Fun with Calculus

Stefan Boettcher

(Dated: 18th November 2005)

Contents

I. FUNDAMENTALS

A. Sums

We define

$$\sum_{n=n_0}^{n_1} a_n = a_{n_0} + a_{n_0+1} + a_{n_0+2} + \ldots + a_{n_1-1} + a_{n_1} \tag{1}$$

as the "sum" of the numbers a_n for $n$ranging over all integers from (including) n_0 to n_1. Typically, $n_0 \leq n_1$are integers; if either is a real number, the sum starts at the next-larger integer to n_0 and ends at the next-smaller integer to n_1. If $n_0 > n_1$, the sum evaluates to 0by definition. Note that n is a "dummy" index, which could be replaced by any other symbol, i. e. $\sum_{n=n_0}^{n_1} a_n = \sum_{\heartsuit=n_0}^{n_1} a_\heartsuit$.

EXAMPLES:

1. With $a_n = n^3$:$\sum_{n=-2}^{1} n^3 = -8 - 1 + 0 + 1$,
2. with $a_i = i$: $\sum_{i=0}^{n} i = 0 + 1 + 2 + \ldots + n = \frac{n(n+1)}{2}$,
3. with $a_i = x^i$(Geometric Series):

$$G_n(x) = \sum_{i=0}^{n} x^i = 1 + x + x^2 + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x}. \tag{2}$$

There are some basic operations for sums (similar to integration, but more limited):

1. Distributive Law: for any expression C, $\sum_{n=n_0}^{n_1} C a_n = C \sum_{n=n_0}^{n_1} a_n$.
2. Extraction: $\sum_{n=n_0}^{n_1} a_n = a_{n_0} + \sum_{n=n_0+1}^{n_1} a_n$, $\qquad \sum_{n=n_0}^{n_1} a_n = \sum_{n=n_0}^{n_1-1} a_n + a_{n_1}$.
3. Shift: $n \to n' = n + \Delta n$ results in $\sum_{n=n_0}^{n_1} a_n = \sum_{n'=n_0+\Delta n}^{n_1+\Delta n} a_{n'-\Delta n}$.
4. Reflection: $n \to n' = \Delta n - n$ results in $\sum_{n=n_0}^{n_1} a_n = \sum_{n'=\Delta n - n_1}^{\Delta n - n_0} a_{\Delta n - n'}$ (note interchange of upper and lower limit!).

Infinite sums (where one of the limits is $\pm\infty$) have the restriction that terms must remain in order, unless the sum is absolutely convergent.

EXAMPLES:

- Infinite Geometric Series:

$$\begin{aligned} G_\infty(x) &= \sum_{i=0}^{\infty} x^i = 1 + x + x^2 + \ldots \\ &= 1 + \sum_{i=1}^{\infty} x^i \\ &= 1 + \sum_{i'=0}^{\infty} x^{i'+1} \\ &= 1 + x \sum_{i'=0}^{\infty} x^{i'} \\ &= 1 + x G_\infty(x), \end{aligned} \tag{3}$$

hence,

$$G_\infty(x) = \sum_{i=0}^{\infty} x^i = \frac{1}{1-x}. \tag{4}$$

- Riemann Integral:

$$\lim_{N\to\infty} \sum_{i=1}^{N} \frac{b-a}{N} f\left(a + i\frac{b-a}{N}\right) = \int_a^b dx\, f(x) \tag{5}$$

EXERCISE:

1. Derive the result for the geometric series, $G_n(x)$.

2. How would you order the first 10 terms in the harmonic series,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n},$$

such that the sum can converge to 0, or 1, or...

B. Taylor-Series:

If $f(x)$is sufficiently regular near some point x, i. e. all derivatives $f^{(n)}(x)$ for all $n >$ 0exist, we can represent the function locally (near x) by

$$f(x + \Delta x) = a_0 + a_1\Delta x + a_2\Delta x^2 + a_3\Delta x^3 + a_4\Delta x^4 + \ldots. \tag{6}$$

Then, differentiating this expression n times and setting $\Delta x = 0$:

$$f(x) = a_0, \quad f'(x) = a_1, \quad f''(x) = 2a_2 \quad f^{(3)}(x) = 6a_3, \ldots, f^{(n)}(x) = n!a_n.$$

Hence,

$$f(x + \Delta x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} \Delta x^n. \tag{7}$$

EXAMPLES:

The exponential function $f(x) = e^x$satisfies the differential equation $f'(x) = f(x)$, $f(0) = 1$, so at $x = 0$we get the *MacLaurin series* (a Taylor series at $x = 0$):

$$f(\Delta x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \Delta x^n, \tag{8}$$

in particular with $f(0) = f'(0) = f''(0) = \ldots = 1$,

$$e^{\Delta x} = \sum_{n=0}^{\infty} \frac{1}{n!} \Delta x^n. \tag{9}$$

Note, that using "∂_x" as the instruction ("operator") of differentiating everything to the right by x, we can write

$$f(x + \Delta x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x)}{n!} \Delta x^n = \sum_{n=0}^{\infty} \frac{\Delta x^n}{n!} \partial_x^n f(x) = \left[\sum_{n=0}^{\infty} \frac{(\Delta x \partial_x)^n}{n!}\right] f(x) = e^{\Delta x \partial_x} f(x). \tag{10}$$

In this sense, one may interpret a Taylor series as an operator that "generates" a translation in the argument of a function from $f(x) \to f(x + \Delta x)$. Clearly, the (transcendental!) exponential of an operator only makes sense in terms of the infinite series that defines it. Hence, the above relation serves no purpose except as an efficient, suggestive shorthand.

For the trigonometric functions $f(x) = \sin(x), \cos(x)$, we know that $f^{(n+2)}(x) = -f^{(n)}(x)$ for all $n \geq 0$. Since $\sin(0) = 0$and $\cos'(0) = 0$, all terms in the Taylor series at x=0 vanish for even n for $\sin(x)$, and all odd n for $\cos(x)$. Thus,

$$f(\Delta x) = \sin(\Delta x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \Delta x^n = \sum_{n=0; n\, odd}^{\infty} \frac{f^{(n)}(0)}{n!} \Delta x^n,$$

and substituting $n = 2i + 1$ to parameterize all odd terms (with $i = 0, 1, 2, \ldots$),

$$\sin(\Delta x) = \sum_{i=0}^{\infty} \frac{f^{(2i+1)}(0)}{(2i+1)!} \Delta x^{2i+1}.$$

Using $f^{(1)}(0) = -f^{(3)}(0) = +f^{(5)}(0) = \ldots = (-1)^i f^{(2i+1)}(0) = \ldots = 1$, it is finally

$$\sin(\Delta x) = \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i+1)!} \Delta x^{2i+1}. \tag{11}$$

EXERCISE:

1. Show that

$$\cos(\Delta x) = \sum_{i=0}^{\infty} \frac{(-1)^i}{(2i)!} \Delta x^{2i}. \tag{12}$$

PRODUCT OF TWO SERIES:

For any two series, it is

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{m=0}^{\infty} b_m\right) = \left(\sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_n\right) b_m\right) = \sum_{m=0}^{\infty}\left(\sum_{n=0}^{\infty} a_n b_m\right) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a_n b_m.$$

Since we are summing over all pairs of discrete indices (m, n), we can reorganize the terms (without rearranging the order of summation in m for fixed n, or in n for fixed m!) by setting $i = m + n$ $(0 \le i \le \infty)$ and $j = n$ $(0 \le j \le i)$ and get

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{m=0}^{\infty} b_m\right) = \sum_{i=0}^{\infty}\sum_{j=0}^{i} a_j b_{i-j}. \tag{13}$$

Note that, unlike for the double-sum above, the order of the two sums here matters because the inner sum depends on the index of the outer sum.

This result is particularly useful for finding Taylor series for the product of two functions, for which each Taylor series is know. With $a_n = A_n x^n$ and $b_m = B_m x^m$

$$\left(\sum_{n=0}^{\infty} A_n x^n\right)\left(\sum_{m=0}^{\infty} B_m x^m\right) = \sum_{i=0}^{\infty}\sum_{j=0}^{i} \left(A_j x^j\right)\left(B_{i-j} x^{i-j}\right) = \sum_{i=0}^{\infty}\sum_{j=0}^{i} A_j B_{i-j} x^{j+i-j}.$$

As $x^{j+i-j} = x^i$ does no longer depend on the summation index $j$of the inner sum, it is merely a factor that can be pulled in front of that sum:

$$\left(\sum_{n=0}^{\infty} A_n x^n\right)\left(\sum_{m=0}^{\infty} B_m x^m\right) = \sum_{i=0}^{\infty} x^i \left(\sum_{j=0}^{i} A_j B_{i-j}\right) = \sum_{i=0}^{\infty} c_i x^i \tag{14}$$

with $c_i = \sum_{j=0}^{i} A_j B_{i-j}$. Now, the ith coefficient in the series is a sum itself, but at least it is finite (and may even have a closed form).

EXAMPLES:

$$\frac{e^x}{1-x} = \left(\sum_{n=0}^{\infty} \frac{1}{n!} x^n\right)\left(\sum_{m=0}^{\infty} x^m\right) = \sum_{i=0}^{\infty} x^i \left(\sum_{j=0}^{i} \frac{1}{j!}\right)$$

with $A_n = \frac{1}{n!}$ and $B_m = 1$.

Sometimes, these formal manipulations lead to non-trivial relations:

$$\frac{1}{1-x} = \left(\frac{e^x}{1-x}\right) e^{-x} = \left[\sum_{n=0}^{\infty} x^n \left(\sum_{k=0}^{n} \frac{1}{k!}\right)\right] \left[\sum_{m=0}^{\infty} \frac{(-1)^m}{m!} x^m\right] = \sum_{i=0}^{\infty} x^i \left(\sum_{j=0}^{i} \frac{(-1)^{i-j}}{(i-j)!} \sum_{k=0}^{j} \frac{1}{k!}\right).$$

Since this expression must be the geometric series, we know that

$$\sum_{j=0}^{i} \sum_{k=0}^{j} \frac{(-1)^{i-j}}{k!(i-j)!} = 1, \qquad (i \geq 1).$$

EXERCISE:

1. Use the Taylor series for $e^{(p+q)x}$ and compare it term-by-term in powers of x with the product series of $e^{px}e^{qx}$to obtain the binomial series,

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}. \tag{15}$$

C. Complex Algebra:

ORIGIN: Desire to be able to solve general algebraic equations, such as quadratic equations:

$$x^2 + ax + b = 0,$$

completing the square:

$$x^2 + ax + \frac{a^2}{4} - \frac{a^2}{4} + b = 0,$$

then,

$$\left(x + \frac{a}{2}\right)^2 + b - \frac{a^2}{4} = 0. \tag{16}$$

Finally, solving for x:

$$x_{\pm} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b}. \tag{17}$$

Hence, the problem arises when the discriminate is $\frac{a^2}{4} - b < 0$. For example, when $a = -2$ and $b = 2$, then

$$x_{\pm} = 1 \pm \sqrt{-1}.$$

Amazingly, any algebraic relation like that can in principle be "solved" by simply introducing the "imaginary number"

$$i := \sqrt{-1} \tag{18}$$

and demanding that any expression can be written in "complex" form

$$z = a + ib, \tag{19}$$

where $a = Re\{z\}$ is called the "real" part of z, and $b = Im\{z\}$ is called the "imaginary" part of z, but both a, b are entirely real expressions. Note that

$$i^2 = -1, \quad i^3 = i\, i^2 = -i, \quad i^4 = i^2 i^2 = (-1)^2 = 1, \quad \frac{1}{i} = \frac{i}{i^2} = \frac{i}{-1} = -i,$$

and so on.

A useful definition is is the so-called "complex conjugate" of z, denoted by

$$z \longrightarrow z^* or\ \bar{z} = a - ib. \tag{20}$$

Complex conjugation essentially means to replace $i \to -i$ everywhere. So, while

$$z^2 = (a + ib)^2 = \left(a^2 - b^2\right) + i(2ab)$$

is also a complex number with $Re\{z^2\} = a^2 - b^2$ and $Im\{z^2\} = 2ab$, the product of

$$z\, z^* = (a + ib)(a - ib) = a^2 + b^2$$

is entirely real, i. e. $Im\{z\, z^*\} = 0$, and positive! Thus, it is justified to define

$$|z| = \sqrt{z\, z^*} \geq 0 \tag{21}$$

as the norm or "modulus" of any expression z.

