Séminaire Lotharingien de Combinatoire **44** (2000), Article B44d

Mathemagics* (A Tribute to L. Euler and R. Feynman)

Pierre Cartier**

CNRS, Ecole Normale Supérieure de Paris, 45 rue d'Ulm, 75230 Paris Cedex 05

To the memory of Gian-Carlo Rota,
modern master of mathematical magical tricks

Table of contents

---

* Lectures given at a school held in Chapelle des Bois (April 5-10, 1999) on "Noise, oscillators and algebraic randomness".

** cartier@ihes.fr

1 Introduction

The implicit philosophical belief of the working mathematician is today the Hilbert-Bourbaki formalism. Ideally, one works within a **closed system**: the basic principles are clearly enunciated once for all, including (that is an addition of twentieth century science) the formal rules of logical reasoning clothed in mathematical form. The basic principles include precise definitions of all mathematical objects, and the coherence between the various branches of mathematical sciences is achieved through reduction to basic models in the universe of sets. A very important feature of the system is its **non-contradiction** ; after Gödel, we have lost the initial hopes to establish this non-contradiction by a formal reasoning, but one can live with a corresponding belief in non-contradiction. The whole structure is certainly very appealing, but the illusion is that it is eternal, that it will function for ever according to the same principles. What history of mathematics teaches us is that the principles of mathematical deduction, and not simply the mathematical theories, have evolved over the centuries. In modern times, theories like General Topology or Lebesgue's Integration Theory represent an almost perfect model of precision, flexibility and harmony, and their applications, for instance to probability theory, have been very successful.

My thesis is: **there is another way of doing mathematics, equally successful, and the two methods should supplement each other and not fight.**

This other way bears various names: symbolic method, operational calculus, operator theory ... Euler was the first to use such methods in his extensive study of infinite series, convergent as well as divergent. The calculus of differences was developed by G. Boole around 1860 in a symbolic way, then Heaviside created his own symbolic calculus to deal with systems of differential equations in electric circuitry. But the modern master was R. Feynman who used his diagrams, his disentangling of operators, his path integrals ... The method consists in stretching the formulas to their extreme consequences, resorting to some internal feeling of coherence and harmony. They are obvious pitfalls in such methods, and only experience can tell you that for the Dirac δ-function an expression like $x\delta(x)$ or $\delta'(x)$ is lawful, but not $\delta(x)/x$ or $\delta(x)^2$. Very often, these so-called symbolic methods have been substantiated by later rigorous developments, for instance Schwartz distribution theory gives a rigorous meaning to $\delta(x)$, but physicists used sophisticated formulas in “momentum space” long before Schwartz codified the Fourier transformation for distributions. The Feynman “sums over histories” have been immensely successful in many problems, coming from physics as well from mathematics, despite the lack of a comprehensive rigorous theory.

To conclude, I would like to offer some remarks about the word “formal”. For the mathematician, it usually means “according to the standard of formal rigor, of formal logic”. For the physicists, it is more or less synonymous with “heuristic” as opposed to “rigorous”. It is very often a source of misunderstanding between these two groups of scientists.

2 A new look at the exponential

2.1 The power of exponentials

The multiplication of numbers started as a shorthand for repeated additions, for instance 7 times 3 (or rather “seven taken three times”) is the sum of three terms equal to 7

$$7 \times 3 = \underbrace{7 + 7 + 7}_{3 \text{ times}}.$$

In the same vein 7^3 (so denoted by Viete and Descartes) means $\underbrace{7 \times 7 \times 7}_{3 \text{ factors}}$.
There is no difficulty to define x^2 as xx or x^3 as xxx for any kind of multiplication (numbers, functions, matrices ...) and Descartes uses interchangeably xx or x^2, xxx or x^3.

In the exponential (or power) notation, the exponent plays the role of an operator. A great progress, taking approximately the years from 1630 to 1680 to accomplish, was to generalize a^b to new cases where the operational meaning of the exponent b was much less visible. By 1680, a well defined meaning has been assigned to a^b for a, b real numbers, $a > 0$. Rather than to retrace the historical route, we shall use a formal analogy with vector algebra. From the original definition of a^b as $a \times \ldots \times a$ (b factors), we deduce the fundamental rules of operation, namely

$$(a \times a')^b = a^b \times a'^b,\ a^{b+b'} = a^b \times a^{b'},\ (a^b)^{b'} = a^{bb'},\ a^1 = a. \tag{1}$$

The other rules for manipulating powers are easy consequences of the rules embodied in (1). The fundamental rules for **vector algebra** are as follows:

$$\begin{aligned}(\mathbf{v} + \mathbf{v}').\lambda = \mathbf{v}.\lambda + \mathbf{v}'.\lambda,\ \mathbf{v}.(\lambda + \lambda') = \mathbf{v}.\lambda + \mathbf{v}.\lambda',\\ (\mathbf{v}.\lambda).\lambda' = \mathbf{v}.(\lambda\lambda'),\ \mathbf{v}.1 = \mathbf{v}.\end{aligned} \tag{2}$$

The analogy is striking provided we compare the product $a \times a'$ of numbers to the sum $\mathbf{v} + \mathbf{v}'$ of vectors, and the exponentiation a^b to the scaling $\mathbf{v}.\lambda$ of the vector $\mathbf{v}$ by the scalar λ.

In modern terminology, to define a^b for a, b real, $a > 0$ means that we want to consider the set $\mathbf{R}_+^\times$ of real numbers $a > 0$ as a vector space over the field of real numbers $\mathbf{R}$. But to vectors, one can assign coordinates: if the coordinates of the vector $\mathbf{v}(\mathbf{v}')$ are the $v_i(v_i')$, then the coordinates of $\mathbf{v} + \mathbf{v}'$ and $\mathbf{v}.\lambda$ are respectively $v_i + v_i'$ and $v_i.\lambda$. Since we have only one degree of freedom in $\mathbf{R}_+^\times$, we should need one coordinate, that is a bijective map L from $\mathbf{R}_+^\times$ to $\mathbf{R}$ such that

$$L(a \times a') = L(a) + L(a'). \tag{3}$$

Once such a logarithm L has been constructed, one defines a^b in such a way that $L(a^b) = L(a).b$. It remains the daunting task to construct a logarithm. With hindsight, and using the tools of calculus, here is the simple definition of "natural logarithms"

$$\ln(a) = \int_1^a dt/t \quad \text{for } a > 0. \tag{4}$$

In other words, the logarithm function $\ln(t)$ is the primitive of $1/t$ which vanishes for $t = 1$. The inverse function $\exp s$ (where $t = \exp s$ is synonymous to $\ln(t) = s$) is defined for all real s, with positive values, and is the unique solution to the differential equation $f' = f$ with initial value $f(0) = 1$. The final definition of powers is then given by

$$a^b = \exp(\ln(a).b). \tag{5}$$

If we denote by e the unique number with logarithm equal to 1 (hence $e = 2.71828\ldots$), the exponential is given by $\exp a = e^a$.

The main character in the exponential is the exponent, as it should be, in complete reversal from the original view where 2 in x^2, or 3 in x^3 are mere markers.

2.2 Taylor's formula and exponential

We deal with the expansion of a function $f(x)$ around a fixed value x_0 of x, in the form

$$f(x_0 + h) = c_0 + c_1 h + \cdots + c_p h^p + \cdots. \tag{6}$$

This can be an infinite series, or simply a finite order expansion (include then a remainder). If the function $f(x)$ admits sufficiently many derivatives, we can deduce from (6) the chain of relations

$$\begin{aligned} f'(x_0+h) &= c_1 + 2c_2h + \cdots \\ f''(x_0+h) &= 2c_2 + 6c_3h + \cdots \\ f'''(x_0+h) &= 6c_3 + 24c_4h + \cdots \ . \end{aligned}$$

By putting $h = 0$, deduce

$$f(x_0) = c_0, \quad f'(x_0) = c_1, \quad f''(x_0) = 2c_2, \ldots$$

and in general $f^{(p)}(x_0) = p!c_p$. Solving for the c_p's and inserting into (6) we get Taylor's expansion

$$f(x_0+h) = \sum_{p \geq 0} \frac{1}{p!} f^{(p)}(x_0) h^p. \tag{7}$$

Apply this to the case $f(x) = \exp x$, $x_0 = 0$. Since the function f is equal to its own derivative f', we get $f^{(p)} = f$ for all p's, hence $f^{(p)}(0) = f(0) = e^0 = 1$. The result is

$$\exp h = \sum_{p \geq 0} \frac{1}{p!} h^p. \tag{8}$$

This is one of the most important formulas in mathematics. The idea is that this series can now be used to define the exponential of large classes of mathematical objects: complex numbers, matrices, power series, operators. For the modern mathematician, a natural setting is provided by a complete normed algebra A, with norm satisfying $||ab|| \leq ||a|| \cdot ||b||$. For any element a in A, we define exp a as the sum of the series $\sum_{p \geq 0} a^p/p!$, and the inequality

$$||a^p/p!|| \leq ||a||^p/p! \tag{9}$$

shows that the series is absolutely convergent.

But this would not exhaust the power of the exponential. For instance, if we take (after Leibniz) the step to denote by $\mathbf{D}f$ the derivative of f, $\mathbf{D}^2 f$ the second derivative, etc... (another instance of the exponential notation!), then Taylor's formula reads as

$$f(x+h) = \sum_{p \geq 0} \frac{1}{p!} h^p \mathbf{D}^p f(x). \tag{10}$$

This can be interpreted by saying that the **shift operator** $\mathbf{T}_h$ taking a function $f(x)$ into $f(x+h)$ is equal to $\sum_{p\geq 0} \frac{1}{p!} h^p \mathbf{D}^p$, that is, to the exponential $\exp h\mathbf{D}$ (question: who was the first mathematician to cast Taylor's formula in these terms?). Hence the obvious operator formula $\mathbf{T}_{h+h'} = \mathbf{T}_h \cdot \mathbf{T}_{h'}$ reads as

$$\exp(h + h')\mathbf{D} = \exp h\mathbf{D} \cdot \exp h'\mathbf{D}. \tag{11}$$

Notice that for numbers, the logarithmic rule is

$$\ln(a \cdot a') = \ln(a) + \ln(a') \tag{12}$$

according to the historical aim of reducing via logarithms the multiplications to additions. By inversion, the exponential rule is

$$\exp(a + a') = \exp(a) \cdot \exp(a'). \tag{13}$$

Hence formula (11) is obtained from (13) by substituting $h\mathbf{D}$ to a and $h'\mathbf{D}$ to a'.

But life is not so easy. If we take two matrices A and B and calculate $\exp(A + B)$ and $\exp A \cdot \exp B$ by expansion we get

$$\exp(A + B) = I + (A + B) + \frac{1}{2}(A + B)^2 + \frac{1}{6}(A + B)^3 + \cdots \tag{14}$$

$$\begin{aligned} \exp A. \exp B = I + (A + B) + \frac{1}{2}(A^2 + 2AB + B^2) \\ + \frac{1}{6}(A^3 + 3A^2B + 3AB^2 + B^3) + \cdots . \end{aligned} \tag{15}$$

If we compare the terms of degree 2 we get

$$\frac{1}{2}(A + B)^2 = \frac{1}{2}(A^2 + AB + BA + B^2) \tag{16}$$

in (14) and not $\frac{1}{2}(A^2+2AB+B^2)$. Harmony is restored if A and B commute: indeed $AB = BA$ entails

$$A^2 + AB + BA + B^2 = A^2 + 2AB + B^2 \tag{17}$$

and more generally the **binomial formula**

$$(A + B)^n = \sum_{i=0}^{n} \binom{n}{i} A^i B^{n-i} \tag{18}$$

for any $n \geq 0$. By summation one gets

$$\exp(A+B) = \exp A \cdot \exp B \tag{19}$$

if A and B commute, but not in general. The success in (11) comes from the obvious fact that $h\mathbf{D}$ commutes to $h'\mathbf{D}$ since numbers commute to (linear) operators.

2.3 Leibniz's formula

Leibniz's formula for the higher order derivatives of the product of two functions is the following one

$$\mathbf{D}^n(fg) = \sum_{i=0}^{n} \binom{n}{i} \mathbf{D}^i f . \mathbf{D}^{n-i} g. \tag{20}$$

The analogy with the binomial theorem is striking and was noticed early. Here are possible explanations. For the shift operator, we have

$$\mathbf{T}_h = \exp h\mathbf{D} \tag{21}$$

by Taylor's formula and

$$\mathbf{T}_h(fg) = \mathbf{T}_h f \cdot \mathbf{T}_h g \tag{22}$$

by an obvious calculation. Combining these formulas we get

$$\sum_{n \geq 0} \frac{1}{n!} h^n \mathbf{D}^n(fg) = \sum_{i \geq 0} \frac{1}{i!} h^i \mathbf{D}^i f . \sum_{j \geq 0} \frac{1}{j!} h^j \mathbf{D}^j g; \tag{23}$$

equating the terms containing the same power h^n of h, one gets

$$\mathbf{D}^n(fg) = \sum_{i+j=n} \frac{n!}{i!j!} \mathbf{D}^i f \cdot \mathbf{D}^j g \tag{24}$$

that is, Leibniz's formula.

Another explanation starts from the case $n = 1$, that is

$$\mathbf{D}(fg) = \mathbf{D}f \cdot g + f \cdot \mathbf{D}g. \tag{25}$$

In a heuristic way it means that $\mathbf{D}$ applied to a product fg is the sum of two operators $\mathbf{D}_1$ acting on f only and $\mathbf{D}_2$ acting on g only. These actions being independent, $\mathbf{D}_1$ commutes to $\mathbf{D}_2$ hence the binomial formula

$$\mathbf{D}^n = (\mathbf{D}_1 + \mathbf{D}_2)^n = \sum_{i=0}^{n} \binom{n}{i} \mathbf{D}_1^i \cdot \mathbf{D}_2^{n-i}. \tag{26}$$

By acting on the product fg and observing that $\mathbf{D}_1^i \cdot \mathbf{D}_2^j$ transforms fg into $\mathbf{D}^i f \cdot \mathbf{D}^j g$, one recovers Leibniz's formula. In more detail, to calculate $\mathbf{D}^2(fg)$, one applies $\mathbf{D}$ to $\mathbf{D}(fg)$. Since $\mathbf{D}(fg)$ is the sum of two terms $\mathbf{D}f \cdot g$ and $f \cdot \mathbf{D}g$ apply $\mathbf{D}$ to $\mathbf{D}f \cdot g$ to get $\mathbf{D}(\mathbf{D}f)g + \mathbf{D}f \cdot \mathbf{D}g$ and to $f \cdot \mathbf{D}g$ to get $\mathbf{D}f \cdot \mathbf{D}g + f \cdot \mathbf{D}(\mathbf{D}g)$, hence the sum

$$\begin{aligned} &\mathbf{D}(\mathbf{D}f) \cdot g + \mathbf{D}f \cdot \mathbf{D}g + \mathbf{D}f \cdot \mathbf{D}g + f \cdot \mathbf{D}(\mathbf{D}g) \\ &\quad = \mathbf{D}^2 f \cdot g + 2\mathbf{D}f \cdot \mathbf{D}g + f \cdot \mathbf{D}^2 g. \end{aligned}$$

This last proof can rightly be called "formal" since we act on the formulas, not on the objects: $\mathbf{D}_1$ transforms $f \cdot g$ into $\mathbf{D}f \cdot g$ but this doesn't mean that from the equality of functions $f_1 \cdot g_1 = f_2 \cdot g_2$ one gets $\mathbf{D}f_1 \cdot g_1 = \mathbf{D}f_2 \cdot g_2$ (counterexample: from $fg{=}gf$, we cannot infer $\mathbf{D}f \cdot g = \mathbf{D}g \cdot f$). The modern explanation is provided by the notion of tensor products: if V and W are two vector spaces (over the real numbers as coefficients, for instance), equal or distinct, there exists a new vector space $V \otimes W$ whose elements are formal finite sums $\sum_i \lambda_i (v_i \otimes w_i)$ (with scalars λ_i and v_i in V, w_i in W); we take as basic rules the consequences of the fact that $v \otimes w$ is bilinear in v, w, but nothing more. Taking V and W to be the space $C^\infty(I)$ of the functions defined and indefinitely differentiable in an interval I of $\mathbf{R}$, we define the operators $\mathbf{D}_1$ and $\mathbf{D}_2$ in $C^\infty(I) \otimes C^\infty(I)$ by

$$\mathbf{D}_1(f \otimes g) = \mathbf{D}f \otimes g, \quad \mathbf{D}_2(f \otimes g) = f \otimes \mathbf{D}g. \tag{27}$$

The two operators $\mathbf{D}_1\mathbf{D}_2$ and $\mathbf{D}_2\mathbf{D}_1$ transform $f \otimes g$ into $\mathbf{D}f \otimes \mathbf{D}g$, hence $\mathbf{D}_1$ and $\mathbf{D}_2$ commute. Define $\bar{\mathbf{D}}$ as $\mathbf{D}_1 + \mathbf{D}_2$ hence

$$\bar{\mathbf{D}}(f \otimes g) = \mathbf{D}f \otimes g + f \otimes \mathbf{D}g. \tag{28}$$

We can now calculate $\bar{\mathbf{D}}^n = (\mathbf{D}_1 + \mathbf{D}_2)^n$ by the binomial formula as in (26) with the conclusion

$$\bar{\mathbf{D}}^n(f \otimes g) = \sum_{i=0}^{n} \binom{n}{i} \mathbf{D}^i f \otimes \mathbf{D}^{n-i} g. \tag{29}$$

The last step is to go from (29) to (20). The rigorous reasoning is as follows. There is a linear operator μ taking $f \otimes g$ into $f \cdot g$ and mapping $C^\infty(I) \otimes C^\infty(I)$ into $C^\infty(I)$; this follows from the fact that the product $f \cdot g$ is bilinear in f and g. The formula (25) is expressed by $\mathbf{D} \circ \mu = \mu \circ \bar{\mathbf{D}}$ in operator terms, according to the diagram:

$$\begin{array}{ccc} C^\infty(I) \otimes C^\infty(I) & \xrightarrow{\mu} & C^\infty(I) \\ \bar{\mathbf{D}} \downarrow & & \downarrow \mathbf{D} \\ C^\infty(I) \otimes C^\infty(I) & \xrightarrow{\mu} & C^\infty(I). \end{array}$$

An easy induction entails $\mathbf{D}^n \circ \mu = \mu \circ \bar{\mathbf{D}}^n$, and from (29) one gets

$$\begin{aligned} \mathbf{D}^n(fg) &= \mathbf{D}^n(\mu(f \otimes g)) = \mu(\bar{\mathbf{D}}^n(f \otimes g)) \\ &= \mu(\sum_{i=0}^{n} \binom{n}{i} \mathbf{D}^i f \otimes \mathbf{D}^{n-i} g) = \sum_{i=0}^{n} \binom{n}{i} \mathbf{D}^i f \cdot \mathbf{D}^{n-i} g. \end{aligned} \tag{30}$$

In words: first replace the ordinary product $f \cdot g$ by the neutral tensor product $f \otimes g$, perform all calculations using the fact that $\mathbf{D}_1$ commutes with $\mathbf{D}_2$, then restore the product $\cdot$ in place of $\otimes$.

