Superior Mathematics from an Elementary point of view

2018/2019 Edition

Jack D'Aurizio

www.matemate.it

“Superior Mathematics from an Elementary point of view” course notes

Undergraduate course, 2017-2018, University of Pisa

Jack D’Aurizio

Contents

0 Introduction

This course has been designed to serve University students of the first and second year of Mathematics. The purpose of these notes is to give elements of both *strategy* and *tactics* in problem solving, by explaining ideas and techniques willing to be *elementary* and *powerful* at the same time. We will not focus on a single subject among Calculus, Algebra, Combinatorics or Geometry: we will just try to enlarge the "toolbox" of any professional mathematician wannabe, by starting from humble requirements:

- knowledge of number sets $(\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C})$ and their properties;
- knowledge of mathematical terminology and notation;
- knowledge of the main mathematical functions;
- knowledge of the following concepts: limits, convergence, derivatives, Riemann integral;
- knowledge of basic Combinatorics and Arithmetics.

A collection of problems in Analysis and Advanced integration techniques, kindly provided by Tolaso J. Kos and Zaid Alyafeai, are excellent sources of exercises to match with the study of these notes.

1 Creative Telescoping and DFT

It is soon evident, during the study of Mathematics, that the bijectivity of some function f does not grant that the explicit computation of $f^{-1}(y)$ is "just as easy" as the explicit computation of $f(x)$. Some examples are related to the ease of multiplication, against the hardness of factorization; the possibility of computing derivatives in a algorithmic fashion, against the lack of a completely algorithmic way to find indefinite integrals; the determination of a Galois group of an irreducible polynomial over $\mathbb{Q}$, against the difficult task of finding a polynomial having a given Galois group. In the present section we will outline two interesting techniques for solving (or getting arbitrarily close to an actual "solution") a peculiar *inverse problem*, that is the computation of series.

Definition 1. We give the adjective *telescopic* to objects of the form $a_1 + a_2 + \ldots + a_n$, where each a_i can be written as $b_i - b_{i+1}$ for some sequence $b_1, b_2, \ldots, b_n, b_{n+1}$. With such assumption we have:

$$a_1 + a_2 + \ldots + a_n = (b_1 - b_2) + (b_2 - b_3) + \ldots + (b_n - b_{n+1}) = b_1 - b_{n+1}.$$

Essentially, every telescopic sum is *simple to compute*, just as any convergent series with terms of the form $b_i - b_{i+1}$. A peculiar example is provided by *Mengoli series*: the identity

$$\sum_{n=1}^{N} \frac{1}{n(n+1)} = \sum_{n=1}^{N} \left(\frac{1}{n} - \frac{1}{n+1} \right) = 1 - \frac{1}{N+1}$$

grants we have

$$\sum_{n \geq 1} \frac{1}{n(n+1)} = 1.$$

The following generalization is also straightforward:

Lemma 2.

$$\forall k \in \mathbb{N}^+, \qquad \sum_{n \geq 1} \frac{1}{n(n+1) \cdot \ldots \cdot (n+k)} = \frac{1}{k \cdot k!}.$$

In a forthcoming section we will also see a proof not relying on telescoping, but on properties of Euler's Beta function.

The first *technical issue* we meet in this framework is related to the fact that recognizing a contribution of the form $b_i - b_{i+1}$ in the general term of a series is not always easy, just like in the continuous analogue: if the task is to find $\int_a^b f(x)\, dx$, it is not always easy to devise a function g such that $f(x) = g'(x)$. Here there are some non-trivial examples:

Lemma 3.

$$\sum_{n \geq 1} \arctan \left(\frac{1}{1+n+n^2} \right) = \frac{\pi}{4}.$$

Proof. If we use "backwards" the sum/subtraction formulas for the tangent function, we have that

$$\tan(x \pm y) = \frac{\tan(x) \pm \tan(y)}{1 \mp \tan(x)\tan(y)}$$

implies:

$$\arctan \frac{1}{n} - \arctan \frac{1}{n+1} = \arctan \left(\frac{\frac{1}{n} - \frac{1}{n+1}}{1 + \frac{1}{n(n+1)}} \right) = \arctan \left(\frac{1}{1+n+n^2} \right)$$

so the given series is telescopic and it converges to $\arctan(1) = \frac{\pi}{4}$. □

Exercise 4. The sequence of Fibonacci numbers $\{F_n\}_{n\geq 0}$ is defined through $F_0=0, F_1=1$ and $F_{n+2}=F_{n+1}+F_n$ for any $n\geq 0$. Show that:

$$\sum_{n\geq 1}\arctan\left(\frac{(-1)^{n+1}}{F_{n+1}(F_n+F_{n+2})}\right)=\arctan(\sqrt{5}-2).$$

Exercise 5. Show that:

$$\sum_{n\in\mathbb{Z}}\arctan\left(\frac{\sinh 1}{\cosh(2n)}\right)=\frac{\pi}{2},\qquad \sum_{n\geq 1}\arctan\left(\frac{1}{8n^2}\right)=\frac{\pi}{4}-\arctan\tanh\frac{\pi}{4}.$$

Exercise 6. Prove that:

$$\sum_{n=0}^{N}\frac{1}{4^n}\binom{2n}{n}=\frac{N+1}{2^{2N+1}}\binom{2N+2}{N+1},$$

for instance by considering that by De Moivre's formula we have

$$A_n=\frac{1}{4^n}\binom{2n}{n}=\frac{1}{2\pi}\int_{-\pi}^{\pi}\cos^{2n}(x)\,dx$$

since $\cos(x)^{2n}=\frac{1}{4^n}\sum_{j=0}^{2n}\binom{2n}{j}e^{(2n-j)ix}e^{-jix}$ and $\int_{-\pi}^{\pi}e^{kix}\,dx=2\pi\cdot\delta(k)$, so the only non-vanishing contribution is related to the $j=n$ term, and

$$\sum_{n=0}^{N}\frac{1}{4^n}\binom{2n}{n}=\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1-\cos^{2N+2}(x)}{\sin^2(x)}\,dx=(2N+2)\,A_{N+1}=\frac{N+1}{2^{2N+1}}\binom{2N+2}{N+1}$$

follows from the integration by parts formula ($\int\frac{dx}{\sin^2 x}=-\cot x$). Prove also that

$$\sum_{k\geq 0}\binom{2k}{k}\frac{1}{(k+1)4^k}=1$$

by recognizing in the main term a telescopic contribution. Give a probabilistic interpretation to the proved identities, by considering random paths on a infinite grid ($\mathbb{Z}\times\mathbb{Z}$) where only unit movements towards North or East are allowed.

We now outline the first (really) interesting idea, namely **creative telescoping**: even if we are not able to write the main term of a series in the b_i-b_{i+1} form, it is not unlikely there is an accurate approximation of the main term that can be represented in such a telescopic form. By subtracting the accurate telescopic approximation from the main term, the original problem boils down to computing/approximating a series that is likely to converge faster than the original one, and the same approximation-by-telescopic-series trick can be performed again. For instance, we might employ **creative telescoping** for producing very accurate approximations of the series

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

that will be the main character of a forthcoming section.
In particular, for any $n>1$ the term $\frac{1}{n^2}$ is quite close to the telescopic term $\frac{1}{n^2-\frac{1}{4}}$:

$$\frac{1}{n^2-\frac{1}{4}}=\frac{4}{(2n-1)(2n+1)}=2\left(\frac{1}{2n-1}-\frac{1}{2n+1}\right)$$

and we have $\frac{1}{n^2} - \frac{1}{n^2-\frac{1}{4}} = -\frac{1}{(2n-1)n^2(2n+1)}$, so:

$$\begin{aligned}\zeta(2) = 1 + \sum_{n\geq 2}\frac{1}{n^2} &= 1 + \sum_{n\geq 2}\frac{1}{n^2-\frac{1}{4}} - \sum_{n\geq 2}\frac{1}{(2n-1)n^2(2n+1)} \\ &= 1 + \frac{2}{3} - \sum_{n\geq 2}\frac{1}{(2n-1)n^2(2n+1)}\end{aligned}$$

gives us $\zeta(2) < \frac{5}{3}$ (that we will prove to be equivalent to $\pi^2 < 10$), and the magenta "residual series" can be manipulated in the same fashion (by extracting the first term, approximating the main term with a telescopic contribution, considering the residual series) or simply bounded above by:

$$\sum_{n\geq 2}\frac{1}{(2n-1)n^2(2n+1)} < \sum_{n\geq 2}\frac{1}{(2n-1)\left(n-\frac{1}{2}\right)\left(n+\frac{1}{2}\right)(2n+1)} = \frac{3}{2}\zeta(2) - \frac{22}{9}$$

from which the lower bound $\zeta(2) > \frac{74}{45}$ follows. We may notice that the difference between $\frac{74}{45}$ and $\frac{5}{3}$ is already pretty small. In this framework the iteration of *creative telescoping* leads to two interesting consequences: the identity

$$\zeta(2) = \sum_{n\geq 1}\frac{1}{n^2} = \sum_{n\geq 1}\frac{3}{n^2\binom{2n}{n}},$$

providing a remarkable acceleration of the series defining $\zeta(2)$, and **Stirling's inequality**:

$$\left(\frac{n}{e}\right)^n\sqrt{2\pi n}\exp\left(\frac{1}{12n+1}\right) \leq n! \leq \left(\frac{n}{e}\right)^n\sqrt{2\pi n}\exp\left(\frac{1}{12n}\right)$$

further details will be disclosed soon.
It is also possible to employ creative telescoping for proving that:

$$\zeta(3) = \sum_{n\geq 1}\frac{1}{n^3} = \frac{5}{2}\sum_{n\geq 1}\frac{(-1)^{n+1}}{n^3\binom{2n}{n}}$$

a key identity in Apery's proof of $\zeta(3) \notin \mathbb{Q}$.
We may notice that:

$$\frac{1}{n^3} = \frac{1}{(n-1)n(n+1)} + \frac{(-1)}{(n-1)n^3(n+1)}$$

$$\frac{1}{(n-1)n^3(n+1)} = \frac{1}{(n-2)(n-1)n(n+1)(n+2)} + \frac{-2^2}{(n-2)(n-1)n^3(n+1)(n+2)}$$

Continuing on telescoping we get that:

$$\frac{1}{n^3} = \frac{(-1)^m m!^2}{(n-m)\dots n^3\dots(n+m)} + \sum_{j=1}^{m}\frac{(-1)^{j-1}(j-1)!^2}{(n-j)\dots(n+j)}$$

So by setting $m = n-1$:

$$\frac{1}{n^3} = \frac{(-1)^{n-1}(n-1)!^2}{n^2(2n-1)!} + \sum_{j=1}^{n-1}\frac{(-1)^{j-1}(j-1)!^2}{(n-j)\dots(n+j)}$$

The terms of the last series can be managed through partial fraction decomposition:

$$\frac{1}{(n-j)\dots(n+j)} = \frac{1}{(2j)!(n-j)} - \frac{1}{(2j-1)!1!(n-j+1)} + \frac{1}{(2j-2)!2!(n-j+2)} - \dots$$

$$\frac{(n-j-1)!}{(n+j)!}=\sum_{k=0}^{2j}\frac{(-1)^k}{(2j-k)\,k!\,(n-j+k)}=\frac{1}{(2j)!}\sum_{k=0}^{2j}\frac{(-1)^k\binom{2j}{k}}{n-j+k}$$

and since:

$$\sum_{n>j}\sum_{k=0}^{2j}\frac{(-1)^k\binom{2j}{k}}{n-j+k}=\sum_{h=1}^{2j}\frac{(-1)^{h-1}\binom{2j-1}{h-1}}{h}=\int_0^1(1-x)^{2j-1}\,dx=\frac{1}{2j}$$

we get:

$$\zeta(3)=\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}n!^2}{n^4(2n-1)!}+\sum_{j=1}^{+\infty}\sum_{n>j}\frac{(-1)^{j-1}(j-1)!^2}{(n-j)\ldots(n+j)}$$

$$\zeta(3)=\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}n!^2}{n^4(2n-1)!}+\sum_{j=1}^{+\infty}\frac{(-1)^{j-1}j!^2}{2j^3\,(2j)!}=\frac{5}{2}\sum_{n=1}^{+\infty}\frac{(-1)^{n-1}}{n^3\binom{2n}{n}}$$

as wanted.

Exercise 7. Prove that the following identity (about the acceleration of an "almost-geometric" series) holds.

$$\sum_{n\geq 2}\frac{1}{2^n-1}=\frac{1}{4}+\sum_{m\geq 2}\frac{8^m+1}{(2^m-1)2^{m^2+m}}.$$

As proved by Tachiya, this kind of *acceleration tricks* provide a simple way for proving the irrationality of $\sum_{n\geq 1}\frac{1}{q^n+1}$ and $\sum_{n\geq 1}\frac{1}{q^n-1}$ for any $q\in\mathbb{Z}$ such that $|q|\geq 2$.

Creative telescoping can also be used for a humble purpose, like proving the divergence of the harmonic series. By recalling that the n-th *harmonic number* H_n is defined through

$$H_n=\sum_{k=1}^{n}\frac{1}{k}$$

and by recalling that over the interval $(0,1]$ we have:

$$x<2\,\text{arctanh}\left(\frac{x}{2}\right)=\log\left(\frac{1+\frac{x}{2}}{1-\frac{x}{2}}\right)$$

it follows that:

$$H_n<\sum_{k=1}^{n}\log\left(\frac{2k+1}{2k-1}\right)=\log(2n+1).$$

On the other hand $2\,\text{arctanh}\left(\frac{x}{2}\right)-x=O(x^3)$ in a neighbourhood of the origin, and the series $\sum_{k\geq 1}\frac{1}{k^3}=\zeta(3)$ is convergent, so there is an absolute constant C granting $H_n\geq\log(2n+1)-C$ for any $n\geq 1$. In a similar way, by defining the n-th generalized harmonic number $H_n^{(j)}$ through

$$H_n^{(j)}=\sum_{k=1}^{n}\frac{1}{k^j},$$

we may easily check that the sequence $\{a_n\}_{n\geq 1}$ defined by

$$a_n=2\sqrt{n}-H_n^{(1/2)}=2\sqrt{n}-\sum_{k=1}^{n}\frac{1}{\sqrt{k}}$$

is increasing and never exceeds a constant close to $1+\frac{1}{\sqrt{5}}$. About $a_{n+1}\geq a_n$ we have:

$$a_{n+1}-a_n=2\sqrt{n+1}-2\sqrt{n}-\frac{1}{\sqrt{n+1}}=\frac{1}{\frac{1}{2}\left(\sqrt{n}+\sqrt{n+1}\right)}-\frac{1}{\sqrt{n+1}}=\frac{1}{\sqrt{n+1}\left(\sqrt{n}+\sqrt{n+1}\right)^2}>0$$

and

$$a_n = \sum_{m=0}^{n} \frac{1}{\sqrt{m+1}\left(\sqrt{m}+\sqrt{m+1}\right)^2} \xrightarrow[n\to+\infty]{} \sum_{n\geq 0} \frac{1}{\sqrt{n+1}\left(\sqrt{n}+\sqrt{n+1}\right)^2}.$$

The claim then follows from considering that the main term of the last series is well-approximated by the telescopic term

$$\frac{1}{\sqrt{4n+1}} - \frac{1}{\sqrt{4n+5}}$$

for any $n \geq 1$. In similar contexts, by exploiting creative telescoping and the Cauchy-Schwarz inequality we may get surprising results, like the following one:

$$\sum_{k=1}^{n} \frac{1}{n+k} < \sum_{k=1}^{n} \frac{1}{\sqrt{n+k-1}\sqrt{n+k}} \stackrel{CS}{\leq} \sqrt{n\sum_{k=1}^{n}\left(\frac{1}{n+k-1} - \frac{1}{n+k}\right)} = \frac{1}{\sqrt{2}}$$

but the limit of the LHS for $n \to +\infty$ is $\log(2)$, hence $\log(2) \leq \frac{1}{\sqrt{2}}$. In general, by mixing few ingredients among creative telescoping, the Cauchy-Schwarz inequality, convexity arguments and Weierstrass products we may achieve short and elegant proofs of highly non-trivial claims, like:

Lemma 8. The sequence $\{a_n\}_{n\geq 1}$ defined through

$$a_n = \binom{2n}{n} \frac{\sqrt{\pi\left(n+\frac{1}{4}\right)}}{4^n}$$

is increasing and convergent to 1, due to the identity

$$(2n+1)^2(4n+5) - 4(n+1)^2(4n+1) = 1.$$

That implies

$$\frac{\Gamma\left(x+\frac{1}{2}\right)}{\Gamma(x)} \sim \frac{x}{\sqrt{x+1/4}}$$

for any $x > 0$, that is a strengthening of Gautschi's inequality.

Creative telescoping is also a key element in the Wilf and Zeilberger algorithm for the symbolic computation of binomial sums (http://mathworld.wolfram.com/Wilf-ZeilbergerPair.html), further extended by Gosper to the hypergeometric case and by Risch (https://en.wikipedia.org/wiki/Risch_algorithm) to the symbolic computation of elementary antiderivatives.

Exercise 9. Prove by creative telescoping that for any $k \in \{2,3,4,\ldots\}$ we have:

$$\sum_{n\geq 1} \frac{1}{n(n+1)^k} = k - \zeta(2) - \ldots - \zeta(k)$$

where $\zeta(m) = \sum_{n\geq 1} \frac{1}{n^m}$.

The following exercise is particularly *exemplary*, since it stresses some interesting relations among creative telescoping, the Cauchy-Schwarz inequality, the Maclaurin series of $\arcsin^2(z)$, $\zeta(2)$ and Catalan numbers: all these topics will be deeply investigated in the following sections.

Exercise 10. Prove that the value of the series

$$S=\sum_{n\geq 1}\frac{1}{(n+1)\sqrt{n}}$$

is extremely close to $\frac{1}{2}+\frac{\pi}{4}\sqrt{3}$.

Proof. It is not difficult to realize that

$$S\approx\frac{1}{2}+\sqrt{\pi}\sum_{n\geq 2}\frac{\binom{2n}{n}}{4^n(n+1)}=\frac{1}{2}+\frac{3}{4}\sqrt{\pi}$$

since $\frac{1}{4^n}\binom{2n}{n}\approx\frac{1}{\sqrt{\pi n}}$ is a pretty good approximation for any $n\geq 1$ and the generating function for Catalan numbers is fairly well-known. This can be improved by exploting the more accurate

$$\frac{1}{\sqrt{n}}\approx\frac{\sqrt{\pi}}{4^n}\binom{2n}{n}\left(1+\frac{1}{8n}+\frac{1}{128n(n+2)}\right).$$

Creative telescoping provides us a more elementary approach: indeed, $\frac{1}{(n+1)\sqrt{n}}<\frac{2}{\sqrt{n}}-\frac{2}{\sqrt{n+1}}$ immediately proves $S<2$, and the more accurate $\frac{1}{(n+1)\sqrt{n}}\approx\frac{2}{\sqrt{n+\frac{1}{6}}}-\frac{2}{\sqrt{n+\frac{7}{6}}}$ gives $S\approx\frac{1}{2}+2\sqrt{\frac{6}{13}}$. On the other hand we may also combine the approximation through central binomial coefficients with the Cauchy-Schwarz inequality to get an exceptionally simple and very accurate approximation:

$$S\leq\frac{1}{2}+\sqrt{\left(\sum_{n\geq 2}\frac{4^n}{n(n+1)\binom{2n}{n}}\right)\left(\sum_{n\geq 2}\frac{\binom{2n}{n}}{(n+1)4^n}\right)}=\frac{1}{2}+\sqrt{\frac{\pi^2}{4}\cdot\frac{3}{4}}$$

gives $S\approx\frac{1}{2}+\frac{\sqrt{3}}{4}\pi$, whose absolute error is less than $4\cdot 10^{-4}$. □

Before introducing a second tool (the discrete Fourier transform, *DFT*), it might be interesting to consider an application of creative telescoping to the computation of an integral.

Exercise 11. Prove that the following identity holds:

$$\int_0^1\frac{\log(x)\log^2(1-x)}{x}\,dx=-\frac{1}{2}\sum_{n\geq 1}\frac{1}{n^4}=-\frac{\zeta(4)}{2}.$$

Proof. The dilogarithm function is defined, for any $x\in[0,1]$, through:

$$\text{Li}_2(x)=\sum_{n\geq 1}\frac{x^n}{n^2}.$$

We may notice that $\text{Li}_2'(x)=-\frac{\log(1-x)}{x}$, so, by integration by parts:

$$\begin{aligned}\int_0^1\frac{\log(x)\text{Li}_2(x)}{1-x}\,dx &= \int_0^1\log(1-x)\left[\frac{\text{Li}_2(x)}{x}+\log(x)\text{Li}_2'(x)\right]dx\\ &= -\int_0^1\text{Li}_2'(x)\text{Li}_2(x)\,dx-\int_0^1\frac{\log(x)\log^2(1-x)}{x}\,dx\end{aligned}$$

In particular the opposite of our integral equals:

$$\begin{aligned}-\int_0^1\frac{\log(x)\log^2(1-x)}{x}\,dx &= \frac{1}{2}\text{Li}_2^2(1)+\int_0^1\sum_{n\geq 1}\frac{x^n}{n^2}\sum_{k\geq 0}x^k\log(x)\,dx\\ &= \frac{1}{2}\zeta(2)^2-\sum_{n\geq 1}\frac{1}{n^2}\sum_{m>n}\frac{1}{m^2}\end{aligned}$$

where, by symmetry:

$$\sum_{m>n\geq 1}\frac{1}{m^2n^2}=\frac{1}{2}\left[\left(\sum_{n\geq 1}\frac{1}{n^2}\right)^2-\sum_{n\geq 1}\frac{1}{n^4}\right]$$

and the claim readily follows. We may notice that:

$$\frac{\log^2(1-x)}{x}=\sum_{n\geq 0}\frac{2H_n}{(n+1)}x^n,$$

since:

$$-\log(1-x)=\sum_{n\geq 1}\frac{x^n}{n}\qquad\frac{-\log(1-x)}{1-x}=\sum_{n\geq 1}H_n\,x^n\qquad\frac{1}{2}\log^2(1-x)=\sum_{n\geq 1}\frac{H_n}{n+1}\,x^{n+1}$$

By termwise integration (through $\int_0^1(-\log x)x^n\,dx=\frac{1}{(n+1)^2}$) the proved identities lead to:

$$\sum_{n\geq 1}\frac{H_n}{(n+1)^3}=\frac{1}{4}\sum_{n\geq 1}\frac{1}{n^4}=\frac{\zeta(4)}{4}.$$

□

A keen reader might ask why this virtuosity[1] has been included in the creative telescoping section. The reason is the following: in order to make the magic work, we actually do not need the dilogarithm function (a mathematical function *with the sense of humour*, according to D.Zagier) or integration by parts. As a matter of fact:

$$H_n=\sum_{k=1}^{n}\frac{1}{k}=\sum_{m\geq 1}\left(\frac{1}{m}-\frac{1}{m+n}\right)=\sum_{m\geq 1}\frac{n}{m(m+n)}$$

hence it follows that:

$$\begin{aligned}\sum_{n\geq 1}\frac{H_n}{(n+1)^3} &= \sum_{n\geq 1}\left(\frac{H_{n+1}}{(n+1)^3}-\frac{1}{(n+1)^4}\right)=-\zeta(4)+\sum_{n\geq 1}\frac{H_n}{n^3}\\ &= -\zeta(4)+\sum_{n,m\geq 1}\frac{1}{mn^2(m+n)}=-\zeta(4)+\frac{1}{2}\sum_{m,n\geq 1}\left(\frac{1}{mn^2(m+n)}+\frac{1}{m^2n(m+n)}\right)\\ &= -\zeta(4)+\frac{1}{2}\sum_{m,n\geq 1}\frac{1}{m^2n^2}=-\zeta(4)+\frac{1}{2}\zeta(2)^2\end{aligned}$$

and by comparing the last identity to the identities we already know, we get that $\zeta(4)=\frac{2}{5}\zeta(2)^2$.
Some questions might naturally arise at this point: *is it possible, in a similar fashion, to relate the value of* $\zeta(2^{k+1})$ *to the value of* $\zeta(2^k)$*?* Or: *is it possible to find the explicit value of* $\zeta(2)$ *by simply squaring the Taylor series at the origin of the arctangent function?* Answers to such questions are postponed.

We directly introduce the DFT through a problem.

Exercise 12. Let A be a finite set with cardinality ≥ 4. Let P_0 be the set of subsets of A with $3j$ elements, let P_1 be the set of subsets of A with $3k+1$ elements, let P_2 be the set of subsets of A with $3h+2$ elements. Prove that any two numbers among $|P_0|,|P_1|,|P_2|$ differ at most by 1, no matter what $|A|$ is.

[1] It is worth mentioning that just like $\int_0^1\frac{\log(1-x)\log^2(x)}{1-x}\,dx$ is associated to an *Euler sum* with weight 4, namely $\sum_{n\geq 1}\frac{H_n}{n^3}$, the similar integral $\int_0^1\frac{\log(1+x)\log^2(x)}{1+x}\,dx$ is associated to the alternating series $\sum_{n\geq 1}\frac{H_n}{n^3}(-1)^{n+1}$. On the other hand, while the first series is clearly given by the values of the Riemann ζ function at $s=2$ or $s=4$, the alternating series has a much more involved closed form:

$$\sum_{n\geq 1}\frac{H_n}{(n+1)^3}(-1)^{n+1}=\frac{1}{2}\int_0^1\frac{\log(1+x)\log^2(x)}{1+x}=\frac{1}{48}\left[-\pi^4-4\pi^2\log^2(2)+4\log^4(2)+96\,\mathrm{Li}_2\left(\tfrac{1}{2}\right)+84\log(2)\zeta(3)\right]$$

has been proved by De Doelder in 1991. See also *Flajolet and Salvy, Euler sums and contour integral representations.*

The claim appears to be a (more or less) direct generalization of a well-known fact: the number of subsets of $I = \{1, 2, \dots, n\}$ with even/odd cardinality is the same. In that framework, we may consider the map sending $B \subseteq I$ in $B \setminus \{1\}$ when $1 \in B$, and in $B \cup \{1\}$ when $1 \notin B$ ("*if there is* 1*, we remove it, otherwise we insert it*"): such map is an involution and provides a bijection between the subsets with even cardinality and the subsets with odd cardinality. As an alternative, by recalling that in I we have $\binom{n}{k}$ subsets with k elements, we may simply check that

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k = 0$$

holds as a trivial consequence of the binomial Theorem applied to $(1-1)^n$.
In the ternary case we have to compare the sums

$$|P_0| = \sum_{k \equiv 0 \pmod 3} \binom{n}{k}, \qquad |P_1| = \sum_{k \equiv 1 \pmod 3} \binom{n}{k}, \qquad |P_2| = \sum_{k \equiv 2 \pmod 3} \binom{n}{k}$$

and we would like to have a tool allowing us to isolate the contributions given by elements in particular positions (positions given by an arithmetic progression) in a sum. The DFT is *precisely* such a tool.

Lemma 13 (DFT)**.** If $n \geq 2$ is a natural number and $\omega = \exp\left(\frac{2\pi i}{n}\right)$, the function $f : \mathbb{Z} \to \mathbb{C}$ given by

$$f(m) = \frac{1}{n} \sum_{k=0}^{n-1} \omega^{km}$$

is the indicator function of $n\mathbb{Z}$. As a consequence,

$$\chi_h(m) = \frac{1}{n} \sum_{k=0}^{n-1} \omega^{-hk} \omega^{km}$$

is the indicator function of the integers $\equiv h \pmod n$.

The possibility of writing some indicator functions as weighted power sums has deep consequences.
In our case, if we take ω as a primitive third root of unity, we have:

$$|P_0| = \sum_{k=0}^{n} \binom{n}{k} \chi_0(k) = \frac{1}{3} \sum_{k=0}^{n} \binom{n}{k} \left(1^k + \omega^k + \omega^{2k}\right) = \frac{(1+1)^n + (1+\omega)^n + (1+\omega^2)^n}{3}$$

due to the binomial Theorem. Since both $(1+\omega)$ and its conjugate $(1+\omega^2)$ lie on the unit circle, we have that $|P_0|$ is an integer number whose distance from $\frac{2^n}{3}$ is bounded by $\frac{1}{3}$. The reader can easily check the same holds for $|P_1|$ and $|P_2|$ and the claim readily follows. The discrete Fourier transform proves so the reasonable proposition claiming the almost-uniform distribution of the cardinality $\pmod 3$ of subsets of $\{1, \dots, n\}$. *Perfect* uniformity is clearly not possible, since $|P_0| + |P_1| + |P_2| = 2^n$ never belongs to $3\mathbb{Z}$.
Exercise left to the reader: prove the claim of **Exercise 11** by induction on $|A|$.

Exercise 14. Find the explicit value of the series:

$$S = \sum_{n \geq 0} \frac{1}{(3n)!}.$$

We may consider that the complex exponential function

$$e^z = \sum_{n \geq 0} \frac{z^n}{n!}$$

is defined by an everywhere-convergent power series, then apply a ternary DFT to such series and get, like in the previous exercise (here we are manipulating an infinite sum, but there is no issue since e^z is an entire function):

$$\sum_{n\geq 0}\frac{z^{3n}}{(3n)!}=\sum_{n\equiv 0 \pmod 3}\frac{z^n}{n!}=\sum_{n\geq 0}\frac{z^n}{n!}\chi_0(n)=\frac{1}{3}\sum_{n\geq 0}\frac{z^n+(\omega z)^n+(\omega^2 z)^n}{n!}=\frac{1}{3}\left(e^z+e^{\omega z}+e^{\omega^2 z}\right)$$

and since $\omega=\frac{-1+i\sqrt{3}}{2}$ and $\omega^2=\frac{-1-i\sqrt{3}}{2}$, for any $z\in\mathbb{C}$ we have:

$$\sum_{n\geq 0}\frac{z^{3n}}{(3n)!}=\frac{1}{3}\left(e^z+2e^{-z/2}\cos\frac{z\sqrt{3}}{2}\right),$$

that by an evaluation at $z=1$ leads to:

$$S=\frac{e}{3}+\frac{2}{3\sqrt{e}}\cos\frac{\sqrt{3}}{2}.$$

Was the introduction of the complex z variable really necessary? It clearly was not: a viable alternative would have been to just re-write 1 as 1^n in the definition of S. Besides the identity $1=1^n$ being really obvious, the idea of tackling the original problem through such identity and the DFT is not obvious at all: similar situations explain just fine the subtle difference between the adjectives *elementary* and *easy* in a mathematical context. Another famous application of the DFT is related to the *Frobenius coin problem*:

Exercise 15. Given $n\in\mathbb{N}$, let U_n be the number of natural solutions of the (diophantine) equation $a+2b+3c=n$, i.e.

$$U_n=\left|\left\{(a,b,c)\in\mathbb{N}^3 : a+2b+3c=n\right\}\right|.$$

Prove that for any n, U_n equals the closest integer to $\frac{(n+3)^2}{12}$.

This claim will be proved in the section about Analytic Combinatorics, since few elements of Complex Analysis and manipulation of formal power series are required. However we remark that the key idea is the same key idea of Hardy-Littlewood's circle method, a really important tool in Additive Number Theory: for instance, it has been used for proving that any odd natural number large enough is the sum of three primes (Chen's theorem, also known as *ternary Goldbach*). Now we will focus on a typical application of the DFT in Arithmetics, i.e. a proof of a particular case of Dirichlet's Theorem.

Theorem 16 (Dirichlet)**.** If a and b are coprime positive integers, there are infinite prime numbers $\equiv a \pmod b$.

The particular case we are going to study is the proof of the existence of infinite primes of the form $6k+1$. We recall that the infinitude of primes of the form $6k-1$ follows from a minor variation on Euclid's proof of the infinitude of primes:

Let us assume the set of primes of the form $6k-1$ is finite and given by $\{p_1=5,11,17,\ldots,p_M\}=E$.
Let us consider the huge number

$$N=-1+6\prod_{m=1}^{M}p_m.$$

By construction, no element of E divides N. On the other hand, N is a number of the form $6K-1$, hence it must have some prime divisor $\equiv -1 \pmod 6$. Such contradiction leads to the fact that the set of primes of the form $6k-1$ is not finite (aka infinite).

We may notice that a number of the form $6k+1$ *is not* compelled to have a prime divisor of the same form (for instance, $55 = 5 \cdot 11$), so the previous argument is not well-suited for covering the $6k+1$ case, too. [2] We then take a step back and a step forward: we provide an alternative proof of the infinitude of primes, then prove it can be adjusted to prove the existence of infinite primes of the form $6k+1$, too. Let us recall the main statement in Analytic Number Theory:

Theorem 17 (Euler's product for the ζ function)**.** If $\mathcal{P}$ is the set of prime numbers and s is a complex number with real part greater than one,

$$\prod_{p\in\mathcal{P}}\left(1-\frac{1}{p^s}\right)^{-1} = \sum_{n\geq 1}\frac{1}{n^s} = \zeta(s).$$

Since $\left(1-\frac{1}{p^s}\right)^{-1} = 1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \ldots$, such result is just the analytic counterpart of the Fundamental Theorem of Arithmetics, stating that $\mathbb{Z}$ is a UFD.

In such framework the following argument is pretty efficient: *if there were just a finite number of primes, given Euler's product the harmonic series would be convergent. But we know it is not, so there have to be an infinite number of primes.* The Theorem just outlined has an interesting generalization:

Theorem 18 (Euler's product for Dirichlet's L-functions)**.** If $\mathcal{P}$ is the set of prime numbers, s is a complex number with real part greater than one and $\chi(n)$ is a totally multiplicative function (i.e. a function such that $\chi(nm) = \chi(n)\chi(m)$ holds for any couple (n, m) of positive integers), we have:

$$\prod_{p\in\mathcal{P}}\left(1-\frac{\chi(p)}{p^s}\right)^{-1} = \sum_{n\geq 1}\frac{\chi(n)}{n^s} = L(s,\chi).$$

We may consider a simple totally multiplicative function: the function that equals 1 over natural numbers $\equiv 1 \pmod 6$, -1 over natural numbers $\equiv -1 \pmod 6$ and zero otherwise. Such function is the *non-principal (Dirichlet) character* $\pmod 6$. We may notice that:

$$\sum_{n\geq 1}\frac{z^n}{n} = -\log(1-z)$$

for any complex number z having modulus less than one. By applying the DFT with respect to a primitive sixth root of unity:

$$L(1,\chi) = \sum_{k\geq 0}\left(\frac{1}{6k+1}-\frac{1}{6k+5}\right) = \frac{\pi}{2\sqrt{3}}.$$

As an alternative:

$$\begin{aligned}
L(1,\chi) &= \sum_{k\geq 0}\left(\frac{1}{6k+1}-\frac{1}{6k+5}\right) = \sum_{k\geq 0}\int_0^1 (x^{6k}-x^{6k+4})\,dx \\
&= \int_0^1 (1-x^4)\sum_{k\geq 0}x^{6k}\,dx = \int_0^1 \frac{1-x^4}{1-x^6}\,dx \\
&= \int_0^1 \frac{1+x^2}{1+x^2+x^4}\,dx = \frac{\pi}{2\sqrt{3}}
\end{aligned}$$

[2] However, there is a light that never goes out: the infinitude of primes of the form $6k+1$ can be proved in a algebraic fashion by considering cyclotomic polynomials. For instance, every prime divisor of $\Phi_6(3n) = 9n^2 - 3n + 1$ is a number of the form $6k+1$.

Let us assume that prime numbers $\equiv 1 \pmod 6$ are finite and consider Euler's product for $L(s,\chi)$:

$$\begin{aligned}
L(s,\chi) &= \prod_{p\equiv 1 \pmod 6}\left(1-\frac{1}{p^s}\right)^{-1}\prod_{p\equiv -1 \pmod 6}\left(1+\frac{1}{p^s}\right)^{-1}\\
&\leq \prod_{p\equiv 1 \pmod 6}\left(1-\frac{1}{p}\right)^{-1}\prod_{p\equiv -1 \pmod 6}\left(1+\frac{1}{p^s}\right)^{-1}\\
&= C\prod_{p\neq 2,3}\left(1+\frac{1}{p^s}\right)^{-1}\\
&= D\prod_{p}\left(1+\frac{1}{p^s}\right)^{-1}\\
&= D\,\frac{\zeta(2s)}{\zeta(s)}
\end{aligned}$$

for some constant $D>0$. From the divergence of the harmonic series we would have:

$$\lim_{s\to 1^+} L(s,\chi)=0,$$

but we already know that $L(1,\chi)>0$ (we computed its explicit value).
Such contradiction leads to the fact that the set of primes of the form $6k+1$ is infinite.
We underline some points in the proof just outlined:

- we used Euler's product, analytic counterpart of the Fundamental Theorem of Arithmetics, for studying the distribution of primes in the arithmetic progressions $\pmod 6$. It looks highly unlikely that there is just a finite number of primes $\equiv 1 \pmod 6$ and infinite primes $\equiv -1 \pmod 6$, so we just need to show that such awkward "imbalance" does not really occur;
- through the DFT, we may compute the value of $L(1,\chi)$ (with χ being the non-principal character $\pmod 6$) in a explicit way, and check it is a positive number;
- from Euler's product we have that the previous "imbalance" would lead to $L(1,\chi)=0$. It is not so, hence there is no "imbalance".

At last, we mention that both the DFT and the existence of Dirichlet's characters are instances of *Pontryagin's duality* (https://en.wikipedia.org/wiki/Pontryagin_duality). The DFT is also of great importance for algorithms, since it gives methods for the fast multiplication of polynomials (or integers): in such a context it is also known as FFT (Fast Fourier Transform). Summarizing:

- The key idea is to exploit interpolation/extrapolation. A polynomial with degree m is fixed by its values at $m+1$ distinct points. If we assume to have $a(x)$ and $b(x)$ and we need to compute $c(x)=a(x)\cdot b(x)$, we may...
- compute in a explicit way the values of a and b at the 2^n-th roots of unity, then the values of c at such points...
- and compute the coefficients of $c(x)$ through such values. Nicely, both the evaluation process and the extrapolation process are associated with a matrix-vector-product problem, where the involved matrix is Vandermonde's matrix given by the 2^n-th roots of unity;
- the structure of such matrix depends in a simple way from the structure of Vandermonde's matrix given by the 2^{n-1}-th roots of unity, hence the needed matrix-vector-products can be computed through a recursive, *divide et impera* approach, with a significant improvement in computational costs.

For further details, please see http://en.wikipedia.org/wiki/Cooley-Tukey_FFT_algorithm.

The formulas of Koecher, Leshchiner and Bailey-Borwein-Bradley.
We have studied how to use the creative telecoping machinery for producing fast-convergent series representing $\zeta(2)$ or $\zeta(3)$. Three formulas provide a wide generalization of such statement. The first one is due to Koecher (1979):

$$\sum_{n\geq 0}\zeta(2n+3)a^{2n}=\sum_{k\geq 1}\frac{1}{k(k^2-a^2)}=\frac{1}{2}\sum_{k\geq 1}\frac{(-1)^{k+1}}{k^3\binom{2k}{k}}\cdot\frac{5k^2-a^2}{k^2-a^2}\prod_{m=1}^{k-1}\left(1-\frac{a^2}{m^2}\right),$$

the second one is due to Leschiner (1981):

$$\sum_{n\geq 0}\left(1-\frac{1}{2^{2n+1}}\right)\zeta(2n+2)a^{2n}=\sum_{n\geq 1}\frac{(-1)^{n+1}}{n^2-a^2}=\frac{1}{2}\sum_{k\geq 1}\frac{1}{k^2\binom{2k}{k}}\cdot\frac{3k^2+a^2}{k^2-a^2}\prod_{m=1}^{k-1}\left(1-\frac{a^2}{m^2}\right),$$

the third one is due to Bailey-Borwein-Bradley (2006):

$$\sum_{n\geq 0}\zeta(2n+2)a^{2n}=\sum_{k\geq 1}\frac{1}{k^2-a^2}=3\sum_{k\geq 1}\frac{1}{\binom{2k}{k}(k^2-a^2)}\prod_{m=1}^{k-1}\frac{m^2-4a^2}{m^2-a^2}.$$

They hold for any $a\in(-1,1)$: by comparing the coefficients of a^h in the LHS/RHS one gets that $\zeta(m)$, for any $m\geq 2$, can be represented as a fast-convergent series involving central binomial coefficients and generalized harmonic numbers. It is straightforward to recover the well-known results

$$\zeta(2)=3\sum_{n\geq 1}\frac{1}{n^2\binom{2n}{n}},\qquad \zeta(3)=\frac{5}{2}\sum_{n\geq 1}\frac{(-1)^{n+1}}{n^3\binom{2n}{n}}$$

together with the lesser known results

$$\zeta(4)=\frac{36}{17}\sum_{k\geq 1}\frac{1}{k^4\binom{2k}{k}},\qquad G=\sum_{k\geq 0}\frac{(-1)^k}{(2k+1)^2}=\frac{1}{2}\sum_{n\geq 0}\frac{2^n}{(2n+1)\binom{2n}{n}}\sum_{k=0}^{n}\frac{1}{2k+1}.$$

The last identity can also be proved by computing integrals involving the $\arcsin^2(x)$ function or by computing the *binomial transform* of $\frac{1}{(2k+1)^2}$.

Exercise 19. Prove the following identity:

$$\sum_{n\geq 2}\operatorname{arctanh}\left(\frac{1}{n^3}\right)=\frac{\log(3)-\log(2)}{2},$$

trivially leading to $\zeta(3)<1+\frac{1}{2}\log\frac{3}{2}$.

Exercise 20. By exploiting Euler's product prove that

$$\forall s>1,\qquad \sum_{m,n\geq 1}\frac{1}{\operatorname{lcm}(m,n)^s}=\frac{\zeta(s)^3}{\zeta(2s)}.$$

Proof. For any $M\in\mathbb{N}^+$ of the form $M=p_1^{\alpha_1}\cdots p_k^{\alpha_k}$, the number of solutions of $\operatorname{lcm}(n,m)=M$ is given by $(2\alpha_1+1)\cdots(2\alpha_k+1)$. It follows that the given series equals

$$\sum_{M\geq 1}\frac{1}{M^s}\prod_{p|M}(2\nu_p(M)+1)$$

and since $M \mapsto \prod_{p|M}(2\nu_p(M)+1)$ clearly is a multiplicative function, by Euler's product

$$\sum_{m,n\geq 1}\frac{1}{\text{lcm}(m,n)^s} = \prod_{p\in\mathcal{P}}\left(1+\frac{3}{p^s}+\frac{5}{p^{2s}}+\frac{7}{p^{3s}}+\ldots\right) = \prod_{p\in\mathcal{P}}\frac{p^s(p^s+1)}{(p^s-1)^2} = \prod_{p\in\mathcal{P}}\frac{1-\frac{1}{p^{2s}}}{\left(1-\frac{1}{p^s}\right)^3} = \frac{\zeta(s)^3}{\zeta(2s)}.$$

□

Exercise 21. Prove the following identity:

$$\int_{-1}^{1}\frac{\log^2(1-x)}{\sqrt{1-x^2}} = \pi\sum_{n\geq 1}\frac{\binom{2n}{n}H_{2n-1}}{n4^n} = \frac{\pi^3}{3}+\pi\log^2(2).$$

Exercise 22. The analytic continuation for the Riemann ζ function to the region $\text{Re}(s)>0$ gives us the identity

$$\zeta\left(\tfrac{1}{2}\right) = -2+\sum_{k\geq 1}\frac{1}{\sqrt{k}(\sqrt{k}+\sqrt{k+1})^2}.$$

Use creative telescoping to show that:

$$\begin{aligned}\zeta\left(\tfrac{1}{2}\right) &= -\frac{3}{2}+\frac{1}{2}\sum_{k\geq 1}\frac{1}{\sqrt{k}\sqrt{k+1}(\sqrt{k}+\sqrt{k+1})^3}\\ &= -\frac{35}{24}-\frac{7}{96}\sum_{k\geq 1}\frac{1}{k\sqrt{k}(k+1)\sqrt{k+1}(\sqrt{k}+\sqrt{k+1})^3}+\frac{1}{96}\sum_{k\geq 1}\frac{1}{k\sqrt{k}(k+1)\sqrt{k+1}(\sqrt{k}+\sqrt{k+1})^7}.\end{aligned}$$

Exercise 23. Find a rational approximation of $\prod_{r\geq 1}\left(1+\frac{1}{2^r}\right)$ within $\frac{1}{100}$ from the exact value.

Proof. We may check that for any $x\in(0,1)$ we have

$$1+x > \left(\frac{1+\frac{2x}{3}}{1+\frac{x}{3}}\right)^3$$

hence it follows that

$$\prod_{r\geq 1}\left(1+\frac{1}{2^r}\right) > \frac{3}{2}\cdot\frac{5}{4}\cdot\prod_{r\geq 3}\left(\frac{1+\frac{1}{3\cdot 2^{r-1}}}{1+\frac{1}{3\cdot 2^r}}\right)^3 = \frac{15}{8}\left(1+\frac{1}{12}\right)^3 > \frac{50}{21}.$$

Similarly

$$1+x < \frac{1+2x+\frac{4}{3}x^2+\frac{8}{7}x^3}{1+x+\frac{1}{3}x^2+\frac{1}{7}x^3}$$

implies

$$\prod_{r\geq 1}\left(1+\frac{1}{2^r}\right) < \frac{15}{8}\left[1+2x+\frac{4}{3}x^2+\frac{8}{7}x^3\right]_{x=1/8} < \frac{74}{31}.$$

The difference between the upper bound and the lower bound is already less than $7\cdot 10^{-3}$. □

Exercise 24. Let $\Omega(n)$ be the function returning the number of prime factors of n, counted according to their multiplicity, such that, for instance, $\Omega(24) = \Omega(2^3 \cdot 3) = 3 + 1 = 4$. Prove that

$$\sum_{\substack{n\geq 1\\ \Omega(n)\text{ is even}}} \frac{1}{n^2} = \frac{7\pi^2}{60}.$$

Proof. By Euler's product we have

$$\prod_{p\in\mathcal{P}}\left(1-\frac{1}{p^2}\right)^{-1} = \sum_{n\geq 1}\frac{1}{n^2}, \qquad \prod_{p\in\mathcal{P}}\left(1+\frac{1}{p^2}\right)^{-1} = \sum_{n\geq 1}\frac{(-1)^{\Omega(n)}}{n^2},$$

hence by averaging/DFT we have

$$\sum_{\substack{n\geq 1\\ \Omega(n)\text{ is even}}} \frac{1}{n^2} = \frac{1}{2}\left[\zeta(2)+\frac{\zeta(4)}{\zeta(2)}\right] = \frac{7\pi^2}{60}.$$

□

2 Convolutions and ballot problems

We start this section by recalling a well-known identity:

Lemma 25.

$$\sum_{k=0}^{n}\binom{n}{k}^2 = \binom{2n}{n}.$$

Proof. The first proof we provide is based on a **double counting** argument. Let us assume to have a parliament with n politicians in the left wing and n politicians in the right wing, and to be asked to count how many committees with n politicians we may have. It is pretty clear such number is given by $\binom{2n}{n}$, i.e. the number of subsets with n elements in a set with $2n$ elements. On the other hand, we may count such committees according to the number of politicians from the left wing ($k \in [0, n]$) in them. There are $\binom{n}{k}$ ways for choosing k politicians of the left wing from the n politicians we have. If in a committee there are k politicians from the left wing, there are $n-k$ politicians from the right wing, and we have $\binom{n}{n-k} = \binom{n}{k}$ for selecting them. It follows that:

$$\binom{2n}{n} = \sum_{k=0}^{n}\binom{n}{k}\binom{n}{n-k} = \sum_{k=0}^{n}\binom{n}{k}^2$$

as wanted. The second proof is based on the fact that

$$(f * g)(n) \stackrel{\text{def}}{=} \sum_{k=0}^{n} f(k)\cdot g(n-k)$$

is the **convolution** between f and g.
Since $\binom{n}{k}$ is the coefficient of x^k in the Taylor series of $(1+x)^n$ at the origin[3],

$$(1+x)^n = \sum_{k=0}^{n}\binom{n}{k}x^k \qquad \Longrightarrow \qquad [x^k](1+x)^n = \binom{n}{k}$$

[3]The notation $[x^k]\,f(x)$ stands for the coefficient of x^k in the Taylor/Laurent series of $f(x)$ at the origin.

implies that:

$$\sum_{k=0}^{n}\binom{n}{k}\binom{n}{n-k}=\sum_{k=0}^{n}[x^k](1+x)^n\cdot[x^{n-k}](1+x)^n=[x^n]\left[(1+x)^n\cdot(1+x)^n\right]=[x^n](1+x)^{2n}=\binom{2n}{n}.$$

□

The second approach leads to a nice generalization of the first identity in the current section:

Theorem 26 (Vandermonde's identity)**.**

$$\sum_{j+k=n}\binom{a}{j}\binom{b}{k}=\binom{a+b}{n}.$$

In the introduced convolution context the last identity simply follows from the trivial $(1+x)^a\cdot(1+x)^b=(1+x)^{a+b}$. We may notice that

$$\left(\sum_{n\geq 0}c(n)\,x^n\right)\cdot\left(\sum_{n\geq 0}d(n)\,x^n\right)=\sum_{n\geq 0}(c*d)(n)\,x^n$$

is the **Cauchy product** between two power series. The interplay between analytic and combinatorial arguments allows us to prove interesting things. For instance we may consider the function $f(x)=\sqrt{1-x}=(1-x)^{1/2}$, analytic in a neighbourhood of the origin. It is not difficult to compute its Taylor series by the extended binomial theorem. Moreover $f(x)^2=(1-x)$ has a trivial Taylor series, hence by defining $a(n)$ as the coefficient of x^n in the Taylor series of $f(x)$, $(a*a)(n)$ always takes values in $\{-1,0,1\}$.

$$\begin{aligned}(1-x)^{1/2} &= \sum_{n\geq 0}\binom{1/2}{n}(-1)^n x^n=\sum_{n\geq 0}(-1)^n\frac{\frac{1}{2}\cdot\left(\frac{1}{2}-1\right)\cdot\ldots\cdot\left(\frac{1}{2}-n+1\right)}{n!}\,x^n\\ &= \sum_{n\geq 0}(-1)^n\frac{1\cdot(1-2)\cdot\ldots(1-2n+2)}{2^n n!}\,x^n\\ &= 1-\sum_{n\geq 1}\frac{(2n-1)!!}{(2n-1)\cdot(2n)!!}\,x^n=1-\sum_{n\geq 1}\frac{(2n)!}{(2n-1)\cdot(2n)!!^2}\,x^n\\ &= 1-\sum_{n\geq 1}\binom{2n}{n}\frac{x^n}{4^n(2n-1)}\end{aligned}$$

By differentiating with respect to x,

$$\frac{1}{\sqrt{1-x}}=\sum_{n\geq 0}\binom{2n}{n}\frac{x^n}{4^n}$$

follows, and since $\frac{1}{1-x}=1+x+x^2+x^3+\ldots$, if we set $a(n)=\frac{1}{4^n}\binom{2n}{n}$ we have $(a*a)(n)=1$, i.e.:

Lemma 27.

$$\sum_{k=0}^{n}\binom{2k}{k}\binom{2n-2k}{n-k}=4^n.$$

We may also notice that

$$\frac{1}{1-x}\sum_{n\geq 0}a(n)x^n=\sum_{n\geq 0}(a*1)(n)\,x^n=\sum_{n\geq 0}\left(\sum_{k=0}^{n}a(k)\right)x^n$$

where the LHS equals

$$\frac{1}{(1-x)\sqrt{1-x}} = 2\frac{d}{dx}\left(\frac{1}{\sqrt{1-x}}\right) = 2\sum_{n\geq 1}\binom{2n}{n}\frac{nx^{n-1}}{4^n} = \sum_{n\geq 0}\binom{2n+2}{n+1}\frac{2n+2}{4^{n+1}}x^n.$$

By comparing the last two RHSs we have an alternative proof of an identity claimed by **Exercise 6**:

$$\sum_{n=0}^{N}\binom{2n}{n}\frac{1}{4^n} = \binom{2N+2}{N+1}\frac{N+1}{2^{2N+1}}.$$

Exercise 28 (Stars and bars)**.** Prove that for any $k \in \mathbb{N}$ we have:

$$\frac{1}{(1-x)^{k+1}} = \sum_{n\geq 0}\binom{n+k}{k}x^n.$$

Proof. We may tackle this question both in a combinatorial and in an analytic way. The coefficient of x^n in the product of $(k+1)$ terms of the form $(1+x+x^2+\ldots)$ is given by the number of ways for writing n as the sum of $k+1$ natural numbers. By *stars and bars* we know the number of ways for writing n as the sum of $k+1$ *positive* natural numbers is $\binom{n-1}{k}$ and it is not difficult to finish from there. As an alternative, we may proceed by induction on k. The claim is trivial in the $k=0$ case, and since

$$\frac{1}{(1-x)^{k+2}} = \frac{1}{1-x}\cdot\frac{1}{(1-x)^{k+1}} = \sum_{n\geq 0}\left[\binom{n+k}{k} * 1\right]x^n$$

the inductive step follows from the *hockey stick identity*

$$\sum_{j=0}^{n}\binom{k+j}{k} = \binom{n+k+1}{k+1}.$$

□

Exercise 29. Given the sequences of Fibonacci and Lucas numbers $\{F_n\}_{n\geq 0}$ and $\{L_n\}_{n\geq 0}$, prove the following convolution identity:

$$\sum_{k=0}^{n} F_k F_{n-k} = \frac{nL_n - F_n}{5}.$$

Proof. Since Fibonacci numbers fulfill the relation $F_{n+2} = F_{n+1} + F_n$, by defining their **generating function** as

$$f(x) = \sum_{n\geq 0} F_n x^n$$

we have that $(1-x-x^2)\cdot f(x)$ is a linear polynomial (a similar idea leads to the Berkekamp-Massey algorithm). On the other hand, if

$$f(x) = \sum_{n\geq 0} F_n x^n = \frac{ax+b}{1-x-x^2}$$

$b=0$ has to hold to grant $f(0) = F_0 = 0$ and $a=1$ has to hold to grant $f'(0) = F_1 = 1$. It follows that:

$$f(x) = \frac{x}{1-x-x^2} = \frac{1}{\sqrt{5}}\left(\frac{1}{1-\varphi x} - \frac{1}{1-\bar{\varphi}x}\right), \qquad \varphi = \frac{1+\sqrt{5}}{2},\ \bar{\varphi} = \frac{1-\sqrt{5}}{2}$$

and by computing the Taylor series (that is a geometric series) of $\frac{1}{1-\varphi x}$ and $\frac{1}{1-\bar{\varphi}x}$ we immediately recover Binet's formula

$$F_n = \frac{\varphi^n - \bar{\varphi}^n}{\sqrt{5}}.$$

The identity $L_n = \varphi^n + \bar{\varphi}^n$ has a similar proof. By starting the convolution machinery:

$$\sum_{k=0}^{n} F_k F_{n-k} = [x^n]\left(\frac{x}{1-x-x^2}\right)^2 = \frac{1}{5}[x^n]\left(\frac{1}{(1-\varphi x)^2} + \frac{1}{(1-\bar{\varphi}x)^2} - \frac{2}{(1-\varphi x)(1-\bar{\varphi}x)}\right)$$

and the claim follows from simple fraction decomposition. To find a combinatorial proof is an exercise left to the reader: we recall that Fibonacci numbers are related to subsets of $\{1, 2, \ldots, n\}$ without consecutive elements. □

A note in mathematical folklore: Alon's **Combinatorial Nullstellensatz** has further tightened the interplay between combinatorial arguments and generating functions arguments. We invite the reader to delve into the bibliography to find a generalization of Cauchy-Davenport's theorem, once known as Kneser's conjecture, now known as Da Silva-Hamidoune's Theorem:

Theorem 30 (Da Silva, Hamidoune)**.** If $A \subseteq \mathbb{F}_p$ and $A \oplus A \stackrel{\text{def}}{=} \{a + a' : a, a' \in A, a \neq a'\}$, we have:

$$|A \oplus A| \geq \min(p, 2|A| - 3).$$

The convolution machinery applies very well to another kind of coefficients given by **Catalan numbers**. We introduce them in a combinatorial fashion, assuming to have two people involved in a ballot and to check the votes one by one.

Theorem 31 (Bertrand's ballot problem)**.** If the winning candidate gets A votes and the loser gets B votes (so we are clearly assuming $A > B$), the probability that the winning candidate had the lead during the whole scrutiny equals:

$$\frac{A-B}{A+B}$$

Proof. The final outcome is so simple due to a slick symmetry argument, applied in a *double counting* framework: instead of trying to understand what happens or might happen once a single vote is checked, it is more effective to consider which orderings of the votes favour A or not. Let us consider just the first vote: if it is a vote for B, at some point of the scrutiny there must be a tie, since the winning candidate is A. If the first vote is for A and at some point of the scutiny there is a tie, by switching the votes for A and for B till the tie we return in the previous situation. It is pretty clear that the probability the first vote is a vote for B is $\frac{B}{A+B}$. It follows that the probability of a tie happening during the scrutiny is $\frac{2B}{A+B}$, and the probability that A always leads is:

$$1 - \frac{2B}{A+B} = \frac{A-B}{A+B}.$$

□

Theorem 32 (Catalan numbers)**.** The number of strings made by n characters 0 and n characters 1, with the further property that no initial substring has more 1s than 0s, is:

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

Proof. Any string with the given property can be associated (in a bijective way) with a path on a $n \times n$ grid, starting in the bottom left corner and ending in the upper right corner, made by unit steps towards East (for each 1 character) or North (for each 0 character) and never crossing the SW-NE diagonal (this translates the substrings constraint). These paths can be associated in a bijective way with ballots that end in a tie, in which at every moment of the scrutiny the votes for B are $\leq$ than the votes for A. If a *deus ex machina* adds an extra vote for A *before* the scrutiny begins, we have a situation in which A gets $n+1$ votes, B gets n votes and A is always ahead of B. There are $\binom{2n+1}{n} = \binom{2n+1}{n+1}$ possible scrutinies in which A gets $n+1$ votes and B gets n votes: by the previous result (Bertrand's ballot problem) the number of the wanted strings is given by:

$$\frac{(n+1)-n}{(n+1)+n}\binom{2n+1}{n} = \frac{1}{2n+1}\binom{2n+1}{n} = \frac{1}{n+1}\binom{2n}{n}.$$

For a slightly different perspective on the same subject, the reader is invited to have a look at Josef Rukavicka's "third proof" on the Wikipedia page about Catalan numbers. □

Theorem 33. Given two natural numbers a and b with $a \geq b$, the following identity holds:

$$\sum_{k=0}^{b} \frac{1}{k+1}\binom{2k}{k}\frac{a-b}{a+b-2k}\binom{a+b-2k}{b-k} = \frac{1+a-b}{1+a+b}\binom{a+b+1}{b}.$$

Proof. It is enough to count scrutinies for a ballot between two candidates A and B, with A getting a votes, B getting b votes and A being always ahead of B, without excluding the chance of a tie at some point. We get the RHS by adding an extra vote for A before the scrutiny begins and mimicking the previous proof. On the other hand we may count such scrutinies according to the last moment in which we have a tie. If the last tie happens when $2k$ votes have been checked, we simply need to assign $a-k$ votes for A and $b-k$ votes for B: what happens before the tie can be accounted through Catalan numbers and what happens next through Bertrand's ballot problem. This leads to the LHS. □

The last identity is a particular case of a remarkable generalization of Vandermonde's identity $\binom{m+n}{r} = \sum_{k=0}^{r}\binom{m}{k}\binom{n}{r-k}$:

Theorem 34 (Rothe-Hagen)**.**

$$\sum_{k=0}^{n} \frac{x}{x+kz}\binom{x+kz}{k}\frac{y}{y+(n-k)z}\binom{y+(n-k)z}{(n-k)} = \frac{x+y}{x+y+nz}\binom{x+y+nz}{n}.$$

Proof. This identity is usually proved through generating functions and that approach is not terribly difficult. We may point that a purely combinatorial proof is also possible, by following the lines of the previous proof. It is enough to slightly modify the constraint *at any point, the votes for B are $\leq$ than the votes for A* by replacing it with something involving the ratio of such votes. This is surprising both for experts and for newbies: Micheal Spivey has written an interesting lecture about it **on his blog**. □

The following problems are equivalent:

Exercise 35 (Balanced parenthesis)**.** How many strings with $2n$ characters over the alphabet $\Sigma = \{(,)\}$ have as many open parenthesis as closed parenthesis, and in every initial substring the number of closed parenthesis is always $\leq$ the number of open parenthesis?

Exercise 36 (Sub-diagonal paths)**.** Let us consider the paths from $(0;0)$ to $(n;n)$ where each step is a unit step towards East or North. How many such paths belong to the region $y \leq x$?

Exercise 37 (Triangulations of a polygon)**.** Given a convex polygon, a **triangulation** of such polygon is a partition in almost-disjoint triangles, with the property that every triangle has its vertices on the boundary of the original polygon. How many triangulations are there for a convex polygon with $n+2$ sides?

Exercise 38 (Complete binary trees)**.** A **tree** is a connected, undirected and acyclic graph. It is said **binary** and **complete** if each vertex has two neighbours (in such a case it is an **inner** node) or no neighbours (in such a case it is a **leaf**). How many complete binary trees with n inner nodes are there?

It is not difficult to prove the above claiming by exhibiting three combinatorial bijections:

$$\text{TP} \longleftrightarrow \text{CBT} \longleftrightarrow \text{BP} \longleftrightarrow \text{SDP}.$$

In each case the answer is given by the **Catalan number**

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

where the sub-diagonal paths interpretation proves the identity

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}$$

in a very straightfoward way. Such identity is a **convolution formula**: if we set $c(x) = \sum_{n\geq 0} C_n x^n$, we have $c(x) = 1 + x \cdot c(x)^2$. By solving such quadratic equation in $c(x)$ we get:

$$c(x) = \frac{1-\sqrt{1-4x}}{2x}$$

(the square root sign is chosen in such a way that $c(x)$ is continuous at the origin) hence the coefficients of the power series $c(x)$ can be computed from the extended binomial theorem, *extended* since it is applied to $(1-x)^\alpha$ with $\alpha = \frac{1}{2} \notin \mathbb{N}$.

Exercise 39. Prove that:

$$\sum_{k=0}^{n}\binom{2n-2k}{n-k}\binom{2k}{k}\frac{1}{2k+1} = \frac{16^n}{(2n+1)\binom{2n}{n}} = 4^n\int_0^{\pi/2}\cos(x)^{2n+1}\,dx.$$

The last identity is not trivial at all, and it has very deep consequences. A possible proof of such identity (that is not the most elementary one: to find an elementary proof is an exercise we leave to the reader) exploits a particular class of orthogonal polynomials. We have not introduced the L^2 space yet, nor the usual techniques for dealing with square-integrable objects, so such "advanced" proof is postponed to the end of this section, in a dedicated box. What we can say through the convolution machinery is that the above identity is related to a Taylor coefficient in the product between $\arcsin(x)$ and its derivative $\frac{1}{\sqrt{1-x^2}}$. By termwise integration it implies:

$$\arcsin^2(x) = \frac{1}{2}\sum_{n\geq 1}\frac{(4x^2)^n}{n^2\binom{2n}{n}}$$

and by evaluating the last identity at $x = \frac{1}{2}$ we get that:

$$\sum_{n\geq 1}\frac{3}{n^2\binom{2n}{n}} = 6\arcsin^2\left(\frac{1}{2}\right) = \frac{\pi^2}{6}$$

where the LHS equals $\zeta(2)$ by creative telescoping, as seen in the previous section, where we proved $\zeta(4) = \frac{2}{5}\zeta(2)^2$ too. It follows that:

$$\zeta(2)=\sum_{n\geq 1}\frac{1}{n^2}=\frac{\pi^2}{6},\qquad \zeta(4)=\sum_{n\geq 1}\frac{1}{n^4}=\frac{\pi^4}{90}.$$

We have just solved *Basel problem* through a very creative approach, i.e. by combining creative telescoping with convolution identities for Catalan-ish numbers. Plenty of other approaches are presented in a forthcoming section.

Theorem 40 (Euler's acceleration technique)**.**

$$\sum_{n\geq 0}(-1)^n a_n=\sum_{n\geq 0}(-1)^n\frac{\Delta^n a_0}{2^{n+1}},\qquad \Delta^n a_0=\sum_{k=0}^{n}(-1)^k\binom{n}{k}a_{n-k}.$$

This identity is simple to prove and it is really important in series manipulation and numerical computation: for instance, it is the core of Van Wijngaarden's algorithm for the numerical evaluating of series with alternating signs. Let us study a consequence of Euler's acceleration technique, applied to:

$$\sum_{n\geq 0}\frac{(-1)^n}{2n+1}=\sum_{n\geq 0}\int_0^1(-1)^n x^{2n}\,dx=\int_0^1\frac{dx}{1+x^2}=\arctan(1)=\frac{\pi}{4}$$

We may compute $\Delta^n a_0$ in a explicit way:

$$\begin{aligned}\Delta^n a_0 &= \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}\frac{1}{2k+1}=(-1)^n\int_0^1\sum_{k=0}^{n}\binom{n}{k}(-x^2)^k\,dx\\ &= (-1)^n\int_0^1(1-x^2)^n\,dx=(-1)^n\frac{4^n}{(2n+1)\binom{2n}{n}}\end{aligned}$$

and derive that:

$$\pi=\sum_{n\geq 0}\frac{2^{n+1}}{(2n+1)\binom{2n}{n}}$$

where the main term of the last series behaves like $\frac{1}{2^n}\sqrt{\frac{\pi}{n}}$ for $n\to+\infty$, with a significative boost for the convergence speed of the original series [4]. About the series defining $\zeta(2)$, Euler's acceleration technique leads to:

$$\sum_{n\geq 1}\frac{H_n}{n2^n}=\frac{\zeta(2)}{2}=\frac{\pi^2}{12}.$$

By recalling $H_n=\sum_{m\geq 1}\frac{n}{m(m+n)}$, the last identity turns out to be equivalent to

$$\frac{\pi^2}{12}=\int_0^1\frac{-\log(1-x^2)}{x}\,dx,$$

that is trivial by applying termwise integration to the Taylor series at the origin of $-\frac{\log(1-x^2)}{x}$.

[4] By applying Euler's acceleration technique to the Taylor series of the arctangent function we get something equivalent to the functional identity

$$\arctan x=\arcsin\left(\frac{x}{\sqrt{1+x^2}}\right).$$

A convolution involving the Riemann ζ function.