Complex conjugates have many uses, for instance, to show that the inverse of z can also be written in complex form:

$$\frac{1}{z} = \frac{z^*}{z\, z^*} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}.$$

Further insight is provided, when we represent a complex number as a 2-component vector in the "complex plane", $z = (a, b)$, where the first component, $a = Re\{z\}$, measures the extent along the "real axis" (horizontal), and the second component, $b = Im\{z\}$, measures the extent along the "imaginary axis" (vertical). This plane geometry suggest a new, polar representation of a complex number as

$$z = r\cos(\phi) + ir\sin(\phi)$$

with $\{a = r\cos(\phi), b = r\sin(\phi)\}$ or conversely, $\{r = \sqrt{a^2+b^2} = |z|, \phi = arctan\left(\frac{b}{a}\right)\}$. This polar representation obtains its value from the following realization: Consider the Taylor series for the exponential wit imaginary argument,

$$e^{i\phi} = \sum_{n=0}^{\infty} \frac{i^n\phi^n}{n!} = \sum_{n=0;n\,even}^{\infty} \frac{i^n\phi^n}{n!} + \sum_{n=0;n\,odd}^{\infty} \frac{i^n\phi^n}{n!},$$

and substituting $n = 2k$ for even n, or $n = 2k+1$ for odd n, respectively, we get

$$e^{i\phi} = \sum_{k=0}^{\infty} \frac{i^{2k}\phi^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{i^{2k+1}\phi^{2k+1}}{(2k+1)!} = \sum_{k=0}^{\infty} \frac{(-1)^k\phi^{2k}}{(2k)!} + i\sum_{k=0}^{\infty} \frac{(-1)^k\phi^{2k+1}}{(2k+1)!}, \tag{22}$$

where we have used the fact that $i^{2k} = \left(i^2\right)^k = (-1)^k$ and $i^{2k+1} = i\,\left(i^2\right)^k = i\,(-1)^k$. Now, we can easily identify these two series as the MacLaurin series for $\sin\phi$ and $\cos\phi$:

$$e^{i\phi} = \cos(\phi) + i\sin(\phi), \tag{23}$$

This result, called "Euler's Formula", provides a very compact and suggestive form for complex expressions:

$$z = re^{i\phi}. \tag{24}$$

Euler's formula itself leads to a number of fascinating relations such as$e^{i\frac{\pi}{2}} = i, \quad e^{i\pi} = -1$, combining the imaginary number i with the irrational numbers e and π. Furthermore, it is for all integers $k = 0, \pm1, \pm2, \ldots$

$$e^{2\pi ik} = \cos(2\pi k) + i\sin(2\pi k) = 1.$$

Its use for the polar representation of a complex number also provides a geometric description of basic operations in (complex) algebra. For example, it is well-known that $\sqrt{1}$ has two roots (solutions), namely ±1, since $(\pm1)^2 = 1$. But what are the roots of $1^{\frac{1}{3}}$? One root clearly is 1 itself, but we would expect three roots. Is the root at 1 simply a triple root? The answer is: No! And to see this it is best to write "1" as a complex number in polar form, as above: $1 = e^{2\pi ik}$. Then

$$1^{\frac{1}{3}} = \left(e^{2\pi ik}\right)^{\frac{1}{3}} = e^{\frac{2k}{3}\pi i},$$

seemingly yielding an infinity of roots, one for each integer k. Now, for $k = 0$ we get the known root, 1. For $k = 1$ we obtain a new, non-trivial root, $e^{\frac{2}{3}\pi i} = \cos(\frac{2}{3}\pi) + i\sin(\frac{2}{3}\pi) = -\frac{1}{2} + i\,\frac{1}{2}\sqrt{3}$. It is straightforward to test that this is a cube-root of 1 via $\left(-\frac{1}{2} + i\,\frac{1}{2}\sqrt{3}\right)^3 = 1$. Also, for $k = 2$ we obtain another, non-trivial root, $e^{\frac{4}{3}\pi i} = \cos(\frac{4}{3}\pi) + i\sin(\frac{4}{3}\pi) = -\frac{1}{2} - i\,\frac{1}{2}\sqrt{3}$, the complex conjugate of the previous root. But for $k = 3$ we obtain $e^{\frac{6}{3}\pi i} = e^{2\pi i} = 1$ again, and similarly for any other integer k, we merely reproduce one of the 3 already known roots. This is a quite general result: taking a power corresponds to rotating a number around the origin in the complex plane. For instance, a real positive number sitting on the Re-axis at $\phi = 2\pi k$ taken to an integer power n rotates to $\phi = 2\pi kn$, i. e. it remains on the positive Re-axis. A number on the negative Re-axis, at $\phi = \pi + 2\pi k$ taken to an integer power n rotates to the positive Re-axis, if n is even, and to the negative Re-axis, if n is odd. In general, a number at ϕ is rotated to $n\phi \, mod \, 2\pi$. This gets complicated when the power is rational, say, n/m (as in the cube-root of 1 above: $n = 1, m = 3$). First, $\phi \to \phi' = n\phi$ uniquely, then $\phi' \to (n\phi/m + 2\pi k/m) \, mod \, 2\pi$, which leads to m distinct values. Now, when the power is irrational, α, then $\phi \to (\alpha\phi + 2\pi k\alpha) \, mod \, 2\pi$, which has infinitely many values for all integer values of k.

Finally, there are a few interesting relations derived form Euler's formula:

$$\sin(x) = \frac{e^{ix} - e^{-ix}}{2i}, \qquad \cos(x) = \frac{e^{ix} + e^{-ix}}{2},$$

and setting $x \to ix$ defines the "hyperbolic" functions

$$\sinh(x) = \frac{e^{x} - e^{-x}}{2} = -i\sin(ix), \qquad \cosh(x) = \frac{e^{x} + e^{-x}}{2} = \cos(ix).$$

EXERCISE:

1. Use these relations to find

$$\sum_{n=0}^{\infty} \cos(n) = \frac{1}{2}.$$

(Hint: Remember the geometric series!)

II. GAUSSIAN INTEGRALS

This chapter is short because it is so important! I. e. there is not too much to say here because we will return to its implications many times in the following chapters. The "Gaussian" function $e^{-x^2} = \exp(-x^2)$ is pervasive in many areas of applied mathematics. It does not have an elementary integral (which is simply tabulated as the "error function" $erf(x) = \int_{-\infty}^{x} d\xi\, e^{-\xi^2}$, named due to its use in statistics). Even the definite integral of the Gaussian is not trivial. One writes:

$$I = \int_{-\infty}^{\infty} dx\, e^{-x^2} \rightarrow I^2 = \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{\infty} dxdy\, \exp\left[-\left(x^2+y^2\right)\right], \tag{25}$$

and transforms to polar coordinates via $(x = r\cos\phi, y = r\sin\phi)$, with $dxdy \rightarrow rdrd\phi$, hence

$$I^2 = \int_{\phi=0}^{2\pi} \int_{r=0}^{\infty} rdrd\phi\, e^{-r^2} = \left[\int_0^{2\pi} d\phi\right] \left[\int_0^{\infty} dr\, re^{-r^2}\right]. \tag{26}$$

As these two integrals magically factorize, leaving the first integral trivial, $\int_0^{2\pi} d\phi = 2\pi$, and with the substitution $rdr = \frac{1}{2}d(r^2) = d\xi$,

$$I^2 = 2\pi\frac{1}{2}\int_{\xi=0}^{\infty} d\xi\, e^{-\xi} = \pi,$$

we have solved the Gaussian integral,

$$I = \int_{-\infty}^{\infty} dx\, e^{-x^2} = \sqrt{\pi}. \tag{27}$$

We can readily generalize, using a well-chosen linear substitution of the integration variable, to

$$\int_{-\infty}^{\infty} dx\, e^{-ax^2+bx} = \sqrt{\frac{\pi}{a}}e^{\frac{b^2}{4a}} \qquad (a > 0, \text{ any } b). \tag{28}$$

Exercise:

1. Derive Eq. (28).

2. Assume that b in Eq. (28) can also be complex, and use that to solve the integral

$$\int_{-\infty}^{\infty} dx\, \cos(x)\, e^{-x^2}.$$

III. INTEGRATION-BY-DIFFERENTIATION

A. Gaussian Distributions

The Gaussian is import particularly in statistics, where one is interested in "moments" of a particular distribution $p(x)$. The meaning of these terms I will discuss below for the more intuitive, discrete binomial distribution. Here, I simply define the nth "moment" of the "Gaussian distribution

$$p(x) = \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \tag{29}$$

as

$$\langle x^n \rangle = \int_{-\infty}^{\infty} dx\, x^n p(x) = \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}}\, x^n e^{-\frac{x^2}{2}}. \tag{30}$$

These integrals are also quite difficult, and as for the basic Gaussian integral, integration-by-parts will not help (why?). We could play the same trick as above and multiply with $1 = \int_{-\infty}^{\infty} dy\, e^{-y^2/2}/\sqrt{2\pi}$ and transform again to polar coordinates, in which case we will have to solve a number of integrals involving nth order polynomials in r and $\cos\phi$ multiplying the exponential. While doable, this is awkward! Instead, we will play a powerful trick, that has a number of applications beyond this case. Note that $1 = [e^{ax}]_{a=0}$is an identity for any regular (i. e. not singular) expression x, where $[\ldots]_{a=0}$ means "evaluate everything inside of the bracket in the end at $a \to 0^+$." (We will typically assume $a > 0$, and take the limit to 0 from above.) Using the notion of the derivative introduced in Sec. I, we can more generally write in the same spirit the identity

$$x^n = [\partial_a^n e^{ax}]_{a=0}. \tag{31}$$

The power of this identity becomes apparent, when we apply it to our moment problem above:

$$\begin{aligned}
\langle x^n \rangle &= \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}}\, [\partial_a^n e^{ax}]_{a=0}\, e^{-\frac{x^2}{2}} \\
&= \left[\int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}}\, \partial_a^n e^{ax} e^{-\frac{x^2}{2}} \right]_{a=0} \\
&= \left[\partial_a^n \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}}\, e^{-\frac{1}{2}x^2 + ax} \right]_{a=0}.
\end{aligned}$$

The first transformation of pulling the brackets $[\ldots]_{a=0}$ outward is obvious, since everything depending on a is still inside. The second step is quite profound, and may be hard to justify in general: We have to assume that the integral converges for any a near 0 and that it is always OK to differentiate n times to be able to interchange the order of integration and differentiation! Yet, given that, the integral is now straightforward (see Sec. II), and we obtain

$$\langle x^n \rangle = \left[\partial_a^n e^{\frac{a^2}{2}} \right]_{a=0},$$

having reduced the integral to an nth-fold derivative (hence the name "integration-by-differentiation"). For instance, we can now obtain

$$\begin{aligned}
\left\langle x^0 \right\rangle &= \left[e^{\frac{a^2}{2}} \right]_{a=0} = 1, \\
\left\langle x^1 \right\rangle &= \left[\partial_a e^{\frac{a^2}{2}} \right]_{a=0} = \left[a e^{\frac{a^2}{2}} \right]_{a=0} = 0, \\
\left\langle x^2 \right\rangle &= \left[\partial_a^2 e^{\frac{a^2}{4}} \right]_{a=0} = \left[\left(a^2+1\right) e^{\frac{a^2}{2}} \right]_{a=0} = 1,
\end{aligned}$$

which are, respectively, the "norm", the "mean" ($\mu =< x >$) and the "variance" $\left(\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2\right)$ of this Gaussian distribution in Eq. (29) In effect, what we have derived is an important property of any distribution, which is often referred to in statistics, called the "characteristic function"

$$\Phi(a) = \langle e^{ax} \rangle, \qquad \langle x^n \rangle = \left[\partial_a^n \Phi(a)\right]_{a=0}, \tag{32}$$

from which we can obtain any moment by mere differentiation. For this Gaussian,

$$\Phi(a) = e^{a^2/2}. \tag{33}$$

Alternatively, we could have realized immediately that only even moments $n = 2k$ in this Gaussian distribution are non-zero (why?), and used a variant of the same trick:

$$\begin{aligned}
\left\langle x^{2k} \right\rangle &= \int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} \left(x^2\right)^k e^{-\frac{x^2}{2}} \\
&= \left[\int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} \left(-\partial_a\right)^k e^{-ax^2} \right]_{a=\frac{1}{2}} \\
&= \left[\left(-\partial_a\right)^k \left(\int_{-\infty}^{\infty} \frac{dx}{\sqrt{2\pi}} e^{-ax^2} \right) \right]_{a=\frac{1}{2}} \\
&= \frac{1}{\sqrt{2}} \left[\left(-\partial_a\right)^k \frac{1}{\sqrt{a}} \right]_{a=\frac{1}{2}}.
\end{aligned}$$

Can you see, why we need $a = 1/2$ in the end here? Clearly, we get the same results: $\langle x^0 \rangle = [1/\sqrt{2a}]_{a=1/2} = 1$ and $\langle x^2 \rangle = \sqrt{1/2}[-\partial_a(a^{-1/2})]_{a=1/2} = 1$.

EXERCISE:

1. Rewrite the generalize Gaussian distribution

$$p(x) = ae^{-bx^2+cx}$$

by replacing the constants a, b, and c by the mean $\mu = \langle x \rangle$ and the "standard deviation" $\sigma = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$. (Hint: calculate $\langle x \rangle$ and $\langle x^2 \rangle$ and require proper normalization, i. e. enforce $\langle x^0 \rangle = 1$, to obtain relations between these constants.) You should find that

$$p(x) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}}. \tag{34}$$

Note that if we transform to the "reduced" variable ξ, $x \to \xi = (x-\mu)/\sigma$, in the generalized Gaussian distribution in Eq. (34), the new distribution $p(\xi) = \sigma p(x(\xi))$ has zero mean $\langle \xi \rangle = 0$ and unit variance, $\langle \xi^2 \rangle = 1$. In turn, its characteristic function is always $\phi(a) = \left\langle e^{a\xi} \right\rangle = e^{a^2/2}$.

B. Solving integrals with integration-by-differentiation

This method can be used more generally to solve all kinds of integrals. Take, for instance, the identity

$$\left[\partial_\lambda^n x^\lambda\right]_{\lambda=0} = \left[\partial_\lambda^n e^{\lambda \ln x}\right]_{\lambda=0} = \left[(\ln x)^n e^{\lambda \ln x}\right] = (\ln x)^n\,, \tag{35}$$

which can be useful in solving integrals involving powers of logarithms, e. g.

$$\begin{aligned} \int_0^1 dx\ (\ln x)^2 &= \int_0^1 dx\ \left[\partial_\lambda^2 x^\lambda\right]_{\lambda=0} = \left[\partial_\lambda^2 \int_0^1 dx\, x^\lambda\right]_{\lambda=0} \\ &= \left[\partial_\lambda^2 \frac{1}{1+\lambda}\right]_{\lambda=0} = \left[\frac{2}{(1+\lambda)^3}\right]_{\lambda=0} = 2. \end{aligned}$$

Clearly, we could have solved this integral also by other means, such as integration-by-parts. But this method is more elegant and straightforward; we always know how to differentiate while integration always involves some guess-work.

Exercise:

1. Determine

$$I = \int_0^1 dx\ (\ln x)^2\, x^{\frac{5}{6}}.$$

C. Binomial Distributions

Initially more intuitive than a Gaussian is the "binomial distribution"

$$p_k^n = \binom{n}{k} q^k (1-q)^{n-k} \tag{36}$$

for $0 \leq q \leq 1$ and any integers $0 \leq k \leq n$. It answers the following question: If an individual event has a probability of q to be realized (i. e. $1-q$ to be not realized), what is the probability that in n trials I see k such realizations of the event. For instance, if I do n coin tosses with the event "heads" having an elementary probability $q = 1/2$, about, then observing k such events has probability p_k^n. Similarly, we could be describing a drunkard that on each step is randomly swaying to the right with probability q and to the left with $1-q$. That drunkard would reach a distance $x_k^n = kL + (n-k)(-L) = (2k-n)L$ after $n$steps of length $L$from her origin at $x_0^0 = 0$ with probability p_k^n.