When the vector spaces V and W consist of functions of one variable, the tensor product $f \otimes g$ can be interpreted as the function $f(x)g(y)$ in two variables x, y; moreover $\mathbf{D}_1 = \partial/\partial x$, $\mathbf{D}_2 = \partial/\partial y$ and μ takes a function $F(x, y)$ of two variables into the one-variable function $F(x, x)$ hence $f(x)g(y)$ into $f(x)g(x)$ as it should. Formula (25) reads now

$$\frac{\partial}{\partial x}(f(x)g(x)) = \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) f(x)g(y)\Big|_{y=x}. \tag{31}$$

The previous "formal" proof is just a rephrasing of a familiar proof using Schwarz's theorem that $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ commute.

Starting from the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ of two vector spaces, one can iterate and obtain

$$\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3, \quad \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3 \otimes \mathcal{H}_4, \ldots.$$

Using once again the exponential notation, $\mathcal{H}^{\otimes n}$ is the tensor product of n copies of $\mathcal{H}$, with elements of the form $\sum \lambda \cdot (\psi_1 \otimes \ldots \otimes \psi_n)$. In quantum physics, $\mathcal{H}$ represents the state vectors of a particle, and $\mathcal{H}^{\otimes n}$ represents the state vectors of a system of n independent particles of the same kind. If H is an operator in $\mathcal{H}$ representing for instance the energy of a particle, we define n operators H_i in $\mathcal{H}^{\otimes n}$ by

$$H_i(\psi_1 \otimes \ldots \otimes \psi_n) = \psi_1 \otimes \cdots \otimes H\psi_i \otimes \cdots \otimes \psi_n \tag{32}$$

(the energy of the i-th particle). Then $H_1, \ldots, H_n$ commute pairwise and $H_1 + \cdots + H_n$ is the total energy if there is no interaction. Usually, there is a pair interaction represented by an operator V in $\mathcal{H} \otimes \mathcal{H}$; then the total energy is given by $\sum_{i=1}^n H_i + \sum_{i<j} V_{ij}$ with

$$V_{12}(\psi_1 \otimes \psi_2 \otimes \cdots \otimes \psi_n) = V(\psi_1 \otimes \psi_2) \otimes \psi_3 \otimes \cdots \tag{33}$$

$$V_{23}(\psi_1 \otimes \cdots \otimes \psi_n) = \psi_1 \otimes V(\psi_2 \otimes \psi_3) \otimes \cdots \otimes \psi_n \tag{34}$$

etc... There are obvious commutation relations like

$$\begin{aligned} &H_i H_j = H_j H_i \\ &H_i V_{jk} = V_{jk} H_i \quad \text{if} \quad i, j, k \quad \text{are distinct.} \end{aligned}$$

This is the so-called "locality principle": if two operators A and B refer to disjoint collections of particles (**a**) for A and (**b**) for B, they commute.

Faddeev and his collaborators made an extensive use of this notation in their study of quantum integrable systems. Also, Hirota introduced his so-called bilinear notation for differential operators connected with classical integrable systems (solitons).

2.4 Exponential vs. logarithm

In the case of real numbers, one usually starts from the logarithm and invert it to define the exponential (called antilogarithm not so long ago). Positive numbers have a logarithm; what about the logarithm of -1 for instance?

Things are worse in the complex domain. For a complex number z, define its exponential by the convergent series

$$\exp z = \sum_{n \geq 0} \frac{1}{n!} z^n. \tag{35}$$

From the binomial formula, using the commutativity $zz' = z'z$ one gets

$$\exp(z + z') = \exp z \cdot \exp z' \tag{36}$$

as before. Separating real and imaginary part of the complex number $z = x + iy$ gives Euler's formula

$$\exp(x + iy) = e^x(\cos y + i \sin y) \tag{37}$$

subsuming trigonometry to complex analysis. The trigonometric lines are the "natural" ones, meaning that the angular unit is the radian (hence $\sin\delta \simeq \delta$ for small δ).

From an intuitive view of trigonometry, it is obvious that the points of a circle of equation $x^2 + y^2 = R^2$ can be uniquely parametrized in the form

$$x = R\ \cos\theta, \quad y = R\ \sin\theta \tag{38}$$

with $-\pi < \theta \leq \pi$, but the subtle point is to show that the geometric definition of $\sin\theta$ and $\cos\theta$ agree with the analytic one given by (37). Admitting this, every complex number $u \neq 0$ can be written as an exponential $\exp z_0$, where $z_0 = x_0 + iy_0$, x_0 real and y_0 in the interval $]-\pi, \pi]$. The number z_0 is called the principal determination of the logarithm of u, denoted by $\mathrm{Ln}(u)$. But the general solution of the equation $\exp z = u$ is given by $z = z_0 + 2\pi i n$ where n is a rational integer. Hence a nonzero complex number has infinitely many logarithms. The functional property (36) of the exponential cannot be neatly inverted: for the logarithms we can only assert that $\mathrm{Ln}(u_1 \cdots u_p)$ and $\mathrm{Ln}(u_1) + \ldots + \mathrm{Ln}(u_p)$ differ by the addition of an integral multiple of $2\pi i$.

The exponential of a (real or complex) square matrix A is defined by the series

$$\exp A = \sum_{n\geq 0} \frac{1}{n!} A^n. \tag{39}$$

There are two classes of matrices for which the exponential is easy to compute:

a) Let A be diagonal $A = \mathrm{diag}(a_1, \ldots, a_n)$. Then $\exp A$ is diagonal with elements $\exp a_1, \ldots, \exp a_n$. Hence any **complex** diagonal matrix with non zero diagonal elements is an exponential, hence admits a logarithm, and even infinitely many ones.

b) Suppose that A is a special upper triangular matrix, with zeroes on the diagonal, of the type

$$A = \begin{pmatrix} 0 & a & b & c \\ & 0 & d & e \\ & & 0 & f \\ & & & 0 \end{pmatrix}.$$

Then $A^d = 0$ if A is of size $d \times d$. Hence $\exp A$ is equal to $I + B$ where B is of the form $A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \cdots + \frac{1}{(d-1)!}A^{d-1}$. Hence B is again a special upper triangular matrix and A can be recovered by the formula

$$A = B - \frac{B^2}{2} + \frac{B^3}{3} - \cdots + (-1)^d \frac{B^{d-1}}{d-1}. \tag{40}$$

This is just the truncated series for $\ln(I + B)$ (notice $B^d = 0$). Hence in the case of these special triangular matrices, exponential and logarithm are inverse operations.

In general, A can be put in triangular form $A = UTU^{-1}$ where T is upper triangular. Let $\lambda_1, \ldots, \lambda_d$ be the diagonal elements of T, that is the eigenvalues of A. Then

$$\exp A = U \cdot \exp T \cdot U^{-1} \tag{41}$$

where $\exp T$ is triangular, with the diagonal elements $\exp \lambda_1, \ldots, \exp \lambda_d$. Hence

$$\det(\exp A) = \prod_{i=1}^{d} \exp \lambda_i = \exp \sum_{i=1}^{d} \lambda_i = \exp(\mathrm{Tr}(A)). \tag{42}$$

The determinant of $\exp A$ is therefore nonzero. Conversely **any complex matrix M with a nonzero determinant is an exponential**: for the proof, write M in the form UTU^{-1} where T is composed of Jordan blocks of the form

$$T_s = \begin{pmatrix} \lambda & 1 & . & . & 0 \\ & . & . & . & . \\ & 0 & & . & 1 \\ . & . & . & . & \lambda \end{pmatrix} \quad \text{with } \lambda \neq 0 \quad .$$

From the existence of the complex logarithm of λ and the study above of triangular matrices, it follows that T_s is an exponential, hence T and $M = UTU^{-1}$ are exponentials.

Let us add a few remarks:

a) A complex matrix with nonzero determinant has infinitely many logarithms; it is possible to normalize things to select one of them, but the conditions are rather artificial.

b) A real matrix with nonzero determinant is not always the exponential of a real matrix; for example, choose $M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. This is not surprising since -1 has no real logarithm, but many complex logarithms of the form $k\pi i$ with k odd.

c) The noncommutativity of the multiplication of matrices implies that in general $\exp(A+B)$ is not equal to $\exp A \cdot \exp B$. Here the logarithm of a product cannot be the sum of the logarithms, whatever normalization we choose.

2.5 Infinitesimals and exponentials

There are many notations in use for the higher order derivatives of a function f. Newton uses $\dot{f}, \ddot{f}, \ldots$, the customary notation is $f', f'', \ldots$. Once again, the exponential notation can be systematized, $f^{(m)}$ or $\mathbf{D}^m f$ denoting the m-th derivative of f, for $m = 0, 1, \ldots$. This notation emphasizes that the derivation is a functional operator, hence

$$(f^{(m)})^{(n)} = f^{(m+n)}, \quad \text{or} \quad \mathbf{D}^m(\mathbf{D}^n f) = \mathbf{D}^{m+n} f. \tag{43}$$

In this notation, it is cumbersome to write the chain rule for the derivative of a composite function

$$\mathbf{D}(f \circ g) = (\mathbf{D}f \circ g) \cdot \mathbf{D}g. \tag{44}$$

Leibniz's notation for the derivative is dy/dx if $y = f(x)$. Leibniz was never able to give a completely rigorous definition of the infinitesimals $dx, dy, \ldots$[1]. His explanation of the derivative is as follows: starting from x, increment it by an infinitely small amount dx; then $y = f(x)$ is incremented by dy, see Figure 1.

$$f(x+dx) = y + dy. \tag{45}$$

Then the derivative is $f'(x) = dy/dx$, hence according to (45),

$$f(x+dx) = f(x) + f'(x)dx. \tag{46}$$

This cannot be literally true, otherwise the function $f(x)$ would be linear. The true formula is

$$f(x+dx) = f(x) + f'(x)dx + o(dx) \tag{47}$$

[1] In modern times, Abraham Robinson has vindicated them using the tools of formal logic. There has been many interesting applications of his nonstandard analysis, but one has to admit that it remains too cumbersome to provide a viable alternative to the standard analysis. Maybe in the 21st century!

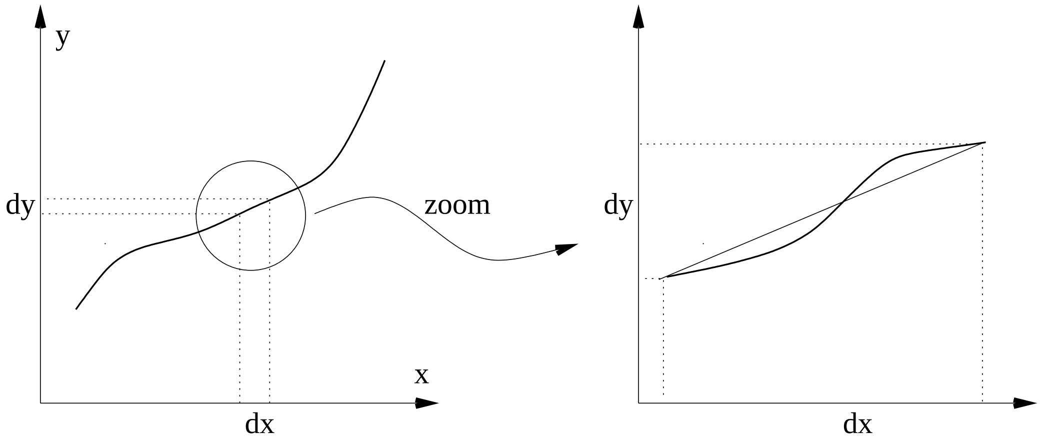

Fig. 1. Geometrical description: an infinitely small portion of the curve $y = f(x)$, after zooming, becomes infinitely close to a straight line, our function is "smooth", not fractal-like.

with an error term $o(dx)$ which is infinitesimal, of a higher order than dx, meaning $o(dx)/dx$ is again infinitesimal. In other words, the derivative $f'(x)$, independent of dx, is infinitely close to $\frac{f(x+dx)-f(x)}{dx}$ for all infinitesimals dx. The modern definition, as well as Newton's point of view of fluents, is a **dynamical** one: when dx goes to 0, $\frac{f(x+dx)-f(x)}{dx}$ tends to the limit $f'(x)$. Leibniz's notion is **statical**: dx is a given, fixed quantity. But there is a hierarchy of infinitesimals: η is of higher order than ϵ if η/ϵ is again infinitesimal. In the formulas, equality is always to be interpreted up to an infinitesimal error of a certain order, not always made explicit.

We use these notions to describe the logarithm and the exponential. By definition, the derivative of $\ln x$ is $\frac{1}{x}$, hence

$$\frac{d\ln x}{dx} = \frac{1}{x}, \quad \text{that is } \ln(x+dx) = \ln(x) + \frac{dx}{x}.$$

Similarly for the exponential

$$\frac{d\exp x}{dx} = \exp x, \quad \text{that is } \exp(x+dx) = (\exp x)(1+dx).$$

This is a rule of compound interest. Imagine a fluctuating daily rate of interest, namely $\epsilon_1, \epsilon_2, \ldots, \epsilon_{365}$ for the days of a given year, every daily rate being of the order of 0.0003. For a fixed investment C, the daily reward is $C\epsilon_i$ for day i, hence the capital becomes $C + C\epsilon_1 + \ldots + C\epsilon_{365} = C \cdot (1 + \sum_i \epsilon_i)$, that is approximately $C(1+.11)$. If we reinvest every day our profit, invested capital changes according to the rule:

$$\begin{array}{ccccc} C_{i+1} = & C_i & + & C_i\epsilon_i = & C_i(1+\epsilon_i). \\ \uparrow & \uparrow & & \uparrow & \\ \text{capital} & \text{capital} & & \text{profit} & \\ \text{at day } i+1 & \text{at day } i & & \text{during day } i & \end{array}$$

At the end of the year, our capital is $C \cdot \prod_i (1+\epsilon_i)$. We can now formulate the "bankers rule":

$$\text{if } S = \epsilon_1 + \ldots + \epsilon_N, \text{ then } \exp S = (1+\epsilon_1)\cdots(1+\epsilon_N). \qquad (B)$$

Here N is infinitely large, and $\epsilon_1, \ldots, \epsilon_N$ are infinitely small; in our example, $S = 0.11$, hence $\exp\ S = 1 + S + \frac{1}{2}S^2 + \ldots$ is equal to $1.1163\ldots$: **by reinvesting daily, the yearly profit of** 11% **is increased to** 11.63%.

Formula (B) is not true without reservation. It certainly holds if all ϵ_i are of the same sign, or more generally if $\sum_i |\epsilon_i|$ is of the same order as $\sum \epsilon_i = x$. For a counter-example, take $N = 2p^2$ with half of the ϵ_i being equal to $+\frac{1}{p}$, and the other half to $-\frac{1}{p}$ (hence $\sum_i \epsilon_i = 0$ while $\prod_i(1+\epsilon_i)$ is infinitely close to $1/e = \exp(-1)$).

To connect definition (B) of the exponential to the power series expansion $\exp S = 1 + S + \frac{1}{2!}S^2 + \cdots$ one can proceed as follows: by algebra we get

$$\prod_{i=1}^{N}(1+\epsilon_i) = \sum_{k=0}^{N} S_k, \tag{48}$$

where $S_0 = 1, S_1 = \epsilon_1 + \ldots + \epsilon_N = S$, and generally

$$S_k = \sum_{i_1 < \ldots < i_k} \epsilon_{i_1} \cdots \epsilon_{i_k}. \tag{49}$$

We have to compare S_k to $\frac{1}{k!}S^k = \frac{1}{k!}(\epsilon_1 + \cdots + \epsilon_N)^k$. Developing the k-th power of S by the multinomial formula, we obtain S_k plus error terms each containing at least one of the ϵ_i's to a higher power, $\epsilon_i^2, \epsilon_i^3, \ldots$, hence infinitesimal compared to the ϵ_i's. The general **principle of compensation of errors**[2] is as follows: given a sum of infinitesimals

$$\Sigma = \eta_1 + \cdots + \eta_M \tag{50}$$

[2] This terminology was coined by Lazare Carnot in 1797. Our formulation is more precise than his!

and new summands $\eta'_j = \eta_j + o(\eta_j)$ with an error $o(\eta_j)$ of higher order than η_j, we obtain that

$$\Sigma' = \eta'_1 + \cdots + \eta'_M \tag{51}$$

is equal to Σ plus an error term $o(\eta_1) + \cdots + o(\eta_M)$. **If the η_j are of the same sign, the error is $o(\Sigma)$, that is negligible compared to Σ.**

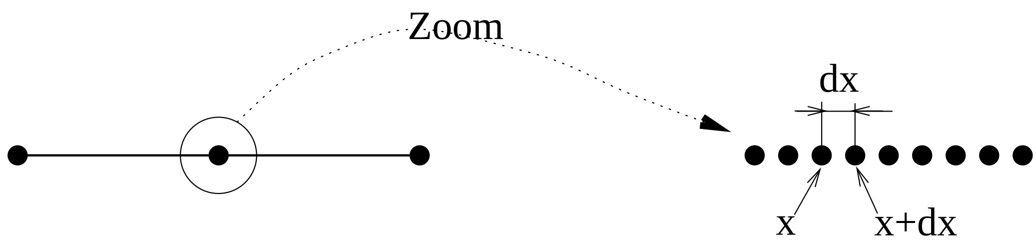

Fig. 2. Leibniz' continuum: by zooming, a finite segment of line is made of a large number of atoms of space: a fractal.

The implicit view of the continuum underlying Leibniz's calculus is as follows: a finite segment of a line is made of an infinitely large number of geometric atoms of space which can be arranged in a succession, each atom x being separated by dx from the next one. Hence in the definition of the logarithm

$$\ln a = \int_1^a \frac{dx}{x} \quad \text{(for } a > 1\text{)}, \tag{52}$$

we really have $\sum_{1 \le x \le a} \frac{dx}{x}$. Similarly, the bankers rule (B) should be interpreted as

$$\exp a = \prod_{0 \le x \le a} (1 + dx) \quad \text{(for } a > 0\text{)}. \tag{53}$$

2.6 Differential equations

The previous formulation of the exponential suggests a method to solve a differential equation, for instance $y' = ry$. In differential form

$$dy = r(x) y dx, \tag{54}$$

that is

$$y + dy = (1 + r(x)dx)y. \tag{55}$$

The solution is

$$y(b) = \prod_{a \leq x \leq b} (1 + r(x)dx) \cdot y(a). \tag{56}$$

What is the meaning of this product? Putting $\epsilon(x) = r(x)dx$, an infinitesimal, and expanding the product as in (48), we get

$$\prod_x (1 + \epsilon(x)) = \sum_{k \geq 0} \sum_{a \leq x_1 < \ldots < x_k \leq b} \epsilon(x_1) \cdots \epsilon(x_k); \tag{57}$$

reinterpreting the multiple sum as a multiple integral, this is

$$\sum_{k \geq 0} \int \cdots \int_{\Delta_k} r(x_1) \cdots r(x_k) dx_1 \cdots dx_k. \tag{58}$$

The domain of integration Δ_k is given by the inequalities

$$a \leq x_1 \leq x_2 \leq \ldots \leq x_k \leq b. \tag{59}$$

The classical solution to the differential equation $y' = ry$ is given by

$$y(b) = (\exp \int_a^b r(x)dx) \cdot y(a). \tag{60}$$

Let us see how to go from (58) to (60). Geometrically, consider the hypercube C_k given by

$$a \leq x_1 \leq b, \cdots, a \leq x_k \leq b \tag{61}$$

in the euclidean space $\mathbf{R}^k$ of dimension k with coordinates $x_1, \ldots, x_k$. The group S_k of the permutations σ of $\{1, \ldots, k\}$ acts on $\mathbf{R}^k$, by transforming the vector $\mathbf{x}$ with coordinates $x_1, \ldots, x_k$ into the vector $\sigma.\mathbf{x}$ with coordinates $x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(k)}$. Then the cube C_k is the union of the $k!$ transforms $\sigma(\Delta_k)$. Since the function $r(x_1) \ldots r(x_k)$ to be integrated is symmetrical in the variables $x_1, \ldots, x_k$ and moreover two distinct domains $\sigma(\Delta_k)$ and $\sigma'(\Delta_k)$ overlap by a subset of dimension $< k$, hence of volume 0, we see that the integral of $r(x_1) \cdots r(x_k)$ over C_k is $k!$ times the integral over Δ_k. That is

$$\int \cdots \int_{\Delta_k} r(x_1) \cdots r(x_k) dx_1 \cdots dx_k =$$
$$\frac{1}{k!} \int_a^b dx_1 \cdots \int_a^b dx_k \; r(x_1) \cdots r(x_k) = \frac{1}{k!} (\int_a^b r(x)dx)^k.$$

Summing over k, and using the definition of an exponential by a series, we conclude

$$\sum_{k\geq 0}\int\cdots\int_{\Delta_k} r(x_1)\cdots r(x_k)dx_1\cdots dx_k = \exp\int_a^b r(x)dx. \tag{62}$$

as promised.