We now present a result about an ubiquitous **Euler sum**:

Theorem 41. For any $q \in \{3,4,5,\ldots\}$ the following identity holds:

$$\sum_{m=1}^{\infty}\frac{H_m}{m^q} = \frac{q+2}{2}\zeta(q+1) - \frac{1}{2}\sum_{j=1}^{q-2}\zeta(q-j)\zeta(j+1).$$

Proof.

$$\begin{aligned}
&& &\sum_{j=0}^{k}\zeta(k+2-j)\zeta(j+2)\\
&\text{(expand }\zeta) &=& \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\sum_{j=0}^{k}\frac{1}{m^{k+2-j}n^{j+2}}\\
&\left(\substack{\text{pull out the terms for } m=n \text{ and use the}\\ \text{formula for finite geometric sums on the rest}}\right) &=& (k+1)\zeta(k+4) + \sum_{\substack{m,n=1\\ m\neq n}}^{\infty}\frac{1}{m^2n^2}\frac{\frac{1}{m^{k+1}}-\frac{1}{n^{k+1}}}{\frac{1}{m}-\frac{1}{n}}\\
&\text{(simplify terms)} &=& (k+1)\zeta(k+4) + \sum_{\substack{m,n=1\\ m\neq n}}^{\infty}\frac{1}{nm^{k+2}(n-m)} - \frac{1}{mn^{k+2}(n-m)}\\
&\text{(exploit symmetry)} &=& (k+1)\zeta(k+4) + \sum_{m=1}^{+\infty}\sum_{n=m+1}^{+\infty}\frac{1}{nm^{k+2}(n-m)} - \frac{1}{mn^{k+2}(n-m)}\\
&(n\mapsto n+m \text{ and switch sums}) &=& (k+1)\zeta(k+4) + 2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{(n+m)m^{k+2}n} - \frac{1}{m(n+m)^{k+2}n}
\end{aligned}$$

By exploiting $\frac{1}{mn} = \frac{1}{n(m+n)} + \frac{1}{m(m+n)}$ we get:

$$(k+1)\zeta(k+4) + 2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left(\frac{1}{m^{k+3}n} - \frac{1}{(m+n)m^{k+3}}\right) - 2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}\left(\frac{1}{m(n+m)^{k+3}} + \frac{1}{n(n+m)^{k+3}}\right)$$

and since $H_m = \sum_{n\geq 1}\left(\frac{1}{n} - \frac{1}{n+m}\right)$, by exploiting the symmetry of $\frac{1}{m(n+m)^{k+3}} + \frac{1}{n(n+m)^{k+3}}$ we get:

$$\begin{aligned}
&& &(k+1)\zeta(k+4) + 2\sum_{m=1}^{\infty}\frac{H_m}{m^{k+3}} - 4\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{1}{n(n+m)^{k+3}}\\
&(m\mapsto m-n) &=& (k+1)\zeta(k+4) + 2\sum_{m=1}^{\infty}\frac{H_m}{m^{k+3}} - 4\sum_{n=1}^{\infty}\sum_{m=n+1}^{\infty}\frac{1}{nm^{k+3}}\\
&\text{(reintroducing terms)} &=& (k+1)\zeta(k+4) + 2\sum_{m=1}^{\infty}\frac{H_m}{m^{k+3}} - 4\sum_{n=1}^{\infty}\sum_{m=n}^{\infty}\frac{1}{nm^{k+3}} + 4\zeta(k+4)\\
&\text{(switching sums)} &=& (k+5)\zeta(k+4) + 2\sum_{m=1}^{\infty}\frac{H_m}{m^{k+3}} - 4\sum_{m=1}^{\infty}\sum_{n=1}^{m}\frac{1}{nm^{k+3}}\\
&&=& (k+5)\zeta(k+4) + 2\sum_{m=1}^{\infty}\frac{H_m}{m^{k+3}} - 4\sum_{m=1}^{\infty}\frac{H_m}{m^{k+3}}\\
&\text{(combining sums)} &=& (k+5)\zeta(k+4) - 2\sum_{m=1}^{\infty}\frac{H_m}{m^{k+3}}
\end{aligned}$$

Letting $q = k+3$ and reindexing $j \mapsto j-1$ yields

$$\sum_{j=1}^{q-2}\zeta(q-j)\zeta(j+1) = (q+2)\zeta(q+1) - 2\sum_{m=1}^{\infty}\frac{H_m}{m^q}$$

and the claim is proved. □

Exercise 42. Do we get something interesting (like an approximated functional identity for the ζ function) from the previous convolution identity, by recalling that

$$H_m = \log(m) + \gamma + \frac{1}{2m} - \frac{1}{12m^2} + \frac{1}{120m^4} - \frac{1}{252m^6} + \ldots$$

and that

$$\sum_{m\geq 1} \frac{\log m}{m^q} = -\zeta'(q) = \sum_{m\geq 1} \frac{1}{m^q} \sum_{d|m} \Lambda(d)$$

?

Exercise 43. By exploiting $H_m = \sum_{n\geq 1}\left(\frac{1}{n} - \frac{1}{n+m}\right)$ and symmetry prove that

$$\sum_{m\geq 1} \frac{H_m}{m^2} = 2\,\zeta(3) = 2\sum_{m\geq 1}\frac{1}{m^3}.$$

Proof.

$$\sum_{m\geq 1}\frac{H_m}{m^2} = \sum_{m\geq 1}\sum_{n\geq 1}\frac{1}{(m+n)mn} = \int_0^1 \sum_{m,n\geq 1}\frac{x^{m+n-1}}{mn}\,dx = \int_0^1 \frac{\log^2(1-x)}{x}\,dx$$

by the substitution $x \mapsto 1-x$ turns into

$$\int_0^1 \frac{\log^2(x)}{1-x}\,dx = \sum_{n\geq 0}\int_0^1 x^n \log^2(x)\,dx = \sum_{n\geq 0}\frac{2}{(n+1)^3}.$$

□

Exercise 44. By generalizing the previous approach prove that

$$\sum_{n\geq 1}\frac{H_n^2}{n(n+1)} = 3\,\zeta(3).$$

Vandermonde's identity and Bessel functions. Bessel functions are important mathematical functions: they are associated with the coefficients of the Fourier series of some inverse trigonometric functions and they arise in the study of the diffusion of waves, like in the vibrating drum problem. Bessel functions of the first kind with integer order can be simply defined by giving their Taylor series at the origin:

$$J_n(z) = \sum_{l\geq 0}\frac{(-1)^l}{2^{2l+n}\,l!(m+l)!}\,z^{2l+n}$$

from which it is trivial that $J_n(z)$ is an entire function, a solution of the differential equation $z^2 f'' + zf' + (z^2 - n^2)f = 0$ and much more. In this paragraph we will see how Vandermonde's identity plays a major role in dealing with the square of a Bessel function of the first kind.

Exercise 45. Prove the identity:

$$J_n^2(z) = \frac{2}{\pi}\int_0^{\frac{\pi}{2}} J_{2n}(2z\cos(\theta))d\theta.$$

Proof. Let us try the brute-force approach. We have:

$$J_n(x) = \sum_{a\geq 0}\frac{(-1)^a}{a!(a+n)!}\left(\frac{x}{2}\right)^{2a+n}$$

hence:

$$J_n^2(x) = \sum_{m\geq 0}\sum_{a+b=m}\frac{(-1)^m(x/2)^{2m+2n}}{a!b!(a+n)!(b+n)!}$$

where:

$$\begin{aligned}\sum_{a+b=m}\frac{1}{a!b!(a+n)!(b+n)!} &= \frac{1}{m!(m+2n)!}\sum_{a+b=m}\binom{m}{a}\binom{m+2n}{b+n}\\ &= \frac{1}{m!(m+2n)!}[x^{m+n}](1+x)^m(1+x)^{m+2n}\\ &= \frac{1}{m!(m+2n)!}\binom{2m+2n}{m+n} \qquad (\clubsuit)\end{aligned}$$

leads to:

$$J_n^2(x) = \sum_{m\geq 0}\frac{(-1)^m(x/2)^{2m+2n}}{m!(m+2n)!}\binom{2m+2n}{m+n}. \qquad (\heartsuit)$$

Since $\frac{2}{\pi}\int_0^{\pi/2}\cos^{2h}(\theta)\,d\theta = \frac{1}{4^h}\binom{2h}{h}$ follows from De Moivre's formula, in order to prove the claim it is enough to expand $J_{2n}(2z\cos\theta)$ as a power series in $2z\cos\theta$, perform terwmwise integration and exploit ($\heartsuit$). The claim is ultimately a consequence of Vandermonde's identity proved in ($\clubsuit$). □

This technique also shows that the Laplace transform (an important tool we will introduce soon) of $J_0^2(x)$ is related to the complete elliptic integral of the first kind (another object we will study in a forthcoming section) through the identity

$$\left(\mathcal{L}\,J_0^2\right)(s) = \frac{2}{\pi s}K\left(-\frac{4}{s^2}\right).$$

Exercise 46. Prove that the inequality

$$\binom{2n}{n} \geq \frac{4^n}{n+1}$$

is a trivial consequence of the Cauchy-Schwarz inequality.

Exercise 47. Prove that if φ is the *golden ratio* $\frac{1+\sqrt{5}}{2}$, we have:

$$6!\cdot\log^2\varphi < 167.$$

Hint: it might be useful to consider the rapidly convergent series $\sum_{n\geq 1}\frac{(-1)^{n+1}}{2n^2\binom{2n}{n}}$.

Exercise 48. Prove that:

$$\sum_{n\geq 0}\left(\frac{1}{(6n+1)^2}+\frac{1}{(6n+5)^2}\right)=\frac{\pi^2}{9}.$$

Exercise 49. Prove that:

$$\int_{0}^{\pi/2}\log\left(1+\frac{1}{2}\sin\theta\right)\frac{d\theta}{\sin\theta}=\frac{5\pi^2}{72}.$$

Exercise 50. Prove that $\log(3)>\frac{346}{315}$ follows from computing/approximating the integral

$$\int_{0}^{1}\frac{x^4(1-x^2)^2}{4-x^2}\,dx.$$

There is another kind of *convolution*, that is known as *multiplicative* convolution or *Dirichlet's* convolution. We say that a function $f:\mathbb{Z}^+\to\mathbb{C}$ is *multiplicative* when $\gcd(a,b)=1$ grants $f(ab)=f(a)\cdot f(b)$. A multiplicative function has to fulfill $f(1)=1$, and since $\mathbb{Z}$ is a UFD the values of a multiplicative function are fixed by the values of such function over prime powers. Common examples of multiplicative functions are the constant 1, the divisor function $d(n)=\sigma_0(n)$ and Euler's totient function $\varphi(n)$. Given two multiplicative functions $f,g:\mathbb{Z}^+\to\mathbb{C}$, their Dirichlet convolution is defined through

$$(f*g)(n)=\sum_{d|n}f(d)\cdot g\left(\frac{n}{d}\right).$$

It is a simple but interesting exercise to prove that the convolution between two multiplicative functions still is a multiplicative function. Just like additive convolutions are related to products of power series, multiplicative convolutions are related with products of Dirichlet series. If we state that

$$L(f,s)=\sum_{n\geq 1}\frac{f(n)}{n^s}$$

is the Dirichlet series associated with f, the following analogue of Cauchy's product holds:

$$L(f*g,s)=\sum_{n\geq 1}\frac{(f*g)(n)}{n^s}=\sum_{n\geq 1}\sum_{d|n}\frac{f(d)\,g\left(\frac{n}{d}\right)}{n^s}=L(f,s)\cdot L(g,s).$$

The main difference between additive and multiplicative convolutions is that in the multiplicative context, given $f(n)$ and $H(n)=\sum_{d|n}h(d)$, we can always find a multiplicative function g such that $f*g=H$, and such function is unique. There is no additive analogue of Möbius inversion formula, allowing us to solve such problem. The extraction process of a coefficient from a given power/Dirichlet series is similar and relies on the residue theorem, applied to the original function multiplied by $\frac{1}{x^h}$ or to the Laplace transform of the original function. Let us see how to use this machinery for solving actual problems.

Exercise 51. Prove that for any $n\in\mathbb{Z}^+$ the following identity holds:

$$n=\sum_{d|n}\varphi(d).$$

Proof. The usual combinatorial proof starts by considering the n-th roots of unity on the unit circle. Every n-th root of unity is a **primitive** d-th root of unity for some $d \mid n$, and the number of primitive d-th roots of unity is exactly $\varphi(d)$, so the claim follows from checking that no overcounting or undercounting occurs. With the multiplicative convolution machinery, we do not have to find a combinatorial interpretation for both sides, we just have to find the Dirichlet series associated with both sides. In equivalent terms, in order to show that $\text{Id} = \varphi * \mathbf{1}$ it is enough to compute:

$$L(\text{Id}, s) = \sum_{n\geq 1} \frac{n}{n^s} = \sum_{n\geq 1} \frac{1}{n^{s-1}} = \zeta(s-1),$$

$$L(\mathbf{1}, s) = \sum_{n\geq 1} \frac{1}{n^s} = \zeta(s)$$

then prove that $L(\varphi, s) = \frac{\zeta(s-1)}{\zeta(s)}$. By Euler's product:

$$L(\varphi, s) = \prod_p \left(1 + \frac{\varphi(p)}{p^s} + \frac{\varphi(p^2)}{p^{2s}} + \frac{\varphi(p^3)}{p^{3s}} + \ldots\right) = \prod_p \frac{p^s - 1}{p^s - p} = \prod_p \frac{1 - \frac{1}{p^s}}{1 - \frac{1}{p^{s-1}}},$$

$$\zeta(s) = L(\mathbf{1}, s) = \prod_p \left(1 + \frac{1}{p^s} + \frac{1}{p^{2s}} + \frac{1}{p^{3s}} + \ldots\right) = \prod_p \left(1 - \frac{1}{p^s}\right)^{-1}$$

and the claim is proved. □

The keen reader might observe the $\frac{1}{\zeta(s)}$ function played an import role in the previous proof, and ask about the multiplicative function associated to such Dirichlet series. Well, by defining $\omega(n)$ as the number of distinct prime factors of n and $\mu(n)$ as

$$\mu(n) = \begin{cases} (-1)^{\omega(n)} & \text{if } n \text{ is square-free} \\ 0 & \text{otherwise} \end{cases}$$

we have that μ (Möbius' function) is a multiplicative function and

$$L(\mu, s) = \sum_{n\geq 1} \frac{\mu(n)}{n^s} = \prod_p \left(1 + \frac{\mu(p)}{p^s}\right) = \frac{1}{\zeta(s)}$$

as wanted. Then the trivial $1 = \zeta(s) \cdot \frac{1}{\zeta(s)}$ leads to the following convolution identity:

$$\sum_{d\mid n} \mu(d) = \begin{cases} 1 & \text{if } n = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Theorem 52 (Möbius inversion formula)**.** If we have

$$f(n) = \sum_{d\mid n} g(d)$$

then

$$g(n) = \sum_{d\mid n} \mu(d) \cdot f\left(\frac{n}{d}\right)$$

holds.

Proof. Let us denote with ε the multiplicative function that equals 1 at $n = 1$ and zero otherwise. Since $f = g * \mathbf{1}$,

$$\mu * g = \mu * (f * \mathbf{1}) = \mu * (\mathbf{1} * f) = (\mu * \mathbf{1}) * f = \varepsilon * f = f$$

as wanted, by just exploiting $a * b = b * a$ and the associativity of $*$. □

Corollary 53.

$$F(n)=\prod_{d\mid n}f(d)\quad\Longrightarrow\quad f(n)=\prod_{d\mid n}F(d)^{\mu(n/d)}.$$

The last identity encodes an algebraic equivalent of the inclusion-exclusion principle. For instance, by denoting through $\Phi_m(x)$ the m-th cyclotomic polynomial (i.e. the minimal polynomial over $\mathbb{Q}$ of a primitive m-th root of unity) we have

$$x^n-1=\prod_{d\mid n}\Phi_d(x)$$

and from Möbius inversion formula it follows that:

$$\Phi_n(x)=\prod_{d\mid n}(x^d-1)^{\mu(n/d)}.$$

By comparing the degrees of the LHS and RHS we also get:

$$\varphi(n)=\sum_{d\mid n}\frac{n}{d}\cdot\mu(d)=n\sum_{d\mid n}\frac{\mu(d)}{d}$$

corresponding to $\varphi=\text{Id}*\mu$. The explicit formula for cyclotomic polynomials has many interesting consequences, for instance:

$$\forall n>1,\qquad \Phi_n(0)=(-1)^{(\mu*\mathbf{1})(n)}=1$$

that also follows from the fact that $\Phi_n(x)$ is a palindromic polynomial (if ξ is a root of Φ_n, ξ^{-1} is a root of Φ_n too). We also have

$$\frac{\Phi_n'(z)}{\Phi_n(z)}=\frac{d}{dz}\log\Phi_n(z)=\sum_{d\mid n}\mu\left(\frac{n}{d}\right)\frac{dx^{d-1}}{x^d-1}$$

hence for any $n>1$:

$$[z^1]\Phi_n(z)=\Phi_n'(0)=\Phi_n(0)\frac{\Phi_n'(0)}{\Phi_n(0)}=\lim_{x\to 0}\sum_{d\mid n}\mu\left(\frac{n}{d}\right)\frac{dx^{d-1}}{x^d-1}=-\mu(n)$$

since only the term $d=1$ may provide a non-zero contribution to the limit. By putting together the following facts:

- $\Phi_n(z)$ is a palindromic polynomial with degree $\varphi(n)$;
- by Vieta's formulas, for a monic polynomial with degree q the sum of the roots equals the opposite of the coefficient of x^{q-1};
- we know where the roots of $\Phi_n(z)$ lie

it follows that $\mu(n)$ can be represented as an exponential sum, too:

$$\mu(n)=\sum_{\substack{1\leq m\leq n\\ \gcd(m,n)=1}}\exp\left(\frac{2\pi i m}{n}\right).$$

This sum is a particular case of Ramanujan sum. Thanks to Srinivasa Ramanujan we also know that the σ_3 function, $\sigma_3(n)=\sum_{d\mid n}d^3$, fulfills at the same time an additive convolution identity (due to the fact that the Eisenstein series $E_4(\tau)$ depends on σ_3) and a multiplicative convolution identity:

$$\sum_{m=1}^{n-1}\sigma_3(m)\,\sigma_3(n-m)=\frac{\sigma_7(n)-\sigma_3(n)}{120},\qquad \sum_{d\mid n}\sigma_3(d)\,\sigma_3\left(\frac{n}{d}\right)=\underbrace{(1*1*\ldots*1*1)}_{8\text{ times}}=\sigma_7(n).$$

Thanks to Giuseppe Melfi and his work on the modular group $\Gamma(3)$ we also know that:

$$\sum_{k=0}^{n}\sigma_1(3k+1)\sigma_1(3n-3k+1)=\frac{1}{9}\sigma_3(3n+2).$$

Exercise 54. Prove that for any $M \in \{3, 4, 5, \ldots\}$ the following identity holds:

$$\sum_{\substack{1 \leq n \leq M \\ \gcd(n,M)=1}} \sin^2\left(\frac{\pi n}{M}\right) = \frac{\varphi(M) - \mu(M)}{2}.$$

Hint: convert the LHS into something depending on the roots of a cyclotomic polynomial, then recall the representation of the Möbius function as an exponential sum.

A proof of the identity presented in Exercise 32.
If we define the sequence of **Legendre polynomials** through **Rodrigues' formula**

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n}(x^2 - 1)^n = \frac{1}{2^n} \sum_{k=0}^{n} \binom{n}{k}^2 (x-1)^{n-k}(x+1)^k$$

we have that these polynomials, as a consequence of integration by parts,
give a orthogonal and complete base of $L^2(-1, 1)$ with respect to the usual inner product:

$$\int_{-1}^{1} P_n(x)\, P_m(x)\, dx = \frac{2\delta(n,m)}{2n+1}.$$

Additionally, their generating function is pretty simple:

$$\frac{1}{\sqrt{1 - 2xt + t^2}} = \sum_{n \geq 0} P_n(x) t^n.$$

In a similar way, the **shifted Legendre polynomials** $\tilde{P}_n(x) = P_n(2x - 1)$
can be defined through Rodrigues' formula

$$\tilde{P}_n(x) = \frac{1}{n!} \frac{d^n}{dx^n}(x - x^2)^n = (-1)^n \sum_{k=0}^{n} \binom{n}{k}\binom{n+k}{k}(-x)^k$$

and they give a complete and orthogonal base of $L^2(0, 1)$ with respect to the usual inner product:

$$\int_{0}^{1} \tilde{P}_n(x) \tilde{P}_m(x)\, dx = \frac{\delta(m,n)}{2n+1}.$$

By integration by parts it follows that:

$$\int_{-1}^{1} x^n P_n(x)\, dx = \frac{1}{2^n} \int_{-1}^{1} (1 - x^2)^n\, dx = \frac{2^{n+1} n!^2}{(2n+1)!}.$$

In particular:

$$\int_{-1}^{1} \frac{dx}{\sqrt{1 - 2x^2 t + x^2 t^2}} = \sum_{n \geq 0} \left(\int_{-1}^{1} x^n P_n(x)\, dx \right) t^n,$$

$$\frac{4 \arcsin \sqrt{\frac{t}{2}}}{\sqrt{t(2-t)}} = \sum_{n \geq 0} \frac{2^{n+1} n!^2}{(2n+1)!} t^n,$$

$$2 \arcsin^2(t) = \sum_{n \geq 1} \frac{(4t^2)^n}{n^2 \binom{2n}{n}}.$$

Exercise 55. Prove that for any $x \in (0,1)$ we have:

$$\frac{1}{\sqrt{1-x}} = \sqrt{2}\sum_{n\geq 0} P_n(2x-1), \qquad -\log(1-x) = 1 + \sum_{n\geq 1}\frac{2n+1}{n(n+1)}P_n(2x-1).$$

The acceleration of the $\zeta(2)$ series from another perspective.
We already mentioned that $\sum_{n\geq 1}\frac{1}{n^2} = \sum_{n\geq 1}\frac{3}{n^2\binom{2n}{n}}$ can be proved by creative telescoping. In this paragraph we prove it is a consequence of a change of variable in a integral, in particular the *tangent half-angle substitution* (sometimes known as *Weierstrass' substitution*, apparently for no reason) $x = 2\arctan\left(\frac{t}{2}\right)$, sending $(0,\pi/2)$ into $(0,1)$ and $\frac{dx}{\sin x}$ in $\frac{dt}{t}$. Let us set

$$I = -\int_0^{\pi/2}\log\left(1-\frac{1}{4}\sin^2 x\right)\frac{dx}{\sin x}.$$

By expanding $-\log\left(1-\frac{1}{4}\sin^2 x\right)$ as a Taylor series in $\sin x$ we have that

$$I = \sum_{n\geq 1}\frac{1}{n4^n}\int_0^{\pi/2}\sin(x)^{2n-1}\,dx = \sum_{n\geq 1}\frac{1}{4n(2n-1)\binom{2n-2}{n-1}} = \frac{1}{2}\sum_{n\geq 1}\frac{1}{n^2\binom{2n}{n}}$$

and by applying the tangent half-angle substitution we get:

$$I = \int_0^1 -\log\left(1-\left(\frac{t}{1+t^2}\right)^2\right)\frac{dt}{t}$$

where the rational function $1-\left(\frac{t}{1+t^2}\right)^2$ can be written in terms of products and ratios of polynomials of the form $1-t^m$. *Eureka*, since:

$$I_m = -\int_0^1\frac{\log(1-t^m)}{t} = \sum_{n\geq 1}\frac{1}{n}\int_0^1 t^{mn-1}\,dt = \sum_{n\geq 1}\frac{1}{mn^2} = \frac{\zeta(2)}{m}$$

implies:

$$I = (I_2 - 2I_4 + I_6) = \frac{1}{6}\zeta(2).$$

Exercise 56. Is it possible to prove the identity

$$\zeta(3) = \sum_{n\geq 1}\frac{1}{n^3} = \frac{5}{2}\sum_{n\geq 1}\frac{(-1)^{n+1}}{n^3\binom{2n}{n}}$$

through a change of variable in a integral?

The Taylor series of $\arcsin^2(x)$ from the Complex Analysis point of view.
The function $f(z) = \sin(z)$ is an entire function of the form $z + o(z)$, hence it gives a conformal map between two neighbourhoods of the origin (this crucial observation is the same leading to the **Buhrmann-Lagrange inversion formula**). In particular,

$$a_{2n-1} = [x^{2n-1}]\frac{\arcsin(x)}{\sqrt{1-x^2}} = \frac{1}{2\pi i}\oint\frac{\arcsin(z)}{z^{2n}\sqrt{1-z^2}}\,dz = \frac{1}{2\pi i}\oint\frac{z}{\sin(z)^{2n}}\,dz$$

and

$$a_{2n-1} = \text{Res}\left(\frac{z}{\sin(z)^{2n}}, z=0\right) = -\text{Res}\left(\int \frac{dz}{\sin(z)^{2n}}, z=0\right).$$

Since $\frac{d}{dz}\cot(z) = \frac{1}{\sin^2 z}$, the last integral can be computed through repeated integration by parts:

$$\int \frac{dz}{\sin(z)^{2n}} = -\sum_{k=0}^{n-1} \frac{\cot(z)}{(2n)\sin(z)^{2n-2k}} \prod_{j=0}^{k}\left(1+\frac{1}{2n-2j-1}\right)$$

and for every $m \geq 1$ we have:

$$\text{Res}\left(\frac{\cot z}{\sin^{2m}(z)}, z=0\right) = \text{Res}\left(\frac{\cos z}{\sin^{2m+1}(z)}, z=0\right) = \text{Res}\left(\frac{1}{z^{2m+1}}, z=0\right) = 0,$$

so there is a single term really contributing to the value of the residue at the origin, and since $\text{Res}(\cot(z), z=0) = 1$,

$$-\text{Res}\left(\int \frac{dz}{\sin(z)^{2n}}, z=0\right) = \frac{1}{2n}\prod_{j=0}^{n}\left(1+\frac{1}{2n-2j-1}\right) = \frac{(2n)!!}{(2n)\cdot(2n-1)!!} = \frac{4^n}{2n\binom{2n}{n}}$$

from which we get:

$$\frac{\arcsin(x)}{\sqrt{1-x^2}} = \sum_{n\geq 1}\frac{4^n}{2n\binom{2n}{n}}x^{2n-1}, \qquad \arcsin^2(x) = \sum_{n\geq 1}\frac{(4x^2)^n}{2n^2\binom{2n}{n}}.$$

Exercise 57. Prove that the value of the rapidly convergent series

$$\frac{13}{8} - \sum_{n\geq 0}\frac{(-1)^n(2n+1)!}{n!(n+2)!4^{2n+3}}$$

is the golden ratio $\frac{1+\sqrt{5}}{2}$.

Exercise 58. Given $f(x) = \sum_{k\geq 0}\frac{x^k}{k!}\sqrt{k}$, prove that $\lim_{x\to+\infty}\frac{f(x)}{e^x\sqrt{x}} = 1$. Hint: compute the series of $f(x)^2$ and exploit the inequality $2z(1-z) \leq \sqrt{z(1-z)} \leq \frac{1}{2}$, holding for any $z \in [0,1]$.

Exercise 59. Prove that the series

$$\sum_{k\geq 0}\frac{(-1)^k}{2k+1}\sum_{m=0}^{\lfloor k/2\rfloor}\binom{2m}{m}\frac{(-1)^m}{4^m}$$

is convergent to $\frac{1}{\sqrt{2}}\log(1+\sqrt{2})$.

Proof.

$$\sum_{m\geq 0}\frac{(-1)^m}{4^m}\binom{2m}{m} = \frac{2}{\pi}\int_0^{\pi/2}\frac{d\theta}{1+\cos^2\theta} = \frac{1}{\sqrt{2}}$$

and Dirichlet's test ensure that the given series is convergent. If k is even (say $k=2n$) we have

$$\sum_{m=0}^{n}\binom{2m}{m}\frac{(-1)^m}{4^m} = [x^n]\frac{1}{(1-x)\sqrt{1+x}}$$

and if k is odd (say $k = 2n + 1$) we have the same identity, where $[x^n]f(x)$ stands for the coefficient of x^n in the Maclaurin series of $f(x)$. In particular the original series can be written as

$$\sum_{n\geq 0} \frac{1}{(2n+1)(2n+2)} \cdot [x^{2n}] \frac{1}{(1-x^2)\sqrt{1+x^2}} = \int_0^1 \frac{1-x}{(1-x^2)\sqrt{1+x^2}}\,dx$$

since $\int_0^1 x^{2n}(1-x)\,dx = \frac{1}{(2n+1)(2n+2)}$. It turns out that the original series is just

$$\int_0^1 \frac{dx}{(1+x)\sqrt{1+x^2}} = \frac{\operatorname{arcsinh}(1)}{\sqrt{2}} = \frac{\log(1+\sqrt{2})}{\sqrt{2}}.$$

□

3 Chebyshev and Legendre polynomials

Lemma 60. For any $n \in \mathbb{N}$ there exists a polynomial $T_n(x) \in \mathbb{Z}[x]$ such that:

$$\cos(n\theta) = T_n(\cos\theta).$$

It is not difficult to prove the claim by induction on n. It is trivial for $n = 0$ ed $n = 1$, and by the cosine addition formulas:

$$\cos((n+2)\theta) + \cos(n\theta) = 2\cos(\theta)\cos((n+1)\theta)$$

such that:

$$T_{n+2}(x) = 2x \cdot T_{n+1}(x) - T_n(x)$$

for any $n \geq 0$. The $T_n(x)$ polynomials are **Chebyshev polynomials of the first kind** and they have many properties, simple to prove:

- (uniform boundedness)

$$\forall x \in [-1,1], \qquad |T_n(x)| \leq 1$$

- (distibution of roots)

$$T_n(x) = 2^{n-1} \prod_{k=1}^{n} \left(x - \cos\frac{(2k-1)\pi}{2n}\right)$$

- (orthogonality)

$$\int_{-1}^{1} \frac{T_n(x)\,T_m(x)}{\sqrt{1-x^2}}\,dx = \frac{\pi}{2}\delta(m,n)(1+\delta(n,0))$$

- (a simple generating function)

$$\frac{1-xt}{1-2xt+t^2} = \sum_{n\geq 0} T_n(x)t^n$$

- (an explicit representation)

$$T_m(x) = \frac{1}{2}\left[(x+i\sqrt{1-x^2})^m + (x-i\sqrt{1-x^2})^m\right].$$

Chebyshev polynomials of the second kind, $U_n(x)$, are similarly defined through $\frac{\sin((n+1)\theta)}{\sin\theta} = U_n(\cos\theta)$: they share with Chebyshev polynomials of the first kind the recurrence relation $U_{n+2}(x) = 2x\,U_{n+1}(x) - U_n(x)$ and similar properties:

- (boundedness)

$$\forall x \in [-1,1], \qquad |U_n(x)| \leq (n+1)$$

- (distribution of roots)

$$U_n(x) = 2^n \prod_{k=1}^{n} \left(x - \cos\frac{k\pi}{n+1}\right)$$

- (orthogonality)

$$\int_{-1}^{1} U_n(x)\,U_m(x)\sqrt{1-x^2}\,dx = \frac{\pi}{2}\,\delta(m,n)$$

- (a simple generating function)

$$\frac{1}{1-2xt+t^2} = \sum_{n\geq 0} U_n(x)t^n$$

- (an explicit representation)

$$U_m(x) = \sum_{r=0}^{\lfloor m/2 \rfloor} (-1)^r \binom{m-r}{r} (2x)^{m-2r}.$$

By combining Vieta's formulas (about the interplay between roots and coefficients of a polynomial) with the explicit form of the roots of $U_n(x)$ or $T_n(x)$, we have that many trigonometric sums or products can be easily evaluated through Chebyshev polynomials.

Lemma 61.

$$\sum_{k=1}^{n-1} \sin^2\frac{\pi k}{n} = \frac{n}{2}, \qquad \sum_{k=1}^{n-1} \frac{1}{\sin^2\frac{\pi k}{n}} = \frac{n^2-1}{3}, \qquad \prod_{k=1}^{n-1} \sin\left(\frac{\pi k}{n}\right) = \frac{2n}{2^n}.$$

The last identity is related to the combinatorial *broken stick problem* and it provides an unexpected way for tackling (through Riemann sums!) the following integral:

Lemma 62.

$$\int_0^{\pi} \log\sin(x)\,dx = -\pi\log 2.$$

The first proof we provide relies on a "hidden symmetry":

$$\begin{aligned}
\int_0^{\pi} \log\sin(x)\,dx &= 2\int_0^{\pi/2} \log\sin(x)\,dx = \int_0^{\pi/2} \log\sin^2(x)\,dx = \int_0^{\pi/2} \log\cos^2(x)\,dx \\
&= \int_0^{\pi/2} \log\left[\sin(x)\cos(x)\right]\,dx = \int_0^{\pi/2} \log\frac{\sin(2x)}{2}\,dx \\
(2x=z) \quad &= -\frac{\pi}{2}\log(2) + \frac{1}{2}\int_0^{\pi} \log\sin(z)\,dz.
\end{aligned}$$

Then we exploit the previous closed form for a trigonometric product:

$$\begin{aligned}
\int_0^{\pi} \log\sin(x)\,dx &= \lim_{n\to+\infty} \frac{\pi}{n}\sum_{k=1}^{n-1} \log\sin\left(\frac{\pi k}{n}\right) = \lim_{n\to+\infty} \frac{\pi}{n}\log\prod_{k=1}^{n-1}\sin\left(\frac{\pi k}{n}\right) \\
&= \lim_{n\to+\infty} \frac{\pi}{n}\log\left(\frac{2n}{2^n}\right) = -\pi\log 2.
\end{aligned}$$

Exercise 63. Prove the following identities:

$$\sum_{k=1}^{2N}\frac{1}{\cos^2\left(\frac{\pi k}{2n+1}\right)}=4N(N+1),\qquad \sum_{k=1}^{N}\frac{1}{1-\cos\left(\frac{\pi k}{N}\right)}=\frac{2N^2+1}{6}.$$

Another famous application of Chebyshev polynomials is related to the determination of the spectrum of tridiagonal Toeplitz matrices. Due to the Laplace expansion and the recurrence relation for Chebyshev polynomials

$$\det\begin{pmatrix}2x & 1 & 0 & \dots & 0\\ 1 & 2x & 1 & \dots & 0\\ 0 & 1 & \ddots & \ddots & 0\\ \vdots & \vdots & \ddots & \ddots & 1\\ 0 & 0 & 0 & 1 & 2x\end{pmatrix}=U_n(x)$$

so the spectrum of the $n\times n$ matrix with C on the diagonal, 1 on the sup- and sub-diagonal and zero anywhere else is given by:

$$\lambda_k=C+2\cos\frac{\pi k}{n+1},\qquad k=1,2,\dots n.$$

These matrices are deeply involved in the numerical solution of differential equations depending on the Laplacian operator and in extensions of the rearrangement inequality, like:

$$|a_1a_2+a_2a_3+\dots+a_{n-1}a_n|\leq\left(a_1^2+\dots+a_n^2\right)\cos^2\frac{\pi}{n+1},$$

that combined with the *shoelace formula* can be used to prove the *isoperimetric inequality* in the polygonal case. Another important (but lesser-known) application of Chebyshev polynomials is the proof of the uniform convergence of the Weierstrass products for the sine and cosine functions, over compact subsets of $\mathbb{R}$:

Theorem 64 (Weierstrass)**.** For any $x\in\mathbb{R}$ the following identities holds

$$\operatorname{sinc}(x)=\prod_{n\geq 1}\left(1-\frac{x^2}{\pi^2n^2}\right),\qquad \cos(x)=\prod_{n\geq 0}\left(1-\frac{4x^2}{(2n+1)^2\pi^2}\right)$$

and the convergence is uniform over any compact $K\subset\mathbb{R}$.

Exercise 65 (Uniform convergence of the Weierstrass product for the cosine function)**.** Let $I=[a,b]\subseteq\mathbb{R}$ and $\{f_n(x)\}_{n\in\mathbb{N}}$ the sequence of real polynomials defined through:

$$f_n(x)=\prod_{j=0}^{n}\left(1-\frac{4\,x^2}{(2j+1)^2\,\pi^2}\right).$$

Prove that on I the sequence of functions $\{f_n(x)\}_{n\in\mathbb{N}}$ is uniformly convergent to $\cos x$.

Proof. The factorization of Chebyshev polynomials of the first and second kind leads to the following identities:

$$\frac{\sin x}{(2n+1)\,\sin\frac{x}{2n+1}}=\prod_{k=1}^{n}\left(1-\frac{\sin^2\frac{x}{2n+1}}{\sin^2\frac{k\pi}{2n+1}}\right),\qquad \cos x=\prod_{j=0}^{n-1}\left(1-\frac{\sin^2\frac{x}{2n}}{\sin^2\frac{(2j+1)\pi}{4n}}\right),$$

holding for every $x \in \mathbb{R}$ and every $n \in \mathbb{Z}^+$.
We may assume without loss of generality that $0 < x < m < n$ holds, with m and n being positive natural numbers. Since for every θ in the interval $\left(0, \frac{\pi}{2}\right)$ we have $\frac{2\theta}{\pi} < \sin\theta < \theta$, it follows that:

$$1 > \prod_{k=m+1}^{n}\left(1-\frac{\sin^2\frac{x}{2n}}{\sin^2\frac{(2k+1)\pi}{4n}}\right) > \prod_{k=m+1}^{n}\left(1-\frac{x^2}{(2k+1)^2}\right) > 1-x^2\sum_{k=m+1}^{n}\frac{1}{(2k+1)^2} > 1-\frac{x^2}{4m},$$

and by defining $H_m(x)$ as

$$H_m(x) = \prod_{j=0}^{m}\left(1-\frac{\sin^2\frac{x}{2n}}{\sin^2\frac{(2j+1)\pi}{4n}}\right),$$

$\cos x$ belongs to the interval:

$$\left(\left(1-\frac{x^2}{4m}\right)H_m(x),\quad H_m(x)\right).$$

By sending n towards $+\infty$ we get that $\cos x$ belongs to the interval:

$$\left[\left(1-\frac{x^2}{4m}\right)\prod_{j=0}^{m}\left(1-\frac{4\,x^2}{(2j+1)^2\pi^2}\right),\ \prod_{j=0}^{m}\left(1-\frac{4\,x^2}{(2j+1)^2\pi^2}\right)\right],$$

so, by sending m towards $+\infty$, the pointwise convergence of the Weierstrass product for the cosine function is proved. Additionally, by the last line it follows that:

$$|f_m(x)-\cos x| \leq |\cos x|\,\frac{4x^2/m}{1-4x^2/m} \leq \frac{4x^2}{m-4x^2},$$

and such inequality proves the uniform convergence. The proof of the uniform convergence (over compact subsets of the real line) of the Weierstrass product for the sine function is analogous. □

Chebyshev polynomials can also be employed to prove the following statement (a first *density* result in Functional Analysis) through an approach due to Lebesgue.

Theorem 66 (Weierstrass approximation Theorem)**.** If $f(x)$ is a continuous function on the interval $[a,b]$, for any $\varepsilon > 0$ there exists a polynomial $p_\varepsilon(x)$ such that:

$$\forall x \in [a,b], \qquad |f(x)-p_\varepsilon(x)| \leq \varepsilon.$$

We may clearly assume $[a,b] = [-1,1]$ without loss of generality. Moreover, any continuous function over a compact interval of the real line is uniformly continuous, so f can be uniformly approximated by a piecewise-linear function of the form

$$g_n(x) = \sum_{k=-n}^{n} c_k\left|x-\frac{k}{n}\right|$$

and it is enough to prove the statement for the function $f(x) = |x|$ on the interval $[-1,1]$. For such a purpose, we may consider the projection of $f(x)$ on the subspace of $L^2(-1,1)$ (equipped with the "Chebyshev" inner product $\langle u(x), v(x)\rangle = \int_{-1}^{1}\frac{u(x)\,v(x)}{\sqrt{1-x^2}}\,dx$) spanned by $T_0(x), T_1(x), \ldots, T_{2N}(x)$. Since

$$\int_{-1}^{1}\frac{|x|\,T_{2k+1}(x)}{\sqrt{1-x^2}}\,dx = 0, \qquad \int_{-1}^{1}\frac{|x|\,T_{2k}(x)}{\sqrt{1-x^2}}\,dx = \int_{0}^{\pi}|\cos\theta|\,\cos(2k\theta)\,d\theta$$

such projection is given by:

$$p_N(x) = \frac{2}{\pi}+\frac{4}{\pi}\sum_{k=1}^{N}\frac{(-1)^{k+1}}{4k^2-1}\,T_{2k}(x)$$

and it can be proved that the maximum difference, in absolute value, between $p_N(x)$ and $|x|$ occurs at $x = 0$ and equals:

$$p_N(0) = \frac{2}{\pi} - \frac{2}{\pi}\sum_{k=1}^{N}\left(\frac{1}{2k-1} - \frac{1}{2k+1}\right) = \frac{4}{\pi}\sum_{k>N}\frac{1}{4k^2-1} = \frac{2}{\pi(2N+1)},$$

so the sequence of polynomials $\{p_N(x)\}_{N\geq 0}$ provides a uniform approximation of $|x|$ as wanted.
As an alternative we may consider:

$$q_N(x) = 1 - \sum_{n=1}^{N}\binom{2n}{n}\frac{(1-x^2)^n}{4^n(2n-1)}$$

from the truncation of the Taylor series at the origin of $\sqrt{1-z}$, evaluated at $z = 1-x^2$. In such a case it is trivial that $||x| - q_N(x)|$ achieves its maximum value at the origin, but the approximation we get this way, according to the degree of the approximating polynomial, is worse than the approximation we got through Chebyshev polynomials, since $|q_N(0)| \approx \frac{1}{\sqrt{\pi N}}$.

The projection technique on $L^2(-1,1)$ (equipped with the non-canonical inner product $\langle u(x), v(x)\rangle = \int_{-1}^{1}\frac{u(x)\,v(x)}{\sqrt{1-x^2}}\,dx$) is also known as **Fourier-Chebyshev series expansion**. About applications, it is important to mention that many functions have a pretty simple Fourier-Chebyshev series expansion:

Lemma 67. For any $x \in (-1,1)$,

$$\begin{aligned} -\log(1-x) &= \log(2) + 2\sum_{n\geq 1}\frac{T_n(x)}{n} \\ \sqrt{1-x^2} &= \frac{2}{\pi} - \frac{4}{\pi}\sum_{n\geq 1}\frac{T_{2n}(x)}{4n^2-1} \\ \arcsin(x) &= \frac{4}{\pi}\sum_{n\geq 1}\frac{T_{2n-1}(x)}{(2n-1)^2}. \end{aligned}$$

We many notice that the last identity provides an interesting way for the explicit evaluation of $\zeta(2)$ and $\zeta(4)$. Since $T_{2n-1}(1) = 1$,

$$\frac{3}{4}\zeta(2) = \sum_{n\geq 1}\frac{1}{(2n-1)^2} = \frac{\pi}{4}\arcsin(1) = \frac{\pi^2}{8}.$$

Additionally, due to orthogonality relations (i.e. by Parseval's theorem),

$$\frac{8}{\pi}\sum_{n\geq 1}\frac{1}{(2n-1)^4} = \int_{-1}^{1}\frac{\arcsin^2(x)}{\sqrt{1-x^2}}\,dx = \int_{-\pi/2}^{\pi/2}\theta^2\,d\theta = \frac{\pi^3}{12}$$

so:

$$\zeta(4) = \frac{16}{15}\sum_{n\geq 1}\frac{1}{(2n-1)^2} = \frac{16}{15}\cdot\frac{\pi^4}{96} = \frac{\pi^4}{90}.$$

We will now study some applications of Chebyshev polynomials in Arithmetics.

Lemma 68. If $q \in \mathbb{Q}$ and $\cos(\pi q) \in \mathbb{Q}$, then

$$\cos(\pi q) \in \left\{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\right\}.$$

Proof. We may assume without loss of generality $q = \frac{a}{b}$ with $a, b \in \mathbb{Z}^+$ and $\gcd(a, b) = 1$.
Let us study the case in which b is odd first. With such assumption:

$$T_b(\cos(\pi q)) = \cos(b\pi q) = \cos(a\pi) = (-1)^a$$

so $\cos(\pi q)$ is a root of $T_b(x) - (-1)^a$. The value at the origin of such polynomial is ± 1 and the leading term is 2^{b-1}: due to the rational root Theorem,

$$\cos(\pi q) \in \mathbb{Q} \quad \Longrightarrow \quad \cos(\pi q) = \pm \frac{1}{2^k}.$$

However, due to the cosine duplication formula, if $\alpha = \cos(\pi q)$ is a rational number, $2\alpha^2 - 1 = \cos(2\pi q)$ is a rational number too. Such observation leads to a proof of the given claim when b is odd. If $\nu_2(b) \geq 1$, it is enough to consider that by the duplication/bisection formulas $\left(\cos(\theta) = \pm\sqrt{\frac{1+\cos\theta}{2}}\right)$ we have:

$$\cos(\pi q) \in \mathbb{Q} \quad \Longrightarrow \quad \cos(2^{\nu_2(b)}\pi q) \in \mathbb{Q} \quad \Longrightarrow \quad \cos(2^{\nu_2(b)}\pi q) \in \left\{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\right\}$$

so $\cos(\pi q) \in \{-1, 0, 1\}$. □

Lemma 69. If $n \geq 3$, the number $\alpha = \cos\left(\frac{2\pi}{n}\right)$ is an algebraic number over $\mathbb{Q}$ with degree $\frac{\varphi(n)}{2}$.
Additionally, the Galois group of its minimal polynomial over $\mathbb{Q}$ is cyclic.

Proof. It is enough to check that the algebraic conjugates of α are given by

$$\alpha_k = \cos\left(\frac{2\pi k}{n}\right), \qquad 1 \leq k < \frac{n}{2}, \;\; \gcd(k, n) = 1.$$

□

We may notice that the previous result is actually just a corollary of this Lemma.

Exercise 70. Prove that

$$\alpha = \cos\frac{2\pi}{19} - \cos\frac{3\pi}{19} - \cos\frac{5\pi}{19}$$

is an algebraic number over $\mathbb{Q}$ with degree 3, i.e. it is a root of a cubic polynomial with integer coefficients.

Legendre polynomials share many properties with Chebyshev polynomials. Our opinion is that the most efficient way for introducing Legendre polynomials is to do that through **Rodrigues formula**:

Definition 71.

$$P_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2 - 1)^n$$

Such definition provides a simple way for proving the following statements about Legendre polynomials:

- (orthogonality)

$$\int_{-1}^{1} P_n(x)\, P_m(x)\, dx = \frac{2\,\delta(m, n)}{2n + 1}$$

- (Legendre differential equation)

$$\frac{d}{dx}\left[(1 - x^2)\frac{d}{dx}P_n(x)\right] + n(n + 1)P_n(x) = 0$$

- (a simple generating function)

$$\frac{1}{\sqrt{1-2xt+t^2}} = \sum_{n\geq 0} P_n(x)t^n$$

- (an interesting convolution formula)

$$U_n(x) = \sum_{l=0}^{n} P_l(x)P_{n-l}(x)$$

- (Bonnet's recursion formula)

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - P_{n-1}(x), \qquad (2n+1)P_n(x) = \frac{d}{dx}\left(P_{n+1}(x) - P_{n-1}(x)\right)$$

- (an explicit representation)

$$P_n(x) = \sum_{k=0}^{n}(-1)^k \binom{n}{k}^2 \left(\frac{1+x}{2}\right)^{n-k} \left(\frac{1-x}{2}\right)^k.$$

It is very practical to introduce **shifted Legendre polynomials** too, defined by $\tilde{P}_n(x) = P_n(2x-1)$.
Due to such affine transform, shifted Legendre polynomials fulfill the following properties:

- (Rodrigues formula)

$$\tilde{P}_n(x) = \frac{1}{n!}\frac{d^n}{dx^n}(x^2-x)^n$$

- (orthogonality)

$$\int_0^1 \tilde{P}_n(x)\,\tilde{P}_m(x)\,dx = \frac{\delta(m,n)}{2n+1}$$

- (a simple generating function)

$$\frac{1}{\sqrt{(t+1)^2-4tx}} = \sum_{n\geq 0} \tilde{P}_n(x)t^n$$

- (Bonnet's recursion formula)

$$(n+1)\tilde{P}_{n+1}(x) = (2n+1)(2x-1)\tilde{P}_n(x) - \tilde{P}_{n-1}(x), \qquad (4n+2)\tilde{P}_n(x) = \frac{d}{dx}\left(\tilde{P}_{n+1}(x) - \tilde{P}_{n-1}(x)\right)$$

- (an explicit representation)

$$P_n(x) = (-1)^n \sum_{k=0}^{n} \binom{n}{k}\binom{n+k}{k}(-x)^k.$$

The explicit representation and orthogonality are really important. Shifted Legendre polynomials give an orthogonal base (with respect to the canonical inner product $\langle f, g\rangle = \int_0^1 f(x)\,g(x)\,dx$) of the space of square-integrable functions over $(0,1)$, and every C^1 function over $(0,1)$ has a representation of the form:

$$f(x) = \sum_{n\geq 0} c_n\,\tilde{P}_n(x), \qquad c_n = (2n+1)\int_0^1 f(x)\,\tilde{P}_n(x)\,dx.$$

We may define a **Fourier-Legendre series expansion** just like we did for the Fourier-Chebyshev series expansion: we just have to change the integration range into $(0,1)$ and the inner product into the canonical one, since the existence of a complete orthogonal base of polynomials is unchanged.
Let us study an application of the Fourier-Legendre series expansion.

Exercise 72. Prove that for any $f \in C^1[-1,1]$ we have

$$2\int_{-1}^{1} f(x)^2\,dx \geq 3\left(\int_{-1}^{1} x\,f(x)\,dx\right)^2 + \left(\int_{-1}^{1} f(x)\,dx\right)^2$$

and find all the functions for which equality occurs.

Proof. We may assume to have

$$f(x) = \sum_{n\geq 0} c_n\, P_n(x)$$

and by the orthogonality relations we have:

$$\int_{-1}^{1} f(x)^2\,dx = 2\sum_{n\geq 0}\frac{c_n^2}{2n+1},\qquad \int_{-1}^{1} f(x)\,dx = 2c_0,\qquad \int_{-1}^{1} P_1(x)\,f(x)\,dx = \frac{2c_1}{3},$$

so the inequality to prove is equivalent to:

$$4\sum_{n\geq 0}\frac{c_n^2}{2n+1} \geq \frac{4c_1^2}{3} + 4c_0^2,$$

that is trivial and holds as an equality iff $c_2 = c_3 = c_4 = \ldots = 0$, i.e. iff $f(x)$ is a linear polynomial of the form $ax + b$. □

Like in the Chebyshev case, the Fourier-Legendre series expansion of many functions can be simply derived by manipulating the generating function for our sequence of polynomials. For any $x \in (0, 1)$ we have, for instance:

$$\frac{1}{\sqrt{1-x}} = \sqrt{2}\sum_{n\geq 0} P_n(x),\qquad -\log(x) = 1 + \sum_{n\geq 1}\frac{(-1)^n(2n+1)}{n(n+1)}\,\tilde{P}_n(x),\qquad -\log(1-x) = 1 + \sum_{n\geq 1}\frac{(2n+1)}{n(n+1)}\,\tilde{P}_n(x)$$

and the lesser known:

$$K(k) = \int_0^{\pi/2}\frac{d\theta}{\sqrt{1-k^2\sin^2\theta}} = 2\sum_{n\geq 0}\frac{P_n(2k^2-1)}{2n+1}$$

about the complete elliptic integral of the first kind.
Legendre and Chebyshev polynomials share the uniform boundedness property

$$\forall x \in [-1, 1],\quad |P_n(x)| \leq 1$$

but that is not entirely trivial for Legendre polynomials.
A possible proof relies on the application of Cauchy-Schwarz inequality to the integral representation

$$P_n(x) = \frac{1}{\pi}\int_0^{\pi}\left(x + i\sqrt{1-x^2}\cos\theta\right)^n\,d\theta$$

that is a consequence of the generating function. An improved inequality is due to Tricomi:

$$\forall x \in (-1, 1),\qquad |P_n(x)| \leq \frac{2}{\sqrt{(2n+1)\pi}}\cdot\frac{1}{\sqrt[4]{1-x^2}}.$$

Readers can find a sketch of its (highly non-trivial) proof on the Whittaker&Watson book or on MSE (thanks to M.Spivey). From these remarks (and/or from Bonnet's recursion formula) it follows that the Legendre polynomial $P_n(x)$, just like the Chebyshev polynomial $T_n(x)$, has only real roots.

Theorem 73 (Turán)**.** For any x in the interval $(-1, 1)$ the following inequality holds:

$$P_n(x)^2 > P_{n-1}(x)\,P_{n+1}(x).$$

We will later see a proof of this brilliant result (a key ingredient for the Askey-Gasper inequality, that led De Branges in 1985 to the proof of Bieberbach conjecture) based on the following remarks:

- due to the integral representation, $\{P_n(x)\}_{n\geq 0}$ is a *sequence of moments*;

- due to the Cauchy-Schwarz inequality, every sequence of moments is log-convex.

In this paragraph we will outline an alternative and really elegant approach, due to Szegö.
We just need few preliminary lemmas.

Lemma 74 (Newton's inequality)**.** If $a_1, \ldots, a_n$ are distinct real numbers and σ_k is the k-th elementary symmetric function of $a_1, \ldots, a_n$, by setting $S_k = \sigma_k \binom{n}{k}^{-1}$ it follows that $S_k^2 > S_{k-1}S_{k+1}$.

Lemma 75 (Pólya, Schur)**.** If

$$f(z) = \sum_{n\geq 0} \frac{a_n}{n!} z^n$$

is an entire function and the zeroes of f are real and simple, as soon as these zeroes $\zeta_1, \zeta_2, \zeta_3, \ldots$ fulfill

$$\sum_{k\geq 1} \frac{1}{\zeta_k^2} < +\infty$$

it happens that $a_{n+1}^2 > a_n a_{n+2}$.

Lemma 76 (Gauss, Lucas)**.** The Bessel function of the first kind

$$J_0(z) = \sum_{n\geq 0} \frac{(-1)^n z^{2n}}{4^n n!^2}$$

fulfills the hypothesis of the previous Lemma.

Szegö's remark is that the following identity is easy to derive from the generating function for Legendre polynomials:

$$\sum_{n\geq 0} \frac{P_n(x)}{n!} z^n = e^{zx} J_0(z\sqrt{1-x^2})$$

hence Turán's inequality is a straightforward consequence of Pólya-Schur's Lemma. Szegö's idea is quite deep since it implies the existence of many "Turán-type" inequalities, not only for Legendre polynomials, but for many solutions of second-order differential equations with polynomial coefficients: Chebyshev, Hermite, Laguerre, Jacobi polynomials, Bessel functions ...

Exercise 77. Prove the identity:

$$\sum_{j=0}^{n} \binom{n}{j}\binom{n+j}{j} \frac{(-1)^{n+j}}{(j+1)^2} = \frac{(-1)^n}{n(n+1)}.$$

Exercise 78 (A Fejér-Jackson-like inequality)**.** Prove that for any natural number n,

$$\forall x \in (-1,1), \qquad \sum_{k=0}^{n} P_k(x) > 0.$$

Exercise 79. Prove the following combinatorial identity:

$$\sum_{l=0}^{n} \binom{n}{l}^2 (x+y)^{2l}(x-y)^{2n-2l} = \sum_{l=0}^{n} \binom{2l}{l}\binom{2n-2l}{n-l} x^{2l} y^{2n-2l}.$$

Exercise 80 (Ramanujan-like formulas for $\frac{1}{\pi}$ and $\frac{1}{\pi^2}$)**.** Use Rodrigues' formula or the generating function for Legendre polynomials to show that

$$\frac{1}{\sqrt{x(1-x)}} \stackrel{L^2(0,1)}{=} \pi \sum_{n\geq 0} \frac{4n+1}{16^n}\binom{2n}{n}^2 P_{2n}(2x-1).$$

Use Bonnet's formula to deduce how the Fourier-Legendre series of a function $g(x)$ changes if $g(x)$ is replaced by $g(1-x)$ or $x \cdot g(x)$. Use such transformations for showing that

$$\sqrt{x(1-x)} \stackrel{L^2(0,1)}{=} \frac{\pi}{8} \sum_{n\geq 0} \frac{4n+1}{(n+1)(1-2n)16^n}\binom{2n}{n}^2 P_{2n}(2x-1).$$

Compute $P_{2n}(0)$ in explicit terms, then use the evaluation of the previous line at $x = \frac{1}{2}$ and Parseval's identity for showing that

$$\frac{4}{\pi} = \sum_{n\geq 0} \frac{(4n+1)(-1)^n}{(n+1)(1-2n)64^n}\binom{2n}{n}^3,$$

$$\frac{32}{3\pi^2} = \sum_{n\geq 0} \frac{(4n+1)}{(n+1)^2(2n-1)^2 256^n}\binom{2n}{n}^4.$$

A note on Delannoy numbers and their asymptotic behaviour. Let us assume to travel in $\mathbb{Z}\times\mathbb{Z}$, having the origin as a starting point and the allowed steps

$$(n,m)\mapsto(n+1,m) \quad \text{or} \quad (n,m)\mapsto(n,m+1) \quad \text{or} \quad (n,m)\mapsto(n+1,m+1).$$

Let us denote with D_n the number of paths from the origin to (n,n).
By this way we define the sequence of *Delannoy numbers*

$$\{D_n\}_{n\geq 0} = \{1, 3, 13, 63, 321, 1683, 8989, \ldots\}$$

which are straightforward to describe through a bivariate generating function:

$$D_n = [x^n y^n]\frac{1}{1-x-y-xy} = [x^n y^n]\frac{1}{2-(1+x)(1+y)} = \sum_{k\geq 0}\frac{1}{2^{k+1}}\binom{k}{n}^2.$$

Through Cauchy's integral theorem or the orthogonality relations in $L^2(0,2\pi)$ the RHS of the previous line can be written as

$$D_n = \frac{1}{2\pi}\int_0^{2\pi}\frac{d\theta}{(3-2\sqrt{2}\cos\theta)^{n+1}}$$

giving that $\{D_n\}_{n\geq 0}$ is a sequence of moments, hence a log-convex sequence.
This integral representation leads to the ordinary generating function

$$\sum_{n\geq 0} D_n x^n = \frac{1}{\sqrt{1-6x+x^2}}$$

and by the Hayman/Laplace method we have the asymptotic approximation

$$D_n \sim \frac{1}{\sqrt{3n}}(1+\sqrt{2})^{2n}$$

as $n\to+\infty$. Given the generating function for Legendre polynomials, we have $D_n = P_n(3)$.

Exercise 81. The n-th Fibonacci number can be computed in at most $(2+\varepsilon)\log_2(n)$ integer multiplications by exploiting the relations between F_n and L_n and the duplication formulas

$$L_{2n} = L_n^2 + 2(-1)^{n+1}, \qquad L_{2n+1} = L_n L_{n+1} + (-1)^{n+1}.$$

Is it possible to compute the n-th Delannoy number in $O(\log n)$ integer multiplications?

Exercise 82. Prove that if we restrict the paths defining D_n to lie in the region $m \leq n$, we get the sequence of Schroeder numbers $\{S_n\}_{n\geq 0}$, with ordinary generating function

$$\sum_{n\geq 0} S_n x^n = \frac{1 - x - \sqrt{1-6x+x^2}}{2x}$$

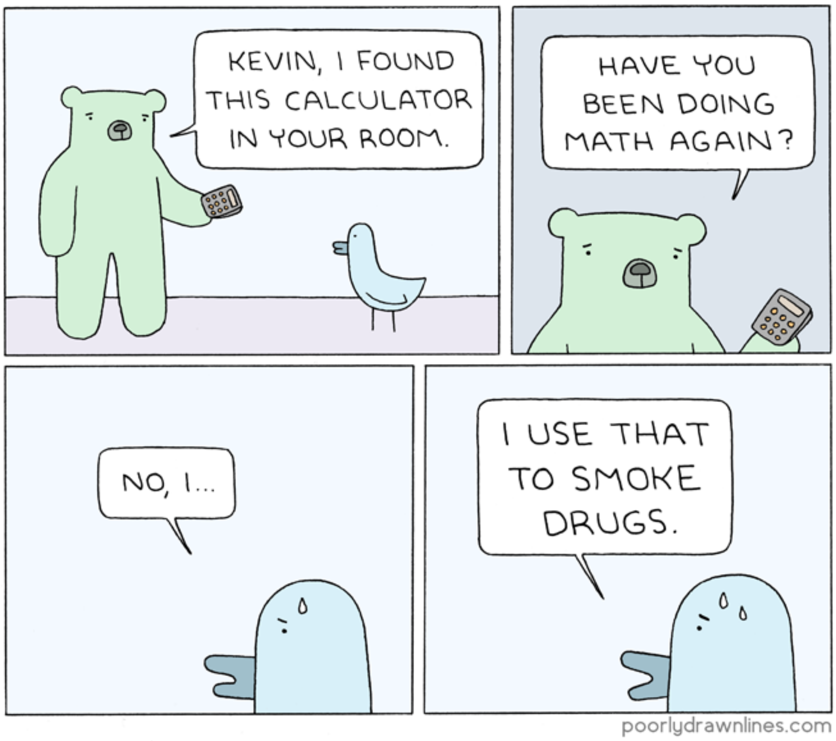
and related to Legendre polynomials via

$$S_n = \frac{1}{2}\left(-P_{n-1}(3) + 6P_n(3) - P_{n+1}(3)\right).$$

Through the last identity, find the asymptotic behaviour of S_n as $n \to +\infty$.

Exercise 83. Prove that for any prime p we have

$$D_{p-1} \equiv 1 \pmod{p}.$$



4 The glory of Fourier, Laplace, Feynman and Frullani

In the computation of many limits it is often very practical to exploit Taylor series and Landau notation, i.e., loosely speaking, the fact that a wide class of functions is well-approximated by polynomials. A key observation in Analysis (due to Joseph Fourier) is that the same holds for *trigonometric* polynomials: every real function that is 2π-periodic and regular enough can be written as a combination of terms of the form $\sin(nx)$ or $\cos(mx)$ (in Signal Processing these terms are often called *harmonics*). For every couple (m, n) of positive natural numbers the following orthogonality relations, with respect to the canonical inner product $\langle f, g\rangle = \int_0^{2\pi} f(x)\,g(x)\,dx$, hold:

$$\int_0^{2\pi} \sin(nx)\sin(mx)\,dx = \int_0^{2\pi} \cos(nx)\cos(mx)\,dx = \pi\,\delta(m,n), \qquad \int_0^{2\pi} \sin(nx)\cos(mx)\,dx = 0.$$

In particular, the coefficients of the involved harmonics can be easily computed through integrals: the determination of such coefficients is a problem equivalent to finding the coordinates of a vector in a infinite-dimensional vector space, equipped with an inner product and an orthogonal base. We perform such computation in three examplary cases.

Lemma 84 (Fourier series of the sawtooth wave)**.**
Let $g(x)$ be the 2π-periodic function that equals $\frac{\pi-x}{2}$ on the interval $(0, 2\pi)$.

$$\forall x \in \mathbb{R}\setminus 2\pi\mathbb{Z}, \qquad f(x) = \sum_{n\geq 1} \frac{\sin(nx)}{n}.$$

Lemma 85 (Fourier series of the rectangle wave)**.**
Let $f(x)$ be the 2π-periodic function that equals 1 on the interval $(0, \pi)$ and -1 on the interval $(\pi, 2\pi)$.

$$\forall x \in \mathbb{R}\setminus \pi\mathbb{Z}, \qquad f(x) = \frac{4}{\pi}\sum_{n\geq 0} \frac{\sin((2n+1)x)}{2n+1}.$$

Lemma 86 (Fourier series of the triangle wave)**.**
Let $h(x)$ be the 2π-periodic function that equals $\pi - |x|$ on the interval $[-\pi, \pi]$.

$$\forall x \in \mathbb{R}, \qquad f(x) = \frac{\pi}{2} + \frac{4}{\pi}\sum_{n\geq 0} \frac{\cos((2n+1)x)}{(2n+1)^2}.$$

We may notice that 85 follows from 84 by mapping $g(x)$ into $g(x) - g(x+\pi)$, while 86 follows from 85 by integration. We remark that the computation of **Fourier coefficients** for a 2π-periodic and piecewise-polynomial function is always pretty simple, but there might be converge issues nevertheless:

- The Fourier series of $f(x)$ might fail to be pointwise convergent to $f(x)$ for some $x \in [0, 2\pi]$: for instance, such lack-of-convergence phenomenon occurs for sure if $\lim_{x\to 0^+} f(x)$ and $\lim_{x\to 0^-} f(x)$ exist but do not agree;
- Even if pointwise convergence holds, the Fourier series of f might fail to be uniformly convergent to $f(x)$ on $[0, 2\pi]$: that happens for sure when **Gibbs phenomenon** occurs:

$$\lim_{N\to+\infty} \sup_{x\in(\pi/(N+1),\pi)} \left| \frac{\pi - x}{2} - \sum_{n=1}^{N} \frac{\sin(nx)}{n} \right| = \int_0^1 \frac{\sin(\pi x)}{\pi x}\,dx \neq 0.$$

Moreover, given a trigonometric series, it might be extremely difficult to find a 2π-periodic function with the given series as a Fourier series. These subtleties will be investigated in full detail during Calculus courses. Now we are just interested in proving some consequences of the above results, by starting from this observation: if $f(x)$ is a 2π-periodic function with mean zero, it is a continuous function on $(0,2\pi)$ and its Fourier series is pointwise convergent to $f(x)$ on $\mathbb{R}\setminus\pi\mathbb{Z}$, by performing a termwise integration on the Fourier series of $f(x)$ we get the Fourier series of an antiderivative for $f(x)$, and the convergence becomes uniform. In particular, for any $x\in(0,2\pi)$ we have:

$$\frac{\pi x}{2}-\frac{x^2}{4}=\int_0^x\frac{\pi-y}{2}\,dy=\sum_{n\geq 1}\frac{1-\cos(nx)}{n^2}=c_0-\sum_{n\geq 1}\frac{\cos(nx)}{n^2}$$

where c_0 has to be the *mean value* of the function $\frac{\pi x}{2}-\frac{x^2}{4}$ on the interval $(0,2\pi)$, since every term of the form $\cos(nx)$ has mean zero. It follows that:

$$c_0=\zeta(2)=\frac{1}{8\pi}\int_0^{2\pi}\left(2\pi x-x^2\right)\,dx=\pi^2\int_0^1\left(x-x^2\right)\,dx=\pi^2\left(\frac{1}{2}-\frac{1}{3}\right)=\frac{\pi^2}{6}$$

and such identity further leads to:

$$\forall x\in(0,1),\qquad \sum_{n\geq 1}\frac{\cos(2\pi nx)}{n^2}=\frac{\pi^2}{6}(1-6x+6x^2).$$

Due to the orthogonality relations

$$\int_0^1\cos(2\pi mx)\cos(2\pi nx)\,dx=\frac{1}{2}\delta(m,n)$$

we also have that:

$$\zeta(4)=\sum_{n\geq 1}\frac{1}{n^4}=2\int_0^1\left[\frac{\pi^2}{6}(1-6x+6x^2)\right]^2\,dx=\frac{\pi^4}{18}\int_0^1(1-6x+6x^2)^2\,dx=\frac{\pi^4}{90}.$$

The approach just outlined has an interesting generalization:

Lemma 87. By defining the sequence of **Bernoulli polynomials** through $B_0(x)=1$ and

$$\forall n\geq 0,\qquad B_{n+1}(x)=\kappa_{n+1}+(n+1)\int_0^x B_n(y)\,dy,$$

where the κ_{n+1} constant is chosen in such a way that $B_{n+1}(x)$ has mean zero over $(0,1)$, we have that for any $n\geq 1$ the value of $\zeta(2n)$ is a rational multiple of $\pi^{2n}\int_0^1 B_n(x)^2\,dx$, hence:

$$\frac{\zeta(2n)}{\pi^{2n}}\in\mathbb{Q}.$$

If we apply the same technique to the Fourier series of the triangle wave $\sum_{n\geq 0}\frac{(-1)^n\sin((2n+1)x)}{(2n+1)^2}$, we may prove in a similar way that:

$$\forall m\in\mathbb{N},\qquad \sum_{n\geq 0}\frac{(-1)^n}{(2n+1)^{2m+1}}\in\pi^{2m+1}\mathbb{Q}.$$

In particular, from the identity

$$\sum_{n\geq 0}\frac{(-1)^n}{(2n+1)^3}=\frac{\pi^3}{32}$$

it follows that $\pi^3\approx 31$ is a pretty accurate approximation. Let us see how to prove the last identity through a very powerful tool, the **Laplace transform**. Let us assume that $f(x)$ is a continuous and "vaguely integrable" function on $\mathbb{R}^+$, meaning that the following limit

$$\lim_{M\to+\infty}\int_0^M f(x)e^{-sx}\,dx$$

is finite for any $s \in \mathbb{R}^+$. Its Laplace transform, denoted by $\mathcal{L}f$, is defined through:

$$\forall s \in \mathbb{R}^+, \qquad (\mathcal{L}f)(s) = \int_0^{+\infty} f(x)e^{-sx}\,dx.$$

In the given hypothesis, the map sending a continuous and vaguely integrable function into its Laplace transform is injective, so if $\mathcal{L}g = f$ holds we may say that g is the **inverse Laplace transform** of f, by using the notation $g = \mathcal{L}^{-1}f$.
If $f(x) = x^k$ with $k \in \mathbb{N}$, we may notice that

$$\mathcal{L}f(s) = \frac{k!}{s^{k+1}},$$

hence by exploiting the linearity of the Laplace transform and the integration by parts formula we may state that:

$$\mathcal{L}^{-1}\left(\frac{1}{(2x+1)^3}\right) = \frac{s^2}{16}e^{-s/2}.$$

That allows us to convert the series $\sum_{n\geq 0}\frac{(-1)^n}{(2n+1)^3}$ into an (indefinite) integral:

$$\sum_{n\geq 0}\frac{(-1)^n}{(2n+1)^3} = \int_0^{+\infty}\sum_{n\geq 0}\frac{s^2}{16}e^{-s/2}(-1)^n e^{-ns}\,ds = \frac{1}{32}\int_0^{+\infty}\frac{s^2}{\cosh(s/2)}\,ds$$

and the equality between the last integral and π^3 is a straightforward consequence of the residue Theorem. In a similar way,

$$\frac{3}{4}\zeta(2) = \sum_{n\geq 0}\frac{1}{(2n+1)^2} = \frac{1}{8}\int_0^{+\infty}\frac{s\,ds}{\sinh(s/2)}.$$

The trick of converting a series into an integral through the (inverse) Laplace transform is widespread in Analytic Number Theory: for instance, it is the key ingredient of Ankeny and Wolf's proof of the fact that, by assuming the generalized Riemann hypothesis (GRH), for any prime p large enough the minimum quadratic non-residue (mod p) is $\leq 2\log^2(p)$. It is interesting to underline that the sharpest, GRH-independent upper bound actually known, arising from the combination of Burgess inequality with Vinogradov's amplification trick, is extremely weaker: $\eta_p \ll p^{\frac{1}{4\sqrt{e}}}$.
Another crucial property of the Laplace transform is the following one:

Lemma 88. Under suitable (and not particularly restrictive) hypothesis on the regularity of f and g and their speed of decay, we have:

$$\int_0^{+\infty} f(x)\cdot g(x)\,dx = \int_0^{+\infty}(\mathcal{L}f)(s)\cdot(\mathcal{L}^{-1}g)(s)\,ds.$$

There are many famous applications of this Lemma, that can be seen as a "regularized" version of the integration by parts formula. For instance, **Dirichlet** and **Fresnel**'s integrals are very simple to compute by using the Laplace transform.