Note that the binomial distribution is closely related to the binomial series introduced in Sec. I with $p = 1-q$, from which immediately follows

$$\sum_{k=0}^{n} p_k^n = \sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k} = [q + (1-q)]^n = 1,$$

i. e. p_k^n is properly normalized, with the probability of "any" outcome being certainty (=1). In terms of this distribution, the meaning of "moments" becomes more clear. For example, we may consider the mean number of "heads" μ arising from a (large) number N of experiments of n coin tosses. In N_0 of those experiments, we found 0 "heads", N_1 with 1 "heads",..., N_n with n "heads". If N is large enough, we would naturally expect that $N_k/N \to p_k^n$. To get the mean number of "heads" one adds up all the individual experimental outcomes, $(0+0+\ldots)+(1+1+\ldots)+\ldots+(n+n+\ldots) = 0N_0 + 1N_1 + 2N_2 + \ldots + nN_n$, and divides by the total number of experiments, $N = \sum_{k=0}^{n} N_k$, to get

$$\mu = \frac{1}{N}\sum_{k=0}^{n} kN_k = \sum_{k=0}^{n} k\frac{N_k}{N} = \sum_{k=0}^{n} kp_k^n = \langle k \rangle \,,$$

in close correspondence to the definition of moments for the Gaussian distribution above. While p_k^n is a "discrete" distribution for a countable (discrete) set of outcomes k, the Gaussian is a continuous distribution, for which we can measure the likelihood outcomes only over closed intervals Δx, $p_k^n \approx p(x)\Delta x$. In the limit of large n, as we will see, the above expression will take on the character of a Riemann sum with infinitesimal intervals, $\Delta x \to dx$.

Furthermore, we can ask ourselves, how much do an experiment typically deviates (or "fluctuates") around the mean μ, since each individual experiment of n trials with give exactly the mean number of outcomes. (The mean may even be realizable, such as for $n = 3$ coin tosses with mean $\mu = 3/2$ non-integer!) Since we don't care whether the deviation $k_i - \langle k \rangle$ in the ith experiment was above or below average, we consider the squared deviation, $(k_i - \langle k \rangle)^2 = (k - \mu)^2$, so that a below-average fluctuation in one experiment can not cancel out another above-average one. Now, we simply average those squared-deviations themselves to obtain the "variance"

$$\sigma^2 = \left\langle (k-\mu)^2 \right\rangle = \left\langle k^2 - 2k\mu + \mu^2 \right\rangle = \left\langle k^2 \right\rangle - 2\mu \left\langle k \right\rangle + \mu^2 = \left\langle k^2 \right\rangle - \mu^2. \tag{37}$$

Therefore, it becomes apparent that higher moments (here, the 2nd moment $\langle k^2 \rangle$) are essential to probe details of a distribution.

Even more than for continuous distributions with integrals, the evaluation of sums to obtain moments of discrete distributions can be tedious. For example, we can calculate the first moment of the (relatively simple) binomial distribution:

$$\mu = \langle k \rangle = \sum_{k=0}^{n} k p_k^n = \sum_{k=1}^{n} \frac{n!k}{k!(n-k)!} q^k (1-q)^{n-k},$$

where we have used the definition of $\binom{n}{k}$ and the fact that the $k = 0$ term vanishes. Canceling the k by the factorial to get $(k-1)!$ in the denominator, we can shift $k = i + 1$ in the sum to get

$$\begin{aligned} \mu &= \sum_{i=0}^{n-1} n \frac{(n-1)!}{i![(n-1)-i]!} q^{i+1} (1-q)^{(n-1)-i} \\ &= nq \sum_{i=0}^{n-1} \binom{n-1}{i} q^i (1-q)^{(n-1)-i} \\ &= nq[q + (1-q)]^{n-1} \\ &= nq \end{aligned}$$

This is a familiar result, of course, since we expect that n coin-tosses with $q = 1/2$ should yield $n/2$ "heads", on average (after many experiments). Similarly, for a dice with elementary probability $q = 1/6$ to roll "6", we would expect $n/6$ times a "6" out of n rolls.

EXERCISE:

1. Use similar steps to obtain the 2nd moment, $\langle k^2 \rangle$.

To avoid the tedium of such evaluations, we can use the same trick as above for continuum distributions, using integration-by-differentiation. But instead of just the first moment, lets go after the characteristic function immediately:

$$\begin{aligned}
\phi(a) = \left\langle e^{ak} \right\rangle &= \sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k} e^{ak} \\
&= \sum_{k=0}^{n} \binom{n}{k} \left(q e^a\right)^k (1-q)^{n-k} \\
&= \left(q e^a + 1 - q\right)^n . \qquad (38)
\end{aligned}$$

Now, it is a straightforward matter to obtain moments:

$$\begin{aligned}
\mu &= \left[\partial_a \phi(a)\right]_{a=0} \\
&= \left[\partial_a (q e^a + 1 - q)^n\right]_{a=0} \\
&= \left[n q e^a (q e^a + 1 - q)^{n-1}\right]_{a=0} \\
&= nq \\
\left\langle k^2 \right\rangle &= \left[\partial_a^2 \phi(a)\right]_{a=0} \\
&= \left[\partial_a n q e^a (q e^a + 1 - q)^{n-1}\right]_{a=0} \\
&= \left[n q e^a (q e^a + 1 - q)^{n-1} + n(n-1) q^2 e^{2a} (q e^a + 1 - q)^{n-2}\right]_{a=0} \\
&= nq + n(n-1)q^2,
\end{aligned}$$

hence

$$\sigma = \sqrt{\langle k^2 \rangle - \mu^2} = \sqrt{nq + n^2 q^2 - n q^2 - n^2 q^2} = \sqrt{nq(1-q)}. \qquad (39)$$

This result says that if you toss a coin n times, you can expect the number of "heads" will typically deviate from the mean result at most by (roughly) $\pm\sqrt{n}$ with a certain degree of certainty. (This statement can be made more precise: In general, the outcome would be between $\mu - \sigma$ and $\mu + \sigma$ with probability $\approx 2/3$; that's called a "standard deviation" and is often the width of an error-bar in a plot of data.) So, if you do $n = 100$ coin tosses, you can expect to find somewhere between $k = 90 - 110$ "heads". Note that the relative size of fluctuations gets smaller and smaller with increasing n, $\sigma/n \approx 1/\sqrt{n}$, making the mean value μ better and better defined.

In all this, there is a deep connection between the binomial distribution (and many others) on one hand and the Gaussian distribution on the other hand. Many of these distributions have a Gaussian as a limiting case for large n. This is roughly the content of the "law of large numbers", which gives the Gaussian its fundamental importance. We will now demonstrate this limit for the binomial distribution. To make this connection, we want to calculate the characteristic function of the binomial distribution for the "reduced" variable, $\xi = (k-\mu)/\sigma$. (It would make no sense to compare two distributions, if they don't have at least the same mean and variance!) Then, in a calculation similar to the above,

$$\phi(a) = \langle e^{a\xi} \rangle = e^{-\frac{a\mu}{\sigma}} \sum_{k=0}^{n} \binom{n}{k} q^k (1-q)^{n-k} e^{\frac{a}{\sigma}k} = e^{-\frac{1\mu}{\sigma}} \left(q e^{\frac{a}{\sigma}} + 1 - q \right)^n .$$

Now, with $\mu/\sigma = \sqrt{nq/(1-q)}$, so $e^{-a\mu/\sigma} = \left(e^{-a\sqrt{q/[(1-q)n]}} \right)^n$, we get

$$\phi(a) = \left[e^{-a\sqrt{\frac{1}{(1-q)n}}} \left(q e^{\frac{a}{\sqrt{nq(1-q)}}} + 1 - q \right) \right]^n .$$

At this point, it is not at all clear that this expression would reach an interesting limiting form, independent of n for $n \to \infty$. The only thing to do here is to Taylor-expand both exponentials in the small variable $1/\sqrt{n}$:

$$\phi(a) \approx \left[\left(1 - a\sqrt{\frac{q}{(1-q)n}} + \frac{a^2 q}{2(1-q)n} + \ldots \right) \left(q + a\sqrt{\frac{q}{(1-q)n}} + \frac{a^2}{2(1-q)n} + \ldots + 1 - q \right) \right]^n .$$

Note first that the terms in q cancel in the second parenthesis. After multiplying out term-by-term, keeping only terms not smaller than $1/n$, a miraculous cancellation occurs of all terms in $1/\sqrt{n}$, and we are left with

$$\begin{aligned} \phi(a) &= \left[1 + \frac{a^2 q}{2(1-q)n} + \frac{a^2}{2(1-q)n} - \left(a\sqrt{\frac{q}{(1-q)n}} \right)^2 + \ldots \right]^n \\ &= \left[1 + \frac{a^2}{2n} + \ldots \right]^n , \end{aligned}$$

from which all q-dependence has now canceled! Using $\lim_{n\to\infty}(1 + x/n)^n = \lim_{n\to\infty} \exp[n \ln(1 + x/n)] = e^x$ (since $\ln(1+\epsilon) \approx \epsilon$ for $\epsilon \to 0$ from the Taylor series for the logarithm), we get finally for $n \to \infty$,

$$\phi(a) = e^{\frac{a^2}{2}},$$

i. e. we re-obtained the same characteristic function as for the Gaussian, Eq. (33). If two distributions share the same characteristic function, they share all moments, obviously. While this is not a proof here, it can be shown that both distributions are identical.

IV. GAMMA FUNCTIONS $\Gamma(x)$

A. Surface of a D-Dimensional Sphere

Let's consider the following question: What is the surface area $S_D(r)$ of the sphere in D dimensions? We remember that $S_3(r) = 4\pi r^2$ in $D = 3$ [$V_3(r) = 4\pi r^3/3$]. What is a sphere in $D = 2$? It's the $D = 2$ object that, when rotated into the third dimension, gives a $D = 3$ sphere. That object is the circle, of course. Its "volume" is the circle's area, $V_2(r) = \pi r^2$, hence, its surface is the circles circumference, $S_2(r) = 2\pi r$. By the same token, the plain line between $-1 \le x \le 1$ must be the corresponding unit "sphere" in $D = 1$. Clearly, its volume is $V_1(r) = 2r$, and its surface is a constant, $S_1(r) = 2$.

How can this quantity be generalized? It appears, that a consideration of physical units requires that $V_D(r) \propto r^D$ and $S_D(r) = \Omega_D r^{D-1}$; there is only one characteristic length for a sphere, its radius r. But that only solves half of the problem. What remains is a determination of the surface for a *unit* ($r = 1$) sphere, Ω_D. Our way of finding Ω_D is, in fact, a generalization of the trick we used to solve the Gaussian integral in Sec. II. There, we relied in the *spherical symmetry* of the Gaussian integrand. In trust of the generalization of the Euclidean metric into higher dimensions,

$$r_D^2 = x_1^2 + x_2^2 + \ldots + x_D^2,$$

we can write:

$$\begin{aligned}\left(\sqrt{\pi}\right)^D &= \left(\int_{-\infty}^{\infty} dx\, e^{-x^2}\right)^D \\ &= \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 \ldots \int_{-\infty}^{\infty} dx_D \, \exp\left[-\left(x_1^2 + x_2^2 + \ldots + x_D^2,\right)\right] \\ &= \int_{r=0}^{\infty} dV_D(r)\, e^{-r^2} \end{aligned} \tag{40}$$

Note that any spherically symmetric function could be used, in principle, but the Gaussian has the big advantage that it goes to zero very rapidly near $\pm\infty$. At this point, replacing the infinitesimal cubic-like volume element in D-dimensional Cartesian coordinates with the corresponding volume element in spherical-like coordinates,

$$dx_1 dx_2 \ldots dx_D \to dV_D(r),$$

is merely a formal statement. To fill it with meaning, we "observe" (if that is the right term for imagining something in arbitrary dimensions!) that we can tile the volume of D-dimensional space just as well with a sequence of infinitesimal, concentric, spherical shells as with Cartesian cubes, $dV_D = dx_1 dx_2 \dots dx_D$. Each such shell has a radius r from the origin and an infinitesimal width dr. Since dr gets arbitrarily small, the inner and outer surface area of the shell is about the same, $S_D(r)$. Hence, the infinitesimal volume of each shell is

$$dV_D(r) = S_D(r)dr = \Omega_D r^{D-1} dr, \tag{41}$$

which, inserted in Eq. (40), leads to

$$\pi^{\frac{D}{2}} = \Omega_D \int_0^\infty dr\, r^{D-1} \exp\left(-r^2\right). \tag{42}$$

Now, our problem is solved, almost! We reduced it to a discussion of a Gaussian integral, which is of general interest, and which we will now consider at length before we return to the D-sphere.

B. Factorials for non-integers

Let us consider the properties of the following integral, which we will *define* as the "$\Gamma-$function"

$$\Gamma(x) = \int_0^\infty dt\, t^{x-1} e^{-t}. \tag{43}$$

This integral exists for all $x > 0$ (why not $x \leq 0$?) and has a few remarkable properties. Integration-by-parts, using $u = e^{-t} \to u' = -e^{-t}$ and $v' = t^{x-1} \to v = t^x/x$ yields

$$\begin{aligned}
\Gamma(x) &= \int_0^\infty dt\, t^{x-1} e^{-t} \\
&= \left[-\frac{1}{x} t^x e^{-t}\right]_{t=0}^\infty + \frac{1}{x}\int_0^\infty dt\, t^x e^{-t} \\
&= \frac{1}{x}\Gamma(x+1), \\
\Gamma(x+1) &= x\Gamma(x).
\end{aligned} \tag{44}$$

This last, "recursion" equation for Γ is very suggestive. It is easy to integrate $\Gamma(1) = 1$. Then by Eq. (44), $\Gamma(2) = 1\Gamma(1) = 1$ and $\Gamma(3) = 2\Gamma(2) = 2 \times 1$, $\Gamma(4) = 3\Gamma(3) = 3 \times 2 \times 1$, $\Gamma(5) = 4\Gamma(4) = 4 \times 3 \times 2 \times 1$, and so on. It is, in fact, easy to show that Eq. (44) for non-negative integers $x = n$ is solved by the factorial function, i. e. $\Gamma(n+1) = n!$, indicating the close connection of the Γ-function to anything having to do with "counting" (combinatorics). In turn, the Γ-function accomplishes an amazing feat: it continues the factorial function off the non-negative integers! Such "analytic continuations" are used quite commonly. This continuation is not unique, though, as $\Gamma(x)[1+a\sin(2x\pi)]$ does the same thing *and* satisfies the recursion relation. Other, obviously desirable properties, like monotonicity for sufficiently large x, are needed to make $\Gamma(x)$ unique. Now, it makes sense to ask, for instance: What is $(\pi!)$, or $\left(-\frac{1}{2}!\right)$? The former is just some irrational number, but the latter is easy to find indeed:

$$\begin{aligned}\left(-\frac{1}{2}\right)! = \Gamma\left(\frac{1}{2}\right) &= \int_0^\infty dt\, t^{-\frac{1}{2}} e^{-t} \\ &= 2\int_0^\infty d\left(\sqrt{t}\right)\, e^{-t} \\ &= 2\int_0^\infty du\, e^{-u^2} \\ &= \sqrt{\pi}, \end{aligned} \tag{45}$$

where we have used the Gaussian integral in Eq. (27).