The same method applies to the linear systems of differential equations. We cast them in the matrix form

$$y' = A\cdot y, \tag{63}$$

that is the differential form

$$dy = A(x)ydx. \tag{64}$$

Here $A(x)$ is a matrix depending on the variable x, and $y(x)$ is a vector (or matrix) function of x. From (64) we get

$$y(x+dx) = (I + A(x)dx)y(x). \tag{65}$$

Formally the solution is given by

$$y(b) = \prod_{a\leq x\leq b}(I + A(x)dx)\cdot y(a). \tag{66}$$

We have to take into account the **noncommutativity** of the products $A(x)A(y)A(z)\ldots$. Explicitly, if we have chosen intermediate points

$$a = x_0 < x_1 < \ldots < x_N = b,$$

with infinitely small spacing

$$dx_1 = x_1 - x_0, dx_2 = x_2 - x_1, \ldots, dx_N = x_N - x_{N-1},$$

the product in (66) is

$$(I + A(x_N)dx_N)(I + A(x_{N-1})dx_{N-1})\cdots(I + A(x_1)dx_1).$$

We use the notation $\overleftarrow{\prod}_{1\leq i\leq N} U_i$ for a **reverse product** $U_N U_{N-1}\cdots U_1$; hence the previous product can be written as $\overleftarrow{\prod}_{1\leq i\leq N}(I + A(x_i)dx_i)$ and we should replace $\prod$ by $\overleftarrow{\prod}$ in equation (66). The noncommutative version of equation (48) is

$$\overleftarrow{\prod}_{1\leq i\leq N}(I+A_i)=\sum_{k=0}^{N}\sum_{i_1>\cdots>i_k}A_{i_1}\cdots A_{i_k}. \tag{67}$$

Let us define the resolvant (or propagator) as the matrix

$$U(b,a)=\overleftarrow{\prod}_{a\leq x\leq b}(I+A(x)dx). \tag{68}$$

Hence the differential equation $dy=A(x)ydx$ is solved by $y(b)=U(b,a)y(a)$ and from (67) we get

$$U(b,a)=\sum_{k\geq 0}\int\cdots\int_{\Delta_k}A(x_k)\cdots A(x_1)dx_1\cdots dx_k \tag{69}$$

with the factors $A(x_i)$ in **reverse order**

$$A(x_k)\cdots A(x_1)\quad\text{for } x_1<\ldots<x_k. \tag{70}$$

One owes to R. Feynman and F. Dyson (1949) the following notational trick. If we have a product of factors $U_1,\cdots,U_N$, each attached to a point x_i on a line, we denote by $T(U_1\cdots U_N)$ (or more precisely by $\overleftarrow{T}(U_1\cdots U_N)$) the product $U_{i_1}\cdots U_{i_N}$ where the permutation $i_1\ldots i_N$ of $1\ldots N$ is such that $x_{i_1}>\cdots>x_{i_N}$. Hence in the rearranged product the abscisses attached to the factors increase from right to left. We argue now as in the proof of (62) and conclude that

$$\begin{aligned}&\int\cdots\int_{\Delta_k}A(x_k)\cdots A(x_1)dx_1\cdots dx_k\\ &=\frac{1}{k!}\int_a^b dx_1\cdots\int_a^b dx_k T(A(x_1)\cdots A(x_k)).\end{aligned} \tag{71}$$

We can rewrite the propagator as

$$U(b,a)=T\exp\int_a^b A(x)dx, \tag{72}$$

with the following interpretation:

a) First use the series $\exp S=\sum_{k\geq 0}\frac{1}{k!}S^k$ to expand $\exp\int_a^b A(x)dx$.

b) Expand $S^k = (\int_a^b A(x)dx)^k$ as a multiple integral

$$\int_a^b dx_1 \cdots \int_a^b dx_k\ A(x_1)\cdots A(x_k).$$

c) Treat T as a linear operator commuting with series and integrals, hence

$$\begin{aligned} T\exp S &= \sum_{k\geq 0}\frac{1}{k!}T(S^k) = \sum_{k\geq 0}\frac{1}{k!}T\{\int_a^b dx_1\cdots\int_a^b dx_k\ A(x_1)\cdots A(x_k)\} \\ &= \sum_{k\geq 0}\frac{1}{k!}\int_a^b dx_1\cdots\int_a^b dx_k\ T(A(x_1)\cdots A(x_k)). \end{aligned}$$

We give a few properties of the T (or time ordered) exponential:

a) Parallel to the rule

$$\int_a^c A(x)dx = \int_a^b A(x)dx + \int_b^c A(x)dx \quad (\text{for } a < b < c) \tag{73}$$

we get

$$T\exp\int_a^c A(x)dx = T\exp\int_b^c A(x)dx\cdot T\exp\int_a^b A(x)dx. \tag{74}$$

Notice that, in (73), the two matrices

$$L = \int_a^b A(x)dx, \quad M = \int_b^c A(x)dx$$

do not commute, hence $\exp(L+M)$ is in general different from $\exp L\cdot\exp M$. Hence formula (74) is not in general valid for the ordinary exponential.

b) The next formula embodies the classical method of "variation of constants" and is known in the modern literature as a "gauge transformation". It reads as

$$S(b)\cdot T\exp\int_a^b A(x)dx\cdot S(a)^{-1} = T\exp\int_a^b B(x)dx \tag{75}$$

with

$$B(x) = S(x)A(x)S(x)^{-1} + S'(x)S(x)^{-1}, \tag{76}$$

where $S(x)$ is an invertible matrix depending on the variable x. The general formula (75) can be obtained by "taking a continuous reverse product" $\overleftarrow{\prod}_{a\leq x\leq b}$ over the infinitesimal form

$$S(x+dx)(I + A(x)dx))S(x)^{-1} = I + B(x)dx \tag{77}$$

(for the proof, write $S(x+dx) = S(x) + S'(x)dx$ and neglect the terms proportional to $(dx)^2$). We leave it as an exercise to the reader to prove (75) from the expansion (69) for the propagator.

c) There exists a complicated formula for the T-exponential $T\exp\int_a^b A(x)\,dx$ when $A(x)$ is of the form $\frac{A_1(x)+A_2(x)}{2}$. Neglecting terms of order $(dx)^2$, we get

$$I + A(x)dx = \left(I + A_2(x)\frac{dx}{2}\right)\left(I + A_1(x)\frac{dx}{2}\right) \tag{78}$$

and we can then perform the product $\overleftarrow{\prod}_{a \leq x \leq b}$. This formula is the foundation of the **multistep method** in numerical analysis: starting from the value $y(x)$ at time x of the solution to the equation $y' = Ay$, we split the infinitesimal interval $[x, x+dx]$ into two parts

$$I_1 = [x, x + \frac{dx}{2}], \quad I_2 = [x + \frac{dx}{2}, x + dx];$$

we move at speed $A_1(x)y(x)$ during I_1 and then at speed $A_2(x)y(x+\frac{dx}{2})$ during I_2. Let us just mention one corollary of this method, the so-called **Trotter-Kato-Nelson** formula:

$$\exp(L+M) = \lim_{n\to\infty}(\exp(L/n)\exp(M/n))^n. \tag{79}$$

d) If the matrices $A(x)$ pairwise commute, the T-exponential of $\int_a^b A(x)dx$ is equal to the ordinary exponential. In the general case, the following formula holds

$$T\exp\int_a^b A(x)dx = \exp V(b,a) \tag{80}$$

where $V(b,a)$ is explicitly calculated using integration and iterated Lie brackets. Here are the first terms

$$\begin{aligned} V(b,a) &= \int_a^b A(x)dx + \frac{1}{2}\int\int_{\Delta_2}[A(x_2), A(x_1)]dx_1dx_2 \\ &+\frac{1}{3}\int\int\int_{\Delta_3}[A(x_3),[A(x_2),A(x_1)]]dx_1dx_2dx_3 \\ &-\frac{1}{6}\int\int\int_{\Delta_3}[A(x_2),[A(x_3),A(x_1)]]dx_1dx_2dx_3 + \cdots \quad . \end{aligned} \tag{81}$$

The higher-order terms involve integrals of order $k \geq 4$. As far as I can ascertain, this formula was first enunciated by K. Friedrichs around 1950 in his work on the foundations of Quantum Field Theory. A corollary is the **Campbell-Hausdorff formula**:

$$\exp L \cdot \exp M = \exp(L + M + \frac{1}{2}[L,M] + \frac{1}{12}[L,[L,M]] + \frac{1}{12}[M,[M,L]] + \cdots). \quad (82)$$

It can be derived from (80) by putting $a = 0$, $b = 2$, $A(x) = M$ for $0 \leq x \leq 1$ and $A(x) = L$ for $1 \leq x \leq 2$.

The T-exponential found lately numerous geometrical applications. If C is a curve in a space of arbitrary dimension, the line integral $\int_C A_\mu(x) dx^\mu$ is well-defined and the corresponding T-exponential

$$T\exp \int_C A_\mu(x) dx^\mu \quad (83)$$

is closely related to the parallel transport along the curve C.

3 Operational calculus

3.1 An algebraic digression: umbral calculus

We first consider the classical **Bernoulli numbers**. I claim that they are defined by the equation

$$(B+1)^n = B^n \quad \text{for } n \geq 2, \quad (1)$$

together with the initial condition $B^0 = 1$. The meaning is the following: expand $(B+1)^n$ by the binomial theorem, then replace the power B^k by B_k. Hence $(B+1)^2 = B^2$ gives $B^2 + 2B^1 + B^0 = B^2$, that is after lowering the indices $B_2 + 2B_1 + B_0 = B_2$, that is $2B_1 + B_0 = 0$. Treating $(B+1)^3 = B^3$ in a similar fashion gives $3B_2 + 3B_1 + B_0 = 0$. We write the first equations of this kind

$$\begin{array}{ll} n = 2 & 2B_1 + B_0 = 0 \\ n = 3 & 3B_2 + 3B_1 + B_0 = 0 \\ n = 4 & 4B_3 + 6B_2 + 4B_1 + B_0 = 0 \\ n = 5 & 5B_4 + 10B_3 + 10B_2 + 5B_1 + B_0 = 0. \end{array}$$

Starting from $B_0 = 1$ we get successively

$$B_1 = -\frac{1}{2},\ B_2 = \frac{1}{6},\ B_3 = 0,\ B_4 = -\frac{1}{30}, \dots$$

Using the same kind of formalism, define the Bernoulli polynomials by

$$B_n(X) = (B + X)^n. \tag{2}$$

According to the previous rule, we first expand $(B+X)^n$ using the binomial theorem, then replace B^k by B_k. Hence we get explicitly

$$B_n(X) = \sum_{k=0}^{n} \binom{n}{k} B_{n-k} X^k. \tag{3}$$

Since $\frac{d}{dX}(X + c)^n = n(X + c)^{n-1}$ for any c independent of X, we expect

$$\frac{d}{dX} B_n(X) = n B_{n-1}(X). \tag{4}$$

This is easy to check on the explicit definition (3). Here is a similar calculation

$$(B + (X + Y))^n = ((B + X) + Y)^n = \sum_{k=0}^{n} \binom{n}{k} (B + X)^{n-k} Y^k,$$

from which we expect to find

$$B_n(X + Y) = \sum_{k=0}^{n} \binom{n}{k} B_{n-k}(X) Y^k. \tag{5}$$

Indeed from (4) we get

$$\left(\frac{d}{dX}\right)^k B_n(X) = \frac{n!}{(n-k)!} B_{n-k}(X) \tag{6}$$

by induction on k, hence (5) follows from Taylor's formula $B_n(X + Y) = \sum_{k \geq 0} \frac{1}{k!} (\frac{d}{dX})^k B_n(X) Y^k$.

We deduce now a generating series for the Bernoulli numbers. Formally

$$\begin{aligned}(e^S - 1)e^{BS} &= e^S e^{BS} - e^{BS} = e^{(B+1)S} - e^{BS} \\ &= \sum_{n \geq 0} \frac{1}{n!} S^n ((B + 1)^n - B^n) = S((B + 1)^1 - B^1) = S.\end{aligned}$$

Since $e^{BS} = \sum_{n \geq 0} \frac{1}{n!} B^n S^n$, we expect

$$\sum_{n \geq 0} B_n S^n / n! = \frac{S}{e^S - 1}. \tag{7}$$

Again this can be checked rigorously.

What is the secret behind these calculations?

We consider functions $F(B, X, \ldots)$ depending on a variable B and other variables $X, \ldots$. Assume that $F(B, X, \ldots)$ can be expanded as a polynomial or power series in B, namely

$$F(B, X, \ldots) = \sum_{n \geq 0} B^n F_n(X, \ldots). \tag{8}$$

Then the "mean value" with respect to B is defined by

$$< F(B, X, \ldots) > = \sum_{n \geq 0} B_n F_n(X, \ldots), \tag{9}$$

where the B_n's are the Bernoulli numbers: this corresponds to the rule "lower the index in B^n". If the function $F(B, X, \ldots)$ can be expanded into a series $\sum_i F_i(B, X, \ldots) G_i(X, \ldots)$ where the G_i's are independent of B, then obviously[3]

$$< F(B, X, \ldots) > = \sum_i < F_i(B, X, \ldots) > G_i(X, \ldots). \tag{10}$$

The formal calculations given above are justified by this simple rule **which affords also a probabilistic interpretation** (see Section 3.7).

The previous method is loosely described as "umbral calculus". We insisted on speaking of "mean values" to keep in touch with physical applications. From a purely mathematical point of view, it is just applying a linear functional acting on polynomials in B, mapping B^n into B_n for all n's.

3.2 Binomial sequences of polynomials

These are sequences of polynomials $U_0(X), U_1(X), \ldots$ in one variable X satisfying the following relations:

[3] So far we considered only identities linear in the B_n's. If we want to treat nonlinear terms, like products $B_m \cdot B_n$, we need to introduce two independent symbols B and B' and use the umbral rule to replace $B^m B'^n$ by $B_m B_n$. In probabilistic terms (see Section 3.7), we introduce two independent random variables and take the mean value with respect to both simultaneously.

a) $U_0(X)$ is a constant;
b) for any $n \geq 1$, one gets

$$\frac{d}{dX}U_n(X) = nU_{n-1}(X). \tag{11}$$

By induction on n it follows that $U_n(X)$ is of degree $\leq n$. The binomial sequence is **normalized** if furthermore $U_0(X) = 1$, in which case every $U_n(X)$ is a monic polynomial of degree n, that is

$$U_n(X) = X^n + c_1 X^{n-1} + \ldots + c_n.$$

Applying Taylor's formula as above (derivation of formula (5)), one gets

$$U_n(X+Y) = \sum_{k=0}^{n} \binom{n}{k} U_{n-k}(X) Y^k. \tag{12}$$

We introduce now a numerical sequence by $u_n = U_n(0)$ for $n \geq 0$. Putting $X = 0$ in (12) and then replacing Y by X (as a variable), we get

$$U_n(X) = \sum_{k=0}^{n} \binom{n}{k} u_{n-k} X^k. \tag{13}$$

Conversely, given any numerical sequence $u_0, u_1, \ldots$, and defining the polynomials $U_n(X)$ by (13), one derives immediately the relations

$$\frac{d}{dX}U_n(X) = nU_{n-1}(X), \quad U_n(0) = u_n. \tag{14}$$

The exponential generating series for the constants u_n is given by

$$u(S) = \sum_{n \geq 0} u_n S^n / n!. \tag{15}$$

From (13), one obtains the exponential generating series

$$U(X,S) = \sum_{n \geq 0} U_n(X) S^n / n!$$

for the polynomials $U_n(X)$, namely in the form

$$U(X,S) = u(S) e^{XS}. \tag{16}$$

This could be expected. Writing $\partial_X,\ \partial_S \dots$ for the partial derivatives, the basic relation $\partial_X U_n = nU_{n-1}$ translates as $(\partial_X - S)U(X,S) = 0$ or equivalently as

$$\partial_X(e^{-XS}U(X,S)) = 0. \tag{17}$$

Hence $e^{-XS}U(X,S)$ depends only on S, and putting $X = 0$ we obtain the value $U(0,S) = u(S)$.