Lemma 89 (Dirichlet)**.**

$$\lim_{M\to+\infty}\int_0^M \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$$

Proof. Since $\mathcal{L}^{-1}\left(\frac{1}{x}\right) = 1$ and, due to the integration by parts formula, $\mathcal{L}(\sin x) = \frac{1}{s^2+1}$, we have:

$$\lim_{M\to+\infty}\int_0^M \frac{\sin x}{x}\,dx = \int_0^{+\infty}\frac{ds}{s^2+1} = \lim_{N\to+\infty}\arctan(N) = \frac{\pi}{2}.$$

□

Lemma 90 (Fresnel)**.**

$$\lim_{M\to+\infty}\int_0^M \sin(x^2)\,dx = \lim_{M\to+\infty}\int_0^M \cos(x^2)\,dx = \sqrt{\frac{\pi}{8}}.$$

Proof. Since the equality $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$ holds (that can be proved by considering the integral of a gaussian function or, equivalently, as a consequence of Legendre's duplication formula for the Γ function), we have that $\mathcal{L}^{-1}\left(\frac{1}{\sqrt{x}}\right)=\frac{1}{\sqrt{\pi s}}$ (i.e. $\frac{1}{\sqrt{x}}$ is an *eigenfunction* for the Laplace transform). In particular:

$$\int_0^{+\infty}\sin(x^2)\,dx = \int_0^{+\infty}\frac{\sin x}{2\sqrt{x}}\,dx = \frac{1}{2\sqrt{\pi}}\int_0^{+\infty}\frac{ds}{\sqrt{s}(s^2+1)} = \frac{1}{\sqrt{\pi}}\int_0^{+\infty}\frac{dt}{t^4+1},$$
$$\int_0^{+\infty}\cos(x^2)\,dx = \int_0^{+\infty}\frac{\cos x}{2\sqrt{x}}\,dx = \frac{1}{2\sqrt{\pi}}\int_0^{+\infty}\frac{\sqrt{s}\,ds}{(s^2+1)} = \frac{1}{\sqrt{\pi}}\int_0^{+\infty}\frac{t^2\,dt}{t^4+1}.$$

We remark that the integrals of rapidly oscillating functions have been converted into integrals of positive and very well-behaved functions. Additionally, from Sophie Germain's identity

$$1+4u^4=(1-2u+2u^2)(1+2u+2u^2)$$

we can derive the partial fraction decomposition of $\frac{1}{t^4+1},\frac{t^2}{t^4+1}$ and the computation of Fresnel integrals boils down to the computation of some values of the arctangent function, like in the previous case. □

We also have that the Laplace transform (or the Fourier transform, that we have not introduced yet) can be used to define derivatives of fractional order, or *fractional derivatives*. We have indeed that due to the integration by parts formula, the Laplace transform of $f^{(n)}(x)$ depends in a very simple way on $s^n\left(\mathcal{L}f\right)(s)$, hence we may define a half-differentiation operator (also known as *semiderivative*) in the following way:

$$D^{1/2}f(x)=\mathcal{L}^{-1}\left[\sqrt{s}\left(\mathcal{L}f\right)(s)\right](x)$$

As soon as all the involved transforms and inverse transforms are well-defined, we have $D^{1/2}\circ D^{1/2}=\frac{d}{dx}$.
For instance, by setting

$$C=\int_{-\infty}^{+\infty}e^{-x^2}\,dx$$

we have $\mathcal{L}(\sqrt{x})=\frac{C}{2s^{3/2}}$ and $D^{1/2}x=\frac{2}{C}\sqrt{x}$. Additionally $\mathcal{L}\left(\frac{1}{\sqrt{x}}\right)=\frac{C}{\sqrt{s}}$ and

$$D^{3/2}x=\frac{d}{dx}D^{1/2}x=\frac{1}{C\sqrt{x}}.$$

However this is not the *only* way for defining fractional derivatives.
For instance, for every C^∞ function on the whole real line we have that:

$$f^{(n)}(x)=\lim_{h\to0^+}\frac{1}{h^n}\sum_{k\geq0}\binom{n}{k}(-1)^k f(x-kh)$$

where the binomial coefficient $\binom{n}{k}$ is well-defined also if $n\notin\mathbb{N}$.
So we might introduce the semiderivative of f (in the Grünwald-Letnikov sense) also as

$$\lim_{h\to0^+}\frac{1}{\sqrt{h}}\sum_{k\geq0}\binom{1/2}{k}(-1)^k f(x-kh)=\lim_{h\to0^+}\frac{1}{\sqrt{h}}\left[f(x)-\sum_{k\geq1}\binom{2k}{k}\frac{f(x-kh)}{4^k(2k-1)}\right]$$

but in order that the involved series is (at least conditionally) convergent, the f function has to be smooth and with a sufficiently rapid decay to zero. For functions in the Schwartz space $\mathcal{S}(\mathbb{R})$, the two given definitions of semiderivative are equivalent, but they are not in full generality.

Exercise 91. Investigate about the interplay between the semiderivative of the sine function (defined through the Laplace transform) and Fresnel integrals.

Exercise 92. Prove that:

$$\lim_{\alpha\to+\infty} 2\alpha\int_0^{+\infty}\sin(x^\alpha)\,dx=\pi,\qquad \int_0^{+\infty}\left(\frac{\sin x}{x}\right)^4 dx=\frac{\pi}{3}.$$

A series related to the Trigamma function.

Lemma 93. For any $a>0$, the integration by parts formula leads to:

$$\frac{1}{a}\int_0^{+\infty}\cos(nx)\,e^{-ax}\,dx=\frac{1}{a^2+n^2}=\int_0^{+\infty}\frac{\sin(nx)}{n}\,e^{-ax}\,dx.$$

Lemma 94. The series

$$\sum_{n=1}^{+\infty}\frac{\sin(nx)}{n}$$

is pointwise convergent on $\mathbb{R}\setminus 2\pi\mathbb{Z}$ to the function:

$$f(x)=\pi\left(\frac{1}{2}-\left\{\frac{x}{2\pi}\right\}\right)$$

where $\{x\}$ stands for the fractional part of x, i.e. $x-\lfloor x\rfloor$.

Due to the dominated convergence Theorem we have:

$$\sum_{n=1}^{+\infty}\frac{1}{a^2+n^2}=\pi\int_0^{+\infty}\left(\frac{1}{2}-\left\{\frac{x}{2\pi}\right\}\right)e^{-ax}\,dx,$$

and by partitioning $[0,+\infty)$ as $[0,2\pi)\cup[2\pi,4\pi)\cup\ldots$ it follows that:

$$\sum_{n=1}^{+\infty}\frac{1}{a^2+n^2}=\frac{e^{2a\pi}}{e^{2a\pi}-1}\int_0^{2\pi}\frac{\pi-x}{2}e^{-ax}dx=\frac{\pi a\coth(\pi a)-1}{2a^2}.$$

By computing the limit of both sides as $a\to 0^+$, we find $\zeta(2)=\frac{\pi^2}{6}$ again.

Exercise 95. Prove the following identity through the Laplace transform:

$$\int_0^{+\infty}e^{-t}\sin^2(t)\frac{dt}{t}=\frac{1}{4}\log 5.$$

Exercise 96. Compute the explicit values of these indefinite integrals:

$$\int_0^{+\infty}\frac{\sin(s)}{e^s-1}\,ds,\qquad \int_0^{+\infty}\frac{\sin(x)-x\cos(x)}{x^2}\,dx.$$

Exercise 97. By considering the Laplace transform of the indicator function of $(0,a)$, prove that for any $a>0$ we have:

$$\arctan(a)=\int_0^{+\infty}\left(1-e^{-at}\right)\frac{\sin t}{t}\,dt.$$

Exercise 98. Prove that for any $u\in(0,1)$ we have:

$$\int_0^{+\infty}\frac{1}{x+1-u}\cdot\frac{dx}{\pi^2+\log^2(x)}=\frac{1}{u}+\frac{1}{\log(1-u)}.$$

Exercise 99. By the Laplace transform and the Cauchy-Schwarz inequality, prove that:

$$\int_0^{+\infty}\frac{\sin x}{x(x+1)}\,dx\leq 1.$$

Exercise 100 (The Laplace transform as an acceleration trick)**.**
Prove that, due to the Laplace transform:

$$\sum_{n\geq 0}\frac{(-1)^n}{\sqrt{2n+1}}=\frac{1}{\sqrt{\pi}}\int_0^{+\infty}\frac{ds}{\cosh(s^2)}$$

where a slowly convergent series has been converted into the integral of a function in $\mathcal{S}(\mathbb{R}^+)$.
Notice that it is possible to state:

$$\frac{1}{\sqrt{\pi}}\int_0^{+\infty}\frac{ds}{\exp\left(\frac{s^4}{2}\right)}\leq\sum_{n\geq 0}\frac{(-1)^n}{\sqrt{2n+1}}\leq\frac{1}{\sqrt{\pi}}\int_0^{+\infty}\frac{ds}{1+\frac{s^4}{2}}.$$

The Laplace transform also encodes an important Theorem known as **differentiation under the integral sign** or **Feynman's trick**. In his biography *Surely you're joking, Mr. Feynman*, the American physicist Richard Feynman talks about instruments learned by reading Woods *Advanced Calculus* book, that in Feynman's opinion did not get proper credit in University courses: differentation under the integral sign granted some fame to Feynman, since he proved many logarithmic integrals can be tackled in a very slick way by exploiting such instrument.

Lemma 101. As soon as the hypothesis of the dominated convergence Theorem are met,

$$\frac{d}{dy}\int_E f(x,y)\,dx=\int_E\frac{\partial f}{\partial y}(x,y)\,dx.$$

The very deep consequences of the fact that, under suitable assumptions, the operators $\frac{\partial}{\partial y}$ and $\int_E dx$ commute might not be evident at first sight. Practical applications are needed to fully understand the hidden "magic".

Exercise 102. Prove that

$$I = \int_0^1 \frac{\arctan(x)}{x\sqrt{1-x^2}}\,dx = \frac{\pi}{2}\log(1+\sqrt{2}).$$

Proof. If we define $I(k)$ as

$$I(k) = \int_0^{\pi/2} \frac{\arctan(k\sin\theta)}{\sin\theta}\,d\theta$$

we have that $I = I(1)$ and

$$I'(k) = \int_0^{\pi/2} \frac{1}{1+k^2\sin^2\theta} = \frac{\pi}{2\sqrt{1+k^2}}$$

holds by the substitution $\theta = \arctan(t)$. Since $\lim_{k\to 0^+} I(k) = 0$, it follows that:

$$I = I(1) = \frac{\pi}{2}\int_0^1 \frac{dk}{\sqrt{1+k^2}} = \frac{\pi}{2}\operatorname{arcsinh}(1) = \frac{\pi}{2}\log(1+\sqrt{2})$$

as wanted. □

Exercise 103. Prove that for any $k \in \mathbb{R}$

$$\int_0^{\pi/2} \log(\sin^2\theta + k^2\cos^2\theta)\,d\theta = \pi\log\left(\frac{1+|k|}{2}\right).$$

Proof. By denoting as $I(k)$ the integral in the LHS and exploiting $\int_0^{\pi/2}\log\tan(\theta)\,d\theta = 0$ we have that:

$$I(k) = \int_0^{\pi/2} \log(k^2+\tan^2\theta)\,d\theta = \int_0^{+\infty} \frac{\log(k^2+t^2)}{1+t^2}\,dt$$

and by differentiation under the integral sign:

$$I'(k) = \int_0^{+\infty} \frac{2k\,dt}{(k^2+t^2)(1+t^2)} = \frac{\pi}{k+\operatorname{Sign}(k)}.$$

Since $I(1) = 0$, the claim follows in a straightforward way. □

Exercise 104. Prove that for any couple (a, b) of distinct and positive natural numbers we have:

$$I(a,b) = \int_0^{+\infty} \frac{\log x}{(x+a)(x+b)}\,dx = \frac{\log^2(b)-\log^2(a)}{2(b-a)}.$$

Proof. By exploiting partial fraction decomposition and integration by parts,

$$\begin{aligned}\int_0^{+\infty} \frac{\log x}{(x+a)(x+b)}\,dx &= \frac{1}{b-a}\int_0^{+\infty}\left(\frac{x}{x+a}-\frac{x}{x+b}\right)\frac{\log(x)}{x}\,dx \\ &= \frac{1}{2(b-a)}\int_0^{+\infty}\left(\frac{b}{(x+b)^2}-\frac{a}{(x+a)^2}\right)\log^2(x)\,dx.\end{aligned}$$

If we define $J(c)$ as

$$J(c)=\int_0^{+\infty}\frac{c\log^2(x)}{(x+c)^2}\,dx=\int_0^{+\infty}\frac{\log^2(cz)}{(z+1)^2}\,dz$$

and notice that

$$\int_0^{+\infty}\frac{\log^2(c)}{(z+1)^2}\,dz=\log^2(c),\qquad\int_0^{+\infty}\frac{2\log(c)\log(z)}{(z+1)^2}\,dz=0$$

we immediately have:

$$J(c)=\log^2(c)+\int_0^{+\infty}\frac{\log^2(z)}{(z+1)^2}\,dz=\log^2(c)+2\int_0^1\frac{\log^2(z)}{(z+1)^2}\,dz$$

and the claim, since the last integral does not depend on c.
Anyway we may observe that by the integration by parts formula

$$\int_0^1\frac{\log^2(z)}{(1+z)^2}\,dz=-\int_0^1\frac{2\log z}{(1+z)}\,dz=2\int_0^1\frac{\log(1+z)}{z}\,dz=2\sum_{n\geq1}\int_0^1\frac{(-1)^{n+1}z^{n-1}}{n}\,dz=2\sum_{n\geq1}\frac{(-1)^{n+1}}{n^2}=\zeta(2).$$

□

Exercise 105. Prove that:

$$\int_0^{+\infty}\left[1-\cos\left(t\sqrt{e-1}\right)\right]\frac{dt}{te^t}=\frac{1}{2}.$$

Proof. We may introduce the parametric integral

$$I(\omega)=\int_0^{+\infty}\frac{1-\cos(\omega t)}{t}e^{-t}\,dt.$$

It is trivial that $\lim_{\omega\to0^+}I(\omega)=0$. We also have:

$$I'(\omega)=\omega\int_0^{+\infty}\sin(\omega t)e^{-t}\,dt=\frac{\omega}{1+\omega^2}$$

hence for any $\omega>0$

$$I(\omega)=\int_0^{\omega}\frac{u}{1+u^2}\,du=\frac{1}{2}\log(1+\omega^2)$$

and by considering $\omega=\sqrt{e-1}$ the claim follows. □

Exercise 106. Prove that, if $a>b>0$,

$$I(a,b)=\int_0^{\pi}\frac{d\theta}{(a+b\cos\theta)^2}=\frac{\pi a}{(a^2-b^2)^{3/2}}.$$

Proof. We may notice that:

$$I(a,b)=\int_0^{\pi/2}\frac{d\theta}{(a+b\cos\theta)^2}+\int_0^{\pi/2}\frac{d\theta}{(a-b\cos\theta)^2}=2\int_0^{\pi/2}\frac{a^2+b^2\cos^2\theta}{(a^2-b^2\cos^2\theta)^2}\,d\theta.$$

By enforcing the substitution $\theta=\arctan t$ we get:

$$I(a,b)=2\int_0^{+\infty}\frac{a^2(1+t^2)+b^2}{\left(a^2(1+t^2)-b^2\right)^2}\,dt=\frac{2}{a(a^2-b^2)^{3/2}}\int_0^{+\infty}\frac{(a^2+b^2)+(a^2-b^2)t^2}{(1+t^2)^2}\,dt$$

so $I(a,b)$ only depends on the integrals $H_0=\int_0^{+\infty}\frac{dt}{(1+t^2)^2}$ and $H_2=\int_0^{+\infty}\frac{t^2\,dt}{(1+t^2)^2}=\frac{\pi}{2}-H_0$.
On the other hand, for any $\beta>0$

$$\int_0^{+\infty}\frac{dt}{(\beta+t^2)^2}=-\int_0^{+\infty}\frac{\partial}{\partial\beta}\frac{1}{\beta+t^2}\,dt=-\frac{d}{d\beta}\int_0^{+\infty}\frac{dt}{\beta+t^2}=-\frac{d}{d\beta}\frac{\pi}{2\sqrt{\beta}}=\frac{\pi}{4\beta^{3/2}}$$

hence $H_0 = H_2 = \frac{\pi}{4}$ and we have:

$$I(a,b) = \frac{2}{a(a^2-b^2)^{3/2}} \cdot \frac{\pi a^2}{2} = \frac{\pi a}{(a^2-b^2)^{3/2}}$$

as wanted. The given problem can be solved through an interesting geometric approach, too. If a simple closed curve around the origin is regular and has a parametrization of the form $(\rho(\theta)\cos\theta, \rho(\theta)\sin\theta)$ for $\theta \in [0, 2\pi]$, the enclosed area is given by:

$$A = \frac{1}{2}\int_0^{2\pi} \rho(\theta)^2 \, d\theta.$$

In our case the curve decribed by $\rho(\theta) = \frac{1}{a+b\cos\theta}$ is an ellipse and

$$I(a,b) = \int_0^{\pi} \frac{d\theta}{(a+b\cos\theta)^2} = \frac{1}{2}\int_0^{2\pi} \rho(\theta)^2 \, d\theta$$

is precisely the area enclosed by such ellipse. Since affine maps preserve the ratios of areas, the area enclosed by an ellipse is π times the product between the lengths of semi-axis. The length of the major axis is given by $\frac{1}{a-b} - \frac{1}{a+b} = \frac{2a}{a^2-b^2}$ and the ratio between the square of the minor axis and the major axis (also known as *semi-latus rectum*) equals $\frac{1}{a}$. It follows that:

$$I(a,b) = A = \pi \cdot \frac{a}{a^2-b^2} \cdot \frac{1}{\sqrt{a^2-b^2}}$$

just like we found before. □

Exercise 107. Prove that for any couple (A, B) of positive real numbers the following identity holds:

$$I(A,B) = \int_0^{+\infty} \frac{dx}{\left(1+\frac{x^2}{A^2}\right)\left(1+\frac{x^2}{B^2}\right)} = \frac{\pi}{4}\,\text{HM}(A,B)$$

where $\text{HM}(A,B)$ is the **harmonic mean** of A and B, i.e. $\frac{2AB}{A+B}$.

Exercise 108. Find an accurate numerical approximation of the following integral:

$$\int_0^{+\infty} \frac{\sin(3x)\sin(4x)\sin(5x)}{x\sin^2(x)\cosh(x)}\,dx.$$

Proof. It is useful to notice that:

$$I(n) \stackrel{\text{def}}{=} \int_0^{+\infty} \frac{\sin(2nx)}{x\cosh(x)}\,dx = 2\arctan\left(\tanh\frac{\pi n}{2}\right).$$

By expanding $\frac{1}{\cosh x}$ as a geometric series we get

$$\frac{1}{\cosh x} = \frac{2}{e^x+e^{-x}} = \frac{2e^{-x}}{1+e^{-2x}} = 2\left(e^{-x} - e^{-3x} + e^{-5x} - e^{-7x} + \ldots\right)$$

and it follows that

$$I(n) = 2\sum_{k\geq 0}(-1)^k \int_0^{+\infty} \frac{\sin(2nx)}{x} e^{-(2k+1)x}\,dx.$$

By differentiating with respect to the n parameter:

$$\begin{aligned} I'(n) = 4\sum_{k\geq 0}(-1)^k \int_0^{+\infty} \cos(2nx)e^{-(2k+1)x}\,dx &= 4\sum_{k\geq 0}\frac{(2k+1)(-1)^k}{(2k+1)^2+4n^2} \\ &= 2\sum_{k\geq 0}(-1)^k\left(\frac{1}{(2k+1)+2in} + \frac{1}{(2k+1)-2in}\right). \end{aligned}$$

The last series can be interpreted as a logarithmic derivative, and it leads to the identity

$$I'(n) = \frac{\pi}{\cosh(\pi n)}.$$

Since $\lim_{n\to 0^+} I(n) = 0$, or by exploiting $\lim_{n\to+\infty} I(n) = 0$ that follows from the Riemann-Lebesgue Lemma, we get:

$$I(n) = \int_0^n \frac{\pi}{\cosh(\pi u)}\,du = 2\arctan\left(\tanh\frac{\pi n}{2}\right).$$

On the other hand, by exploiting Chebyshev polynomials of the second kind it is not difficult to check that:

$$\frac{\sin(3x)\sin(4x)\sin(5x)}{\sin(x)^2} = 2\sin(2x) + 3\sin(4x) + 3\sin(6x) + 2\sin(8x) + \sin(10x),$$

$$\int_0^{+\infty} \frac{\sin(3x)\sin(4x)\sin(5x)}{x\sin^2(x)\cosh(x)}\,dx = 2\,I(1) + 3\,I(2) + 3\,I(3) + 2\,I(4) + I(5) \approx \frac{11\pi}{2}$$

where the last approximation follows from the fact that $I(n)$, for $n \gg 1$, converges quite fast to the value $\frac{\pi}{2}$. □

Exercise 109. Prove that for any $t > 0$ we have:

$$I(t) = \int_0^1 \frac{x^t - 1}{\log x}\,dx = \log(t+1).$$

Proof. Since $x^t - 1 = \exp(t\log x) - 1 = t\log(x) + O\left(t^2\log^2 x\right)$ for $t \to 0^+$ and $-\log(x)$ is a positive and integrable function over the interval $(0,1)$, we have that $\lim_{t\to 0^+} I(t) = 0$. In order to prove the claim it is enough to show that:

$$\frac{d}{dt}\int_0^1 \frac{x^t - 1}{\log x}\,dx = \frac{1}{t+1} = \frac{d}{dt}\log(t+1)$$

which is a trivial consequence of $\frac{\partial}{\partial t}\frac{x^t-1}{\log x} = x^t$ and $\int_0^1 x^t\,dx = \frac{1}{t+1}$. □

Exercise 110. Compute the following integral by using Riemann sums:

$$\int_0^\pi \log(7 + \cos\theta)\,d\theta.$$

Proof. We may notice that for any $n \in \mathbb{N}^+$:

$$x^{2n} - 1 = \prod_{k=1}^{2n}\left(x - \exp\frac{\pi i k}{n}\right) = (x^2 - 1)\prod_{k=1}^{n-1}\left(x^2 + 1 - 2x\cos\frac{\pi k}{n}\right).$$

If we choose x in such a way that $\frac{x^2+1}{2x} = -7$ holds, for instance through $x = 4\sqrt{3} - 7$, we get:

$$\int_0^\pi \log(7 + \cos\theta)\,d\theta = \lim_{n\to+\infty}\frac{\pi}{n}\log\prod_{k=1}^{n-1}\left(7 + \cos\tfrac{\pi k}{n}\right) = \pi\log\left(\frac{7}{2} + 2\sqrt{3}\right).$$

□

Exercise 111. Prove that for any $a \in (-1, 1)$

$$\int_0^{\pi} \log(1 - 2a\cos x + a^2)\, dx = 0.$$

Exercise 112. Prove that

$$\int_0^1 \frac{\arctan x}{x+1} = \frac{\pi}{8}\log 2.$$

Exercise 113. Prove that

$$\int_0^{\pi/2} \log(\sin x)\log(\cos x)\, dx = \frac{\pi}{2}\log^2(2) - \frac{\pi^3}{48}.$$

At this point it should be clear that Feynman's trick is a really powerful tool.
In the following formal manipulation

$$\int_E f(x)\, dx = \int_E \tilde{f}(x, 1)\, dx = \int_0^1 \int_E \frac{\partial}{\partial\alpha}\tilde{f}(x,\alpha)\, dx = \int_0^1 g(\alpha)\, d\alpha$$

we have a complete freedom in choosing *where* to introduce the dummy parameter α in the definition of f, in such a way that $\tilde{f}(x,\alpha)|_{\alpha=1} = f(x)$ holds. In particular, Feyman's trick is really effective in the computation of logarithmic integrals, since $\log(x) = \frac{d}{d\alpha}x^\alpha\big|_{\alpha=0^+}$, or in proving identities like

$$\sum_{n\geq 1}\frac{\log n}{n^2} = -\frac{d}{d\alpha}\sum_{n\geq 1}\frac{1}{n^\alpha}\Bigg|_{\alpha=2} = -\zeta'(2).$$

Additionally differentiation under the integral sign, the Laplace transform and **Frullani's Theorem** provide interesting integral representations for logarithms.

Theorem 114 (Frullani)**.** If $f \in C^1(\mathbb{R}^+)$ and $\lim_{x\to+\infty} f(x) = 0$, for any couple (a, b) of positive real numbers we have:

$$\int_0^{+\infty} \frac{f(ax) - f(bx)}{x}\, dx = \log\left(\frac{b}{a}\right)\cdot \lim_{x\to 0^+} f(x).$$

The most typical case of Frullani's Theorem is related to the function $f(x) = e^{-x}$:

Lemma 115. If $b \geq a > 0$,

$$\int_0^{+\infty} \frac{e^{-ax} - e^{-bx}}{x}\, dx = \log\left(\frac{b}{a}\right).$$

Let us prove this Corollary through Feynman's trick and the Laplace transform.
It is enough to show that:

$$\forall a > 0, \qquad I(a) = \int_0^{+\infty} \frac{e^{-x} - e^{-ax}}{x}\, dx = \log(a).$$

It is pretty clear that $\lim_{a\to 1} I(a) = 0$. Since

$$I'(a) = \int_0^{+\infty} e^{-ax}\,dx = \frac{1}{a}$$

we have:

$$I(a) = \int_1^a I'(\xi)\,d\xi = \int_1^a \frac{d\xi}{\xi} = \log(a).$$

As an alternative, by exploiting $\mathcal{L}^{-1}\left(\frac{1}{x}\right) = 1$ and $\mathcal{L}(e^{-ax}) = \frac{1}{a+s}$,

$$I(a) = \int_0^{+\infty} \left(\frac{1}{s+1} - \frac{1}{a+s}\right) ds = \lim_{M\to+\infty} [\log(M+1) - \log(M+a) + \log(a)] = \log(a).$$

We remark the strong analogy between the last identity and $H_n = \sum_{m\geq 1}\left(\frac{1}{m} - \frac{1}{m+n}\right)$. Frullani's Theorem can be used to provide integral representations for the **Euler-Mascheroni constant** γ or the constant $\log\frac{\pi}{2}$:

Lemma 116 (An integral representation for the Euler-Mascheroni constant)**.**

$$\gamma \stackrel{\text{def}}{=} \lim_{n\to+\infty} (H_n - \log n) = \sum_{n\geq 1}\left[\frac{1}{n} - \log\left(1+\frac{1}{n}\right)\right] = \int_0^{+\infty}\left(\frac{1}{e^x-1} - \frac{1}{xe^x}\right) dx.$$

Proof. Due to Frullani's Theorem:

$$\frac{1}{n} - \log\left(1+\frac{1}{n}\right) = \int_0^{+\infty}\left(e^{-nx} - \frac{e^{-nx} - e^{-(n+1)}}{x}\right) dx$$

and the claim follows by summing both sides on $n \geq 1$.
We may notice that by Weierstrass product for the Γ function we have $\gamma = -\Gamma'(1)$.
It follows that by Feynman's trick we also have:

$$\gamma = -\frac{d}{ds}\int_0^{+\infty} x^{s-1}e^{-x}\,dx\Big|_{s=1} = -\int_0^{+\infty}\frac{\log x}{\exp(x)}\,dx.$$

□

Lemma 117 (An integral representation for the $\log\frac{\pi}{2}$ constant)**.**

$$\log\frac{\pi}{2} = \int_0^{+\infty}\frac{e^x-1}{e^x+1}\cdot\frac{dx}{xe^x}.$$

Proof. By Weierstrass product for the sinc function it is simple to prove that:

$$\frac{2}{\pi} = \prod_{n\geq 1}\left(1 - \frac{1}{4n^2}\right).$$

By switching to logarithms and exploiting Frullani's Theorem,

$$\log\frac{\pi}{2} = \sum_{n\geq 1}(-1)^{n+1}[\log(n+1) - \log(n)] = \int_0^{+\infty}\sum_{n\geq 1}(-1)^{n+1}\frac{e^{-nx} - e^{-(n+1)x}}{x}\,dx = \int_0^{+\infty}\frac{e^x-1}{e^x+1}\cdot\frac{dx}{xe^x}.$$

□

Exercise 118. Prove that

$$\sum_{n\geq 1}\left(n\log\frac{2n+1}{2n-1}-1\right)=\int_0^{+\infty}\frac{\sinh x-x}{2x\sinh^2(x)}\,dx=\frac{1-\log 2}{2}.$$

Hint: the first equality follows from the Laplace transform, and the equality between the first term and the last one is a consequence of Stirling's inequality, since:

$$\begin{aligned}S_N=\sum_{n=1}^{N}\left(n\log\frac{2n+1}{2n-1}-1\right) &= -N+\sum_{n=1}^{N}n\left[\log(2n+1)-\log(2n-1)\right]\\ &= -N+N\log(2N+1)-\sum_{n=1}^{N-1}\log(2n+1)\\ &= -N+N\log(2N+1)-\log\left[(2N-1)!!\right]\\ &= -N+N\log(2N+1)-\log\left[(2N)!\right]+N\log 2+\log\left[N!\right].\end{aligned}$$

As an alternative, we may notice that:

$$S_\infty=\sum_{n\geq 1}\left(2n\,\text{arctanh}\,\frac{1}{2n}-1\right)=\sum_{n\geq 1}\sum_{m\geq 1}\frac{1}{(2m+1)(2n)^{2m}}=\sum_{m\geq 1}\frac{\zeta(2m)}{4^m(2m+1)}$$

and recognize in the last series the integral $\int_0^{1/2}\left[1-\pi x\cot(\pi x)\right]\,dx$ that is simple to compute by integration by parts: by this way the problem boils down to the computation of the well-known integral $\int_0^{\pi/2}\log\sin(\theta)\,d\theta=-\frac{\pi}{2}\log(2)$.

We now introduce another very powerful tool for dealing with series, i.e. the discrete equivalent of the integration by parts formula.

Theorem 119 (Abel, Summation by parts formula)**.**
($\Sigma\Sigma$) version: if $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$ are two sequences of complex numbers, by setting $A_n=a_1+\ldots+a_n$ we have:

$$\sum_{n=1}^{N}a_nb_n=A_Nb_N-\sum_{n=1}^{N-1}A_n(b_{n+1}-b_n).$$

($\Sigma\int$) version: if $\phi(x)$ is a C^1 function on $\mathbb{R}^+$, for any sequence $\{a_n\}_{n\geq 1}$ of complex numbers we have:

$$\sum_{1\leq n\leq x}a_n\phi(n)=A(x)\phi(x)-\int_1^x A(u)\phi'(u)\,du$$

where

$$A(x)=\sum_{1\leq n\leq x}a_n.$$

The existence of a discrete-discrete and a discrete-continuous analogue of the integration by parts formula is very practical, almost *comforting*. Abel's theorem allows to study the convergence of a power series on the boundary of the region of convergence, or to regularize series whose convergence is uncertain at first sight. The following Lemma, for instance, is a straightforward consequence of the summation by parts formula:

Lemma 120 (Dirichlet's test)**.**
(Discrete version) If $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$ are two sequence of real numbers, such that $A_n = a_1 + \ldots + a_n$ is bounded and b_n is decreasing towards zero, the series

$$\sum_{n\geq 1} a_n b_n$$

is convergent.
(Continuous version) If $f(x), g(x) \in C^0(\mathbb{R}^+)$, where $g(x)$ is weakly decreasing towards zero and $\int_0^x f(u)\,du$ is bounded, the improper Riemann integral

$$\int_0^{+\infty} f(x)\,g(x)\,dx$$

is convergent.

The convergence of

$$\sum_{n\geq 1}\frac{\sin n}{n},\qquad \sum_{n\geq 1}\frac{\cos n}{n},\qquad \int_0^{+\infty}\frac{\sin x}{x}\,dx$$

immediately follows from Dirichlet's test, once we show two simple results on the behaviour of particular trigonometric sums.

Lemma 121. For any couple (k, N) of positive real numbers we have:

$$\left|\sum_{n=1}^{N}\sin(kn)\right| \leq \tfrac{1}{2}\left|\cot\tfrac{k}{4}\right|,\qquad \left|\sum_{n=1}^{N}\cos(kn)\right| \leq \frac{1}{2}+\frac{1}{\left|2\sin\frac{k}{2}\right|}.$$

Proof. Both the mentioned sums are telescopic sums in disguise.
As a matter of fact, by the sine/cosine addition formulas:

$$\sin\left(\tfrac{k}{2}\right)\sum_{n=1}^{N}\sin(kn) = \tfrac{1}{2}\sum_{n=1}^{N}\left[\cos\left(\left(n-\tfrac{1}{2}\right)k\right)-\cos\left(\left(n+\tfrac{1}{2}\right)k\right)\right] = \tfrac{1}{2}\left[\cos\tfrac{k}{2}-\cos\tfrac{(2N+1)k}{2}\right],$$

$$\sin\left(\tfrac{k}{2}\right)\sum_{n=1}^{N}\cos(kn) = \tfrac{1}{2}\sum_{n=1}^{N}\left[\sin\left(\left(n-\tfrac{1}{2}\right)k\right)-\sin\left(\left(n+\tfrac{1}{2}\right)k\right)\right] = \tfrac{1}{2}\left[\sin\tfrac{k}{2}-\sin\tfrac{(2N+1)k}{2}\right],$$

so, in particular:

$$\left|\sum_{n=1}^{N}\sin(kn)\right| \leq \left|\frac{1+\cos\frac{k}{2}}{2\sin\frac{k}{2}}\right| = \tfrac{1}{2}\left|\cot\tfrac{k}{4}\right|,\qquad \left|\sum_{n=1}^{N}\cos(kn)\right| \leq \left|\frac{1+\sin\frac{k}{2}}{2\sin\frac{k}{2}}\right| \leq \frac{1}{2}+\frac{1}{\left|2\sin\frac{k}{2}\right|}.$$

□

Exercise 122 (about Van Der Corput's trick and Weyl's inequality)**.** Prove that the series

$$\sum_{n\geq 1}\frac{\sin(n^2)}{n}$$

is convergent. Hint: it is enough to show that for any N large enough the inequality

$$\left|\sum_{n=1}^{N}\exp(in^2)\right| \leq 4\sqrt{N}\log^2(N)$$

holds, then apply summation by parts.

Let us see how to compute $\sum_{n\geq 1}\frac{\sin n}{n}$ and $\sum_{n\geq 1}\frac{\cos n}{n}$ in a explicit way, now that we know they are convergent series. Since we have:

$$\sum_{n\geq 1}\frac{z^n}{n} = -\log(1-z)$$

for any complex number z with modulus less than 1, it is reasonable to expect that, by evaluating both sides at $z = e^i$,

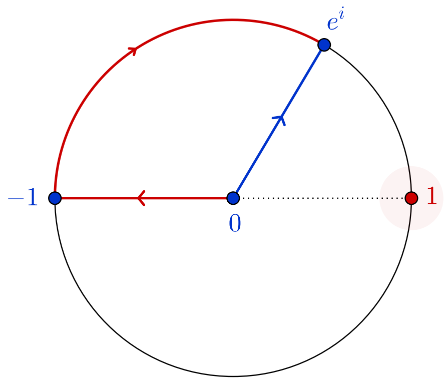
$$\sum_{n\geq 1}\frac{\sin n}{n} = -\text{Im}\log(1-e^i) = \frac{\pi-1}{2}, \qquad \sum_{n\geq 1}\frac{\cos n}{n} = -\text{Re}\log(1-e^i) = -\log\left(2\sin\tfrac{1}{2}\right)$$

holds. That is correct, indeed, but we have to explain why we are allowed to evaluate the Taylor series $\sum_{n\geq 1}\frac{z^n}{n}$ at a point on the boundary of the convergence region $|z|<1$. We may notice that the function $-\log(1-z)$ is continuous (and much more, actually: holomorphic) in a neighbourhood of $z = e^i$: that is enough to justify the "wild" evaluation we performed. Such observation is also known as **Abel's Lemma**. We may also proceed through a **deformation** of an integration path. Let us consider the function $g(z) = \frac{1}{1-z}$: it is a meromorphic function with a simple pole at $z = 1$, and at the interior of the unit disk we have:

$$\frac{1}{1-z} = 1 + z + z^2 + z^3 + \ldots = \sum_{n\geq 0} z^n.$$

By termwise integration it follows that:

$$-\log(1-z) = \int_0^z \frac{du}{1-u} = \sum_{n\geq 1}\frac{z^n}{n}, \qquad \sum_{n\geq 1}\frac{\sin n}{n} = \text{Im}\int_0^{e^i} g(z)\,dz.$$



Since $g(z)$ is holomorphic far from $z = 1$, the integral along the straight segment joining the origin with e^i equals the integral along the path given by the concatenation of the straight segment joining $z = 0$ with $z = -1$ and the depicted circle arc γ joining -1 with e^i. In particular:

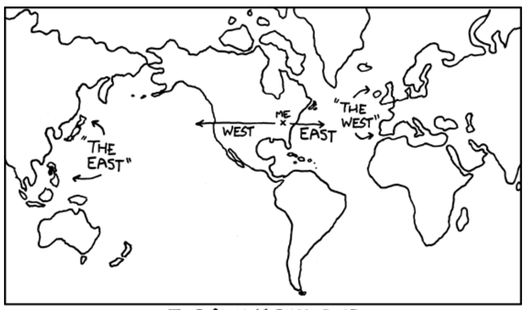
$$\sum_{n\geq 1}\frac{\sin n}{n} = \text{Im}\int_0^{-1}\frac{dz}{1-z} + \text{Im}\int_\gamma \frac{dz}{1-z}$$

and the first integral appearing in the RHS is a real number.

As a consequence:

$$\sum_{n\geq 1}\frac{\sin n}{n} = \text{Im}\int_{-1}^{e^i}\frac{dz}{1-z} = \text{Im}\int_\pi^1 \frac{ie^{i\theta}}{1-e^{i\theta}}\,d\theta = \text{Re}\int_1^\pi \frac{e^{i\theta/2}\,d\theta}{e^{i\theta/2}-e^{-i\theta/2}} = \int_1^\pi \frac{\sin\frac{\theta}{2}}{2\sin\frac{\theta}{2}}\,d\theta = \frac{\pi-1}{2}.$$

The same argument (finally!) proves the pointwise convergence of the Fourier series of $\frac{\pi-x}{2}$ on the interval $(0,\pi)$. Moreover, both the series $\sum_{n\geq 1}\frac{\sin n}{n}$ and the integral $\int_0^{+\infty}\frac{\sin x}{x}\,dx$ can be computed through the properties of the **Fejér kernel**.



THIS ALWAYS BUGGED ME.

The Fejér kernel. In Functional Analysis **density** and **regularization** (or "mollification") tricks are often used. They are based on the fact that the convolution $(f*g)(x)$, defined through

$$(f*g)(x)=\int_{-\infty}^{+\infty}f(\tau)\,g(x-\tau)\,d\tau$$

inherits **the best** regularity between the behaviours of $f(x)$ and $g(x)$. In particular, if $g(x)$ is a non-negative function with unit integral, belonging to C^k and concentrated enough around the origin, $(f*g)(x)$ is an excellent approximation of $f(x)$ that may be way more regular than $f(x)$. Functions $g(x)$ fulfilling the previous constraints are regular approximations of the Dirac δ distribution, and they are said *approximated identities* or **convolution kernels**. The Fejér kernel is a classical example.

Definition 123.

$$\begin{aligned}F_N(x)\stackrel{\text{def}}{=}\sum_{|j|\leq N}\left(1-\frac{|j|}{N}\right)e^{ijx}=\frac{1}{N}\sum_{k=0}^{N-1}\sum_{s=-k}^{k}e^{isx}\quad&=\quad\frac{1}{N}\sum_{k=0}^{N-1}\left(1+2\sum_{s=1}^{k}\cos(sx)\right)\\&=\quad\frac{1}{N}\left(\frac{1-\cos(Nx)}{1-\cos x}\right)=\frac{1}{N}\left(\frac{\sin\frac{Nx}{2}}{\sin\frac{x}{2}}\right)^2.\end{aligned}$$

This particular trigonometric function is non-negative due to the last identity (for short: it is a square). Moreover $\lim_{x\to 0}F_N(x)=N$ and $\int_{-\pi}^{\pi}F_N(x)\,dx=2\pi$ for any $N\geq 1$. By termwise integration,

$$\int_0^1 F_N(x)\,dx=1+\frac{1}{N}\sum_{k=1}^{N-1}\sum_{s=1}^{k}\frac{\sin s}{s}=1+2\sum_{s=1}^{N-1}\frac{\sin s}{s}\left(1-\frac{s}{N}\right).$$

On the other hand

$$\sum_{s=1}^{N}\frac{\sin s}{s}-\sum_{s=1}^{N}\frac{\sin s}{s}\left(1-\frac{s}{N}\right)=\frac{1}{N}\sum_{s=1}^{N}\sin(s)=O\left(\frac{1}{N}\right),$$

so:

$$\sum_{s\geq 1}\frac{\sin s}{s}=\frac{1}{2}\lim_{N\to+\infty}\left(-1+\int_0^1 F_N(x)\,dx\right)=\frac{\pi-1}{2}$$

since $\lim_{N\to+\infty}\int_1^{\pi}F_N(x)\,dx=0$. It is not difficult to locate the stationary points of the function $\frac{\sin(Nx/2)}{\sin(x/2)}$ and state that for any large enough N the ratio $\left(\frac{\sin(Nx/2)}{\sin(x/2)}\right)^2$ is bounded by an absolute constant on the interval $(1,\pi)$, from which it follows that $\int_1^{\pi}F_N(x)\,dx=O\left(\frac{1}{N}\right)$. With the same approach based on termwise integration and by exploiting a Riemann sum we have:

$$\lim_{N\to+\infty}\int_0^{\frac{k}{N}}F_N(x)\,dx=2\cdot\text{Si}(k)=2\int_0^k\frac{\sin x}{x}\,dx$$

hence by considering the limit as $k\to+\infty$ we get:

$$\int_0^{+\infty}\frac{\sin x}{x}\,dx=\frac{\pi}{2}.$$

The pointwise convergence of the Fourier series of the sawtooth wave can also be studied through the Laplace transform. If $f(x)$ is a bounded and vaguely integrable function over $\mathbb{R}^+$, the dominated convergence theorem ensures

$$\lim_{x\to 0^+}f(x)=\lim_{s\to+\infty}\int_0^{+\infty}se^{-sx}f(x)\,dx=\lim_{s\to+\infty}s\cdot\mathcal{L}f(s).$$

If this manipulation (*convolution with an approximate identity*) is applied to $\sum_{n\geq 1}\frac{\sin(nx)}{n}$, it produces:

$$\lim_{x\to 0^+}\sum_{n\geq 1}\frac{\sin(nx)}{n}=\lim_{s\to+\infty}\sum_{n\geq 1}\frac{s}{s^2+n^2}$$

where the last series equals $\frac{-1+\pi s\coth(\pi s)}{2s}$. But even ignoring this identity, Riemann sums grant

$$\lim_{s\to+\infty}\sum_{n\geq 1}\frac{s}{s^2+n^2}=\int_0^{+\infty}\frac{dx}{1+x^2}=\frac{\pi}{2}.$$

Poisson summation formula.

Lemma 124 (Poisson)**.** If $f\in C^2(\mathbb{R})$, $|f(x)|\leq\frac{C}{1+x^2}$ and the first two derivatives of f are integrable on $\mathbb{R}$,

$$\sum_{n=-\infty}^{+\infty}f(n)=\sum_{n=-\infty}^{+\infty}\hat{f}(n),\qquad \hat{f}(\nu)\stackrel{\text{def}}{=}\int_{-\infty}^{+\infty}f(x)e^{-2\pi i\nu x}\,dx.$$

This result has a great importance in Harmonic Analysis and Number Theory. It follows from the fact that the distribution known as *Dirac comb* is a fixed point of the Fourier transform, and the identity

$$\forall a>0,\qquad \sum_{n=-\infty}^{+\infty}\exp(-\pi a n^2+2bin)=\frac{1}{\sqrt{a}}\sum_{n=-\infty}^{+\infty}\exp\left(-\frac{\pi}{a}(n-b)^2\right)$$

leads to the reflection formula for the Riemann ζ function. It is possible to employ the Poisson summation formula to prove interesting identities related to modular forms, like:

$$\sum_{n\geq 1}\frac{\coth(\pi n)}{n^3}=\frac{7\pi^3}{180}$$

or identities already proved, like:

$$-1+2\sum_{n\geq 1}\frac{1}{n^2+1}=\sum_{n\in\mathbb{Z}}\frac{1}{n^2+1}=\sum_{n\in\mathbb{Z}}\pi e^{-2\pi|n|}=\pi\coth(\pi),$$

$$\sum_{n\geq 1}\frac{1}{n^2+(n+1)^2}=-1+\frac{\pi}{2}\tanh\left(\frac{\pi}{2}\right).$$

The Poisson kernel. Since $\log\|z\|=\operatorname{Re}\log z$ and $\int_0^{2\pi}e^{ni\theta}\,d\theta=2\pi\delta(n)$, for any $r\in\mathbb{R}$ we have:

$$\int_0^{2\pi}\log\left\|1-re^{i\theta}\right\|\,d\theta=2\pi\log\max(1,|r|).$$

Such identity can be re-written in terms of the integral of a real function:

$$\forall r\in\mathbb{R},\qquad \int_0^{2\pi}\log(1+r^2-2r\cos\theta)\,d\theta=4\pi\log\max(1,|r|).$$

By differentiating with respect to the r parameter, it follows that:

$$\forall r \in \mathbb{R} \setminus \{-1,+1\}, \qquad \int_0^{2\pi} \frac{r-\cos\theta}{1+r^2-2r\cos\theta}\,d\theta = \begin{cases} \frac{2\pi}{r} & \text{if} \quad |r|>1 \\ 0 & \text{if} \quad |r|<1, \end{cases}$$

$$\forall r \in \mathbb{R} \setminus \{-1,+1\}, \qquad \int_0^{2\pi} \frac{1-r\cos\theta}{1+r^2-2r\cos\theta}\,d\theta = \begin{cases} 2\pi & \text{if} \quad |r|<1 \\ 0 & \text{if} \quad |r|>1. \end{cases}$$

That is not entirely surprising: the last result also follow from the behavior of complex homographies or **Cayley transforms**, since

$$\forall r \in [0,1), \qquad P_r(\theta) = \sum_{n\in\mathbb{Z}} r^{|n|} e^{in\theta} = \text{Re}\left(\frac{1+re^{i\theta}}{1-re^{i\theta}}\right) = \frac{1-r^2}{1+r^2-2r\cos\theta}.$$

Exercise 125. Prove that the function

$$f(r) = \frac{1}{2\pi}\int_0^{2\pi} \frac{(1+3r^2)-(5r+r^3)\cos(\theta)+(1+r^2)\cos(2\theta)}{(1+r^2-2r\cos\theta)^2}\,d\theta$$

equals 1 on the interval $(-1,1)$ and $\frac{1}{r^2}$ outside the previous interval.

The Fourier series of $\log\Gamma$.

$$\log\Gamma(z) = \left(\tfrac{1}{2}-z\right)(\gamma+\log 2) + (1-z)\ln\pi - \frac{1}{2}\log\sin\pi z + \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{\sin 2\pi nz\cdot\log n}{n}, \qquad 0<z<1$$

is an important identity that can be derived through Weierstrass product for the Γ function and the Laplace transform. For a long time it was credited to Ernst Kummer, that proved it in 1847. Only recently Iaroslav Blagouchine has pointed out the same result was known to the Swedish mathematician Carl Johan Malmsten since 1842.

Exercise 126. $\{a_n\}_{n\geq 0}$ is a sequence defined by $a_0 = 0$ and

$$a_{n+1} = \frac{a_n + \sqrt{a_n^2 + \frac{1}{4^n}}}{2}$$

for any $n \geq 0$. Prove that:

$$\lim_{n\to+\infty} a_n \leq \sqrt{\frac{5}{12}}, \qquad \lim_{n\to+\infty} a_n = \frac{2}{\pi}.$$

Exercise 127 (Kronecker's Lemma)**.** A sequence $\{a_n\}_{n\geq 1}$ of real numbers is such that

$$\lim_{N\to+\infty}\sum_{n=1}^{N}\frac{a_n}{n} = C < +\infty.$$

Prove that $\{a_n\}_{n\geq 1}$ is necessarily a sequence with mean zero, i.e.

$$\lim_{N\to+\infty}\frac{a_1+a_2+\ldots+a_N}{N}=0.$$

Proof. Let us set $A_M=\frac{a_1}{1}+\frac{a_2}{2}+\ldots+\frac{a_M}{M}$. By summation by parts:

$$\frac{1}{n}\sum_{k=1}^{n}a_k=\frac{1}{n}\sum_{k=1}^{n}k\cdot\frac{a_k}{k}=A_n-\frac{1}{n}\sum_{k=1}^{n-1}A_k$$

and the convergence of the original series ensures that for any $\varepsilon>0$, there is some N such that $|A_n-C|\leq\varepsilon$ holds for any $n\geq N$. By picking an ε and considering the associated N, the previous RHS can be written as

$$A_n-\frac{1}{n}\sum_{k=1}^{N-1}A_k-\frac{n-N}{n}C-\frac{1}{n}\sum_{k=N}^{n-1}(A_k-C).$$

If now we consider the limit as $n\to+\infty$, the first term goes to C, which cancels with the third term; the second term goes to zero (as the sum is a fixed value) and the last term is bounded in absolute value by $\frac{n-N}{n}\varepsilon\leq\varepsilon$. □

A remark about *nuking mosquitoes*: since the sequence $1,1,1,\ldots$ has not mean zero, the harmonic series is divergent.

Exercise 128. Find a function $f\in C^0(\mathbb{R}^+)$ such that the following equality holds for any $t>0$:

$$f(t)=e^{-3t}+e^{-t}\int_0^t e^{-\tau}f(\tau)d\tau$$

Proof. Let $g(s)=(\mathcal{L}f)(s)$. The Laplace transform of $e^{-x}f(x)$ is given by $g(s+1)$ and the Laplace transform of $\int_0^x e^{-u}f(u)\,du$ is given by $\frac{1}{s}g(s+1)$, hence the given differential equation can be written in terms of g as

$$g(s)=\frac{1}{s+3}+\frac{g(s+2)}{s+1}$$

leading to:

$$g(s)=\frac{1}{s+3}+\frac{1}{s+1}\left(\frac{1}{s+5}+\frac{1}{s+3}\left(\frac{1}{s+7}+\frac{1}{s+5}\left(\frac{1}{s+9}+\ldots\right)\right)\right)$$

$$g(s)=\frac{1}{s+3}+\frac{1}{(s+1)(s+5)}+\frac{1}{(s+1)(s+3)(s+7)}+\frac{1}{(s+1)(s+3)(s+5)(s+9)}+\ldots$$

It is clear that $g(s)$ is a meromorphic function with poles at the negative odd integers. Additionally it is simple to compute the closed form of $R_\xi=\text{Res}_{s=\xi}\,g(s)$ for any $\xi\in\Xi=\{-1,-3,-5,\ldots\}$, then to consider the inverse Laplace transform of g:

$$f(x)=\sum_{\xi\in\Xi}R_\xi e^{\xi x}.$$

The closed form for $g(s)$ is related to the incomplete Γ function:

$$\begin{aligned}g(s)&=\frac{1+2^{\frac{1+s}{2}}\sqrt{e}\left[-\Gamma\left(\frac{3+s}{2}\right)+\Gamma\left(\frac{3+s}{2},\frac{1}{2}\right)\right]}{1+s}\\&=\frac{1}{1+s}\left[1-2^{\frac{1+s}{2}}\sqrt{e}\int_0^{1/2}u^{\frac{s+1}{2}}e^{-u}\,du\right]\\&=\frac{2-\sqrt{e}}{s+1}+\frac{\frac{1}{2}\sqrt{e}}{s+3}-\frac{\frac{1}{8}\sqrt{e}}{s+5}+\frac{\frac{1}{48}\sqrt{e}}{s+7}-\frac{\frac{1}{384}\sqrt{e}}{s+9}+\ldots\end{aligned}$$

and from the last line it is straightforward to recover a solution $f(x)$:

$$\begin{aligned} f(x) &= (2-\sqrt{e})e^{-x} + \sqrt{e}\sum_{n\geq 1}\frac{(-1)^{n+1}}{(2n)!!}e^{-(2n+1)x} \\ &= e^{-x}\left(2 - e^{e^{-x}\sinh(x)}\right). \end{aligned}$$

□

The last entry of this section is a surprising consequence of the Laplace transform and the residue Theorem:

Theorem 129 (Ramanujan's Master Theorem)**.** Under suitable regularity assumptions for ϕ, the identity

$$f(x) = \sum_{n\geq 0}\frac{\phi(n)}{n!}(-x)^n$$

implies:

$$\int_0^{+\infty} x^{s-1} f(x)\,dx = \Gamma(s)\,\phi(-s).$$

In particular, given the series

$$\log\Gamma(1+x) = -\gamma x + \sum_{k\geq 2}\frac{\zeta(k)}{k}(-x)^k$$

following from the Weierstrass product for the Γ function, we have:

$$\forall s\in(0,1),\qquad \int_0^{+\infty}(\gamma x + \log\Gamma(1+x))\frac{dx}{x^{s+2}} = \frac{\pi}{\sin(\pi s)}\cdot\frac{\zeta(1+s)}{1+s}.$$

Exercise 130 (The Russian Integral)**.** Prove that for any $a\in\mathbb{R}^+$ and any $b\in(0,2)$ the following identity holds:

$$\int_0^{+\infty}\frac{x^{-ia}}{x^2+bx+1}\,dx = \frac{2\pi}{\sqrt{4-b^2}}\cdot\frac{\sinh\left(a\arccos\frac{b}{2}\right)}{\sinh(a\pi)}.$$

Proof. We may immediately notice that

$$\mathcal{L}(x^{-ia})(s) = s^{ia-1}\Gamma(1-ia),\qquad \mathcal{L}^{-1}\left(\frac{1}{x^2+bx+1}\right)(s) = \frac{e^{Bs}-e^{\overline{B}s}}{\sqrt{b^2-4}}$$

where B is the root of x^2+bx+1 with a positive imaginary part. By the properties of the Laplace transform, the original integral is converted into

$$\frac{\Gamma(1-ia)}{\sqrt{b^2-4}}\int_0^{+\infty} s^{ia-1}\left(e^{Bs}-e^{\overline{B}s}\right)\,ds$$

which can be evaluated in terms of the Γ function. Due to the reflection formula, the final outcome simplifies into

$$\frac{\left(B^{-ia}-\overline{B}^{-ia}\right)\pi}{\sinh(\pi a)\sqrt{4-b^2}}$$

and we may notice that $B = \exp\left(i\arccos\frac{b}{2}\right)$ allows a further simplification, proving the claim. □

Exercise 131. Find a closed form for $\sum_{n\geq 1}\frac{\cos(2nx)}{2^n}$ and use it to prove the following identity:

$$\int_0^{\pi/2}\frac{d\theta}{1+8\sin^2(\tan\theta)}=\frac{\pi}{6}\cdot\frac{2e^2+1}{2e^2-1}.$$

Remark. By the inverse Laplace transform, for any $p>0$,

$$\eta(p)=\sum_{n\geq 1}\frac{(-1)^{n+1}}{n^p}=\int_0^{+\infty}\frac{s^{p-1}e^{-s}}{\Gamma(p)}\cdot\frac{ds}{1+e^{-s}}=\mathbb{E}\left[\frac{1}{1+\exp(-X)}\right]$$

where X is a random variable with a $\Gamma(p,1)$ distribution. Given the RHS, the inequality $\frac{1}{2}<\eta(p)<1$ is trivial.

The inverse Laplace transform of the central binomial coefficients.
Given the identity

$$\frac{1}{4^n}\binom{2n}{n}=\frac{2}{\pi}\int_0^{\pi/2}\left(\sin^2\theta\right)^n\,d\theta=\frac{1}{\pi}\int_0^{+\infty}\frac{e^{-nx}}{\sqrt{e^x-1}}\,dx$$

we have the equality $\mathcal{L}^{-1}\left(\frac{1}{4^s}\binom{2s}{s}\right)=\frac{1}{\pi\sqrt{e^x-1}}$, hence the asymptotic behaviour of $\frac{1}{4^n}\binom{2n}{n}$ depends on the Maclaurin series of $\sqrt{\frac{x}{e^x-1}}$. On the other hand

$$\frac{x}{e^x-1}=\frac{x}{2}\left(\coth\frac{x}{2}-1\right)=1-\frac{x}{2}-\sum_{n\geq 1}\frac{\zeta(2n)}{(2\pi)^{2n}}(-1)^n x^{2n}$$

implies

$$\log\left(\frac{x}{e^x-1}\right)=-\frac{x}{2}+\sum_{n\geq 1}\frac{\zeta(2n)}{n(2\pi)^{2n}}(-1)^n x^{2n}$$

and

$$\sqrt{\frac{x}{e^x-1}}=\exp\left[-\frac{x}{4}+\sum_{n\geq 1}\frac{\zeta(2n)}{2n(2\pi)^{2n}}(-1)^n x^{2n}\right].$$

From the local approximation

$$\mathcal{L}^{-1}\left(\frac{1}{4^s}\binom{2s}{s}\right)\approx\frac{e^{-x/4}}{\pi\sqrt{x}}\left(1-\frac{x^2}{48}\right)$$

it is straightforward to recover the very accurate asymptotic approximation (which actually is a lower bound)

$$\frac{1}{4^n}\binom{2n}{n}\sim\frac{(8n+1)(8n+3)}{2\sqrt{\pi}(4n+1)^{5/2}}.$$

The evaluation of both sides at $n=2$ produces $\frac{1292}{729}$ as a rational approximation of $\sqrt{\pi}$, whose absolute error is about $1.63\cdot 10^{-4}$. The evaluation at $n=6$ produces the approximation $\frac{60928}{34375}$, whose absolute error is less than $3\cdot 10^{-6}$. An interesting exercise is to check that the approximation produced via $\mathcal{L},\mathcal{L}^{-1}$ outperforms the approximation produced via creative telescoping:

$$\begin{aligned}\left[\frac{1}{4^n}\binom{2n}{n}\right]^2 &= \frac{1}{4}\prod_{k=2}^{n}\left(1-\frac{1}{2k}\right)^2=\frac{1}{4n}\prod_{k=2}^{n}\left(1-\frac{1}{(2k-1)^2}\right)^{-1}=\frac{1}{\pi n}\prod_{k\geq n}\left(1-\frac{1}{(2k+1)^2}\right)\\ &= \frac{1}{\pi n}\exp\sum_{k\geq n}\sum_{m\geq 1}\frac{-1}{m(2k+1)^{2m}}\overset{\text{CT}}{\approx}\frac{1}{\pi n}\prod_{k\geq n}\frac{8k-1}{8k+7}\cdot\frac{8k+9}{8k+1}=\frac{1}{\pi n}\cdot\frac{8n-1}{8n+1}.\end{aligned}$$

On the other hand a refinement of this approach leads to a tighter upper bound:

$$\frac{(8n+1)^2(8n+3)^2}{4\pi(4n+1)^5}\overset{\mathcal{L}}{\leq}\left[\frac{1}{4^n}\binom{2n}{n}\right]^2\overset{\text{CT}}{\leq}\frac{1}{\pi n}\cdot\frac{64n^2-8n+3}{64n^2+8n+3}.$$

5 The Basel problem

The **Basel problem** has been posed by Pietro Mengoli in 1644 and solved by Leonhard Euler in 1735 with a not entirely rigorous argument, fixed in 1741. This section is dedicated to many classical and alternative approaches for showing that:

$$\zeta(2) = \sum_{n\geq 1}\frac{1}{n^2} = \frac{\pi^2}{6}.$$

Proof # 1 (Euler, 1735-1741).
Let us consider the function of complex variable usually denoted as $\text{sinc}(z)$, i.e. the function that equals 1 at the origin and $\frac{\sin z}{z}$ anywhere else. Due to the formula [5] $\sin(z) = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right)$, all the zeroes of the function $\text{sinc}(z)$ belong to the real line, are simple and exactly located at the elements of $\pi\mathbb{Z}\setminus\{0\}$. We may recall that by the Fundamental Theorem of Algebra, any even polynomial $p(z)$ with simple zeroes, such that $p(0) = 1$, can be written in the form

$$p(z) = \prod_{k=1}^{n}\left(1 - \frac{z^2}{\zeta_k^2}\right)$$

where $\pm\zeta_1, \ldots, \pm\zeta_n$ stands for the zeroes of $p(z)$. Given such identity,

$$\sum_{k=1}^{n}\frac{1}{\zeta_k^2} = -[z^2]p(z)$$

holds as a consequence of Vieta's formulas.
By assuming to be allowed to deal with $\text{sinc}(z)$ *as a "polynomial with an infinite degree"*, from the identity

$$\text{sinc}(z) = \prod_{k\geq 1}\left(1 - \frac{z^2}{\pi^2 k^2}\right)$$

it follows that:

$$\zeta(2) = -\pi^2[z^2]\,\text{sinc}(z) = -\pi^2[z^2]\left(1 - \frac{z^2}{6} + \frac{z^4}{120} - \ldots\right) = \frac{\pi^2}{6}.$$

Quoting Sullivan, *you may now sit back and smile smugly at his brilliance.*
However, besides the last proof being absolutely brilliant and incredibly efficient, it is based on a unproven assumption: the statement that we are allowed to deal with $\text{sinc}(z)$ by regarding it as an *infinite-degree polynomial.* Which is true, indeed, since the Weierstrass product for the entire function $\text{sinc}(z)$ has no exponential part, also as a consequence of Mittag-Leffler's Theorem. However this part of Complex Analysis has been developed only in the middle nineteenth century, and it was most certainly unknown to Euler. Such "flaw" was probably the reason for Euler to fix his original proof, based on such *inspired guess.* In 1741, by starting from the property of uniform convergence of the Weierstrass product for the $\text{sinc}(z)$ function, he proved that in a neighbourhood of the origin we have:

$$\log\text{sinc}(\pi z) = \sum_{n\geq 1}\log\left(1 - \frac{z^2}{n^2}\right)$$

uniformly. Once the uniform convergence of derivatives is proved, it follows that

$$\frac{1}{z} - \pi\cot(\pi z) = \sum_{n\geq 1}\frac{2z}{n^2 - z^2}$$

holds for any $z \neq 0$ close enough to the origin. In particular:

$$\zeta(2) = \lim_{z\to 0}\sum_{n\geq 1}\frac{1}{n^2 - z^2} = \frac{1}{2}\lim_{z\to 0}\left(\frac{1}{z^2} - \frac{\pi\cos(\pi z)}{z\sin(\pi z)}\right) = \frac{\pi^2}{6}$$

[5] usually attributed to De Moivre, even if we are pretty sure Newton knew it yet, back in 1676.

and rigour is safe. We may state that considering the logarithmic derivative $\frac{d}{dz}\log(\cdot)$ is a big detour on the original idea, nevertheless Euler's second proof exploits an instrument that few years later will become essential in Complex Analysis (deeply related with the *topological degree* of curves), and is able to show that **all** the values of the ζ function at positive even integers can be computed through the coefficients of the Taylor series of $z\cot z$ at the origin:

$$\frac{1-\pi z\cot(\pi z)}{2}=\sum_{n\geq 1}\zeta(2n)\,z^{2n}.$$

In other terms, Euler's "fixed" proof exhibits a generating function for the sequence $\{\zeta(2n)\}_{n\geq 1}$. To produce deeper results with less than two pages of written math is barely human: this is one of the reasons for which, since the nineteenth century, concise and elementary proofs, able to prove highly non-trivial statements, have been mentioned as *Eulerian*.