But there is more! Although the integral in Eq. (43) only defines $\Gamma(x)$ for $x > 0$, the recursion relation in Eq. (44) *uniquely* continues $\Gamma(x)$ to all real x. (Actually, the defining integral also converges for complex x with $Re\{x\} > 0$, hence the recursion then extends $\Gamma(x)$ to *all* complex x!) For example, $\Gamma\left(\frac{1}{2}\right) = -\frac{1}{2}\Gamma\left(-\frac{1}{2}\right)$, so $\left(-\frac{3}{2}\right)! = \Gamma\left(-\frac{1}{2}\right) = -2\Gamma\left(\frac{1}{2}\right) = -2\sqrt{\pi}$, and so on. In fact, if we knew $\Gamma(x)$ only on $a \le x \le a+1$ for any a, the recursion would provide $\Gamma(x)$ everywhere!

We know the factorial function as monotonic and rapidly rising but non-singular. The secret life of its continuation, $\Gamma(x)$, is somewhat surprising. Consider, for example, for a small $\epsilon \to 0$

$$\Gamma(\epsilon) = \frac{1}{\epsilon}\Gamma(1+\epsilon),$$

by virtue of the recursion relation in Eq. (44) again. It is easy to see from the defining integral in Eq. (43) that $\lim_{\epsilon\to 0^\pm}\Gamma(1+\epsilon) = 1$. Thus, near the origin, $\Gamma(\epsilon) = 1/\epsilon$ has a simple pole, i. e. it diverges to $+\infty$ for $\epsilon \to 0^+$, and to $-\infty$ for $\epsilon \to 0^-$. And the recursion relation implies that such simple poles occur in $\Gamma(x)$ for all non-positive integers x. For instance, consider $\Gamma(\epsilon - 1) = \Gamma(\epsilon)/(\epsilon - 1)$ for $\epsilon \to 0^\pm$: Since $1/(\epsilon - 1) \to -1$, $\Gamma(\epsilon - 1) \approx -\Gamma(\epsilon)$ inherits the pole of $\Gamma(0)$ but with a minus sign. Its behavior for $x \le 0$ is not unlike that of

$1/\sin(x\pi)$, and one can actually show the "reflection" relation

$$\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin(x\pi)}. \tag{46}$$

Returning to the D-dimensional sphere, we can finally write

$$\Omega_D = \frac{2\pi^{\frac{D}{2}}}{\Gamma\left(\frac{D}{2}\right)}, \qquad S_D(r) = \Omega_D r^{D-1}, \tag{47}$$

and, incidentally,

$$V_D(r) = \Omega_D \frac{1}{D} r^D = \frac{\pi^{\frac{D}{2}} r^D}{\Gamma\left(\frac{D}{2}+1\right)} \tag{48}$$

D	Ω_D	V_D
0	0	1
1	2	2
2	2π	π
3	4π	$\frac{4}{3}\pi$
4	$2\pi^2$	$\frac{1}{2}\pi^2$

Table I: Surface Ω_D and Volume V_D of the $D-$dimensional unit sphere for some integer dimensions.

Exercises:

1. Express the moments $\langle x^{2n}\rangle$ of the Gaussian distribution in Eq. (30) in terms of Γ-functions.

2. Sketch, by hand, $\Gamma(x)$ for $-3 < x < 3$.

3. Use integration-by-differentiation, the identity $1/A = \int_0^\infty dt\, e^{-At}$, and Eq. (46) to show that

$$I_n = \int_0^\infty dx \frac{\ln^n(x)}{1+x^2} = \frac{\pi}{2}\left[\partial_\lambda^n \frac{1}{\cos\left(\frac{\pi}{2}\lambda\right)}\right]_{\lambda=0}.$$

V. SADDLE POINT EXPANSIONS

A. The Traveling Salesperson Problem

We have see, how the Γ-function provides us with a continuation of the factorial function to any real number. Its utility, though, extends even beyond that! Consider a few every-day problems: You run a delivery service and your truck has to make a hundred drops a day in your area, moving from one drop point to the next. How many different paths are there to route your truck? Or you have to drill a thousand wholes into a board. How many ways are there to move your drill from whole-to-whole? This is but one, and the simplest, of a family of combinatorial problems. This particular one is called the "Traveling Salesperson Problem" (TSP).

The actual goal of the TSP is to minimize travel-time (or -cost). Here, we only focus on the much simpler question of counting the space of all solutions (paths). This is done best by "induction". Say, we already have counted the number of paths, a_n, for the problem with n locations. If we add one more location to just one of those a_n paths, how many new paths do we get? Well, any path consists of a closed n-sided polygon with an *ordered* sequence of those n locations on its corners. The new, $n+1$-st, location can be inserted on any side of the polygon, leading to n new, distinct paths. Hence, for $n+1$ locations, there are

$$a_{n+1} = n a_n \tag{49}$$

possible paths. This "recursion relation" in fact solves our problem: We already know that the factorial is a solution to the recursion relation, $a_n = n!$.

So, in the delivery business, we have to select the best of $N = 100!$ routes, and in the drill problem, the best of $N = 1000!$ paths. What *are* these numbers? Can we even discern their *order of magnitude?* How do we evaluate the factorial for large n?

Dealing with such large numbers, it may be helpful to look at their logarithm, which turns very rapidly varying functions in to more slowly moving ones. In particular, we can take the logarithm of Eq. (49) and set $b_n = \ln(a_n)$ to get

$$\begin{aligned}
b_{n+1} &= b_n + \ln(n)\\
\frac{b_{n+1} - b_n}{(n+1) - n} &= \ln(n)\\
\frac{db_n}{dn} &\approx \ln(n)\\
b_n &\approx n\ln(n) - n + B\\
a_n &\approx A n^n e^{-n}.
\end{aligned}$$

This somewhat venturous calculation, considering $dn = (n+1) - n = 1$ as *infinitesimal* compared to $n \gg 1$ and simply integrating the equation, in fact, provides a very useful result, which is probably not wrong by more than a factor of 10-100 (i. e. off by 1-2 orders of magnitude!) on the above problems. In this estimate, it doesn't even matter then that we simply replace the unknown integration constants by $B \approx 0$ or $A \approx 1$ and $e^3 \approx 20$:

$$a_{1000} \approx 1000^{1000} e^{-1000} \approx 10^{3000} 2^{-333} 10^{-333} \approx 10^{2667} \left(2^{10}\right)^{-33} \approx 10^{2567},$$

where we have also used $2^{10} \approx 10^3$. This result is too high, since we have underestimated both, 2^{10} and e^3. But how wrong are we? Thanks to the Γ-function, this question can be answered to any desired precision. (In reality, one gets $a_{1000} \approx 5\,10^{2565}$.)

B. Sharply peaked functions

To get a more precise value for $n!$, we first make a few general observations. Let us consider the simple polynomial function

$$\phi(t) = t^2.$$

This function varies quite slowly and has a minimum at $t = 0$, where $\phi'(t = 0) = 0$ and $\phi"(t = 0) = 2 > 0$. Its integral is straightforward and requires no approximation. Now, let us consider the function

$$e^{-x\phi(t)} \qquad (x \to +\infty).$$

This exponential now varies quite rapidly and has a sharp maximum at $t = 0$ for large x. For example, with, say, $x = 9$, the function at $t = \pm 1$ is $e^{-x} \approx 20^{-3} \approx 10^{-4}$, or four orders of magnitude smaller than at $t = 0$. So, let us look at its integral

$$\begin{aligned} I(x) &= \int_{-1}^{1} dt\, e^{-x\phi(t)} \quad (x \to +\infty). \qquad (50) \\ &= \int_{-\infty}^{\infty} dt\, e^{-xt^2} - \int_{-\infty}^{-1} dt\, e^{-xt^2} - \int_{1}^{\infty} dt\, e^{-xt^2} \\ &= \sqrt{\frac{\pi}{x}} - 2\int_{1}^{\infty} dt\, e^{-xt^2}, \end{aligned}$$

where we have used Eq. (28). As it turns out, for large x, the correction to this integral stemming from the domains $-\infty < t < -1$ and $1 < t < \infty$ are entirely negligible! Note that if $t > 1$ then $t < t^2$ and $e^{-xt} > e^{-xt^2} > 0$, hence

$$\int_{1}^{\infty} dt\, e^{-xt^2} < \int_{1}^{\infty} dt\, e^{-xt} = \frac{e^{-x}}{x} \ll \sqrt{\frac{\pi}{x}}.$$

Accordingly, we can write

$$I(x) \approx \sqrt{\frac{\pi}{x}}$$

with only an *exponentially* small error, say, 1 in 10^4at $x \approx 10$, as above.

How close to the minimum of $\phi(x)$ do we need to be to be able to neglect any corrections? Well, let us look at the same integral but with some extra parameter, $\epsilon > 0$, that measure the closeness to the minimum:

$$\begin{aligned} I(x) &= \int_{-\epsilon}^{\epsilon} dt\, e^{-x\phi(t)} \quad (x \to +\infty). \\ &= \int_{-\infty}^{\infty} dt\, e^{-xt^2} - \int_{-\infty}^{-\epsilon} dt\, e^{-xt^2} - \int_{\epsilon}^{\infty} dt\, e^{-xt^2} \\ &= \sqrt{\frac{\pi}{x}} - 2\int_{\epsilon}^{\infty} dt\, e^{-xt^2}. \end{aligned}$$

Now, substitute $t = \epsilon s$ and get for the correction integral

$$\begin{aligned} \int_{\epsilon}^{\infty} dt\, e^{-xt^2} &= \epsilon \int_{1}^{\infty} ds\, e^{-(x\epsilon^2)s^2} \\ &< \epsilon \int_{1}^{\infty} ds\, e^{-(x\epsilon^2)s} \\ &= \frac{e^{-x\epsilon^2}}{x\epsilon}. \end{aligned}$$

So, for the correction to be exponentially small, we only need an

$$\epsilon \gg 1/\sqrt{x}. \tag{51}$$

That means, at $x = 10$, we should choose $\epsilon \gg 1/3$. Say, already with $\epsilon = 1/2$, the correction is ≈ 0.01, missing $I(x) \approx \sqrt{\pi/x}$ by only $\approx 2\%$.

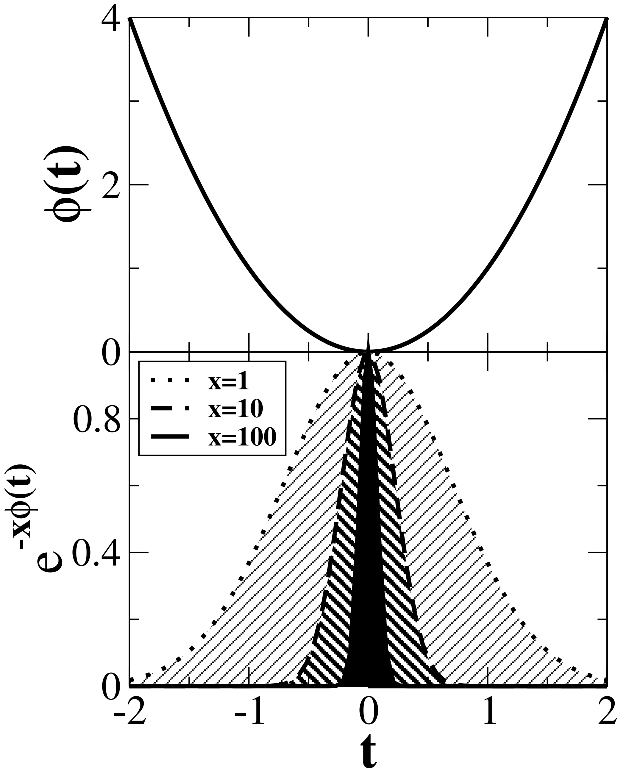

Figure 1: Plot of $\phi(t) = t^2$ and $e^{-x\phi(t)}$ for various x. While the upper graph only varies slowly, its exponential varies very rapidly for larger x. For increasing x, the area under each curve (i. e. its integral) in the lower plot is significant only over a decreasing interval in t.