The umbral calculus can be successfully applied to our case. Hence $U_n(X)$ can be interpreted as $\langle (X+U)^n\rangle$ provided $\langle U^n\rangle = u_n$. Similarly $u(S)$ is equal to $\langle e^{US}\rangle$ and $U(X,S)$ to $\langle e^{(X+U)S}\rangle$. The symbolic derivation of (16) is as follows

$$U(X,S) = \langle e^{(X+U)S}\rangle = \langle e^{XS}\cdot e^{US}\rangle = e^{XS}\langle e^{US}\rangle = e^{XS}u(S).$$

We describe in more detail the three basic binomial sequences of polynomials:

a) The sequence $I_n(X) = X^n$ satisfies obviously (11). In this (rather trivial) case, we get

$$i_0 = 1,\ i_1 = i_2 = \ldots = 0,\ I(S) = 1,\ I(X,S) = e^{XS}.$$

b) **The Bernoulli polynomials** obey the rule (11) (see formula (4)). I claim that they are characterized by the normalization $B_0(X) = 1$ and the further property

$$\int_0^1 B_n(x)dx = 0 \quad \text{for } n \geq 1. \tag{18}$$

Indeed, introducing the exponential generating series

$$B(X,S) = \sum_{n\geq 0} B_n(X)S^n/n!, \tag{19}$$

the requirement (18) is equivalent to the integral formula

$$\int_0^1 B(x,S)dx = 1. \tag{20}$$

According to the general theory of binomial sequences, $B(X,S)$ is of the form $b(S)e^{XS}$, hence

$$\int_0^1 B(x,S)dx = \int_0^1 b(S)e^{xS}dx = b(S)\left(\frac{e^S-1}{S}\right).$$

Solving (20) we get $b(S) = S/(e^S - 1)$ and from (7) this is the exponential generating series for the Bernoulli numbers. The exponential generating series for the Bernoulli polynomials is therefore

$$B(X, S) = \frac{Se^{XS}}{e^S - 1}. \tag{21}$$

Here is a short table:

$$\begin{aligned}
B_0(X) &= 1 \\
B_1(X) &= X - \frac{1}{2} \\
B_2(X) &= X^2 - X + \frac{1}{6} \\
B_3(X) &= X^3 - \frac{3}{2}X^2 + \frac{1}{2}X.
\end{aligned}$$

c) We come to the **Hermite polynomials** which form the normalized binomial sequence of polynomials characterized by

$$\int_{-\infty}^{+\infty} H_n(x) d\gamma(x) = 0 \quad \text{for } n \geq 1, \tag{22}$$

where $d\gamma(x)$ denotes the normal probability law, that is

$$d\gamma(x) = (2\pi)^{-1/2} e^{-x^2/2} dx. \tag{23}$$

We follow the same procedure as for the Bernoulli polynomials. Hence for the exponential generating series

$$H(X, S) = \sum_{n \geq 0} H_n(X) S^n / n! = h(S) e^{XS} \tag{24}$$

we get

$$\int_{-\infty}^{+\infty} H(x, S) d\gamma(x) = 1, \tag{25}$$

that is

$$1/h(S) = \int_{-\infty}^{+\infty} e^{xS} d\gamma(x). \tag{26}$$

The last integral being easily evaluated, we conclude

$$h(S) = e^{-S^2/2}. \tag{27}$$

From this relation, we can evaluate $H(X,S)$ namely

$$H(X,S) = e^{XS-S^2/2} = e^{X^2/2}e^{-(X-S)^2/2} \tag{28}$$

and using Taylor's expansion for $e^{-(X-S)^2/2}$, we get

$$H_n(X) = (-1)^n e^{X^2/2}\left(\frac{d}{dX}\right)^n e^{-X^2/2}. \tag{29}$$

In the spirit of operator calculus, use the identity

$$e^{X^2/2}\frac{d}{dX}e^{-X^2/2} = \frac{d}{dX} - X, \tag{30}$$

whence

$$H_n(X) = \left(X - \frac{d}{dX}\right)^n .1. \tag{31}$$

This is tantamount to a recursion formula

$$H_{n+1}(X) = XH_n(X) - \frac{d}{dX}H_n(X) = XH_n(X) - nH_{n-1}(X). \tag{32}$$

The following table is then easily derived:

$$\begin{aligned}
&H_0(X) = 1\\
&H_1(X) = X\\
&H_2(X) = X^2 - 1\\
&H_3(X) = X^3 - 3X\\
&H_4(X) = X^4 - 6X^2 + 3.
\end{aligned}$$

3.3 Transformation of polynomials

We use the standard notation $\mathbf{C}[X]$ to denote the vector space of polynomials in the variable X with complex coefficients. Since the monomials X^n form a basis of $\mathbf{C}[X]$, a linear operator $\mathbf{U}:\ \mathbf{C}[X] \to \mathbf{C}[X]$ is completely determined by the sequence of polynomials $U_n(X)$ defined as the image $\mathbf{U}[X^n]$ of X^n under $\mathbf{U}$. Here are a few examples:

$\mathbf{I}$ identity operator	$I_n(X) = X^n$
$\mathbf{D}$ derivation $\dfrac{d}{dX}$	$D_n(X) = nX^{n-1}$
$\mathbf{T}_c$ translation operator	$T_{c,n}(X) = (X+c)^n$.

Notice that in general $\mathbf{T}_c$ transforms a polynomial $P(X)$ into $P(X+c)$ and Taylor's formula amounts to

$$\mathbf{T}_c = e^{c\mathbf{D}}; \tag{33}$$

furthermore $\mathbf{T}_0 = \mathbf{I}$. From the definition of the derivative, one gets

$$\mathbf{D} = \lim_{c\to 0}(\mathbf{T}_c - \mathbf{I})/c. \tag{34}$$

We can reformulate the properties of binomial sequences:
- the definition $\mathbf{D}U_n(X) = nU_{n-1}(X)$ amounts to $\mathbf{DU} = \mathbf{UD}$;
- the exponential generating series $U(X,S)$ is nothing else than $\mathbf{U}[e^{XS}]$;
- formula (12), after substituting c to Y reads as

$$U_n(X+c) = \sum_{k=0}^{n} \binom{n}{k} U_{n-k}(X)c^k$$

that is

$$\mathbf{T}_c\mathbf{U}[X^n] = \sum_{k=0}^{n} \binom{n}{k} \mathbf{U}[X^{n-k}]c^k = \mathbf{U}[(X+c)^n] = \mathbf{U}\mathbf{T}_c[X^n].$$

Hence this formula expresses that $\mathbf{U}$ commutes to $\mathbf{T}_c$

$$\mathbf{T}_c\mathbf{U} = \mathbf{U}\mathbf{T}_c; \tag{35}$$

- formula (13) can be rewritten as

$$\mathbf{U}[X^n] = \sum_{k\geq 0} \frac{1}{k!} u_k \mathbf{D}^k[X^n]. \tag{36}$$

From the definition (15) of the exponential generating series, we obtain

$$\mathbf{U} = u(\mathbf{D}). \tag{37}$$

To sum up, our operators are characterized by the following equivalent properties:
a) $\mathbf{U}$ commutes with the derivative $\mathbf{D}$;
b) $\mathbf{U}$ commutes with the translation operators $\mathbf{T}_c$;
c) $\mathbf{U}$ can be expressed as a power series $u(\mathbf{D})$ in $\mathbf{D}$.
Furthermore, since $\mathbf{D}$ acts on e^{XS} by multiplication by S, the operator $\mathbf{U} = u(\mathbf{D})$ multiplies e^{XS} by $u(S)$. Hence we recover formula (16).

3.4 Expansion formulas

As we saw before, $B_n(X)$ and $H_n(X)$ are monic polynomials and therefore the sequences $(B_n(X))_{n\geq 0}$ and $(H_n(X))_{n\geq 0}$ are two bases of the vector space $\mathbf{C}[X]$. Hence an arbitrary polynomial $P(X)$ can be expanded as a linear combination of the Bernoulli polynomials, as well as of the Hermite polynomials. Our aim is to give explicit formulas.

Consider a general binomial sequence $(U_n(X))_{n\geq 0}$ such that $u_0 \neq 0$, with exponential generating series $U(X,S) = u(S)e^{XS}$. Introduce the inverse series $v(S) = 1/u(S)$; explicitly

$$v(S) = \sum_{n\geq 0} v_n S^n/n!,$$

and the coefficients v_n are defined inductively by

$$v_0 = 1/u_0, \ v_n = -\frac{1}{u_0}\sum_{k=1}^{n}\binom{n}{k} u_k v_{n-k}. \tag{38}$$

In the spirit of umbral calculus, let us define the linear form ϕ_0 on $\mathbf{C}[X]$ by $\phi_0[X^n] = v_n$. I claim that **the expansion of an arbitrary polynomial in terms of the U_n's is given by**

$$P(X) = \sum_{n\geq 0} \frac{1}{n!}\phi_0[\mathbf{D}^n P] \cdot U_n(X). \tag{39}$$

Before giving a proof, let us examine the three basic examples:

a) If $U_n(X) = X^n$, then $u(S) = 1$, hence $v(S) = 1$. That is $v_0 = 1$ and $v_n = 0$ for $n \geq 1$. The linear form ϕ_0 is given by $\phi_0[P] = P(0)$ and formula (39) reduces to MacLaurin's expansion

$$P(X) = \sum_{n\geq 0} \frac{1}{n!}(\mathbf{D}^n P)(0) \cdot X^n. \tag{40}$$

b) For the Bernoulli polynomials we know that the series $1/b(S)$ is equal to $(e^S - 1)/S$, hence $v_n = \frac{1}{n+1}$. The linear form ϕ_0 is defined by $\phi_0[X^n] = \frac{1}{n+1}$, that is

$$\phi_0[P] = \int_0^1 P(x)dx. \tag{41}$$

Hence

$$P(X) = \sum_{n\geq 0} \frac{1}{n!} \left(\int_0^1 \mathbf{D}^n P(x) dx \right) \cdot B_n(X). \tag{42}$$

c) In the case of Hermite polynomials, we know that $1/h(S)$ is equal to $e^{S^2/2}$, hence

$$v_{2m} = \frac{(2m)!}{m! 2^m}, \quad v_{2m+1} = 0. \tag{43}$$

According to (26), we get $v_n = \int_{-\infty}^{+\infty} x^n d\gamma(x)$, hence

$$\phi_0[P] = \int_{-\infty}^{+\infty} P(x) d\gamma(x) = (2\pi)^{-1/2} \int_{-\infty}^{+\infty} P(x) e^{-x^2/2} dx. \tag{44}$$

In these three cases, the formula for $\phi_0[P]$ takes a similar form, namely

$$\phi_0[P] = \int_a^b P(x) w(x) dx \tag{45}$$

with the following prescriptions:

$a = -\infty, \ b = +\infty, \ \ w(x) = \delta(x)$ in case a),

$a = 0, \ b = 1, \ \ w(x) = 1$ in case b),

$a = -\infty, \ b = +\infty, \ \ w(x) = (2\pi)^{-1/2} e^{-x^2/2}$ in case c).

The normalization $\phi_0[1] = 1$ amounts to $\int_a^b w(x) dx = 1$, that is $w(x)$ is the probability density of a random variable taking values in the interval $[a, b]$ (see Section 3.7).

There is a peculiarity in case c). Namely, according to the general formula (39), an arbitrary polynomial $P(X)$ can be expanded in a series of Hermite polynomials $\sum_{n\geq 0} c_n H_n(X)$ where c_n is equal to $\frac{1}{n!} \int_{-\infty}^{+\infty} \mathbf{D}^n P(x) d\gamma(x)$. Integrating by parts and taking into account the definition (29) of $H_n(X)$, we obtain

$$c_n = \frac{1}{n!} \int_{-\infty}^{+\infty} P(x) H_n(x) d\gamma(x). \tag{46}$$

This amounts to the **orthogonality relation**

$$\int_{-\infty}^{+\infty} H_m(x) H_n(x) d\gamma(x) = \delta_{mn} n! \tag{47}$$

for the Hermite polynomials. There is no such orthogonality relation for the Bernoulli polynomials.

One final word about the proof of (39). By linearity, it suffices to consider the case $P = U_m$, that is to prove the **biorthogonality relation**

$$\phi_0[\mathbf{D}^n U_m] = n!\delta_{mn}. \tag{48}$$

We first calculate $\phi_0[U_m]$. From formula (13), we obtain

$$\phi_0[U_m] = \sum_{k=0}^{m} \binom{m}{k} u_{m-k}\phi_0[X^k] = \sum_{k=0}^{m} \binom{m}{k} u_{m-k}v_k,$$

and from (38), $\phi_0[U_m]$ is 0 for $m \geq 1$. Since $\mathbf{D}^n U_m$ is proportional to U_{m-n} according to the basic formula $\mathbf{D}U_m = mU_{m-1}$, one gets $\phi_0[\mathbf{D}^n U_m] = 0$ for $m \neq n$. Finally $\mathbf{D}^m U_m = m!u_0$, hence $\phi_0[\mathbf{D}^m U_m] = v_0 m! u_0 = m!$.

3.5 Signal transforms

A transmission device transforms a suitable **input** f into an **output** F. Both are evolving in time and are represented by functions of time $f(t)$ and $F(t)$.[4] We assume the device to be **linear** and in a **stationary** regime, that is there is a linear operator $\mathbf{V}$ taking $f(t)$ into $F(t)$ (linearity) and $f(t+\tau)$ into $F(t+\tau)$ for any fixed τ (stationarity).

Here are the main types of response:

Input	*Output*
$\delta(t)$	$I(t)$
e^{pt}	$\Theta(t)e^{pt}$
t^n	$V_n(t)$

In the first case, $\delta(t)$ is a Dirac singular function, that is a **pulse**, and $I(t)$ is the **impulse response**. By stationarity, $\mathbf{V}$ transforms $\delta(t-\tau)$ into $I(t-\tau)$; an arbitrary input f can be represented as a superposition of pulses

$$f(t) = \int_{-\infty}^{+\infty} f(\tau)\delta(t-\tau)d\tau \tag{49}$$

hence by linearity the output

$$F(t) = \int_{-\infty}^{+\infty} f(\tau)I(t-\tau)d\tau = \int_{-\infty}^{+\infty} f(t-\tau)I(\tau)d\tau. \tag{50}$$

[4] For simplicity, we restrict to the case where input and output are scalars and not vectors.

In the **non-anticipating case**, $I(t)$ is zero before the pulse $\delta(t)$ occurs, that is $I(t) = 0$ for $t < 0$. In this case the output is given by

$$F(t) = \int_{-\infty}^{t} f(\tau) I(t-\tau) d\tau. \tag{51}$$

In the case of the exponential input $f(t) = e^{pt}$, the output is equal to $\int_{-\infty}^{+\infty} e^{p\tau} I(t-\tau) d\tau = \int_{-\infty}^{+\infty} e^{p(t-\tau)} I(\tau) d\tau$ according to (50), that is, to $\Theta(p) e^{pt}$ with the **spectral gain**

$$\Theta(p) = \int_{-\infty}^{+\infty} e^{-p\tau} I(\tau) d\tau. \tag{52}$$

We can give an **a priori** argument: $f_p(t) = e^{pt}$ is a solution of the differential equation $\mathbf{D} f_p = p f_p$; since the operator $\mathbf{V}$ is stationary, that is, commutes with the translation operators $\mathbf{T}_c$, it commutes with $\mathbf{D} = \lim_{c \to 0} (\mathbf{T}_c - \mathbf{I})/c$. Hence the output F_p corresponding to the input f_p is a solution of the differential equation $\mathbf{D} F_p = p F_p$, hence is proportional to e^{pt}.

In a similar way, the monomials t^n satisfy the cascade of differential equations

$$\mathbf{D}[t] = 1, \quad \mathbf{D}[t^2] = 2t, \quad \mathbf{D}[t^3] = 3t^2, \ldots$$

Since $\mathbf{V}$ commutes with $\mathbf{D}$ and the constants are the solutions of the differential equation $\mathbf{D}(f) = 0$, it follows that the images $V_n(t) = \mathbf{V}[t^n]$ form a binomial sequence of polynomials. Explicitly

$$V_n(t) = \int_{-\infty}^{+\infty} (t-\tau)^n I(\tau) d\tau = \sum_{k=0}^{n} \binom{n}{k} v_{n-k} t^k$$

with the constants

$$v_n = (-1)^n \int_{-\infty}^{+\infty} I(\tau) \tau^n d\tau = V_n(0). \tag{53}$$

Comparing (52) to (53) we conclude

$$\Theta(p) = \sum_{n \geq 0} v_n p^n / n!. \tag{54}$$

More generally, since e^{pt} is equal to $\sum_{n \geq 0} \frac{1}{n!} p^n t^n$, application of the linear operator $\mathbf{V}$ gives

$$\mathbf{V}[e^{pt}] = \sum_{n \geq 0} \frac{1}{n!} p^n V[t^n], \tag{55}$$

that is

$$\Theta(p)e^{pt} = \sum_{n\geq 0} \frac{1}{n!} p^n V_n(t). \tag{56}$$

Up to the change in notation (p for S, and t for X), the spectral gain $\Theta(p)$ is nothing else than the numerical exponential generating series associated to the binomial sequence $(V_n(t))_{n\geq 0}$.

Comparing with the results obtained in Section 3.3, it is tempting to write the operator $\mathbf{V}$ as $\Theta(\mathbf{D})$. According to (54), $\Theta(\mathbf{D})$ can be interpreted as $\sum_{n\geq 0} v_n \mathbf{D}^n/n!$, but it is known that infinite order differential operators are not so easily dealt with. A better interpretation is obtained via Laplace or Fourier transform. Indeed since $\mathbf{D}$ multiplies e^{pt} by p, any function $F(\mathbf{D})$ ought to multiply e^{pt} by $F(p)$, and the rule $\mathbf{V}[e^{pt}] = \Theta(p)e^{pt}$ is in agreement with the interpretation $\mathbf{V} = \Theta(\mathbf{D})$. If the input can be represented as a **Laplace transform**

$$f(t) = \int e^{pt}\phi(p)dp, \tag{57}$$

then $\mathbf{V} = \Theta(\mathbf{D})$ transforms it into the output

$$F(t) = \int e^{pt}\Theta(p)\phi(p)dp. \tag{58}$$

Similarly, if the input $f(t)$ is given by its **spectral resolution** (or Fourier transform)

$$f(t) = \int_{-\infty}^{+\infty} \hat{f}(\omega)e^{i\omega t}d\omega, \tag{59}$$

then the output is given by

$$F(t) = \int_{-\infty}^{+\infty} \Theta(i\omega)\hat{f}(\omega)e^{i\omega t}d\omega. \tag{60}$$

This is **Heaviside's magic trilogy**:

symbolic	p	
operator	$\mathbf{D}$	
spectral	$i\omega = 2\pi i\nu$	(ν frequency, $\omega = 2\pi\nu$ pulsation)

$$\Theta(p) \longleftrightarrow \Theta(\mathbf{D}) \longleftrightarrow \Theta(2\pi i\nu).$$

Recall that in the Laplace transform, p is a complex variable, while in the Fourier transform ω and ν are real variables (see example in the next section).

3.6 The inverse problem

This is the problem of recovering the input, knowing the output. In operator terms, we have to compute the inverse $\mathbf{U}$ of the operator $\mathbf{V}$ (if it exists!). Since $\mathbf{V}$ is stationary, so is $\mathbf{U}$, and at the level of polynomial inputs and outputs, $\mathbf{U}$ corresponds to a binomial sequence of polynomials $U_n(t)$:

$$\begin{cases} \text{input} \quad U_n(t) \\ \text{output } t^n. \end{cases}$$

Together with the numerical sequence $v_n = V_n(0)$, we have to consider the numerical sequence $u_n = U_n(0)$. Introducing the exponential generating series

$$u(S) = \sum_{n \geq 0} \frac{1}{n!} u_n S^n, \;\; v(S) = \sum_{n \geq 0} \frac{1}{n!} v_n S^n = \Theta(S), \tag{61}$$

we can write $\mathbf{U} = u(\mathbf{D})$ and $\mathbf{V} = v(\mathbf{D})$ at least when acting on polynomials. Since $\mathbf{U}$ and $\mathbf{V}$ are inverse operators, we expect the relation $u(S)v(S) = 1$, equivalent to the chain of relations

$$u_0 v_0 = 1, \;\; \sum_{k=0}^{n} \binom{n}{k} u_k v_{n-k} = 0 \;\text{ for } n \geq 1, \tag{62}$$

to hold. Indeed, this is easily checked (see Section 3.4, formula (38)).

Since the input $U_n(t)$ corresponds to the output t^n, a Taylor-MacLaurin expansion of the output corresponds to an expansion of the input in terms of the $U_n(t)$. Fix an epoch t_0 and use the Taylor expansion of the output

$$F(t) = \sum_{n \geq 0} \frac{1}{n!} \mathbf{D}^n F(t_0)(t - t_0)^n. \tag{63}$$

Applying the operator $\mathbf{U}$, we get

$$f(t) = \sum_{n \geq 0} \frac{1}{n!} \mathbf{D}^n F(t_0) \cdot U_n(t - t_0), \tag{64}$$

since $\mathbf{U}$ transforms $(t - t_0)^n$ into $U_n(t - t_0)$ by stationarity. The reader is invited to compare this formula to formula (39).