We remark that, by following the Eulerian approach of differentiating the logarithm of a Weierstrass product, then exploiting the substitution $z\mapsto iz$, we have:

$$\sum_{n\geq 1}\frac{1}{n^2+z^2}=\frac{-1+\pi z\coth(\pi z)}{2z^2}$$

for any $z\in\mathbb{C}$. Additionally, by differentiating both sides with respect to z we have that:

$$\sum_{n\geq 1}\frac{1}{(n^2+z^2)^2}=\frac{1}{4z^4}\left(-2+\frac{\pi z\cosh(\pi z)}{\sinh(\pi z)}+\frac{\pi^2 z^2}{\sinh^2(\pi z)}\right).$$

Before studying other "classical attacks" to Basel problem, we outline a further proof due to Euler, related to manipulations of the squared arcsine function.

Proof # 2 (Euler, 174?). It is trivial that:

$$\int_0^1\frac{\arcsin(x)}{\sqrt{1-x^2}}\,dx=\frac{1}{2}\arcsin^2(1)=\frac{\pi^2}{8}.$$

Additionally the Taylor series at the origin of $\arcsin(x)$ and $\frac{1}{\sqrt{1-x^2}}$ are well-known. It follows that:

$$\begin{aligned}\frac{\pi^2}{6}=\frac{4}{3}\int_0^1\frac{\arcsin x}{\sqrt{1-x^2}}\,dx&=\frac{4}{3}\int_0^1\frac{x+\sum_{n=1}^{\infty}\frac{(2n-1)!!}{(2n)!!}\frac{x^{2n+1}}{2n+1}}{\sqrt{1-x^2}}\,dx\\&=\frac{4}{3}\int_0^1\frac{x}{\sqrt{1-x^2}}\,dx+\frac{4}{3}\sum_{n=1}^{\infty}\frac{(2n-1)!!}{(2n)!!(2n+1)}\int_0^1 x^{2n}\frac{x}{\sqrt{1-x^2}}\,dx\\&=\frac{4}{3}+\frac{4}{3}\sum_{n=1}^{\infty}\frac{(2n-1)!!}{(2n)!!(2n+1)}\left[\frac{(2n)!!}{(2n+1)!!}\right]\\&=\frac{4}{3}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}\\&=\frac{4}{3}\left(\sum_{n=1}^{\infty}\frac{1}{n^2}-\frac{1}{4}\sum_{n=1}^{\infty}\frac{1}{n^2}\right)\\&=\sum_{n=1}^{\infty}\frac{1}{n^2}=\zeta(2).\end{aligned}$$

We remark we already exploited a similar idea: we showed it is possible to prove the identity

$$\zeta(2)=\sum_{n\geq 1}\frac{3}{n^2\binom{2n}{n}}$$

by creative telescoping or by exploiting the tangent half-angle substitution in a logarithmic integral. Once the coefficients of the Taylor series of $\arcsin^2(x)$ have been computed, one may recognize in the RHS the quantity $\frac{4}{3}\arcsin^2(1)$.

Proof # 3 (Cauchy, 1821) The following proof comes from its author's *Cours d'Analyse*, note VIII.
Let $x \in \left(0, \frac{\pi}{2}\right)$ and let n be an odd natural number. By De Moivre's formula and the definition of cotangent we have:

$$\begin{aligned}\frac{\cos(nx) + i\sin(nx)}{(\sin x)^n} &= \frac{(\cos x + i\sin x)^n}{(\sin x)^n} \\ &= \left(\frac{\cos x + i\sin x}{\sin x}\right)^n \\ &= (\cot x + i)^n.\end{aligned}$$

Due to the binomial Theorem we also have:

$$\begin{aligned}&(\cot x + i)^n \\ &= \binom{n}{0}\cot^n x + \binom{n}{1}(\cot^{n-1} x)i + \ldots + \binom{n}{n-1}(\cot x)i^{n-1} + \binom{n}{n}i^n \\ &= \left[\binom{n}{0}\cot^n x - \binom{n}{2}\cot^{n-2} x \pm \ldots\right] + i\left[\binom{n}{1}\cot^{n-1} x - \binom{n}{3}\cot^{n-3} x \pm \ldots\right].\end{aligned}$$

By comparing the equations above and considering the imaginary parts of the involved terms:

$$\frac{\sin(nx)}{(\sin x)^n} = \left[\binom{n}{1}\cot^{n-1} x - \binom{n}{3}\cot^{n-3} x \pm \ldots\right].$$

Given the last identity, we may fix a positive integer m, set $n = 2m+1$ and define $x_r = \frac{r\pi}{2m+1}$ per $r = 1, 2, \ldots, m$. Since nx_r is an integer multiple of π, we have $\sin(nx_r) = 0$. As a consequence:

$$0 = \binom{2m+1}{1}\cot^{2m} x_r - \binom{2m+1}{3}\cot^{2m-2} x_r \pm \cdots + (-1)^m\binom{2m+1}{2m+1}$$

holds for $r = 1, 2, \ldots, m$. The numbers $x_1, \ldots, x_m$ are distinct elements of the interval $\left(0, \frac{\pi}{2}\right)$ and the function $\cot^2(x)$ is injective on such interval, hence the numbers $t_r = \cot^2 x_r$ (per $r = 1, 2, \ldots, m$) are distinct. Given the previous equation, these m numbers are the roots of the following polynomial with degree m:

$$p(t) = \binom{2m+1}{1}t^m - \binom{2m+1}{3}t^{m-1} \pm \cdots + (-1)^m\binom{2m+1}{2m+1}.$$

Due to Vieta's formulas we may compute the sum of the roots of p through its coefficients:

$$\cot^2 x_1 + \cot^2 x_2 + \cdots + \cot^2 x_m = \frac{\binom{2m+1}{3}}{\binom{2m+1}{1}} = \frac{2m(2m-1)}{6}.$$

Since $\csc^2 x = \cot^2 x + 1$ we further have:

$$\csc^2 x_1 + \csc^2 x_2 + \cdots + \csc^2 x_m = \frac{2m(2m-1)}{6} + m = \frac{2m(2m+2)}{6}.$$

By considering the inequality $\cot^2(x) < \frac{1}{x^2} < \csc^2(x)$ and summing its terms on $x_r = \frac{r\pi}{2m+1}$ we get:

$$\frac{2m(2m-1)}{6} < \left(\frac{2m+1}{\pi}\right)^2 + \left(\frac{2m+1}{2\pi}\right)^2 + \ldots + \left(\frac{2m+1}{m\pi}\right)^2 < \frac{2m(2m+2)}{6}.$$

By multiplying both sides by $\left(\frac{\pi}{2m+1}\right)^2$:

$$\frac{\pi^2}{6}\left(\frac{2m}{2m+1}\right)\left(\frac{2m-1}{2m+1}\right) < \frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{m^2} < \frac{\pi^2}{6}\left(\frac{2m}{2m+1}\right)\left(\frac{2m+2}{2m+1}\right).$$

If we consider the limits of the RHS and LHS as $m \to +\infty$, both limits equal $\frac{\pi^2}{6}$, hence

$$\zeta(2) = \sum_{k\geq 1} \frac{1}{k^2} = \lim_{m\to\infty} \left(\frac{1}{1^2} + \frac{1}{2^2} + \ldots + \frac{1}{m^2}\right) = \frac{\pi^2}{6}$$

follows by (*squeezing*). Cauchy's proof can be shortened a bit by recalling that the Cayley transform $z \mapsto \frac{1-z}{1+z}$ is an involution on $\mathbb{C} \setminus \{-1\}$. In particular a solution of $\left(\frac{1-z}{1+z}\right)^n = e^{i\theta}$ is given by

$$z = \frac{1 - e^{i\theta/n}}{1 + e^{i\theta/n}} = -i \tan \tfrac{\theta}{2n}$$

and the minimal polynomial of $\tan \frac{\pi}{n}$ or $\cot \frac{\pi}{n}$ over $\mathbb{Q}$ is simple to derive. A proof relying on squeezing, but not requiring the explicit construction of the minimal polynomial of $\cot^2 \frac{\pi}{2m+1}$, is presented in *Proofs from The Book*:

Proof # 3 (Aigner, Ziegler, 1998).
Assuming $0 < x < \pi/2$ we have that

$$\frac{1}{\tan^2 x} < \frac{1}{x^2} < \frac{1}{\sin^2 x}$$

and we may notice that $\frac{1}{\tan^2 x} = \frac{1}{\sin^2 x} - 1$. If we partition the interval $\left(0, \frac{\pi}{2}\right)$ in 2^n equal subintervals, then sum both sides of the previous inequality evaluated at $x_k = \frac{\pi}{2} \cdot \frac{k}{2^n}$, we get:

$$\sum_{k=1}^{2^n-1} \frac{1}{\sin^2 x_k} - \sum_{k=1}^{2^n-1} 1 < \sum_{k=1}^{2^n-1} \frac{1}{x_k^2} < \sum_{k=1}^{2^n-1} \frac{1}{\sin^2 x_k}.$$

By denoting the RHS with S_n we may state that:

$$S_n - (2^n - 1) < \sum_{k=1}^{2^n-1} \left(\frac{2 \cdot 2^n}{\pi}\right)^2 \frac{1}{k^2} < S_n.$$

At first sight, an explicit computation of S_n might seem to be difficult, however:

$$\frac{1}{\sin^2 x} + \frac{1}{\sin^2(\frac{\pi}{2} - x)} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x \cdot \sin^2 x} = \frac{4}{\sin^2 2x}.$$

As a consequence, by coupling terms appearing in S_n, except the contribution given by term associated with $\frac{\pi}{4}$, we get 4 times a sum with the same structure, but with a doubled "step" and a halved number of terms. The contribution provided to S_n by the term associated with $\frac{\pi}{4}$ equals $\frac{1}{\sin^2(\pi/4)} = 2$, hence we have the following recurrence formula:

$$S_n = 4S_{n-1} + 2$$

that together with the initial condition $S_1 = 2$ produces the following explicit formula:

$$S_n = \frac{2(4^n - 1)}{3}.$$

As a consequence, the following inequality holds:

$$\frac{2(4^n - 1)}{3} - (2^n - 1) \leq \frac{4^{n+1}}{\pi^2} \sum_{k=1}^{2^n-1} \frac{1}{k^2} \leq \frac{2(4^n - 1)}{3}$$

and by considering the limit as $n \to +\infty$ we reach, like in Cauchy's proof, the wanted identity $\zeta(2) = \frac{\pi^2}{6}$.

Many solutions to Basel problem are based on trigonometric identities and bits of Theory of Hilbert spaces, in particular Parseval's Theorem.

Proof # 4 (Fourier, 1817) On the interval $(0, \pi)$ the following identity holds pointwise:

$$\frac{\pi - x}{2} = \sum_{n\geq 1} \frac{\sin(nx)}{n}.$$

Additionally the convergence is uniform on any compact subinterval. Due to the orthogonality relations

$$\int_0^{2\pi} \sin(nx)\sin(mx)\,dx = \pi\,\delta(m,n)$$

we have that:

$$\begin{aligned}\frac{\pi^3}{6} &= \int_0^{2\pi} \left(\frac{\pi - x}{2}\right)^2 dx \\ &= \sum_{n\geq 1} \int_0^{2\pi} \left(\frac{\sin(nx)}{n}\right)^2 dx \\ &= \sum_{n\geq 1} \frac{\pi}{n^2} = \pi\,\zeta(2).\end{aligned}$$

There are many orthogonal bases of $L^2(0,1)$, equipped with the canonical inner product $\langle f, g\rangle = \int_0^1 f(x)\,g(x)\,dx$, so the above proof admits many variations. For instance, we may consider the Fourier-Legendre expansions of $K(m)$ (the complete elliptic integral of the first kind, having the elliptic modulus as a variable) and $\frac{1}{\sqrt{1-m}}$:

$$K(m) = \sum_{n\geq 0} \frac{2}{2n+1} P_n(2m-1), \qquad \frac{1}{\sqrt{1-m}} = \sum_{n\geq 0} 2 \cdot P_n(2m-1).$$

These expansions lead to $3\,\zeta(2) = \int_0^1 \frac{K(m)}{\sqrt{1-m}}\,dm$. On the other hand, by the Taylor series of K,

$$\int_0^1 \frac{K(m)}{\sqrt{1-m}}\,dm = \frac{\pi}{2} \sum_{n\geq 0} \frac{\binom{2n}{n}^2}{16^n} \int_0^1 \frac{x^n}{\sqrt{1-x}}\,dx = \pi \sum_{n\geq 0} \frac{\binom{2n}{n}}{4^n(2n+1)} = \pi \int_0^1 \frac{dx}{\sqrt{1-x^2}}$$

hence $3\,\zeta(2) = \pi \arcsin(1)$. There is a deep connection between Fourier-analytique proofs of the identity $\zeta(2) = \frac{\pi^2}{6}$ and Euler's proof #2: $f(x) = \arcsin(\sin x)$ *is* a triangle wave.

Proof # 5 (Ritelli at al., 2001)
By combining different ideas from Apostol, Pace and Ritelli, the identity $\zeta(2) = \frac{\pi^2}{6}$ turns out to be a consequence of simple manipulations of a double integral:

$$\begin{aligned}\zeta(2) &= \frac{4}{3} \sum_{n=0}^{+\infty} \frac{1}{(2n+1)^2} = \frac{4}{3} \int_0^1 \frac{\log y}{y^2-1} dy \\ &= \frac{2}{3} \int_0^1 \frac{1}{y^2-1} \left[\log\left(\frac{1+x^2y^2}{1+x^2}\right)\right]_{x=0}^{+\infty} dy \\ &= \frac{4}{3} \int_0^1 \int_0^{+\infty} \frac{x}{(1+x^2)(1+x^2y^2)} dx\,dy \\ &= \frac{4}{3} \int_0^1 \int_0^{+\infty} \frac{dx\,dz}{(1+x^2)(1+z^2)} = \frac{4}{3} \cdot \frac{\pi}{4} \cdot \frac{\pi}{2} = \frac{\pi^2}{6}.\end{aligned}$$

Proof # 6 (D'Aurizio, 2015)
The identity

$$\frac{3}{4}\zeta(2) = \sum_{n\geq 0} \frac{1}{(2n+1)^2} = \frac{\pi}{2} \sum_{k\geq 0} \frac{(-1)^k}{2k+1} = \frac{\pi}{2} \cdot \frac{\pi}{4}$$

holds as a consequence of the residue Theorem. The integral

$$I = \int_{-\infty}^{\infty} e^y \left(\frac{e^y - 1}{y^2} - \frac{1}{y}\right) \frac{1}{e^{2y}+1}\,dy$$

is clearly real, hence the imaginary part of the sum of residues of the integrand function equals zero:

$$I = 2\pi i\left(\left(\frac{1}{\pi}\sum_{n=0}^{\infty}\frac{(-1)^n}{2n+1} - \frac{2}{\pi^2}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^2}\right) - \frac{2i}{\pi^2}\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)^2}\right)$$

This argument first appeared at **this** MSE thread, where the identity

$$\int_0^{+\infty}\left(\frac{x-1}{\log^2 x} - \frac{1}{\log x}\right)\frac{dx}{x^2+1} = \frac{4}{\pi}\sum_{n\geq 0}\frac{(-1)^n}{(2n+1)^2}$$

is also shown. We may find a similar symmetry trick in a further proof due to Euler.

Proof # 7 (Euler, 174?)
If we consider the reflection formula for the Γ function

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$$

and apply the operator $\frac{d^2}{dz^2}\log(\cdot)$ to both sides, we get:

$$\psi'(z) + \psi'(1-z) = \frac{\pi^2}{\sin^2(\pi z)}$$

From which:

$$\pi^2 = 2\,\psi'\left(\frac{1}{2}\right) = 2\sum_{n\geq 1}\frac{1}{\left(n-\frac{1}{2}\right)^2} = 6\,\zeta(2).$$

It is interesting to remark that the previous approach through residues can be encoded in a combinatorial Lemma not making any explicit mention of residues:

Lemma 132. If $\{a_n\}_{n\geq 0}$ is a weakly decreasing sequence of positive numbers and $\sum_{n\geq 0} a_n^2$ is convergent, the following series are also convergent

$$s \stackrel{\text{def}}{=} \sum_{n=0}^{\infty}(-1)^n a_n, \qquad \delta_k \stackrel{\text{def}}{=} \sum_{n=0}^{\infty} a_n a_{n+k}, \qquad \Delta \stackrel{\text{def}}{=} \sum_{k=1}^{\infty}(-1)^{k-1}\delta_k$$

and we have:

$$\sum_{n\geq 0} a_n^2 = s^2 + 2\Delta.$$

Proof # 8 (Knopp, 1950) If we consider the sequence defined through $a_n = \frac{1}{2n+1}$, we have that:

$$s = \sum_{n\geq 0}(-1)^n\frac{1}{2n+1} = \frac{\pi}{4}$$

$$\delta_k = \sum_{n\geq 0}\frac{1}{(2n+1)(2n+2k+1)} = \frac{1}{2k}\sum_{n\geq 0}^{\infty}\left(\frac{1}{2n+1} - \frac{1}{2n+2k+1}\right) = \frac{1}{2k}\left(1 + \frac{1}{3} + \ldots + \frac{1}{2k-1}\right).$$

In particular:

$$\sum_{n\geq 0}\frac{1}{(2n+1)^2} = \left(\frac{\pi}{4}\right)^2 + \sum_{k\geq 1}\frac{(-1)^{k-1}}{k}\left(1 + \frac{1}{3} + \ldots + \frac{1}{2k-1}\right) = \frac{\pi^2}{16} + \frac{\pi^2}{16} = \frac{\pi^2}{8}$$

from which it follows that:

$$\zeta(2) = \frac{4}{3}\sum_{n\geq 0}\frac{1}{(2n+1)^2} = \frac{\pi^2}{6}.$$

A crucial step has been hidden "under the carpet": the equality of blue terms is not entirely trivial. A proof of such equality (not really efficient, but hopefully interesting) has been included in a box at the end of this section. We finish by presenting a proof that follows from the contents of the previous sections in a very straightforward way.

Proof # 9 (MSE, 2016) We proved that from the identity

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

by switching to logarithms and considering Riemann sums, it follows that:

$$\int_0^{\pi/2}\log\sin(\theta)\,d\theta=-\frac{\pi}{2}\log(2),$$

hence by the substitution $\theta\mapsto\frac{\pi}{2}-\theta$ we have:

$$0=\int_0^{\pi/2}\log\left(2\cos\theta\right)\,d\theta.$$

Due to De Moivre's formula, $2\cos\theta=e^{i\theta}+e^{-i\theta}=e^{i\theta}\left(1-e^{-2i\theta}\right)$.
By exploiting the Taylor series of $\log(1+x)$ at the origin and termwise integration we have:

$$\begin{aligned}\int_0^{\pi/2}\operatorname{Im}\log\left(2\cos\theta\right)\,d\theta &= \int_0^{\pi/2}\theta\,d\theta+\operatorname{Im}\sum_{n\geq1}\frac{(-1)^{n+1}}{n}\int_0^{\pi/2}e^{-2ni\theta}\,d\theta\\ &= \frac{\pi^2}{8}-\operatorname{Im}\sum_{n\geq1}\frac{(-1)^{n+1}i\left(1-e^{-\pi in}\right)}{2n^2}\\ &= \frac{\pi^2}{8}-\sum_{n\geq1}\frac{(-1)^{n+1}\left(1-(-1)^n\right)}{2n^2}\end{aligned}$$

but since this integral has to be zero,

$$\frac{\pi^2}{8}=\sum_{m\geq0}\frac{1}{(2m+1)^2}.$$

2017 addendum. **Proof # 10 (MSE, 2017)** Here it is another crazy approach *by symmetry*. For any $s>1$ we have

$$\zeta(s)=\sum_{n\geq1}\frac{1}{n^s}=\left(1-\frac{2}{2^s}\right)^{-1}\sum_{n\geq1}\frac{(-1)^{n+1}}{n^s}\stackrel{\mathcal{L}^{-1}}{=}\frac{1}{\Gamma(s)}\left(1-\frac{2}{2^s}\right)^{-1}\int_0^{+\infty}\frac{t^{s-1}}{e^t+1}\,dt$$

where the RHS is converging for any $s>0$, providing an analytic continuation of the LHS over such region. By applying integration by parts twice, we get the following integral representation for the ζ function over the region $s>-2$:

$$\zeta(s)=\frac{1}{\Gamma(s+2)}\left(1-\frac{2}{2^s}\right)^{-1}\int_0^{+\infty}\frac{t^{s+1}e^t(e^t-1)}{(e^t+1)^3}\,dt$$

and due to the reflection formula $\frac{\zeta(1-s)}{\zeta(s)}=\frac{2\,\Gamma(s)}{(2\pi)^s}\cos\left(\frac{\pi s}{2}\right)$ we have

$$\zeta(s)=\frac{(2\pi)^s}{2\,\Gamma(3-s)\,\Gamma(s)\cos\frac{\pi s}{2}}\left(1-2^s\right)^{-1}\int_0^{+\infty}\frac{t^{2-s}e^t(e^t-1)}{(e^t+1)^3}\,dt$$

for any $s<3$. By evaluating the previous line at $s=2$ and by enforcing the substitution $t=\log u$ we get:

$$\zeta(2)=\frac{2\pi^2}{3}\int_1^{+\infty}\frac{u-1}{(u+1)^3}\,du\stackrel{u\mapsto\frac{1}{1-v}}{=}\frac{2\pi^2}{3}\int_0^1\frac{v}{(2-v)^3}\,dv=\frac{2\pi^2}{3}\cdot\frac{1}{4}=\frac{\pi^2}{6}.$$

In equivalent terms, the value of $\zeta(2)$ can be derived from the value of $\zeta(-1)$, which on its turn is related to a Bernoulli number.

2018 addendum. **Proof # 11**. We start with the statement of the **MacMahon master theorem**.
Let $A = (a_{ij})_{m\times m}$ be a complex matrix, and let $x_1, \ldots, x_m$ be formal variables. Consider a coefficient

$$G(k_1, \ldots, k_m) = \left[x_1^{k_1} \cdots x_m^{k_m}\right] \prod_{i=1}^{m} \left(a_{i1}x_1 + \cdots + a_{im}x_m\right)^{k_i}.$$

Let $t_1, \ldots, t_m$ be another set of formal variables, and let $T = (\delta_{ij}t_i)_{m\times m}$ be a diagonal matrix. Then

$$\sum_{(k_1,\ldots,k_m)} G(k_1, \ldots, k_m)\, t_1^{k_1} \cdots t_m^{k_m} = \frac{1}{\det(I_m - TA)},$$

where the sum runs over all nonnegative integer vectors $(k_1, \ldots, k_m)$, and I_m denotes the identity matrix of size m. By considering the matrix

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}$$

it is not difficult to derive an important combinatorial identity due to Dixon:

$$\sum_{k\in\mathbb{Z}} (-1)^k \binom{a+b}{a+k}\binom{b+c}{b+k}\binom{c+a}{c+k} = \frac{(a+b+c)!}{a!b!c!}$$

holding for any $a, b, c \in \mathbb{N}^+$. By logarithmic convexity and the Bohr-Mollerup theorem, the range of validity of Dixon's identity can be extended to $a, b, c \in (-1, +\infty)$. If we consider the instance $a = b = c = \frac{1}{2}$ we get:

$$\sum_{k\in\mathbb{Z}} \frac{1}{(1-4k^2)^3} = \frac{3\pi^2}{32},$$

where the LHS, by partial fraction decomposition, can be written as:

$$\sum_{k\in\mathbb{Z}} \left[\frac{1}{8}\cdot\frac{1}{(2k+1)^3} - \frac{1}{8}\cdot\frac{1}{(2k-1)^3} + \frac{3}{16}\cdot\frac{1}{(2k+1)^2} + \frac{3}{16}\cdot\frac{1}{(2k-1)^2} + \frac{3}{16}\cdot\frac{1}{2k+1} - \frac{3}{16}\cdot\frac{1}{2k-1}\right]$$

which by cancellation and symmetry boils down to $\frac{3}{4}\sum_{k\geq 0}\frac{1}{(2k+1)^2}$.
The identity $\zeta(2) = \frac{\pi^2}{6}$ is a straightforward consequence.

Proof # 12. We may consider that $\mathcal{J} = \int_0^{+\infty} \frac{\arctan x}{1+x^2}\,dx = \left[\frac{1}{2}\arctan^2 x\right]_0^{+\infty} = \frac{\pi^2}{8}$.
On the other hand, by Feynman's trick or Fubini's theorem

$$\mathcal{J} = \int_0^{+\infty}\int_0^1 \frac{x}{(1+x^2)(1+a^2x^2)}\,da\,dx = \int_0^1 \frac{-\log a}{1-a^2}\,da$$

and since $\int_0^1 -\log(x)x^n\,dx = \frac{1}{(n+1)^2}$, by expanding $\frac{1}{1-a^2}$ as a geometric series we have

$$\frac{\pi^2}{8} = \mathcal{J} = \sum_{n\geq 0}\frac{1}{(2n+1)^2}.$$

Exercise 133. In order to stress the importance of the dominated convergence Theorem, prove that:

$$\sum_{n\geq 1}\sum_{m\geq 1}\frac{m^2-n^2}{(m^2+n^2)^2}=\frac{\pi}{4},\qquad \sum_{m\geq 1}\sum_{n\geq 1}\frac{m^2-n^2}{(m^2+n^2)^2}=-\frac{\pi}{4}.$$

It is possible to tackle the given problem through

$$\sum_{m\geq 1}\frac{m^2-n^2}{(m^2+n^2)^2}=\frac{1}{2}\left(\frac{1}{n^2}-\frac{\pi^2}{\sinh^2(\pi n)}\right)$$

and by exploiting the identity in the next exercise, or by exploiting Poisson summation formula.

Exercise 134 (First steps towards modular forms)**.** Prove the identity:

$$\sum_{n\geq 1}\frac{1}{\sinh^2(\pi n)}=\frac{1}{6}-\frac{1}{2\pi}.$$

Hint: apply the $\frac{d^2}{dz^2}\log(\cdot)$ operator to both sides of the equality representing the Weierstrass product for the $\frac{\sinh z}{z}$ function.

Exercise 135 (First steps towards modular forms)**.** Prove the identity:

$$\sum_{n\geq 1}\frac{n(-1)^{n+1}}{\sinh(\pi n)}=\frac{1}{4\pi}.$$

Exercise 136 (Ramanujan's first steps towards modular forms)**.** Prove the identity:

$$\frac{1^{13}}{e^{2\pi}-1}+\frac{2^{13}}{e^{4\pi}-1}+\frac{3^{13}}{e^{6\pi}-1}+\ldots=\frac{1}{24}.$$

The identity under the carpet. As promised, we are going to prove that:

$$\sum_{k\geq 1}\frac{(-1)^{k+1}}{k}\left(1+\frac{1}{3}+\ldots+\frac{1}{2k-1}\right)=\frac{\pi^2}{16}.$$

According to the terminology introduced by Gioffrè, Iandoli and Scandone, the following proof is a "*level* 4 *proof*", since at some point four consecutive symbols for series or integrals appear:

$$\begin{aligned}
\sum_{k\geq 1}\frac{(-1)^{k+1}}{k}\left(1+\frac{1}{3}+\ldots+\frac{1}{2k-1}\right) &= \sum_{k\geq 1}(-1)^{k+1}\sum_{n\geq 0}\frac{2}{(2n+1)(2n+2k+1)}\\
&= 2\int_0^1\int_0^1\sum_{k\geq 1}\sum_{n\geq 0}(-1)^{k+1}x^{2n}y^{2n+2k}\,dx\,dy\\
&= 2\iint_{(0,1)^2}\frac{y^2\,dx\,dy}{(1+y^2)(1-x^2y^2)}\\
&= 2\iint_{(0,1)^2}\frac{dx\,dy}{1-x^2y^2}-2\iint_{(0,1)^2}\frac{dx\,dy}{(1+y^2)(1-x^2y^2)}\\
&= 2\sum_{n\geq 0}\frac{1}{(2n+1)^2}-2\int_0^1\frac{\text{arctanh}(y)}{y(y^2+1)}\,dy\\
(y\mapsto\tanh u)\qquad &= 2\sum_{n\geq 0}\frac{1}{(2n+1)^2}-2\int_0^{+\infty}\frac{u\coth(u)\,du}{\cosh(2u)}\\
(u\mapsto -\log z)\qquad &= 2\sum_{n\geq 0}\frac{1}{(2n+1)^2}+4\int_0^1\frac{z(1+z^2)^2\log z}{1-z^8}\,dz\\
(\text{Feynman's trick})\qquad &= 2\sum_{n\geq 0}\frac{1}{(2n+1)^2}-4\sum_{k\geq 0}\left(\frac{1}{(8k+2)^2}+\frac{1}{(8k+6)^2}+\frac{2}{(8k+4)^2}\right)\\
&= \frac{1}{2}\sum_{n\geq 0}\frac{1}{(2n+1)^2}=\tfrac{3}{8}\zeta(2).
\end{aligned}$$

In a simpler way, we may set $\mathscr{H}_n=\sum_{k=0}^{n}\frac{1}{2k+1}$ and notice that:

$$\sum_{n\geq 0}\mathscr{H}_n x^{2n+1}=\frac{\text{arctanh}\,(x)}{1-x^2},\qquad \sum_{n\geq 0}(-1)^n\mathscr{H}_n x^{2n+1}=\frac{\arctan x}{1+x^2},$$

from which it immediately follows that:

$$\sum_{n\geq 0}\frac{(-1)^n}{n+1}\mathscr{H}_n=2\int_0^1\frac{\arctan x}{1+x^2}\,dx=\arctan^2(1).$$

Exercise 137. Prove that:

$$\int_{-\infty}^{+\infty}\frac{dx}{(1+x^2)\cosh^2(\pi x)}=\pi-\frac{8}{\pi}.$$

Exercise 138. Prove the following identity:

$$\sum_{m\geq 1}\sum_{n\geq 1}\frac{1}{m^2n^2(m+n)^2}=\frac{\pi^6}{2835}.$$

6 Special functions and special products

This section is dedicated to an introduction the Γ function and its properties. We start by investigating about the interplay among values of the Riemann ζ function, Bernoulli numbers and power series. We recall that, by considering the logarithmic derivative of the Weierstrass product for the sine function (*Euler docet*):

$$\begin{aligned}\frac{1}{2z}-\frac{\pi}{2}\cot(\pi z) &= z\sum_{n\geq 1}\frac{1}{n^2}\sum_{h\geq 0}\frac{z^{2h}}{n^{2h+2}}=\sum_{h\geq 0}z^{2h+1}\zeta(2h+2)=\sum_{k\geq 1}\zeta(2k)z^{2k-1}\\ \zeta(2k) &= \frac{1}{(2k-1)!}\left[\frac{d^{2k-1}}{dz^{2k-1}}\left(\frac{1}{2z}-\frac{\pi}{2}\cot(\pi z)\right)\right]_{z=0}\end{aligned}$$

By introducing the **Bernoulli numbers** B_n through their (exponential[6]) generating function

$$\frac{z}{e^z-1}=\sum_{n\geq 0}\frac{B_n}{n!}z^n$$

we immediately have that every Bernoulli number with odd index equals zero, with the only exception of $B_1=-\frac{1}{2}$. This follows from the fact that

$$\frac{z}{e^z-1}+\frac{z}{2}=\frac{z}{2}\coth\frac{z}{2}$$

is clearly an odd function. Additionally, since coth is the logarithmic derivative of

$$\sinh(z)=z\prod_{n\geq 1}\left(1+\frac{z^2}{\pi^2 n^2}\right),$$

Bernoulli numbers with an even index, $B_0=1, B_2=\frac{1}{2}, B_4=-\frac{1}{30}, B_6=\frac{1}{42}, B_8=-\frac{1}{30}, B_{10}=\frac{5}{66}, B_{12}=-\frac{691}{2730},\ldots$, have alternating signs from B_2 onward. Due to the ordinary generating function for $\{\zeta(2n)\}_{n\geq 1}$ introduced in the Basel problem section,

$$\zeta(2k)=\frac{(2\pi)^{2k}|B_{2k}|}{2\cdot(2k)!}\in\pi^{2k}\mathbb{Q}.$$

Since for any $k\geq 1$ we have $1\leq\zeta(2k)\leq\zeta(2)$, the previous formula allows a simple estimation of the magnitude of $|B_n|$. We may further notice that:

$$\begin{aligned}1=\frac{e^x-1}{x}\cdot\frac{x}{e^x-1} &= \left(\sum_{n\geq 1}\frac{1}{(n+1)!}x^n\right)\cdot\left(\sum_{n\geq 0}\frac{B_n}{n!}x^n\right)\\ &= \sum_{n\geq 0}\left(\sum_{k=0}^{n}\frac{B_k}{k!(n+1-k)!}\right)x^n\\ &= \sum_{n\geq 1}\left(\sum_{k=0}^{n}\binom{n+1}{k}B_k\right)\frac{x^n}{(n+1)!}\end{aligned}$$

implies that for any $n>1$ we have:

[6] According to the usual terminology, the *ordinary* and *exponential* generating functions for a sequence of complex numbers $\{a_n\}_{n\geq 0}$ are respectively given by

$$\text{(OGF)}\ \sum_{n\geq 0}a_n z^n,\qquad \text{(EGF)}\ \sum_{n\geq 0}\frac{a_n}{n!}z^n.$$

$$B_n = -\frac{1}{n+1}\sum_{k=0}^{n-1}\binom{n+1}{k}B_k,$$

an identity allowing the evaluation of Bernoulli numbers in a recursive fashion. Due to such relation, Bernoulli numbers play a major role in problems related to (finite) power sums:

Theorem 139 (Faulhaber's formula)**.** For any natural number p,

$$\sum_{k=1}^{n} k^p = \frac{1}{p+1}\sum_{j=0}^{p}(-1)^j\binom{p+1}{j}B_j n^{p+1-j}.$$

Since for any $k \geq 1$ we have $B_k = -k\,\zeta(1-k)$, the previous statement can also be written as:

$$\sum_{k=1}^{n} k^p = \frac{n^{p+1}}{p+1} - \sum_{j=0}^{p-1}\binom{p}{j}\zeta(-j)n^{p-j}$$

from which it follows that $\zeta(s)$ for $s < 1$ can be computed through the *regularization* of a divergent series. [7]For instance we proved that the sequence given by $a_n = 2\sqrt{n} - H_n^{(1/2)}$ is convergent as $n \to +\infty$. Due to Faulhaber's formula, we may further state that:

$$\lim_{n\to+\infty}\left(2\sqrt{n} - H_n^{(1/2)}\right) = (1+\sqrt{2})\sum_{n\geq 1}\frac{(-1)^{n+1}}{\sqrt{n}} = -\zeta\left(\tfrac{1}{2}\right)$$

then use the (inverse) Laplace transform to get an integral representation for $-\zeta\left(\frac{1}{2}\right)$:

$$-\zeta\left(\tfrac{1}{2}\right) = (1+\sqrt{2})\int_0^{+\infty}\sum_{n\geq 1}\frac{(-1)^{n+1}e^{-ns}}{\sqrt{\pi s}}\,ds = \frac{2+2\sqrt{2}}{\sqrt{\pi}}\int_0^{+\infty}\frac{ds}{1+\exp(s^2)}.$$

A *de facto* equivalent technique is to apply Ramanujan's Master Theorem to the exponential generating function of Bernoulli numbers.

The recurrence relation fulfilled by Bernoulli numbers, together with Lucas Theorem on the behaviour of binomial coefficients (mod p), leads to an interesting consequence: the denominator of the rational number B_{2k} is a squarefree integer, given by the product of primes p such that $p-1$ is a divisor of $2k$:

Theorem 140 (Von Staudt-Clausen)**.** By denoting as $\mathcal{P}$ the set of prime numbers,

$$\forall n \geq 1, \qquad B_{2n} + \sum_{\substack{p\in\mathcal{P}\\(p-1)|(2n)}}\frac{1}{p} \in \mathbb{Z}.$$

The core of this section is summarized in the following four points.

[7]The famous "claim" $\sum_{n\geq 1} n = -\frac{1}{12}$ is preposterous since the LHS is a divergent series. $\zeta(-1) = -\frac{1}{12}$ holds, but $\zeta(s)$ for $s < 1$ is defined through an analytic continuation and not directly through the series $\sum_{n\geq 1}\frac{1}{n^s}$, which is convergent only if $\text{Re}(s) > 1$.

Theorem 141. If s is a complex number with a positive real part, the following definitions are equivalent:

- **(A)**(Integral representation)
$$\Gamma(s)=\int_{0}^{+\infty}x^{s-1}e^{-x}\,dx$$
- **(B)**(Euler product)
$$\Gamma(s)=\lim_{n\to+\infty}\frac{n!n^s}{s(s+1)\cdot\ldots\cdot(s+n)}$$
- **(C)**(Weierstrass product)
$$\Gamma(s)=\frac{e^{-\gamma s}}{s}\prod_{n\geq 1}\left(1+\frac{s}{n}\right)^{-1}e^{s/n},\qquad \gamma=\lim_{n\to+\infty}\left(H_n-\log n\right)$$
- **(D)**(Bohr-Mollerup characterization) On the positive real semiaxis, there is a unique function with a convex logarithm that fulfills $\Gamma(1)=1$ and $\Gamma(s+1)=s\,\Gamma(s)$ for any $s>0$.

Sketch of proof. **(A)** $\leftrightarrow$ **(B)**. Due to the integration by parts formula, the function defined by **(A)** fulfills $\Gamma(1)=1$ and $\Gamma(s+1)=s\,\Gamma(s)$. Due to the dominated convergence Theorem we have:

$$\int_{0}^{+\infty}x^{s-1}e^{-x}\,dx=\lim_{n\to+\infty}\int_{0}^{n}x^{s-1}\left(1-\frac{x}{n}\right)^{n}\,dx$$

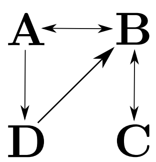
where the integral appearing in the RHS can be computed by integration by parts too, proving **(B)** and **(B)** $\to$ **(A)**.
(B) $\leftrightarrow$ **(C)** follows from simple algebraic manipulations, since any $n\in\mathbb{Z}^+$ can be written as a telescopic product:

$$n=\prod_{k=1}^{n-1}\left(1+\frac{1}{k}\right).$$

(A) $\to$ **(D)**. If s is a positive real number, from (A) it follows that $\Gamma(s)$ is a **moment** for a positive random variable, i.e. an integral of the form $\int_{\mathbb{R}^+}x^s\omega(x)\,dx$ with $\omega(x)$ being a positive and locally integrable function. In particular, $\Gamma(s)$ is a continuous positive function and due to the Cauchy-Schwarz inequality:

$$\Gamma(s_1)\Gamma(s_2)=\int_{0}^{+\infty}x^{s_1}\frac{dx}{xe^x}\int_{0}^{+\infty}x^{s_2}\frac{dx}{xe^x}\geq\left(\int_{0}^{+\infty}x^{\frac{s_1+s_2}{2}}\frac{dx}{xe^x}\right)^2=\Gamma\left(\frac{s_1+s_2}{2}\right)^2.$$

It follows that $\log\Gamma$ is a *midpoint-convex* function, and since it is a continuous function on $\mathbb{R}^+$, it is *tout court* **convex**.
(D) $\to$ **(B)**. Due to logarithmic convexity, the function of real variable defined through **(D)** has a representation as an infinite product and it is not only continuous, but analytic. Due to the **analytic continuation principle**, the function defined by **(D)** can be extended to the whole half-plane $\mathrm{Re}(s)>0$ and the statement **(B)** holds in that region too. Summarizing:



We may also notice that once the $\Gamma(s)$ function is defined on the half-plane $\mathrm{Re}(s)>0$, the functional identity $\Gamma(s+1)=s\,\Gamma(s)$ allows to extend Γ to the whole complex plane. In particular, from the Weierstrass product it follows that:

- $\Gamma(s)$ is a meromorphic function on $\mathbb{C}$, not vanishing at any point;

- the singularities of $\Gamma(s)$ are simple poles and they are located at the elements of $\{0\} \cup \mathbb{Z}^-$. Additionally:

$$\forall n \in \mathbb{N}, \qquad \operatorname{Res}\left(\Gamma(s), s=-n\right) = \frac{(-1)^n}{n!}$$

- the Γ function fulfills the following **reflection formula**:

$$\forall z \notin \mathbb{Z}, \qquad \Gamma(z)\,\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}.$$

The logarithmic convexity allows to produce tight approximations of the Γ function, like:

Theorem 142 (Gautschi's inequality)**.** For any real number $x > 0$ and for any $s \in (0,1)$ we have:

$$x^{1-s} < \frac{\Gamma(x+1)}{\Gamma(x+s)} < (x+1)^{1-s}.$$

To have extended the factorial function allows us to define binomial coefficients also for non-integer parameters. In particular, from $n! = \Gamma(n+1)$ it follows that:

$$\binom{2n}{n} = \frac{(2n)!}{n!^2} = \frac{\Gamma(2n+1)}{\Gamma(n+1)^2}.$$

Given these important notions, we are ready to study the behavior of central binomial coefficients.
We may start by noticing that for any $n \in \mathbb{Z}^+$,

$$\frac{1}{4^n}\binom{2n}{n} = \frac{(2n)!}{(2n)!!^2} = \frac{(2n-1)!!}{(2n)!!} = \prod_{k=1}^{n}\left(1-\frac{1}{2k}\right).$$

By squaring both sides and performing some algebraic manipulations:

$$\left[\frac{1}{4^n}\binom{2n}{n}\right]^2 = \frac{1}{4}\prod_{k=2}^{n}\left(1-\frac{1}{k}+\frac{1}{4k^2}\right) = \frac{1}{4}\prod_{k=2}^{n}\left(1-\frac{1}{k}\right)\prod_{k=2}^{n}\left(1+\frac{1}{4k(k-1)}\right) = \frac{1}{4n}\prod_{k=2}^{n}\left(1-\frac{1}{(2k-1)^2}\right)^{-1}.$$

If now we denote as W the infinite product $\prod_{k\geq 2}\left(1-\frac{1}{(2k-1)^2}\right)^{-1}$ we have:

$$\left[\frac{1}{4^n}\binom{2n}{n}\right]^2 = \frac{W}{4n}\prod_{k>n}\left(1-\frac{1}{(2k-1)^2}\right) = \frac{W}{4n}\prod_{k>n}\left(1+\frac{1}{4k(k-1)}\right)^{-1}.$$

Since in a right neighbourhood of the origin we have $e^x > 1+x$ (truth to be told, such inequality holds for any $x \in \mathbb{R}^*$ by convexity), we may state:

$$\left[\frac{1}{4^n}\binom{2n}{n}\right]^2 > \frac{W}{4n}\exp\sum_{k>n}\frac{-1}{4k(k-1)} = \frac{W}{4n}\exp\left(-\frac{1}{4n}\right)$$

as well as:

$$\left[\frac{1}{4^n}\binom{2n}{n}\right]^2 < \frac{W}{4n}\exp\sum_{k>n}\frac{-1}{(2k-1)^2} = \frac{W}{4n}\exp\left(-\frac{1}{4n+2}\right)$$

by creative telescoping. Essentially, it is enough to find an explicit value for W (also known as **Wallis product**) to have accurate approximations of central binomial coefficients. Due to the Weierstrass product for the cosine function,

$$\cos(\pi z) = \prod_{k\geq 0}\left(1-\frac{4z^2}{(2k+1)^2}\right)$$

we have:

$$\frac{1}{W} = \prod_{k\geq 1}\left(1-\frac{1}{(2k+1)^2}\right) = \lim_{z\to 1/2}\frac{\cos(\pi z)}{1-4z^2} \stackrel{\text{d.H.}}{=} \frac{\pi}{4},$$

hence:

$$\frac{1}{4^n}\binom{2n}{n} = \frac{1}{\sqrt{\pi n}}\left(1-\frac{1}{8n}+O\left(\frac{1}{n^2}\right)\right)$$

where the asymptotic expansion still holds if $n \geq 1$ is not an integer. By the reflection formula we clearly have $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$, leading to a remarkable consequence:

Lemma 143.

$$\int_0^{+\infty} e^{-x^2}\,dx = \int_0^{+\infty}\frac{e^{-t}}{2\sqrt{t}}\,dt \overset{\textbf{(A)}}{=} \frac{1}{2}\,\Gamma\left(\tfrac{1}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

In this framework we will see soon that the area of the unit circle being equal to $\Gamma\left(\frac{1}{2}\right)^2$ is not accidental at all. Before introducing the Beta function and the multiplication formula for the Γ function, let us investigate about the Eulerian thought "*if something is defined by an infinite product, it might be the case to consider its logarithmic derivative* ". For any complex number with a positive real part, let us set:

$$\psi(s) \overset{\text{def}}{=} \frac{d}{ds}\log\Gamma(s).$$

The ψ function is also known as **Digamma** function. Due to the Weierstrass product for the Γ function, we have:

$$\psi(s)=\frac{\Gamma'(s)}{\Gamma(s)}=\frac{d}{ds}\left(-\log s-\gamma s+\sum_{k\geq 1}\left(\frac{s}{k}-\log\left(1+\frac{s}{k}\right)\right)\right)=-\frac{1}{s}-\gamma+\sum_{k\geq 1}\left(\frac{1}{k}-\frac{1}{k+s}\right)$$

where the function relation $\Gamma(s+1)=s\,\Gamma(s)$ translates into $\psi(s+1)=\frac{1}{s}+\psi(s)$ for the Digamma function. Such identity allows an analytic continuation of the ψ function to the whole complex plane. The Digamma function turns out to be a meromorphic function with simple poles with residue -1 at each non-positive integer. Far enough from the singularities, the following identities hold:

Main properties of the Digamma function

$$\begin{cases}
\psi(b)-\psi(a) &=& \sum_{n\geq 0}\left(\frac{1}{(n+a)}-\frac{1}{(n+b)}\right)\\
\frac{\psi(b)-\psi(a)}{b-a} &=& \sum_{n\geq 0}\frac{1}{(n+a)(n+b)}\\
\sum_k \psi(-a_k)\text{Res}_{z=a_k}\left(\frac{1}{\prod_k(z-a_k)}\right) &=& \sum_{n\geq 0}\frac{1}{\prod_k(n-a_k)}\\
\sum_i \frac{(-1)^{p_i}}{p_i!}\frac{\psi^{(p_i-1)}(-\zeta_i)}{\prod_{j\neq i}(\zeta_i-\zeta_j)} &=& \sum_{n\geq 0}\frac{1}{\prod_i(n-\zeta_i)^{p_i}}\\
\psi'(a) &=& \sum_{n\geq 0}\frac{1}{(n+a)^2}\\
\frac{\psi^{(n)}(a)}{n!} &=& (-1)^{n+1}\sum_{k\geq 0}\frac{1}{(k+a)^{n+1}}\\
\psi(z)-\psi(1-z) &=& \pi\cot(\pi z)\\
\psi(nz) &=& \log(n)+\frac{1}{n}\sum_{k=0}^{n-1}\psi\left(z+\frac{k}{n}\right)\\
\psi(1/2),\ \psi(1),\ \psi(n) &=& -\gamma-\log 4,\ -\gamma,\ H_{n-1}-\gamma.
\end{cases}$$

They are clearly interesting from a combinatorial point of view. On the half-plane $\text{Re}(s)>0$ we also have the following integral representations:

$$\psi(s)=\int_0^{+\infty}\left(\frac{e^{-t}}{t}-\frac{e^{-st}}{1-e^{-t}}\right)dt,\qquad \psi(s+1)=-\gamma+\int_0^1\frac{1-x^s}{1-x}\,dx$$

and the Taylor series of $\psi(x+1)$ at the origin is really simple:

$$\psi(x+1) = -\gamma - \sum_{k\geq 1} \zeta(k+1)(-z)^k$$

just like the asymptotic expansion for $x \gg 1$:

$$\psi(x) = \log(x) - \frac{1}{2x} - \sum_{n\geq 1} \frac{B_{2n}}{2nx^{2n}}.$$

The last identity is a typical consequence of the **Euler-McLaurin summation formula** or, as we will see, of techniques based on creative telescoping. We remark that by combining the discrete Fourier transform with the multiplication and reflection formulas for the Γ and ψ functions we get a deep result:

Theorem 144 (Gauss Digamma Theorem)**.** If r, m are positive integers and $r < m$, we have:

$$\psi\left(\frac{r}{m}\right) = -\gamma - \log(2m) - \frac{\pi}{2}\cot\frac{\pi r}{m} + 2\sum_{n=1}^{\lfloor \frac{m-1}{2} \rfloor} \cos\left(\frac{2\pi n r}{m}\right) \log\sin\left(\frac{\pi n}{m}\right)$$

allowing an explicit evaluation of $\psi(s)$ for any $s \in \mathbb{Q}$. A proof can be found **on PlanetMath**.

The multiplication formulas for the Γ function follow from the multiplication formulas for the ψ function in a very straightforward way. Since:

$$\begin{aligned}\frac{\log\Gamma(ns) - \log\Gamma(n)}{n} = \int_1^s \psi(nz)\,dz &= (s-1)\log n + \sum_{k=0}^{n-1}\int_1^s \psi\left(z + \frac{k}{n}\right)dz \\ &= (s-1)\log n + \sum_{k=0}^{n-1}\log\Gamma\left(s+\frac{k}{n}\right) - \sum_{k=0}^{n-1}\log\Gamma\left(1+\frac{k}{n}\right)\end{aligned}$$

by exponentiating both sides we get another result due to Gauss:

Theorem 145 (Multiplication formulas for the Γ function)**.** For any $s \in \mathbb{C}$ with positive real part and for any $n \in \mathbb{Z}^+$,

$$\Gamma(ns) = (2\pi)^{\frac{1-n}{2}} n^{ns-1/2} \prod_{k=0}^{n-1} \Gamma\left(s + \frac{k}{n}\right).$$

The $n = 2$ case is also known as **Legendre duplication formula**:

$$\Gamma(2z) = \frac{2^{2z-1}}{\sqrt{\pi}}\Gamma(z)\Gamma\left(z + \frac{1}{2}\right).$$

These identities can also be proved through a slick technique known as **Herglotz trick**: if two meromorphic functions $f(z), g(z)$ only have simple poles at the same points with the same residues, they share a value at a regularity point and they fulfill the same functional equation, they are the same function. In particular, for any $z \in \mathbb{C}$ with positive real part we may define

$$g(z) \stackrel{\text{def}}{=} \frac{\Gamma(2z)}{\Gamma(z)\Gamma(z+1/2)}$$

and check that $g(z)$ is an entire function such that $g(1) = \frac{2}{\sqrt{\pi}}$ and $g(z+1) = 4\,g(z)$. Legendre duplication formula and, in a similar fashion, Gauss multiplication formulas are so straightforward consequences.
Back to central binomial coefficients, we may notice that:

$$\frac{1}{4^n}\binom{2n}{n} = \frac{\Gamma(2n+1)}{4^n\Gamma(n+1)^2} = \frac{\Gamma(2n)}{2^{2n-1}\Gamma(n+1)^2} = \frac{\Gamma\left(n+\frac{1}{2}\right)}{\sqrt{\pi}\,\Gamma(n+1)}$$

hence the identity $\prod_{k\geq 1}\left(1-\frac{1}{(2k+1)^2}\right) = \frac{\pi}{4}$ **is equivalent** to the identity $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$, and loosely speaking:

$$\text{Wallis product} \quad \longleftrightarrow \quad \Gamma\left(\tfrac{1}{2}\right) = \sqrt{\pi} \quad \longleftrightarrow \quad \text{Integral of the Gaussian function.}$$

Theorem 146 (Stirling's approximation)**.** For any $n > 0$ (without assuming $n \in \mathbb{Z}$) the following inequality holds:

$$\left(\frac{n}{e}\right)^n \sqrt{2\pi n}\exp\left(\frac{1}{12n+1}\right) \leq n! \leq \left(\frac{n}{e}\right)^n \sqrt{2\pi n}\exp\left(\frac{1}{12n}\right).$$

Proof. Since $\frac{1}{n^2} - \frac{1}{n(n+1)} = \frac{1}{n^2(n+1)}$, for any $m \in \mathbb{Z}^+$ we have:

$$\begin{aligned}\sum_{n\geq m}\frac{1}{n^2} &= \sum_{n\geq m}\left(\frac{1}{n}-\frac{1}{(n+1)}\right) + \frac{1}{2}\sum_{n\geq m}\left(\frac{1}{n^2}-\frac{1}{(n+1)^2}\right)\\ &+ \frac{1}{6}\sum_{n\geq m}\left(\frac{1}{n^3}-\frac{1}{(n+1)^3}\right) - \frac{1}{6}\sum_{n\geq m}\frac{1}{n^3(n+1)^3}\end{aligned}$$

hence:

$$\psi'(m) = \sum_{n\geq m}\frac{1}{n^2} \leq \frac{1}{m} + \frac{1}{2m^2} + \frac{1}{6m^3}$$

and the inequality still holds if $m \geq 1$ does not belong to $\mathbb{Z}$. Additionally, in a similar fashion:

$$\psi'(m) \geq \frac{1}{m} + \frac{1}{2m^2} + \frac{1}{6m^3} - \frac{1}{30m^5}.$$

By integrating both sides with respect to the m variable twice, we get that $\log\Gamma(m)$ has the following behaviour:

$$\log\Gamma(m) \approx \left(m-\frac{1}{2}\right)\log(m) - \alpha m + \beta + \frac{1}{12m}$$

where $\alpha = 1$ follows from the functional relation $\log\Gamma(m+1) - \log\Gamma(m) = \log m$. That gives Stirling's approximation up to a multiplicative constant: $\beta = \log\sqrt{2\pi}$ then follows from Legendre duplication formula and the identity $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$. □

The coefficients found through creative telescoping, due to Faulhaber's formula, directly depend on Bernoulli numbers. In particular the *Trigamma function* $\psi'(z)$ has the following asymptotic expansion:

$$\psi'(z) = \sum_{m\geq 0}\frac{1}{(z+m)^2} = \frac{1}{z} + \frac{1}{2z^2} + \sum_{t\geq 1}\frac{B_{2t}}{z^{2t+1}}$$

and by termwise integration:

$$\psi(z) = \log(z) - \frac{1}{2z} - \sum_{t\geq 1}\frac{B_{2t}}{2tz^{2t}},$$

$$\begin{aligned}\log\Gamma(z) &= \left(z-\frac{1}{2}\right)\log(z)-z+\log\sqrt{2\pi}+\sum_{t\geq 1}\frac{B_{2t}}{2t(2t-1)z^{2t-1}}\\ &= \left(z-\frac{1}{2}\right)\log(z)-z+\log\sqrt{2\pi}+\int_0^{+\infty}\frac{2\arctan\frac{t}{z}}{e^{2\pi t}-1}\,dt.\end{aligned}$$

where the last identity, following from the (inverse) Laplace transform, is also known as **second Binet's Theorem** for $\log\Gamma$.

We now consider the problem of computing

$$I(a,b)=\int_0^1 x^{a-1}(1-x)^{b-1}\,dx$$

for positive integer values of a and b. By exploiting the substitution $x=e^{-t}$ and the identity

$$\mathcal{L}^{-1}(1-e^{-t})^{b-1}=\sum_{k=0}^{b-1}(-1)^k\frac{\binom{b-1}{k}}{s+k}=\frac{(b-1)!}{s(s+1)\cdots(s+b-1)}$$

we have:

$$I(a,b)=\int_0^{+\infty}e^{-ta}(1-e^{-t})^b\,dt=\frac{(b-1)!}{a(a+1)\cdots(a+b-1)}=\frac{(a-1)!(b-1)!}{(a+b-1)!}=\frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)}.$$

The last identity holds in a more general context:

Theorem 147 (Eulero)**.** If a,b are complex numbers with positive real parts,

$$B(a,b)\stackrel{\text{def}}{=}\int_0^1 x^{a-1}(1-x)^{b-1}\,dx=\frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

Proof. The function $B(a,b)$ is clearly continuous and non-vanishing on its domain of definition.
Through the substitution $x\mapsto(1-x)$ we have $B(a,b)=B(b,a)$ and due to the integration by parts formula:

$$B(a+1,b)=\frac{a}{a+b}\,B(a,b).$$

$B(1,1)=\int_0^1 1\,dx=1$ holds and on the positive real line both $B(\cdot,b)$ and $B(a,\cdot)$ are log-convex, since they are *moments*. The claim hence follows from the Bohr-Mollerup characterization: $B(a,b)$ is known as Euler's **Beta function**. □

Euler's Beta function immediately gives extremely useful integral representations. For instance:

Lemma 148. If α and β are complex numbers with real parts greater than -1,

$$\int_0^{\pi/2}\sin^\alpha(\theta)\cos^\beta(\theta)\,d\theta=\frac{\Gamma\left(\frac{\alpha+1}{2}\right)\,\Gamma\left(\frac{\beta+1}{2}\right)}{2\,\Gamma\left(\frac{\alpha+\beta+2}{2}\right)}.$$

Lemma 149. If m and n are natural numbers, we have:

$$\int_0^{\pi}\sin^{2m}(\theta)\cos(2n\theta)\,d\theta=\frac{\pi(-1)^n}{4^n}\binom{2m}{m+n}.$$

Lemma 150. If $\text{Re}(a)>-1$ and $\text{Re}(b)>\text{Re}(a)+1$ we have:

$$\int_0^{+\infty}\frac{t^a\,dt}{1+t^b}=\frac{\pi}{b\,\sin\left(\frac{\pi(a+1)}{b}\right)}.$$

By combining Euler's Beta function with Feynman's trick, we get that many non-trivial integrals can be computed in an explicit way. We immediately study a *highly* non-trivial example.

Exercise 151. Prove that:

$$\int_0^{\pi/2} \log^3(\sin\theta)\, d\theta = -\frac{\pi}{8}\left[\pi^2\log 2 + 4\log^3 2 + 6\,\zeta(3)\right].$$

Proof. By differentiation under the integral sign we have that:

$$\int_0^{\pi/2} \log^3(\sin\theta)\, d\theta = \frac{\sqrt{\pi}}{2}\,\frac{d^3}{d\alpha^3}\,\frac{\Gamma\left(\frac{1}{2}+\frac{\alpha}{2}\right)}{\Gamma\left(1+\frac{\alpha}{2}\right)}\Bigg|_{\alpha=0}.$$

In particular the value of our integral just depends on the values of $\Gamma, \Gamma', \Gamma''$ and Γ''' at the points $\frac{1}{2}$ and 1.
We know that $\Gamma(1) = 1, \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ and by differentiating both sides of $\Gamma'(x) = \Gamma(x)\,\psi(x)$ multiple times we get:

$$\begin{aligned}
\Gamma'(x) &= \Gamma(x)\psi(x),\\
\Gamma''(x) &= \Gamma(x)\psi(x)^2 + \Gamma(x)\psi'(x),\\
\Gamma'''(x) &= \Gamma(x)\psi(x)^3 + 3\Gamma(x)\psi(x)\psi'(x) + \Gamma(x)\psi''(z)
\end{aligned}$$

The problem boils down to computing the values of ψ, ψ' and ψ'' at the points $\frac{1}{2}$ and 1.
By Gauss Theorem we have:

$$\psi(1) = -\gamma, \qquad \psi\left(\tfrac{1}{2}\right) = -\gamma - 2\log 2$$

and we also know that:

$$\psi'(a) = \sum_{n\geq 0}\frac{1}{(n+a)^2}$$

from which $\psi'(1) = \zeta(2) = \frac{\pi^2}{6}$ and $\psi'\left(\frac{1}{2}\right) = 3\,\zeta(2) = \frac{\pi^2}{2}$ follow.
By differentiating again with respect to the a variable, we get:

$$\psi''(a) = -2\sum_{n\geq 0}\frac{1}{(n+a)^3}$$

from which it is simple to derive $\psi''(1) = -2\,\zeta(3)$ and $\psi''\left(\frac{1}{2}\right) = -14\,\zeta(3)$.
At this point, in order to prove the claim it is enough to trust in a Computer Algebra System to perform the needed simplifications, or just perform them by hand with a bit of patience. In a similar way it is possible to prove:

$$\begin{aligned}
\int_0^{\pi/2}\log(\sin\theta)\log(\cos\theta)\, d\theta &= \frac{\pi}{2}\log^2(2) - \frac{\pi^3}{48},\\
\int_0^{\pi/2}\log^2(\sin\theta)\, d\theta &= \frac{\pi^3}{24} + \frac{\pi}{2}\log^2(2),\\
\int_0^{\pi/2}\log^4(\sin\theta)\, d\theta &= \frac{19\pi^5}{480} + \frac{\pi^3}{4}\log^2(2) + \frac{\pi}{2}\log^4(2) + 3\pi\log(2)\,\zeta(3).
\end{aligned}$$

□

As mentioned before, we investigate now about the interplay between the Γ function and the area of the unit circle. Since the graph of $f(x) = \sqrt{1-x^2}$ over $[-1, 1]$ is a half-circle,

$$\begin{aligned}
\pi &= 4\int_0^1 \sqrt{1-x^2}\, dx = 2\int_0^1 \frac{1}{\sqrt{u}}\sqrt{1-u}\, du\\
&= 2\int_0^1 u^{\frac{1}{2}-1}(1-u)^{\frac{3}{2}-1}\, du = 2\,B\left(\tfrac{1}{2},\tfrac{3}{2}\right) = 2\,\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{3}{2}\right)}{\Gamma(2)} = \Gamma\left(\tfrac{1}{2}\right)^2.
\end{aligned}$$

The last identity has an interesting generalization: let $V_n(\rho)$ and $A_n(\rho)$ be the volume and the surface area of the Euclidean ball with radius $\rho \geq 0$ in $\mathbb{R}^n$. For starters, we may notice that:

$$V_n(\rho) = \rho^n V_n(1), \qquad A_n(\rho) = \rho^{n-1} A_n(1), \qquad A_n(\rho) = \frac{d}{d\rho} V_n(\rho) = n\rho^{n-1} V_n(1),$$

then we consider the integral

$$I_n = \int_{\mathbb{R}^n} \exp\left[-(x_1^2 + x_2^2 + \ldots + x_n^2)\right]\, d\mu.$$

Due to Fubini's Theorem, the following identity clearly holds:

$$I_n = \left(\int_{\mathbb{R}} e^{-x^2}\, dx\right)^n = \pi^{n/2}.$$

Due to Cavalieri's principle we also have:

$$I_n = \int_0^{+\infty} A_n(\rho) e^{-\rho^2}\, d\rho = A_n(1)\int_0^{+\infty} \rho^{n-1} e^{-\rho^2}\, d\rho = \frac{A_n(1)}{2}\int_0^{+\infty} u^{\frac{n}{2}-1} e^{-u}\, du$$

hence we get the identity:

$$2\pi^{n/2} = A_n(1)\Gamma(n/2)$$

and finally:

$$V_n(\rho) = \frac{\pi^{n/2}}{\Gamma(1+n/2)}\, \rho^n, \qquad A_n(\rho) = \frac{2\pi^{n/2}}{\Gamma(n/2)}\, \rho^{n-1}.$$

As n ranges over $\mathbb{N}^+$, $V_n(1)$ attains a maximum at $n = 5$ and $A_n(1)$ attains a maximum at $n = 7$. Additionally, if $n \geq 13$ the volume of the unit balls *is less* than the volume of the unit cube.

Theorem 152 (Bailey-Borwein-Plouffe, BBP formula)**.**

$$\pi = \sum_{k\geq 0}\left[\frac{1}{16^k}\left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right)\right].$$

This quite recent result (1995) has opened new frontiers in the problem of finding in a efficient way the digits in the binary expansion of π. The proof of such identity through the instruments so far acquired is pretty simple:

$$\begin{aligned}
\sum_{k\geq 0}\left[\frac{1}{16^k}\left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right)\right] &= \int_0^1 \sum_{k\geq 0} \frac{4x^{8k} - 2x^{8k+3} - x^{8k+4} - x^{8k+5}}{16^k}\, dx \\
&= 16\int_0^1 \frac{1-x}{4-4x+2x^3-x^4}\, dx \\
(x \mapsto 1-x) \qquad &= 16\int_0^1 \frac{x\, dx}{(1+x^2)(1+2x-x^2)} \\
&= 4\int_0^1 \frac{1+x}{1+x^2}\, dx - 4\int_0^1 \frac{x}{2-x^2}\, dx \\
&= (\pi + \log 4) - \log 4 = \pi.
\end{aligned}$$

The author conjectures the BBP formula (or series expansions with a similar structure) might be the key for proving the base-2 *normality* of the π constant, i.e., loosely speaking, the fact that any binary string appears in the binary expansion of π an infinite number of times, possibly with a regular frequency. On its behalf, the base-2 normality of π is deeply related to the (actually unknown) convergence properties of series like $\sum_{n\geq 1}\frac{\sin(2^n)}{n}$ and similar ones, linked with *universally bad averaging sequences.*

Exercise 153. Prove that:

$$\sum_{n\geq 0}\frac{n!}{(2n+1)!} = 2e^{1/4}\int_0^{1/2} e^{-x^2}\,dx \leq \frac{1}{3}+\frac{2}{3}e^{1/4}.$$

Proof. We may notice that:

$$\begin{aligned}
\sum_{n\geq 0}\frac{n!}{(2n+1)!} = \sum_{n\geq 0}\frac{\Gamma(n+1)}{\Gamma(2n+2)} &= \sum_{n\geq 0}\frac{B(n+1,n+1)}{n!}\\
&= \int_0^1 \sum_{n\geq 0}\frac{x^n(1-x)^n}{n!}\,dx\\
&= \int_0^1 \exp\left[x(1-x)\right]\,dx\\
&= 2\int_0^{1/2} \exp\left[x(1-x)\right]\,dx\\
&= \int_0^1 \exp\left[\frac{1-x^2}{4}\right]\,dx.
\end{aligned}$$

The last inequality follows from the fact that, by convexity, on the interval $[0,1]$ we have:

$$\exp\left[\frac{1-x^2}{4}\right] \leq 1+(e^{1/4}-1)(1-x^2).$$

□

Exercise 154 (Reflection formula for the Dilogarithm function)**.** Prove that for any $x\in(0,1)$ we have:

$$\zeta(2)-\log(x)\log(1-x) = \mathrm{Li}_2(x)+\mathrm{Li}_2(1-x)$$

where $\mathrm{Li}_2(x)=\sum_{n\geq 1}\frac{x^n}{n^2}$.