Of course, the function $\phi(t) = t^2$ is somewhat trivial, but it demonstrates the *generic* behavior of so-called Laplace Integrals, Eq. (50). Let's us do a more complicated example,

$$\begin{aligned} \phi(t) &= \sin(t), \\ I(x) &= \int_{-\pi}^{0} dt\, e^{-x\sin(t)} \qquad (x \to \infty). \end{aligned} \tag{52}$$

This is in general a very hard integral, which has no simple solution, unless x is large. The $\sin(t)$ function has minima at $\ldots - 5\pi/2, -\pi/2, 3\pi/2, \ldots$ and any shifts by multiples of 2π thereof, but for this integral, only the minimum at $t_0 = -\pi/2$matters (why?). From the previous calculation, we expect that for large x only a small interval around the minimum of $\phi(t)$ near $t = t_0 = -\pi/2$ contributes to $I(x)$, so

$$\begin{aligned} I(x) &= \int_{t_0-\epsilon}^{t_0+\epsilon} dt\, e^{-x\sin(t)} + \int_{t_0+\epsilon}^{0} dt\, e^{-x\sin(t)} + \int_{-\pi}^{t_0-\epsilon} dt\, e^{-x\phi\sin(t)} \\ &= \int_{t_0-\epsilon}^{t_0+\epsilon} dt\, e^{-x\sin(t)} + \mathit{exponential\ smaller\ terms\ in}\ x \\ &\approx \int_{-\epsilon}^{\epsilon} ds\, e^{-x\sin(t_0+s)} \end{aligned} \tag{53}$$

where we have substituted $t = t_0 + s$, $dt = ds$. What have we gained? Well, in such a small environment, $-\epsilon < s < \epsilon$, we can Taylor-expand

$$\phi(t_0+s) \approx \phi(t_0) + s\phi'(t_0) + \frac{1}{2}s^2\phi"(t_0) + \frac{1}{6}s^3\phi^{(3)}(t_0) + \frac{1}{24}s^4\phi^{(4)}(t_0) + \ldots \tag{54}$$

$$\begin{aligned} \sin\left(-\frac{\pi}{2}+s\right) &\approx \sin\left(-\frac{\pi}{2}\right) + s\cos\left(-\frac{\pi}{2}\right) - \frac{1}{2}s^2\sin\left(-\frac{\pi}{2}\right) \\ &\quad -\frac{1}{6}s^3\cos\left(-\frac{\pi}{2}\right) + \frac{1}{24}s^4\sin\left(-\frac{\pi}{2}\right) + \ldots \\ &\approx -1 + \frac{1}{2}s^2 - \frac{1}{24}s^4 + \ldots . \end{aligned} \tag{55}$$

Inserting this result into Eq. (53), we get

$$\begin{aligned} I(x) &\approx \int_{-\epsilon}^{\epsilon} ds\, e^{-x\left(-1+\frac{1}{2}s^2-\frac{1}{24}s^4+\ldots\right)} \\ &\approx e^x \int_{-\epsilon}^{\epsilon} ds\, e^{-\frac{1}{2}xs^2+\frac{1}{24}xs^4+\ldots} \end{aligned}$$

It is key now to substitute $u = s\sqrt{x}$, $ds = du/\sqrt{x}$, then

$$I(x) \approx \frac{e^x}{\sqrt{x}} \int_{-\epsilon\sqrt{x}}^{\epsilon\sqrt{x}} du\, e^{-\frac{1}{2}u^2+\frac{1}{24x}u^4\ldots}.$$

In this way, it is apparent that we can neglect the u^4-term (and any higher-order terms), since $1/x$ is small. Furthermore, because of Eq. (51), $\epsilon\sqrt{x}$ is a *large* number, so that

$$\begin{aligned} I(x) &\approx \frac{e^x}{\sqrt{x}} \int_{-\infty}^{\infty} du\, e^{-\frac{1}{2}u^2} \\ &\approx \sqrt{\frac{2\pi}{x}}e^x. \end{aligned}$$

We can easily generalize this result, also adding another, x-independent function $f(t)$ in the integrand. As in Eq. (53), we obtain in general

$$\begin{aligned} I(x) &= \int_a^b dt\, f(t)\, e^{-x\phi(t)} \qquad (56) \\ &= \int_{t_0-\epsilon}^{t_0+\epsilon} dt\, f(t)\, e^{-x\phi(t)} + \textit{exponential smaller terms in } x \\ &\approx \int_{-\epsilon}^{\epsilon} ds\, f(t_0+s)\, e^{-x\phi(t_0+s)}, \\ &\approx \int_{-\epsilon}^{\epsilon} ds\, \left[f(t_0) + sf'(t_0) + \ldots\right]\, e^{-x\left[\phi(t_0)+s\phi'(t_0)+\frac{1}{2}s^2\phi''(t_0)+\frac{1}{6}s^3\phi^{(3)}(t_0)+\ldots\right]}, \qquad (57) \end{aligned}$$

Note that since $\phi(t)$ is *assumed* to have an absolute minimum at $t_0 \in (a,b)$, we have $\phi'(t_0) = 0$, hence,

$$\begin{aligned} I(x) &\approx e^{-x\phi(t_0)} \int_{-\epsilon}^{\epsilon} ds\, \left[f(t_0) + sf'(t_0) + \ldots\right]\, e^{-\frac{1}{2}xs^2\phi''(t_0)+\frac{1}{6}xs^3\phi^{(3)}(t_0)+\ldots} \\ &\approx \frac{f(t_0)}{\sqrt{x}} e^{-x\phi(t_0)} \int_{-\epsilon\sqrt{x}}^{\epsilon\sqrt{x}} du\, \left[1 + \frac{u}{\sqrt{x}}\frac{f'(t_0)}{f(t_0)} + \ldots\right]\, e^{-\left[\frac{1}{2}\phi''(t_0)\right]u^2+\frac{1}{\sqrt{x}}\left[\frac{1}{6}\phi^{(3)}(t_0)\right]u^3+\ldots} \qquad (58) \end{aligned}$$

Neglecting small terms in $1/\sqrt{x}$, $1/x$, etc, and letting $\epsilon\sqrt{x} \to \infty$, it is finally

$$\begin{aligned} I(x) &\approx \frac{f(t_0)}{\sqrt{x}}\, e^{-x\phi(t_0)} \int_{-\infty}^{\infty} du\, e^{-\left[\frac{1}{2}\phi''(t_0)\right]u^2} \\ \int_a^b dt\, e^{-x\phi(t)} &\approx f(t_0)\sqrt{\frac{2\pi}{x\phi''(t_0)}}\, e^{-x\phi(t_0)} \qquad (x \to \infty), \qquad (59) \end{aligned}$$

where t_0 is defined by the global minimum on (a,b) with

$$\phi'(t_0) = 0.$$

It should be noted that at this stage the result in Eq. (59) is only accurate up to terms of order $1/\sqrt{x}$, because of the terms we dropped in Eq. (58). But those terms could be further incorporated systematically by simply Tayloring in powers of $1/\sqrt{x}$ in Eq. (58). Any exponentially small corrections from the tails of the integral can *always* be neglected.

EXERCISES:

1. Find the approximation to

$$I(x) = \int_{-\infty}^{\infty} dt\, e^{-x\cosh(at)} \qquad (x \to \infty).$$

2. Find the approximation to

$$I(x) = \int_0^{\infty} dt\, t\, e^{-x\left(gt^4 - m^2t^2\right)} \qquad (x \to \infty).$$

3. Find the approximation to

$$I(x) = \int_0^\infty dt\, e^{-x\left(t^4 - \frac{7}{8}t^2 + \frac{3}{8}t\right)} \qquad (x \to \infty).$$

C. Γ-Function for Large Arguments

Now we have all the tools needed to obtain a *systematic* approximation to $\Gamma(x)$ for large x. Although, we still have to be clever! Starting from the Γ-integral,

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

which is *not at all* in the form of the integral in Eq. (56). We need the x in the exponential, so let us shift $t = xs$, $dt = xds$, to get

$$\begin{aligned} \Gamma(x) &= x^x \int_0^\infty ds\, s^{x-1} e^{-xs} \\ &= x^x \int_0^\infty \frac{ds}{s}\, e^{-xs + x\ln(s)}. \end{aligned}$$

Now, we have an integral that has exactly the form of Eq. (59) with

$$\begin{aligned} \phi(s) &= s - \ln(s) \\ \phi'(s) &= 1 - \frac{1}{s} \\ \phi"(s) &= \frac{1}{s^2}. \end{aligned}$$

Clearly, $\phi(s)$ has one unique minimum, where$\phi'(s_0) = 0$, on $(0, \infty)$ at $s_0 = 1$. There, it is $\phi(s_0) = 1$ $\phi"(s_0) = 1$ and $f(s_0) = 1/s_0 = 1$. Applying Eq. (59), we finally get

$$\Gamma(x) \approx x^x \sqrt{\frac{2\pi}{x}}\, e^{-x} \qquad (x \to \infty). \tag{60}$$

This is called the "Stirling Approximation" for the Γ-function.

Returning to our above problem of calculating the number of paths, $a_{1000} = 1000!$. First, note that

$$n! = \Gamma(n+1) = n\Gamma(n) \approx \sqrt{2\pi n}\, n^n\, e^{-n}. \tag{61}$$

Hence, our above estimate was off by a factor of $\sqrt{2\pi 1000} \approx 80$. Estimates aside, the actual relative error between 1000! and the Stirling approximation is less than 0.1%.

EXERCISE:

1. Determine the relative error between 10! and its Stirling approximation.

VI. DIFFERENCE EQUATIONS

A. Mortgage Equation

Let us consider how interest-payments on loans are calculated. Say, you are taking out a loan amount L at an periodic interest rate of I. You agree to repay your debt, including all interest, after T periods. What would be your periodic payment P? Well, we can proceed inductively again, and write down a recursion relation [as for the factorial function in Eq. (49)] for your current debt D_n during period n. Clearly, at the beginning, $n = 0$, we have $D_0 = L$, and in the end, $n = T$, we want $D_T = 0$. What is our debt during the next period, D_{n+1}, assuming that our current debt is D_n? During that period, our debt increases by the interest on that debt, ID_n, and decreases by our payment P from its current level. Hence,

$$D_{n+1} = D_n + ID_n - P. \tag{62}$$

This is called a "difference equation", due to its close similarity to differential equations. To see this similarity, we define $D_n = d(t_0 + n\Delta t)$ for a continuous time-variable t and rewrite Eq. (62) as

$$\frac{d(t_0 + (n+1)\Delta t) - d(t_0 + n\Delta t)}{[t_0 + (n+1)\Delta t] - [t_0 + n\Delta t]} = \left(\frac{I}{\Delta t}\right) d(t_0 + n\Delta t) - \left(\frac{P}{\Delta t}\right).$$

Defining $i = I/\Delta t$ and $p = P/\Delta t$ as the interest rate per unit time and the payment per unit time, respectively, we obtain in the limit of $\Delta t \to 0$,

$$d'(t) = id(t) - p, \tag{63}$$

where $t = t_0 + n\Delta t$ is now a continuous variable. Note that if we wanted to solve such a differential equation numerically on a digital computer, which is inherently discrete, we in fact would have to proceed exactly backwards from Eq. (63) to Eq. (62) by introducing a sufficiently small but finite Δt. (The smaller Δt, the more accurate the numerical result would be, but the more steps T would be required in the computation to cut a preset time interval into slices of size Δt.)

Eq. (63) is an inhomogeneous differential equation of first order, for those familiar with differential equations. There are standard methods to solve those equations in generality and, not surprisingly, many of these methods have their correspondents for difference equations. For our purposes here, let us just note that we can reduce Eq. (63) to the differential equation for the exponential by a simple shift by a constant, $d(t) = \bar{d}(t) + p/i$, since $d'(t) = \bar{d}'(t)$. Similarly for Eq. (62), we can shift $D_{n(+1)} = \bar{D}_{n(+1)} + P/I$ to find

$$\bar{D}_{n+1} = \bar{D}_n(1+I). \tag{64}$$

A typical way to obtain a solution for a difference equation, which is linear in the dependent variable D_n and has only constant (n-independent) coefficients, is with the "Ansatz" (German for "try")

$$D_n = Ar^n \qquad (A, r\,\text{constants}).$$

Inserting that into Eq. (64), we obtain $r = 1 + I$, leading to $\bar{D}_n = A(1+I)^n$. Thus we find as general solution for Eq. (62)

$$D_n = A(1+I)^n + \frac{P}{I}. \tag{65}$$

To make this general solution apply to our particular problem, we have to determine the arbitrary constant A and the desired payment P in terms of our specific requirements, i. e. that $D_0 = L$ and $D_T = 0$. First,

$$L = D_0 = A + \frac{P}{I} \Rightarrow A = L - \frac{P}{I},$$

which leads to the solution to our problem

$$D_n = \left(L - \frac{P}{I}\right)(1+I)^n + \frac{P}{I}. \tag{66}$$

With that, we can now determine our periodic payments when the repayment is stretched over T periods:

$$0 = D_T = \left(L - \frac{P}{I}\right)(1+I)^T + \frac{P}{I} \Rightarrow P = \frac{IL(1+I)^T}{1-(1+I)^T}. \tag{67}$$

EXERCISES:

1. Mortgages are typically paid over monthly periods, although market interest rates are quoted as "annual percentage rates" (APR). If your mortgage has an APR of 6%, what is I?

2. "Interest-only loans" are quite popular currently, for which the "principle" loan amount L never gets paid off, i. e. $D_n = L$ for all $n \geq 0$. (This may be useful – compared to renting – if one *must* sell the property again within a few years, but hopefully not at a loss!) What is the monthly payment P on a loan of $L = 200,000$ at 5% APR?

3. Consider the choice between a 30-year (360 months) mortgage at 6% APR and a 15-year (180 months) mortgage at 5% APR for a loan of $L = 200,000$. What are the respective monthly payments P, what is the total excess pay (difference between total amount paid and the "equity" owned, where your equity is $E_n = L - D_n$) after 15 years for (1) the 15-year mortgage, (2) the 30-year mortgage, and (3) the interest-only loan from 2. above?

B. Random Walks

We have already referred to the walk of a drunkard in the context of the binomial distribution in Sec. III C. This problem, called a *random walk*, is in fact of great importance, since it describes general *diffusion* processes. For instance, we could replace the staggering drunkard by a molecule in solution that is randomly kicked into anywhich direction by the surrounding molecules. If we imagine a droplet full of a pigment, which contains many molucules, immersed in clear water, we would soon find that the pigmentation has spread throughout the water.

Let us consider a discrete "lattice" in one dimension, i. e. the line of all integers $-\infty < i < \infty$, and discretized time steps $t = 0, 1, 2, \ldots$. We assume to have at $t = 0$ a walker at the origin $i = 0$. At each subsequent step, the walker moves to the left ($i \to i - 1$) or right ($i \to i + 1$) with equal probability. We want to find the number of possible paths $N_{i,t}$ that this walker could have taken to reach a point (i, t). For example, at time $t = 1$, there is only one path each to reach the points ($i = -1, t = 1$) and ($i = 1, t = 1$), so $N_{-1,1} = N_{1,1} = 1$ and $N_{i,1} = 0$ for all other $i \neq \pm 1$. At time $t = 2$, there is again only one path each to reach $i = \pm t$, generally $N_{\pm t,t} = 1$, since it leaves the walker no choice but to *always* walk left or right, resp. But there are now two paths to return to the origin, $N_{0,2} = 2$, for a walker going right-left or left-right; all other $N_{i,2} = 0$. Continuing on, one notices that the numbers for the paths generated are in fact those of the Pascal triangle. For each time t, there are 2^tpossible path, so the probability to reach any point (i, t) is simply $N_{i,t}/2^t$, since all path are independent. Due to the close relation to the Pascal triangle, we can easily deduce from Sec. III C that $p^t_{(t+i)/2} = N_{i,t}/2^t$ for $t + i$ even, refering to the binomial distribution p^n_k in Eq. (36). Note that no path can reach a site i at time t, if $i + t$ is odd; this "checkerboard pattern" makes the equations for $N_{i,t}$ unnecessarily difficult. Instead, we consider the numbers in the familiar Pascal triangle, $\Pi_{n,k} = 2^n p_k^n$, directly.