We give one illustrating example. Let the output be a moving average of the input

$$F(t) = \int_{t-1}^{t} f(s)ds, \tag{65}$$

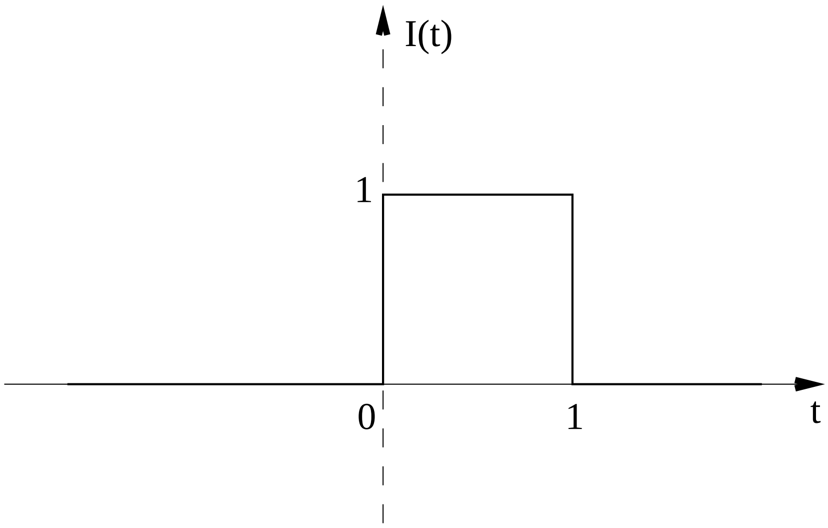

Fig. 3. The impulse response corresponding to (65)

corresponding to the impulse response shown in Figure 3.

The spectral gain is

$$\Theta(p) = \int_0^1 e^{-p\tau} d\tau = \frac{1 - e^{-p}}{p}. \tag{66}$$

The inverse series $u(p) = 1/\Theta(p)$ is given by

$$u(p) = \frac{p}{1 - e^{-p}}. \tag{67}$$

Since $u(-p) = \frac{p}{e^p - 1}$ is the exponential generating series of the Bernoulli numbers, the polynomials $U_n(t)$ are easily identified,

$$U_n(t) = (-1)^n B_n(-t) = B_n(t) + nt^{n-1} = B_n(t+1). \tag{68}$$

Notice that $\Theta(p)$ vanishes for $p \neq 0$ of the form $p = 2\pi in$ with an integer n; equivalently, the inverse function $u(p)$ has poles for $p \neq 0,\ p = 2\pi in$. Hence not every output is admissible since (65) entails $\sum_n F(t+n) = \int_{-\infty}^{+\infty} f(s)ds$. That is, an output satisfies the necessary (and sufficient) condition

$$\sum_n F(t+n) = c \quad (c \text{ a constant}) \tag{69}$$

and the input $f(t)$ can be reconstructed from the output $F(t)$ up to the addition of a function $f_0(t)$ with

$$f_0(t) = f_0(t+1), \quad \int_0^1 f_0(t)dt = 0. \tag{70}$$

Exercise a) Derive from (64) the relation

$$f(t) = \sum_{n \geq 0} \frac{1}{n!} u_n \mathbf{D}^n F(t), \tag{71}$$

for a general transmission device, where $1/\Theta(p) = \sum_{n \geq 0} u_n p^n / n!$.

b) In the particular case (65), one gets $u_n = B_n(1)$, hence $u_n = B_n$ if $n \geq 2$, and $u_0 = 1$, $u_1 = \frac{1}{2}$.

c) Deduce from relation (7) that $B_n = 0$ for $n \geq 3$, n odd.

d) Derive the Euler-MacLaurin summation formula

$$\begin{aligned} &\frac{1}{2}(f(t) + f(t-1)) \\ &= \int_{t-1}^t f(s)ds + \sum_{m \geq 1} \frac{1}{(2m)!} B_{2m}[\mathbf{D}^{2m-1} f(t) - \mathbf{D}^{2m-1} f(t-1)]. \end{aligned} \tag{72}$$

3.7 A probabilistic application

We consider a random variable ξ. In general, we denote by $\langle X \rangle$ the mean value of a random variable X. We want to define a probabilistic version of the so-called Wick Powers in Quantum Field Theory.

The goal is to associate to ξ a sequence of random variables $:\xi^n:$ such that

a) **the mean value of** $:\xi^n:$ **is** 0 for $n \geq 1$;

b) **there exists a normalized binomial sequence of polynomials** $\Pi_n(X)$ **such that** $:\xi^n: = \Pi_n(\xi)$.

Let $w(x)$ be the probability density associated to ξ, hence $w(x) \geq 0$ and $\int_{-\infty}^{+\infty} w(x)dx = 1$. Moreover, for any (non random) function $f(x)$ of a real variable x, the random variable $f(\xi)$ has a mean value given by

$$\langle f(\xi) \rangle = \int_{-\infty}^{+\infty} f(x)w(x)dx. \tag{73}$$

Hence the conditions a) and b) amount to

$$0 = \langle \Pi_n(\xi) \rangle = \int_{-\infty}^{+\infty} \Pi_n(x) w(x) dx \quad \text{for } n \geq 1. \tag{74}$$

Using the same method as in Section 3.2, we introduce the exponential generating series $\Pi(X, S) = \sum_{n \geq 0} \Pi_n(X) S^n / n!$, hence the relation (74) translates into

$$\int_{-\infty}^{+\infty} \Pi(x, S) w(x) dx = 1. \tag{75}$$

Putting $\pi(S) = \Pi(0, S)$, hence $\Pi(X, S) = \pi(S) e^{XS}$, we derive

$$1/\pi(S) = \int_{-\infty}^{+\infty} e^{xS} w(x) dx. \tag{76}$$

We translate these relations into probabilistic jargon: replace S by p and X by ξ to get

$$1/\pi(p) = \langle e^{p\xi} \rangle \tag{77}$$

$$\Pi(\xi, p) = \sum_{n \geq 0} \frac{1}{n!} p^n : \xi^n : \tag{78}$$

$$\Pi(\xi, p) = \pi(p) e^{p\xi}. \tag{79}$$

Extending the definition of $: \quad :$ by linearity to $e^{p\xi} = \sum_{n \geq 0} \frac{1}{n!} p^n \xi^n$, we rewrite (78) as $\Pi(\xi, p) =: e^{p\xi} :$. **Here is the conclusion**

$$: e^{p\xi} : \quad = \quad \frac{e^{p\xi}}{\langle e^{p\xi} \rangle}. \tag{80}$$

Let us specialize our results in the case of the binomial sequences considered so far:

a) If $\xi = 0$, then $\langle e^{p\xi} \rangle = 1$, hence $: e^{p\xi} := e^{p\xi} = 1$. That is $: \xi^n := 0$ for $n \geq 1$.

b) Suppose that ξ is **uniformly distributed** in the interval $[0, 1]$, that is $w(x) = 1$ if $0 \leq x \leq 1$, and $w(x) = 0$ otherwise. Then

$$\langle e^{p\xi} \rangle = \int_0^1 e^{px} dx = \frac{e^p - 1}{p}. \tag{81}$$

We get

$$\sum_{n \geq 0} \frac{1}{n!} p^n : \xi^n : \quad = \quad : e^{p\xi} : \quad = \quad \frac{p e^{p\xi}}{e^p - 1}, \tag{82}$$

that is

$$:\xi^n: \quad = \quad B_n(\xi) \tag{83}$$

where $B_n(X)$ is the Bernoulli polynomial of degree n. In particular

$$\begin{aligned}
:\xi: \quad &= \quad \xi - \langle \xi \rangle = \xi - \frac{1}{2} \\
:\xi^2: \quad &= \quad \xi^2 - \xi + \frac{1}{6} \\
:\xi^3: \quad &= \quad \xi^3 - \frac{3}{2}\xi^2 + \frac{1}{2}\xi, \quad \text{etc}\ldots
\end{aligned}$$

c) Assume now that ξ is **normalized**: $\langle \xi \rangle = 0$, $\langle \xi^2 \rangle = 1$, and follows a **Gaussian law**. Then $w(x) = (2\pi)^{-1/2} e^{-x^2/2}$ and

$$\langle e^{p\xi} \rangle = (2\pi)^{-1/2} \int_{-\infty}^{+\infty} e^{px - x^2/2} dx = e^{p^2/2}. \tag{84}$$

Reasoning as above, we obtain

$$:\xi^n: \quad = \quad H_n(\xi) \tag{85}$$

where $H_n(X)$ is the Hermite polynomial of degree n. Explicitly

$$\begin{aligned}
:\xi: \quad &= \quad \xi \\
:\xi^2: \quad &= \quad \xi^2 - 1 \\
:\xi^3: \quad &= \quad \xi^3 - 3\xi.
\end{aligned}$$

To get a general formula, apply (80) to obtain the pair of relations

$$:e^{p\xi}: \quad = \quad e^{-p^2/2} e^{p\xi}, \ e^{p\xi} = e^{p^2/2} :e^{p\xi}: . \tag{86}$$

Equating equal powers of p, we derive

$$:\xi^n: \quad = \quad \sum_{0 \leq k \leq n/2} (-1)^k \frac{n!}{2^k k! (n-2k)!} \xi^{n-2k}, \tag{87}$$

and conversely

$$\xi^n = \sum_{0 \leq k \leq n/2} \frac{n!}{2^k k! (n-2k)!} :\xi^{n-2k}: . \tag{88}$$

Notice that the orthogonality relation (47) for the Hermite polynomials translates in probabilistic terms into

$$\langle :\xi^m: \quad :\xi^n:\rangle = m!\delta_{mn}, \tag{89}$$

hence the sequence $1, :\xi:, :\xi^2:, \ldots$ is derived from the natural sequence $1,\ \xi,\ \xi^2, \ldots$ by orthogonalization.

To conclude, we can use the reflected probability density $w(-x)$ as an impulse response and define the input-output relation by

$$F(t) = \int f(t+\tau)w(\tau)d\tau, \tag{90}$$

that is

$$F(t) = \langle f(t+\xi)\rangle \tag{91}$$

in probabilistic terms. The interpretation is that the input is spoiled by random delay in transmission. Then $\Pi_n(t)$ is the input corresponding to the output t^n. Analytically this is expressed by

$$\int_{-\infty}^{+\infty} \Pi_n(t+\tau)w(\tau)d\tau = t^n, \tag{92}$$

and probabilistically by

$$\langle :(\xi+t)^n:\rangle = t^n. \tag{93}$$

3.8 The Bargmann-Segal transform

Let us consider again the input-output transformation in the Gaussian case. It is then called the Bargmann-Segal transform (or B-transform), denoted by $\mathbf{B}$. According to (90), we have

$$\mathbf{B}f(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x+z)e^{-x^2/2}dx \tag{94}$$

or

$$\mathbf{B}f(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(x)e^{-(z-x)^2/2}dx. \tag{95}$$

Comparing formulas (24) and (28), one obtains

$$e^{-(z-x)^2/2} = e^{-x^2/2} \sum_{n\geq 0} H_n(x)z^n/n!, \tag{96}$$

and by integrating term by term in the expression (95) one concludes

$$\mathbf{B}f(z) = \sum_{n\geq 0} \Gamma_n(f) z^n / n! \tag{97}$$

with

$$\Gamma_n(f) = \int_{-\infty}^{+\infty} H_n(x) f(x) d\gamma(x). \tag{98}$$

Taking into account the orthogonality property (47) namely

$$\int_{-\infty}^{+\infty} H_m(x) H_n(x) d\gamma(x) = \delta_{mn} n!,$$

one derives $\Gamma_n(f) = n! c_n$ for f given by a series $\sum_{n\geq 0} c_n H_n(x)$. That is, **the B-transform takes $\sum_{n\geq 0} c_n H_n(x)$ into $\sum_{n\geq 0} c_n z^n$.**

To be more precise, we need to introduce some function spaces. The natural one is $L^2(d\gamma)$ consisting of the (measurable) functions $f(x)$ for which the integral $\int_{-\infty}^{+\infty} |f(x)|^2 d\gamma(x)$ is finite; the scalar product is given by

$$\langle f_1 | f_2 \rangle = \int_{-\infty}^{+\infty} \overline{f_1(x)} f_2(x) d\gamma(x). \tag{99}$$

In this space, the functions $He_n(x) := H_n(x)/(n!)^{1/2}$ (for $n = 0, 1, \ldots$) form an orthonormal basis[5]. The B-transform takes this space onto the space of series $\sum c_n z^n$ with $\sum_{n\geq 0} n! |c_n|^2 < \infty$.

In its original form (94), the transformation $\mathbf{B}$ requires z to be real, but the form (95) extends to the case of a complex number z. Indeed, from the property that $\sum_{n\geq 0} n! |c_n|^2$ is finite, it follows that the series $\sum_{n\geq 0} c_n z^n$ has an infinite radius of convergence, hence represents an entire function of the complex variable z. The space of such entire functions is denoted $\mathcal{F}(\mathbf{C})$ and called the **Fock space** (in one degree of freedom, see Section 3.9).

The elements of $L^2(d\gamma)$ can be interpreted as the random variables of the form $X = f(\xi)$ with $\langle |X|^2 \rangle$ finite, where ξ is a normalized Gaussian random variable. We saw that $\mathbf{B}$ takes $H_n(\xi) = \, : \xi^n :$ into z^n. Hence it is tempting to denote by $: \quad :$ the map inverse to $\mathbf{B}$, so that $: z^n : \, = H_n(x)$. We have **a**

[5] The orthonormality condition $\langle He_m | He_n \rangle = \delta_{mn}$ is nothing else than the orthogonality condition (47). But it requires a proof to show that this system is **complete**, that is that any function in the Hilbert space $L^2(d\gamma)$ can be approximated by polynomials (in the norm convergence).

new kind of operational calculus

$$L^2(d\gamma) \underset{:\ :}{\overset{\mathbf{B}}{\rightleftharpoons}} \mathcal{F}(\mathbf{C}),$$

where B transforms a function $f(x)$ **or the random variable** $X = f(\xi)$ **into the entire function**

$$\mathbf{B}X(z) = e^{-z^2/2}\langle X \cdot e^{z\xi}\rangle \tag{100}$$

and the inverse map takes an entire function $\Phi(z) = \sum_{n\geq 0} c_n z^n$ **in the Fock space into the random variable** $:\Phi(\xi) := \sum_{n\geq 0} c_n : \xi^n :$.

According to the definition (94), $\mathbf{B}$ takes the function e^{px} into $e^{pz+p^2/2}$, that is it acts as $e^{\mathbf{D}^2/2}$ where $\mathbf{D}$ is the derivation, followed by the change of variable x into z. This result can be reformulated as follows. Using the exponential generating series

$$\sum_{n\geq 0} H_n(x)p^n/n! = e^{px-p^2/2} \tag{101}$$

for the Hermite polynomials, and noting that $e^{\mathbf{D}^2/2}$ applied to $e^{pz-p^2/2}$ gives e^{px}, we conclude that $e^{\mathbf{D}^2/2}$ takes $H_n(x)$ into x^n, that is **B coincides on the polynomials with the differential operator** $e^{\mathbf{D}^2/2}$ **of infinite degree**.

We would like to conclude the general rule

$$\mathbf{B} = e^{\mathbf{D}^2/2}, \tag{102}$$

hence

$$:\ : \quad = e^{-\mathbf{D}^2/2}. \tag{103}$$

One way to substantiate these claims is to consider the heat (or diffusion) equation

$$\partial_s F(s,x) = \frac{1}{2}\partial_x^2 F(s,x) \tag{104}$$

with initial value

$$F(0,x) = f(x). \tag{105}$$

Since $\mathbf{D} = \partial_x$, the solution of equation (104) can be written formally as $F(s,x) = e^{s\mathbf{D}^2/2}f(x)$, hence $e^{\mathbf{D}^2/2}f(x)$ **represents the value for** $s = 1$ **of**

the solution of equation (104) which agrees for $s = 0$ **with** $f(x)$. But we know an explicit solution to the heat equation

$$F(s, x) = \frac{1}{\sqrt{2\pi s}} \int_{-\infty}^{+\infty} f(x+u) e^{-u^2/2s} du. \tag{106}$$

Comparing with (94), we obtain

$$\mathbf{B} f(x) = F(1, x), \tag{107}$$

and this relation is the true expression of $\mathbf{B} = e^{\mathbf{D}^2/2}$.

The operator $e^{\mathbf{D}^2/2}$ (or $\mathbf{B}$) is **smoothing**. That is, if we simply assume that f belongs to $L^2(d\gamma)$ (that is that the integral $\int_{-\infty}^{+\infty} |f(x)|^2 e^{-x^2/2} dx$ is finite), then the function $F(1, x) = e^{\mathbf{D}^2/2} f(x)$ extends to an entire function in the complex domain. Conversely, the **Wick operator** $: \quad : = e^{-\mathbf{D}^2/2}$ makes sense only for the functions $g(x)$ (for x real) which extend in the complex domain into a function $\Phi(z)$ (for z complex) belonging to the Fock space $\mathcal{F}(\mathbf{C})$.

3.9 The quantum harmonic oscillator

Let us rewrite the definition of the B-transform as an integral operator

$$\mathbf{B} f(z) = \int_{-\infty}^{+\infty} B(z, x) f(x) dx \tag{108}$$

with a kernel

$$B(z, x) = (2\pi)^{-1/2} e^{-(z-x)^2/2}. \tag{109}$$

It is often more convenient to replace the Hilbert space $L^2(d\gamma)$ by the Hilbert space $L^2(\mathbf{R})$ [6]. Defining the function $u_0(x)$ by $(2\pi)^{-1/4} e^{-x^2/4}$, we get

$$d\gamma(x) = u_0(x)^2 dx, \tag{110}$$

hence $\int |f(x)|^2 d\gamma(x)$ is finite if and only if $\int |f(x) u_0(x)|^2 dx$ is finite. That is, the multiplication by the function $u_0(x)$ gives an isometry of $L^2(d\gamma)$ onto

[6] consisting of the (measurable) functions $\phi(x)$ such that $\int_{-\infty}^{+\infty} |\phi(x)|^2 dx$ be finite, with scalar product $\langle \phi_1 | \phi_2 \rangle = \int_{-\infty}^{+\infty} \overline{\phi_1(x)} \phi_2(x) dx$.