Proof. Let us define:

$$f(x) = \log(x)\log(1-x)+\mathrm{Li}_2(x)+\mathrm{Li}_2(1-x).$$

We are interested in showing that f is constant, hence we compute f':

$$f'(x) = \left(\frac{\log(1-x)}{x}-\frac{\log(x)}{1-x}\right)-\frac{\log(1-x)}{x}+\frac{\log x}{1-x} = 0.$$

In order to prove the claim it is enough to compute $f(x)$ at a point of regularity, or to compute the limit:

$$\lim_{x\to 1^-} f(x) = \zeta(2)+\lim_{x\to 1^-}\log(x)\log(1-x) = \zeta(2).$$

We have an interesting corollary:

$$\sum_{n\geq 1}\frac{1}{2^n n^2} = \mathrm{Li}_2\left(\frac{1}{2}\right) = \frac{1}{2}\left(f\left(\frac{1}{2}\right)-\log^2 2\right) = \frac{\pi^2}{12}-\frac{\log^2 2}{2}.$$

□

Exercise 155 ("Continuous" binomial theorem)**.** Prove that for any $n \in \mathbb{N}$ we have:

$$\int_{-\infty}^{+\infty} \binom{n}{x}\, dx = 2^n.$$

Proof. By exploiting the reflection formula for the Γ function we have:

$$\binom{n}{x} = \frac{n!}{\pi} \cdot \frac{\sin(\pi x)}{(n-1)(n-1-x)\cdot \ldots \cdot (1-x)x}$$

and since

$$\frac{1}{(n-1)(n-1-x)\cdot \ldots \cdot (1-x)x} = \sum_{k=0}^{n} \frac{c_k}{x-k}, \qquad \int_{-\infty}^{+\infty} \frac{\sin(\pi x)}{x-k} = (-1)^k \pi$$

follows from computing a partial fraction decomposition through the residue Theorem, $c_k = \frac{(-1)^k}{n!}\binom{n}{k}$, we have:

$$\int_{-\infty}^{+\infty} \binom{n}{x}\, dx = \sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

□

Exercise 156. By recalling that $\operatorname{sinc}(x)$ is the function that equals 1 at $x = 0$ and $\frac{\sin x}{x}$ anywhere else, prove that for any couple (α, β) of real numbers in $(0, 1)$ we have:

$$\sum_{n\in\mathbb{Z}} \operatorname{sinc}(n\alpha)\operatorname{sinc}(n\beta) = \frac{\pi}{\max(\alpha,\beta)}.$$

Proof. Since

$$\sum_{n\in\mathbb{Z}} \operatorname{sinc}(n\alpha)\operatorname{sinc}(n\beta) = 1 + \frac{2}{\alpha\beta}\sum_{n\geq 1} \frac{\sin(n\alpha)\sin(n\beta)}{n^2},$$

due to the addition formulas for the sine and cosine functions it is enough to prove the equality

$$\forall \theta \in [0, 2\pi], \qquad g(\theta) \stackrel{\text{def}}{=} \sum_{n\geq 1} \frac{\cos(n\theta)}{n^2} = \frac{\pi^2}{6} - \frac{\theta(2\pi-\theta)}{4}$$

following by termwise integration of the Fourier series of a sawtooth wave. In particular,

$$\begin{aligned}
\sum_{n\in\mathbb{Z}} \operatorname{sinc}(n\alpha)\operatorname{sinc}(n\beta) &= 1 + \frac{g(|\alpha-\beta|) + g(\alpha+\beta)}{\alpha\beta} \\
&= 1 + \frac{1}{4\alpha\beta}\left[(\alpha-\beta)^2 - (\alpha+\beta)^2 + 2\pi(\alpha+\beta-|\alpha-\beta|)\right] \\
&= 1 + \frac{1}{4\alpha\beta}\left[-4\alpha\beta + 4\pi\min(\alpha,\beta)\right] = \frac{\pi}{\max(\alpha,\beta)}.
\end{aligned}$$

□

Exercise 157. If for any $n \geq 1$ we define

$$A_n = \prod_{k=1}^{n} \frac{3k}{6k-4},$$

what is the value of $\sum_{n\geq 1} A_n$?

Proof. Since $A_n = \frac{n}{2^n} \cdot B\left(n, \frac{1}{3}\right)$, we have:

$$\begin{aligned} \sum_{n\geq 1} A_n &= \sum_{n\geq 1} \frac{n}{2^n} \int_0^1 x^{n-1}(1-x)^{-2/3}\,dx \\ &= \int_0^1 \frac{2(1-x)^{-2/3}}{(2-x)^2}\,dx \\ (x\mapsto 1-x) \quad &= \int_0^1 \frac{2x^{-2/3}}{(1+x)^2}\,dx \\ (x\mapsto u^3) \quad &= \int_0^1 \frac{6\,du}{(1+u^3)^2} \end{aligned}$$

and the last integral can be (tediously) computed through partial fraction decomposition:

$$\sum_{n\geq 1} A_n = 1 + \frac{4\pi}{3\sqrt{3}} + \frac{4\log 2}{3}.$$

□

Exercise 158. Prove that by setting

$$f(x) \stackrel{\text{def}}{=} \sum_{n\geq 1} \frac{\sin(n^2 x)}{n}$$

we have that:

$$\lim_{x\to 0^+} f(x) = \frac{\pi}{2}.$$

Proof. $f(x)$ is defined by a pointwise convergent series by **Dirichlet's test** and by **Weyl's inequality**. It follows that the wanted limit can be computed through a convolution trick, i.e. by expoiting an approximated identity:

$$\lim_{x\to 0^+} f(x) = \lim_{m\to+\infty} \int_0^{+\infty} f(x) m e^{-mx}\,dx = \lim_{m\to+\infty} \sum_{n\geq 1} \frac{mn}{m^2+n^4} = \lim_{m\to+\infty} \sum_{n\geq 1} \frac{2m^2 n}{4m^4+n^4}.$$

Due to Sophie Germain's identity we have:

$$\frac{2m^2 n}{4m^4+n^4} = \frac{m}{2}\left(\frac{1}{2m^2+2mn+n^2} - \frac{1}{2m^2-2mn+n^2}\right),$$

hence:

$$\begin{aligned} \lim_{m\to+\infty} \sum_{n\geq 1} \frac{2m^2 n}{4m^4+n^4} &= \lim_{m\to+\infty} \frac{i}{4}\left(H_{-m(1+i)} - H_{m(-1+i)} - H_{m(1-i)} + H_{m(1+i)}\right) \\ &= \frac{i}{4}\left[\log(-1-i) - \log(-1+i) - \log(1-i) + \log(1+i)\right] = \frac{\pi}{4}. \end{aligned}$$

In a similar way it is possible to show that:

$$\forall k \in \mathbb{Z}^+, \qquad \lim_{x\to 0^+} \sum_{n\geq 1} \frac{\sin(n^k x)}{n} = \frac{\pi}{2k}.$$

□

Exercise 159 (Raabe's $\log\Gamma$ Theorem)**.** Prove that for any positive real number α we have:

$$\int_0^1 \log\Gamma(x+\alpha)\,dx = \alpha\log\alpha - \alpha + \log\sqrt{2\pi}.$$

Exercise 160 (Glasser's Master Theorem)**.** Prove that if F and $F \circ \varphi$ are integrable functions on the real line, where

$$\varphi(x) = |a|x - \sum_{k=1}^{N} \frac{|a_k|}{x - b_k},$$

then the following identity holds:

$$\int_{-\infty}^{+\infty} F(x)\, dx = \int_{-\infty}^{+\infty} F(\varphi(x))\, dx.$$

For the proof presented here, the author is really grateful to *achillehui*.

Proof. Let $\phi(z)$ be any meromorphic function over $\mathbb{C}$ which

1. preserve the extended real line $\mathbb{R}^* = \mathbb{R} \cup \{\infty\}$ in the sense:

$$\begin{cases} \phi(\mathbb{R}) \subset \mathbb{R}^* \\ \phi^{-1}(\mathbb{R}) \subset \mathbb{R} \end{cases} \implies P \stackrel{def}{=} \phi^{-1}(\infty) = \left\{ p \in \mathbb{C} : p \text{ poles of } \phi(z) \right\} \subset \mathbb{R}$$

2. Split $\mathbb{R} \setminus P$ as a countable union of its connected components $\bigcup_n (a_n, b_n)$. Each connected component is an open interval (a_n, b_n) and on such an interval, $\phi(z)$ increases from $-\infty$ at a_n^+ to ∞ at b_n^-.

3. There exists an ascending chain of Jordan domains $D_1, D_2, \ldots$ that cover $\mathbb{C}$,

$$\{0\} \subset D_1 \subset D_2 \subset \cdots \quad \text{with} \quad \bigcup_{k=1}^{\infty} D_k = \mathbb{C}$$

whose boundaries ∂D_k are "well behaved", "diverge" to infinity and $|z - \phi(z)|$ is bounded on the boundaries. More precisely, let

$$\begin{cases} R_k \stackrel{def}{=} \inf \left\{ |z| : z \in \partial D_k \right\} \\ L_k \stackrel{def}{=} \int_{\partial D_k} |dz| < \infty \\ M_k \stackrel{def}{=} \sup \left\{ |z - \phi(z)| : z \in \partial D_k \right\} \end{cases} \quad \text{and} \quad \begin{cases} \lim_{k\to\infty} R_k = \infty \\ \lim_{k\to\infty} \frac{L_k}{R_k^2} = 0 \\ M = \sup_k M_k < \infty \end{cases}$$

Given such a meromorphic function $\phi(z)$ and any Lebesgue integrable function $f(x)$ on $\mathbb{R}$, we have the following identity:

$$\int_{-\infty}^{\infty} f(\phi(x))dx = \int_{-\infty}^{\infty} f(x)dx \qquad (*1)$$

In order to prove this, we split our integral into a sum over the connected components of $\mathbb{R} \setminus P$.

$$\int_{\mathbb{R}} f(\phi(x))dx = \int_{\mathbb{R}\setminus P} f(\phi(x))dx = \sum_n \int_{a_n}^{b_n} f(\phi(x))dx$$

For any connected component (a_n, b_n) of $\mathbb{R} \setminus P$ and $y \in \mathbb{R}$, consider the roots of the equation $\phi(x) = y$. Using properties (1) and (2) of $\phi(z)$, we find there is a unique root for the equation $y = \phi(x)$ over (a_n, b_n). Let us call this root as $r_n(y)$. Enforcing the substitution $y = \phi(x)$ the integral becomes

$$\sum_n \int_{-\infty}^{\infty} f(y) \frac{dr_n(y)}{dy} dy = \int_{-\infty}^{\infty} f(y) \left(\sum_n \frac{dr_n(y)}{dy} \right) dy.$$

We can use the obvious fact $\frac{dr_n(y)}{dy} \geq 0$ and dominated convergence theorem to justify the switching of order of summation and integral.

This means to prove (∗1), one only need to show

$$\sum_n \frac{dr_n(y)}{dy} \stackrel{?}{=} 1 \qquad (*2)$$

For any $y \in \mathbb{R}$, let $R(y) = \phi^{-1}(y) \subset \mathbb{R}$ be the collection of roots of the equation $\phi(z) = y$. Over any Jordan domain D_k, we have the following expansion

$$\frac{\phi'(z)}{\phi(z) - y} = \sum_{r \in R(y) \cap D_k} \frac{1}{z-r} - \sum_{p \in P \cap D_k} \frac{1}{z-p} + \text{something analytic}$$

This leads to

$$\sum_{r \in R(y) \cap D_k} r - \sum_{p \in P \cap D_k} p = \frac{1}{2\pi i} \int_{\partial D_k} z \left(\frac{\phi'(z)}{\phi(z) - y} \right) dz$$

As long as $R(y) \cap \partial D_k = \emptyset$, we can differentiate both sides and get

$$\begin{aligned}\sum_{r_n(y) \in D_k} \frac{dr_n(y)}{dy} &= \frac{1}{2\pi i} \int_{\partial D_k} z \left(\frac{\phi'(z)}{(\phi(z) - y)^2} \right) dz = -\frac{1}{2\pi i} \int_{\partial D_k} z \frac{d}{dz} \left(\frac{1}{\phi(z) - y} \right) dz \\ &= \frac{1}{2\pi i} \int_{\partial D_k} \frac{dz}{\phi(z) - y}\end{aligned}$$

For those k large enough to satisfy $R_k > 2(M + |y|)$, we can expand the integrand in last line as

$$\frac{1}{\phi(z) - y} = \frac{1}{z - (y + z - \phi(z))} = \frac{1}{z} + \sum_{j=1}^{\infty} \frac{(y + z - \phi(z))^j}{z^{j+1}}$$

and obtain a bound

$$\left| \left(\sum_{r_n(y) \in D_k} \frac{dr_n(y)}{dy} \right) - 1 \right| \le \frac{1}{2\pi} \sum_{j=1}^{\infty} \int_{\partial D_k} \frac{(|y| + |z - \phi(z)|)^j}{|z|^{j+1}} |dz| \le \frac{(M + |y|)L_k}{2\pi R_k^2} \sum_{j=0}^{\infty} \left(\frac{M + |y|}{R_k} \right)^j \le \frac{M + |y|}{\pi} \frac{L_k}{R_k^2}$$

Since $\lim_{k \to \infty} \frac{L_k}{R_k^2} = 0$, this leads to

$$\sum_n \frac{dr_n(y)}{dy} = \lim_{k \to \infty} \sum_{r_n(y) \in D_k} \frac{dr_n(y)}{dy} = 1$$

This justifies (∗2) and hence (∗1) is proved. Notice all the $\frac{dr_n(y)}{dy}$ are positive, there is no issue in rearranging the order of summation in last line.

□

Corollary 161.

$$\int_0^{+\infty} \frac{x^2\,dx}{x^4 + x^2 + 1} \stackrel{\text{parity}}{=} \frac{1}{2} \int_{-\infty}^{+\infty} \frac{dx}{\left(x - \frac{1}{x}\right)^2 + 3} \stackrel{\text{G.M.T.}}{=} \frac{1}{2} \int_{-\infty}^{+\infty} \frac{dx}{x^2 + 3} = \frac{\pi}{2\sqrt{3}}.$$

Corollary 162. Observe that, for $x \neq n\pi$, $n = 0, \pm 1, \pm 2, \ldots$, we have

$$\cot x = \lim_{N \to +\infty} \left(\frac{1}{x} + \frac{1}{x + \pi} + \frac{1}{x - \pi} + \cdots + \frac{1}{x + N\pi} + \frac{1}{x - N\pi} \right)$$

leading to:

$$\int_{-\infty}^{+\infty} f\left(x - \cot x\right)\,dx = \int_{-\infty}^{+\infty} f(x)\,dx.$$

By the substitution $x \mapsto \frac{\pi}{2} - x$ it follows that:

$$\int_{-\infty}^{+\infty} \frac{dx}{1 + (x + \tan x)^2} = \int_{-\infty}^{+\infty} \frac{dx}{1 + x^2} = \pi.$$

Corollary 163. For any $a, b \in \mathbb{R}^+$,

$$\int_0^{+\infty} e^{-ax^2+\frac{b}{x^2}}\,dx = \sqrt{\frac{\pi}{4a}}e^{-2\sqrt{ab}}.$$

The Riemann ζ function and its analytic continuation. For any complex number with real part greater than one the Riemann ζ function is defined through the absolutely convergent series

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

that through the (inverse) Laplace transform admits the following integral representation:

$$\zeta(s) = \frac{1}{\Gamma(s)}\int_0^{+\infty}\frac{x^{s-1}}{e^x-1}\,dx.$$

In the region $\text{Re}(s) > 1$ we further have:

$$\eta(s) \stackrel{\text{def}}{=} \sum_{n\geq 1}\frac{(-1)^{n+1}}{n^s} = \left(1-\frac{2}{2^s}\right)\zeta(s),$$

but the series defining $\eta(s)$ has a larger domain of (conditional) convergence, $\text{Re}(s) > 0$.
Due to such fact we are allowed to extend the ζ function to the half-plane $\text{Re}(s) > 0$ through

$$\begin{aligned}
\zeta(s) &= \left(1-\frac{2}{2^s}\right)^{-1}\sum_{n\geq 1}\frac{(-1)^{n+1}}{n^s} \\
(\text{by } \mathcal{L}^{-1}) &= \frac{1}{\Gamma(s)}\left(1-\frac{2}{2^s}\right)^{-1}\int_0^{+\infty}\frac{x^{s-1}}{e^x+1}\,dx \\
(\text{by parts}) &= \frac{4^s}{2(2^s-2)\Gamma(s+1)}\int_0^{+\infty}\frac{x^s\,dx}{\cosh^2(x)} \\
(\text{by substitution}) &= \frac{4^s}{2(2^s-2)\Gamma(s+1)}\int_0^1 \text{arctanh}(x)^s\,dx \\
&= \frac{2^{s-1}}{(2^s-2)\Gamma(s+1)}\int_0^1 \log^s\left(\frac{1+x}{1-x}\right)dx.
\end{aligned}$$

These integral representations shed a good amount of light on the tight interplay between logarithmic integrals and values of the ζ function. They also underline that $s = 1$ is a simple pole with residue 1, that $\zeta(s)$ attains negative values for any $s \in [0, 1)$ and that $\zeta(0) = -\frac{1}{2}$. By exploiting Euler's product

$$\forall s : \text{Re}(s) > 0, \qquad \zeta(s) = \prod_{p\in\mathcal{P}}\left(1-\frac{1}{p^s}\right)^{-1}$$

and its logarithmic derivative, we get that there is a very deep correspondence between the distribution of zeroes of the ζ function in the region $0 < \text{Re}(s) < 1$ and the distribution of prime numbers:

$$\sum_{p^m<X}\log p = X-(b+1)-\lim_{T\to+\infty}\sum_{|\text{Im}(\rho)|<T}\frac{X^\rho}{\rho}+\sum_{n\geq 1}\frac{X^{-2n}}{2n}.$$

In particular, the **Prime Number Theorem** (PNT)

$$\lim_{n\to+\infty}\frac{|\mathcal{P}\cap[1,n]|\cdot\log n}{n}=1$$

is substantially equivalent to the statement *the ζ function is non-vanishing on the line $Re(s)=1$*, statement that was almost simultaneously (but independently) proved by Hadamard and de la Vallée-Poussin through a trigonometric trick. By defining $\pi(x)$ as the number of prime numbers in the interval $[1,x]$, the following strengthening of the PNT

$$\left|\pi(x)-\int_2^x\frac{dt}{\log t}\right|\ll\sqrt{x}\log^2(x)$$

is substantially equivalent to **Riemann Hypothesis** (RH): *all the zeroes of the $\zeta(s)$ function in the region $0<Re(s)<1$ lie on the* critical *line $Re(s)=\frac{1}{2}$*. Conversely, inequalities of the form

$$\left|\pi(x)-\int_2^x\frac{dt}{\log t}\right|\ll x^\alpha$$

for some $\alpha\in\left(\frac{1}{2},1\right)$ imply the absence of zeroes of the ζ function in subsets of the **critical line** $0<\text{Re}(s)<1$. The best result actually known about the zero-free region for the ζ function is due to Korobov and Vinogradov:

$$\sum_{p^m<X}\log p=X+O\left(X\exp\left(-c\,\frac{\log^{3/5}X}{(\log\log X)^{1/5}}\right)\right).$$

This result (pitifully very far from RH) comes from a sophisticated combinatorial manipulation of exponential sums (variants on **Van Der Corput's trick**), combined with classical inequalities in Complex Analysis (Hadamard, Borel-Caratheodory). The reader can find on **Terence Tao's blog** a very detailed lecture. The ζ function can be further extended to the whole complex plane. By setting

$$\theta(z)=\sum_{n\in\mathbb{Z}}e^{\pi i n^2 z}$$

for any $z\in\mathbb{C}$ with positive real part, we have

$$\theta(z)=\frac{1}{\sqrt{-iz}}\,\theta\left(\frac{1}{z}\right)$$

as a consequence of Poisson summation formula, hence the function

$$\xi(s)\stackrel{\text{def}}{=}\pi^{-s/2}\Gamma\left(\frac{s}{2}\right)\zeta(s)$$

has the following integral representation

$$\xi(s)=-\frac{1}{s(1-s)}+\int_1^{+\infty}\frac{\theta(iy)-1}{2}\left(y^{s/2}+y^{(1-s)/2}\right)\frac{dy}{y}$$

and ξ turns out to be a meromorphic function (with simple poles at $s=0$ and $s=1$) such that $\xi(s)=\xi(1-s)$. This is known as the *reflection formula* for the ζ function.

Exercise 164 (Exploiting Euler's product in a quantitative way)**.**
By defining $\mathcal{P}$ as the set of prime numbers, prove that for any real number $s>1$ we have:

$$\prod_{p\in\mathcal{P}}\left(1-\frac{1}{p^s}\right)^{-1}=\zeta(s)=\sum_{n\geq 1}\frac{1}{n^s}$$

then prove the following inequality:

$$\sum_{p\in\mathcal{P}} \frac{1}{p^2} \leq \log\sqrt{\frac{5}{2}}.$$

Hint: for any $x \in (0,1)$ we have $x \leq \operatorname{arctanh} x$.

Exercise 165. Prove that:

$$\sum_{n\geq 1} (\zeta(2n)-1) = \frac{3}{4}.$$

Proof. By the generating function for the sequence $\{\zeta(2n)\}_{n\geq 1}$,

$$\sum_{n\geq 1} (\zeta(2n)-1) = \lim_{z\to 1^-} \left(\frac{1-\pi z\cot(\pi z)}{2} - \frac{z^2}{1-z^2}\right) \overset{\text{d.H.}}{=} \frac{3}{4}.$$

As an alternative,

$$\sum_{n\geq 1} (\zeta(2n)-1) = \sum_{n\geq 1} \int_0^{+\infty} \frac{x^{2n-1}}{(2n-1)!}\left(\frac{1}{e^x-1} - \frac{1}{e^x}\right) dx = \int_0^{+\infty} \frac{\sinh(x)}{e^x(e^x-1)}\,dx = \int_1^{+\infty} \frac{t+1}{2t^3}\,dt = \frac{3}{4}$$

or simply:

$$\sum_{n\geq 1} (\zeta(2n)-1) = \sum_{n\geq 1}\sum_{m\geq 2} \frac{1}{m^{2n}} = \sum_{m\geq 2} \frac{1}{m^2-1} = \frac{1}{2}\sum_{m\geq 2}\left(\frac{1}{m-1} - \frac{1}{m+1}\right) = \frac{H_2}{2}.$$

□

Exercise 166. Prove that for any $s \geq 1$ the following inequality holds:

$$\int_0^{+\infty} \frac{dx}{(1+x^2)^s} \leq \sqrt{\frac{\pi}{4s-3}}$$

and notice that it implies $\pi < \sqrt{10}$ (Hint: evaluate both sides at $s=3$).

Exercise 167. Let A be the set $(\mathbb{Z}\times\mathbb{Z})\setminus\{0,0\}$. Prove that:

$$\sum_{(m,n)\in A} \frac{1}{m^4+n^4} \leq 4\,\zeta(4) + \pi\sqrt{2}\,\zeta(3).$$

Exercise 168. Prove that:

$$\frac{\pi}{2} = \sum_{n\geq 0} \frac{(-1)^n}{2n+1}\cdot\frac{\zeta(4n+2)}{4^n}.$$

Exercise 169. Prove the inequality:
$$\sum_{n\geq 1}\log^2\left(1+\frac{1}{n}\right)<1.$$

Proof. By recalling $\log^2(1-z)=\sum_{n\geq 1}\frac{2H_{n-1}}{n}z^n$ and the integral representation for harmonic numbers we have:

$$\begin{aligned}
S=\sum_{m\geq 1}\log^2\left(1+\frac{1}{m}\right) &= \sum_{m\geq 1}\sum_{n\geq 1}\frac{2H_{n-1}(-1)^n}{nm^n}\\
&= \int_0^1\sum_{m\geq 1}\sum_{n\geq 1}\frac{2H_{n-1}(z^{n-1}-1)(-1)^n}{nm^n(z-1)}\,dz\\
&= 2\int_0^1\sum_{m\geq 1}\frac{\log\left(1+\frac{z}{m}\right)-z\log\left(1+\frac{1}{m}\right)}{z(1-z)}\,dz\\
\text{(by Euler's product for the }\Gamma\text{ function)}\quad &= -2\int_0^1\frac{\log\Gamma(1+z)}{z(1-z)}\,dz\\
\text{(by integration by parts)}\quad &= 2\int_0^1\psi(z+1)\log\left(\frac{z}{1-z}\right)\,dz\\
\text{(by the reflection formula for }\psi)\quad &= 2\int_0^1\left(\frac{1}{z}-\frac{1}{1-z}-\pi\cot(\pi z)\right)\log(z)\,dz
\end{aligned}$$

The magic now comes from studying the function $g(z)=\left(\frac{1}{z}-\frac{1}{1-z}-\pi\cot(\pi z)\right)$ over the interval $(0,1)$. It is extremely close to $2z-1$ by "cancellation of singularities", hence
$$S\approx 2\int_0^1(2z-1)\log(z)\,dz=1.$$
$g(z)$ is symmetric with respect to $z=\frac{1}{2}$ and $\log(z)$ is negative and increasing over $(0,1)$, hence $\approx$ is indeed a $<$.

□

Exercise 170. Prove that the previous statement implies
$$\sum_{n\geq 2}\log^2\left(\frac{n+1}{n-1}\right)\leq 2\log 2+2-\log^2 2.$$

Exercise 171 (Putnam 2016, Problem B6)**.** Prove that:
$$\sum_{k\geq 1}\frac{(-1)^{k+1}}{k}\sum_{n\geq 0}\frac{1}{k2^n+1}=1.$$

Proof. The dominated convergence Theorem allows us to perform the following manipulations:
$$\sum_{k\geq 1}\frac{(-1)^{k+1}}{k}\sum_{n\geq 0}\frac{1}{k2^n+1}=\sum_{n\geq 0}\sum_{k\geq 1}\frac{(-1)^k}{k}\left(\frac{1}{k2^n}-\frac{1}{k^2 4^n}+\frac{1}{k^3 8^n}-\ldots\right)=\sum_{h\geq 1}{}'(-1)^{h+1}\frac{2^h}{2^h-1}\sum_{k\geq 1}\frac{(-1)^{k+1}}{k^{h+1}}$$
$$=\sum_{h\geq 1}{}'(-1)^{h+1}\zeta(h+1)=\int_0^{+\infty}\frac{1}{e^x-1}\sum_{h\geq 1}\frac{(-1)^{h+1}x^h}{h!}\,dx=\int_0^{+\infty}e^{-x}\,dx=1$$
where $\sum'$ denotes a regularized sum in the Cesàro sense.

□

An exercise on generalized Euler sums, logarithmic integrals and values of ζ.

Exercise 172. Prove that

$$\sum_{k=2}^{\infty}\frac{(-1)^k}{k^2}H_kH_{k-1}=\frac{3}{16}\zeta(4).$$

Proof. We are going to see a generalization of the approach used (11) to show that $\zeta(4)=\frac{2}{5}\zeta(2)^2$. This is also an opportunity to make a tribute to Pieter J. de Doelder (1919-1994) from Eindhoven University of Technology, who evaluated in closed form the given series in a somewhat famous paper published in 1991. One may start by using the following identity coming from the Cauchy product,

$$\ln^2(1+x)=2\sum_{n=1}^{\infty}(-1)^{n-1}\frac{H_n}{n+1}\,x^{n+1}$$

giving

$$\int_0^1\frac{\ln(1-x)\ln^2(1+x)}{x}\,dx=2\sum_{n=1}^{\infty}(-1)^{n-1}\frac{H_n}{n+1}\int_0^1 x^n\ln(1-x)\,dx,$$

then using the standard evaluation

$$\int_0^1 x^n\ln(1-x)\,dx=-\frac{H_{n+1}}{n+1},\quad n\geq 0,$$

one gets

$$\int_0^1\frac{\ln(1-x)\ln^2(1+x)}{x}\,dx=2\sum_{n=2}^{\infty}(-1)^{n-1}\frac{H_nH_{n-1}}{n^2}.$$

Here are the main steps which de Doelder took to evaluate the related integral. We clearly have

$$\int_0^1\ln^3\left(\frac{1+x}{1-x}\right)\frac{dx}{x}=\int_0^1\frac{\ln^3(1+x)}{x}\,dx-3\int_0^1\frac{\ln(1-x)\ln^2(1+x)}{x}\,dx$$
$$+3\int_0^1\frac{\ln^2(1-x)\ln(1+x)}{x}\,dx-\int_0^1\frac{\ln^3(1-x)}{x}\,dx$$

and

$$\int_0^1\frac{\ln^3(1-x^2)}{x}\,dx=\int_0^1\frac{\ln^3(1+x)}{x}\,dx+3\int_0^1\frac{\ln(1-x)\ln^2(1+x)}{x}\,dx$$
$$+3\int_0^1\frac{\ln^2(1-x)\ln(1+x)}{x}\,dx+\int_0^1\frac{\ln^3(1-x)}{x}\,dx,$$

subtracting the two equalities,

$$6\int_0^1\frac{\ln(1-x)\ln^2(1+x)}{x}dx=\int_0^1\frac{\ln^3(1-x^2)}{x}dx-\int_0^1\ln^3\left(\frac{1+x}{1-x}\right)\frac{dx}{x}-2\int_0^1\frac{\ln^3(1-x)}{x}dx$$
$$=I_1-I_2-2I_3.$$

It is easy to obtain

$$\begin{aligned} I_1 = \int_0^1 \frac{\ln^3\left(1-x^2\right)}{x} dx &= \frac{1}{2}\int_0^1 \frac{\ln^3(1-u)}{u} du \quad (u = x^2) \\ &= \frac{1}{2}\int_0^1 \frac{\ln^3 v}{1-v} dv \quad (v = 1-u) \\ &= \frac{1}{2}\sum_{n=0}^{\infty}\int_0^1 v^n \ln^3 v \, dv \\ &= -3\sum_{n=1}^{\infty}\frac{1}{n^4} = -\frac{\pi^4}{30}, \end{aligned}$$

similarly

$$I_3 = \int_0^1 \frac{\ln^3(1-x)}{x} dx = -\frac{\pi^4}{15}.$$

By the change of variable, $u = \dfrac{1-x}{1+x}$, one has $\dfrac{dx}{x} = \dfrac{-2\,du}{1-u^2}$ getting

$$\begin{aligned} I_2 = \int_0^1 \ln^3\left(\frac{1+x}{1-x}\right)\frac{dx}{x} &= -2\int_0^1 \frac{\ln^3 u}{1-u^2} du \\ &= -2\sum_{n=0}^{\infty}\int_0^1 u^{2n}\ln^3 u \, dv \\ &= 12\sum_{n=0}^{\infty}\frac{1}{(2n+1)^4} = \frac{\pi^4}{8}. \end{aligned}$$

Then,

$$\int_0^1 \frac{\ln(1-x)\ln^2(1+x)}{x} \, dx = -\frac{\pi^4}{240}$$

and

$$\sum_{n=2}^{\infty}(-1)^n\frac{H_n H_{n-1}}{n^2} = \frac{3}{16}\zeta(4) = \frac{\pi^4}{480}$$

as wanted.

□

An interesting integral related to Lambert's function.

Exercise 173. Prove that

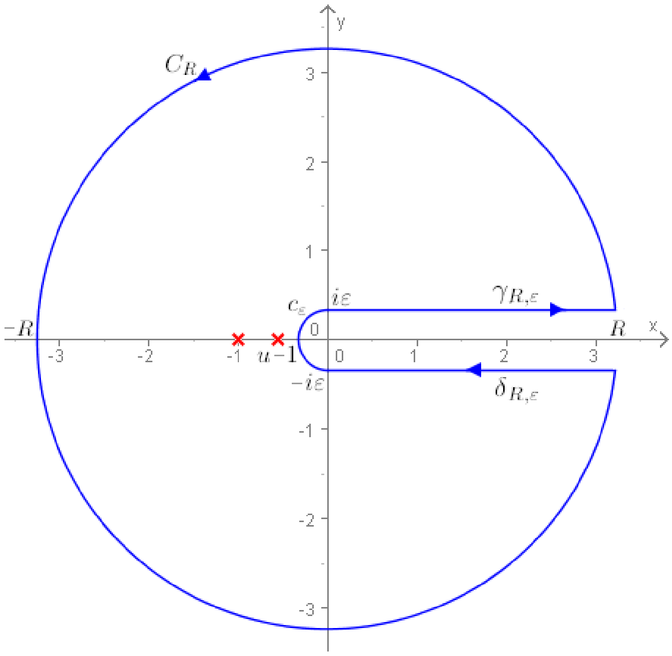
$$I_1 = \int_{-\infty}^{\infty}\frac{e^x+1}{(e^x-x+1)^2+\pi^2}\, dx = 1.$$

Proof. We are going to exploit the following Lemma:

Lemma 174. If $a > 0$ and $b \in \mathbb{R}$,

$$\int_{-\infty}^{+\infty}\frac{a^2\, dx}{(e^x-ax-b)^2+(a\pi)^2} = \frac{1}{1+W\left(\frac{1}{a}e^{-b/a}\right)}$$

where W is Lambert's function, i.e. the principal branch of the inverse function of $x \mapsto xe^x$.



Let us consider the function $f(z) = \frac{1}{a\log(-z)+b-z} \cdot \frac{1}{z}$ on the region $D = \mathbb{C} \setminus \mathbb{R}^{\geq 0}$. For any $z \in D$ it is possible to pick some $r > 0$ and some $\varphi \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ such that $-z = re^{i\varphi}$. This leads to:

$$a\log(-z) + b - z = a\log\left(re^{i\varphi}\right) + b + re^{i\varphi}$$

$$= a\log r + ia\varphi + b + r\cos\varphi + ir\sin\varphi$$

$$= (a\log r + r\cos\varphi + b) + i(a\varphi + r\sin\varphi)$$

where the imaginary part of the RHS has the same sign as φ and the real part is a monotonic function of the r variable. It follows that $z = -a \cdot W\left(\frac{1}{a}\, e^{-b/a}\right)$ is the only pole of f, attained at $\varphi = 0$ and $r = a \cdot W\left(\frac{1}{a}\, e^{-b/a}\right)$. A simple computation gives

$$\operatorname{Res}\left(f(z), z = -a \cdot W\left(\frac{1}{a}\, e^{-b/a}\right)\right) = \frac{1}{a} \cdot \frac{1}{1 + W\left(\frac{1}{a}\, e^{-b/a}\right)},$$

therefore, according to the residue Theorem:

$$\int_{\gamma_{R,\varepsilon}} f\, dz + \int_{C_R} f\, dz + \int_{\delta_{R,\varepsilon}} f\, dz + \int_{c_\varepsilon} f\, dz = \frac{1}{a} \cdot \frac{2\pi i}{1 + W\left(\frac{1}{a}\, e^{-b/a}\right)}$$

From $L(C_R) \sim 2\pi R$ and $M(C_R) = \max_{z \in C_R} |f(z)| \sim \frac{1}{R^2}$ it follows that $\left|\int_{C_R} f\, dz\right| \leq L(C_R)\, M(C_R) \sim \frac{2\pi}{R}$. Therefore the contribution of $\int_{C_R} f\, dz$ is negligible as $R \to +\infty$. From $L(c_\varepsilon) = \pi\varepsilon$ and $M(c_\varepsilon) = \max_{z \in c_\varepsilon} |f(z)| \sim \frac{-1}{a\log\varepsilon} \cdot \frac{1}{\varepsilon}$ it follows that $\left|\int_{c_\varepsilon} f\, dz\right| \leq L(c_\varepsilon)\, M(c_\varepsilon) \sim \frac{-\pi}{a\log\varepsilon}$, hence the contribution of $\int_{c_\varepsilon} f\, dz$ is negligible as $r \to 0^+$. By letting $R \to +\infty$ and $\varepsilon \to 0^+$ we get the equality

$$\int_0^\infty \left(\frac{1}{a\log(-x - i0^+) + b - x} - \frac{1}{a\log(-x + i0^+) + b - x}\right)\frac{dx}{x} = \frac{1}{a} \cdot \frac{2\pi i}{1 + W\left(\frac{1}{a}\, e^{-b/a}\right)}$$

and since:

$$\frac{1}{a(\log x - i\pi) + b - x} - \frac{1}{a(\log x + i\pi) + b - x} = \frac{2\pi i \cdot a}{(a\log x + b - x)^2 - (i\pi a)^2},$$

the previous identity has the following consequence:

$$\int_0^\infty \frac{a^2}{(a\log x + b - x)^2 + (a\pi)^2} \cdot \frac{dx}{x} = \frac{1}{1 + W\left(\frac{1}{a}\, e^{-b/a}\right)}.$$

The claim then follows by enforcing the substitution $x \mapsto e^x$, leading to:

$$\int_{-\infty}^{\infty} \frac{a^2}{(ax + b - e^x)^2 + (a\pi)^2}\, dx = \frac{1}{1 + W\left(\frac{1}{a}\, e^{-b/a}\right)}.$$

□

In our case, by choosing $a = 1$ and $b = -1$ we get that I_1 depends on $W(e) = 1$ and equals 1.

Due to the identity

$$\begin{aligned} I_1 &= \int_0^{+\infty} \frac{u+1}{u} \cdot \frac{du}{(u+1-\log u)^2+\pi^2} \\ &= \frac{1}{\pi}\int_0^{+\infty}\int_0^{+\infty} (u+1)u^{x-1}e^{-(u+1)x}\sin(\pi x)\,dx\,du \\ &= \frac{2}{\pi}\int_0^{+\infty} e^{-x}x^{-x}\Gamma(x)\sin(\pi x)\,dx \\ &= 2\int_0^{+\infty} \frac{e^{-x}x^{-x}}{\Gamma(1-x)}\,dx \end{aligned}$$

the previous Lemma also proves the highly non-trivial identity

$$\int_0^{+\infty} \frac{(ex)^{-x}}{\Gamma(1-x)}\,dx = \frac{1}{2} \qquad \text{(HNT)}$$

equivalent to the claim $I_1 = 1$. It would be interesting to find an independent proof of (HNT), maybe based on Glasser's master theorem, Ramanujan's master theorem or Lagrange inversion. There also is an interesting discrete analogue of (HNT),

$$\sum_{n\geq 1} \frac{n^n}{n!(4e)^{n/2}} = 1$$

that comes from the Lagrange inversion formula.

A remark on some series involving $\sinh$ **or** $\cosh$.
From the Weierstrass product

$$\cosh(\pi x/2) = \prod_{m\geq 0}\left(1+\frac{z^2}{(2m+1)^2}\right)$$

by applying $\frac{d^2}{dz^2}\log(\cdot)$ to both sides we get:

$$\frac{\pi^2}{8\cosh^2(\pi x/2)} = \sum_{m\geq 0} \frac{(2m+1)^2-z^2}{((2m+1)^2+z^2)}$$

If we replace z with $(2n+1)$ and sum over $n \geq 0$,

$$\sum_{n\geq 0} \frac{\pi^2}{8\cosh^2(\pi(2n+1)/2)} = \sum_{n\geq 0}\sum_{m\geq 0} \frac{(2m+1)^2-(2n+1)^2}{((2m+1)^2+(2n+1)^2)^2}$$

where the RHS of (3) can also be written as

$$\sum_{n\geq 0}\sum_{m\geq 0}\int_0^{+\infty} \cos((2n+1)x)xe^{-(2m+1)x}\,dx = \sum_{n\geq 0}\int_0^{+\infty} \frac{x\cos((2n+1)x)}{2\sinh(x)}\,dx$$

or, by exploiting integration by parts,

$$\sum_{n\geq 0}\int_0^{+\infty} \frac{x\cosh(x)-\sinh(x)}{2\sinh(x)}\cdot\frac{\sin((2n+1)x)}{2n+1}\,dx$$

On the other hand, $\sum_{n\geq 0}\frac{\sin((2n+1)x)}{2n+1}$ is the Fourier series of a 2π-periodic rectangle wave that equals $\frac{\pi}{4}$ over $(0,\pi)$ and $-\frac{\pi}{4}$ over $(\pi, 2\pi)$. That implies, by massive cancellation:

$$\sum_{n\geq 0} \frac{1}{\cosh^2(\pi(2n+1)/2)} = \frac{1}{2\pi}.$$

On the other hand, the Fourier transform of $\frac{1}{\cosh^2(\pi x)}$ is given by by $\frac{s\sqrt{8\pi}}{\sinh(\pi s)}$. By Poisson's summation formula,

$$\sum_{n\geq 1}\frac{(-1)^{n+1}n}{\sinh(\pi n)}=\frac{1}{4\pi}.$$

Exercise 175. Prove that

$$\int_0^{\pi/4}\arctan\left(\frac{\sqrt{2}\cos(3\phi)}{(3+2\cos(2\phi))\sqrt{\cos(2\phi)}}\right)d\phi=0.$$

Proof. By replacing ϕ with $\arctan(t)$, then using integration by parts, we have:

$$I=\int_0^1\frac{1}{1+t^2}\arctan\left(\frac{\sqrt{2}(1-3t^2)}{(5+t^2)\sqrt{1-t^2}}\right)dt=\frac{\pi^2}{8}-\int_0^1\frac{3\sqrt{2}\,t\arctan(t)}{(3-t^2)\sqrt{1-t^2}}\,dt.$$

Now comes the magic. Since:

$$\int\frac{3\sqrt{2}\,t}{(3-t^2)\sqrt{1-t^2}}\,dt=-3\arctan\sqrt{\frac{1-t^2}{2}}$$

integrating by parts once again we get:

$$I=\frac{\pi^2}{8}-3\int_0^1\frac{1}{1+t^2}\arctan\sqrt{\frac{1-t^2}{2}}\,dt$$

hence we just need to prove that:

$$\int_0^1\frac{dt}{1+t^2}\arctan\sqrt{\frac{1-t^2}{2}}=\int_0^{\frac{1}{\sqrt{2}}}\frac{\arctan\sqrt{1-2t^2}}{1+t^2}\,dt=\frac{\pi^2}{24}$$

and this is not difficult since both

$$\int_0^1\frac{dt}{1+t^2}(1-t^2)^{\frac{2m+1}{2}},\qquad\int_0^{\frac{1}{\sqrt{2}}}\frac{(1-2t^2)^{\frac{2m+1}{2}}}{1+t^2}\,dt$$

can be computed through the residue theorem or other techniques. For instance:

$$\int_0^1\frac{(1-t)^{\frac{2m+1}{2}}}{t^{\frac{1}{2}}(1+t)}\,dt=\sum_{n\geq 0}(-1)^n\int_0^1(1-t)^{\frac{2m+1}{2}}t^{n-\frac{1}{2}}\,dt=\sum_{n\geq 0}(-1)^n\frac{\Gamma\left(m+\frac{3}{2}\right)\Gamma\left(n+\frac{1}{2}\right)}{\Gamma(m+n+2)}$$

or just:

$$\int_0^1\frac{\sqrt{\frac{1-t^2}{2}}}{(1+t^2)\left(1+\frac{1-t^2}{2}u^2\right)}\,dt=\frac{\pi}{2(1+u^2)}\left(1-\frac{1}{\sqrt{2+u^2}}\right)$$

from which:

$$\int_0^1\frac{dt}{1+t^2}\arctan\sqrt{\frac{1-t^2}{2}}=\frac{\pi}{2}\int_0^1\frac{du}{1+u^2}\left(1-\frac{1}{\sqrt{2+u^2}}\right)=\frac{\pi^2}{24}$$

as wanted, since:

$$\int\frac{du}{(1+u^2)\sqrt{2+u^2}}=\arctan\frac{u}{\sqrt{2+u^2}}.$$

□

Exercise 176. Prove that for any $a > -2$ the following identity holds:

$$\int_0^1 \frac{1-x}{1+x} \cdot \frac{dx}{\sqrt{x^4+ax^2+1}} = \frac{1}{\sqrt{a+2}} \log\left(1+\frac{\sqrt{a+2}}{2}\right).$$

Proof. Simple algebraic manipulations allow us to write the given integral in the following form:

$$\int_0^{+\infty} \left(-1+\frac{2}{\sqrt{4+u^2}}\right) \frac{du}{u\sqrt{u^2+a+2}}$$

and by setting $u = \sqrt{a+2}\sinh\theta$ we get:

$$\frac{1}{\sqrt{a+2}} \int_0^{+\infty} \left(-1+\frac{2}{\sqrt{4+(a+2)\sinh^2\theta}}\right) \frac{d\theta}{\sinh\theta}$$

We may get rid of the last term through the "hyperbolic Weierstrass substitution"

$$\theta = 2\,\text{arctanh}(e^{-v}) = \log\left(\frac{e^v+1}{e^v-1}\right)$$

that wizardly gives

$$\frac{1}{\sqrt{a+2}} \int_0^{+\infty} \left(-1+\frac{2}{\sqrt{4+\frac{a+2}{\sinh^2 v}}}\right) dv$$

i.e. a manageable integral through differentiation under the integral sign.
To fill in the missing details is an exercise we leave to the reader. □

Exercise 177 (Ahmed's integral)**.** Prove that:

$$\int_0^1 \frac{\arctan\sqrt{x^2+2}}{(x^2+1)\sqrt{x^2+2}}\,dx = \frac{5\pi^2}{96}.$$

Proof. In 2001-2002, Zahar Ahmed proposed the above integral in the American Mathematical Monthly (AMM). Here we present his maiden solution. Let us call the given integral as I and use $\arctan z = \frac{\pi}{2} - \arctan\frac{1}{z}$ to split I as $I_1 - I_2$. Using the substitution $x = \tan\theta$, we can write

$$I_1 = \frac{\pi}{2} \int_0^{\pi/4} \frac{\cos\theta}{\sqrt{2-\sin^2\theta}}\,d\theta$$

which can be evaluated as $I_1 = \frac{\pi^2}{12}$ by using the substitution $\sin\theta = \sqrt{2}\sin\varphi$. Next we use the representation

$$\frac{1}{a}\arctan\frac{1}{a} = \int_0^1 \frac{dx}{x^2+a^2}$$

to express

$$I_2 = \int_0^1 \int_0^1 \frac{dx\,dy}{(1+x^2)(2+x^2+y^2)}.$$

By partial fraction decomposition I_2 can be re-written as:

$$I_2 = \int_0^1 \int_0^1 \frac{dx\,dy}{(1+x^2)(1+y^2)} - \int_0^1 \int_0^1 \frac{dx\,dy}{(1+x^2)(2+x^2+y^2)}.$$

Using the symmetry of the integrands and the domains for x and y, the second integral in the RHS of the last identity equals I_2 itself. This leads to:

$$I = \frac{\pi^2}{12} - \frac{1}{2}\left(\int_0^1 \frac{dx}{1+x^2}\right)^2 = \frac{\pi^2}{12} - \frac{\pi^2}{32} = \frac{5\pi^2}{96}.$$

□

Exercise 178. Prove the following identities:

$$\int_0^{\pi/2} \operatorname{arccot}\sqrt{1+\csc\theta}\,d\theta = \frac{\pi^2}{8}, \qquad \int_0^{\pi/2} \operatorname{arccsc}\sqrt{1+\cot\theta}\,d\theta = \frac{\pi^2}{12}.$$

Exercise 179. Prove the following identity:

$$\int_0^{1/\sqrt{2}} \frac{\arcsin(x^2)}{(2x^2+1)\sqrt{x^2+1}}\,dx = \frac{\pi^2}{144}.$$

Exercise 180. Prove the following identity:

$$\zeta(2) = \int_0^1 \arctan\left(\frac{88\sqrt{21}}{215+36x^2}\right)\frac{dx}{\sqrt{1-x^2}}.$$

Exercise 181. Prove the following identity:

$$\sum_{n\geq 1}\frac{H_n^2}{n^2} = \frac{17\pi^4}{360}.$$

Proof. It is very practical to recall that

$$-\log^3(1-z) = \sum_{n\geq 1}\frac{3\left(H_n^2 - H_n^{(2)}\right)}{n+1}z^{n+1},$$

immediately leading to the following intermediate identity:

$$\sum_{n\geq 1}\frac{H_n^2 - H_n^{(2)}}{(n+1)^2} = \frac{1}{3}\int_0^1 \frac{-\log^3(1-z)}{z}\,dz = \frac{1}{3}\int_0^1 \frac{-\log^3 z}{1-z}\,dz = \frac{\pi^4}{45}.$$

We may notice that

$$\begin{aligned} S = \sum_{n\geq 1}\frac{H_n^2}{n^2} &= \sum_{n\geq 1}\frac{H_{n-1}^2}{n^2} + 2\sum_{n\geq 1}\frac{H_{n-1}}{n^3} + \zeta(4) \\ &= \sum_{n\geq 1}\frac{H_{n-1}^2 - H_{n-1}^{(2)}}{n^2} + \sum_{n\geq 1}\frac{H_{n-1}^{(2)}}{n^2} + \frac{\pi^4}{60} \end{aligned}$$

since the value of $\zeta(4)$ is known and the Euler sum $\sum_{n\geq 1}\frac{H_{n-1}}{n^3}$ can be tackled through the Theorem (41). On the other hand, by symmetry:

$$\sum_{n\geq 1}\frac{H_{n-1}^{(2)}}{n^2}=\sum_{\substack{m,n\geq 1,\\ m<n}}\frac{1}{m^2n^2}=\frac{1}{2}\left[\zeta(2)^2-\zeta(4)\right]=\frac{\pi^4}{120}$$

from which it follows that

$$S=\frac{\pi^4}{45}+\frac{\pi^4}{120}+\frac{\pi^4}{60}=\frac{17\pi^4}{360}$$

as wanted. □

Exercise 182. Prove that by applying Feynman's trick to an integral representation for $\zeta(s)$ in the region $\text{Re}(s)\in(0,1)$ we have:

$$\int_0^{+\infty}\frac{\log(x)}{e^{2\pi x}+1}\,dx=-\frac{\log(2)\log(8\pi^2)}{4\pi}.$$

Exercise 183. By exploiting Euler's Beta function, convolutions and the discrete Fourier transform, prove that for any $n\in\mathbb{N}$ the following identity holds:

$$\sum_{k=0}^{2n}(-1)^k\binom{4n}{2k}\binom{2n}{k}^{-1}=\frac{1}{1-2n}.$$

Proof.

$$\begin{aligned}
S(n) &= (2n+1)\sum_{k=0}^{2n}(-1)^k\binom{4n}{2k}\int_0^1(1-x)^k x^{2n-k}\,dx\\
&= (2n+1)\sum_{k=0}^{2n}(-1)^k\binom{4n}{2k}\int_0^1 2z(1-z^2)^k z^{4n-2k}\,dx\\
&= -(2n+1)\int_0^{\pi/2}\sin(\theta)\cos(\theta)\left[e^{4ni\theta}+e^{-4ni\theta}\right]\,d\theta\\
&= -(2n+1)\int_0^{\pi/2}\sin(2\theta)\cos(4n\theta)\,d\theta\\
&= -\frac{2n+1}{4n^2-1}=-\frac{1}{2n-1}.
\end{aligned}$$

□

Exercise 184. By using suitable substitutions, the Laplace transform and special values of Γ',ζ,ζ', prove that:

$$\int_0^1\left(\frac{1}{\log t}+\frac{1}{1-t}\right)\frac{dt}{(1+t)^2}=\frac{1}{2}+\frac{\log 2}{3}-\frac{\log\pi}{4}-\frac{3}{2\pi^2}\sum_{n\geq 1}\frac{\log n}{n^2}.$$

Exercise 185. By enforcing the substitution $x\mapsto\frac{1-t}{1+t}$, check that

$$\int_0^1\frac{\arctan x}{\sqrt{x(1-x^2)}}\,dx=\sqrt{\tfrac{\pi}{2}}\,\Gamma\left(\tfrac{5}{4}\right)^2.$$

7 The Cauchy-Schwarz inequality and beyond

Definition 186. A **Hilbert space** is a vector space H (real or complex) equipped with a positive definite inner product $\langle\cdot,\cdot\rangle$ such that $(H, \|\cdot\|)$, where $\|u\|^2 = \langle u, u\rangle$, is a **complete** metric space.

Hilbert spaces like ℓ^2 or L^2 (respectively the space of square-summable sequences and the space of square-integrable functions) are the most natural places to extend the theory of inner products and orthogonal projections on $\mathbb{R}^n$: Fourier series and Fourier transforms are so *natural products* of this viewpoint. The context of Hilbert spaces is also the typical framework for important inequalities like Bessel's inequality (becoming Parseval's identity under the completeness assumption) and the **Cauchy-Schwarz inequality**:

Theorem 187 (Cauchy-Schwarz inequality)**.** If $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are n-uples of real numbers, we have:

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leq \left(\sum_{k=1}^{n} a_k^2\right) \cdot \left(\sum_{k=1}^{n} b_k^2\right)$$

and equality holds if and only if a n-uple is a real multiple of the other one.

The integral form of Cauchy-Schwarz inequality is sometimes mentioned as **Bunyakovskii's inequality**:

Theorem 188. If f, g are real, square-integrable functions over (a, b),

$$\left(\int_a^b f(x)\, g(x)\, dx\right)^2 \leq \left(\int_a^b f(x)^2 dx\right) \cdot \left(\int_a^b g(x)^2 dx\right)$$

and equality holds if and only if $\frac{f(x)}{g(x)} = \lambda$, or $\frac{g(x)}{f(x)} = \lambda$, almost everywhere in (a, b).

We start this section by studying multiple proofs of Cauchy-Schwarz inequalities, among "classical" approaches and less typical ones. The first proof depends on a **amplification** trick.
Proof # 1. Le us assume that u, v are non-zero vectors. From the trivial inequality

$$\|u - v\|^2 \geq 0$$

it follows immediately:

$$\langle u, v\rangle \leq \tfrac{1}{2}\|u\|^2 + \tfrac{1}{2}\|v\|^2$$

where by considering some $\lambda > 0$, the inner product $\langle u, v\rangle$ stays unchanged if u is replaced by λu and v is replaced by $\frac{v}{\lambda}$, implying:

$$\forall \lambda > 0, \qquad \langle u, v\rangle \leq \tfrac{\lambda^2}{2}\|u\|^2 + \tfrac{1}{2\lambda^2}\|v\|^2$$

and we have complete freedom in choosing λ in such a way that the RHS is as small as possible.
With the optimal choice $\lambda = \frac{\|v\|}{\|u\|}$ we get:

$$\langle u, v\rangle \leq \|u\| \cdot \|v\|$$

that is precisely Cauchy-Schwarz inequality: *the inner product between two vectors is bounded by the product of their lengths* [8]. It is pretty clear that equality holds only if u and v are linearly dependent, and this is even more evident as a consequence of the next proof.

[8] This is without doubt a really efficient mnemonic trick for recalling if CS holds as $\leq$ or $\geq$ when needed.

Theorem 189 (Lagrange's identity)**.**

$$\left(\sum_{i=1}^{n} a_i^2\right)\cdot\left(\sum_{i=1}^{n} b_i^2\right) = \left(\sum_{i=1}^{n} a_i b_i\right)^2 + \sum_{j\neq k}(a_j b_k - b_j a_k)^2.$$

To check the last identity is straightforward, and what is the best way for proving an inequality, than deriving it from an identity? Given Lagrange's identity, it is clear that Cauchy-Schwarz inequality holds as an equality if and only if $a_j b_k = a_k b_j$ for any $k \neq j$. [9]

The most well-known proof exploits the concept of **discriminant** for a quadratic polynomial. The function

$$p(t) = \sum_{i=1}^{n}(a_i - tb_i)^2$$

is clearly a quadratic polynomial, not assuming any negative value for any $t \in \mathbb{R}$ (as a sum of squares).
That implies the discriminant of $p(t)$ is non-positive, and from

$$[t^2]p(t) = \sum_{i=1}^{n} b_i^2, \qquad [t^1]p(t) = -2\sum_{i=1}^{n} a_i b_i, \qquad [t^0]p(t) = \sum_{i=1}^{n} a_i^2$$

Cauchy-Schwarz inequality immediately follows. Another approach, usually proposed to students as a tedious exercise, is to prove Cauchy-Schwarz inequality by induction on n. Such approach is not tedious anymore (quite the opposite, indeed) if we bring the Cauchy-Schwarz inequality into a (almost) equivalent form:

Lemma 190 (Titu)**.** If $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are n-uples of positive real numbers, we have:

$$\sum_{k=1}^{n} \frac{a_k^2}{b_k} \geq \frac{(a_1 + \ldots + a_n)^2}{b_1 + \ldots + b_n}.$$

We may firstly notice that Titu's Lemma is a consequence of the Cauchy-Schwarz inequality:

$$(a_1 + \ldots + a_n)^2 = \left(\frac{a_1}{\sqrt{b_1}}\cdot\sqrt{b_1} + \ldots + \frac{a_n}{\sqrt{b_n}}\cdot\sqrt{b_n}\right)^2 \stackrel{CS}{\leq} \left(\sum_{k=1}^{n}\frac{a_k^2}{b_k}\right)\cdot\left(\sum_{k=1}^{n} b_k\right).$$

Conversely, Titu's Lemma implies that Cauchy-Schwarz inequality holds for positive n-uples.
On the other hand the $n = 1$ case of Titu's Lemma is trivial and the $n = 2$ case

$$\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} \geq \frac{(a_1 + a_2)^2}{b_1 + b_2}$$

is equivalent to the inequality $(a_1 b_2 - a_2 b_1)^2 \geq 0$. By combining the cases $n = 2$ and $n = N$,

$$\begin{aligned} \sum_{k=1}^{N+1}\frac{a_k^2}{b_k} &= \frac{a_{N+1}^2}{b_{N+1}} + \sum_{k=1}^{N}\frac{a_k^2}{b_k} \\ \text{(case } n = N) \quad &\geq \frac{a_{N+1}^2}{b_{N+1}} + \frac{(a_1 + \ldots a_N)^2}{b_1 + \ldots + b_N} \\ \text{(case } n = 2) \quad &\geq \frac{(a_1 + \ldots + a_{N+1})^2}{b_1 + \ldots + b_{N+1}} \end{aligned}$$

the proof of Titu's lemma turns out to be straightforward.

[9] This rises a hystorical/epistemological caveat: since the Cauchy-Schwarz inequality immediately follows from Lagrange's identity, why such inequality was not attributed to Lagrange, whose work came before Cauchy's work by at least thirty years?

We may notice that the triangle inequality is a consequence of the Cauchy-Schwarz inequality:

$$\langle u, v\rangle \leq \|u\| \cdot \|v\| \longrightarrow \|u+v\|^2 \leq \|u\|^2 + \|v\|^2 + 2\|u\| \cdot \|v\| \longrightarrow \|u+v\| \leq \|u\| + \|v\|.$$

The Cauchy-Schwarz inequality can also be used to reconcile the typical definitions of *inner product* in Physics and Linear Algebra. By assuming that $u, v \in \mathbb{R}^2$, $u = (u_x, u_y)$, $v = (v_x, v_y)$, $\langle u, v\rangle$ can be equivalently defined as $u_x v_x + u_y v_y$ or as $\|u\|\|v\| \cos\theta$, where θ is the angle between the u and v vectors. How they come to be equivalent? Simple: on one hand, $\|u\|\|v\| \cos\theta$ is the product between the length of v and the length of the projection of u along v, that we may name w. w is a vector of the form λv minimizing $\|u - \lambda v\|$ or, equivalently:

$$\|u - \lambda v\|^2 = \|u\|^2 - 2\lambda(u_x v_x + u_y v_y) + \lambda^2 \|v\|^2.$$

Such quadratic polynomial in λ attains its minimum value at $\lambda = \frac{u_x v_x + u_y v_y}{\|v\|^2}$, reconciling the previous definitions and proving that two vectors have a zero inner product if and only if they are orthogonal.

Exercise 191. Given $A, B \in \mathbb{R}^+$, prove that

$$\max_\theta (A \sin\theta + B \cos\theta) = \sqrt{A^2 + B^2}.$$

Proof. Due to Cauchy-Schwarz inequality,

$$(A \sin\theta + B \cos\theta)^2 \leq A^2 + B^2$$

and equality holds as soon as $\tan\theta = \frac{A}{B}$. □

Exercise 192. Prove that for any $N \geq 1$ we have $H_N \leq \sqrt{2N}$.

Proof. Due to Cauchy-Schwarz inequality

$$H_N = \sum_{n=1}^{N} \frac{1}{N} \leq \sqrt{N \cdot \sum_{n=1}^{N} \frac{1}{n^2}} \leq \sqrt{N\,\zeta(2)}.$$

□

Exercise 193. Prove that if

$$\sum_{n\geq 1} a_n$$

is a convergent series with positive terms, then

$$\sum_{n\geq 1} \frac{1}{n^2 a_n}$$

is a divergent series.

Proof. Due to Cauchy-Schwarz inequality,

$$\left(\sum_{n=1}^{N} a_n\right) \cdot \left(\sum_{n=1}^{N} \frac{1}{n^2 a_n}\right) \geq \left(\sum_{n=1}^{N} \frac{1}{n}\right)^2 \geq \log^2(N).$$

□

Exercise 194. Prove that $\sqrt{3}\log(3) < 2$.

Proof.

$$\log(3) = \int_0^2 \frac{dx}{1+x} \overset{CS}{<} \sqrt{\int_0^2 \frac{dx}{(1+x)^2} \int_0^2 dx} = \frac{2}{\sqrt{3}}.$$

□

Exercise 195. Prove that

$$I = \int_0^{+\infty} \frac{\sin x}{x+1}\, dx \leq \sqrt{\frac{\pi}{8}}.$$

Proof. We are dealing with the integral of an oscillating function: in order to improve the convergence speed, it is very practical to consider that $\mathcal{L}(\sin x) = \frac{1}{s^2+1}$ and $\mathcal{L}^{-1}\left(\frac{1}{x+1}\right) = e^{-s}$, from which:

$$I = \int_0^{+\infty} \frac{e^{-s}}{s^2+1}\, ds \overset{CS}{\leq} \sqrt{\int_0^{+\infty} e^{-2s}\, ds \int_0^{+\infty} \frac{ds}{(1+s^2)^2}} = \sqrt{\frac{\pi}{8}}.$$

This approximation is quite accurate, since the functions e^{-2s} ed $\frac{1}{(1+s^2)^2}$ have a very similar behaviour in a right neighbourhood of the origin. □

Exercise 196. In a acute-angled triangle ABC, L_A, L_B, L_C are the feet of the angle bisectors from A, B, C. By denoting as r the inradius, prove that:

$$AL_A + BL_B + CL_C \geq 9r.$$

Proof. By naming I the incenter of ABC, from Van Obel's theorem it follows that:

$$\frac{AI}{IL_A} = \frac{b+c}{a}$$

and $IL_A \geq r$ is trivial. By combining these observations:

$$AL_A + BL_B + CL_C \geq r \cdot (a+b+c) \cdot \left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c}\right) \overset{CS}{\geq} 9r.$$

□

Exercise 197. Prove that if $p(x) \in \mathbb{R}[x]$ is a polynomial with non-negative coefficients, for any couple (x, y) of positive real numbers we have:

$$p(\sqrt{xy}) \leq \sqrt{p(x)p(y)}.$$

With greater generality, the concept of **convexity** is a cornerstone in the theory of inequalities.

Theorem 198 (Arithmetic-geometric inequality, AM-GM)**.** If $a_1, \ldots, a_n$ are non-negative real numbers, we have:

$$\mathrm{GM}(a_1, \ldots, a_n) \stackrel{\text{def}}{=} \sqrt[n]{a_1 \cdot \ldots \cdot a_n} \leq \frac{a_1 + \ldots + a_n}{n} \stackrel{\text{def}}{=} \mathrm{AM}(a_1, \ldots, a_n)$$

and equality holds iff $a_1 = \ldots = a_n$.

Proof. We may clearly assume that all the variables are positive without loss of generality.
In such a case, by setting $b_k = \log a_k$, the arithmetic-geometric inequality turns out to be equivalent to

$$\exp\left(\frac{1}{n}\sum_{k=1}^{n} b_k\right) \leq \frac{1}{n}\sum_{k=1}^{n} \exp(b_k)$$

that is Jensen's inequality for the exponential function, holding since $\frac{d^2}{dx^2}e^x = e^x > 0$. An alternative and really interesting approach is due to Cauchy. In his *Cours d'analyse* he observes the arithmetic-geometric inequality can be proved through an "atypical induction":

- AM-GM is trivial in the $n = 2$ case;
- if the AM-GM inequality holds for n variables, it also holds for $2n$ variables;
- if the AM-GM inequality holds for $(n+1)$ variables, it also holds for n variables.

□

Lemma 199 (Superadditivity of the geometric mean)**.** If $a_1, \ldots, a_n$ are positive and distinct numbers,

$$\forall \tau > 0, \qquad \mathrm{GM}(a_1 + \tau, \ldots, a_n + \tau) > \tau + \mathrm{GM}(a_1, \ldots, a_n).$$

Proof. The given inequality is equivalent to:

$$\prod_{k=1}^{n}(\tau + a_k) > (\tau + \sqrt[n]{a_1 \cdot \ldots \cdot a_n})^n$$

and by setting $b_k = \frac{a_k}{\tau}$ and $c_k = \log b_k$ it is also equivalent to:

$$\frac{1}{n}\sum_{k=1}^{n}\log(1 + e^{c_k}) > \log\left(1 + e^{\frac{c_1 + \ldots + c_n}{n}}\right),$$

i.e. Jensen's inequality for $f(x) = \log(1 + e^x)$. In order to prove the claim it is enough to notice that:

$$\frac{d^2}{dx^2}\log(1 + e^x) = \frac{e^x}{(e^x + 1)^2} > 0.$$

In a similar way we have that, if $(a_1, \ldots, a_n)$ and $(b_1, \ldots, b_n)$ are n-uples of non-negative numbers,

$$\mathrm{GM}(a_1 + b_1, \ldots, a_n + b_n) \geq \mathrm{GM}(a_1, \ldots, a_n) + \mathrm{GM}(b_1, \ldots, b_n).$$

□

Exercise 200 (Huygens inequality)**.** Prove that for any $\theta \in [0, 1]$ we have:

$$\tan\theta + 2\sin\theta \geq 3\theta.$$

Proof.
$$\tan\theta + 2\sin\theta = \int_0^\theta \left(\frac{1}{\cos^2 u} + \cos u + \cos u\right) du \overset{\text{AM-GM}}{\geq} 3\int_0^\theta 1\,d\theta = 3\theta.$$

□

Exercise 201 (Weitzenbock's inequality)**.** ABC is a triangle with area Δ and side lengths a, b, c. Prove that the equality $a^2 + b^2 + c^2 = 4\Delta\sqrt{3}$ implies that ABC is an equilateral triangle.