For the construction of the Pascal triangle, we can write down

$$\Pi_{n,k} = \Pi_{n-1,k} + \Pi_{n-1,k-1}, \qquad (n, k > 0), \tag{68}$$

and $\Pi_{n,0} = 1$ for all $n \geq 0$. Again, we have obtained a recursion relation, but in this case it depends to *two* indices, n and k. While this equation could just as well be solved by other means, we want to use it here to introduce an important method, called *generating functions*, used to solve even more general recursion relations. Defining the generating function

$$g_k(x) = \sum_{n=0}^{\infty} \Pi_{n,k} x^n, \qquad g_0(x) = \sum_{n=0}^{\infty} \Pi_{n,0} x^n = \frac{1}{1-x}, \tag{69}$$

we can multiply Eq. (68) with x^n and sum over all $n > 0$ to get

$$\begin{aligned} \sum_{n=1}^{\infty} \Pi_{n,k} x^n &= \sum_{n=1}^{\infty} \Pi_{n-1,k} x^n + \sum_{n=1}^{\infty} \Pi_{n-1,k-1} x^n, \qquad (k > 0), \\ \sum_{n=0}^{\infty} \Pi_{n,k} x^n - \Pi_{0,k} &= x \sum_{n=0}^{\infty} \Pi_{n,k} x^n + x \sum_{n=0}^{\infty} \Pi_{n,k-1} x^n, \qquad (k > 0), \\ g_k(x) &= x g_k(x) + x g_{k-1}(x), \qquad (k > 0), \end{aligned} \tag{70}$$

since $\Pi_{0,k} = 0$ for $k > 0$. Thus, the generating function has reduced the problem to a simple, ordinary recursion relation in one variable, k, with x as a parameter. We rearrange Eq. (70) and solve:

$$\begin{aligned} g_k(x) &= \frac{x}{1-x}\, g_{k-1}(x), \qquad (k > 0), \\ g_k(x) &= g_0(x) \left(\frac{x}{1-x} \right)^k = \frac{x^k}{(1-x)^{k+1}}, \qquad (k \geq 0). \end{aligned}$$

Now we have to "pay" for the convenience that the generating function $g_k(x)$ provided. To convert back to $\Pi_{n,k}$, we have to determine the Taylor coefficients of $g_k(x)$. In this case, the task is quite easy, but in general, the conversion may be very difficult. Here, we observe from the binomial expansion in Eq. (15), analytically continued to arbitrary exponent n that

$$\begin{aligned} (1-x)^n &= \sum_{i=0}^{n} \frac{n!}{i!(n-i)!} (-x)^i, \\ &= \sum_{i=0}^{\infty} \frac{\Gamma(n+1)}{i!\Gamma(n-i+1)} (-x)^i, \\ &= \sum_{i=0}^{\infty} \frac{(-x)^i}{i!} \frac{\Gamma(i-n) \sin\left[\pi\,(n+1-i)\right]}{\Gamma(-n) \sin\left[\pi\,(n+1)\right]}, \end{aligned}$$

where we have used the reflection formula for Γ−functions in Eq. (46) twice. With integer i and arbitrary n, $\sin\left[\pi\left(n+1-i\right)\right] = \sin\left[\pi\left(n+1\right)\right]\cos(\pi i) - \sin(\pi i)\cos\left[\pi\left(n+1\right)\right] = (-1)^i \sin\left[\pi\left(n+1\right)\right]$, so

$$
\begin{aligned}
(1-x)^n &= \sum_{i=0}^{\infty} \frac{\Gamma(i-n)}{i!\Gamma(-n)}\, x^i, && (71)\\
\frac{x^k}{(1-x)^{k+1}} &= \sum_{i=0}^{\infty} \frac{\Gamma(i+k+1)}{i!\Gamma(k+1)}\, x^{i+k},\\
g_k(x) &= \sum_{i=k}^{\infty} \frac{\Gamma(i+1)}{(i-k)!\Gamma(k+1)}\, x^i. && (72)
\end{aligned}
$$

Hence, (replacing the dummy index $i \to n$)

$$\Pi_{n,k} = \binom{n}{k}, \qquad (n \geq k \geq 0),$$

the familiar result for the Pascal triangle. Thus, we obtain for the number of paths for the random walker to get to site i at time t,

$$N_{i,t} = 2^t p^t_{\frac{i+t}{2}} = \Pi_{t,\frac{i+t}{2}} = \binom{t}{\frac{i+t}{2}}, \qquad \left(0 \leq \frac{i+t}{2} \leq t,\ i+t\,\text{even}\right), \tag{73}$$

and there are no paths if $i+t$ is odd.

C. Poisson Distribution

Another example of an interesting difference equation is given by Poisson process. Consider you are observing a sequence of events, say, the decay of nuclear material or the arrival of costumers into a queue. The elementary events have a specific rate of occurence λ, and all events are independent of each other. The question is, How many events n will have occured after a time t has passed since we started counting? We define the probability to have see exactly n events after time t as $P_n(t)$. The rate of change, $\partial_t P_n(t)$, is easily determined: At rate λ, the $n+1$st event deminishes $P_n(t)$, while with rate λ, the nth event occurs, adding $\lambda P_{n-1}(t)$ to $P_n(t)$. Hence,

$$\partial_t P_n(t) = -\lambda P_n(t) + \lambda P_{n-1}(t), \qquad (n > 0), \tag{74}$$

starting from $P_n(0) = 0$ for $n > 0$ and $P_0(0) = 1$. Note that we have to require also that

$$\partial_t P_0(t) = -\lambda P_0(t) \tag{75}$$

for the initial process. Eq.(74-75) is a mixed, differential-difference equation. To solve the differential equation first, we define a generating function as

$$g(x,t) = \sum_{n=0}^{\infty} P_n(t)\, x^n, \tag{76}$$

with $g(x,0) = 1$. We simply multiply x^n through Eq. (74) and sum for all $n > 0$ to get

$$\begin{aligned}
\partial_t \sum_{n=1}^{\infty} P_n(t)\, x^n &= -\lambda \sum_{n=1}^{\infty} P_n(t)\, x^n + \lambda \sum_{n=1}^{\infty} P_{n-1}(t)\, x^n, \\
\partial_t g(x,t) - \partial_t P_0(t) &= -\lambda \left[g(x,t) - P_0(t)\right] + \lambda x g(x,t), \\
\partial_t g(x,t) + \lambda(1-x) g(x,t) &= \partial_t P_0(t) + \lambda P_0(t) = 0,
\end{aligned}$$

where the last equality follows from Eq. (75). This ordinary differential equation for $g(x,t)$ in t (x is merely a paramter here) is easily solved by an exponential:

$$\begin{aligned}
g(x,t) &= g(x,0) e^{-\lambda(1-x)t}, && (77)\\
&= e^{-\lambda(1-x)t}, && (78)\\
&= \sum_{n=0}^{\infty} \frac{t^n}{n!}\, e^{-\lambda t}\, x^n. && (79)
\end{aligned}$$

Thus, by comparison with Eq. (76), we finally get

$$P_n(t) = \frac{t^n}{n!}\, e^{-\lambda t}. \tag{80}$$

D. How to convert Generating Functions

Consider again the generating function $g_k(x)$ in Eq. (69) and rewrite it with a *complex* argument $x = e^{i\phi}$, to get

$$\begin{aligned}
g_k\left(e^{i\phi}\right) &= \sum_{n=0}^{\infty} \Pi_{n,k} e^{in\phi}, \\
\int_0^{2\pi} d\phi\, g_k\left(e^{i\phi}\right) e^{-im\phi} &= \sum_{n=0}^{\infty} \Pi_{n,k} \int_0^{2\pi} d\phi\, e^{i(n-m)\phi}, && (81)
\end{aligned}$$

where m is another integer. The integral on the right-hand-side is easily evaluated, with a surprising result: If $m = n$, the integrant is simply unity, so the integral equals 2π; if $m \neq n$, we obtain by straightforward integration:

$$\begin{aligned}\int_0^{2\pi} d\phi\, e^{i(n-m)\phi} &= \frac{1}{i(n-m)}\left[e^{i(n-m)\phi}\right]_{\phi=0}^{2\pi},\\ &= \frac{1}{i(n-m)}\left(e^{2\pi i(n-m)}-1\right),\\ &= 0,\end{aligned}$$

as $e^{2\pi i k}=1$ for any integer k, see Sec. I C. Therefore, the sum on the right-hand-side of Eq. (81) collapses, leaving only the term for $m=n$:

$$\Pi_{m,k} = \frac{1}{2\pi}\int_0^{2\pi} d\phi\, g_k\left(e^{i\phi}\right)e^{-im\phi}. \tag{82}$$

Clearly, in general, this integration may be quite difficult.

VII. PERTURBATION EXPANSIONS

In many areas of science and engineering, almost everything we know derives from the power of *perturbative expansions*. Well, with the advent of computation, things have improved somewhat; simulations even helped to discover mathematical solutions to difficult problems previously only accessible to perturbations. One area where our knowledge almost exclusively depended on perturbative expansions until some 30 years ago was the theory of sub-atomic particles, for instance.

A. Perturbing a Differential Equation

Perturbative expansions are mostly used in the context of differential equations. Since this is not a course in differential equations, we consider only a very simple example here: Say, you want to solve the problem

$$y'(x) = y(x) - 2\ln\left[y(x)\right], \qquad y(0)=1. \tag{83}$$

Clearly, this being a ordinary differential equation of first order, it is easy to write down the answer in terms of a implicit integral expression. But there is no closed form solution, and without having a computer around to solve it numerically, that solution is not very helpful. Instead, we note that it is easy to solve Eq. (83), if the logarithmic term were not present, $y'(x)=y(x)$, i. e. $y(x)=e^x$. How can we utilize what we know to obtain information about the real problem? The idea is to *interpolate* between the original, hard problem and a suitable(!) "easy" problem. Hence, we have to deal with a one-parameter family of problems,

$$y'(x,\epsilon) = y(x,\epsilon) - 2\epsilon \ln\left[y(x,\epsilon)\right], \qquad y(0,\epsilon) = 1, \tag{84}$$

where we obtain the original problem for $\epsilon = 1$ and the easy problem for $\epsilon = 0$. The idea is to construct a Taylor series (see Sec. I B) in the parameter ϵaround the easy problem,

$$y(x,\epsilon) \approx y(x,0) + \epsilon\partial_\epsilon y(x,\epsilon)|_{\epsilon=0} + \frac{1}{2}\epsilon^2\partial_\epsilon^2 y(x,\epsilon)|_{\epsilon=0} + \frac{1}{6}\epsilon^3\partial_\epsilon^3 y(x,\epsilon)|_{\epsilon=0} + \ldots \tag{85}$$

$$\approx y_0(x) + \epsilon y_1(x) + \epsilon^2 y_2(x) + \epsilon^3 y_3(x) + \ldots \tag{86}$$

Ostensibly, then,

$$y(x) = y(x,1) \approx y_0(x) + y_1(x) + y_2(x) + y_3(x) + \ldots$$

provides an ever improving solution to our problem, but we will deal with that question later. (Note that since $y(0) = 1$, we want $y_0(0) = 1$, $y_1(0) = 0$, $y_2(0) = 0, \ldots$.) First, we analyze what happens to first order in ϵ when we insert Eq. (86) into Eq. (84):

$$\begin{aligned} y_0' + \epsilon y_1' + \ldots &\approx y_0 + \epsilon y_1 + \ldots - 2\epsilon \ln\left[y_0 + \epsilon y_1 + \ldots\right] \\ &\approx y_0 + \epsilon y_1 + \ldots - 2\epsilon\left(\ln[y_0] + \epsilon\frac{y_1}{y_0} + \ldots\right) \\ 0 &= (y_0' - y_0) + \epsilon\left(y_1' - y_1 + 2\ln[y_0]\right) + \ldots, \end{aligned}$$

where we have dropped all terms of higher order in ϵ. Since ϵ is arbitrary, the expressions in each bracket $(\ldots)$ must vanish separately, i. e.

$$\begin{aligned} \epsilon^0: \qquad y_0' - y_0 &= 0, \\ \epsilon^1: \qquad y_1' - y_1 &= -2\ln[y_0], \\ &\ldots. \end{aligned} \tag{87}$$

We are thus faced with having to solve a *infinite hierarchy* of easy problems to approximate our hard problem, Eq. (84). The equation for y_0 in Eq. (87) has the expected solution (by design) of $y_0(x) = e^x$. For the first correction, we obtain

$$y_1' - y_1 = -2x,$$

a simple linear, homogeneous problem with the general solution $y_1(x) = A_1 e^x + 2x + 2$, where we find $A_1 = -2$ for $y_1(0) = 0$. Therefore, our first approximant to the hard problem in Eq. (84) is

$$y(x,\epsilon) \approx e^x + \epsilon\left[2x + 2 - 2e^x\right] + \ldots \to y(x) = y(x,1) \approx 2x + 2 - e^x. \tag{88}$$

This being a Taylor series, one would actually have some freedom in interpreting this approximant, for instance, we could write with equal justification

$$y(x,\epsilon) \approx \frac{e^x}{1 - 2\epsilon\left[(x+1)e^{-x} - 1\right]} + \ldots \to y(x) = y(x,1) \approx \frac{e^x}{3 - 2(x+1)e^{-x}}. \tag{89}$$

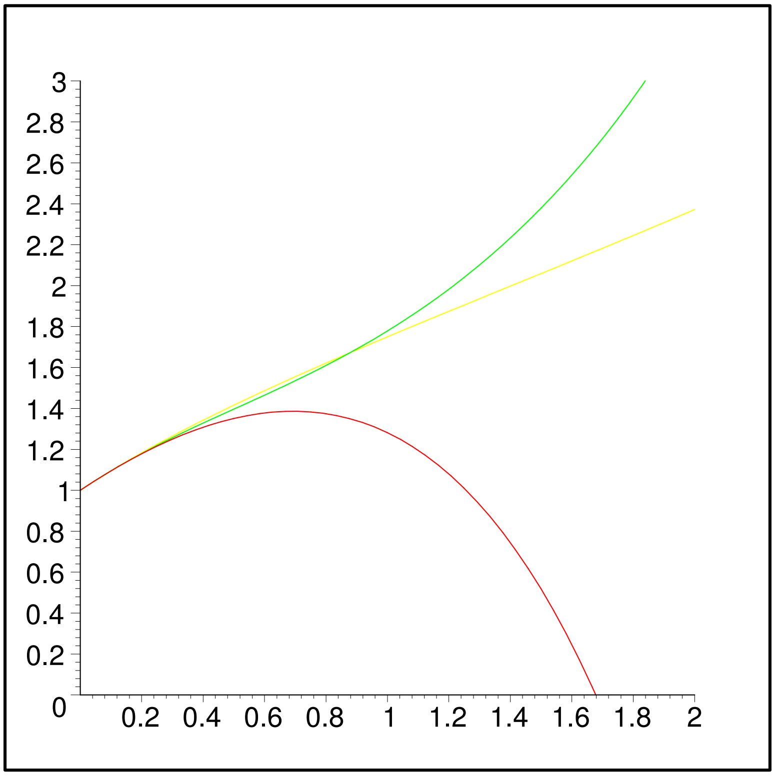

Figure 2: Plot of the numerical solution of Eq. (83) (top curve), the Taylor approximant in Eq. (88) (bottom curve), and the rational approximant in Eq. (89) (middle curve).

