\prod_{k=1}^{\infty} \rho = \rho^{\zeta(0)} = \rho^{-\frac{1}{2}}$$

so that

$$\left(\prod_{k=1} \left[\frac{2\pi}{t} k \right] \right)^2 = \frac{t}{2\pi} 2\pi = t$$

Thus

$$\sqrt{\det \frac{A}{2\pi}} = 2 \sinh \frac{t}{2}.$$

$$\text{tr} e^{-\hat{H}t} = \int_{\phi(0)=\phi(t)} e^{-\frac{1}{2} \int \dot{\phi}^2(s) ds - \int V(\phi(s)) ds} \mathcal{D}\phi = \frac{1}{2 \sinh \frac{t}{2}}.$$

The usual spectrum $E_n = n + \frac{1}{2}$, $n = 0, 1, 2, \dots$ follows from this.