Proof. Due to Heron's formula
$$4\Delta = \sqrt{(a^2 + b^2 + c^2)^2 - 2(a^4 + b^4 + c^4)},$$
hence $a^2 + b^2 + c^2 = 4\Delta\sqrt{3}$ implies
$$(a^2 + b^2 + c^2)^2 = (1 + 1 + 1)(a^4 + b^4 + c^4),$$
that is the equality case in $(a^2 + b^2 + c^2)^2 \overset{\text{CS}}{\leq} (1 + 1 + 1)(a^4 + b^4 + c^4)$.
For other interesting proofs have a look at this thread on MSE. □

Exercise 202. Prove that for any $N \geq 1$ we have
$$\tfrac{1}{2}\left(\tfrac{3}{5} + \log N\right)^2 \geq \sum_{n=1}^{N}\left(n^{1/n} - 1\right) \geq \tfrac{1}{2}\log^2 N.$$

Proof. Since
$$n^{1/n} = \text{GM}\left(1, \frac{2}{1}, \ldots, \frac{n}{n-1}\right) \overset{\text{AM-GM}}{<} 1 + \frac{H_{n-1}}{n}$$
and
$$\sum_{n=1}^{N}\frac{H_{n-1}}{n} = \sum_{1\leq k<n\leq N}\frac{1}{kn} = \frac{H_N^2 - H_N^{(2)}}{2}$$
the inequality on the left follows from $H_N \leq \gamma + \log N$ and $\gamma < \frac{3}{5}$.
In order to prove the inequality on the right it is enough to show that for any $N \geq 1$ we have:
$$(N+1)^{\frac{1}{N+1}} - 1 \geq \tfrac{1}{2}\left[\log^2(N+1) - \log^2(N)\right].$$
To complete the proof is a task left to the reader. It might be useful to exploit:
$$\tfrac{1}{2}\left[\log^2(N+1) - \log^2(N)\right] = \int_N^{N+1}\frac{\log x}{x}\,dx.$$

□

Exercise 203. Prove that the sequence $\{a_n\}_{n\geq 1}$ defined by
$$a_n = \left(1 + \frac{1}{n}\right)^n$$
is increasing and bounded.

Proof. We have $a_{n+1} > a_n$ by AM-GM:

$$\sqrt[n+1]{1 \cdot a_n} = \text{GM}\left(1, 1+\frac{1}{n}, \ldots, 1+\frac{1}{n}\right) \overset{\text{AM-GM}}{<} \frac{1}{n+1}\left[1 + n\cdot\left(1+\frac{1}{n}\right)\right] = 1 + \frac{1}{n+1}.$$

Additionally:

$$\frac{a_{2n}}{a_n} = \left(1 + \frac{1}{4n(n+1)}\right)^n \leq \frac{1}{1 - \frac{1}{4(n+1)}} = 1 + \frac{1}{4n+3}$$

hence for any $N \geq 1$ we have:

$$a_N \leq a_1 \prod_{k\geq 0}\left(1 + \frac{1}{4\cdot 2^k + 3}\right) = \frac{16}{7}\prod_{k\geq 1}\left(1 + \frac{1}{4\cdot 2^k+3}\right) \leq \frac{16}{7}\prod_{k\geq 1}\frac{1+\frac{1}{2^{k+1}}}{1+\frac{1}{2^{k+2}}} = \frac{20}{7}.$$

□

Exercise 204 (from Baby Rudin, 1)**.** Let $\{a_n\}_{n\in\mathbb{N}}$ be an increasing and unbounded sequence of positive real numbers. Prove that there is a sequence $\{b_n\}_{n\in\mathbb{N}}$ of positive real numbers such that the series $\sum_{n\in\mathbb{N}} b_n$ is convergent but the series $\sum_{n\in\mathbb{N}} a_n b_n$ is divergent.

Exercise 205 (from Baby Rudin, 2)**.** Let $\{a_n\}_{n\in\mathbb{N}}$ be a sequence of positive real numbers, such that the series $\sum_{n\in\mathbb{N}} a_n$ is convergent. Prove the existence of a sequence $\{b_n\}_{n\in\mathbb{N}}$ of positive real numbers, increasing and unbounded, such that the series $\sum_{n\in\mathbb{N}} a_n b_n$ is convergent.

Proof. In the first exercise we may consider the sequence $\{b_n\}_{n\in\mathbb{N}}$ defined through:

$$b_0 = \frac{1}{\sqrt{a_0}}, \quad b_n = \frac{1}{\sqrt{a_{n-1}}} - \frac{1}{\sqrt{a_n}}.$$

The sequence $\{b_n\}_{n\in\mathbb{N}}$ is clearly increasing and we have:

$$\sum_{n\leq N} b_n = \frac{1}{\sqrt{a_0}} + \sum_{n=1}^{N}\left(\frac{1}{\sqrt{a_{n-1}}} - \frac{1}{\sqrt{a_n}}\right) = \frac{2}{\sqrt{a_0}} - \frac{1}{\sqrt{a_N}}.$$

Since the sequence $\{a_n\}_{n\in\mathbb{N}}$ is unbounded, the previous identity proves the series $\sum_{n\in\mathbb{N}} b_n$ is convergent to $\frac{2}{\sqrt{a_0}}$.

We further have:

$$a_n b_n = a_n \cdot \frac{\sqrt{a_n} - \sqrt{a_{n-1}}}{\sqrt{a_{n-1}a_n}} \geq \sqrt{a_n} - \sqrt{a_{n-1}}.$$

The last identity proves $\sum_{n=1}^{N} a_n b_n \geq \sqrt{a_N} - \sqrt{a_0}$, hence the series $\sum_{n\in\mathbb{N}} a_n b_n$ is divergent.
About the second exercise, we may assume $\sum_{n\in\mathbb{N}} a_n = 1$ without loss of generality, by multiplying every element of $\{a_n\}_{n\in\mathbb{N}}$ by a suitable constant. Under such assumption, we may consider the sequence $\{b_n\}_{n\in\mathbb{N}}$ defined through:

$$b_0 = c_0 = 1, \qquad b_n = \frac{1}{\sqrt{c_n}}, \qquad c_n = 1 - \sum_{m=0}^{n-1} a_m.$$

Since $\sum_{n\in\mathbb{N}} a_n$ is convergent to 1 and has positive terms, $\{c_n\}_{n\in\mathbb{N}}$ is a sequence with positive terms decreasing to 0, hence $\{b_n\}_{n\in\mathbb{N}}$ is positive and unbounded. We also have $a_n = c_n - c_{n+1}$, from which:

$$a_n b_n = \frac{c_n - c_{n+1}}{\sqrt{c_n}} = \frac{\left(\sqrt{c_n} - \sqrt{c_{n+1}}\right)\cdot\left(\sqrt{c_n} + \sqrt{c_{n+1}}\right)}{\sqrt{c_n}} \leq 2\left(\sqrt{c_n} - \sqrt{c_{n+1}}\right).$$

It follows that:

$$\sum_{n\leq N} a_n b_n \leq a_0 + 2\cdot\sum_{n\leq N}\left(\sqrt{c_n} - \sqrt{c_{n+1}}\right) = a_0 + 2\left(1 - \sqrt{c_{N+1}}\right),$$

hence the series $\sum_{n\in\mathbb{N}} a_n b_n$ is convergent to some value $\leq (a_0+2)$.
Both the exercises studied here follow from a more general inequality:

$$\forall\beta\in(0,1),\ \forall N\geq 1,\qquad \sum_{n=1}^{N} a_n\left(\sum_{m\geq n} a_m\right)^{-\beta} < \frac{1}{\beta}\left(\sum_{n=1}^{N} a_n\right)^{1-\beta},$$

holding for any sequence $\{a_n\}_{n=1}^{+\infty}$ with positive terms such that $\sum_{n=1}^{+\infty} a_n$ is convergent. □

Exercise 206. $\{a_n\}_{n\geq 1}$ is a sequence of real numbers with the following property: for any sequence of real numbers $\{b_n\}_{n\geq 1}$ such that $\sum_{n\geq 1} b_n^2$ is convergent, $\lim_{N\to+\infty}\sum_{n=1}^{N} a_n b_n$ exists and it is finite. Prove that $\{a_n\}_{n\geq 1}$ is square-summable, i.e. $\sum_{n\geq 1} a_n^2$ is convergent. *Hint*: assume that $\sum_{n\geq 1} a_n^2$ is divergent and consider $b_n = \frac{a_n}{A_n}$, where $A_N = a_1 + a_2 + \ldots + a_N$.

Exercise 207 (Hermite-Hadamard inequality)**.** Prove that if $f:[a,b]\to\mathbb{R}$ is a convex function,

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(x)\,dx \leq \frac{f(a)+f(b)}{2}.$$

In other terms: for convex (or concave) functions it is not difficult to estimate the magnitude of the error in computing $\int_a^b f(x)\,dx$ through the rectangle or trapezoid numerical methods.

Corollary 208. Since $f(x)=\log(x)$ is a concave function on $\mathbb{R}^+$,

$$n\log n - n + 1 = \int_1^n \log(x)\,dx \geq f(1)+f(2)+\ldots+f(n) = \log(n!).$$

Corollary 209. For any $n\in\mathbb{N}^+$,

$$\frac{\pi^2}{6} - \frac{1}{n+\frac{1}{2}} \leq H_n^{(2)} \leq \frac{\pi^2}{6} - \frac{1}{n+1}.$$

Exercise 210. Given $b>a>0$, provide an accurate lower bound for $\frac{\log b-\log a}{b-a}$ in terms of a rational function.

Proof. As a rule of thumb, when dealing with objects like $f(b)-f(a)$ it always is the case to wonder if such difference has a nice/practical integral representation. That is certainly the case here:

$$\delta(a,b) \stackrel{\text{def}}{=} \frac{\log b-\log a}{b-a} = \frac{1}{b-a}\int_a^b \frac{dx}{x} = \frac{1}{b-a}\int_0^{b-a}\frac{dx}{x+a} = \int_0^1 \frac{du}{(1-u)a+ub}$$

where by symmetry (i.e. by enforcing the substitution $u\to 1-u$):

$$\delta(a,b) = \frac{a+b}{2}\int_0^1 \frac{du}{((1-u)a+ub)\,(ua+(1-u)b)}.$$

This "*folding trick*" leads to an integrand function that is convex, almost-constant on $[0,1]$ and symmetric with respect to $u=\frac{1}{2}$. By applying the same trick a second time we get:

$$\delta(a,b) = \int_0^1 \frac{2(a+b)}{(a+b)^2-(b-a)^2t^2}\,dt,$$

hence by denoting $g_{a,b}(t) = \frac{2(a+b)}{(a+b)^2-(b-a)^2t^2}$ and by exploiting the Hermite-Hadamard inequality we get:

$$\delta(a,b) \geq g_{a,b}\left(\tfrac{1}{2}\right) = \frac{8(a+b)}{(a+3b)(3a+b)}.$$

As a straightforward corollary, $\log(2) \geq \frac{24}{35}$ holds, and such inequality is pretty tight. □

Theorem 211 (Karamata, Hardy-Littlewood)**.**
Let $(a_1, \ldots, a_k)$ and $(b_1, \ldots, b_k)$ be sequences of real numbers with the following properties:

- $\forall i < j,\ a_i \geq a_j$ and $b_i \geq b_j$ (weakly decreasing);
- $\forall i \in [1,k],\ A_i \stackrel{\text{def}}{=} \sum_{j=1}^{i} a_j \geq \sum_{j=1}^{i} b_j \stackrel{\text{def}}{=} B_i$ (the first sequence majorizes the second one);
- $\sum_{j=1}^{k}(a_j - b_j) = 0$ (same sum).

Karamata's inequality, also known as **Hardy-Littlewood's inequality** in its integral form, states that the previous assumptions grant that for every real convex function f:

$$\sum_{i=1}^{k} f(a_i) \geq \sum_{i=1}^{k} f(b_i).$$

Proof. If f is convex, the function

$$\delta_f(a,b) = \frac{f(b)-f(a)}{b-a}$$

is symmetric in its arguments and increasing with respect to each argument.
In the given hypothesis we may define

$$c_i = \delta_f(a_i, b_i),$$

then notice that:

$$\sum_{i=1}^{k}(f(a_i)-f(b_i)) = \sum_{i=1}^{k} c_i(a_i-b_i) = \sum_{i=1}^{k} c_i(A_i - A_{i-1} - B_i + B_{i-1}) = \sum_{i=1}^{k-1}(c_i - c_{i+1})(A_i - B_i),$$

so

$$c_i = \delta_f(a_i, b_i) \geq \delta_f(b_i, a_{i+1}) \geq \delta_f(a_{i+1}, b_{i+1}) = c_{i+1}$$

and the claim is proved. □

Theorem 212 (Young's inequality)**.** If a and b are positive real numbers, $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, we have:

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

where equality holds iff $a^p = b^q$.

Proof. Due to the concavity of the function $f(x) = \log x$ we have:

$$\log(ab) = \frac{1}{p}\log(a^p) + \frac{1}{q}\log(b^q) \leq \log\left(\frac{1}{p}a^p + \frac{1}{q}b^q\right),$$

then the claim follows by exponentiation. We may notice that, if $a^p \neq b^q$, the last inequality is tight.
As an alternative, one might consider that on $\mathbb{R}^+$ the function $f(x) = x^{p-1}$ has the inverse function $g(x) = x^{q-1}$.
By a well-known theorem on the integration of inverse functions (or simply by integration by parts) it follows that:

$$\frac{a^p}{p} + \frac{b^q}{q} = \int_0^a f(x)\,dx + \int_0^b g(x)\,dx \geq ab.$$

□

Hölder's inequality can be obtained from Young's inequality through an amplification trick:

Theorem 213 (Hölder's inequality)**.** If $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ are non-negative real numbers, $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$, we have:

$$\sum_{i=1}^{n} x_i y_i \leq \left(\sum_{i=1}^{n} x_i^p\right)^{1/p} \cdot \left(\sum_{i=1}^{n} y_i^q\right)^{1/q},$$

where equality holds if and only if for any $k \in [1, n]$ we have $x_k = \lambda y_k$.

Proof. By setting

$$\|x\|_p = \left(\sum_{i=1}^{n} x_i^p\right)^{1/p},$$

Hölder's inequality immediately follows from Young's inequality, since:

$$\frac{\sum_{i=1}^{n} x_i y_i}{\|x\|_p \cdot \|y\|_q} = \sum_{i=1}^{n} \frac{x_i}{\|x\|_p} \frac{y_i}{\|y\|_q} \leq \frac{1}{p} \frac{\|x\|_p^p}{\|x\|_p^p} + \frac{1}{q} \frac{\|y\|_q^q}{\|y\|_q^q} = \frac{1}{p} + \frac{1}{q} = 1.$$

Morever, it is not difficult to prove the following generalization: if p, q, r are positive real numbers and $\frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1$, then:

$$\sum_i |x_i y_i z_i| \leq \left(\sum_i |x_i|^p\right)^{\frac{1}{p}} \cdot \left(\sum_i |x_i|^q\right)^{\frac{1}{q}} \cdot \left(\sum_i |x_i|^r\right)^{\frac{1}{r}}.$$

□

Theorem 214 (Minkowski's inequality, triangle inequality for the L_p norm)**.** If $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ are non-negative real numbers and $p > 1$, by setting

$$\|x\|_p = \left(\sum_{i=1}^{n} x_i^p\right)^{1/p}$$

we have:

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p.$$

Proof.

$$\|x + y\|_p^p = \|(x + y)^p\|_1 \leq \|x(x + y)^{p-1}\|_1 + \|y(x + y)^{p-1}\|_1,$$

and due to Hölder's inequality we have:

$$\|x(x + y)^{p-1}\|_1 \leq \|x\|_p \cdot \|(x + y)^{p-1}\|_q = \|x\|_p \cdot \|x + y\|_p^{p-1},$$

from which:

$$\|x + y\|_p^p \leq (\|x\|_p + \|y\|_p) \|x + y\|_p^{p-1}.$$

□

Theorem 215 (Carleman's inequality). If $\{a_n\}_{n=1}^{+\infty}$ is a sequence of positive real numbers and the serie

$$\sum_{n=1}^{+\infty} a_n$$

is convergent to some $C \in \mathbb{R}$, we have:

$$\sum_{n=1}^{+\infty} \left(\prod_{i=1}^{n} a_i \right)^{1/n} \leq e\,C.$$

Proof. By denoting through GM the geometric mean, through AM the arithmetic mean :

$$\text{GM}(a_1, \ldots, a_n) = \text{GM}(a_1, 2a_2, \ldots, na_n)(n!)^{-1/n} \leq \text{AM}(a_1, 2a_2, \ldots, na_n)(n!)^{1/n},$$

where by Stirling's inequality we have $(n!)^{1/n} \leq \frac{e}{n+1}$ for any $n \geq 1$, hence:

$$\text{GM}(a_1, \ldots, a_n) \leq \frac{e}{n(n+1)} \sum_{k=1}^{n} k a_k.$$

It follows that:

$$\sum_{n=1}^{+\infty} \text{GM}(a_1, \ldots, a_n) \leq e \sum_{k=1}^{+\infty} \left(\sum_{n \geq k} \frac{1}{n(n+1)} \right) k a_k = e \sum_{k=1}^{+\infty} a_k.$$

We may notice that such inequality always holds as a strict inequality. As a matter of fact, by assuming

$$\text{GM}(a_1, 2a_2, \ldots, na_n) = \text{AM}(a_1, 2a_2, \ldots, na_n),$$

we have $a_k = \frac{D}{k}$ for some constant D, but the harmonic series is divergent. □

Exercise 216 (Indam test 2014, Exercise B3). Given $\sum_{n \geq 1} a_n$, convergent series with positive terms, prove that

$$\sum_{n \geq 1} (a_n)^{\frac{n-1}{n}}$$

is also a convergent series.

Proof. We may notice that $(a_n)^{\frac{n-1}{n}} \leq \frac{1+(n-1)a_n}{n}$ by AM-GM, but $\sum_{n \geq 1} \frac{1}{n}$ is a divergent series, so such argument does not prove the claim directly. By tweaking it a bit:

$$a_n^{\frac{n-1}{n}} = \text{GM}\left(\frac{1}{n}, 2a_n, \frac{3}{2} a_n, \ldots, \frac{n}{n-1} a_n \right)$$

leads, by AM-GM, to:

$$a_n^{\frac{n-1}{n}} \leq \frac{1}{n} \left(\frac{1}{n} + a_n \sum_{k=1}^{n-1} \frac{k+1}{k} \right) \leq \frac{1}{n^2} + \left(1 + \frac{\log n}{n} \right) a_n$$

hence

$$\sum_{n \geq 1} a_n^{\frac{n-1}{n}} \leq \frac{\pi^2}{6} + \left(1 + \frac{1}{e} \right) \sum_{n \geq 1} a_n.$$

An alternative approach is the following one: let $\alpha > 1$ some real number and let

$$S = \left\{ n : a_n \leq \frac{1}{\alpha^n} \right\}, \qquad L = \left\{ n : a_n > \frac{1}{\alpha^n} \right\}.$$

Our purpose is to prove that both $\sum_{n\in S}\frac{a_n}{\sqrt[n]{a_n}}$ and $\sum_{n\in L}\frac{a_n}{\sqrt[n]{a_n}}$ are convergent. By comparison with a geometric series, $\sum_{n\in S}\frac{a_n}{\sqrt[n]{a_n}} \leq \sum_{n\geq 1}\frac{1}{\alpha^{n-1}} = \frac{\alpha}{\alpha-1}$. On the other hand, if $n\in L$ then $a_n^{-1/n}\leq\alpha$, hence:

$$\sum_{n\geq 1}(a_n)^{\frac{n-1}{n}} \leq \frac{\alpha}{\alpha-1}+\alpha\sum_{n\geq 1}a_n$$

and by minimizing the RHS on α,

$$\sum_{n\geq 1}(a_n)^{\frac{n-1}{n}} \leq \left(1+\sqrt{\sum_{n\geq 1}a_n}\right)^2.$$

□

Theorem 217 (Hilbert's inequality)**.** If $\{a_m\}_{m\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ are two sequences in $\ell^2(\mathbb{R})$ the following inequality holds:

$$\sum_{n=1}^{+\infty}\sum_{m=1}^{+\infty}\frac{a_m b_n}{n+m} \leq \pi\sqrt{\sum_{m=1}^{+\infty}a_m^2}\cdot\sqrt{\sum_{n=1}^{+\infty}b_n^2}.$$

Proof. The proof we are going to present is essentially due to Schur, whose key idea is to prove the claim by a weighted version of the Cauchy-Schwarz inequality. For any positive real number λ we have:

$$\left(\sum_{m=1}^{+\infty}\sum_{n=1}^{+\infty}\frac{a_m b_n}{n+m}\right)^2 \leq \sum_{m=1}^{+\infty}\sum_{n=1}^{+\infty}\frac{a_m^2}{n+m}\left(\frac{m}{n}\right)^{2\lambda}\cdot\sum_{n=1}^{+\infty}\sum_{m=1}^{+\infty}\frac{b_n^2}{n+m}\left(\frac{n}{m}\right)^{2\lambda}.$$

We may notice that the first term in the RHS can be written as

$$\sum_{m=1}^{+\infty}a_m^2\sum_{n=1}^{+\infty}\frac{1}{m+n}\left(\frac{m}{n}\right)^{2\lambda},$$

so, by symmetry, it is enough to prove the existence of some $\lambda>0$ such that, for any $m\geq 1$:

$$\sum_{n=1}^{+\infty}\frac{1}{m+n}\left(\frac{m}{n}\right)^{2\lambda}\leq\pi.$$

Since the terms of the series on the LHS are positive and decreasing, the following inequality holds:

$$\sum_{n=1}^{+\infty}\frac{1}{m+n}\left(\frac{m}{n}\right)^{2\lambda} \leq \int_0^{+\infty}\frac{dx}{m+x}\frac{m^{2\lambda}}{x^{2\lambda}} = \int_0^{+\infty}\frac{dy}{(1+y)\,y^{2\lambda}} = \frac{\pi}{\sin(2\pi\lambda)},$$

so it is enough to choose $\lambda=\frac{1}{4}$ to finish. The π constant is optimal:
we leave to the reader to check that, if $a_n=b_n=n^{-1/2-\varepsilon}$,

$$\sum_{n=1}^{+\infty}\sum_{m=1}^{+\infty}\frac{a_m b_n}{n+m} = (\pi-O(\varepsilon))\sqrt{\sum_{m=1}^{+\infty}a_m^2}\cdot\sqrt{\sum_{n=1}^{+\infty}b_n^2}.$$

□

Exercise 218 (Hilber's inequality, ℓ^p version)**.**
Prove that if $p>1$, $\frac{1}{p}+\frac{1}{q}=1$, $\{a_m\}_{m\in\mathbb{N}}\in\ell^p(\mathbb{R})$, $\{b_n\}_{n\in\mathbb{N}}\in\ell^q(\mathbb{R})$, then

$$\left|\sum_{n=1}^{+\infty}\sum_{m=1}^{+\infty}\frac{a_m b_n}{n+m}\right| \leq \frac{\pi}{\sin\left(\frac{\pi}{p}\right)}\left(\sum_{m=1}^{+\infty}a_m^p\right)^{1/p}\cdot\left(\sum_{n=1}^{+\infty}b_n^q\right)^{1/q}.$$

Exercise 219 (A variant of Hilbert's inequality)**.**
Prove that if $\{a_n\}_{n\geq 1}$ and $\{b_n\}_{n\geq 1}$ are two sequences in $\ell^2(\mathbb{R}^+)$, then:

$$\left|\sum_{m\geq 1}\sum_{n\geq 1}\frac{a_m b_n}{\sqrt{m^2+n^2}}\right| \leq \frac{\Gamma\left(\frac{1}{4}\right)^2}{2\sqrt{\pi}}\,\|a\|_2\,\|b\|_2.$$

Theorem 220 (Hardy's inequality)**.** Let $p>1$ and let $a_1,\ldots,a_N$ be positive real numbers. By setting $A_k=\sum_{i=1}^{k}a_i$, we have:

$$\sum_{n=1}^{N}\left(\frac{A_n}{n}\right)^p \leq \left(\frac{p}{p-1}\right)^p\sum_{n=1}^{N}a_n^p.$$

In integral form: for every non-negative function in $L^p(\mathbb{R}^+)$,

$$\int_0^{+\infty}\left(\frac{1}{x}\int_0^x f(u)\,du\right)^p dx \leq \left(\frac{p}{p-1}\right)^p\int_0^{+\infty}f(x)^p dx.$$

Proof. The proof we are going to present is essentially due to Eliott.
We may firstly prove:

$$(\spadesuit)\qquad \sum_{n=1}^{N}\frac{A_n^p}{n^p}\leq\frac{p}{p-1}\sum_{n=1}^{N}\frac{a_n\,A_n^{p-1}}{n^{p-1}}.$$

Let $B_n=\frac{A_n}{n}$ and let Δ_n be the difference between the n-th terms in the RHS and LHS of ($\spadesuit$). We have:

$$\Delta_n \stackrel{\text{def}}{=} B_n^p-\frac{p}{p-1}\,a_nB_n^{p-1}=B_n^p-\frac{p}{p-1}\left(nB_n-(n-1)B_{n-1}\right)B_n^{p-1},$$

or:

$$\Delta_n = B_n^p\left(1-\frac{np}{p-1}\right)+\frac{p(n-1)}{p-1}B_{n-1}B_n^p.$$

As a consequence of Young's inequality (212) we have:

$$B_{n-1}B_n^{p-1}\leq\frac{B_{n-1}^p}{p}+(p-1)\frac{B_n^p}{p},$$

from which it follows that:

$$\Delta_n\leq\frac{n-1}{p-1}\,B_{n-1}^p-\frac{n}{p-1}\,B_n^p,$$

and by creative telescoping we may state:

$$\sum_{n=1}^{N}\Delta_n\leq-\frac{NB_N^p}{p-1}<0,$$

proving ($\spadesuit$). By applying Hölder's inequality (213) to the RHS of ($\spadesuit$) we have:

$$\sum_{n=1}^{N}\frac{A_n^p}{n^p}\leq\frac{p}{p-1}\sum_{n=1}^{N}\frac{a_nA_n^{p-1}}{n^{p-1}}\leq\frac{p}{p-1}\left(\sum_{n=1}^{N}a_n^p\right)^{1/p}\left(\sum_{n=1}^{N}\frac{A_n^p}{n^p}\right)^{(p-1)/p},$$

immediately proving the claim. □

Exercise 221. Let $\{p_n\}_{n=1}^{+\infty}$ be a sequence of positive real numbers such that the series $\sum_{n=1}^{+\infty}\frac{1}{p_n}$ is convergent. Prove that such assumptions grant that the series

$$\sum_{n=1}^{+\infty} p_n \frac{n^2}{(p_1+\ldots+p_n)^2}$$

is convergent as well.

Proof. For the sake of brevity, let us set:

$$C^2 = \sum_{n=1}^{+\infty}\frac{1}{p_n}, \qquad P_N = \sum_{i=1}^{N} p_i, \qquad S_N = \sum_{n=1}^{N} p_n \frac{n^2}{(p_1+\ldots+p_n)^2} = \sum_{n=1}^{N}\frac{n^2\,p_n}{P_n^2}.$$

Since $\{P_N\}_{N=1}^{+\infty}$ is an increasing sequence, we have:

$$S_N = \frac{1}{p_1} + \sum_{n=2}^{N}\frac{n^2(P_n-P_{n-1})}{P_n^2} < \frac{1}{p_1} + \sum_{n=2}^{N}\frac{n^2(P_n-P_{n-1})}{P_nP_{n-1}} = \frac{1}{p_1} + \sum_{n=2}^{N}\left(\frac{n^2}{P_{n-1}} - \frac{n^2}{P_n}\right),$$

$$S_N < \frac{5}{P_1} + \left(\sum_{n=2}^{N-1}\frac{2n+1}{P_n}\right) - \frac{N^2}{P_N} < 5\sum_{n=1}^{N}\frac{n}{P_n}.$$

By exploiting the Cauchy-Schwarz inequality we also have:

$$\sum_{n=1}^{N}\frac{n}{P_n} \leq \sqrt{\sum_{n=1}^{N}\frac{1}{p_n}}\cdot\sqrt{\sum_{n=1}^{N}\frac{n^2\,p_n}{P_N^2}} \leq C\,\sqrt{S_n},$$

from which it follows that:

$$S_N < 5C\,\sqrt{S_N},$$

or:

$$S_N < 25\,C^2.$$

Since such inequality holds for any N, and since the sequence $\{S_N\}_{N=1}^{+\infty}$ is increasing, the following series is convergent by the monotone convergence Theorem:

$$\sum_{n=1}^{+\infty}\frac{n^2\,p_n}{P_n^2}.$$

It is not difficult to prove that we actually have the sharper inequality

$$\sum_{n=1}^{N}\frac{2n+1}{P_n} < 4\sum_{n=1}^{N}\frac{1}{p_n},$$

from which we may derive the more accurate bound:

$$S_N < \frac{2}{a_1} + 4C^2.$$

□

Exercise 222. Prove that if $\sum_{n=1}^{+\infty}\frac{1}{a_n}$ is a convergent series with positive real terms, there exists a constant $C\in\mathbb{R}$ such that:

$$\sum_{n=1}^{+\infty}\frac{n}{a_1+\ldots+a_n} \leq C\sum_{n=1}^{+\infty}\frac{1}{a_n}.$$

Proof. Due to the AM-GM inequality we have that:

$$\sum_{n=1}^{+\infty} \frac{n}{a_1+\ldots+a_n} \leq \sum_{n=1}^{+\infty} \frac{1}{\text{GM}(a_1,\ldots,a_n)} = \sum_{n=1}^{+\infty} \text{GM}\left(\frac{1}{a_1},\ldots,\frac{1}{a_n}\right),$$

where the RHS, by Carleman's inequality (215), is bounded by:

$$e\sum_{n=1}^{+\infty}\frac{1}{a_n},$$

so the given inequality holds for sure by taking $C = e$. In the next exercise we will see that such result can be improved: the given inequality holds for $C = 2$, that is the optimal constant. □

Exercise 223. Prove that if $a_1,\ldots,a_n$ are positive real numbers,

$$\sum_{k=1}^{n} \frac{2k+1}{a_1+a_2+\ldots+a_k} < 4\sum_{k=1}^{n}\frac{1}{a_k}.$$

Proof. We recall that, by exploiting the Cauchy-Schwarz inequality in the form of Titu's Lemma,

Lemma 224. If $r_1, r_2, \alpha, \beta, \gamma$ are positive real numbers and $\gamma = \alpha + \beta$, we have:

$$\frac{\gamma^2}{r_1+r_2} \leq \frac{\alpha^2}{r_1} + \frac{\beta^2}{r_2}.$$

This Lemma implies:

$$\frac{(n+1)^2}{a_1+\ldots+a_n} + \frac{2n-1}{a_1+\ldots+a_{n-1}} \leq \frac{4}{a_n} + \frac{2n-1+(n-1)^2}{a_1+\ldots+a_{n-1}} = \frac{4}{a_n} + \frac{n^2}{a_1+\ldots+a_{n-1}},$$

from which it follows that:

$$\frac{n^2}{a_1+\ldots+a_n} + \sum_{k=1}^{n} \frac{2k+1}{a_1+a_2+\ldots+a_k} \leq \frac{4}{a_n} + \frac{(n-1)^2}{a_1+\ldots+a_{n-1}} + \sum_{k=1}^{n-1}\frac{2k+1}{a_1+a_2+\ldots+a_k},$$

and by induction:

$$\frac{n^2}{a_1+\ldots+a_n} + \sum_{k=1}^{n} \frac{2k+1}{a_1+a_2+\ldots+a_k} \leq \left(\sum_{k=2}^{n}\frac{4}{a_k}\right) + \frac{1}{a_1} + \frac{3}{a_1} = \sum_{k=1}^{n}\frac{4}{a_k}.$$

□

Exercise 225. Prove that for any $p > 1$ and for any $a, b, \alpha, \beta > 0$ we have:

$$\left(\frac{(\alpha+\beta)^{p+1}}{a^p+b^p}\right)^{1/p} \leq \left(\frac{\alpha^{p+1}}{a^p}\right)^{1/p} + \left(\frac{\beta^{p+1}}{b^p}\right)^{1/p}.$$

Proof. If we set $b/a = x$, it is enough to show that the minimum of the function $f : \mathbb{R}^+ \to \mathbb{R}^+$ defined by:

$$f(x) = \alpha^{\frac{p+1}{p}}(1+x^p)^{1/p} + \beta^{\frac{p+1}{p}}(1+x^{-p})^{1/p}$$

is exactly $(\alpha+\beta)^{\frac{p+1}{p}}$. In order to do that, it is enough to check that $f'(x)$ vanishes only at

$$x = \left(\frac{\beta}{\alpha}\right)^{1/p}.$$

□

Theorem 226 (Knopp)**.** For any real number $p \geq 1$ there exists some constant $C_p \in \mathbb{R}^+$ such that for any sequence $a_1, \ldots, a_N$ of positive real numbers,

$$\sum_{n=1}^{N}\left(\frac{n}{a_1^p+\ldots+a_n^p}\right)^{1/p} < C_p \sum_{n=1}^{N}\frac{1}{a_n}.$$

Proof. In a similar way to Exercise 223, we prove there is a positive and increasing function $f:\mathbb{N}_0 \to \mathbb{R}^+$ such that:

$$(\diamondsuit)\ \left(\frac{f(N)}{\sum_{n=1}^{N}a_n^p}\right)^{1/p}+\sum_{n=1}^{N}\left(\frac{n}{a_1^p+\ldots+a_n^p}\right)^{1/p} \leq \frac{C_p}{a_N}+\left(\frac{f(N-1)}{\sum_{n=1}^{N-1}a_n^p}\right)^{1/p}+\sum_{n=1}^{N-1}\left(\frac{n}{a_1^p+\ldots+a_n^p}\right)^{1/p},$$

granting that, by induction, we have:

$$\left(\frac{f(N)}{\sum_{n=1}^{N}a_n^p}\right)^{1/p}+\sum_{n=1}^{N}\left(\frac{n}{a_1^p+\ldots+a_n^p}\right)^{1/p} \leq \frac{1+f(1)^{1/p}}{a_1}+\sum_{n=2}^{N}\frac{C_p}{a_n}.$$

In order that $(\diamondsuit)$ implies the claim it is enough that $f(1) \leq (C_p-1)^p$ holds and:

$$\forall N \geq 2, \qquad \left(\frac{f(N)}{\sum_{n=1}^{N}a_n^p}\right)^{1/p}+\left(\frac{N}{\sum_{n=1}^{N}a_n^p}\right)^{1/p} \leq \frac{C_p}{a_n}+\left(\frac{f(N-1)}{\sum_{n=1}^{N-1}a_n^p}\right)^{1/p}.$$

By exploiting the inequalities proved in 225, by assuming $f(N)^{1/p}+N^{1/p} \geq C_p^{p/(p+1)}$ we have:

$$\frac{f(N)^{1/p}+N^{1/p}}{\left(\sum_{n=1}^{N}a_n^p\right)^{1/p}} \leq \frac{C_p}{a_N}+\frac{\left(\left(f(N)^{1/p}+N^{1/p}\right)^{\frac{p}{p+1}}-C_p^{\frac{p}{p+1}}\right)^{\frac{p+1}{p}}}{\left(\sum_{n=1}^{N-1}a_n^p\right)^{1/p}},$$

hence it is enough to find some function f such that:

$$\left(f(N)^{1/p}+N^{1/p}\right)^{\frac{p}{p+1}} \leq f(N-1)^{\frac{1}{p+1}}+C_p^{\frac{p}{p+1}}.$$

Now we may consider $C_p=(1+p)^{\frac{1}{p}}$, that is the best constant we may put in the RHS of the initial inequality, if $a_n=n$. Then we consider $f(N)=k\cdot N^{p+1}$: the previous inequality becomes:

$$(\spadesuit)\quad k^{\frac{1}{p+1}}N\left(1+\frac{1}{Nk^{1/p}}\right)^{\frac{p}{p+1}} \leq k^{\frac{1}{p+1}}(N-1)+(1+p)^{\frac{1}{p+1}}.$$

Due to Bernoulli's inequality we have:

$$\left(1+\frac{1}{Nk^{1/p}}\right)^{\frac{p}{p+1}} \leq 1+\frac{p}{N(p+1)k^{1/p}},$$

so if we have some k such that:

$$(\heartsuit)\qquad \frac{p}{p+1}k^{-\frac{1}{p(p+1)}}+k^{\frac{1}{p+1}} \leq C_p^{\frac{p}{p+1}}=(p+1)^{\frac{1}{p+1}}$$

the inequality $(\spadesuit)$ is fulfilled. By studying the stationary points of the function $g(x)=Ax^{-\alpha}+x^{\beta}$ it is simple to derive that the choice

$$k=(p+1)^{-p}$$

leads to an equality in $(\heartsuit)$. It just remains to prove that with the choice $f(N)=\frac{N^{p+1}}{(p+1)^p}$ we have $f(1) \leq (C_p-1)^p$, or $C_p \geq 1+\frac{1}{p+1}$, i.e.:

$$(p+1)^{\frac{1}{p}} \geq 1+\frac{1}{p+1}.$$

By multiplying both sides by $(p+1)$ we get that such inequality is equivalent to:

$$(p+1)^{\frac{p+1}{p}} \geq p+2,$$

again a consequence of Bernoulli's inequality, since:

$$(p+1)^{\frac{p+1}{p}} \geq 1+\frac{p+1}{p}\cdot p = p+2.$$

Summarizing, we proved that for any $p \geq 1$ and for any sequence $a_1,\ldots,a_N$
of positive real numbers we have:

$$\frac{N^{\frac{p+1}{p}}}{(p+1)\left(a_1^p+\ldots+a_N^p\right)^{1/p}}+\sum_{n=1}^{N}\left(\frac{n}{a_1^p+\ldots+a_n^p}\right)^{1/p} \leq (1+p)^{\frac{1}{p}}\sum_{n=1}^{N}\frac{1}{a_n},$$

that essentially is a generalization of Hardy's inequality (220) to negative exponents. We leave to the reader to prove that the shown C_p constant is optimal, since by assuming $a_n = n$ and letting $N \to +\infty$, it is clear it cannot be replaced by any smaller real number. □

Theorem 227 (Brunn-Minkowski)**.** If A and B are two compact subsets of $\mathbb{R}^n$ and μ is the n-dimensional Lebesgue measure,

$$\mu(A+B)^{\frac{1}{n}} \geq \mu(A)^{\frac{1}{n}}+\mu(B)^{\frac{1}{n}}.$$

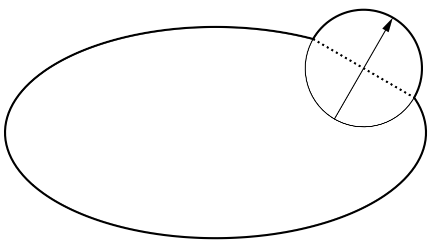
If both A and B are given by cartesian products of closed intervals (let us say *boxes*) the given inequality is trivial by AM-GM. By an ingenious trick known as *Hadwiger-Ohmann's cut* it is possible to show it continues to hold for disjoint unions of *boxes*, hence the claim follows from the regularity of the Lebesgue measure.

Equality is achieved only if A and B are homothetic shapes, i.e. can be brought one into the other by the composition of a uniform dilation and a translation. This inequality is extremely powerful and it can be employed to prove the **isoperimetric inequality** in n dimensions.

Theorem 228 (Isoperimetric inequality in the plane)**.** If γ is a regular, closed and simple curve, we have:

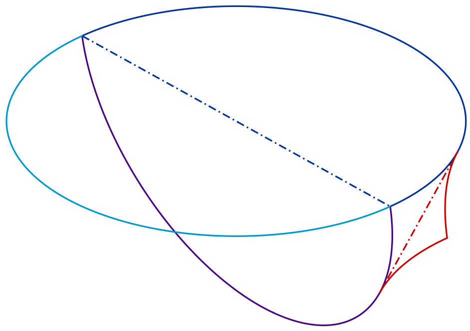
$$4\pi A \leq L^2$$

where L is the length of γ and A is the area enclosed by γ. Equality holds if and only if γ is a circle.



For a start, we show the most classical and elementary approaches. A key observation is that if a simple, regular and closed curve with a given length encloses the maximum area, it has to be convex. Otherwise we might apply a reflection to an arc of such curve, increasing the enclosed area without affecting the length.

On the other hand, given a closed and convex curve, we may consider an arbitrary chord and apply a reflection with respect to the perpendicular bisector of such chord to one of the arcs cut. Through such transform, both the perimeter and the enclosed area stay unchanged. In particular any curve with a given length that is regular, simple, closed and maximizes the enclosed area is mapped by any of such transforms into a convex curve: otherwise it would be possible to increase the enclosed area without affecting the length, as depicted on the right.



Given these considerations it is not difficult to show that the circle is the only solution of the isoperimetric problem in the plane. As an alternative, one may follow a **discretization** approach.

Lemma 229. Among all the simple and closed polygonal lines $P_1P_2\ldots P_n$ ($P_{n+1}=P_1$) having sides $l_i = P_iP_{i+1}$ with fixed lengths, the polygonal line enclosing the greatest area is such that $P_1,\ldots,P_n$ are vertices of a cylic polygon.

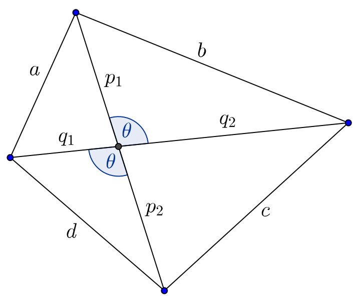
We may prove the claim by induction on n, by considering first the case $n=4$.
That case can be tackled through **Ptolemy's inequality**.

Theorem 230 (Ptolemy)**.** If A, B, C, D (in this ordering) are the vertices of a convex quadrilateral in the plane, we have:

$$AB\cdot CD + BC\cdot DA \geq AC\cdot BD$$

and equality holds if and only if $ABCD$ is a cyclic quadrilateral.

Proof. It is enough to apply a **circle inversion** with respect to a unit circle centered at A, then consider how distances change under circle inversion. □



Let us consider the configuration on the side and set $p = p_1 + p_2$ and $q = q_1 + q_2$. We have:

$$\left\{\begin{array}{rcl} a^2 &=& p_1^2+q_1^2-2p_1q_1\cos(\pi-\theta)\\ b^2 &=& p_1^2+q_2^2-2p_1q_2\cos(\theta)\\ c^2 &=& p_2^2+q_2^2-2p_2q_2\cos(\pi-\theta)\\ d^2 &=& p_2^2+q_1^2-2p_2q_1\cos(\theta)\end{array}\right.$$

from which it follows that:

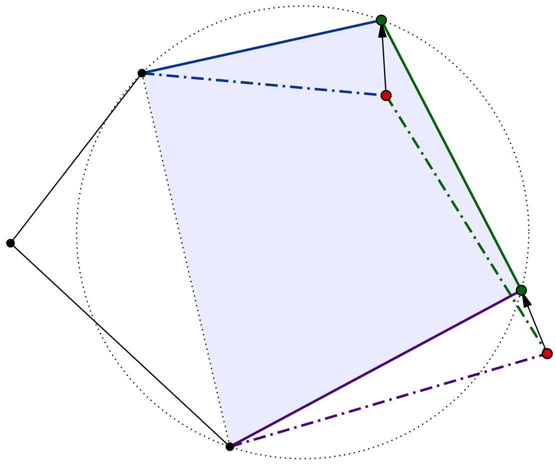
$$(a^2-b^2+c^2-d^2) = 2(p_1+q_1)(p_2+q_2)\cos\theta = 2pq\cos\theta.$$

By denoting as A the area of the depicted quadrilateral we have:

$$\begin{array}{rcl} 16A^2 &=& (2pq\sin\theta)^2\\ &=& (2pq)^2-(2pq\cos\theta)^2\\ &\leq& (2ac+2bd)^2-(a^2-b^2+c^2-d^2)^2\\ &=& (a+b+c-d)(a+b-c+d)(a-b+c+d)(-a+b+c+d)\end{array}$$

and $\leq$ holds as an equality iff the previous quadrilateral is cyclic. Moreover, given some closed and simple polygonal line $P_1, P_2, \ldots, P_n$ ($P_{n+1} = P_1$) having sides with fixed lengths, it is always possible to rearrange its vertices in such a

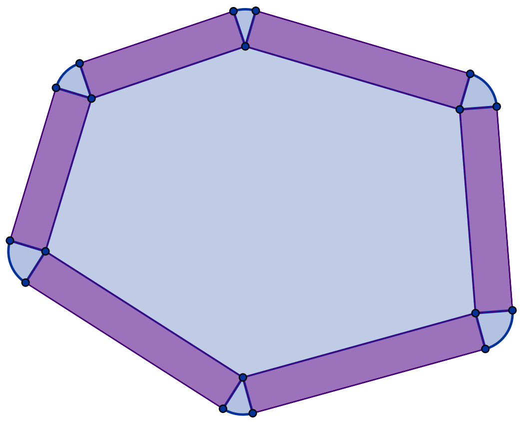
way they lie on the same circle. Indeed we may place $Q_1, Q_2, \ldots, Q_n, Q_{n+1}$ on a circle with a huge radius, in such a way that $Q_1Q_2 = P_1P_2, \ldots, Q_nQ_{n+1} = P_nP_{n+1}$ hold, then slowly decrease the radius of such circle, letting $Q_1, \ldots, Q_{n+1}$ "slide" on it while preserving the mutual distances. By continuity, there exists some radius such that $Q_{n+1} \equiv Q_1$, leading to a cyclic rearrangement of the polygonal line.
In particular, if $P_1P_2 \ldots P_n$ is a closed and simple polygonal line with $n \geq 4$ vertices, enclosing the largest possible area, all the quadrilaterals $P_kP_{k+1}P_{k+2}P_{k+3}$ are cyclic. Otherwise by leaving P_k and P_{k+3} where they are and by rearranging P_{k+1} and P_{k+2} we would increase the enclosed area. This proves Lemma (229).



Cyclic rearrangement of 4 vertices:
the enclosed area increases.

Since every regular curve is uniformly approximated by a polygonal line, and since all the solutions to the isoperimetric problem in the plane (i.e. the curves of fixed length enclosing the largest possible area) are regular functions, the isoperimetric inequality follows from Lemma (229) "by sending n towards $+\infty$". However the Brunn-Minkowski inequality provides a less involved approach.

Lemma 231. If $P_1, P_2, \ldots, P_n$ are vertices of a convex polygon K having perimeter L and area A, and B_r is a circle with radius r, the area of $K + B_r$ is given by $A + Lr + \pi r^2$.



In order to prove the Lemma it is enough to decompose $K + B_r$ as the union of K, n rectangles with height r and bases on the sides of K, n circle sectors corresponding to a partition of B_r.

By approximating regular curves through polygonal lines, if K is the region enclosed[10] by a regular, simple and closed curve with length L, we have:

$$\mu(K + B_r) = \mu(K) + rL + \pi r^2.$$

where μ is the 2-dimensional Lebesgue measure.
Due to Minkowski's inequality with $n = 2$:

$$\sqrt{\mu(K) + rL + \pi r^2} \geq \sqrt{\mu(K)} + \sqrt{\pi r^2},$$

inequality which is equivalent to $L^2 \geq 4\pi \cdot \mu(K)$, i.e. to the isoperimetric inequality.

Equality holds if and only if, up to translations, $K = \lambda B_r$, meaning that K is a circle with radius λr. The last approach can be easily extended to the $n > 2$ case: *if the surface area of the boundary is fixed, the suitable closed balls with respect to the Euclidean norm are convex sets enclosing the maximum volume.* The last proof of the isoperimetric inequality we are going to see has a more *analytic flavour*, and it is deeply related to the **Poincaré-Wirtinger inequality**.

[10]We used the letter K twice on purpose. The solutions to the isoperimetric problem are given by convex sets and, since every regular and convex curve is a limit of boundaries of convex polygons, in the isoperimetric problem the continuous and discrete approaches are equivalent.

Theorem 232 (Wirtinger's inequality for functions). (Version I) If $f \in C^1(\mathbb{R})$ is a 2π-periodic function such that $\int_0^{2\pi} f(\theta)\,d\theta = 0$, we have:

$$\int_0^{2\pi} f(\theta)^2\,d\theta \leq \int_0^{2\pi} f'(\theta)^2\,d\theta$$

where equality holds iff $f(\theta) = a\sin\theta + b\cos\theta$.
(Version II) If f is a function of class C^1 on the interval $[0, 2\pi]$ and $f(0) = f(2\pi) = 0$, we have:

$$\int_0^{2\pi} f(\theta)^2\,d\theta \leq 4\int_0^{2\pi} f'(\theta)^2\,d\theta$$

where equality holds iff $f(\theta) = \kappa\sin\frac{\theta}{2}$.

Proof. By assuming the following identities

$$\begin{aligned} f(\theta) &= \sum_{n\geq 1} s_n \sin(n\theta) + \sum_{n\geq 1} c_n\cos(n\theta), \\ f'(\theta) &= \sum_{n\geq 1} ns_n\cos(n\theta) - \sum_{n\geq 1} nc_n \sin(n\theta) \end{aligned}$$

hold in the L^2-sense, due to Parseval's identity we get:

$$\int_0^{2\pi} f(\theta)^2\,d\theta = \pi\sum_{n\geq 1}(s_n^2 + c_n^2) \leq \pi\sum_{n\geq 1} n^2(s_n^2 + c_n)^2 = \int_0^{2\pi} f'(\theta)^2\,d\theta$$

and $\leq$ holds as an equality iff $c_2 = c_3 = \ldots = s_2 = s_3 = \ldots = 0$. This proves the first version of Wirtinger's inequality. When proving the second version we may assume without loss of generality

$$\begin{aligned} f(\theta) &= \sum_{n\geq 1} s_n\sin(n\theta/2) \\ f'(\theta) &= \tfrac{1}{2}\sum_{n\geq 1} ns_n\cos(n\theta/2) \end{aligned}$$

(in the L^2-sense) and the claim follows again from Parseval's identity. □

We may now consider that any regular, simple and closed curve with length L has an **arc length parametrization**, i.e. a couple of piecewise-C^1, L-periodic functions $x(s), y(s)$ such that

$$\left(\frac{dx}{ds}\right)^2 + \left(\frac{dy}{ds}\right)^2 = 1.$$

By introducing

$$f(\theta) \stackrel{\text{def}}{=} x\left(\frac{L\theta}{2\pi}\right), \qquad g(\theta) \stackrel{\text{def}}{=} y\left(\frac{L\theta}{2\pi}\right)$$

we have that the area enclosed by γ, as a consequence of Green's Theorem, is given by the integral $\int_0^{2\pi} f(\theta)\,g'(\theta)\,d\theta$. By defining $\overline{f}$ as the mean value of f on the interval $[0, 2\pi]$, and by noticing that the mean value of $g'(\theta)$ is zero, we have:

$$\begin{aligned} A = \int_0^{2\pi} f(\theta)\,g'(\theta)\,d\theta &= \int_0^{2\pi} (f(\theta) - \overline{f})\,g'(\theta)\,d\theta \\ &\leq \frac{1}{2}\int_0^{2\pi}\left[(f(\theta) - \overline{f})^2 + g'(\theta)^2\right]\,d\theta \\ &\stackrel{\text{Wirt}}{\leq} \frac{1}{2}\int_0^{2\pi}\left[f'(\theta)^2 + g'(\theta)^2\right]\,d\theta \\ &= \frac{1}{2}\int_0^{2\pi}\frac{L^2 d\theta}{4\pi^2} = \frac{L^2}{4\pi}. \end{aligned}$$

Equality holds if and only if $f(\theta)-\overline{f} = a\sin\theta + b\cos\theta$ and $g'(\theta) = f(\theta)-\overline{f}$, i.e. only when γ is a circle. It is interesting to point out that the given proof can also be reversed, proving that Wirtinger's inequality for functions (or, at least, its first version) is a consequence of the isoperimetric inequality in the plane.

Exercise 233 (Dido's problem)**.** A rich landowner has bought $1\,Km$ of metal fence. He wants to use such fence and a wall of his huge home to enclose the largest possible area for his flock. What is the optimal shape he may choose for the fence, and how large is the largest area he can dedicate to his flock?

Exercise 234 (Lhuilier's inequality)**.** Prove that if P is a convex polygon with n sides, having external angles $\alpha_1, \ldots, \alpha_n$, we have:

$$L(\partial P)^2 \geq 4\,A(P)\sum_{k=1}^{n} \tan\left(\frac{\alpha_i}{2}\right)$$

and equality holds if and only if P is circumscribed to a circle.

Exercise 235. Prove that if P is a convex polygon with n sides containing a unit circle, we have:

$$A(P) \geq n\tan\left(\frac{\pi}{n}\right)$$

and equality holds if and only if P is a regular polygon.

Exercise 236. Prove that if P is a convex polygon with n sides, we have

$$L(\partial P)^2 \geq 4n\tan\left(\frac{\pi}{n}\right) A(P)$$

and equality holds if and only if P is a regular polygon.

Exercise 237 (Bonnesen's inequality)**.** Let γ be a regular, simple and closed curve enclosing a convex region P. Let R, r denote the radii of the circumscribed and inscribed circle. Prove that the following strengthening of the isoperimetric inequality holds:

$$L(\gamma)^2 - 4\pi\,A(P) \geq \pi^2(R-r)^2.$$

The isoperimetric inequality through Lagrange multipliers.
Among the convex polygons enclosing the origin, having sides of fixed lengths, the cyclic polygon encloses the maximum area.

Proof. Setting $\theta_i = \widehat{P_{i+1}P_iO}$, $L_i = d(P_i, P_{i+1})$ and $\phi_i = \widehat{OP_{i+1}P_i}$, we want to find the maximum of

$$A = \frac{1}{2}\sum \frac{\sin\theta_i \sin\phi_i}{\sin(\theta_i+\phi_i)}\,L_i^2$$

subject to the constraint:

$$\sum(\theta_i+\phi_i) = (n-2)\pi.$$

We are in position of applying Lagrange multipliers, from which we have, for any i:

$$\left(\frac{\sin\phi_i}{\sin(\theta_i+\phi_1)}\right)^2 = \frac{\partial}{\partial\theta_i}\left(\frac{\sin\theta_i\sin\phi_i}{\sin(\theta_i+\phi_i)}\right) = \frac{\lambda}{L_i^2} = \frac{\partial}{\partial\phi_i}\left(\frac{\sin\theta_i\sin\phi_i}{\sin(\theta_i+\phi_i)}\right) = \left(\frac{\sin\theta_i}{\sin(\theta_i+\phi_1)}\right)^2.$$

It follows that all the triangles OP_iP_{i+1} are isosceles triangles with vertex at O, hence the cyclic polygon $P_1\dots P_n$ certainly has the greatest area. □

Theorem 238 (Jung, isodiametric inequality)**.**
Every compact set $K\subset\mathbb{R}^n$ of diameter d is contained in some closed ball of radius

$$R\le d\sqrt{\frac{n}{2(n+1)}}$$

with equality attained only by regular n-simplex of side d.

Proof. Given a set of points of diameter d in $\mathbb{R}^n$ it is trivial to see that it can be covered by a ball of radius d.
But the above Theorem by Jung improves the result by a factor of about $\frac{1}{\sqrt{2}}$, and is the best possible.
We first prove this Theorem for sets of points S with $|S|\le n+1$ and then extend it to an arbitrary point set. If $|S|\le n+1$ then the smallest ball enclosing S exists. We assume that its center is the origin and denote its radius by R. Denote by $S'\subseteq S$ the subset of points such that $\|p\|=R$ for $p\in S'$. It is easy to see that S' is in fact non empty.

Observation: The origin must lie in the convex hull of S'. Assuming the contrary, there is a separating hyperplane H such that S' lies on one side and the origin lies on the other side of H (strictly). By assumption, every point in $S\setminus S'$ has a distance strictly less than R from the origin. Move the center of the ball slightly from the origin, in a direction perpendicular to the hyperplane H towards H such that the distances from the origin to every point in $S\setminus S'$ remains less than R. However, now the distance to every point of S' is decreased and so we will have a ball of radius strictly less than R enclosing S which is a contradiction to the minimality of R.

Let $S'=\{p_1,p_2,\dots,p_m\}$ where $m\le n\le d+1$ and because the origin is in the convex hull of S' so we have non-negative λ_i such that,

$$\sum\lambda_i p_i=0,\qquad \sum\lambda_i=1.$$

Fix a k, $1\le k\le m$. Then we have:

$$\begin{aligned} 1-\lambda_k &= \textstyle\sum_{i\ne k}\lambda_i\\ &\ge \tfrac{1}{d^2}\textstyle\sum_{i=1}^m\lambda_i\|p_i-p_k\|^2\\ &= \tfrac{1}{d^2}\left(\textstyle\sum_{i=1}^m\lambda_i(2R^2-2\langle p_i,p_k\rangle)\right)\\ &= \tfrac{1}{d^2}\left(2R^2-2\left\langle\textstyle\sum_{i=1}^m\lambda_i p_i,p_k\right\rangle\right)\\ &= \tfrac{2R^2}{d^2}\end{aligned}$$

Adding up the above inequalities for all values of k, we get

$$m-1\ge\frac{2mR^2}{d^2}$$

Thus we get $\frac{R^2}{d^2}\le\frac{m-1}{2m}\le\frac{n}{2n+2}$ since $m\le n+1$ and the function $\frac{x-1}{2x}$ is monotonic. So we have immediately $R\le d\sqrt{\frac{n}{2n+2}}$. The remainder of the proof uses the beautiful theorem of Helly. So assume S is any set of points of diameter d. With each point as center draw a ball of radius $R=d\sqrt{\frac{n}{2n+2}}$. Clearly any $n+1$ of these balls intersect. This is true because the center of the smallest ball enclosing $n+1$ of the points is at most R away from each of those points. So we have a collection of compact convex sets, any $n+1$ of which have a nonempty intersection. By Helly's theorem all of them have an intersection. Any point of this intersection can be chosen to be the center of a ball of radius R that will enclose all of S.

Theorem 239 (Helly)**.** Let $\{X_1, \ldots, X_d\}$ be a finite collection of convex subsets of $\mathbb{R}^d$, with $n > d$. If the intersection of every $d+1$ of these sets is non-empty, then the whole collection has a nonempty intersection; that is,

$$\bigcap_{j=1}^{n} X_j \neq \varnothing.$$

For infinite collections one has to assume compactness: Let $\{X_\alpha\}$ be a collection of compact convex subsets of $\mathbb{R}^d$, such that every subcollection of cardinality at most $d+1$ has a non-empty intersection, then the whole collection has a non-empty intersection.

Proof. The proof is by mathematical induction:

Base case: Let $n = d+2$. By our assumptions, for every $j = 1, \ldots, n$ there is a point x_j that is in the common intersection of all X_i with the possible exception of X_j. Now we apply Radon's Theorem to the set $A = \{x_1, \ldots, x_n\}$, which furnishes us with disjoint subsets A_1, A_2 of A such that the convex hull of A_1 intersects the convex hull of A_2. Suppose that p is a point in the intersection of these two convex hulls. We claim that

$$p \in \bigcap_{j=1}^{n} X_j.$$

Indeed, consider any $j \in \{1, \ldots, n\}$. We shall prove that $p \in X_j$. Note that the only element of A that may not be in X_j is x_j. If $x_j \in A_1$, then $x_j \notin A_2$, and therefore $X_j \supset A_2$. Since X_j is convex, it then also contains the convex hull of A_2 and therefore also $p \in X_j$. Likewise, if $x_j \notin A_1$, then $X_j \supset A_1$, and by the same reasoning $p \in X_j$. Since p is in every X_j, it must also be in the intersection.
Above, we have assumed that the points $x_1, \ldots, x_n$ are all distinct. If this is not the case, say $x_i = x_k$ for some $i \neq k$, then x_i is in every one of the sets X_j, and again we conclude that the intersection is nonempty. This completes the proof in the case $n = d+2$.

Inductive Step: Suppose $n > d+2$ and that the statement is true for $n-1$. The argument above shows that any subcollection of $d+2$ sets will have nonempty intersection. We may then consider the collection where we replace the two sets X_{n-1} and X_n with the single set $X_{n-1} \cap X_n$. In this new collection, every subcollection of $d+1$ sets will have nonempty intersection. The inductive hypothesis therefore applies, and shows that this new collection has nonempty intersection. This implies the same for the original collection, and completes the proof.

Theorem 240 (Radon, 1921)**.** Any set of $d+2$ points in $\mathbb{R}^d$ can be partitioned into two disjoint sets whose convex hulls intersect. A point in the intersection of these convex hulls is called a *Radon point* of the set.

Proof. Consider any set $\{x_1, x_2, \ldots, x_{d+2}\} \subset \mathbb{R}^d$ of $d+2$ points in a d-dimensional space. Then there exists a set of multipliers $a_1, a_2, \ldots, a_{d+2}$, not all of which are zero, solving the system of linear equations

$$\sum_{i=1}^{d+2} a_i x_i = 0, \qquad \sum_{i=1}^{d+2} a_i = 0,$$

because there are $d+2$ unknowns (the multipliers) but only $d+1$ equations that they must satisfy (one for each coordinate of the points, together with a final equation requiring the sum of the multipliers to be zero). Fix some particular nonzero solution $a_1, a_2, \ldots, a_{d+2}$. Let I be the set of points with positive multipliers, and let J be the set of points with multipliers that are negative or zero. Then I and J form the required partition of the points into two subsets with intersecting convex hulls. The convex hulls of I and J must intersect, because they both contain the point

$$p = \sum_{i \in I} \frac{a_i}{A} x_i = \sum_{j \in J} \frac{-a_j}{A} x_j,$$

where

$$A = \sum_{i \in I} a_i = -\sum_{j \in J} a_j.$$

The left hand side of the formula for p expresses this point as a convex combination of the points in I, and the right hand side expresses it as a convex combination of the points in J. Therefore, p belongs to both convex hulls, completing the proof. □

Van Der Corput's trick for lower bounds. The purpose of this paragraph is to prove that the partial sums of the sequence $\{\sin(n^2)\}_{n\geq 1}$ are not bounded. We have:

$$\begin{aligned} 2\left(\sum_{k=1}^{n}\sin(k^2)\right)^2 &= \sum_{j,k=1}^{n}\cos(j^2-k^2) - \sum_{j,k=1}^{n}\cos(j^2+k^2) \\ &= n + 2\sum_{m=1}^{n^2-1} d_1(m)\cos(m) - 2\sum_{m=2}^{2n^2} d_2(m)\cos m \end{aligned}$$

where $d_1(m)$ accounts for the number of ways to write m as $j^2 - k^2$ with $1 \leq k < j \leq n$ and $d_2(m)$ accounts for the number of ways to write m as $j^2 + k^2$ with $1 \leq j, k \leq n$. Since both these arithmetic functions do not deviate much from their average order (by Dirichlet's hyperbola method $d_1(m)$ behaves on average like $\log m$ and $d_2(m)$ behaves on average like $\frac{\pi}{4}$), It is not terribly difficult to prove that for infinitely many ns

$$\left|\sum_{k=1}^{n}\sin(k^2)\right| \geq C\sqrt{n}$$

holds for some absolute constant $C \approx \frac{1}{\sqrt{2}}$ through summation by parts and the Cauchy-Schwarz inequality.
A detailed exposition on Dirichlet's hyperbola method can be found on Terence Tao's blog.
We recall that Van Der Corput's trick is usually employed to produce upper bounds: for instance

$$\left|\sum_{k=1}^{n}\sin(k^2)\right| \leq D\sqrt{n}\log n,$$

for some absolute constant $D > 0$, holds for any n large enough.
In particular $\sum_{n\geq 1}\frac{\sin(n^2)}{n^\alpha}$ is convergent for any $\alpha > \frac{1}{2}$.

8 Remarkable results in Linear Algebra

This brief section is devoted to some important results in Linear Algebra: we will outline proofs *by density* for the Hamilton-Cayley Theorem and the identity $\text{Tr}(AB) = \text{Tr}(BA)$, then we will outline a recent elementary proof (due to Suk-Geun Hwang) of Cauchy's interlace theorem, admitting Sylvester's criterion as a straightforward corollary.

Theorem 241 (Hamilton-Cayley)**.** If $p \in \mathbb{C}[x]$ is the characteristic polynomial of a $n \times n$ matrix A with complex entries,

$$p(A) = 0.$$

Proof. If A is a diagonalizable matrix the claim is trivial: if $A = J^{-1}DJ$ is the Jordan normal form of A, the diagonal entries of D are the eigenvalues $\lambda_1, \ldots, \lambda_n$ of A. We have $p(\lambda_j) = 0$ by the very definition of characteristic polynomial, and since $p(M^k) = p(M)^k$,

$$p(A) = p(J^{-1}DJ) = J^{-1}p(D)J = 0.$$

On the other hand diagonalizable matrices form a *dense* subspace in the space of $n \times n$ matrices with complex entries. Assuming that A is not a diagonalizable matrix it follows that for any $\varepsilon > 0$ there exist some diagonalizable matrix A_ε such that $\|A - A_\varepsilon\|_2 \leq \varepsilon$ (actually the choice of the Euclidean norm is immaterial, any induced norm does the job equally fine). The eigenvalues of A_ε converge to the eigenvalues of A and p is a continuous function, hence

$$p(A) = \lim_{\varepsilon \to 0} p(A_\varepsilon) = 0.$$

□

Exercise 242. We have that $\{x_n\}_{n\geq 1}$, $\{y_n\}_{n\geq 1}$, $\{z_n\}_{n\geq 1}$ are three sequences of real numbers such that any term among x_n, y_n, z_n can be written as a linear combination of $x_{n-1}, y_{n-1}, z_{n-1}$ with constant coefficients, for instance:

$$x_n = 4y_{n-1} + z_{n-1}, \qquad y_n = 3y_{n-1} - z_{n-1}, \qquad z_n = x_{n-1} + y_{n-1} - 2z_{n-1}.$$

Prove that there exists an order-3 linear recurrence relation simultaneously fulfilled by each one of the sequences $\{x_n\}_{n\geq 1}$, $\{y_n\}_{n\geq 1}$, $\{z_n\}_{n\geq 1}$.

Exercise 243. The sequence $\{a_n\}_{n\geq 1}$ is defined through

$$a_1 = 1, \qquad a_{n+1} = \frac{2a_n}{7 + a_n}.$$

Prove that $\lim_{n\to+\infty} a_n = 0$, then find $\lim_{n\to+\infty} \frac{\log a_n}{\log n}$.

Theorem 244 ("The trace is Abelian")**.** For any couple (A, B) of $n \times n$ matrices with complex entries, the following identity holds:

$$\text{Tr}(AB) = \text{Tr}(BA).$$

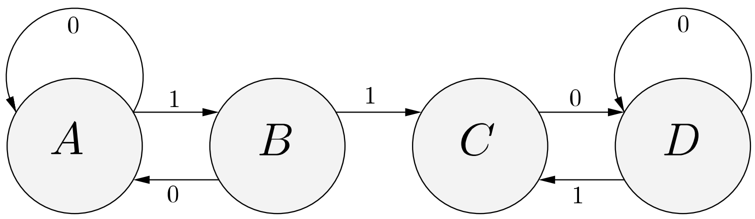
Proof. Assuming A is an invertible matrix, AB and BA share the same characteristic polynomial, since they are conjugated matrices due to $BA = A^{-1}(AB)A$. In particular they have the same trace. Equivalently, they share the same eigenvalues (counted according to their algebraic multiplicity) hence they share the sum of such eigenvalues. On the other hand, if A is a singular matrix then $A_\varepsilon \stackrel{\text{def}}{=} A + \varepsilon I$ is an invertible matrix for any $\varepsilon \neq 0$ small enough. It follows that $\text{Tr}(A_\varepsilon B) = \text{Tr}(BA_\varepsilon)$, and since Tr is a continuous operator, by considering the limits of both sides as $\varepsilon \to 0$ we get $\text{Tr}(AB) = \text{Tr}(BA)$ just as well. □

Corollary 245. If A is a $n \times n$ matrix with real entries and $B = A^T$, A and B have the same rank. Prove this statement by showing that $\text{Tr}(A^k) = \text{Tr}(B^k)$ for any $k \in \mathbb{N}$. *Hint*: notice that $\text{Tr}(M) = \text{Tr}(M^T)$ and that the power sums of the eigenvalues fix the coefficients of the characteristic polynomial of a matrix.