Finally, extracting information from the perturbative series is somewhat of an art, in fact.

B. A simpler Example

Applying perturbative expansions to differential equation is of course very useful, but makes understanding the technique unnecessarily difficult. To understand the nature of perturbative expansions, it is much easier to consider a simple problem, where we know both the "easy" *and* the "hard" problem (and everything in-between) in all detail. And instead of solving for a function $y(x)$, we just want to find the root of an equation, say,

$$x^2 + x = 2. \tag{90}$$

It is obvious that $x = 1$ and $x = -2$ are the two possible solutions to this equation. Even if we introduce a one-parameter family of problems, such as

$$\epsilon x^2 + x = 2, \tag{91}$$

the solution is simple:

$$x_\pm = x_\pm(\epsilon) = -\frac{1}{2\epsilon}\left[1 \pm \sqrt{1+8\epsilon}\right]. \tag{92}$$

Yet, we want to solve the problem by a perturbative expansion,

$$x = x(\epsilon) \approx x_0 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \epsilon^4 x_4 + \ldots, \tag{93}$$

to understand its properties. We insert Eq. (93) into Eq. (91) and obtain

$$\begin{aligned}
\epsilon\left(x_0 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \ldots\right)^2 + x_0 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \epsilon^4 x_4 + \ldots &= 2, \\
(x_0 - 2) + \epsilon\left(x_0^2 + x_1\right) + \epsilon^2\left(2x_0 x_1 + x_2\right) + \epsilon^3\left(x_1^2 + 2x_0 x_2 + x_3\right) & \\
+\epsilon^4\left(2x_1 x_2 + 2x_0 x_3 + x_4\right) + \ldots &= 0.
\end{aligned}$$

Since the last statement is supposed to hold for *any* ϵ (between 0 and 1, at least), one can show that *each* $(\ldots)$ must vanish independently:

$$\begin{aligned}
\epsilon^0 : \quad 0 &= x_0 - 2 \\
\epsilon^1 : \quad 0 &= x_0^2 + x_1 \\
\epsilon^2 : \quad 0 &= 2x_0 x_1 + x_2 \\
\epsilon^3 : \quad 0 &= x_1^2 + 2x_0 x_2 + x_3 \\
\epsilon^4 : \quad 0 &= 2x_1 x_2 + 2x_0 x_3 + x_4 \\
&\ldots
\end{aligned} \tag{94}$$

providing us with a infinite sequence of simple problems, similar to Eq. (87). Some notes are in order:

1. While the "hard" problem was non-linear (here: quadratic), the "easy" problem is linear.

2. Subsequent problems in Eq. (94) express at order ϵ^n the unknown x_n in terms of the already determined quantities $x_0, x_1, \ldots, x_{n-1}$.

3. Each subsequent problem in Eq. (94) has the same structure as the "easy" problem; here this means, it is also linear, $x_n = F_n(x_0, x_1, \ldots, x_{n-1})$. (The functions F_n may not be linear, but since all of its arguments are already known, it is merely a number as far as the unknown variable x_n is concerned!)

Each of these points are generic properties of a well-designed perturbation expansion. "Well-designed" here means that we choose to insert the expansion parameter ϵ in such a way that the easy problem is linear [be it in the algebraic or differential equation sense as in Eq. (84)], or otherwise solvable. In case of the differential equation above, the "easy" problem was a homogeneous linear equation, while each subsequent problem is inhomogeneous. Ever more complex inhomogeneous terms arise order-by-order, which may limit the usefulness of the expansion at some order.

Solving Eq. (94) order-by-order, we obtain

$$
\begin{aligned}
\epsilon^0: \quad & x_0 = 2 \\
\epsilon^1: \quad & x_1 = -x_0^2 = -4 \\
\epsilon^2: \quad & x_2 = -2x_0x_1 = 16 \\
\epsilon^3: \quad & x_3 = -x_1^2 - 2x_0x_2 = -80 \\
\epsilon^4: \quad & x_4 = -2x_1x_2 - 2x_0x_3 = 448 \\
& \ldots
\end{aligned}
\tag{95}
$$

Inserted into our Taylor-series Ansatz in Eq. (93), we obtain as "solution"

$$x(\epsilon) \approx 2 - 4\epsilon + 16\epsilon^2 - 80\epsilon^3 + 448\epsilon^4 + \ldots, \tag{96}$$

which evaluated at $\epsilon = 1$ provides a very poor approximation indeed [to either root of Eq. (90)]! In fact, when we consider a sequence of partial sums by truncating Eq. (96) first after the ϵ^1–term, than after the ϵ^2–term, and so on, it yields $x(1) \approx -2, 14, -66, 382, \ldots$, which is oscillating around the known roots with ever larger amplitude. The series in Eq. (96) is clearly divergent! Since we have (for this very simple problem) the general solution for $x(\epsilon)$ in Eq. (92) available, we can immediately detect the problem: Due to the singularity of the square-root, the radius of convergence for the MacLaurin series for $x(\epsilon)$ is just 1/8! After all this work, what are we to make of such an apparently useless result? The problem has to do with an annoying property of Taylor series in general: Their radius of convergence is limited by the singularity closest to the point around which we are expanding a function in the complex plane, even if this singularity is just an isolated pole, and the underlying function itself is perfectly regular anywhere else. Remember the geometric series in Eq. (4): The MacLaurin series of $G(x) = 1/(1-x)$ has only a unit radius of convergence, yet, $G(-1) = 1/2$ or even $G(10) = -1/9$, say, are perfectly acceptable values, which the series can not produce. In particular, our approximation consisting of a finite number of terms in that series, a polynomial, can never have any singularities. But there is nothing germane to writing that finite number of terms in polynomial form. We may just as well write it as a rational function, say,

$$x = \frac{a + \epsilon b + \epsilon^2 c}{1 + \epsilon d + \epsilon^2 e}, \tag{97}$$

which to 4-th order has a unique relation to Eq. (96). To this end, we simply expand Eq. (97) in ϵ to 4-th order only and match the coefficients term-by-term:

$$\begin{aligned}
x &= \left(a + \epsilon b + \epsilon^2 c\right)\left[1 - (\epsilon d + \epsilon^2 e) + (\epsilon d + \epsilon^2 e)^2 - (\epsilon d + \epsilon^2 e)^3 + (\epsilon d + \ldots)^4 + \ldots\right] \\
&= \left(a + \epsilon b + \epsilon^2 c\right)\left[1 - \epsilon d - \epsilon^2 e + \epsilon^2 d^2 + 2\epsilon^3 de + \epsilon^4 e^2 - \epsilon^3 d^3 - 3\epsilon^4 d^2 e - \ldots + \epsilon^4 d^4 + \ldots\right] \\
&= \left(a + \epsilon b + \epsilon^2 c\right)\left[1 - \epsilon d + \epsilon^2\left(d^2 - e\right) + \epsilon^3\left(2de - d^3\right) + \epsilon^4\left(e^2 - 3d^2 e + d^4\right) + \ldots\right] \\
&= a + \epsilon\,(b - ad) + \epsilon^2\left(ad^2 - ae - bd + c\right) + \epsilon^3\left(2ade - ad^3 + bd^2 - be - cd\right) \\
&\quad + \epsilon^4\left(ae^2 - 3ad^2 e + ad^4 + 2bde - bd^3 + cd^2 - ce\right) + \ldots .
\end{aligned}$$

Eq. (97) is called a "Pade-approximant." In particular, a $[n, m]-$Pade is a rational expression, which has a nth-order polynomial in the numerator and an mth-order polynomial in the denominator. It can be uniquely matched to an $(m + n)$th-order Taylor series. So, what we have in Eq. (97) is a [2,2]-Pade. Best results are typically obtained from the sequence $[0, 1], [1, 1], [1, 2], [2, 2], \ldots$of Pade approximants, although convergence is not guaranteed. The success depends on how well the poles, due to the zeros of the denominator, are able to mimic the actual singularities of the approximated expression.

Matching this series to that in Eq. (96) and solving for the letters appears to lead to a horrible set of nonlinear equations. But solved sequentially, matters again simplify drastically. Clearly, $a = 2$. Then, we have

$$b - 2d = -4, \qquad 2d^2 - 2e - bd + c = 16,$$

where the first relation, $b = 2d - 4$, linearizes the second,

$$4d - 2e + c = 16, \qquad 4de - 2d^3 + bd^2 - be - cd = 2de - 4d^2 + 4e - cd = -80.$$

Now, $c = 16 - 4d + 2e$ linearizes the latter relation,

$$4e - 16d = -80,$$

so, $d = e/4 + 5$ finally leads to $400 + 4e = 448$, from which follows

$$e = 12, \quad d = 8, \quad c = 8, \quad b = 12, \quad a = 2.$$

Hence, we obtain for Eq. (97),

$$\begin{aligned} x(\epsilon) &\approx \frac{2 + 12\epsilon + 8\epsilon^2}{1 + 8\epsilon + 12\epsilon^2} \\ x(1) &\approx \frac{22}{21} \approx 1.048, \end{aligned} \tag{98}$$

an approximation of better than 5% accuracy to the root of Eq. (90) at $x = 1$.

This approximation is clearly much improved over the diverging Taylor series in Eq. (96). It is amazing that this simple *rational* expression in Eq. (98) manages to mimic the *square-root* expression in Eq. (92). Rational expression are merely capable of poles through the zeros of their denominator. Yet, note that the two poles of Eq. (98) at $\epsilon = -1/2, -1/6$ are actually quite near the square-root singularity of Eq. (92) at $\epsilon = -1/8$. In fact, for higher-order rational approximants, their poles would tend to *accumulate* at the singularities of the true solution.

One serious problem of our approximation remains: Instead of the two solutions of the original, we only found one near $x = 1$. What happened to the second solution at $x = 2$? According to Eq. (92), this root corresponds to $x_+(\epsilon)$, which moves to infinity (as $1/\epsilon$) for $\epsilon \to 0$, highlighting the fact that this limit is singular: the "easy" problem at $\epsilon = 0$ in Eq. (91) has only one solution, while there are two for any other value of ϵ. Our naive

Taylor series Ansatz in Eq. (93) does not provide for singular terms such as $1/\epsilon$ (or more complicated singular expressions of ϵ in general) and looses all those solutions. The only remaining solution our Ansatz permits is that which *smoothly interpolates* between the easy and the hard problem, a scenario often suggested by our intuition about the physics at hand in a problem. To obtain other solutions, the singular behavior near $\epsilon \to 0$ would have to be discovered first to start out the perturbative expansion. In this simple example, a series Ansatz with a leading $1/\epsilon-$term would obviously do the trick.

EXERCISE:

1. Repeat the calculation with a leading $1/\epsilon$ factor in the Ansatz in Eq. (93) to be inserted into Eq. (91) to approximate the root at $x = -2$. (Hint: you find two solutions for x_0; choose the one that does not correspond to the already found solution.)

2. Find the perturbative expansion for

$$\sqrt{2}\sin(x) + x = \frac{\pi}{4}$$

to 4-th order, i. e. $x \approx x_0 + \epsilon x_1 + \epsilon^2 x_2 + \epsilon^3 x_3 + \epsilon^4 x_4$ by appropriately inserting a parameter ϵ into the equation. Determine $x_0, x_1, \ldots, x_4$. Express x as a rational function in ϵ (Pade approximant), as in Eq. (97) and evaluate this expression at $\epsilon = 1$. Compare with the exact result, $x = 0.32877\ldots$.