$L^2(\mathbf{R})$. We can transfer the B-transform to $L^2(\mathbf{R})$, as the isometry $\mathbf{B}'$ of $L^2(\mathbf{R})$ onto $\mathcal{F}(\mathbf{C})$ [7] defined by $\mathbf{B}f = \mathbf{B}'(fu_0)$. Explicitly

$$\mathbf{B}'f(z) = \int_{-\infty}^{+\infty} B'(z,x)f(x)dx \tag{111}$$

with

$$B'(z,x) = u_0(x)^{-1}B(z,x) = (2\pi)^{-1/4}e^{-z^2/2+zx-x^2/4}. \tag{112}$$

Many properties are easier to describe in the Fock space. For instance the function 1 is called the **ground state** Ω, the multiplication by z is called the **creation operator**, denoted by $\mathbf{a}^*$, and the derivation $\partial_z = \frac{d}{dz}$ is the **annihilation operator**, denoted by $\mathbf{a}$. The vectors

$$e_n = \frac{1}{\sqrt{n!}}(\mathbf{a}^*)^n\Omega, \tag{113}$$

that is the functions $e_n(z) = \frac{1}{\sqrt{n!}}z^n$, form an orthonormal basis of $\mathcal{F}(\mathbf{C})$ with $e_0 = \Omega$. An easy calculation gives

$$\begin{cases} \mathbf{a}e_n = n^{1/2}e_{n-1} \text{ for } n \geq 1,\ \mathbf{a}e_0 = 0 \\ \mathbf{a}^*e_n = (n+1)^{1/2}e_{n+1}, \end{cases} \tag{114}$$

hence the matrices

$$\mathbf{a} = \begin{pmatrix} 0 & \sqrt{1} & 0 & 0 & \ldots \\ 0 & 0 & \sqrt{2} & 0 & \ldots \\ 0 & 0 & 0 & \sqrt{3} & \ldots \\ 0 & 0 & 0 & 0 & \ldots \\ . & . & . & . & \ldots \\ . & . & . & . & \ldots \end{pmatrix}, \quad \mathbf{a}^* = \begin{pmatrix} 0 & 0 & 0 & 0 & \ldots \\ \sqrt{1} & 0 & 0 & 0 & \ldots \\ 0 & \sqrt{2} & 0 & 0 & \ldots \\ 0 & 0 & \sqrt{3} & 0 & \ldots \\ . & . & . & . & \ldots \\ . & . & . & . & \ldots \end{pmatrix}$$

in the basis $(e_n)_{n\geq 0}$; it follows that **a and $\mathbf{a}^*$ are adjoint to each other**. Moreover from the definitions $\mathbf{a}^* = z$, $\mathbf{a} = \partial_z$ follows the commutation relation

$$\mathbf{a}\mathbf{a}^* - \mathbf{a}^*\mathbf{a} = 1. \tag{115}$$

[7] with the scalar product $\langle \Phi_1 | \Phi_2 \rangle = \sum_{n\geq 0} n!\bar{c}_{n,1}c_{n,2}$ for $\Phi_j(z) = \sum_{n\geq 0} c_{n,j}z^n$ $(j = 1,2)$.

Finally the **number operator** $\mathbf{N} = \mathbf{a}^*\mathbf{a}$ is given by $\mathbf{N} = z\partial_z$, hence is diagonalized in the basis (e_n)

$$\mathbf{N}e_n = ne_n. \tag{116}$$

In the spirit of operational calculus, we transfer these results from the Fock space model to the spaces $L^2(d\gamma)$ and $L^2(\mathbf{R})$. The following table summarizes these translations (where ∂_x is $\frac{d}{dx}$):

Space	$L^2(d\gamma)$	$L^2(\mathbf{R})$	$\mathcal{F}(\mathbf{C})$
Ω	1	$(2\pi)^{-1/4}e^{-x^2/4} = u_0(x)$	1
e_n	$(n!)^{-1/2}H_n(x)$	$(n!)^{-1/2}H_n(x)u_0(x)$	$(n!)^{-1/2}z^n$
$\mathbf{a}^*$	$x - \partial_x$	$x/2 - \partial_x$	z
$\mathbf{a}$	∂_x	$x/2 + \partial_x$	∂_z
$\mathbf{N}$	$x\partial_x - \partial_x^2$	$-\partial_x^2 + x^2/4 - 1/2$	$z\partial_z$

For instance, the fact that $\mathbf{a}^*$ corresponds to $x - \partial_x$ in $L^2(d\gamma)$ is proved as follows: from the definition of $B(z,x)$ one gets

$$(x + \partial_x)B(z,x) = zB(z,x). \tag{117}$$

Multiplying by $f(x)$ and integrating by parts, we get

$$\begin{aligned}&\int B(z,x)(x-\partial_x)f(x)dx = \int (x+\partial_x)B(z,x)f(x)dx \\ &= z\int B(z,x)f(x)dx,\end{aligned}$$

that is $\mathbf{B}((x-\partial_x)f) = z\mathbf{B}f$. The other cases are similar.

We apply these results to the **harmonic oscillator**. In classical mechanics, the harmonic oscillator is described by the **Hamiltonian** $H = \frac{p^2}{2m} + \frac{Kq^2}{2}$ in canonical coordinates p,q. The equation of motion is $\ddot{q} + \omega^2 q = 0$ with the pulsation $\omega = \sqrt{K/m}$, and the momentum $p = m\dot{q}$. To get the corresponding quantum Hamiltonian $\mathbf{H}$, replace p by the operator $\mathbf{p} = -i\hbar\partial_q$ hence

$$\mathbf{H} = -\frac{\hbar^2}{2m}\partial_q^2 + \frac{m\omega^2 q^2}{2}. \tag{118}$$

Introduce the dimensionless coordinate $x = (2m\omega/\hbar)^{1/2}q$. Then $\mathbf{H}$ can be rewritten as

$$\mathbf{H} = \hbar\omega(\mathbf{a}^*\mathbf{a} + \frac{1}{2}) \tag{119}$$

in the model $L^2(\mathbf{R})$. From the diagonalization of $\mathbf{N} = \mathbf{a}^*\mathbf{a}$, we conclude that the energy levels of the quantum harmonic oscillator (that is, the eigenvalues of $\mathbf{H}$) are given by $\hbar\omega(n+\frac{1}{2})$ with $n = 0, 1, 2, \ldots$: that is **Planck's radiation law**, with the correction $\frac{1}{2}$ giving $\frac{1}{2}\hbar\omega$ for the energy of the ground state $u_0 = \Omega$.

4 The art of manipulating infinite series

4.1 Some divergent series

Euler claimed that $S = 1 - 1 + 1 - 1 + \ldots$ is equal to $\frac{1}{2}$. Here is the purported proof:

$$\begin{array}{rllll} S = 1 & -1 & +1 & -1 + \ldots & \\ + S = & & 1 & -1 & +1 - \ldots \\ \hline 2\,S = 1 & +0 & +0 & +0 + \ldots = 1. & \end{array}$$

What is implicit is the use of two rules:

a) If $S = u_0 + u_1 + u_2 + \ldots$, then $S = 0 + u_0 + u_1 + \ldots$

b) If $S = u_0 + u_1 + u_2 + \ldots$ and $S' = u_0 + u_1' + u_2' + \ldots$, then $S + S' = (u_0 + u_0') + (u_1 + u_1') + (u_2 + u_2') + \ldots$.

These rules certainly hold for convergent series but to extend them to divergent series is somewhat hazardous.

Let us repeat the previous calculation in a slightly more general form:

$$\begin{array}{rllll} S = 1 & -t & +t^2 & -t^3 + \ldots & \\ + tS = & & t & -t^2 & +t^3 - \ldots \\ \hline (1+t)S = 1 & +0 & +0 & +0 + \ldots = 1. & \end{array}$$

The result is

$$1 - t + t^2 - t^3 + \ldots = \frac{1}{1+t}, \tag{1}$$

the classical summation of the geometric series. If t is a real number such that $|t| < 1$, the geometric series is convergent, and the use of rules a) and b) is justified. To get Euler's result, take the limiting value $t = 1$ in (1).

What we need is the explicit description of various procedures to define rigorously the sum of certain divergent series (not all at once) and to compare these procedures. **Suppose we want to define the sum**

$$S = u_0 + u_1 + \ldots. \tag{2}$$

Introduce weights $p_{0,t}, p_{1,t}, \ldots$ **and the weighted series**

$$S_t = p_{0,t}u_0 + p_{1,t}u_1 + \ldots. \tag{3}$$

If the series S_t **is convergent for each value of the parameter** t**, and** S_t **approaches a limit** S **when** t **approaches some limiting value** t_0**, then** S **is the sum for this procedure**[8].

The previous procedure is reasonable only when $\lim_{t\to t_0} p_{n,t} = 1$ for $n = 0, 1, \ldots$. Some examples:

a) $p_{0,N} = p_{1,N} = \ldots = p_{N,N} = 1$, $p_{n,N} = 0$ for $n > N$ and $N = 0, 1, 2, \ldots$. Then the weighted sum amounts to the finite sum

$$S_N = u_0 + \ldots + u_N$$

(obviously convergent) and the convergence of S_N towards a limit S corresponds to the convergence of the series $u_0 + u_1 + u_2 + \ldots$ in the standard sense, with the standard sum S.

b) Put $\sigma_N = \frac{1}{N+1}(S_0 + \ldots + S_N)$; this corresponds to the weights

$$p_{n,N} = \begin{cases} 1 - \frac{n}{N+1} & \text{for } 0 \leq n \leq N \\ 0 & \text{for } n > N. \end{cases} \tag{4}$$

If σ_N converges to a limit σ, this is the **Cesaro-sum** of the series $u_0 + u_1 + u_2 + \ldots$.

[8] See Knopp's book [8] for this method.

c) To get the **Abel summation**, we introduce the weights $p_{n,t} = t^n$ for $n = 0, 1, 2, \ldots$ and a real parameter t with $0 < t < 1$. We take therefore the limit for $t = 1$ of the power series $\sum_{n \geq 0} u_n t^n$.

It is known that every convergent series with sum S is Cesaro-summable with the same sum $\sigma = S$. Similarly, Cesaro summation is extended by Abel summation. Euler's example is $u_n = (-1)^n$, hence

$$S_N = \begin{cases} 1 & \text{if } N \geq 0 \text{ is even} \\ 0 & \text{if } N \geq 0 \text{ is odd} \end{cases} \tag{5}$$

and therefore

$$\sigma_N = \begin{cases} \frac{1}{2} & \text{if } N \text{ is odd} \\ \frac{1}{2} + \frac{1}{2N+2} & \text{if } N \text{ is even.} \end{cases} \tag{6}$$

It follows that σ_N converges to $\sigma = \frac{1}{2}$. Hence the series $1 - 1 + 1 - 1 + \ldots$ is Cesaro-summable to $\frac{1}{2}$, and a previous calculation shows that it is Abel-summable to $\frac{1}{2}$ also.

The scope of Abel summation can be extended in various ways. For instance, if the sequence (u_n) is bounded, that is $|u_n| \leq M$ for $n = 0, 1, 2, \ldots$ with a constant M independent of n, then the power series $\sum_{n \geq 0} u_n z^n$ converges for any complex number z with $|z| < 1$ and defines therefore a holomorphic function $U(z)$ in the open disk $|z| < 1$ (see Fig. 4). If the limit $\lim_{r \to 1} U(re^{i\theta})$(for $0 \leq r < 1$) exists, it can be taken as an Abel sum for the series $\sum_{n \geq 0} u_n e^{in\theta}$.

In a slightly more general way, we can assume that the sequence (u_n) is polynomially bounded, that is

$$|u_n| \leq Cn^k$$

for all $n = 1, 2, \ldots$ and some constants $C > 0$ and $k = 1, 2, \ldots$ The radius of convergence of the series $\sum_{n \geq 0} u_n z^n$ is still at least 1, and if $U(1) = \lim_{r \to 1} U(r) = \lim_{r \to 1} \sum_{n \geq 0} u_n r^n$ exists, it is the Abel sum for $u_0 + u_1 + u_2 + \ldots$

We just give one example, namely $u_n = (-1)^n n^k$ for $k = 0, 1, 2, \ldots$. We calculate

$$\begin{aligned} U(z) &= \sum_{n \geq 0} u_n z^n = \sum_{n \geq 0} (-1)^n n^k z^n = \sum_{n \geq 0} n^k (-z)^n \\ &= \sum_{n \geq 0} (z\partial_z)^k (-z)^n = (z\partial_z)^k \sum_{n \geq 0} (-z)^n = (z\partial_z)^k \frac{1}{1+z}. \end{aligned}$$

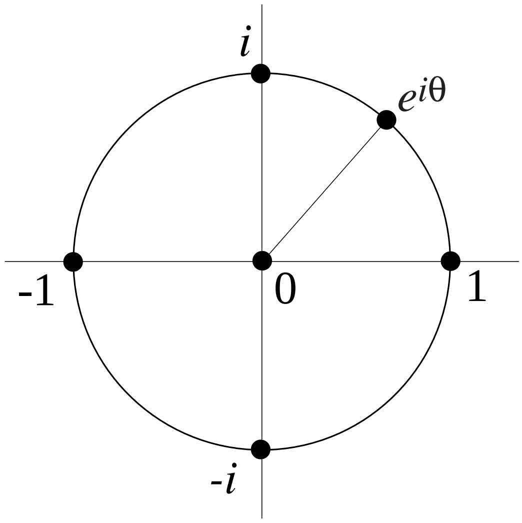

Fig. 4. The open unit disk

Particular cases:

$$
\begin{aligned}
k &= 0, \quad & U(z) &= \frac{1}{1+z}, & U(1) &= \frac{1}{2} \\
k &= 1, \quad & U(z) &= \frac{-z}{(1+z)^2}, & U(1) &= -\frac{1}{4} \\
k &= 2, \quad & U(z) &= \frac{z(z-1)}{(1+z)^3}, & U(1) &= 0 \\
k &= 3, \quad & U(z) &= \frac{-z^3+4z^2-z}{(1+z)^4}, & U(1) &= \frac{1}{8},
\end{aligned}
$$

that is

$$\sum_{n\geq 0}(-1)^n = \frac{1}{2}$$

$$\sum_{n\geq 0}(-1)^n n = -\frac{1}{4}$$

$$\sum_{n\geq 0}(-1)^n n^2 = 0$$

$$\sum_{n\geq 0}(-1)^4 n^3 = \frac{1}{8}.$$

In general, we get $U(1) = (z\partial_z)^k \frac{1}{1+z}|_{z=1}$ that is, after the change of variable $z = e^u$,

$$\sum_{n\geq 0}(-1)^n n^k = \partial_u^k \frac{1}{e^u+1}\Big|_{u=0}. \tag{7}$$

Using the exponential generating series for the Bernoulli numbers in the form

$$\frac{1}{e^u-1} = \frac{1}{u} + \sum_{k\geq 0} \frac{B_{k+1}}{(k+1)!}u^k$$

we obtain

$$\frac{1}{e^u+1} = \frac{1}{e^u-1} - \frac{2}{e^{2u}-1} = \sum_{k\geq 0} \frac{(1-2^{k+1})B_{k+1}}{(k+1)!}u^k, \tag{8}$$

hence Euler's result

$$\sum_{n\geq 0}(-1)^n n^k = \frac{(1-2^{k+1})B_{k+1}}{k+1}. \tag{9}$$

We leave it to the reader to rederive the previous cases $0 \leq k \leq 3$ using the values for B_1, B_2, B_3, B_4 given in Section 3.1. We come back to this result in Section 4.4.

Euler gave formulas for wildly divergent series like $\sum_{n\geq 0}(-1)^n n!$. Using the classical formula

$$n! = \int_0^\infty e^{-t}t^n dt \tag{10}$$

and assuming unrestrictedly as legitimate formal rules both term-by-term integration and summation of a geometric series, we get

$$\begin{aligned}\sum_{n\geq 0}(-1)^n n! &= \sum_{n\geq 0}(-1)^n \int_0^\infty e^{-t}t^n dt\\ &= \int_0^\infty e^{-t}\sum_{n\geq 0}(-t)^n \, dt = \int_0^\infty \frac{e^{-t}}{1+t}dt,\end{aligned}$$

the last integral being convergent. This is just the beginning of the use of Borel transform and Borel summation for divergent series.

4.2 Polynomials of infinite degree and summation of series

It is an important principle that **a polynomial can be reconstructed from its roots**. More precisely, let

$$P(z) = c_n z^n + c_{n-1} z^{n-1} + \ldots + c_1 z + c_0 \tag{11}$$

(with $c_n \neq 0$) be a polynomial of degree n with complex coefficients. If λ_1 is a root of P, that is $P(\lambda_1) = 0$, it is elementary to factorize $P(z) = (z-\lambda_1)P_1(z)$ where $P_1(z)$ is a polynomial of degree $n-1$. Continuing this process, we end up with a factorization

$$P(z) = (z-\lambda_1)\ldots(z-\lambda_m)Q(z), \tag{12}$$

where the polynomial $Q(z)$ of degree $n-m$ has no more roots. According to a highly non-trivial result, first stated by d'Alembert (1746) and proved by Gauß (1797), a polynomial without roots is a constant, hence the factorization (12) takes the form

$$P(z) = c_n(z-\lambda_1)\ldots(z-\lambda_n). \tag{13}$$

By a well known calculation, one derives the following relations between coefficients and roots

$$\begin{aligned} &\lambda_1 + \ldots + \lambda_n = -c_{n-1}/c_n \\ &\sum_{i<j} \lambda_i \lambda_j = c_{n-2}/c_n, \ \text{etc}\ldots \end{aligned}$$

For our purposes, it is better to use the inverses of the roots, assumed to be nonzero. Since the logarithmic derivative transforms product into sum and annihilates constants, we derive

$$\mathbf{D}P(z)/P(z) = \sum_{i=1}^{n} \frac{1}{z-\lambda_i}. \tag{14}$$

Use of the geometric series gives

$$\sum_{i=1}^{n} \frac{1}{z-\lambda_i} = -\sum_{i=1}^{n} \sum_{k \geq 0} z^k/\lambda_i^{k+1}. \tag{15}$$

Introducing the sums of inverse powers of roots

$$\gamma_k = \sum_{i=1}^{n} \lambda_i^{-k}, \tag{16}$$

we conclude from this calculation

$$z\mathbf{D}P(z) + P(z) \sum_{k\geq 1} \gamma_k z^k = 0. \tag{17}$$

Assuming for simplicity $c_0 = 1$ and equating the coefficients of equal powers of z, we obtain the following variant of **Newton's relations**

$$\gamma_k + c_1\gamma_{k-1} + \ldots + c_{k-1}\gamma_1 + kc_k = 0 \tag{18}$$

for $k \geq 1$. It is important to notice that **the degree n of $P(z)$ does not appear explicitly in the relation (18)**, which can be solved inductively

$$\gamma_1 = -c_1 \tag{19}$$
$$\gamma_2 = c_1^2 - 2c_2 \tag{20}$$
$$\gamma_3 = -c_1^3 + 3c_1c_2 - 3c_3 \tag{21}$$
$$\gamma_4 = c_1^4 - 4c_1^2c_2 + 4c_1c_3 - 4c_4 + 2c_2^2. \tag{22}$$

Around 1734, Euler undertook to calculate the sum of the series $S_2 = \sum_{n\geq 1} \frac{1}{n^2}$. This series is slowly convergent, but Euler invented efficient acceleration methods for summing series and calculated the sum $S_2 = 1.64493406\ldots$; he recognized $S_2 = \pi^2/6$. He obtained also the value of $S_4 = \sum_{n\geq 1} 1/n^4$ to be $\pi^4/90$. To establish these results rigorously, he introduced the equation $\sin x = 0$ admitting the solutions $x = 0, \pm\pi, \pm 2\pi, \pm 3\pi, \ldots$ Discarding the root $x = 0$ and using the power series expansion of $\sin x$, we are led to consider the equation

$$1 - \frac{x^2}{6} + \frac{x^4}{120} - \ldots = 0$$

with roots $\pm\pi, \pm 2\pi, \pm 3\pi, \ldots$. With the previous notations we have

$$c_1 = 0,\ c_2 = -\frac{1}{6},\ c_3 = 0,\ c_4 = \frac{1}{120}, \ldots$$

$$\gamma_2 = \sum_{n\geq 1} \left[\frac{1}{(\pi n)^2} + \frac{1}{(-\pi n)^2} \right] = 2S_2/\pi^2$$

$$\gamma_4 = \sum_{n\geq 1} \left[\frac{1}{(\pi n)^4} + \frac{1}{(-\pi n)^4} \right] = 2S_4/\pi^4.$$

Assuming that the relations (20) and (22) still hold, we get

$$2S_2/\pi^2 = \gamma_2 = -2c_2 = \frac{1}{3}$$
$$2S_4/\pi^4 = \gamma_4 = -4c_4 + 2c_2^2 = -\frac{1}{30} + \frac{1}{18} = \frac{1}{45}.$$

The desired relations

$$S_2 = \pi^2/6, \ S_4 = \pi^4/90$$

follow immediately.