Exercise 246. Prove that any matrix T with real entries such that $\text{Tr}(T) = 0$ can be written in the form $T = AB - BA$ for a suitable choice of the matrices A, B.

Unexpected applications of the Hamilton-Cayley theorem: *if you are able to draw it, you also know its asymptotic behaviour.*

Exercise 247. Let T_n be the number of strings over the alphabet $\Sigma = \{0, 1\}$ with length n and exactly one occurrence of the "11" substring. Find an explicit formula for T_n.



Proof. It is pretty simple to construct a finite automaton accepting the strings of our (regular) language. The automaton depicted above has the following transition matrix:

$$M = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Since the starting state is A and the accepting states are C, D, we simply have:

$$T_n = (1\ 0\ 0\ 0)\ M^n\ (0\ 0\ 1\ 1)^T$$

and by the Hamilton-Cayley Theorem the sequence $\{T_n\}_{n \geq 0}$ and the matrix M **share the same characteristic polynomial**. Since the eigenvalues of M are $\frac{1 \pm \sqrt{5}}{2}$ and they both have algebraic multiplicity 2 and geometric multiplicity 1, by the Jordan decomposition of M we have:

$$T_n = (a + bn) \left(\tfrac{1+\sqrt{5}}{2}\right)^n + (c + dn) \left(\tfrac{1-\sqrt{5}}{2}\right)^n$$

and the constants a, b, c, d can be found by interpolation, through $T_0 = T_1 = 0$, $T_2 = 1$, $T_3 = 2$.
In particular we have $T_n = \frac{1}{5}(nL_n - F_n)$. To prove the same in a purely combinatorial fashion is a bit more involved, but certainly not impossible. By *stars and bars*, the number of strings with length $n - 1$ and exactly k non-adjacent 1s is given by $\binom{n-k}{k}$. In particular:

$$T_n = \sum_{k \geq 1} \binom{n-k}{k} k = \sum_{k=1}^{n-1} A_k A_{n-k}$$

where A_k is the number of strings with length k, having no adjacent 1s and starting with a 1. By the convolution machinery the generating function of $\{T_n\}_{n\geq 0}$ is simply given by the square of the generating function of $\{A_n\}_{n\geq 0} = \{F_{n+1}\}_{n\geq 0}$. The latter is a meromorphic function with two simple poles and we may recover the previous closed form by partial fraction decomposition. □

Exercise 248. (☠) Let s be a non-empty string over the alphabet $\Sigma = \{0, 1\}$. Let A_s be the finite automaton accepting the strings over Σ that do not contain s as a substring. Let us define $\text{Spec}(s)$ as the spectrum of the transition matrix of A_s. Investigate about the relations between $\text{Spec}(s.t)$ and $\text{Spec}(s), \text{Spec}(t)$, where . denotes the concatenation of strings.

Exercise 249. (☠) According to the notation introduced in the previous exercise, prove or disprove the existence of a string s such that $2\cos\frac{2\pi}{7} \in \text{Spec}(s)$.

Definition 250. We say that a weakly increasing sequence of real numbers $a_1 \leq a_2 \leq \ldots \leq a_n$ *interlaces* another weakly increasing sequence of real numbers $b_0 \leq b_1 \leq \ldots \leq b_n$ if

$$b_0 \leq a_1 \leq b_1 \leq a_2 \leq \ldots \leq a_n \leq b_n$$

holds.

Theorem 251 (Cauchy's interlace Theorem)**.** The eigenvalues of a Hermitian matrix A of order n are interlaced with those of any principal submatrix of order $n-1$.

Proof. Hermitian matrices have real eigenvalues. Let A be a Hermitian matrix of order n and let B be a principal submatrix of A of order $n-1$. If $\lambda_n \leq \lambda_{n-1} \leq \ldots \leq \lambda_1$ lists the eigenvalues of A and $\mu_n \leq \mu_{n-1} \leq \ldots \leq \mu_2$ the eigenvalues of B, we shall prove that

$$\lambda_n \leq \mu_n \leq \lambda_{n-1} \leq \mu_{n-1} \leq \ldots \leq \lambda_2 \leq \mu_2 \leq \lambda_1.$$

Proofs of this theorem have been based on Sylvester's law of inertia and the Courant-Fischer min-max theorem. Here we will give a simple, elementary proof of the theorem by using the intermediate value theorem. Simultaneously permuting rows and columns, if necessary, we may assume that the submatrix B occupies rows $2, 3, \ldots n$ and columns $2, 3, \ldots, n$, so that A has the form

$$A = \begin{pmatrix} a & y^* \\ y & B \end{pmatrix}$$

where $*$ stands for the conjugate transpose of a matrix. Let $D = \text{diag}(\mu_2, \mu_3, \ldots, \mu_n)$. Then, since B is also Hermitian, by the spectral Theorem there exists a unitary matrix U of order $n-1$ such that $U^*BU = D$. Let $U^*y = z = (z_2, z_3, \ldots, z_n)^T$. We first prove the theorem for the special case where $\mu_n < \mu_{n-1} < \ldots < \mu_3 < \mu_2$ and $z_i = 0$ for $i = 2, 3, \ldots, n$. Let

$$V = \begin{pmatrix} 1 & 0^T \\ 0 & U \end{pmatrix}$$

in which 0 denotes the zero vector. Then V is a unitary matrix and

$$V^*AV = \begin{pmatrix} a & z^* \\ z & D \end{pmatrix}.$$

Let $f(x) = \det(xI - A) = \det(xI - V^*AV)$, where I denotes the identity matrix. Expanding $\det(xI - V^*AV)$ along the first row, we get

$$f(x) = (x-a)(x-\mu_2)(x-\mu_3)\cdots(x-\mu_n) - \sum_{i=2}^{n} f_i(x)$$

where $f_i(x) = |z_i|^2(x-\mu_2)\cdots\widehat{(x-\mu_i)}\cdots(x-\mu_n)$ for $i = 2,3,\ldots,n$, with the hat-sign denoting a missing term. We may notice that $f_i(\mu_j) = 0$ when $j \neq i$ and $f_i(\mu_i)$ is strictly positive or strictly negative according to i being, respectively, even or odd. It follows that $f(\mu_i)$ is positive if i is odd and negative if i is even. Since $f(x)$ is a polynomial of degree n with positive leading coefficient, the intermediate value Theorem ensures the existence of n roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of the equation $f(x) = 0$ such that $\lambda_n < \mu_n < \lambda_{n-1} < \mu_{n-1} < \ldots < \lambda_2 < \mu_2 < \lambda_1$.
For the proof of the general case, let $\varepsilon_1, \varepsilon_2, \ldots$ be a sequence of positive real numbers such that ε_k is decreasing towards zero, $z_i + \varepsilon_k \neq 0$ for $i = 2,3,\ldots,n$ and $k = 1,2,\ldots$ and the diagonal entries of $D + \varepsilon_k \mathrm{diag}(2,3,\ldots,n)$ are distinct for fixed k. For $k = 1,2,\ldots$ let

$$C_k = \begin{pmatrix} a & z(\varepsilon_k)^* \\ z(\varepsilon_k) & D(\varepsilon_k) \end{pmatrix}$$

where $z(\varepsilon_k) = z + \varepsilon_k(1,1,\ldots,1)^T$ and $D(\varepsilon_k) = D + \varepsilon_k \mathrm{diag}(2,3,\ldots,n)$, and let $A_k = VC_kV^*$. Then A_k is Hermitian and A_k converges towards A. Let $\lambda_n^{(k)} \leq \lambda_{n-1}^{(k)} \leq \ldots \leq \lambda_2^{(k)} \leq \lambda_1^{(k)}$ list the eigenvalues of A_k. Then

$$\lambda_n^{(k)} < \mu_n + n\varepsilon_k < \lambda_{n-1}^{(k)} < \mu_{n-1} + (n-1)\varepsilon_k < \ldots < \lambda_2^{(k)} < \mu_2 + 2\varepsilon_k < \lambda_1^{(k)}.$$

Since $\lambda_n^{(k)}, \lambda_{n-1}^{(k)}, \ldots, \lambda_1^{(k)}$ are n distinct roots of $\det(xI - A_k) = 0$ for each k and since the graph of $y = \det(xI - A_k)$ is sufficiently close to that of $y = \det(xI - A)$, it follows that the proof is complete by invoking the implicit function Theorem: $\left(\lambda_n^{(k)}, \lambda_{n-1}^{(k)}, \ldots, \lambda_1^{(k)}\right) \to (\lambda_n, \lambda_{n-1}, \ldots, \lambda_1)$. □

Theorem 252 (Sylvester's criterion)**.** A Hermitian matrix M is positive-definite if and only if the determinants of the leading principal minors are positive.

Proof. If some minor has a negative or zero determinant the original matrix M cannot be positive definite by Cauchy's interlace theorem. This proves that the positivity of the mentioned determinants is a necessary condition. The converse implication can be easily shown by induction on the dimension of M, or by exploiting the Cholesky decomposition $A = B^TB$ with B being a non-singular matrix. □

Theorem 253 (Banach-Steinhaus uniform boundedness theorem)**.** Let $\mathscr{F}$ be a family of bounded linear operators from a Banach space X to a normed linear space Y. If $\mathscr{F}$ is pointwise bounded (i.e., $\sup_{T\in\mathscr{F}} \|Tx\| < \infty$ for all $x \in X$), then $\mathscr{F}$ is norm-bounded (i.e., $\sup_{T\in\mathscr{F}} \|T\| < \infty$).

Proof. The following proof is due to Alan D.Sokal. Let T be a bounded linear operator from a normed linear space X to a normed linear space Y. Then for any $x \in X$ and $r > 0$, we have

$$\sup_{x' \in B(x,r)} \|Tx'\| \geq \|T\|r,$$

where, as usual, $B(x,r) = \{x \in X : \|x' - x\| < r\}$. Indeed, for any $\xi \in X$ we have

$$\max\{\|T(x+\xi)\|, \|T(x-\xi)\|\} \geq \frac{\|T(x+\xi)\| + \|T(x-\xi)\|}{2} \geq \|T\xi\|$$

where the second $\geq$ uses the triangle inequality in the form $\|\alpha-\beta\| \leq \|\alpha\|+\|\beta\|$, and we may consider the supremum over $\xi \in B(0,r)$. If we assume that $\sup_{T\in\mathscr{F}}\|T\| = \infty$, we may construct a sequence $\{T_n\}_{n\geq 1}$ such that $\|T_n\| \geq 4^n$. Setting $x_0 = 0$, we may use the previous lemma on $\sup_{x'\in B(x,r)}\|Tx'\|$ to choose inductively $x_n \in X$ such that $\|x_n - x_{n-1}\| \leq 3^{-n}$ and

$$\|T_n x_n\| \geq \frac{2}{3}3^{-n}\|T_n\|.$$

$\{x_n\}_{n\geq 1}$ is a Cauchy sequence, hence it is convergent to some $x \in X$: it is easy to check that $\|x - x_n\| \leq \frac{1}{2}3^{-n}$, such that $\|T_n x\| \geq \frac{1}{6}3^{-n}\|T_n\| \geq \frac{1}{6}\left(\frac{4}{3}\right)^n \to \infty$, contradicting the pointwise-boundedness of $\mathscr{F}$. □

Corollary 254. There is a 2π-periodic and continuous function f whose Fourier series

$$\sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{inx}, \qquad \widehat{f}(n) = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-nix}\,dx$$

does not converge at $x = 0$.

Proof. In order to invoke the Banach-Steinhaus theorem, we consider the functionals given by the partial sums of the Fourier series of f, evaluated at $x = 0$:

$$\lambda_N(f) = \sum_{|n|\leq N} \widehat{f}(n).$$

There is a simple upper bound, namely

$$|\lambda_N(f)| \leq \int_0^1 \Big|\sum_{|n|\leq N} e^{-2\pi inx}\Big| \cdot |f(x)|\,dx \leq \|f\|_\infty \cdot \Big\|\sum_{|n|\leq N} e^{-2\pi inx}\Big\|_1.$$

If we take $g(x)$ as the sign of the Dirichlet kernel

$$\sum_{|n|\leq N} e^{-2\pi inx} = \frac{\sin\left(2\pi x\left(N+\frac{1}{2}\right)\right)}{\sin(\pi x)}$$

and $\{g_j(x)\}_{j\geq 1}$ as a sequence of periodic continuous functions, such that $|g_j(x)| \leq 1$ and $g_j(x) \to g(x)$, by dominated convergence

$$\lim_{j\to+\infty} \lambda_N(g_j) = \int_0^1 g(x)\sum_{|n|\leq N} e^{-2\pi inx}\,dx = \int_0^1 \Big|\sum_{|n|\leq N} e^{-2\pi inx}\Big|\,dx,$$

hence the previous bound for the norm of λ_N holds as an equality. Since the mean value of $|\sin x|$ is $\frac{2}{\pi}$, integration by parts leads to

$$\Big\|\sum_{|n|\leq N} e^{-2\pi inx}\Big\|_1 \sim \frac{2}{\pi}\log N,$$

hence there is no uniform bound for the L^1-norm of the Dirichlet kernel. By Banach-Steinhaus, there is some f in the unit ball of $C^0(\mathbb{T})$ such that

$$\sup_N |\lambda_N(f)| = +\infty.$$

In fact, the collection of such fs is *dense* in the unit ball, and it is an intersection of a *countable* collection of dense open sets (a G_δ). □

The same phenomenon does not occur if the Dirichlet kernel is replaced by the Fejér kernel: if $f \in C^0(\mathbb{T})$, the sequence of trigonometric polynomials defined by

$$p_N(x) = \sum_{|n|\leq N}\left(1 - \frac{|n|}{N}\right)\widehat{f}(n)e^{-2\pi inx}$$

converges uniformly to $f(x)$.

Corollary 255. If $A = \{a_n\}_{n\geq 1}$ is a sequence of real numbers such that $|\langle\Lambda, A\rangle| = \left|\sum_{n\geq 1}\lambda_n a_n\right|$ is finite for any $\Lambda = \{\lambda_n\}_{n\geq 1} \in \ell^2$, then $B \in \ell^2$.

Proof. We work by contradiction: we assume that $A = \sum_{n\geq 1} a_n^2$ is unbounded and we show that for some sequence $\Lambda \in \ell^2$ the series $\sum_{n\geq 1}\lambda_n a_n$ is positively divergent. By picking the multipliers λ_n with the same sign as the corresponding a_n, we may as well assume that $a_n \geq 0$. For any real number $c > 0$, the subsequence $\{a_{\sigma(n)}\}_{n\geq 1}$ made by the terms $\geq c$ has a finite number of terms: otherwise the inner product with the sequence $\{\lambda_{\sigma(n)} = \frac{1}{n}\}_{n\geq 1} \in \ell^2$ would be unbounded. In particular 0 is the only accumulation point of $\{a_n\}_{n\geq 1}$ and $\lim_{n\to+\infty} a_n = 0$. Let us define $a_0 = 1$,

$$S_n = \sum_{k=0}^{n} a_n^2$$

and let us consider $\lambda_n = \sqrt{\frac{1}{S_{n-1}} - \frac{1}{S_n}}$. In this case $\Lambda \in \ell^2$ by construction, since $\sum_{n\geq 1}\lambda_n^2$ is a telescopic series and $S_n \to +\infty$ by the original assumption $B \notin \ell^2$. If we manage to prove that

$$\sum_{n\geq 1} a_n\sqrt{\frac{1}{S_{n-1}} - \frac{1}{S_n}} = \sum_{n\geq 1}\frac{a_n^2}{\sqrt{S_{n-1}S_n}} \geq \sum_{n\geq 1}\frac{a_n^2}{S_n}$$

is divergent we are done. Let us define, by induction, $\tau(0)$ as the smallest n such that $S_n \geq 2$, $\tau(m)$ as the smallest n such that $S_{\tau(m)}$ is $\geq 2\cdot S_{\tau(m-1)}$. We have

$$\sum_{\tau(m)<n\leq\tau(m+1)}\frac{a_n^2}{S_n} \geq \frac{1}{S_{\tau(m+1)}}\left(S_{\tau(m+1)} - S_{\tau(m)}\right) \geq \frac{1}{2}$$

hence by summing both sides on $m \geq 0$ we have that $\sum_{n\geq 1} a_n\sqrt{\frac{1}{S_{n-1}} - \frac{1}{S_n}}$ is divergent. □

The last proof can be easily adapted to the continuous case through very few adjustements:

Corollary 256. If $f : \mathbb{R}^+ \to \mathbb{R}$ is a function such that $f\cdot g \in L^1(\mathbb{R}^+)$ for any $g \in L^2(\mathbb{R}^+)$, then $f \in L^2(\mathbb{R}^+)$.

Proof. We work by contradiction: we assume that $\int_0^x f(t)^2\,dt$ is unbounded and we show that for some function $g \in L^2(\mathbb{R}^+)$ the integral $\int_0^x f(t)g(t)\,dt$ is positively divergent as $x \to +\infty$. By picking the multipliers $g(x)$ such that f/g has almost everywhere the same sign, we may as well assume that $f(x) \geq 0$. For any real number $c > 0$, the set of $x \in \mathbb{R}^+$ such that $f(x) \geq c$ has finite measure, otherwise the integral of $\frac{f(x)}{\sqrt{x^2+1}}$ over such set would be unbounded. By replacing f with the convolution between f and a non-negative, compact supported and smooth kernel we may also assume that $f(x)$ is continuous on $\mathbb{R}^+$. Let us define

$$F(x) = \int_0^x f(t)^2\,dt$$

and let us consider $g(x) = \sqrt{\frac{d}{dx}\left(-\frac{1}{1+F(x)}\right)}$. In this case $g \in L^2(\mathbb{R}^+)$ by construction, since the integral of $g(x)^2$ can be computed through the fundamental Theorem of Calculus and $F(x)$ is increasing to $+\infty$. If we manage to prove that

$$\int_0^{+\infty} f(x)\sqrt{\frac{d}{dx}\left(-\frac{1}{1+F(x)}\right)}\,dx = \int_0^{+\infty}\frac{f(x)^2}{1+F(x)}\,dx$$

is divergent we are done. Let us define, by induction, $\tau(0)$ as the infimum of the set $\{x : F(x) \geq 2\}$, $\tau(m)$ as the infimum of the set $\{x : F(x) \geq 2F(\tau(m-1))\}$. We have

$$\int_{\tau(m)}^{\tau(m+1)}\frac{f(x)^2}{1+F(x)}\,dx \geq \frac{1}{1+F(\tau(m+1))}\int_{\tau(m)}^{\tau(m+1)} f(x)^2\,dx = \frac{F(\tau(m+1)) - F(\tau(m))}{1+F(\tau(m+1))} \geq \frac{1}{3}$$

hence by summing both sides on $m \geq 0$ we have that $\int_0^{+\infty} f(x)\sqrt{\frac{d}{dx}\left(-\frac{1}{1+F(x)}\right)}\,dx$ is divergent. $\square$

The powerful lemma 255 is a Corollary of the Banach-Steinhaus theorem: given the sequence $A = \{a_n\}_{n\geq 1}$ and any $\Lambda \in \ell^2$, the operators

$$T_N(\Lambda) = \sum_{n=1}^{N} a_n \lambda_n$$

are linear, continuous and pointwise bounded. The uniform boundedness criterion ensures they are norm-bounded, i.e. $A \in \ell^2$. The lemma 255 immediately leads to the fact that ℓ^2 and $L^2(\mathbb{R}^+)$ are complete spaces. Since Fourier series give an isometry between $L^2(0, 2\pi)$ and ℓ^2, such lemma also provides a proof of the completeness of $L^2(I)$ for any bounded interval I.

9 The Fundamental Theorem of Algebra

The purpose of this section is to shortly introduce some elements of Complex Analysis and use them to produce a proof of the Fundamental Theorem of Algebra, stating that $\mathbb{C}$ is an algebraically closed field, or, in layman's terms:

Theorem 257. Any non-constant $p \in \mathbb{C}[z]$ vanishes at some $z \in \mathbb{C}$.

The first serious attempt to such problem is due to Gauss: it was mainly geometric, but it had a topological gap, filled by Alexander Ostrowski in 1920. A rigorous proof was first published by Argand in 1806 (and revisited in 1813). We will actually show many approaches and focus on geometric and analytic insights and their consequences. All proofs below involve some analysis, or at least the topological concept of continuity of real or complex functions. Some also use differentiable or even analytic functions. This fact has led to the remark that the Fundamental Theorem of Algebra is neither fundamental, nor a theorem of algebra.

Definition 258. Suppose we are given a closed, oriented curve in the xy plane, not going through the origin. We can imagine the curve as the path of motion of some object, with the orientation indicating the direction in which the object moves. Then the **winding number** of the curve is equal to the total number of counterclockwise turns that the object makes around the origin. When counting the total number of turns, counterclockwise motion counts as positive, while clockwise motion counts as negative. For example, if the object first circles the origin four times counterclockwise, and then circles the origin once clockwise, then the total winding number of the curve is three.

But given a closed curve $\gamma : [0, 2\pi] \to \mathbb{R}^2 \setminus \{(0,0)\}$ represented by $\gamma(t) = (x(t), y(t))$ (which we may temporarily assume to be smooth, too), how can we find its winding number around the origin? We may notice that in the open first quadrant $\arctan\frac{y(t)}{x(t)}$ gives an "angular displacement" with respect to the origin: in order to compute the winding number of γ we just need to find a *continuous determination* of such angular displacement. For brevity we will not delve into the theory of differential forms, we just mention that such *continuous determination* can be achieved through a step of differentiation and a step of integration, namely:

$$\frac{d}{dt}\arctan\frac{y(t)}{x(t)} = \frac{\frac{d}{dt}\frac{y(t)}{x(t)}}{1+\frac{y(t)^2}{x(t)^2}} = \frac{x(t)y'(t) - y(t)x'(t)}{x(t)^2 + y(t)^2}$$

leading to

$$\frac{1}{2\pi}\int_0^{2\pi} \frac{x(t)y'(t) - y(t)x'(t)}{x(t)^2 + y(t)^2}\,dt$$

as an expression for the winding number of γ around the origin. What if our curve is given by some $f : S^1 \to \mathbb{C}^*$, with f being a holomorphic function? In such a case the previous winding number takes the following form:

$$\frac{1}{2\pi i}\oint_{|z|=1} \frac{f'(z)}{f(z)}\,dz.$$

Now it comes an interesting remark, clarifying the interplay between winding numbers and zeroes of holomorphic functions: in the previous line, the integrand function is formally $\frac{d}{dz}\log f(z)$, hence if $f = hg$ with h and g being holomorphic functions, the winding number of f is just the sum between the winding number of h and the winding number of g. It is very simple to check that for any $m \in \mathbb{N}$ the winding number of z^m is exactly m and for every $w \in \mathbb{C}\setminus S^1$, by setting $f(z) = z - w$ we get that the winding number of f is 1 or 0 according to $|w| < 1$ or $|w| > 1$: in the first case $f(S^1)$ encloses the origin, in the latter it does not. Conversely, if $f(z)$ is a holomorphic and non-vanishing function over D, then $\frac{f'(z)}{f(z)}$ is a holomorphic function over D and the integral $\oint_{\partial D} \frac{f'(z)}{f(z)}\,dz$ equals zero by Stokes' theorem.

Lemma 259. If f is a holomorphic function on $D = \{z \in \mathbb{C} : |z| \leq 1\}$, non-vanishing over ∂D,

$$N_f = \frac{1}{2\pi i}\oint_{\partial D} \frac{f'(z)}{f(z)}\,dz$$

equals the number of zeroes of f in D, counted according to their multiplicity.

This Lemma can be seen both as a consequence of the residue Theorem or as a remark in Differential Geometry that can be used to prove the residue Theorem. Such Lemma has a crucial role in the usual proof of the Jordan curve Theorem, since it gives that a smooth, simple and closed curve cannot partition $\mathbb{R}^2$ in *more* than two connected components. A curve fulfilling such constraints splits $\mathbb{R}^2$ in *at least* two connected components by the existence of a *tubular neighbourhood*, then the chance to drop the previous smoothness assumption is granted by invoking Sard's Theorem.

Theorem 260 (The double leash principle)**.** Assume that you are connected to a thin tree by a leash of fixed length L. Assume that you dog is connected to you by a leash of fixed length $l < L$. If you take a walk and after some time you and your dog return at the starting points, your winding number around the tree and your dog's are the same.

The above statement is usually known as **Rouché Theorem**. We opted for such fancy introduction since we believe the previous formulation might help the reader to grasp the geometric idea faster and better. In a more rigorous way:

Theorem 261 (Rouché)**.** If $f(z)$ and $f(z) + g(z)$ are holomorphic function on the closed unit disk D centered at the origin, and for every $z \in \partial D$ we have $0 < |g(z)| < |f(z)|$, then f and $f + g$ have the same number of zeros inside D, where each zero is counted as many times as its multiplicity. The same holds if D is replaced by some compact region K whose boundary ∂K is a simple, piecewise-smooth and closed curve.

Proof. We have already shown that the wanted number of zeroes is given by a winding number. Since for any $z \in \partial D$ we may write

$$f(z) + g(z) = f(z)\left(1 + \frac{g(z)}{f(z)}\right)$$

the winding number of $f + g$ is given by the sum between the winding number of f and the winding number of the curve

$$h : S^1 \to \mathbb{C}^*, \qquad h(e^{i\theta}) \stackrel{\text{def}}{=} 1 + \frac{g(e^{i\theta})}{f(e^{i\theta})}.$$

However, the winding number of h is clearly zero, since h stays "on the right" of the origin. As a matter of fact,

$$\left|\frac{g(z)}{f(z)}\right|$$

is a continuous function on a compact set (S^1) attaining a maximum $M < 1$, hence $\text{Re}(h) \geq 1 - M > 0$ by the triangle inequality. It follows that f and $f + g$ have the same winding number, hence the same number of zeroes in D. □

We are ready to prove the following statement:

Lemma 262. The Fundamental Theorem of Algebra is a simple consequence of the *double leash principle.*

Proof. We may assume without loss of generality that

$$p(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_0 \in \mathbb{C}[z]$$

is a monic polynomial with degree $n \geq 1$ and $p(0) = a_0 \neq 0$. Our purpose is to prove it has a zero *somewhere.* We prove first that *if it has a zero, it cannot be too far from the origin.* Let us consider

$$M = 1 + |a_{n-1}| + \ldots + |a_0|$$

and show that $|z| \geq M$ implies $p(z) \neq 0$. Since $M > 1$, for any $z \in \mathbb{C}$ such that $|z| \geq M$ we have:

$$\left|a_{n-1}z^{n-1} + \ldots + a_0\right| \leq |a_{n-1}||z|^{n-1} + \ldots + |a_0| \leq (|a_{n-1}| + \ldots + |a_0|)\, |z|^{n-1} = (M-1)|z|^{n-1} < |z|^n$$

hence $p(z) = 0$ cannot occur, since it would imply $|z|^n = \left|a_{n-1}z^{n-1} + \ldots + a_0\right|$. This proves all the complex zeroes of p, if existing, lie in the region $|z| < M$. But the inequality above also shows that $f(z) = z^n$ and $g(z) = a_{n-1}z^{n-1} + \ldots + a_0$ meet the hypothesis of Rouché's Theorem for the the region $K = \{z \in \mathbb{C} : |z| \leq M\}$. In particular the number of zeroes of $p(z) = f(z) + g(z)$ in K equals the number of zeroes of $f(z) = z^n$ in K and we are done:

$$p(z) \text{ has exactly } n \text{ zeroes in the region } |z| < M, \text{ counted according to their multiplicity.}$$

□

The inclusion

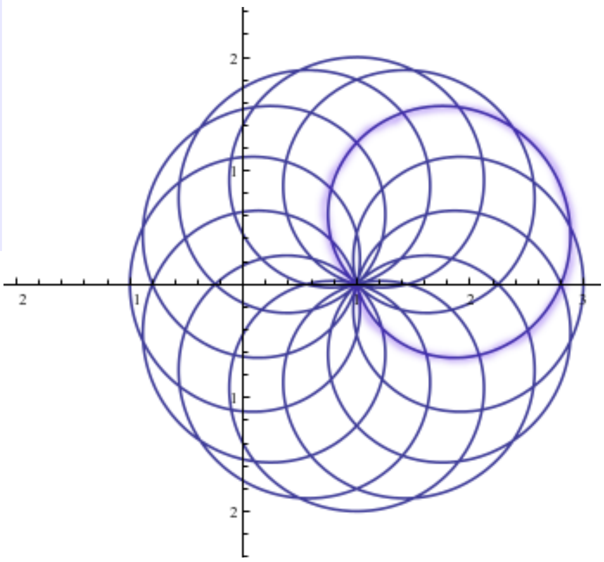
$$\{z \in \mathbb{C} : p(z) = 0\} \subset \{z \in \mathbb{C} : |z| \leq M\}$$

can also be proved by applying the Gershgorin circle Theorem to the companion matrix of p.

Exercise 263. Prove that for any $n \geq 5$ the polynomial

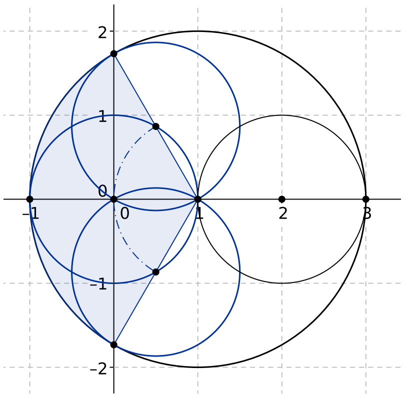
$$p_n(z) = z^n + z + 1$$

has approximately $\frac{n}{3}$ roots in in the region $|z| \leq 1$, with an error not greater than one.



The graph of $e^{13i\theta} + e^{i\theta} + 1$ for $\theta \in [0, 2\pi]$.

Sketch of proof. If $n \equiv 2 \pmod 3$ the primitive third roots of unity ω, ω^2 are also roots of p_n. Conversely, if we denote as D the unit disk centered at the origin, we may notice that ∂D and $-(\partial D + 1)$ intersect only at ω, ω^2: it follows that the $n \equiv 2 \pmod 3$ case is the only case in which $p_n(z)$ has roots at ∂D. The number of roots we want to approximate is given by the winding number of the curve $\gamma_n : [0, 2\pi] \to \mathbb{C}$, $\gamma_n(\theta) = p_n(e^{i\theta})$. The diagram to the right depicts the $n = 13$ case, for instance.



It is pretty clear that the graph of γ_n is given by the union of n approximated circles, completing a revolution around the point $z = 1$ in n steps. Assuming $n \not\equiv 2 \pmod 3$ the number of roots we are interested in is exactly given by the number of the previous approximated circles enclosing the origin. If such portions of γ_n were perfect circles, from the diagram on the left it would be clear that about $\frac{n}{3}$ of them would enclose the origin. To fill in the missing details is a task we leave to the reader. □

We now outline another classical proof of the Fundamental Theorem of Algebra, relying on the following results:

Theorem 264 (Maximum modulus principle)**.** If D is a closed disk in the complex plane and f is a non-constant holomorphic function over D,

$$\max_{z\in D}|f(z)|$$

is attained at ∂D.

Proof. If we assume that $\max_{z\in D}|f(z)|$ is attained at some z_0 belonging to the interior of D we get a contradiction, since by Cauchy's integral formula or termwise integration of a Taylor series we have

$$f(z_0) = \frac{1}{2\pi i}\oint_{|z-z_0|=\varepsilon}\frac{f(z)}{z}\,dz$$

for any $\varepsilon > 0$ small enough, implying

$$|f(z_0)| \leq \frac{1}{2\pi\varepsilon}\oint_{|z-z_0|=\varepsilon}|f(z)|\,dz.$$

If equality holds for any ε small enough then $|f(z)|$ is constant in a neighbourhood of z_0 and $f(z)$ is constant as well. □

Theorem 265 (Liouville)**.** If a holomorphic function over $\mathbb{C}$ is bounded, it is constant.

Proof. The theorem follows from the fact that holomorphic functions are analytic.
If f is an entire function, it can be represented by its Taylor series about 0:

$$f(z) = \sum_{k=0}^{\infty} a_k z^k$$

where by Cauchy's integral formula

$$a_k = \frac{f^{(k)}(0)}{k!} = \frac{1}{2\pi i}\oint_{C_r}\frac{f(\zeta)}{\zeta^{k+1}}\,d\zeta$$

and C_r is the circle about 0 of radius $r > 0$. Suppose f is bounded: i.e. there exists a constant M such that $|f(z)| \leq M$ for all z. We can estimate directly

$$|a_k| \leq \frac{1}{2\pi}\oint_{C_r}\frac{|f(\zeta)|}{|\zeta|^{k+1}}\,|d\zeta| \leq \frac{1}{2\pi}\oint_{C_r}\frac{M}{r^{k+1}}\,|d\zeta| = \frac{M}{2\pi r^{k+1}}\oint_{C_r}|d\zeta| = \frac{M}{2\pi r^{k+1}}2\pi r = \frac{M}{r^k},$$

where in the second inequality we have used the fact that $|z| = r$ on the circle C_r. But the choice of r above is arbitrary. Therefore, letting r tend to infinity gives $a_k = 0$ for all $k \geq 1$. Thus $f(z) = a_0$ and this proves the theorem. □

Corollary 266 (Casorati-Weierstrass)**.** If f is a non-constant entire function, then its range is dense in $\mathbb{C}$.

Proof. If the image of f is not dense, then there is a complex number w and a real number $r > 0$ such that the open disk centered at w with radius r has no element of the image of f. Define $g(z) = \frac{1}{f(z)-w}$. Then g is a bounded entire function, since

$$\forall z \in \mathbb{C}, \quad |g(z)| = \frac{1}{|f(z) - w|} < \frac{1}{r}$$

So, g is constant, and therefore f is constant. □

Corollary 267. If $p(z) \in \mathbb{C}[z]$ is a monic polynomial with degree $n \geq 1$ and $p(0) \neq 0$, it has a complex root.

Proof. Since $|p(z)| \to +\infty$ as $|z| \to +\infty$, there is some closed disk D centered at the origin such that $|p(z)| > |p(0)|$ for any z outside D. Assuming that $p(z)$ is non-vanishing, it follows that $\min_{z\in\mathbb{C}} |p(z)|$ is attained at some $z_0 \in D$ and $q(z) = \frac{1}{p(z)}$ is an entire function such that

$$\forall z \in \mathbb{C}, \qquad |q(z)| \leq \frac{1}{|p(z_0)|}.$$

By Liouville's Theorem we get that both q and p are constant functions, leading to a contradiction. □

A shortened proof. Assume that $p(z) = z^n + a_{n-1}z^{n-1} + \ldots + a_0 \in \mathbb{C}[z]$ with $n \geq 1$ is non-vanishing over $\mathbb{C}$. By the residue Theorem, for any $r > 0$ we have that

$$\oint_{|z|=r} \frac{dz}{z\,p(z)} = \frac{2\pi i}{p(0)} \neq 0$$

but the limit of the LHS as $r \to +\infty$ is clearly 0.

Theorem 268 (Open mapping Theorem)**.** Any non-constant holomorphic function on $\mathbb{C}$ is an *open map*, i.e. it sends open subsets of $\mathbb{C}$ to open subsets of $\mathbb{C}$.

Proof. Assume $f : U \to \mathbb{C}$ is a non-constant holomorphic function and U is a domain of the complex plane. We have to show that every point in $f(U)$ is an interior point of $f(U)$, i.e. that every point in $f(U)$ has a neighbourhood (open disk) which is also in $f(U)$. Consider an arbitrary w_0 in $f(U)$. Then there exists a point z_0 in U such that $w_0 = f(z_0)$. Since U is open, we can find $d > 0$ such that the closed disk B around z_0 with radius d is fully contained in U. Consider the function $g(z) = f(z) - w_0$. Note that z_0 is a root of the function. We know that $g(z)$ is not constant and holomorphic. The roots of g are isolated by the identity theorem, and by further decreasing the radius of the image disk d, we can assure that $g(z)$ has only a single root in B (although this single root may have multiplicity greater than 1). The boundary of B is a circle and hence a compact set, on which $|g(z)|$ is a positive continuous function, so the extreme value Theorem guarantees the existence of a positive minimum e, that is, e is the minimum of $|g(z)|$ for z on the boundary of B and $e > 0$. Denote by D the open disk around w_0 with radius e. By Rouché's theorem, the function $g(z) = f(z) - w_0$ will have the same number of roots (counted with multiplicity) in B as $h(z) \stackrel{\text{def}}{=} f(z) - w_1$ for any w_1 in D. This is because $h(z) = g(z) + (w_0 - w_1)$, and for z on the boundary of B, $|g(z)| \geq e \geq |w_0 - w_1|$. Thus, for every w_1 in D, there exists at least one z_1 in B such that $f(z_1) = w_1$. This means that the disk D is contained in $f(B)$. The image of the ball B, $f(B)$, is a subset of the image of U, $f(U)$. Thus w_0 is an interior point of $f(U)$. Since w_0 was arbitrary in $f(U)$ we know that $f(U)$ is open. Since U was arbitrary, the function f is open. □

Lemma 269. Any non-constant polynomial $p(z) \in \mathbb{C}[z]$ is a closed map.

Proof. We have that $|p(z)| \to +\infty$ as $|z| \to +\infty$. Suppose that $p(z_k) \to w \in C$ as $k \to +\infty$: then $\{z_k\}$ is bounded, so taking a subsequence if necessary, there is $z \in \mathbb{C}$ such that $z_k \to z$. By continuity $p(z_k) \to p(z)$, concluding that $w = p(z)$. □

Corollary 270. If $p(z) \in \mathbb{C}[z]$ is a non-constant polynomial, $p(\mathbb{C})$ is unbounded and simultaneously open and closed. If follows that $p(\mathbb{C}) = \mathbb{C}$, i.e. any non-constant polynomial with complex coefficients is surjective. In particular there is at least a complex solution of $p(z) = 0$.

Exercise 271. Given a non-constant polynomial $p(z) \in \mathbb{C}[z]$ such that $p(0) \neq 0$, prove that the following statements are equivalent forms of the Fundamental Theorem of Algebra:

1. The companion matrix of p has at least an eigenvector in $\mathbb{C}^n$;
2. $q(z) = \frac{1}{p(z)}$ is an analytic function in a neighbourhood of the origin with a finite radius of convergence;
3. If p is the characteristic polynomial of a recurrent sequence $\{a_n\}_{n\geq 0}$, for some $M \in \mathbb{R}^+$ the following limit does not exist:

$$\lim_{n\to+\infty} \sum_{k=0}^{n} a_k M^k.$$

An interesting part of Complex Analysis is related to the problem of extending Rolle's Theorem (if f is a differentiable function on $[a, b]$ and $f(a) = f(b) = 0$ holds, there is some $\xi \in (a, b)$ such that $f'(\xi) = 0$) to the complex case.

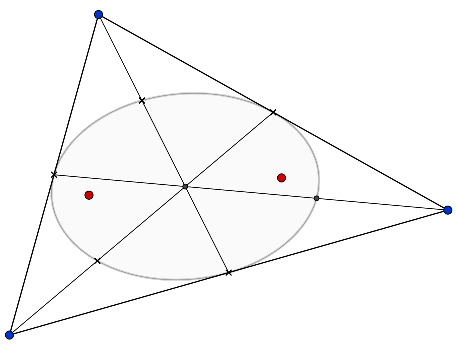
Theorem 272 (Gauss-Lucas)**.** If $\zeta_1, \ldots, \zeta_n \in \mathbb{C}$ ($n \geq 3$) are the roots of a polynomial $p(z) \in \mathbb{C}[z]$, all the roots of $p'(z)$ lie inside the convex hull of $\zeta_1, \ldots, \zeta_n$.

Proof. We may assume without loss of generality that $p(z)$ is a monic polynomial with simple roots. Then by considering the logarithmic derivative of $p(z)$ we have the following identity

$$p'(z) = p(z) \sum_{k=1}^{n} \frac{1}{z - \zeta_k}$$

and $p'(z)$ vanishes iff $s(z) = \sum_{k=1}^{n} \frac{1}{z-\zeta_k}$ vanishes, since $p(z)$ and $p'(z)$ have no common root. Let us assume that $s(z)$ vanishes at a point w lying outside the convex hull of $\zeta_1, \ldots, \zeta_k$, or on its boundary. By the Hann-Banach theorem there is some line ℓ through the origin such that all the vectors $w - \zeta_1, \ldots, w - \zeta_k$ lie on the same side of ℓ. By conjugation, the same applies to the vectors $\frac{1}{w-\zeta_1}, \ldots, \frac{1}{w-\zeta_k}$, hence for some $\theta \in \mathbb{R}$ the complex number $e^{i\theta} s(w)$ has a non-zero real/imaginary part. $s(w) \neq 0$ leads to a contradiction, completing the proof. □

About cubic polynomials, a remarkable and way sharper result is well-known:



Theorem 273 (Marden)**.** If A, B, C are three distinct points in the complex plane and we denote as D, E the roots of

$$\frac{d}{dz}(z - A)(z - B)(z - C),$$

then D, E are the foci of the Steiner inellipse of ABC, centered at $\frac{A+B+C}{3}$ and tangent to the sides of ABC at their midpoints.

What is the best possible improvement of the Gauss-Lucas Theorem still is an open problem in the general case. The following result has only been proved for polynomials having degree ≤ 8 and for some other special cases:

The Ilieff-Sendov conjecture. If all the zeros of a polynomial $p(z)$ lie in $\|z\| \leq 1$ and if r is a zero of $p(z)$, then there is a zero of $p'(z)$ in the disk $\|z - r\| \leq 1$.

Exercise 274. By exploiting the Gauss-Lucas theorem, show that the following entire functions only have real zeroes:

$$\sin(z) + \sqrt{2}\sin(z\sqrt{2}), \qquad \operatorname{Si}(z) = \sum_{n\geq 0} \frac{(-1)^n z^{2n+1}}{(2n+1)\cdot(2n+1)!}, \qquad J_0(z) = \sum_{n\geq 0} \frac{(-z^2)^n}{4^n n!^2}.$$

Exercise 275. By exploiting the Gauss-Lucas theorem, show that for any $\lambda \in [1, +\infty)$ all the solutions of $\cot(x) + \lambda x = 0$ are real numbers.

Yet another way for approaching Calculus. The exponential function is usually introduced by proving that $\lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^n$ exists, then showing that $e^x \stackrel{\text{def}}{=} \lim_{n\to+\infty}\left(1+\frac{x}{n}\right)^n$ is a differentiable function, then noticing that $\frac{d}{dx}e^x = e^x$ leads to very-known Taylor series and to De Moivre's formula. Here we outline a different way for approaching the early stages of Calculus.

1. One may directly introduce the complex exponential function through an everywhere-convergent power series,
$$\forall z \in \mathbb{C}, \qquad e^z \stackrel{\text{def}}{=} \sum_{n\geq 0} \frac{z^n}{n!},$$
2. then check through the convolution machinery that such function fulfills $e^a \cdot e^b = e^{a+b}$ for any $a, b \in \mathbb{C}$;
3. In particular, for any $\theta \in \mathbb{R}$ we have that $e^{i\theta} \in S^1$, since $|e^{i\theta}|^2 = e^{i\theta} \cdot e^{-i\theta} = 1$
4. and the map $\gamma : \mathbb{R} \to S^1$ given by $\gamma(\theta) = e^{i\theta}$ is a parametrization of S^1 with constant speed, since by the series definition $\frac{d}{d\theta}e^{i\theta} = ie^{i\theta}$ and the last quantity has unit modulus by the previous point (Pythagorean Theorem);
5. Since γ is an arc-length parametrization of S^1, we may define $\sin(\theta)$ and $\cos(\theta)$ through
$$\sin(\theta) \stackrel{\text{def}}{=} \operatorname{Im} e^{i\theta}, \qquad \cos(\theta) \stackrel{\text{def}}{=} \operatorname{Re} e^{i\theta}$$
and derive the addition formulas for sin and cos from the point (2.);
6. Since γ travels S^1 counter-clockwise, by defining π as
$$\pi \stackrel{\text{def}}{=} \inf\left\{\theta \in \mathbb{R}^+ : \sin(\theta) = 0\right\}$$
we get that π equals half the length of the unit circle, or, equivalently, the area of the unit circle. Additionally, $e^{i\pi} + 1 = 0$;
7. By the previous points e^z is an entire function and a solution of the differential equation $f'(z) = f(z)$;
8. By the series definition it also follows that
$$\sin(z) = \sum_{n\geq 0} \frac{(-1)^n}{(2n+1)!} z^{2n+1}, \qquad \cos(z) = \sum_{n\geq 0} \frac{(-1)^n}{(2n)!} z^{2n}$$
are entire functions and solutions of the differential equation $f''(z) + f(z) = 0$;

9. From the Pythagorean Theorem $\sin^2\theta + \cos^2\theta = 1$, hence

$$\pi = 2\int_0^1 \frac{dx}{\sqrt{1-x^2}} \overset{x\mapsto\sqrt{u}}{=} \int_0^1 \frac{du}{\sqrt{u(1-u)}} \overset{\text{symmetry}}{=} 4\int_0^{1/\sqrt{2}} \frac{dx}{\sqrt{1-x^2}}$$

and by integrating termwise the Taylor series of $\frac{1}{\sqrt{1-x^2}}$ we get the following series representation for π:

$$\pi = 2\sqrt{2}\sum_{n\geq 0}\frac{\binom{2n}{n}}{8^n(2n+1)} = 3.14159265358979323846264338327950\ldots$$

As an alternative, the integral of $\sqrt{1-x^2}$ over some sub-interval of $[-1,1]$ is clearly related to the area of a circle sector. By computing a Taylor series and applying termwise integration again, we may easily derive Newton's identity

$$\pi = 4 - 4\sum_{n\geq 1}\frac{\binom{2n}{n}}{(4n^2-1)4^n}.$$

The Lagrange inversion theorem in a nutshell.
Let us assume to have a holomorphic function which is $z + o(z)$ in a neighbourhood of the origin, like

$$\sin(z) = \sum_{n\geq 0}\frac{(-1)^n z^{2n+1}}{(2n+1)!} \tag{1}$$

and to want to compute the coefficients of the Maclaurin series of its inverse function $\arcsin(z)$, say the coefficient of z^7. By Cauchy's integral formula

$$[z^7]\arcsin(z) = \frac{1}{2\pi i}\oint_{|z|=\varepsilon}\frac{\arcsin(z)}{z^8}\,dz \tag{2}$$

and something nice happens$^{(*)}$ if we enforce the substitution $z = \sin u$ in the RHS of (2). The simple contour around the origin $|z| = \varepsilon$ is mapped into a similar (homeomorphic) simple contour around the origin by a conformal map, hence

$$[z^7]\arcsin(z) = \frac{1}{2\pi i}\oint_{|u|=\varepsilon}\frac{u\cos(u)}{\sin(u)^8}\,du \tag{3}$$

and the problem boils down to evaluating the residue of $\frac{u\cos u}{\sin(u)^8}$ at the origin, which is a pole of order 7 for such a function. In particular

$$\operatorname*{Res}_{u=0}\frac{u\cos u}{\sin(u)^8} = \lim_{u\to 0}\frac{1}{6!}\frac{d^6}{du^6}\left(u^7\cdot\frac{u\cos u}{\sin(u)^8}\right) = \lim_{u\to 0}\frac{1}{7!}\frac{d^6}{du^6}\left(\frac{u}{\sin u}\right)^7 = \frac{5}{112} \tag{4}$$

and the whole tour proves a connection between the Maclaurin coefficients of arcsin and the derivatives of $\left(\frac{u}{\sin u}\right)^k$ at the origin.

It is pretty natural to wonder if the computation of the derivatives at the origin of $\left(\frac{u}{f(u)}\right)^k$ for some holomorphic $f(u) = u + o(u)$ is a simple task. Well, in general it is not. For instance the Maclaurin series of arcsin can be computed with considerably fewer efforts by applying the extended binomial theorem to $\frac{d}{du}\arcsin(u) = \frac{1}{\sqrt{1-u^2}}$. On the other hand something really nice is produced by this approach by considering $f(u) = ue^u$, i.e. the Maclaurin series of the Lambert W function:

$$W(x) = \sum_{n\geq 1}\frac{(-1)^{n+1}n^{n-1}}{n!}x^n \quad\Longrightarrow\quad \sum_{n\geq 1}\frac{n^{n-1}}{n!e^n} = 1 \tag{!}$$

and the crucial part of the argument $(*)$ can be used for finding the Maclaurin series of $\arcsin^2$, $\arcsin^3$, $\arcsin^4$ etcetera, leading to some non-trivial hypergeometric identities.

Now we introduce a concise form of the statement and another example.

Theorem 276 (Lagrange's inversion formula)**.** If $f(z)$ is a holomorphic function in a neighbourhood of the origin, such that $f(z) = z + o(z)$ as $z \to 0$, we have

$$f^{-1}(z) = \sum_{n\geq 1} \frac{z^n}{n} \cdot [z^{n-1}] \left(\frac{z}{f(z)}\right)^n$$

where $[z^m]g(z)$ stands for the coefficient of z^m in the Maclaurin series of $g(z)$.

More generally, if f, h are holomorphic functions in a neighbourhood of the origin and $f(z) = z + o(z)$,

$$h(f^{-1}(z)) = h(0) + \sum_{n\geq 1} \frac{z^n}{n} \cdot [z^{n-1}] \left(h'(z) \cdot \left(\frac{z}{f(z)}\right)^n\right).$$

Another celebrated application is given by Catalan numbers. It is straightforward to prove in a combinatorial fashion that they fulfill $C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}$, hence their ordinary generating function multiplied by z is given by the inverse function of $f(z) = z - z^2$. By Lagrange's inversion formula

$$f^{-1}(z) = \sum_{n\geq 1} \frac{z^n}{n} [z^{n-1}] \left(\frac{1}{1-z}\right)^n$$

and by *stars and bars* $\frac{1}{(1-z)^n} = \sum_{m\geq 0} \binom{m+n-1}{m} z^n$, hence

$$f^{-1}(z) = \sum_{n\geq 1} \frac{z^n}{n} \binom{2n-2}{n-1}$$

and

$$C_n = \frac{1}{n+1}\binom{2n}{n}.$$

For further references, see A.D. Sokal, *A ridiculously simple and explicit implicit function theorem*, https://arxiv.org/pdf/0902.0069.

Exercise 277. Find the asymptotic behaviour of

$$a_n = \sum_{k=0}^{n} \frac{(n-k)^k}{k!}.$$

A sensible approach is to consider the ordinary generating function of the sequence $\{a_n\}_{n\geq 0}$:

$$\frac{1}{1-xe^x} = \sum_{m\geq 0} x^m e^{mx} = \sum_{m\geq 0}\sum_{k\geq 0} \frac{m^k x^{m+k}}{k!} = \sum_{n\geq 0} x^n \sum_{k=0}^{n} \frac{(n-k)^k}{k!} = \sum_{n\geq 0} a_n x^n$$

then to perform a partial fraction decomposition of the LHS.
By naming Ω the only real solution of $xe^x = 1$ we have

$$\operatorname*{Res}_{x=\Omega} \frac{1}{1-xe^x} = \lim_{x\to\Omega} \frac{1}{-(x+1)e^x} = -\frac{1}{1+e^\Omega} = -\frac{\Omega}{\Omega+1}$$

hence the contribution to a_n coming from the simple pole at Ω exactly equals

$$\frac{1}{\Omega+1}[x^n]\frac{1}{1-\frac{x}{\Omega}} = \frac{e^{\Omega n}}{\Omega+1}.$$

It remains to show that the contribution coming from any other complex pole is negligible. These poles obviously come in conjugated pairs, and if a pole in the upper half-plane occurs at $\sigma + i\tau$ we have

$$\exp(2\sigma) = \frac{1}{\sigma^2+\tau^2}, \quad \tan(\tau) = -\frac{\tau}{\sigma},$$

allowing to state that the k-th couple of complex poles (ranked according to the absolute value of the imaginary part) occurs at $-\log(2\pi k) \pm \left(2k-\frac{1}{2}\right)\pi i + o(1)$, in the *left* half-plane. By the position of the first pair of complex roots

$$a_n = \frac{e^{\Omega n}}{\Omega+1} + O\left(e^{-3n/2}\right).$$

This also gives that $\frac{a_n}{a_{n+1}}$ converges pretty fast to Ω.
For instance, the difference between Ω and $\frac{a_4}{a_5} = \frac{148}{261}$ is already less than 10^{-4}.

A remark on a classical integral.
Let us consider for some $m \in \mathbb{N}^+$,

$$I(m) = \int_{-\infty}^{+\infty} \frac{dx}{x^{2m}+1}.$$

By exploiting parity, the substitution $\frac{1}{x^{2m}+1} = u$, Euler's Beta function and the reflection formula for the Γ function we have $I(m) = \frac{\pi}{m \sin\frac{\pi}{2m}}$. Let us prove this identity through the residue theorem, too. $f(z) = \frac{1}{z^{2m}+1}$ has simple poles at $\zeta_k = \exp\left(\frac{2\pi i}{4m}(2k-1)\right)$ for $k = 1, 2, \ldots, 2m$.

$$\operatorname*{Res}_{z=\zeta_k} \frac{1}{z^{2m}+1} = \lim_{z\to\zeta_k} \frac{z-z_k}{z^{2m}+1} \stackrel{d.H.}{=} \lim_{z\to z_k} \frac{z}{2mz^{2m}} = -\frac{\zeta_k}{2m}$$

hence

$$I(m) = -\frac{\pi i}{m}\sum_{k=1}^{m}\zeta_k = -\frac{\pi i}{m}\sum_{k=1}^{m}\zeta_1^{2k-1} = -\frac{\pi i}{m}\cdot\frac{\zeta_1(\zeta_1^{2m}-1)}{\zeta_1^2-1} = \frac{2\pi i}{m\left(\zeta_1-\zeta_1^{-1}\right)}$$

and by De Moivre's formula we get

$$I(m) = \int_{-\infty}^{+\infty} \frac{dx}{x^{2m}+1} = \frac{\pi}{m\sin\frac{\pi}{2m}}.$$

Exercise 278. Prove that for any $n \in \mathbb{N}$ the following identity holds:

$$\mathcal{S}(n) = \sum_{k=0}^{n} 2^k \binom{2n-k}{n} = 4^n.$$

Proof.

$$\mathcal{S}(n) = \sum_{k=0}^{n} 2^{n-k}\binom{n+k}{n}$$

is the coefficient of x^n in the product between $\sum_{k\geq 0} 2^k x^k = \frac{1}{1-2x}$ (geometric series) and $\sum_{k\geq 0}\binom{n+k}{n}x^k = \frac{1}{(1-x)^{n+1}}$

(stars and bars). In particular

$$\mathcal{S}(n) = \text{Res}\left(\frac{1}{(1-2x)\left[x(1-x)\right]^{n+1}}, x=0\right)$$

but due to the symmetry of the meromorphic function $\frac{1}{(1-2x)[x(1-x)]^{n+1}}$ the residue at 0 and the residue at 1 are the same number. The only other pole is at $x = \frac{1}{2}$ and the sum of the residues is zero, hence $\mathcal{S}(n) = 4^n$ can be proved from the straightforward

$$\text{Res}\left(\frac{1}{(1-2x)\left[x(1-x)\right]^{n+1}}, x=\frac{1}{2}\right) = -2\cdot 4^n.$$

□

10 Quantitative forms of the Weierstrass approximation Theorem

In a previous section about Chebyshev and Legendre polynomials we have already seen a brilliant proof of Weierstrass approximation Theorem, where the problem of finding uniform polynomial approximations for continuous functions has been reduced to finding uniform polynomial approximations for a *single* function, namely $|x|$. In the current section we will investigate about a fascinating subject: by fixing some interval $[a,b]\subset\mathbb{R}$ and recalling that $\|f-g\|_\infty = \sup_{x\in[a,b]}|f(x)-g(x)|$, we ask:

> Given a function $f\in C^0[a,b]$, let us denote as $p_N(x)$ the polynomial with degree N minimizing $\|f-p_N\|_\infty$. Is there some relation between the regularity of f and the speed of convergence towards zero for $\|f-p_N\|_\infty$, as $N\to+\infty$?

Our main goal is to prove the following results:

Theorem 279 (Bernstein)**.** If $f:[0,2\pi]\to\mathbb{C}$ is a 2π-periodic function, n is a natural number, $\alpha\in(0,1)$ and for some constant $C(f)>0$ the sequence of trigonometric polynomials $\{P_n(x)\}_{n\geq 0}$ fulfills

$$\partial P_n = n,\qquad \forall n\geq 1,\quad \|f-P_n\|_\infty \leq \frac{C(f)}{n^{r+\alpha}}$$

then $f(x) = P_0(x)+\varphi(x)$, where $\varphi(x)$ is a function of class C^r on the interval $(0,2\pi)$ and $\varphi^{(r)}(x)$ is a α-Hölder continuous function.

Theorem 280 (Jackson)**.** If $f:[0,2\pi]\to\mathbb{C}$ is a 2π-periodic function of class C^r such that $\left|f^{(r)}(x)\right|\leq 1$ for any $x\in[0,2\pi]$, there exists a sequence $\{P_n(x)\}_{n\geq 1}$ of trigonometric polynomials such that:

$$\partial P_n = n,\qquad \|f-P_n\|_\infty \leq \frac{C(r)}{n^r}$$

where $C(r)$ only depends on r, and it is the Akhiezer-Krein-Favard constant $C(r) = \frac{4}{\pi}\sum_{k\geq 0}\frac{(-1)^{k(r+1)}}{(2k+1)^{r+1}}$.

In other terms, *the degree of regularity of a function is exactly given by the speed of convergence towards zero of the uniform error for the optimal polynomial approximations with respect to the infinity-norm.* The Bernstein Theorem follows from Bernstein's inequality, stating:

Theorem 281 (Bernstein's inequality)**.** If $P(z)$ is a polynomial with complex coefficients and degree n,

$$\max_{|z|\leq 1}|P'(z)| \leq n\cdot\max_{|z|\leq 1}|P(z)|$$

from which it follows that:

$$\max_{|z|\leq 1}\left|P^{(k)}(z)\right| \leq \frac{n!}{(n-k)!}\cdot\max_{|z|\leq 1}|P(z)|\,.$$

Jackson's Theorem, instead, follows from a projection technique in L^2 very similar to the one we saw when computing the Fourier-Chebyshev expansion for $|x|$. Our path starts by presenting Bernstein's proof of Weierstrass approximation Theorem, dating back to 1912. If f is a continuous function on the interval $[0,1]$, the sequence of polynomials defined through

$$\left(B_n(f)\right)(x) = \sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}f\left(\frac{k}{n}\right)$$

is uniformly convergent towards f. An interesting remark is that we need to prove the claim only for the following cases: $f_0(x)=1$, $f_1(x)=x$ and $f_2(x)=x^2$. Indeed we have $B_n(f_0)=f_0$, $B_n(f_1)=f_1$ and $B_n(f_2)=\left(1-\frac{1}{n}\right)f_2+\frac{1}{n}f_1$. Moreover

$$\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^2\binom{n}{k}x^k(1-x)^{n-k} = \frac{x(1-x)}{n}\leq\frac{1}{4n}$$

and by considering some $\delta>0$ and by defining F as the set of ks in $\{0,\ldots,n\}$ such that $\left|\frac{k}{n}-x\right|\geq\delta$, we have:

$$\begin{aligned}\sum_{k\in F}\binom{n}{k}x^k(1-x)^{n-k} &\leq \frac{1}{\delta^2}\sum_{k\in F}\left(\frac{k}{n}-x\right)^2\binom{n}{k}x^k(1-x)^{n-k}\\ &\leq \frac{1}{\delta^2}\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^2\binom{n}{k}x^k(1-x)^{n-k}\leq\frac{1}{4n\delta^2}.\end{aligned}$$

Since f is a continuous function on a compact set, it is uniformly continuous, hence there is some $\delta>0$ which ensures $|f(x)-f(y)|\leq\frac{\varepsilon}{2}$ as soon as $|x-y|<\delta$. Since $\binom{n}{k}x^k(1-x)^{n-k}$ are non-negative numbers adding to 1,

$$\begin{aligned}|f(x)-B_n(f)(x)| &= \left|f(x)-\sum_{k=0}^{n}f\left(\frac{k}{n}\right)\binom{n}{k}x^k(1-x)^{n-k}\right|\\ &= \left|\sum_{k=0}^{n}\left[f(x)-f\left(\frac{k}{n}\right)\right]\binom{n}{k}x^k(1-x)^{n-k}\right|\\ &\leq \sum_{k=0}^{n}\left|f(x)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k}.\end{aligned}$$

If now we fix n (*how* will be clear soon) and denote $\|f\|=\max_{x\in[0,1]}|f(x)|$, by naming as F the set of ks such that $|x-k/n|\geq\delta$, we have $|f(x)-f(k/n)|\leq\frac{\varepsilon}{2}$ for any k in the complement of F and $|f(x)-f(k/n)|\leq 2\|f\|$ in general. It follows that:

$$\begin{aligned}|f(x)-B_n(f)(x)| &\leq \sum_{k\in F}\left|f(x)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k}+\sum_{k\notin F}\left|f(x)-f\left(\frac{k}{n}\right)\right|\binom{n}{k}x^k(1-x)^{n-k}\\ &\leq 2\|f\|\sum_{k\in F}\binom{n}{k}x^k(1-x)^k+\frac{\varepsilon}{2}\sum_{k\notin F}\binom{n}{k}x^k(1-x)^{n-k}\\ &\leq \frac{\|f\|}{2n\delta^2}+\frac{\varepsilon}{2}\end{aligned}$$

and the last term is smaller than ε as soon as $n>\frac{\|f\|}{\varepsilon\delta^2}$. A remarkable feature of Bernstein's approach is that the properties $B_n(\alpha f+\beta g)=\alpha B_n(f)+\beta B_n(g)$ and $B_n(f)\geq 0$ if $f\geq 0$ are enough to lead to a generalization:

Theorem 282 (Bohman, Korovkin, 1952)**.** Assuming that $T_n: C^0[0,1]\to C^0[0,1]$ is a sequence of linear and positive operators, such that $T_n(f)$ is uniformly convergent to f in each one of the following cases

$$f_0(x)=1,\qquad f_1(x)=x,\qquad f_2(x)=x^2,$$

then $T_n(f)$ is uniformly convergent to f for any $f\in C^0[0,1]$.

This result already allows to outline a sketch of proof for Jackson's Theorem: let us assume that f is a function of class C^r, 2π-periodic. In order to approximate f through trigonometric polynomials, the key idea is to approximate $f^{(r)}$ first, then exploit repeated termwise integration. By setting $g = f^{(r)}$, a natural approach we may follow is to consider a truncation of the Fourier series of g, i.e. to consider a convolution between g and Dirichlet's kernel D_N. However, Dirichlet's kernel is not positive or bounded with respect to the L^1-norm (this is the main reason fo the Gibbs phenomenon to arise in peculiar situations), hence the initial *naive* idea requires to be fixed. In particular, instead of considering truncations of Fourier series, it is better to consider suitably weighted Fourier series. The map:

$$g(\theta) = \sum_{n\in\mathbb{Z}} c_n e^{ni\theta} \quad \longrightarrow \quad \tilde{g}(\theta) = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N}\right) c_n e^{in\theta}$$

follows the idea of approximating g through the convolution between g and the Fejér kernel F_N. This kernel is positive and concentrated around the origin, hence $g * F_N$ is uniformly convergent to g on the interval $[0, 2\pi]$. In order to achieve a better control on error with respect to the L^2-norm, a more efficient way is to approximate g through the convolution between g and the square of Fejér kernel, also known as **Jackson's kernel**. By squaring the positivity of such kernel is preserved, and it becomes more concentrated around the origin. Repeated termwise integration then leads to Jackson's theorem in a quite straightforward way. The Akhiezer-Krein-Favard constant comes from the fact that among the 2π-periodic and continuous functions, the triangle wave is in some sense the *worst approximable* function through such convolution trick, but we already computed the Fourier-Chebyshev expansion of the triangle wave.

As an alternative, it is not terribly difficult to prove the following statement, that follows by combining Lagrange interpolation with properties of Chebyshev polynomials:

Theorem 283. If $f(x)$ is a function of class C^{n+1} on the interval $[-1, 1]$ and $C_n(f)$ is the Lagrange interpolating polynomial for f, with respect to the points given by the roots of the Chebyshev polynomial T_{n+1} (also known as *Chebyshev nodes*), we have:

$$|f(x) - C_n(f)(x)| \leq \frac{\left\|f^{(n+1)}\right\|}{2^n(n+1)!}.$$

Additionally, we have that:

Lemma 284. If $p(x)$ is a real polynomial with degree $\leq (n-1)$, we have:

$$p(x) = \sum_{k=1}^{n} p(x_k)(-1)^{k-1}\sqrt{1-x_k^2}\,\frac{T_n(x)}{x-x_k}$$

with $x_1, \ldots, x_n$ being the roots of Chebyshev polynomial $T_n(x)$.
In this case, indeed, Lagrange interpolation leads to an identity, not only an approximation.

Under the same assumptions we also have:

$$\max_{x\in[-1,1]} |p(x)| \leq n \max_{x\in[-1,1]} \left|\sqrt{1-x^2}\,p(x)\right| = M.$$

If $x \in [x_n, x_1]$, i.e. if $|x| \leq \cos\frac{\pi}{2n}$, we have $\sqrt{1-x^2} \geq \sin\frac{\pi}{2n} \geq \frac{1}{n}$, so such inequality is trivial.
On the other hand, if all the terms $(x - x_k)$ have the same sign, we have:

$$|p(x)| \leq \frac{M}{n^2}\left|\sum_{k=1}^{n}\frac{T_n(x)}{x-x_k}\right| = \frac{M}{n^2}\,|T_n'(x)| \leq M.$$

By enforcing the substitution $x \to \cos\theta$, we have that if $S(\theta)$ is an odd trigonometric polynomial having degree n, then:

$$\max_{\theta\in[0,2\pi]}\left|\frac{S(\theta)}{\sin\theta}\right| \leq n \max_{\theta\in[0,2\pi]}|S(\theta)|.$$

We are ready to prove Bernstein's inequality. If $S(\theta)$ is a trigonometric polynomial with degree n, we may define the auxiliary function

$$f(\alpha,\theta) = \frac{S(\alpha+\theta)-S(\alpha-\theta)}{2}$$

and this function (with respect to θ) is an odd trigonometric polynomial with degree $\leq n$. Due to previous results,

$$\max_{\theta\in[0,2\pi]}\left|\frac{f(\alpha,\theta)}{\sin\theta}\right| \leq n \max_{\theta\in[0,2\pi]}|f(\alpha,\theta)| \leq n \max_{\theta\in[0,2\pi]}|S(\theta)|.$$

However

$$S'(\alpha) = \lim_{\theta\to 0}\frac{S(\alpha+\theta)-S(\alpha-\theta)}{2\theta} = \lim_{\theta\to 0}\frac{f(\alpha,\theta)}{\sin\theta}$$

hence for any α, $|S'(\alpha)| \leq n\max_{\theta\in[0,2\pi]}|S(\theta)|$ as wanted.
By reversing the previous substitution, we also get:

Theorem 285 (Markov brothers' inequality)**.** If $p(x)$ is a real polynomial with degree n,

$$\max_{x\in[-1,1]}|p'(x)| \leq n^2 \max_{x\in[-1,1]}|p(x)|.$$

We now prove a *minor version* of Bernstein's Theorem:

Lemma 286. If f is a 2π-periodic function and for some $\alpha\in(0,1)$ the given function can be approximated through trigonometric polynomials with degree n, where the uniform error of such approximations is $\leq \frac{A}{n^\alpha}$, then f is an α-Hölder-continuous function.