VIII. DIRAC'S δ-FUNCTION

A. Definitions and Basic Properties

In the sciences, one often deals with situations in which things happen apparently instantaneous, at least on the time scale one is interested in. For instance, a soccer ball kicked appears to change its state of motion suddenly from stand-still to high velocity. We observe the ball on a scale of seconds, while the detailed (and potentially messy) events by which forces accelerate the ball occur on a milli-second time-scale considered irrelevant for our observation. To this end, the notion of an "impulse" was created in this example. The result is somewhat un-physical (and un-mathematical), since the velocity of the object in this description changes *discontinuously* from zero to finite. Worse yet, the force, being proportional to the time-derivative of the velocity, diverges! In fact, what we have is a very large, finite force over a very small time interval resulting in an ordinary, finite integral – the impulse. Keeping the impulse constant while shrinking the time interval to zero leads to a force function which is non-zero (infinite!) only on a domain of zero support in such a way that the integral of the function is finite. Such an object is not a function in the traditional sense, and is instead called a *distribution*. It is easy to find sequences of perfectly regular functions, which have such distributions as their limit. In our example, we can approximate the force function by a (continuous) triangle-function,

$$F(t) = \Delta p \lim_{n\to\infty} \begin{cases} 0, & |t| > \frac{1}{n}, \\ n\left(1 - n|t|\right), & |t| \leq \frac{1}{n}, \end{cases} \tag{99}$$

which has an integral of $\int_{-\infty}^{\infty} dt\, F(t) = \Delta p = const$ for any n. More abstractly, we can define a so-called "$\delta-$Function" (as due to Dirac) via two axioms,

$$\begin{aligned} (1) \qquad \delta\left(x - x_0\right) &= 0, \qquad \left(x \neq x_0\right), \\ (2) \qquad \int_{x_0-\epsilon}^{x_0+\epsilon} dx\, \delta\left(x - x_0\right) &= 1, \qquad (\text{any}\epsilon > 0). \end{aligned} \tag{100}$$

There are a number of *representations* of this $\delta-$function, each with its own advantages (and dis-advantages):

1. Easiest, but itself discontinuous, is a *rectangular function* sequence,

$$\delta\left(x - x_0\right) = \lim_{n\to\infty} \begin{cases} 0, & |x - x_0| > \frac{1}{2n}, \\ n, & |x - x_0| \leq \frac{1}{2n}. \end{cases} \tag{101}$$

2. At least continuous is the *triangular function* sequence already used above,

$$\delta\left(x - x_0\right) = \lim_{n\to\infty} \begin{cases} 0, & |x - x_0| > \frac{1}{n}, \\ n\left(1 - n\left|x - x_0\right|\right), & |x - x_0| \leq \frac{1}{n}. \end{cases} \tag{102}$$

3. A perfectly regular (analytic!) function sequence often used is the *Gaussian representation*,

$$\delta\left(x - x_0\right) = \lim_{n\to\infty} \frac{n}{\sqrt{\pi}} e^{-n^2 (x-x_0)^2}. \tag{103}$$

4. Another smooth function sequence is the Lorentzian representation,

$$\delta\left(x-x_0\right)=\lim_{n\to\infty}\frac{n}{\pi}\frac{1}{1+n^2\left(x-x_0\right)^2}. \tag{104}$$

5. Even oscillatory functions (with negative values) can make a useful sequence, as in the sin −representation,

$$\delta\left(x-x_0\right)=\lim_{n\to\infty}\frac{\sin\left[n\left(x-x_0\right)\right]}{\pi\left(x-x_0\right)}. \tag{105}$$

6. From the previous sin-representation, we can also derive an often used integral-representation. It is

$$\begin{aligned}\frac{\sin(nx)}{\pi x} &= \frac{1}{2\pi}\int_{-n}^{n} dt\,\cos(xt),\\ &= \frac{1}{2\pi}\int_{-n}^{n} dt\,\left[\cos(xt)\pm i\sin(xt)\right],\\ &= \frac{1}{2\pi}\int_{-n}^{n} dt\,e^{\pm ixt},\end{aligned}$$

where we have used $\int_{-n}^{n} dt\,\sin(xt)=0$ and the Euler relation from Eq. (23). Hence, from Eq. (105),

$$\begin{aligned}\delta\left(x-x_0\right) &= \lim_{n\to\infty}\frac{1}{2\pi}\int_{-n}^{n} dt\,e^{\pm i(x-x_0)t},\\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dt\,e^{\pm i(x-x_0)t}.\end{aligned} \tag{106}$$

One of the most important properties of the δ−function is that for any a,b with $x_0\in(a..b)$ and any function f continuous at x_0, it is

$$\begin{aligned}\int_a^b dx\,f(x)\delta\left(x-x_0\right) &= \lim_{\epsilon\to0^+}\int_{x_0-\epsilon}^{x_0+\epsilon} dx\,f(x)\delta\left(x-x_0\right),\\ &= f\left(x_0\right)\lim_{\epsilon\to0^+}\int_{x_0-\epsilon}^{x_0+\epsilon} dx\,\delta\left(x-x_0\right),\\ &= f\left(x_0\right),\end{aligned} \tag{107}$$

where we have used axiom (2) in Eq. (100).

Of interest is also the variable integral of the δ−function,

$$\theta\left(x-x_0\right) = \int_{-\infty}^{x} dx'\,\delta\left(x'-x_0\right)=\begin{cases}0, & x<x_0,\\ \frac{1}{2}, & x=x_0,\\ 1, & x>x_0,\end{cases} \tag{108}$$

which is called the (unit) *step-function* or Heaviside-function. The result $\theta(0) = \frac{1}{2}$ follows from the symmetry of the $\delta-$function, $\delta(-x) = \delta(x)$, in that

$$\int_{x_0}^{x_0+\epsilon} dx\, \delta\left(x - x_0\right) = \frac{1}{2} = \int_{x_0-\epsilon}^{x_0} dx\, \delta\left(x - x_0\right). \tag{109}$$

EXERCISES:

1. Show for the representations 1.-5. that the axioms in Eq. (100) are satisfied.

2. Using any of the representations to show that

$$\delta(ax) = \frac{1}{|a|}\delta(x). \tag{110}$$

3. Use the previous result, Eq. (110), to argue that

$$\delta\left(x^2 - a^2\right) = \frac{1}{2|a|}\left[\delta\left(x + a\right) + \delta\left(x - a\right)\right]. \tag{111}$$

B. Using the $\delta-$Function

There is a myriad of uses for the $\delta-$function, but most involve the use of differential equations. Instead, we can consider a powerful application in statistics. We want to consider the case of adding random numbers. For example, we know that the roll of one dye X has 6 possible outcomes, each equally likely, $p_1(X = 1) = p_1(X = 2) = \ldots = p_1(X = 6) = \frac{1}{6}$. The same is of course true for a second dye Y. A practicle question (i. e. for players in "Monopoly") is, What is the probability $p_2(Z)$ of rolling both dice and obtaining an outcome $Z = X + Y$? A general way to calculate $p_s(Z)$ is to sum over the entire probability space for all pairs (X, Y) under the constraint $Z = X + Y$,

$$p_2(Z) = \sum_X \sum_Y p_1(X)\, p_1(Y)|_{Z=X+Y}. \tag{112}$$

A simple example is the case of $Z = 3$:

$$\begin{aligned} p_2(3) &= \left[p_1(X = 1) + \ldots + p_1(X = 6)\right]\left[p_1(Y = 1) + \ldots + p_1(Y = 6)\right]|_{X+Y=3}, \\ &= p_1(X = 1)p_1(Y = 2) + p_1(X = 2)p_1(Y = 1), \\ &= \frac{1}{6}\frac{1}{6} + \frac{1}{6}\frac{1}{6} = \frac{1}{18}. \end{aligned}$$

Note that while $p_1(X) \equiv \frac{1}{6}$ was uniform over all outcomes X, $p_2(Z)$ is not. There are 36 possible outcomes (X, Y), only one has $Z = 2$ ("Snake-eyes") or $Z = 12$, i. e. $p_2(Z = 2) = p_2(Z = 12) = \frac{1}{36}$, then $p_2(Z = 3) = p_2(Z = 11) = \frac{1}{18}$, and so on. It is essential for "Monopoly" players to know that the most likely individual outcome is $Z = 7$ with $p_2(Z = 7) = \frac{1}{6}$!

Switching from discrete probabilities to the continuum, a similar formulism can be applied to determine the probability distribution for the sum of two or more *iid* (*i*ndependent, *i*dentically *d*istributed) random numbers. Say, we add two numbers, (X, Y), randomly drawn from a basic distribution p_1, What is the distribution $p_2(Z = X + Y)$? Well, we can simply generalize the notion developed in Eq. (112) and write

$$p_2(Z) = \int_{-\infty}^{\infty} dX \int_{-\infty}^{\infty} dY \, p_1(X) p_1(Y)|_{Z=X+Y}$$

How do we handle the constraint in any actual computation of the integral? What the constraint amounts to is replacing $Y = Z - X$ or $X = Z - Y$ at the expense of the respective integration, i. e.

$$p_2(Z) = \int_{-\infty}^{\infty} dX \, p_1(X) p_1(Z - X). \tag{113}$$

In fact, this is just what a $\delta-$function accomplishes,

$$p_2(Z) = \int_{-\infty}^{\infty} dX \int_{-\infty}^{\infty} dY \, p_1(X) p_1(Y) \, \delta \left(X + Y - Z \right), \tag{114}$$

which leads right back Eq. (113) after integration over Y, according to Eq. (107).

It is easy to extend this idea to the sum of n random numbers (n coin tosses, or n roll of the dice, etc). Eq. (116) simply generalizes to

$$p_n(Z) = \int_{-\infty}^{\infty} dX_1 \int_{-\infty}^{\infty} dX_2 \ldots \int_{-\infty}^{\infty} dX_n \, p_1(X_1) p_1(X_2) \ldots p_1(X_n) \, \delta \left(X_1 + X_2 + \ldots + X_n - Z \right) \tag{115}$$

In fact, we have already shown in Sec. III C by elementary counting that for n coin tosses, $p_n(Z)$ is given by the binomial distribution, and that for large n, $p_n(Z)$ approaches a Normal distribution. We will find this to be a much more general result, virtually independent of the elementary distribution of each random number, $p_1(X)$.

To appreciate the power of the $\delta-$function, let us try to find $p_2(Z)$ and $p_3(Z)$ the "pedestrian" way for the special case of the normal distribution in Eq. (34) for $p_1(X)$, but with zero mean and unit variance ($\mu = 0$ and $\sigma = 1$). Then, Eq. (116) reads:

$$
\begin{aligned}
p_2(Z) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} dX\, e^{-\frac{1}{2}X^2}\, e^{-\frac{1}{2}(Z-X)^2}, \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} dX\, e^{-X^2+ZX}\, e^{-\frac{1}{2}Z^2}, \\
&= \frac{1}{\sqrt{4\pi}}\, e^{-\frac{1}{4}Z^2},
\end{aligned}
\tag{116}
$$

giving us a distribution of *twice* the variance of the original. That calculation was easy enough, but let us consider the corresponding calculation for $n = 3$ in Eq. (115) now:

$$
\begin{aligned}
p_3(Z) &= \frac{1}{(2\pi)^{\frac{3}{2}}}\int_{-\infty}^{\infty} dX_1 \int_{-\infty}^{\infty} dX_2 \int_{-\infty}^{\infty} dX_3\, e^{-\frac{1}{2}X_1^2}\, e^{-\frac{1}{2}X_2^2}\, e^{-\frac{1}{2}X_3^2}\,\delta\left(X_1+X_2+X_3-Z\right), \\
&= \frac{1}{(2\pi)^{\frac{3}{2}}}\int_{-\infty}^{\infty} dX_1 \int_{-\infty}^{\infty} dX_2\, e^{-\frac{1}{2}X_1^2}\, e^{-\frac{1}{2}X_2^2}\, e^{-\frac{1}{2}(Z-X_1-X_2)^2}, \\
&= \frac{e^{-\frac{1}{2}Z^2}}{(2\pi)^{\frac{3}{2}}}\int_{-\infty}^{\infty} dX_1\, e^{-X_1^2+ZX_1} \int_{-\infty}^{\infty} dX_2\, e^{-X_2^2+(Z-X_1)X_2}, \\
&= \frac{e^{-\frac{1}{4}Z^2}}{\sqrt{8\pi}}\int_{-\infty}^{\infty} dX_1\, e^{-\frac{3}{4}X_1^2+\frac{1}{2}ZX_1}, \\
&= \frac{1}{\sqrt{6\pi}}\, e^{-\frac{1}{6}Z^2},
\end{aligned}
\tag{117}
$$

where we have used Eq. (28) repeatedly. Note that we obtain for $p_3(Z)$ a distribution of *trice* the variance of the original.

Well, the calculation in Eq. (117) was not pretty! And the algebraic effort in applying Eq. (28) will only become more messy for larger n. Yet, a clever application of the $\delta-$function that *preserves the symmetry* between $X_1, \ldots, X_n$ in Eq. (115) not only solves the problem for *arbitrary* n, but also provides a much deeper insight into the general problem. To this end we replace the $\delta-$function with its integral representation in Eq. (106), apparently increasing the number of integrations instead of reducing it:

$$
\begin{aligned}
p_n(Z) &= \int_{-\infty}^{\infty} dX_1 \int_{-\infty}^{\infty} dX_2 \ldots \int_{-\infty}^{\infty} dX_n\, p_1(X_1)p_1(X_2)\ldots p_1(X_n) \int_{-\infty}^{\infty}\frac{dt}{2\pi}\, e^{-it(X_1+X_2+\ldots+X_n-Z)}, \\
&= \int_{-\infty}^{\infty}\frac{dt}{2\pi}\, e^{itZ}\left[\int_{-\infty}^{\infty} dX_1\, p_1(X_1)\, e^{-itX_1}\right]\left[\int_{-\infty}^{\infty} dX_2\, p_1(X_2)\, e^{-itX_2}\right]\ldots\left[\int_{-\infty}^{\infty} dX_n\, p_1(X_n)\, e^{-itX_n}\right], \\
&= \int_{-\infty}^{\infty}\frac{dt}{2\pi}\, e^{itZ}\left[\int_{-\infty}^{\infty} dX\, p_1(X)\, e^{-itX}\right]^n.
\end{aligned}
\tag{118}
$$

By symmetry, the $n-$fold integration has magically collapse into a single integral. Of course, it is still not obvious that the determination of $p_n(Z)$ has become much simpler in light of the outer integral with respect to t. The biggest value of the result in Eq. (118) concerns the *explicit* appearance of n, not as an index but as a parameter, for which we can study the limit $n \to \infty$ asymptotically using the methods of Sec. V B. First, though, we can verify also that the $\delta-$function makes life easier for the special case of a normal distribution for $p_1(X)$. We find

$$\begin{aligned} p_n(Z) &= \int_{-\infty}^{\infty} \frac{dt}{2\pi}\, e^{itZ} \left[\int_{-\infty}^{\infty} \frac{dX}{\sqrt{2\pi}}\, e^{-\frac{1}{2}X^2 - itX} \right]^n, \\ &= \int_{-\infty}^{\infty} \frac{dt}{2\pi}\, e^{itZ - \frac{n}{2}t^2}, \\ &= \frac{1}{\sqrt{2n\pi}}\, e^{-\frac{1}{2n}Z^2}, \end{aligned} \tag{119}$$

for *any* $n > 0$, in particular, for $n = 2, 3$ as in Eqs. (116-117). This calculation is undoubtedly easier than those “pedestrian” attempts!