To summarize the method used by Euler:

a) first guess the value from accurate numerical work;

b) consider the function

$$\frac{\sin x}{x} = 1 - \frac{x^2}{6} + \frac{x^4}{120} - \dots$$

as a polynomial of infinite degree, with infinitely many roots $\pm\pi, \pm 2\pi$, $\pm 3\pi, \dots$;

c) since Newton's relations (19) to (22) **do not involve explicitly the degree n of the polynomial**, assume their validity in the case $n = \infty$ as well, and exploit them for $P(x) = (\sin x)/x$.

4.3 The Euler-Riemann zeta function

We use Riemann's definition and notation

$$\zeta(s) = \sum_{n \geq 1} n^{-s}. \tag{23}$$

The series converges absolutely for any complex number s with real part $\Re(s)$ greater than 1. It has been shown by Riemann that $\zeta(s)$ can be analytically continued to the whole complex plane, the only singularity being a pole of order 1 at $s = 1$, that is $\zeta(s) - 1/(s-1)$ is an entire function. Obviously $\zeta(1) = \sum_{n \geq 1} 1/n$ is a divergent series, but $\zeta(s)$ is defined when $s \neq 1$ is an integer (positive or negative). Euler was the first to calculate $\zeta(s)$ when s is an integer.

We consider the case where s is **even and strictly positive**. Euler proved the formula

$$\zeta(2k) = \frac{2^{2k-1}\pi^{2k}|B_{2k}|}{(2k)!}, \tag{24}$$

and in particular

$$\zeta(2) = \frac{2\pi^2|B_2|}{2!} = \frac{\pi^2}{6} \tag{25}$$

$$\zeta(4) = \frac{8\pi^4|B_4|}{4!} = \frac{\pi^4}{90}. \tag{26}$$

Since $\zeta(2) = \sum_{n\geq 1} 1/n^2 = S_2$ and similarly $\zeta(4) = S_4$, we recover the formulas for S_2 and S_4. The method we used in the previous section could be extended to cover the general case (24), but it is simpler to go back to the formula given for the logarithmic derivative in (14). For the function $\sin z$, the logarithmic derivative is $\cot z = \cos z / \sin z$. This function is meromorphic in the whole complex plane, with simple poles of residue 1 at each integral multiple of π. **Euler assumed at first that, in analogy with (14),** $\cot z$ **should be equal to the sum of its polar contributions, that is**

$$\cot z = \sum_{n=-\infty}^{+\infty} \frac{1}{z - n\pi}. \tag{27}$$

Assume this relation for a moment, and derive (24). The series (27) is not absolutely convergent, but can be summed in a symmetrical way by taking $\sum_{n=-\infty}^{+\infty}$ to be $\lim_{N\to\infty} \sum_{n=-N}^{+N}$. Hence

$$\cot z - \frac{1}{z} = \sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2\pi^2}. \tag{28}$$

The right-hand side can be developed using the geometric series; for $|z| < \pi$, the series involved are absolutely convergent, hence

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{2z}{z^2 - n^2\pi^2} &= -\sum_{n=1}^{\infty} 2z \sum_{k=1}^{\infty} \frac{(z^2)^{k-1}}{(n^2\pi^2)^k} = -2\sum_{n=1}^{\infty}\sum_{k=1}^{\infty} \frac{z^{2k-1}}{n^{2k}\pi^{2k}} \\ &= -2\sum_{k=1}^{\infty} \frac{z^{2k-1}}{\pi^{2k}} \sum_{n=1}^{\infty} \frac{1}{n^{2k}},\end{aligned}$$

that is

$$\cot z = \frac{1}{z} - 2\sum_{k\geq 1} \frac{\zeta(2k)}{\pi^{2k}} z^{2k-1}. \tag{29}$$

Another use of the exponential generating series for the Bernoulli numbers yields

$$\begin{aligned}\cot z &= i\frac{e^{2iz}+1}{e^{2iz}-1} = \frac{2i}{e^{2iz}-1} + i \\ &= 2i\Big\{\frac{1}{2iz} - \frac{1}{2} + \sum_{k\geq 1}\frac{B_{2k}}{(2k)!}(2iz)^{2k-1}\Big\} + i,\end{aligned}$$

hence finally

$$\cot z = \frac{1}{z} + \sum_{k\geq 1}\frac{(-1)^k 2^{2k} B_{2k}}{(2k)!} z^{2k-1}. \tag{30}$$

To establish (24), it is enough to compare the expansions (29) and (30) for $\cot z$ and to note that $\zeta(2k) = \sum_{n\geq 1}\frac{1}{n^{2k}}$ is the sum of a convergent series of positive numbers, hence $\zeta(2k) > 0$.

Euler's proof for the expansion (27) of $\cot z$ is reproduced in many text-books. Here is a variant which seems to have been unnoticed so far. Define

$$\Phi(z) = \cot z - \sum_{n=-\infty}^{+\infty}\frac{1}{z-n\pi}. \tag{31}$$

Examining the poles of $\cot z$, we see that $\Phi(z)$ is an entire function of the complex variable z. A simple manipulation yields the functional equation

$$\Phi(z) = \frac{1}{2}\left[\Phi\left(\frac{z}{2}\right) + \Phi\left(\frac{z+\pi}{2}\right)\right]; \tag{32}$$

we have to prove that $\Phi(z) = 0$ for all z.

a) **The function Φ is bounded**: indeed, denote by C_n the set of complex numbers whose modulus is at most $(2^n+1)\pi$. Since C_1 is a compact set and Φ is continuous, there exists a constant $M > 0$ such that $|\Phi(z)| \leq M$ for z in C_1. Assuming the estimate $|\Phi(z)| \leq M$ for z in C_n, we use the functional equation for z in C_{n+1},

$$|\Phi(z)| \leq \frac{1}{2}\left|\Phi\left(\frac{z}{2}\right)\right| + \frac{1}{2}\left|\Phi\left(\frac{z+\pi}{2}\right)\right|, \tag{33}$$

and observe that both $z/2$ and $(z+\pi)/2$ belong to C_n, hence $|\Phi(z/2)| \leq M$, $|\Phi((z+\pi)/2)| \leq M$; from (33) we conclude that $|\Phi(z)| \leq M$ (for z in C_{n+1}). Every complex number belongs to some set C_n, hence $|\Phi(z)| \leq M$ for all z.

b) We appeal now to **Liouville's theorem** to conclude that Φ, being a bounded entire function, is a constant, hence $\Phi(z) = \Phi(0)$.

c) The function Φ is **odd**, that is $\Phi(-z) = -\Phi(z)$, hence $\Phi(0) = 0$.

Liouville's theorem, the main ingredient in this proof, was proved around 1850, a century after Euler worked on these questions. It is interesting to note that the d'Alembert-Gauß theorem is an easy corollary of Liouville's theorem. [**Hint**: if $P(z)$ is a polynomial without roots, the function $\Phi(z) = 1/P(z)$ is entire and bounded, hence a constant; that is, $P(z)$ is a constant.]

4.4 Sums of powers of numbers

The other result of Euler about $\zeta(s)$ can be stated as follows:

$$1^k + 2^k + 3^k + \ldots = -\frac{B_{k+1}}{k+1} \tag{34}$$

for $k = 1, 2, 3, \ldots$ It looks at first odd, since it gives a finite value for an infinite sum of positive numbers, obviously divergent since each term is at least 1. Euler's derivation is more or less as follows.

Formula (9) can be written as

$$1^k - 2^k + 3^k - 4^k + \ldots = -(1 - 2 \cdot 2^k)\frac{B_{k+1}}{k+1}. \tag{35}$$

On the other hand, multiply the left-hand side of (34) by $1 - 2 \cdot 2^k$ and rearrange. This yields

$$\begin{array}{rl}
 & (1 - 2 \cdot 2^k)(1^k + 2^k + 3^k + \ldots) = \\
 & \quad 1^k + 2^k + 3^k + 4^k + 5^k + 6^k + \ldots \\
-2(& \qquad\quad 2^k \qquad + 4^k \qquad + 6^k + \ldots) \\
\hline
= & \quad 1^k - 2^k + 3^k - 4^k + 5^k - 6^k + \ldots
\end{array}$$

and finally (34) is obtained from (35).

This procedure is highly questionable, but can be fixed as follows. We introduce two functions

$$\zeta(s) = \sum_{n \geq 1} n^{-s}, \quad \eta(s) = \sum_{n \geq 1} (-1)^{n-1} n^{-s}. \tag{36}$$

Provided these functions can be continued analytically to the negative integers, formula (34) and (35) read respectively as

$$\zeta(-k) = -\frac{B_{k+1}}{k+1} \tag{34'}$$

$$\eta(-k) = (2^{k+1}-1)\frac{B_{k+1}}{k+1} \tag{35'}$$

for $k = 1, 2, \ldots$ Furthermore

$$\begin{aligned}\eta(s) &= \sum_{\text{n odd}} n^{-s} - \sum_{\text{n even}} n^{-s} = \sum_{n\geq 1} n^{-s} - 2\sum_{\text{n even}} n^{-s} \\ &= \zeta(s) - 2\sum_{m\geq 1}(2m)^{-s}\end{aligned}$$

and finally

$$\eta(s) = (1-2^{1-s})\zeta(s). \tag{37}$$

Our manipulation of series is justified as long as $\Re(s) > 1$, but the final formula remains valid for all s for which both $\zeta(s)$ and $\eta(s)$ are regular (analytic continuation!). In particular

$$\eta(-k) = (1-2^{k+1})\zeta(-k). \tag{38}$$

Hence formulas (34′) and (35′) are equivalent, substantiating Euler's derivation.

Using the known values of the Bernoulli numbers, we deduce

$$\begin{aligned}&\zeta(-2) = \zeta(-4) = \zeta(-6) = \ldots = 0 \\ &\zeta(-1) = -\frac{1}{12},\ \zeta(-3) = \frac{1}{120},\ \zeta(-5) = -\frac{1}{252}, \ldots.\end{aligned}$$

It can also be shown that $\zeta(0) = -1/2$. Hence we get the paradoxical results:

$$\begin{aligned}\zeta(0) &= 1+1+1+\ldots = -\frac{1}{2} \\ \zeta(-1) &= 1+2+3+\ldots = -\frac{1}{12}.\end{aligned}$$

Among the many methods available to construct the analytical continuation of $\zeta(s)$, we select the following one using $\eta(s)$. Indeed, from Euler's definition of the gamma function

$$\Gamma(s) = \int_0^\infty e^{-t}t^{s-1}dt \tag{39}$$

(for $\Re(s) > 1$), we deduce by a simple change of variable

$$\Gamma(s) n^{-s} = \int_0^\infty e^{-nt} t^{s-1} dt. \tag{40}$$

By summation and term by term integration

$$\begin{aligned} \Gamma(s)\eta(s) &= \sum_{n\geq 1} (-1)^{n-1} \Gamma(s) n^{-s} = \sum_{n\geq 1} (-1)^{n-1} \int_0^\infty e^{-nt} t^{s-1} dt \\ &= \int_0^\infty \Big\{ \sum_{n\geq 1} (-1)^{n-1} e^{-nt} \Big\} \cdot t^{s-1} dt, \end{aligned}$$

that is

$$\Gamma(s)\eta(s) = \int_0^\infty \frac{t^{s-1}}{e^t + 1} dt. \tag{41}$$

All the calculations are justified as long as $\Re(s) > 1$. We use now a general principle, unknown to Euler, to deal with integrals of the type

$$\Phi(s) = \int_0^\infty F(t) t^{s-1} dt. \tag{42}$$

We split the integral as $\int_0^1 + \int_0^\infty$.

a) Assuming that $F(t)$ decreases at infinity faster than any power t^{-k} (for $k \geq 0$), then

$$\Phi_{1,\infty}(s) = \int_1^\infty F(t) t^{s-1} dt \tag{43}$$

extends to an entire function.

b) Assuming that $F(t)$ is differentiable to any order on the closed interval $[0, 1]$, then

$$\Phi_{0,1}(s) = \int_0^1 F(t) t^{s-1} dt \tag{44}$$

extends to a meromorphic function in the complex plane $\mathbf{C}$. The only singularities[9] are at $s = 0, -1, -2, \ldots$ with singular part $\frac{\mathbf{D}^k F(0)}{k!(s+k)}$ around $s = -k$.

Applying this principle to the definition (39) of $\Gamma(s)$ we recover the well-known fact that $\Gamma(s)$ extends to a meromorphic function, with poles at $s = 0, -1, -2, \ldots$ and singular part $\frac{(-1)^k}{k!(s+k)}$ around $s = -k$. We use now formula

[9] **Hint**: integrate by parts using $\frac{d}{dt} t^{-s} = -s t^{-s-1}$.

(41) for $\Gamma(s)\eta(s)$. Hence this function extends to a meromorphic function with poles at $s = 0, -1, \ldots$ and singular part $\frac{c_k}{s+k}$ around $s = -k$, where

$$\frac{1}{e^t + 1} = \sum_{k \geq 0} c_k t^k. \tag{45}$$

According to (8), we get

$$c_k = \frac{(1 - 2^{k+1}) B_{k+1}}{(k+1)!}. \tag{46}$$

Dividing $\Gamma(s)\eta(s)$ by $\Gamma(s)$, the poles cancel; hence $\eta(s)$ extends to an entire function, and comparing the singular parts of $\Gamma(s)\eta(s)$ and $\Gamma(s)$ around $s = -k$, we find

$$\eta(-k) = (-1)^k k! c_k. \tag{47}$$

We have to distinguish between several cases:

- $k = 0$ yields $\eta(0) = c_0 = -B_1 = \frac{1}{2}$;
- $k \geq 2$ is even yields $\eta(-k) = 0$ since $B_r = 0$ for $r = k + 1$ odd;
- $k \geq 1$ is odd, then $(-1)^k = -1$ and

$$\eta(-k) = \frac{(2^{k+1} - 1) B_{k+1}}{k+1}. \tag{48}$$

This is Euler's formula (35) or formula (35′). The analytical continuation of $\zeta(s)$ can now be performed by using (37), that is we define

$$\zeta(s) = \frac{\eta(s)}{1 - 2^{1-s}}. \tag{49}$$

Since $\eta(s)$ is entire, the only singularity of $\zeta(s)$ is a pole at $s = 1$, with singular part $\frac{1}{s-1}$. We can now calculate $\zeta(-k)$ from $\eta(-k)$ and get

$$\zeta(-k) = -\frac{B_{k+1}}{k+1} \tag{50}$$

for $k = 1, 2, \ldots$, as expected. Furthermore, from $\eta(0) = \frac{1}{2}$ we get the remaining value

$$\zeta(0) = -\eta(0) = -\frac{1}{2}. \tag{51}$$

4.5 Variation I: Did Euler really fool himself?

Bourbaki wrote (in [2], page VI.29): "Mais la tendance au calcul formel est la plus forte, et l'extraordinaire intuition d'Euler lui-même ne l'empêche pas de tomber parfois dans l'absurde, lorsqu'il écrit par exemple $0 = \sum_{n=-\infty}^{+\infty} x^n$". Did Euler really fool himself?

To keep with our habits (after Cauchy!) denote by z a complex variable and try to evaluate the sum of $I = \sum_{n=-\infty}^{+\infty} z^n$. We break the sum into $I_+ + I_- - 1$, where

$$I_+ = \sum_{n \geq 0} z^n, \quad I_- = \sum_{n \leq 0} z^n.$$

By the geometric series, we get $I_+ = \frac{1}{1-z}$ and $I_- = \frac{1}{1-1/z}$ and simple algebra gives

$$I_+ + I_- = \frac{1}{1-z} + \frac{z}{z-1} = \frac{1-z}{1-z} = 1, \tag{52}$$

hence $I = 0$ as claimed. What is paradoxical is that there is no complex number $z \neq 0$ for which both series I_+ and I_- converge simultaneously, since $\sum_{n\geq 0} z^n$ converges for $|z| < 1$ and $\sum_{n \leq 0} z^n$ converges for $|z| > 1$. We really need analytical continuation: I_+ as a function of z extends from the convergence domain $|z| < 1$ to $\mathbf{C} - \{1\}$ as the rational function $\frac{1}{1-z}$, and one goes from $I_+$to I_- by inverting z (into $1/z$). If both I_+ and I_- are extended in this way to $\mathbf{C} - \{1\}$, the calculation (52) is perfectly valid, hence $I = 0$ in this sense.

Another method to prove $I = 0$ is to observe that multiplication of I by z shifts z^n to z^{n+1}, hence rearranges the series, hence $Iz = I$, hence $I(z-1) = 0$, and by dividing by $z - 1$ we get $I = 0$ for $z \neq 1$. Nevertheless, there is some trouble. Consider the critical region $|z| = 1$ where both I_+ and I_- diverge, and use polar coordinates $z = e^{2\pi i u}$. Then I is the series

$$J(u) = \sum_{n=-\infty}^{+\infty} e^{2\pi i n u}. \tag{53}$$

Playing with Fourier series, introduce a test function $f(u)$ supposed to be smooth (i.e., infinitely differentiable) and periodic, $f(u+1) = f(u)$. We expand it as a Fourier series

$$f(u) = \sum_{n=-\infty}^{+\infty} c_n e^{2\pi i n u} \tag{54}$$

with

$$c_n = \int_0^1 e^{-2\pi i n u} f(u) du. \tag{55}$$

From (54) we get, by putting $u = 0$,

$$f(0) = \sum_{n=-\infty}^{+\infty} c_n \tag{56}$$

hence by (55) and (53)

$$f(0) = \sum_{n=-\infty}^{+\infty} \int_0^1 e^{-2\pi i n u} f(u) du = \int_0^1 \left\{ \sum_{n=-\infty}^{+\infty} e^{-2\pi i n u} \right\} \cdot f(u) du,$$

and finally

$$f(0) = \int_0^1 J(u) f(u) du. \tag{57}$$

Remove now the assumption $f(u+1) = f(u)$ by introducing a smooth function $\phi(u)$ vanishing off some finite interval and by defining

$$f(u) = \sum_{m=-\infty}^{+\infty} \phi(u+m) \tag{58}$$

(an absolutely convergent series). By an easy manipulation, one derives from (57)

$$\sum_{m=-\infty}^{+\infty} \phi(m) = \int_{-\infty}^{+\infty} J(u) \phi(u) du. \tag{59}$$

Using the standard Dirac's function $\delta(u)$, we get by definition

$$\phi(m) = \int_{-\infty}^{+\infty} \phi(u) \delta(u - m) du,$$

hence

$$0 = \int_{-\infty}^{+\infty} \left\{ J(u) - \sum_{m=-\infty}^{+\infty} \delta(u-m) \right\} \phi(u) du. \tag{60}$$

Since the test function ϕ is arbitrary, we can omit it from (60), hence the conclusion

$$J(u) = \sum_{m=-\infty}^{+\infty} \delta(u-m). \tag{61}$$

That is, by substituting $e^{2\pi i u}$ to z, **the series** $I = \sum_{n=-\infty}^{+\infty} z^n$ **is not** 0 **but** $\sum_{m=-\infty}^{+\infty} \delta(u-m)$. So Euler was wrong, but not too much, since $\delta(u-m) = 0$ for $u \neq m$, hence $\sum_{n=-\infty}^{+\infty} z^n$ is 0 for $z \neq 1$.

Recall the other proof, using

$$I(z-1) = 0; \tag{62}$$

division by $z-1$ gives $I = 0$, provided $z \neq 1$, corresponding to $u \notin \mathbf{Z}$ for $z = e^{2\pi i u}$. Formula (62) is equivalent to

$$J(u)(e^{2\pi i u} - 1) = 0, \tag{63}$$

and this suggests a new proof of (61). Indeed, if $f(u)$ is a smooth function with isolated simple zeroes u_m, then $J(u)f(u) = 0$ implies that $J(u)$ is a linear combination of terms $c_m \delta(u - u_m)$. Here $f(u) = e^{2\pi i u} - 1$, hence $u_m = m$ for m in $\mathbf{Z}$, that is $m = 0, \pm 1, \pm 2, \ldots$ hence $J(u) = \sum_{m=-\infty}^{+\infty} c_m \delta(u-m)$ for suitable coefficients c_m. But $J(u+1) = J(u)$, hence all coefficients c_m are equal to some constant c and $J(u) = c\sum_{m=-\infty}^{+\infty} \delta(u-m)$. It remains to calculate the normalization constant c. That kind of argument could be understood by Euler, but it acquires now a rigorous meaning due to Laurent Schwartz's theory of distributions (200 years after Euler!)[10].

Another version of our proof is by using a contour integral (see Fig. 5). Consider a function $\Phi(z)$ holomorphic in a domain containing the annulus $r \leq |z| \leq R$ bounded by C_+ and C_- (beware the orientations). The rational function $R(z) = \frac{1}{z-1}$ is given by a convergent series $\sum_{n=-\infty}^{-1} z^n$ for $|z| > 1$, hence for z in C_+. It follows

$$\int_{C_+} R(z)\Phi(z)dz = \sum_{n=-\infty}^{-1} \int_{C_+} z^n \Phi(z)dz. \tag{64}$$

Similarly

$$\int_{C_-} R(z)\Phi(z)dz = \sum_{n=0}^{\infty} \int_{C_-} z^n \Phi(z)dz, \tag{65}$$

and using the residue formula,

$$\int_{C_+} \sum_{n=-\infty}^{-1} z^n \cdot \Phi(z)dz + \int_{C_-} \sum_{n=0}^{\infty} z^n \cdot \Phi(z)dz = 2\pi i \Phi(1).$$

[10] Our final result can be expressed as $\sum_{n=-\infty}^{+\infty} e^{2\pi i n u} = \sum_{m=-\infty}^{+\infty} \delta(u-m)$. It is equivalent to Poisson's summation formula.

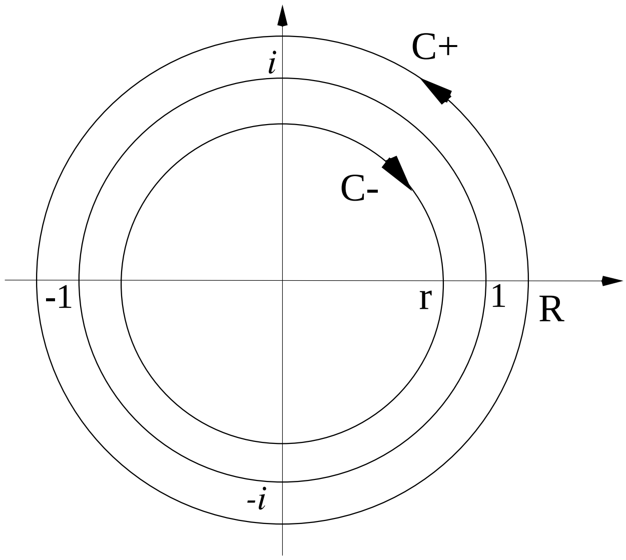

Fig. 5. Path for the contour integral

A shorthand would be

$$\sum_{n=-\infty}^{+\infty} z^n = 2\pi i\delta(z-1) \tag{66}$$

using δ-functions in the complex domain[11].

Let us go back to sums of powers and Bernoulli numbers and polynomials. A classical formula reads as follows:

$$B_k(u) = -k! \sum_{n\neq 0} \frac{e^{2\pi i n u}}{(2\pi i n)^k}. \tag{67}$$

[11] A classical formula is

$$\delta(f(u)) = \sum_m \frac{1}{|f'(u_m)|}\delta(u-u_m),$$

the summation being extended over the solutions of the equation $f(u_m)=0$ (provided $f'(u_m)\neq 0$). In this formula $f(u)$ is a real-valued function of a real variable u. Assuming that it remains valid for $f(u) = e^{2\pi i u}-1$ (with complex values), we derive

$$2\pi i\delta(z-1) = \sum_{m=-\infty}^{+\infty} \delta(u-m)$$

for $z = e^{2\pi i u}$. This brings together our two methods.

A complex version is the following

$$\sum_{n\neq 0} \frac{z^n}{n^k} = -\frac{(2\pi i)^k}{k!} B_k \left(\frac{\log z}{2\pi i}\right) \qquad (68)_k$$

(by the change of variables $z = e^{2\pi i u}$) for $k = 0, 1, 2, \ldots$. For $k = 0$, this reads as Euler's "absurd formula"

$$\sum_{n\neq 0} z^n = -1, \qquad (68)_0$$

since $B_0(x) = 1$. The case $k = 1$ is

$$\sum_{n\neq 0} \frac{z^n}{n} = \pi i - \log z; \qquad (68)_1$$

of course

$$\sum_{n=1}^{\infty} \frac{z^n}{n} = \log \frac{1}{1-z}$$
$$\sum_{n=-\infty}^{-1} \frac{z^n}{n} = -\sum_{m=1}^{\infty} \frac{(z^{-1})^m}{m} = -\log \frac{1}{1-1/z}$$

and $(68)_1$ amounts to

$$\log \frac{1}{1-z} - \log \frac{1}{1-1/z} = \pi i - \log z. \qquad (69)$$

Since $\log(-1) = -\pi i$ and $\frac{1}{1-1/z} = \frac{z}{z-1} = \frac{z\cdot(-1)}{1-z}$, this relation follows from $\log uv = \log u + \log v$, but some care has to be exercized with the multivalued complex logarithm (notice the ambiguity $\log(-1) = \pm\pi i$ for instance).

For the general case, notice the following. Define

$$L_k(z) = \sum_{n\neq 0} \frac{z^n}{n^k}, \quad R_k(z) = -\frac{(2\pi i)^k}{k!} B_k \left(\frac{\log z}{2\pi i}\right). \qquad (70)$$

We know already that $L_0(z) = R_0(z)$ and $L_1(z) = R_1(z)$. Furthermore, it is obvious that

$$z\frac{d}{dz} L_k(z) = L_{k-1}(z), \qquad (71)$$

and from the fact that the derivative of $B_k(x)$ is $kB_{k-1}(x)$, one gets

$$z\frac{d}{dz}R_k(z) = R_{k-1}(z). \tag{72}$$

So we can easily conclude that $L_2(z) - R_2(z)$ is a constant, which has to be shown to be 0 to prove $L_2(z) = R_2(z)$. Then $L_3(z) - R_3(z)$ is a constant, etc...

This line of argument can be made rigorous. Introduce the polylogarithmic functions[12]

$$\mathrm{Li}_k(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^k}. \tag{73}$$

Our formula reads now as follows:

$$\mathrm{Li}_k(z) + (-1)^k \mathrm{Li}_k\left(\frac{1}{z}\right) = -\frac{(2\pi i)^k}{k!} B_k\left(\frac{\log z}{2\pi i}\right). \tag*{$(74)_k$}$$

To make sense of it, we proceed as follows:

a) We cut the complex plane along the real interval $[0, +\infty[$, to get $\Omega_0 = \mathbf{C} - [0, +\infty[$.

b) In the cut plane, we choose the somewhat unusual branch of the logarithm $\log(re^{i\theta}) = \log r + i\theta$ for $0 < \theta < 2\pi$.

c) We define the function $\mathrm{Li}_k(z)$ by the convergent series (73) for $|z| < 1$, and verify that

$$z\frac{d}{dz}\mathrm{Li}_k(z) = \mathrm{Li}_{k-1}(z) \tag{75}$$

for $k = 1, 2, \ldots$ and $\mathrm{Li}_0(z) = \frac{z}{1-z}$. Since the cut plane $\Omega_1 = \mathbf{C} - [1, \infty[$ is simply connected, any holomorphic function in Ω_1 has a primitive, hence by (75), each $\mathrm{Li}_k(z)$ extends analytically to Ω_1.

d) For z in Ω_0, both z and $\frac{1}{z}$ are in Ω_1, hence both $\mathrm{Li}_k(z)$ and $\mathrm{Li}_k(\frac{1}{z})$ are defined for z in Ω_0, and formula $(74)_k$ is asserted for z in Ω_0.

e) The cases $k = 0$ and $k = 1$ are settled as before.

f) From (75) and the rule for the derivative of $B_k(x)$, we get that the validity of $(74)_k$ for the index k implies that of $(74)_{k+1}$ for the index $k + 1$

[12] The dilogarithm $Li_2(z)$ was known to Euler, and further developed in the 19^{th} century in connection with Lobatchevski geometry. Fifteen years ago, the subject was almost forgotten, to be resurrected by geometers and mathematical physicists alike. It is now a hot subject of research.

up to the addition of a constant. To show that it is 0 use the fact that for $k \geq 2$, the series $\sum_{n=1}^{\infty} \frac{z^n}{n^k}$ converges also for $|z| = 1$, and study the limiting value for $z \to 1$, using $B_k(0) = B_k(1)$.

So after all, Euler was right!

Putting $z = 1$ in $(74)_k$ we obtain the value of $\zeta(k) + (-1)^k \zeta(k)$. For k odd, we get $0 = 0$, but for k even, we recover the value of $\zeta(k)$ given by (24).

4.6 Variation II: Infinite products

Suppose we want to calculate $\infty! = 1 \cdot 2 \cdot 3 \cdots$. Going to logarithms we define

$$\infty! = \exp\left(\sum_{n=1}^{\infty} \log n\right). \tag{76}$$

Suppose we have a series $\sum_{n \geq 1} a_n$ with na_n bounded. What is known as the ζ-**summation procedure** fits our general framework at the beginning of Section 4.1: consider the convergent series $\sum_{n \geq 1} a_n n^{-\epsilon}$ for $\epsilon > 0$, and let ϵ tend to 0. To sum the series $\sum_{n \geq 1} \log n$, we should consider $\sum_{n \geq 1} n^{-\epsilon} \cdot \log n$ but this converges for $\epsilon > 1$ only and we cannot go directly to the limit $\epsilon \to 0$. What we have to do is to consider the series $\sum_{n \geq 1} n^{-s} \cdot \log n$ for $\Re(s) > 1$; this is obviously the derivative $-\zeta'(s)$ of the Riemann zeta function, hence it can be analytically continued to the neighborhood of 0. The regularized sum of $\sum_{n \geq 1} \log n$ is then $-\zeta'(0)$ and finally

$$\infty! = e^{-\zeta'(0)}. \tag{77}$$

From the formulas (37) and (41), one derives without much ado $\zeta'(0) = -\frac{1}{2} \log 2\pi$ (a formula more or less equivalent to Stirling's formula). Conclusion:

$$\infty! = \sqrt{2\pi}. \tag{78}$$

General rule: to normalize a divergent product $\Pi_{n \geq 1} a_n$, introduce the series $\sum_{n \geq 1} a_n^{-s} = Z(s)$, make an analytic continuation from the convergence domain $\Re(s) > \sigma_0$ to $s = 0$ and define

$$\prod_{n \geq 1}^{\text{reg}} a_n := e^{-Z'(0)}. \tag{79}$$

Generalizing slightly (78), we can use this method to prove the identity[13]

$$\prod_{n\geq 0}^{\mathrm{reg}} (n+v) = \frac{\sqrt{2\pi}}{\Gamma(v)}. \tag{80}$$

We can compare this to the Weierstraß product expansion for the gamma function

$$\frac{1}{\Gamma(v)} = ve^{\gamma v} \prod_{n\geq 1} \left(1+\frac{v}{n}\right) e^{-v/n}. \tag{81}$$

A careless, but nevertheless instructive, comparison of (80) and (81) is as follows:

$$\begin{aligned}\prod_{n\geq 1} \left(1+\frac{v}{n}\right) e^{-v/n} &= \prod_{n\geq 1} \frac{n+v}{n} e^{-v/n} \\ = \prod_{n\geq 1} (n+v) \Big(\prod_{n\geq 1} n\Big)^{-1} &\exp\Big(-v\sum_{n\geq 1} 1/n\Big)\end{aligned}$$

and regularizing the divergent series $\sum_{n\geq 1} \frac{1}{n}$ by the Euler constant γ, we are through! Notice that the two most important properties of Γ, namely

1) the functional equation $\Gamma(v+1) = v\Gamma(v)$;

2) the function $\frac{1}{\Gamma(v)}$ of a complex variable v is entire with zeroes at $0, -1, -2, \ldots$

can be read off immediately from (80).

According to our general method, the proof of (80) requires to study the function

$$\zeta(s,v) = \sum_{n\geq 0} (n+v)^{-s} \tag{82}$$

known as Hurwitz zeta function (see [9] for more details). We list a few properties:

a) **a particular case** $\zeta(s,1) = \zeta(s)$;

b) **functional equations**:

$$\zeta(s,v+1) = \zeta(s,v) - v^{-s} \tag{83}$$

$$\partial_v \zeta(s,v) = -s\zeta(s+1,v); \tag{84}$$

[13] Formally: $\frac{\infty!}{(\infty+v)!} = \Gamma(v+1)$.

c) **analytic continuation**: for fixed v, $\zeta(s,v)$ can be analytically continued to the complex plane with one singularity at $s=1$, with singular part $\frac{1}{s-1}$; hence $\zeta(s,v)-\zeta(s)$ is an entire function;

d) **special values**:

$$\zeta(-k,v) = -\frac{B_{k+1}(v)}{k+1} \tag{85}$$

for $k = 0, 1, 2, \ldots$

The last relation can be written, in the spirit of Euler, as

$$v^k + (v+1)^k + (v+2)^k + \ldots = -\frac{B_{k+1}(v)}{k+1}. \tag{86}$$

As a particular case we get the surprising identity

$$v^0 + (v+1)^0 + (v+2)^0 + \ldots = \frac{1}{2} - v. \tag{87}$$

5 Conclusion: From Euler to Feynman

Feynman is the modern heir to Euler. Among his many contributions to theoretical physics, the most famous one is his use of diagrams to encode in a very compact way complicated integrals with significance in experiments in high energy physics. His method of diagrams has been generalized by various authors (Cvitanovic, Penrose,...) to provide a very flexible tool for computations in tensor analysis.

His really bold discovery is the use of integrals in function spaces (see for instance [5]), the so-called Feynman path integrals. These (so far) ill-defined integrals are powerful tools to evaluate infinite series and infinite products. We give just one example. Consider the Hilbert space $L^2(0, 2\pi)$ of functions $f(x)$ with $0 < x < 2\pi$ and $\int_0^{2\pi} |f(x)|^2 dx$ finite. The unbounded operator $\Delta = -d^2/dx^2$ can be diagonalized with eigenfunctions $e_n(x) = e^{inx}$ (for $n = 0, \pm 1, \pm 2, \ldots$) corresponding to the eigenvalue n^2. Hence the characteristic determinant $\det(v - \Delta)$ is an entire function with the eigenvalues as zeroes. Using our regularized products, one now defines the regularized determinant as

$$\det{}^{\mathrm{reg}}(v - \Delta) = v \left(\prod_{n \geq 1}^{\mathrm{reg}} (v - n^2)^2 \right) \tag{1}$$

(0 is a simple eigenvalue, and $1^2, 2^2, \ldots$ are eigenvalues of multiplicity 2). This can be evaluated by a formula due to Euler

$$\sin v = v \prod_{n \geq 1} \left(1 - \frac{v^2}{n^2\pi^2}\right), \tag{2}$$

equivalent (via logarithmic derivatives) to the formula

$$\cot v = \frac{1}{v} + \sum_{n \geq 1} \frac{2v}{v^2 - n^2\pi^2} \tag{3}$$

also due to Euler, and considered above.

Feynman's bold step is as follows. From matrix calculus, we learn the following integral formula for a characteristic determinant

$$\det(v - A) = \left[\int_{\mathbf{R}^n} d^n x \ \exp -\pi \left(v \sum_{i=1}^{n} x_i^2 - \sum_{i,j} a_{i,j} x_i x_j\right)\right]^{-2} \tag{4}$$

where $d^n x$ is the volume element $dx_1 \cdots dx_n$ in the Euclidean space $\mathbf{R}^n$, and $A = (a_{i,j})$ is a real symmetric, positive definite, matrix of size $n \times n$. By analogy, Feynman writes $\det(v - \Delta)$ as the square of the inverse of

$$\int_{L^2(0,2\pi)} \mathcal{D}x \cdot \exp(-\pi S(x)), \tag{5}$$

where the so-called action $S(x)$ is defined by

$$S(x) = v \int_0^{2\pi} x(t)^2 dt + \int_0^{2\pi} x'(t)^2 dt \tag{6}$$

(the variable in $[0, 2\pi]$ is denoted by t, the function in $L^2(0, 2\pi)$ by $x(t)$, and its derivative by $x'(t)$). The symbol $\mathcal{D}x$ is formally a volume element in the Hilbert space $L^2(0, 2\pi)$ (infinite-dimensional generalization of the Euclidean space $\mathbf{R}^n$), sometimes written as $C \prod_t dx(t)$. Its rigorous definition is the main problem [5].

Part of these calculations have been put into a rigorous framework, but not all of them.

After all, Feynman shall be right!

Acknowledgements

Ils vont à Michel Planat et Michel Waldschmidt, pour leur impeccable organisation du colloque de Chapelle des Bois, et pour leur conviction que ce texte verrait le jour. Leurs amicales pressions et leur aide matérielle ont grandement contribué à l'heureux dénouement. Merci aussi à Christian Krattenthaler pour une relecture extrêmement soignée de ce texte.

References

1. Borel E. (1902) Leçons sur les séries à termes positifs. Gauthier-Villars.
2. Bourbaki N. (1976) Fonctions d'une variable réelle (chapitres 1 à 7). Hermann, edited by Masson since 1980.
3. Bourbaki N. (1984) Eléments d'histoire des mathématiques. Masson.
4. Cartier P. (1966) Quantum mechanical relations and theta functions, in Algebraic groups and discontinuous subgroups. Proc. Symp. Pure Maths, vol. IX, American Mathematical Society, Providence, 361-383.
5. DeWitt-Morette C., Cartier P. and Folacci P. (editors) (1997) Functional Integration, Basics and Applications, Nato Asi, Series B, vol 361. Plenum Press.
6. Euler L. (1755) Institutiones calculi differentialis cum ejus usu in analysi infinitorum ac doctrina serierum. Berlin, 1910.
7. Hardy G. H. (1924) Orders of infinity. Cambridge Univ. Press (reprinted by Hafner, New York, 1971).
8. Knopp K. (1964) Theorie und Anwendung der unendlichen Reihen. Springer (5th edition).
9. Waldschmidt M. *et al* (1995) From Number Theory to Physics. Springer (see Chapter 1 by P. Cartier: An introduction to zeta functions).
10. Weil A. (1984) Number Theory, An Approach Through History. Birkhäuser.