Proof. For any $n\geq 1$, we may pick a trigonometric polynomial S_n such that $\|f-S_n\|\leq\frac{A}{n^\alpha}$. In particular, $\{S_n\}_{n\geq 1}$ is uniformly convergent. With the assumptions $V_0 = S_1$ and $V_n = S_{2^n} - S_{2^{n-1}}$, V_n turns out to be a trigonometric polynomial having degree $\leq 2^n$ and fulfilling $f = \sum_{n\geq 0} V_n$. Indeed we have

$$\|V_n\| \leq \|S_{2^n} - f\| + \|S_{2^{n-1}} - f\| \leq A\left(2^{-n\alpha} + 2^{\alpha-n\alpha}\right) \leq B\cdot 2^{-n\alpha}$$

and the RHS is summable, hence $\sum_{n\geq 0} V_n$ is uniformly convergent to f. The key idea is now to approximate $|f(x)-f(y)|$ through a finite number of V_ns, number to suitably fix later.

$$\begin{aligned}
|f(x)-f(y)| &\leq \sum_{n\geq 0}|V_n(x)-V_n(y)| \\
&\leq \sum_{n=0}^{m-1}|V_n(x)-V_n(y)| + 2\sum_{n\geq m}\|V_n\| \\
&= |x-y|\sum_{n=0}^{m-1}|V_n'(\xi_n)| + 2B\sum_{n\geq m}2^{-n\alpha} \\
&\overset{\text{Bern}}{\leq} |x-y|\sum_{n=0}^{m-1}2^n\|V_n\| + 2B\sum_{n\geq m}2^{-n\alpha} \\
&\leq C\left[|x-y|\,2^{m(1-\alpha)} + 2^{-m\alpha}\right].
\end{aligned}$$

We want the last term to be bounded by $|x-y|^\alpha$, i.e., by setting $\delta = |x-y|$, we want the following inequality to hold:

$$(2^m\delta)^{1-\alpha} + (2^m\delta)^\alpha \leq D$$

and for such a purpose it is enough to choose m in such a way that $1 < 2^m\delta < 2$ is granted. □

Bernstein's theorem, as presented at the beginning of this section, is a simple generalization of the last statement. There is a non-trivial subtlety: one might conjecture that if f is a 2π-periodic function, admitting approximations through trigonometric polynomials with degree n, whose uniform error is $\leq \frac{A}{n}$, then f is a Lipschitz-continuous function. However that is not the case: we may state, at best, that such assumptions and $|x-y| \leq \delta$ grant $|f(x)-f(y)| \leq K\delta\,|\log\delta|$ for any δ small enough. Essentially, a form of Gibbs phenomenon arises in the proof of Bernstein's Theorem, too.

At last we mention an interesting strengthening of Weierstrass approximation Theorem, which can be interpreted as a deep result for dealing with lacunary Fourier series, or as a deep result in the theory of interpolation.

Theorem 287 (Müntz)**.** If $0 = \lambda_0, \lambda_1, \lambda_2, \ldots$ is an increasing sequence of natural numbers, the subspace of $C^0[0,1]$ spanned by $x^{\lambda_0}, x^{\lambda_1}, x^{\lambda_2}, \ldots$ is dense in $C^0[0,1]$ if and only if the series $\sum_{n\geq 1}\frac{1}{\lambda_n}$ is divergent.

Exercise 288 (Chebyshev equioscillation theorem)**.** Let f be a continuous function on the interval $[a,b]$. Prove that among the polynomials $p \in \mathbb{R}[x]$ having degree $\leq n$, the polynomial minimizing $\|p-f\|_\infty$ fulfills

$$p(x_k) - f(x_k) = \sigma(-1)^k\|p-f\|_\infty$$

at $n+2$ points $a \leq x_0 < x_1 \ldots < x_n \leq b$ with $\sigma \in \{-1,+1\}$.

Exercise 289. $f(x) = ax^3 + bx^2 + cx + d$ is a polynomial with real coefficients fulfilling the following property:

$$\forall x \in [0,1], \quad |f(x)| \leq 1.$$

Prove that $|a|+|b|+|c|+|d| \leq 99$ holds and that such inequality is optimal, i.e. 99 cannot be replaced by any smaller number.

Exercise 290. Find the explicit values of C and D in

$$\int_0^{2\pi}\left|\frac{\pi-x}{2} - \sum_{n=1}^{N}\frac{\sin(nx)}{n}\right|dx \sim \frac{C}{N}\log(DN) \quad \text{as } N \to +\infty.$$

11 Elliptic integrals and the AGM

The flourishing theory of elliptic integrals started its development at the beginning of the nineteenth century, especially thanks to Niels Henrik Abel, whose work was really appreciated by Carl Gustav Jacob Jacobi ("*Abel teaches us: you always have to consider an inversion!*"). Jacobi gave significant contributions to the theory of elliptic integrals,

leading to the foundations of the modern theory of modular forms. In Physics elliptic integrals arise, for instance, in the determination of the exact period of a pendulum.
Let us assume to have a mass-m point at one end of a massless, inextensible rod of length l, with the other end hinged at fixed point. Let us denote as θ the angle between the position of the pendulum and the equilibrium point, in a frictionless environment with a constant gravitational acceleration in the same direction. The mechanical energy of the pendulum equals

$$mgl(1-\cos\theta)+\frac{ml^2}{2}\dot{\theta}^2=E$$

and this quantity is time-invariant. The instants such that $\dot{\theta}=0$ give the amplitude of oscillations, and by assuming $E\leq mgl$ (i.e. by assuming the pendulum has not enough energy to go above the hinge) we have:

$$\theta_{max}=2\arcsin\sqrt{\frac{E}{2mgl}}.$$

The conservation of mechanical energy allows us to write a differential equation for θ:

$$\frac{d\theta}{dt}=\sqrt{\frac{2E}{ml^2}-\frac{2g}{l}(1-\cos\theta)}$$

from which it follows that time can be written as a function of θ, and in particular:

$$T=4\int_0^{\theta_{max}}\frac{d\theta}{\sqrt{\frac{2E}{ml^2}-\frac{2g}{l}(1-\cos\theta)}}=4\sqrt{\frac{l}{g}}\int_0^1\frac{du}{\sqrt{(1-u^2)\left(1-\frac{Eu^2}{2mgl}\right)}}=4\sqrt{\frac{l}{g}}\int_0^{\pi/2}\frac{d\varphi}{\sqrt{1-\frac{E\sin^2\varphi}{2mgl}}}.$$

For small energies an accurate approximation of the period is thus given by:

$$4\sqrt{\frac{l}{g}}\int_0^1\frac{du}{\sqrt{1-u^2}}=2\pi\sqrt{\frac{l}{g}}$$

and the pendulum is approximately *isocronous* (the time needed to complete an oscillation does not depend on the amplitude of such oscillation, just like in the idea harmonic oscillator described by the differential equation $\ddot{x}+\frac{g}{l}x=0$).
In a exact form, by setting $\kappa^2=\frac{E}{2mgl}$:

$$T=4\sqrt{\frac{l}{g}}\int_0^{\pi/2}\frac{d\varphi}{\sqrt{1-\kappa^2\sin^2\varphi}}=2\pi\sqrt{\frac{l}{g}}\left[1+\sum_{m\geq 1}\left(\binom{2m}{m}\frac{1}{4^m}\right)^2\frac{\kappa^{2m}}{4}\right]\leq 2\pi\sqrt{\frac{l}{g}}\left[1-\frac{1}{4\pi}\log(1-\kappa^2)\right].$$

we get that an approximation of the period that is much more accurate than $2\pi\sqrt{\frac{l}{g}}$, also for non-negligible values of E, is provided by:

$$T\approx 2\pi\sqrt{\frac{l}{g}}\left[1+\frac{\log(1-\kappa^2)}{16}\right]=2\pi\sqrt{\frac{l}{g}}\left[1+\frac{\log\cos\frac{\theta_{max}}{2}}{8}\right].$$

In particular, the exact period of a pendulum is given by a **complete elliptic integral of the first kind**.
The adjective *elliptic* has probably been chosen for a geometric reason: let E be an ellipse described by the equation $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$, with major axis $2a$, minor axis $2b$, semi-focal distance $c=\sqrt{a^2-b^2}$ and eccentricity $e=\frac{c}{a}$. Since any affine map preserves the ratios of areas, the area of E equals πab. Let us denote as $L(a,b)$ the ellipse perimeter. By the integral formula for computing the length of a C^1 curve we have:

$$L(a,b)=4\int_0^{\pi/2}\sqrt{a^2\sin^2\theta+b^2\cos^2\theta}\,d\theta=4a\int_0^{\pi/2}\sqrt{1+e^2\sin^2\theta}\,d\theta.$$

In particular, the ellipse perimeter is given by a **complete elliptic integral of the second kind**.
Let us denote as $M_p(a,b)$ the order-p mean between a and b, i.e.:

$$M_p(a,b)=\sqrt[p]{\frac{a^p+b^p}{2}}.$$

Theorem 291.

$$M_1(a,b) \leq \frac{L(a,b)}{2\pi} \leq M_2(a,b).$$

Proof. By the concavity of the logarithm function we have $\sqrt{x}+\sqrt{y} \leq \sqrt{2(x+y)}$, from which:

$$\begin{aligned} L(a,b) &= 4\int_0^{\pi/4} \sqrt{a^2\sin^2\left(\frac{\pi}{4}-\theta\right)+b^2\cos^2\left(\frac{\pi}{4}-\theta\right)}+\sqrt{a^2\sin^2\left(\frac{\pi}{4}+\theta\right)+b^2\cos^2\left(\frac{\pi}{4}+\theta\right)}\,d\theta \\ &\leq 4\int_0^{\pi/4} \sqrt{2(a^2+b^2)}\,d\theta = 2\pi\cdot M_2(a,b). \end{aligned}$$

The first inequality, despite its appearance, is actually more difficult to prove.
We may use the following identity as a starting point:

$$L(a,b) = (a+b)\int_0^{\pi} \sqrt{1+\left(\frac{a-b}{a+b}\right)^2+2\left(\frac{a-b}{a+b}\right)\cos(2\theta)}\,d\theta.$$

Since $1+x^2+2x\cos(2\theta) = (1-x\,e^{2i\theta})\cdot(1+x\,e^{-2i\theta})$ we have:

$$\begin{aligned} \int_0^{\pi} \sqrt{1+x^2+2x\cos(2\theta)}\,d\theta &= \int_0^{\pi} \sqrt{(1-x\,e^{2i\theta})\cdot(1+x\,e^{-2i\theta})}\,d\theta \\ &= \int_0^{\pi} \sum_{m,n\geq 0} \frac{1}{(2m-1)(2n-1)}\cdot\frac{1}{4^{m+n}}\cdot\binom{2m}{m}\cdot\binom{2n}{n}\cdot x^{m+m}\cdot e^{2(m-n)i\theta}\,d\theta \\ &= \pi\sum_{n=0}^{+\infty}\left(\frac{1}{(2n-1)4^n}\cdot\binom{2n}{n}\right)^2\cdot x^{2n}. \end{aligned}$$

The last series has positive terms and by considering only the contribution provided by the $n=0$ term we immediately get:

$$L(a,b) \geq \pi(a+b).$$

□

The last inequality can be substantially stregthened as follows:

$$\frac{1}{4^n}\binom{2n}{n} = \prod_{k=1}^{n}\left(1-\frac{1}{2k}\right) = \frac{1}{2n+1}\prod_{k=1}^{n}\left(1+\frac{1}{2k}\right),$$

$$\left(\frac{1}{4^n}\binom{2n}{n}\right)^2 = \frac{1}{2n+1}\prod_{k=1}^{n}\left(1-\frac{1}{4k^2}\right) = \frac{2}{\pi(2n+1)}\prod_{k>n}\left(1-\frac{1}{4k^2}\right)^{-1},$$

$$\begin{aligned} L(a,b) &= \pi(a+b)\left(1+\sum_{n=0}^{+\infty}\left(\frac{1}{4^{n+1}(2n+1)}\binom{2n+2}{n+1}\right)^2\left(\frac{a-b}{a+b}\right)^{2n+2}\right) \\ &= \pi(a+b)\left(1+\frac{1}{4}\sum_{n=0}^{+\infty}\left(\frac{1}{4^n(n+1)}\binom{2n}{n}\right)^2\left(\frac{a-b}{a+b}\right)^{2n+2}\right), \end{aligned}$$

from which, by setting $\lambda = \frac{a-b}{a+b}$, we get:

Theorem 292 (Lindner, 1904)**.**

$$\frac{L(a,b)}{\pi(a+b)} \geq \left(1+\frac{\lambda^2}{8}\right)^2.$$

The following inequality holds, too:

$$\left(1+\frac{\lambda^2}{8}\right)^3 \geq \frac{(1+\lambda)^{3/2}+(1-\lambda)^{3/2}}{2}.$$

Both sides are analytic real functions on $(-1,1)$ and the associated Taylor expansions at the origin have non-negative coefficients. The RHS $f(\lambda)$ equals $1+\frac{3\lambda^2}{8}+\frac{3\lambda^4}{128}+O(\lambda^6)$, and for any $\lambda\in[0,1]$ we certainly have:

$$f(\lambda)\leq 1+\frac{3\lambda^2}{8}+\frac{3\lambda^4}{128}+\left(f(1)-\left(1+\frac{3}{8}+\frac{3}{128}\right)\right)\lambda^6 < 1+\frac{3\lambda^2}{8}+\frac{3\lambda^4}{128}+\frac{9\lambda^6}{512}.$$

On the other hand, if $\lambda^2\leq 1$:

$$1+\frac{3\lambda^2}{8}+\frac{3\lambda^4}{128}+\frac{9\lambda^6}{512}\leq 1+\frac{3\lambda^2}{8}+\frac{3\lambda^4}{64}+\frac{\lambda^6}{512}=\left(1+\frac{\lambda^2}{8}\right)^3,$$

since $\lambda^2<\frac{3}{2}$. The following result follows:

Theorem 293 (Muir, 1883)**.**

$$\frac{L(a,b)}{2\pi}\geq M_{\frac{3}{2}}(a,b).$$

The problem of finding a constant $\alpha\in\left(\frac{3}{2},2\right]$ such that

$$\frac{L(a,b)}{2\pi}\leq M_\alpha(a,b)$$

holds is equivalent to finding the infimum of the set of αs such that:

$$J(\lambda,\alpha)=\left(\frac{(1-\lambda)^\alpha+(1+\lambda)^\alpha}{2}\right)^{1/\alpha}\geq B(\lambda)=\sum_{n=0}^{+\infty}\left(\frac{1}{(2n-1)4^n}\binom{2n}{n}\right)^2\lambda^{2n}$$
$$=\frac{1}{\pi}\int_0^\pi\sqrt{1+\lambda^2+2\lambda\cos(2\theta)}\,d\theta$$

for any $\lambda\in[0,1]$.
We may notice that a necessary condition for the inequality to hold at $\lambda=1$ is:

$$\alpha\geq\alpha_0=\frac{\log 2}{\log(\pi/2)}=1.534928535\ldots,$$

since $B(1)=\frac{4}{\pi}$. By setting

$$E(m)=\int_0^{\pi/2}\sqrt{1-m\sin^2\theta}\,d\theta$$

we have:

$$B(\lambda)=2\,\frac{\lambda+1}{\pi}\,E\left(\frac{4\lambda}{(\lambda+1)^2}\right)=2\,\frac{1-\lambda}{\pi}\,E\left(-\frac{4\lambda}{(1-\lambda)^2}\right).$$

Therefore the inequality $J(\lambda,\alpha)\geq B(\lambda)$ is equivalent, up to the substitution $\frac{1-\lambda}{1+\lambda}=t$, to the inequality:

$$M_\alpha(1,t)\geq\frac{2}{\pi}E(1-t^2)$$

or the inequality:

$$g(z)=\frac{1+z}{2}\geq f_\alpha(z)=\left(\frac{2}{\pi}E(1-z^{2/\alpha})\right)^\alpha.$$

We remark that the function

$$B(\lambda)=\sum_{n=0}^{+\infty}\left(\frac{1}{(2n-1)4^n}\binom{2n}{n}\right)^2\lambda^{2n}=\frac{1}{\pi}\int_0^\pi\sqrt{1+\lambda^2+2\lambda\cos(2\theta)}\,d\theta$$

is a solution of the differential equation:

$$B = \left(\lambda + \frac{1}{\lambda}\right)\frac{dB}{d\lambda} + \left(1-\lambda^2\right)\frac{d^2B}{d\lambda^2}.$$

Due to the non-negativity of the coefficients in the Taylor expansion of $B(\lambda)$, we have $\frac{d^2B}{d\lambda^2} \geq 0$, from which it follows that $B(\lambda)$ (together with its derivatives) is a convex function and

$$\frac{B'}{B}(\lambda) \leq \frac{\lambda}{\lambda^2+1}$$

holds, from which:

$$B(\lambda) \leq \sqrt{1+\lambda^2}.$$

Since, additionally,

$$\frac{B'}{B}(\lambda) = \frac{\lambda}{(\lambda^2+1) + \frac{1}{1+\lambda\,\frac{B''}{B'}(\lambda)}},$$

we have:

$$\frac{B'}{B}(\lambda) \leq \frac{\lambda}{\lambda^2+2},$$

from which it follows:

$$B(\lambda) \geq \sqrt{1+\frac{\lambda^2}{2}}.$$

Summarizing:

$$\sqrt{1+\frac{\lambda^2}{2}} \leq B(\lambda) \leq \sqrt{1+\lambda^2}.$$

From simple algebraic manipulations we derive:

Theorem 294.

$$\frac{\pi}{4}\sqrt{\frac{3}{2}(2-m)+\sqrt{1-m}} \leq E(m) \leq \frac{\pi}{2}\sqrt{1-\frac{m}{2}}.$$

A recent and quite deep result, providing very tight bounds for the ellipse perimeter, is the following one:

Theorem 295 (Alzer,Qiu, 2004)**.**

$$\frac{L(a,b)}{2\pi} \leq M_{\frac{\log 2}{\log(\pi/2)}}(a,b)$$

where the constant appearing in the RHS cannot be replaced by any smaller number.

The proof exceeds the scope of these notes.

Exercise 296. Prove the following identity:

$$I = \int_0^{2\pi}\iint_{[0,1]^2} xy\sqrt{x^2+y^2-2xy\cos\theta}\,dx\,dy\,d\theta = \frac{64}{45}.$$

Proof. By symmetry, we have:

$$I = \int_0^{2\pi} \iint_{[0,1]^2} xy\sqrt{x^2+y^2-2xy\cos\theta}\,dx\,dy\,d\theta = 2\int_0^{2\pi}\iint_D xy\sqrt{x^2+y^2-2xy\cos\theta}\,dx\,dy\,d\theta$$

where $D = \{(x,y)\in[0,1]^2 : 0\leq y\leq x\}$. It follows that:

$$I = 2\int_0^{2\pi}\int_0^1 x^4 \int_0^1 t\sqrt{t-e^{i\theta}}\sqrt{t-e^{-i\theta}}\,dt\,dx\,d\theta$$

from which we get:

$$I = \frac{4\pi}{5}\int_0^1 \lambda\,B(\lambda)\,d\lambda$$

where:

$$B(\lambda) = \sum_{n\geq 0}\left(\frac{\binom{2n}{n}}{(2n-1)4^n}\right)^2\lambda^{2n} = \frac{1}{2\pi}\int_0^{2\pi}\sqrt{\lambda^2+1-2\lambda\cos\theta}\,d\theta = {}_2F_1\left(-\frac{1}{2},-\frac{1}{2};1;\lambda^2\right)$$

is a solution of the ordinary differential equation:

$$\lambda\,B = (\lambda^2+1)B' + (\lambda-\lambda^3)B''.$$

Since $B(0)=B'(0)=0$ and $B(1)=\frac{4}{\pi}=2\cdot B'(1)$, by the integration by parts formula we get:

$$\begin{aligned}\int_0^1 \lambda\,B(\lambda)\,d\lambda &= \int_0^1 (\lambda^2+1)B'(\lambda)\,d\lambda + \int_0^1(\lambda-\lambda^3)B''(\lambda)\,d\lambda\\ &= 2\cdot B(1) - 2\int_0^1 \lambda\,B(\lambda)\,d\lambda - \int_0^1 (1-3\lambda^2)B'(\lambda)\,d\lambda\end{aligned}$$

hence:

$$\begin{aligned}3\int_0^1 \lambda\,B(\lambda)\,d\lambda &= B(1) + 3\int_0^1 \lambda^2\,B'(\lambda)\,d\lambda\\ &= 4\,B(1) - 6\int_0^1 \lambda\,B(\lambda)\,d\lambda\end{aligned}$$

and finally:

$$\int_0^1 \lambda\,B(\lambda)\,d\lambda = \frac{4}{9}\,B(1) = \frac{16}{9\pi},$$

proving the claim. We may notice that by setting:

$$A_k \stackrel{\text{def}}{=} \int_0^1 x^{2k+1}B(x)\,dx$$

we have $A_0 = \frac{16}{9\pi}$ as proved and:

$$(2k+3)^2 A_k = \frac{16}{\pi} + (2k)^2 A_{k-1}.$$

□

We may notice the previous exercise is related to the celebrated problem of finding the average distance between two random points the region $\{(x,y)\in\mathbb{R}^2 : x^2+y^2\leq 1\}$, picked according to a uniform probability distribution. We may also notice that, in terms of hypergeometric functions, the problem is equivalent to computing

$$\sum_{n\geq 0}\frac{\binom{2n}{n}^2}{16^n(2n-1)^2(n+1)} = {}_2F_1\left(-\tfrac{1}{2},-\tfrac{1}{2};2;1\right)$$

which, by partial fraction decomposition, boils down to computing the following series:

$$S_1 = \sum_{n\geq 0}\frac{\binom{2n}{n}^2}{16^n(n+1)},\qquad S_2 = \sum_{n\geq 0}\frac{\binom{2n}{n}^2}{16^n(2n-1)},\qquad S_3 = \sum_{n\geq 0}\frac{\binom{2n}{n}^2}{16^n(2n-1)^2}$$

or, by reindexing an partial fraction decomposition, to computing the following series:

$$T_1 = \sum_{n\geq 0} \frac{\binom{2n}{n}^2}{16^n(n+1)}, \qquad T_2 = \sum_{n\geq 0} \frac{\binom{2n}{n}^2}{16^n(n+2)}, \qquad T_3 = \sum_{n\geq 0} \frac{\binom{2n}{n}^2}{16^n(n+1)^2}$$

which are given by:

$$T_1 = \frac{2}{\pi}\int_0^1 K(m)\,dm, \qquad T_2 = \frac{2}{\pi}\int_0^1 m\,K(m)\,dm, \qquad T_3 = -\frac{2}{\pi}\int_0^1 K(m)\log(m)\,dm.$$

Since both $K(m)$ and $\log(m)$ have a simple Fourier-Legendre expansion over $(0,1)$ in terms of shifted Legendre polynomials, to check that $T_1 = \frac{4}{\pi}$, $T_2 = \frac{20}{9\pi}$ and $T_3 = -4 + \frac{16}{\pi}$ is actually pretty simple.

Exercise 297. Given a square centered at O, prove that among the ellipses centered at O and tangent to square sides, the circle has the maximum perimeter.

We now consider the following framework: given two non-negative real numbers a_0, b_0, it is possible to define two sequences by the following recurrence relations:

$$a_{n+1} = \text{AM}(a_n, b_n) = \frac{a_n + b_n}{2}, \qquad b_{n+1} = \text{GM}(a_n, b_n) = \sqrt{a_n b_n}.$$

If we assume $b_0 \leq a_0$, due to the AM-GM inequality we have:

$$b_n \leq b_{n+1} \leq a_{n+1} \leq a_n,$$

and by setting $c_n = (a_n - b_n)$ we also have:

$$0 \leq c_{n+1} = \frac{c_n^2}{2(\sqrt{a_n} + \sqrt{b_n})^2},$$

hence the sequences $\{a_n\}_{n\geq 0}, \{b_n\}_{n\geq 0}$ are monotonic sequences and they are rapidly (quadratically) convergent to a common limit, which we call **arithmo-geometric mean** and denote through $\text{AGM}(a_0, b_0)$. The term *mean* is well-used in this context since

$$\min(a,b) \leq \text{AGM}(a,b) = \text{AGM}(b,a) \leq \max(a,b), \qquad \forall \lambda > 0,\ \text{AGM}(\lambda a, \lambda b) = \lambda \cdot \text{AGM}(a,b).$$

Due to the very definition of arithmo-geometric mean, we have:

$$\text{AGM}(a,b) = \text{AGM}\left(\frac{a+b}{2}, \sqrt{ab}\right), \qquad \text{AGM}(1,x) = \frac{1+x}{2}\,\text{AGM}\left(1, \frac{2\sqrt{x}}{1+x}\right).$$

The numerical evaluation of an arithmo-geometric mean is simple due to rapid convergence. An unexpected (at least at first sight) scenario arises from the following remark: the AGM mean has a simple integral representation, hence there are efficient numerical algorithms for the evaluation of a wide class of integrals and many recurrent mathematical constants.

Definition 298. We define the complete elliptic integrals of the first and second kind as:

$$K(k) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{1 - k^2 \sin^2\theta}}, \qquad E(k) = \int_0^{\pi/2} \sqrt{1 - k^2\sin^2\theta}\,d\theta.$$

For any $k \in (0,1)$, usually known as **modulus** of the involved elliptic integral, we further define the **complementary modulus** k' as $\sqrt{1-k^2}$, then we introduce the conventional notation $K'(k) = K(k')$, $E'(k) = E(k')$.

By exploiting the suitable substitutions and Taylor expansions we have that:

$$K(k) = \int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}} = \int_0^{+\infty} \frac{du}{\sqrt{(1+u^2)(1+(1-k^2)u^2)}} = \frac{\pi}{2}\sum_{n\geq 0}\left(\frac{1}{4^n}\binom{2n}{n}\right)^2 k^{2n},$$

$$E(k) = \int_0^1 \sqrt{\frac{1-k^2t^2}{1-t^2}}\,dt = \int_0^{+\infty}\sqrt{(1-k^2)+\frac{k^2}{1+u^2}}\,\frac{du}{1+u^2} = \frac{\pi}{2}\left(1-\sum_{n\geq 0}\left(\frac{1}{4^n}\binom{2n}{n}\right)^2\frac{k^{2n}}{2n-1}\right).$$

Additionally, by differentiation under the integral sign we get the following identities:

Theorem 299 (Legendre)**.**

$$\frac{dE}{dk} = \frac{E-K}{k},\qquad \frac{dK}{dk} = \frac{E-k'^2K}{k(k')^2},$$

$$K(k)E'(k)+K'(k)E(k)-K(k)K'(k) = \frac{\pi}{2}$$

where both $K'(k)$ and $K(k)$ turn out to be solutions of the differential equation:

$$(k^3-k)\frac{d^2y}{dk^2}+(3k^2-1)\frac{dy}{dk}+ky = 0$$

with a strong analogy with the differential equation previously seen for $B(\lambda)$.

Theorem 300 (Gauss)**.**

$$\frac{\pi}{2\ \mathrm{AGM}(a,b)} = \int_0^{+\infty}\frac{dt}{\sqrt{(a^2+t^2)(b^2+t^2)}} = \frac{1}{a}K\left(\sqrt{1-\frac{b^2}{a^2}}\right) = \frac{1}{b}K\left(\sqrt{1-\frac{a^2}{b^2}}\right).$$

Proof. For any couple (a,b) of positive real numbers, let us define

$$T(a,b) = \int_0^{+\infty}\frac{dt}{\sqrt{(a^2+t^2)(b^2+t^2)}}.$$

We may notice that Lagrange's identity implies:

$$(a^2+t^2)(b^2+t^2) = (at+bt)^2+(t^2-ab)^2,$$

hence by setting $u = \frac{1}{2}\left(t-\frac{ab}{t}\right)$ we get that, as t ranges from 0 to $+\infty$, u ranges from $-\infty$ to $+\infty$. Additionally $du = \frac{1}{2}\left(1+\frac{ab}{t^2}\right)dt$ implies:

$$T(a,b) = \int_0^{+\infty}\frac{1}{\sqrt{(a+b)^2+\left(t-\frac{ab}{t}\right)^2}}\cdot\frac{dt}{t} = \int_{-\infty}^{+\infty}\frac{du}{\sqrt{((a+b)^2+4u^2)\,(ab+u^2)}} = T(\mathrm{AM}(a,b),\mathrm{GM}(ab)).$$

By repeating the same trick multiple times, we reach the wanted integral representation for the arithmo-geometric mean:

$$T(a,b) = T(\mathrm{AGM}(a,b),\mathrm{AGM}(a,b)) = \int_0^{+\infty}\frac{dt}{t^2+\mathrm{AGM}(a,b)^2} = \frac{\pi}{2\ \mathrm{AGM}(a,b)}.$$

□

Exercise 301 (So many ks)**.** Prove that:

$$\int_0^1 4k^2\,K(k)\,dk = 1+2\sum_{k\geq 0}\frac{(-1)^k}{(2k+1)^2}.$$

Proof. By exploiting the integral representation for $K(k)$ and Fubini's Theorem (allowing us to switch the involved integrals):

$$\int_0^1 4k^2\,K(k)\,dk = \int_0^{\pi/2}\left(\frac{\theta}{\sin\theta}-\cos\theta\right)\frac{2\,d\theta}{\sin^2(\theta)} = \int_0^{\pi/2}\frac{2\theta-\sin(2\theta)}{\sin^3(\theta)}\,d\theta.$$

At this point we may finish by invoking the residue Theorem, a Fourier-Chebyshev or a Fourier-Legendre expansion. □

Exercise 302. Prove that both $\Gamma\left(\frac{1}{4}\right)$ and $\Gamma\left(\frac{1}{6}\right)$ can be written in terms of π and arithmo-geometric means of algebraic numbers over $\mathbb{Q}$.

Proof. We may consider first the following integral:

$$\begin{aligned}
\frac{\pi}{2\,\text{AGM}(1,\sqrt{2})} = \frac{1}{\sqrt{2}}\,K\!\left(\frac{1}{\sqrt{2}}\right) &= \int_0^{+\infty}\frac{dt}{\sqrt{(1+t^2)(1+2t^2)}}\\
(t\mapsto\sinh u) &= \int_0^{+\infty}\frac{du}{\sqrt{\cosh(2u)}}\\
(u\mapsto-\log z) &= \sqrt{2}\int_0^1\frac{dz}{\sqrt{z^4+1}}\\
(\text{simmetry}) &= \frac{1}{\sqrt{2}}\int_0^{+\infty}\frac{dz}{\sqrt{1+z^4}}\\
(\text{Beta}) &= \frac{1}{4\sqrt{2}}\,B\left(\tfrac{1}{4},\tfrac{1}{4}\right)\\
&= \frac{1}{4\sqrt{2\pi}}\,\Gamma\left(\tfrac{1}{4}\right)^2.
\end{aligned}$$

This chain of equalities tells us that:

$$\Gamma\left(\frac{1}{4}\right)^2 = \frac{(2\pi)^{3/2}}{\text{AGM}(1,\sqrt{2})} = \frac{\sqrt{\pi}}{4}\,K\!\left(\frac{1}{\sqrt{2}}\right).$$

In a similar way, by considering the integral:

$$\begin{aligned}
\int_1^{+\infty}\frac{dx}{\sqrt{x^3-1}} &= \int_0^{+\infty}\frac{dx}{\sqrt{x^3+3x^2+3x}} = \int_0^{+\infty}\frac{2\,dz}{\sqrt{z^4+3z^2+3}}\\
&= \frac{\pi}{\text{AGM}\left(\sqrt{\frac{3+i\sqrt{3}}{2}},\sqrt{\frac{3-i\sqrt{3}}{2}}\right)} = \frac{\pi}{\text{AGM}\left(\frac{1}{2}\sqrt{3+2\sqrt{3}},3^{1/4}\right)}.
\end{aligned}$$

By enforcing the substitution $x=\frac{1}{t}$ and exploiting Euler's Beta function we get:

$$\int_1^{+\infty}\frac{dx}{\sqrt{x^3-1}} = \int_0^1\frac{dt}{\sqrt{t-t^4}} = \int_0^1\frac{2\,du}{\sqrt{1-u^6}} = \frac{1}{3}\,B\left(\frac{1}{6},\frac{1}{2}\right).$$

Due to Legendre's duplication formula we have:

$$\Gamma\left(\frac{1}{6}\right) = \frac{2^{14/9}\,3^{1/3}\,\pi^{5/6}}{\text{AGM}(1+\sqrt{3},\sqrt{8})^{2/3}}.$$

□

Exercise 303. By exploiting the previous identities and Legendre's identity for complete elliptic integrals, prove that:

$$\int_{0}^{\pi/2}\sqrt{1+\sin^2(\theta)}\,d\theta = \frac{1}{4\sqrt{2\pi}}\left[4\,\Gamma^2\left(\tfrac{3}{4}\right)+\Gamma^2\left(\tfrac{1}{4}\right)\right].$$

The interplay between elliptic integrals and the arithmo-geometric mean, together with Legendre's identity, led **Brent** and **Salamin**, in 1976, to develop the following algorithm for the numerical evaluation of π:

- Initialize $x_0=\sqrt{2}$, $\pi_0 = 2+\sqrt{2}$, $y_1 = 2^{1/4}$;
- Set:
$$x_{n+1} = \frac{1}{2}\left(\sqrt{x_n}+\frac{1}{\sqrt{x_n}}\right),\qquad y_{n+1} = \frac{y_n\sqrt{x_n}+1/\sqrt{x_n}}{y_n+1},\qquad \pi_n = \pi_{n-1}\frac{x_n+1}{y_n+1}$$
- The sequence $\{\pi_n\}_{n\geq 0}$ decreases towards π and it is quadratically convergent:
$$\forall n\geq 2,\qquad \pi_n - \pi < 10^{-2^{n+1}}.$$

Exercise 304. Prove that:

$$\int_{0}^{1}\frac{K(k)}{\sqrt{1-k^2}}\,dk = \frac{\pi^2}{4}\sum_{n\geq 0}\left(\frac{1}{4^n}\binom{2n}{n}\right)^3 = K^2\left(\frac{1}{\sqrt{2}}\right).$$

12 Bessel functions and the Gauss circle problem

Bessel functions naturally arise in the solution of the problem $\Delta u = f$ with certain boundary conditions, and they are extremely relevant in Harmonic Analysis. We may introduce them by studying the Fourier sine series of the arcsin function.

$$\begin{aligned}\int_{-1}^{1}\arcsin(x)\sin(\pi n x)\,dx &= \text{Im}\int_{-\pi/2}^{\pi/2} e^{i\pi n\sin z} z\cos(z)\,dz\\ &= \sum_{m\geq 0}\frac{(-1)^m(\pi n)^{2m+1}}{(2m+1)!}\int_{-\pi/2}^{\pi/2} z\cos(z)\,(\sin z)^{2m+1}\,dz\\ &\overset{\text{IBP}}{=} \sum_{m\geq 0}\frac{(-1)^m(\pi n)^{2m+1}}{(2m+1)!}\left(\left[\frac{z\sin(z)^{2m+2}}{2m+2}\right]_{-\pi/2}^{\pi/2} - \int_{-\pi/2}^{\pi/2}(\sin z)^{2m+2}\,dz\right)\\ &= \sum_{m\geq 0}\frac{(-1)^m(\pi n)^{2m+1}}{(2m+1)!}\cdot\frac{\pi}{2m+2}\left[1-\frac{\binom{2m+2}{m+1}}{4^{m+1}}\right]\\ &= \frac{1-\cos(n\pi)}{n} - \pi\sum_{m\geq 0}\frac{(-1)^m(\pi n)^{2m+1}}{4^{m+1}(m+1)!^2}\\ &= \frac{1}{n}\left((-1)^{n+1} + \sum_{m\geq 0}\frac{(-1)^m(\pi n)^{2m}}{4^m m!^2}\right)\end{aligned}$$

hence by defining the following even, entire function

$$J_0(z) = \sum_{n\geq 0}\frac{(-1)^n z^{2n}}{4^n n!^2}$$

we have

$$\arcsin(z) \overset{L^2(-1,1)}{=} \sum_{m\geq 1}\frac{(-1)^{m+1}+J_0(\pi m)}{m}\sin(\pi m z),$$

$$\arcsin(z) - \frac{\pi}{2}z \overset{L^2(-1,1)}{=} \sum_{m\geq 1}\frac{J_0(\pi m)\sin(\pi m z)}{m}.$$

It is straightfoward to check from the power series that J_0 is the unique solution of the following differential equation of order 2:

$$z\,f''(z) + f'(z) + z\,f(z) = 0,\quad f(0)=1, f'(0)=0$$

which allows to state many interesting facts about the behaviour of J_0 far from the origin.
From the power series definition it also follows that

$$(\mathcal{L}J_0)(s) = \frac{1}{\sqrt{1+s^2}}$$

and we have already seen that Vandermonde's identity leads to

$$(\mathcal{L}J_0^2)(s) = \frac{2}{\pi s}K\left(-\frac{4}{s^2}\right) = \frac{1}{\text{AGM}(s,\sqrt{s^2+4})}$$

where the RHS behaves like $-\frac{\log(s)}{\pi}$ in a right neighbourhood of the origin and like $\frac{1}{s}$ in a left neighbourhood of $+\infty$. Anyway the properties of J_0 are better understood by introducing the **Bessel functions of the first kind** J_n and their generating function: for such a purpose, we study the Fourier cosine series of $e^{iz\cos\theta}$, which will lead to the **Jacobi-Anger** expansion.

$$\begin{aligned}\int_0^{2\pi} e^{iz\cos\theta}\cos(m\theta)\,d\theta &= \sum_{n\geq 0}\frac{i^n z^n}{n!}\int_0^{2\pi}(\cos\theta)^n\cos(m\theta)\,d\theta\\ &= \sum_{n\geq 0}\frac{i^n z^n}{2^n n!}\int_0^{2\pi}(e^{i\theta}+e^{-i\theta})^n\cos(m\theta)\,d\theta\\ &= \pi\sum_{n=m+2k}\frac{\pi i^n z^n}{2^n n!}\binom{m+2k}{k}\\ &= \pi i^m (z/2)^m\sum_{k\geq 0}\frac{(-1)^k z^{2k}}{4^k k!(m+k)!}\binom{m+2k}{k}\end{aligned}$$

hence by defining

$$J_m(z)=\sum_{n\geq 0}\frac{(-1)^n (z/2)^{2n+m}}{n!(n+m)!}$$

we have

Theorem 305 (Jacobi-Anger expansion)**.**

$$e^{iz\cos\theta}=J_0(z)+2\sum_{m\geq 1} i^m J_m(z)\cos(m\theta).$$

By considering the real or imaginary part of both sides, we get that Bessel functions of the first kind provide the coefficients of the Fourier series of $\sin(z\cos\theta)$ and $\cos(z\cos\theta)$. By applying Cauchy's integral formula to the Jacobi-Anger expansion we find the following integral representation:

$$J_n(z)=\frac{1}{\pi}\int_0^{\pi}\cos(z\sin\theta-n\theta)\,d\theta$$

with the following equivalent form:

$$J_n(z)=\frac{1}{2\pi i}\oint_{\|z\|=\varepsilon}\exp\left[\frac{z}{2}\left(t-\frac{1}{t}\right)\right]\frac{dt}{t^{n+1}}.$$

Bessel functions share with the families of orthogonal polynomials many interesting properties. The recurrence relations

$$J_0'(z)=-J_1(z),\qquad J_n'(z)=\frac{1}{2}\left(J_{n-1}(z)-J_{n+1}(z)\right),$$

$$J_{n-1}(z)+J_{n+1}(z)=\frac{2n}{z}J_n(z),\qquad \frac{d}{dx}\left(x^m J_m(x)\right)=x^m J_{m-1}(x)$$

are straightforward to prove through the series definition or the integral representation. In the same way the entire function $J_n(z)$ can be checked to be a solution of **Bessel's differential equation**

$$x^2 f''(x)+xf'(x)+(x^2-n^2)f(x)=0$$

whose structure recalls the structure of Legendre's differential equation. The identities

$$1=J_0(x)^2+2\sum_{k\geq 1}J_k(x)^2,\qquad 1=J_0(x)+2\sum_{k\geq 1}J_{2k}(x),\qquad J_n(2z)=\sum_{k=0}^{n}J_k(z)J_{n-k}(z)+2\sum_{k\geq 1}(-1)^k J_k(z)J_{n+k}(z)$$

can be proved through the orthogonality relations in $L^2(-\pi,\pi)$ applied to the Jacobi-Anger expansion. The integral representation or Bessel's differential equation allow to define $J_\nu(x)$ also for non-integer values of ν. For $n\in\mathbb{N}$, the function $J_{-n}(x)$ is defined as $(-1)^n J_n(x)$. According to this convention, Bessel functions of the first kind have a very simple addition formula of the convolution-type:

$$J_n(y+z) = \sum_{m\in\mathbb{Z}} J_m(y) J_{n-m}(z).$$

The substitution $g(x) = \sqrt{x} f(x)$ turns the differential equation defining J_0, namely $xf''(x) + f'(x) + xf(x) = 0$, into:

$$g''(x) + \left(1 + \frac{1}{4x^2}\right) g(x) = 0.$$

Since the term $\frac{1}{4x^2}$ is non-negative and negligible for large values of x, it is reasonable to assume that $g(x) = A(x)\cos(x) + B(x)\sin(x)$ is such that $A(x) \to A$ and $B(x) \to B$ for $x \to +\infty$. In order to find the value of these constants, we may notice that

$$\cos(x)J_0(x) - \sin(x)J_0'(x) = \frac{2\sqrt{2}}{\pi\sqrt{x}} \int_0^{\sqrt{2x}} \cos(t^2)\sqrt{1 - \frac{t^2}{2x}}\, dt,$$

$$\sin(x)J_0(x) + \cos(x)J_0'(x) = \frac{2\sqrt{2}}{\pi\sqrt{x}} \int_0^{\sqrt{2x}} \sin(t^2)\sqrt{1 - \frac{t^2}{2x}}\, dt,$$

hence by the dominated convergence theorem and Fresnel integrals we have $A = B = \frac{1}{\sqrt{\pi}}$.
By studying the action of the operator $\frac{d^2}{dx^2} + \left(1 + \frac{1}{4x^2}\right)$ on the conjectural expression

$$\sqrt{\pi x}\, J_0(x) = \left(A + \frac{A'}{x} + \frac{A''}{x^2} + \ldots\right)\cos(x) + \left(B + \frac{B'}{x} + \frac{B''}{x^2} + \ldots\right)\sin(x)$$

Poisson derived the formal series

$$\left(1 - \frac{1}{8x} - \frac{9}{2\cdot 8^2 x^2} + \frac{9\cdot 25}{2\cdot 3\cdot 8^3 x^3} + \ldots\right)\cos(x) + \left(1 + \frac{1}{8x} - \frac{9}{2\cdot 8^2 x^2} - \frac{9\cdot 25}{2\cdot 3\cdot 8^3 x^3} + \ldots\right)\sin(x)$$

which is not convergent, but whose truncations allow to devise arbitrarily accurate approximations of $J_0(x)$ as $x \to +\infty$. As a side note, we may notice that over the real line the Laplace transform of $J_0(x)$, namely $\frac{1}{\sqrt{1+s^2}}$, and the Laplace transform of $\frac{\sin(x)+\cos(x)}{\sqrt{\pi x}}$, namely $\sqrt{\frac{1}{1+s^2} + \frac{1}{\sqrt{1+s^2}}}$, differ by a term bounded between 0 and $\sqrt{2}-1$ and behaving like $\frac{1}{\sqrt{|s|}}$ for $|s| \to +\infty$.

Theorem 306 (RH for dummies)**.** All the zeroes of the entire functions $J_n(z), J_n'(z)$ in $\mathbb{C}\setminus\{0\}$ are real and simple. The average distance between a zero and the next one approaches π.

Proof. Assuming that for some $z \in \mathbb{C}$ we have $J_n(z) = 0$ and $J_n'(z) = 0$, Bessel's differential equation implies $J_n''(z) = 0$, then $J_n^{(m)}(z) = 0$ by induction, hence $J_n(z) = 0$ by the principle of analytic continuation, which is a contradiction. Since the coefficients of the Maclaurin series of $J_n(z)$ are real, assuming $J_n(z_0) = 0$ with $z_0 \notin \mathbb{R}$ implies that $\overline{z_0}$ is a zero, too. The same holds for J_n'. Now we may consider the identity

$$(a^2 - b^2)\int_0^z tJ_n(at)J_n(bt)\, dt = z\left[bJ_n(az)J_n'(bz) - aJ_n'(az)J_n(bz)\right]$$

which follows from Bessel's differential equation. By considering $a = z_0, b = \overline{z_0}$ and $z = 1$ we get

$$0 = (z_0^2 - \overline{z_0}^2)\int_0^1 t\left|J_n(z_0 t)\right|^2\, dt$$

which can only be true for $z_0 = iy$ with $y \in \mathbb{R}\setminus\{0\}$. In such a case, however,

$$\sum_{m\geq 0} \frac{1}{m!(m+n)!}\left(\frac{y}{2}\right)^{2m}$$

is clearly positive. We may invoke the Gauss-Lucas theorem (the zeroes of $f'(z)$ lies in the convex hull of the zeroes of $f(z)$) to deduce that all the zeroes of $J_n'(z)$ in $\mathbb{C}\setminus\{0\}$ are real and simple. The asymptotic formula

$$J_n(z)\sim\sqrt{\frac{2}{\pi z}}\cos\left(z-\frac{\pi n}{2}-\frac{\pi}{4}\right)\quad\text{for } |z|\to+\infty \text{ with } |\arg z|\leq\frac{\pi}{2}-\varepsilon$$

finishes the proof. □

Exercise 307. Show that $J_{1/2}(x)$ is an elementary function.

Proof. We just have to notice that the Bessel differential equation

$$x^2y''+xy'+\left(x^2-\frac{1}{4}\right)y=0,$$

for which $y=J_{\frac{1}{2}}$ is a solution, is mapped into a well-known homogeneous differential equation with constant coefficients by the substitution $y(x)=\frac{f(x)}{\sqrt{x}}$. □

Exercise 308 (Exploiting the Fourier transform)**.** Compute the value of the integral $\int_0^{+\infty}\left(\frac{J_1(x)}{x}\right)^2\,dx$.

Proof. Since J_1 is a solution of the Bessel differential equation:

$$x^2f''+xf'+x^2f=f$$

by exploiting integration by parts we have that:

$$\int_0^{+\infty}\left(J_1(x)^2+J_1(x)\,J_1''(x)\right)\,dx=\frac{1}{2}\int_0^{+\infty}\left(\frac{J_1(x)}{x}\right)^2\,dx$$

so we just need to recall that the Fourier transform of $\frac{J_1(x)}{x}$ is given by:

$$\mathscr{F}\left(\frac{J_1(x)}{x}\right)(t)=\sqrt{\frac{2}{\pi}}\sqrt{1-t^2}\cdot\mathbb{1}_{(-1,1)}(t)$$

to be able to state:

$$\int_0^{+\infty}\left(J_1(x)^2+J_1(x)\,J_1''(x)\right)\,dx=\frac{1}{\pi}\int_0^1\left(1-t^2\right)\,dt=\frac{2}{3\pi}$$

as a consequence of Parseval's theorem. □

Exploiting the Laplace transform.

$$\begin{aligned}\int_0^{+\infty}\frac{J_1(x)}{x^\alpha}\,dx &= \frac{1}{\Gamma(\alpha)}\int_0^{+\infty}\frac{s^{\alpha-1}}{\sqrt{1+s^2}(s+\sqrt{1+s^2})}\,ds\\ &= \int_0^{\pi/2}\frac{(\tan\theta)^{\alpha-1}}{1+\sin\theta}\,d\theta=\int_0^{\pi/2}\left[(\sin\theta)^{\alpha-1}(\cos\theta)^{-1-\alpha}-(\sin\theta)^{\alpha}(\cos\theta)^{-1-\alpha}\right]\,d\theta\\ &= \frac{\pi\alpha}{2^{\alpha+1}\Gamma\left(1+\frac{\alpha}{2}\right)^2\sin\left(\frac{\pi\alpha}{2}\right)}\end{aligned}$$

holds for any $\alpha\in\left(-\frac{1}{2},2\right)$, as soon as the original integral is intended as an improper Riemann integral. Similar manipulations driven by the Laplace transform of $J_1(x)^2$ lead to

$$\int_0^{+\infty}\frac{J_1(x)^2}{x^\alpha}\,dx=\frac{\sqrt{\pi}(1-\alpha)\Gamma\left(\frac{\alpha}{2}\right)}{2(1+\alpha)\Gamma\left(\frac{1+\alpha}{2}\right)^3\cos\left(\frac{\pi\alpha}{2}\right)}$$

for any $\alpha \in (0,1)$. We also have relations with Fourier-Legendre series and hypergeometric functions. For instance:

$$\begin{aligned}
\int_0^{+\infty} \frac{J_0(x)^2}{\sqrt{x}}\,dx &= \int_0^{+\infty} \frac{2\,K\left(-\frac{4}{s^2}\right)}{(\pi s)^{3/2}}\,ds = \int_0^{+\infty} \frac{ds}{\sqrt{\pi s}\,\mathrm{AGM}(s,\sqrt{s^2+4})}\\
&= \int_0^{+\infty} \frac{ds}{\sqrt{2\pi s}\,\mathrm{AGM}(s,\sqrt{s^2+1})} = \int_0^{+\infty} \frac{ds}{\sqrt{2\pi s(s^2+1)}\,\mathrm{AGM}(1,\sqrt{1-\frac{1}{s^2+1}})}\\
&= \int_0^{+\infty} \frac{\sqrt{2}\,K\left(\frac{1}{s^2+1}\right)\,ds}{\sqrt{\pi^3 s(s^2+1)}} = \int_0^{+\infty} \frac{K\left(\frac{1}{s+1}\right)\,ds}{\sqrt{2\pi^3 s(s+1)}\sqrt{s}} = \frac{1}{\sqrt{2\pi^3}}\int_0^1 \frac{K(s)\,ds}{(s(1-s))^{3/4}}\\
&= \frac{\Gamma\left(\frac{1}{4}\right)}{\sqrt{8\pi^3}}\sum_{n\geq 0}\frac{\Gamma\left(n+\frac{1}{2}\right)\Gamma\left(n+\frac{1}{4}\right)}{\Gamma(n+1)^2} = \frac{\Gamma\left(\frac{1}{4}\right)^2}{\sqrt{8\pi^2}}\cdot {}_2F_1\left(\tfrac{1}{4},\tfrac{1}{2};1;1\right)\\
&= \frac{1}{4\pi^{5/2}}\Gamma\left(\tfrac{1}{4}\right)^4.
\end{aligned}$$

The last integrals are just special instances of the Sonine-Schafheitlin integral formula.
Once the *modified* Bessel functions of the first kind are defined through

$$I_m(x) = \sum_{n\geq 0}\frac{(x/2)^{2n+m}}{n!(n+m)!} = \frac{1}{i^m}\,J_m(ix)$$

the Laplace transform allows a simple evaluation of the integrals

$$I(\alpha,\beta,m) \stackrel{\text{def}}{=} \int_0^{+\infty} x^\alpha e^{-\beta x} I_m(x)\,dx$$

for $\beta > 1$ and $\alpha, m \in \mathbb{N}$. Indeed, from

$$\mathcal{L}(I_m(x))(s) = \frac{\mathbb{1}_{s>1}(s)}{\sqrt{s^2-1}}\left(s-\sqrt{s^2-1}\right)^m$$

we have:

$$\begin{aligned}
I(\alpha,\beta,m) &= \frac{d^\alpha}{ds^\alpha}\left[\frac{\left(s-\sqrt{s^2-1}\right)^m}{\sqrt{s^2-1}}\right]_{s=\beta}\\
&\stackrel{s\mapsto\frac{x^2+1}{2x}}{=} \frac{\alpha!}{2\pi i}\oint_{\|x-(\beta-\sqrt{\beta^2-1})\|=\varepsilon}\frac{x^{m-1}\,dx}{\left(\frac{x^2+1}{2x}-\beta\right)^{\alpha+1}}\\
&= \frac{\alpha!}{2^{\alpha+1}}\operatorname*{Res}_{z=0}\frac{\left(z+\beta-\sqrt{\beta^2-1}\right)^{m+\alpha}}{z^{\alpha+1}\left(z-2\sqrt{\beta^2-1}\right)^{\alpha+1}}\\
&= \frac{\alpha!(-1)^{\alpha+1}}{4^{\alpha+1}\sqrt{\beta^2-1}^{\alpha+1}}\cdot[x^{\alpha+2}]\frac{\left(x+\beta-\sqrt{\beta^2-1}\right)^{m+\alpha}}{\left(1-\frac{z}{2\sqrt{\beta^2-1}}\right)^{\alpha+1}}\\
&= \frac{\alpha!(-1)^{\alpha+1}}{8^{\alpha+1}(\beta^2-1)^{\alpha+1}}\sum_{k=0}^{\alpha+2}\binom{m+\alpha}{k}\binom{2\alpha+2-k}{\alpha}(\beta+\sqrt{\beta^2-1})^{m+\alpha-k}2^{k-1}\sqrt{\beta^2-1}^{k-1}.
\end{aligned}$$

Besides this characterization of the inner products against the elements of $\mathrm{Span}(x^\alpha e^{-\beta x} : \alpha\in\mathbb{N},\beta>0)$, modified Bessel functions of the first kind have a simple integral representation:

$$I_n(\alpha) = \frac{1}{\pi}\int_0^\pi \cos(nx)\,e^{\alpha\cos x}dx.$$

By the cosine addition formula and integration by parts we may easily get the following recurrence relation:

$$I_n = \frac{\alpha}{2n}\left(I_{n-1} - I_{n+1}\right).$$

From this identity, we have that the ratio between I_n and I_{n-1} is the continued fraction:

$$\frac{I_n}{I_{n-1}}(\alpha) = \frac{1}{\frac{2}{\alpha}\,n + \frac{1}{\frac{2}{\alpha}\,(n+1)+\ldots}}.$$

In particular

$$r(x) = \frac{I_1}{I_0}(x) = \frac{1}{\frac{2}{x} + \frac{1}{\frac{4}{x}+\frac{1}{\frac{6}{x}+\frac{1}{\ldots}}}} = \frac{x}{2 + \frac{x^2}{4+\frac{x^2}{6+\frac{x^2}{8+\ldots}}}}.$$

We also have

$$\frac{d}{d\alpha}\,I_n = \frac{1}{2}\left(I_{n-1} + I_{n+1}\right)$$

mimicking the recurrence relation for J_n'.

Corollary 309.

$$[0;1,2,3,4,5,6,7,8,9,\ldots] = \frac{I_1(2)}{I_0(2)} = \frac{\sum_{m\geq 0}\frac{1}{m!(m+1)!}}{\sum_{m\geq 0}\frac{1}{m!^2}} \approx 0.697774657964.$$

In general, if $A, D > 0$ we have

$$[A+D; A+2D, A+3D, A+4D,\ldots] = \frac{I_{A/D}\left(\frac{2}{D}\right)}{I_{1+A/D}\left(\frac{2}{D}\right)}.$$

The asymptotic behaviour of $I_0(x)$ for $x \to +\infty$ can be derived from Bessel's differential equation:

$$I_0(x) = \frac{e^x}{\sqrt{2\pi x}}\left[1 + \frac{1}{1!\cdot 8x} + \frac{9}{2!\cdot 8^2x^2} + \frac{9\cdot 25}{3!\cdot 8^3x^3} + \ldots\right]$$

the identity holds in the Poisson sense. If combined with the continued fraction representation for $\frac{I_{m+1}}{I_m}$, it leads to the asymptotic behavior of $I_m(x)$ as $x \to +\infty$, for any $m \in \mathbb{N}$. For a fixed $x \in \mathbb{R}^+$, the continued fraction representation immediately leads to the upper bound

$$I_n(x) \leq \frac{(x/2)^n}{n!} I_0(x).$$

If one is just interested in the first term of the asymptotic expansion of I_0, it can be recovered in a very simple way: for $s \to 1^+$,

$$(\mathcal{L}I_0)(s) = \frac{1}{\sqrt{s^2-1}} \sim \frac{1}{\sqrt{2}\sqrt{s-1}} = \mathcal{L}\left(\frac{1}{e^x\sqrt{2\pi x}}\right)(s).$$

Relations with the sine and cosine integrals. Since $J_0(x) = \sum_{m=0}^{+\infty} \frac{(-1)^m}{4^m\,m!^2} x^{2m}$ we have:

$$\begin{aligned}
\int_0^{\pi/2} J_0(x\cos\theta)\cos\theta\,d\theta &= \sum_{m=0}^{+\infty}\frac{(-1)^m x^{2m}}{4^m\,m!^2}\int_0^{\pi/2}(\cos\theta)^{2m+1}d\theta\\
&= \sum_{m=0}^{+\infty}\frac{(-1)^m x^{2m}}{4^m\,m!^2}\cdot\frac{4^m\,m!^2}{(2m+1)!}\\
&= \sum_{m=0}^{+\infty}\frac{(-1)^m x^{2m}}{(2m+1)!} = \frac{\sin x}{x}.
\end{aligned}$$

In a similar way $J_1(x)=\sum_{m=0}^{+\infty}\frac{(-1)^m}{2\cdot 4^m\,(m+1)\,m!^2}x^{2m+1}$ gives:

$$\begin{aligned}\int_0^{\pi/2}J_1(x\cos\theta)\,d\theta &= \sum_{m=0}^{+\infty}\frac{(-1)^m\,x^{2m+1}}{2\cdot 4^m\,(m+1)\,m!^2}\int_0^{\pi/2}(\cos\theta)^{2m+1}d\theta\\ &= \sum_{m=0}^{+\infty}\frac{(-1)^m\,x^{2m+1}}{2\cdot 4^m\,(m+1)\,m!^2}\cdot\frac{4^m\,m!^2}{(2m+1)!}\\ &= \sum_{m=0}^{+\infty}\frac{(-1)^m x^{2m+1}}{(2m+2)!}=\frac{1-\cos x}{x}.\end{aligned}$$

By exploiting Fubini's theorem we get:

$$\int_0^{+\infty}\frac{\sin x}{x}\,dx=\int_0^{\pi/2}\cos(\theta)\int_0^{+\infty}J_0(x\cos\theta)\,dx\,d\theta=\int_0^{\pi/2}\frac{\cos\theta}{\cos\theta}\,d\theta=\frac{\pi}{2}$$

and similarly

$$\int_0^{+\infty}\frac{1-\cos x}{x^2}\,dx=\int_0^{\pi/2}\int_0^{+\infty}\frac{J_1(x\cos\theta)}{x}\,dx\,d\theta=\int_0^{\pi/2}1\,d\theta=\frac{\pi}{2}.$$

A curious series. We investigate about the convergence of the series

$$\sum_{n\geq 0}J_0(n)=\sum_{n\geq 0}\frac{1}{2\pi}\int_0^{2\pi}e^{ni\cos\theta}\,d\theta.$$

The partial sums of this series can be written in the following form:

$$S_N=\sum_{n=0}^{N}J_0(n)=\frac{1}{2\pi}\int_0^{2\pi}\frac{1-e^{(N+1)i\cos\theta}}{1-e^{i\cos\theta}}\,d\theta=\frac{1}{2}+\frac{1}{2\pi}\int_0^{\pi}\frac{\sin\left[\left(N+\frac{1}{2}\right)\cos\theta\right]}{\sin\left[\frac{\cos\theta}{2}\right]}\,d\theta$$

by invoking the identities $\cos(\theta+\pi)=-\cos\theta$ and $\frac{1}{1-e^z}+\frac{1}{1-e^{-z}}=1$. Rearranging,

$$S_N=\frac{1}{2}+\frac{1}{2\pi}\int_{-1}^{1}\frac{D_N(x)}{\sqrt{1-x^2}}\,dx$$

where $D_N(x)$ is Dirichlet's kernel. If the coefficients c_n of the Fourier cosine series of $\frac{\mathbb{1}_{(-1,1)}(x)}{\sqrt{1-x^2}}$ were in ℓ^1, we could immediately state

$$(\heartsuit)\qquad \frac{1}{2\pi}\int_{-1}^{1}\frac{D_N(x)}{\sqrt{1-x^2}}\,dx\to\frac{1}{2\pi}\int_{-1}^{1}\frac{dx}{\sqrt{1-x^2}}+\frac{1}{2}\cdot\left.\frac{1}{\sqrt{1-x^2}}\right|_{x=0}=1\qquad\text{as } N\to+\infty.$$

Unluckily, our case is not the case: it can be shown that $|c_n|\leq\frac{8}{\pi^2\sqrt{n}}$ for any $n\geq 1$, but the density of ns such that $\frac{1}{2}\leq\sqrt{n}|c_n|\leq 1$ is positive, hence $\{c_n\}_{n\geq 1}\notin\ell^1$. On the other hand, in order to prove $(\heartsuit)$ it is enough to show that the Fourier series of $\frac{\mathbb{1}_{(-1,1)}(x)}{\sqrt{1-x^2}}$ is pointwise convergent at the origin. Since

$$\mathcal{L}\left(\frac{\mathbb{1}_{(0,1)}(x)}{\sqrt{1-x^2}}\right)(s)=\sum_{n\geq 0}\frac{\binom{2n}{n}}{4^n}\int_0^1 x^{2n}e^{-sx}\,dx=\frac{1-e^{-s}}{s}+\sum_{n\geq 1}\left(\frac{(2n)!}{2^n n!}\right)^2\frac{1}{s^{2n+1}}+w\left(\frac{1}{s}\right)$$

where $w(z)$ is an analytic function fulfilling $w(0)=w'(0)=0$, we have

$$\lim_{s\to+\infty}s\cdot\mathcal{L}\left(\frac{\mathbb{1}_{(0,1)}(x)}{\sqrt{1-x^2}}\right)(s)=1$$

and the pointwise convergent at the origin for the Fourier series of $\frac{1}{\sqrt{1-x^2}}$ is proved.
As a straightforward corollary, we have

$$\sum_{n\geq 0}J_0(n)=\frac{3}{2}.$$

In particular J_0 is an even entire function fulfilling

$$\sum_{n\geq 0} f(n) = +\frac{f(0)}{2} + \int_0^{+\infty} f(x)\,dx$$

while sinc is an even entire function fulfilling

$$\sum_{n\geq 0} f(n) = -\frac{f(0)}{2} + \int_0^{+\infty} f(x)\,dx.$$

The same argument applied to $J_k(x)$ with $k \in \mathbb{N}^+$ leads to $\sum_{n\geq 0} J_k(n) = \int_0^{+\infty} J_k(x)\,dx = 1$.

Series for fixed argument. From the Jacobi-Anger expansion:

$$e^{iz\cos\theta} = J_0(z) + 2\sum_{n=1}^{+\infty} i^n J_n(z)\cos(n\theta)$$

we have, by considering the imaginary part:

$$\sin(z\cos\theta) = 2\sum_{m=0}^{+\infty} (-1)^m J_{2m+1}(z)\cos((2m+1)\theta)$$

and we can remove the cosine-dependent term by exploiting the identities:

$$\int_0^{\pi/2} \cos((2n+1)x)\cos((2m+1)x)dx = \frac{\pi}{4}\delta_{m,n},$$

$$\sum_{m=0}^{+\infty} (-1)^m \cos((2m+1)\theta) = \frac{1}{2\cos\theta},$$

that give (integrating against the proper kernel):

$$\frac{4}{\pi}\int_0^{\pi/2} \frac{\sin(z\cos\theta)}{2\cos\theta}d\theta = 2\sum_{m=0}^{+\infty} J_{2m+1}(z),$$

hence:

$$\begin{aligned}\sum_{m=0}^{+\infty} J_{2m+1}(z) &= \frac{1}{\pi}\int_0^{\pi/2} \frac{\sin(z\cos\theta)}{\cos\theta}d\theta = \frac{1}{\pi}\int_0^1 \frac{\sin(zt)}{t\sqrt{1-t^2}}dt \\ &= \frac{1}{2}\sum_{r=0}^{+\infty} \frac{(-1)^r}{(2r+1)4^r(r!)^2} z^{2r+1} = \frac{1}{2}\int_0^z J_0(u)du.\end{aligned}$$

For the alternating sum, it is sufficient to take $\theta = 0$ in (1) in order to have:

$$\sum_{m=0}^{+\infty} (-1)^m J_{2m+1}(z) = \frac{1}{2}\sin z.$$

There are no issues in exploiting the pointwise convergence of a Fourier series since $g(\theta) = e^{iz\cos\theta}$ is an analytic function.

The integral of $\sin^2 \sin^2$. We have:

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} \sin^2(\sin^2(x))dx &= \int_0^{\frac{\pi}{2}} \frac{1-\cos(2\sin^2(x))}{2}dx \\
&= \frac{\pi}{4} - \frac{1}{2}\int_0^{\frac{\pi}{2}} \sum_{k\geq 0} \frac{(-1)^k 4^k}{(2k)!} \sin^{4k}(x)dx \\
&= \frac{\pi}{4} - \frac{1}{2}\sum_{k\geq 0} \frac{(-1)^k 4^k}{(2k)!} \int_0^{\frac{\pi}{2}} \sin^{4k}(x)dx \\
&= \frac{\pi}{4} - \frac{1}{2}\sum_{k\geq 0} \frac{(-1)^k 4^k}{(2k)!} \frac{(4k)!}{(4^k(2k)!)^2}\frac{\pi}{2} \\
&= \frac{\pi}{4} - \frac{\pi}{4}\sum_{k\geq 0} \frac{(-1)^k}{((2k)!)^3} \frac{(4k)!}{4^k}.
\end{aligned}$$

On the other hand:

$$\begin{aligned}
\int_0^{\frac{\pi}{2}} \sin^2(\sin^2(x))dx &= \int_0^{\frac{\pi}{2}} \frac{1-\cos(2\sin^2(x))}{2}dx = \int_0^{\frac{\pi}{2}} \frac{1-\cos(1-\cos(2x))}{2}dx \\
&= \frac{\pi}{4} - \int_0^{\frac{\pi}{2}} \frac{\cos(1)\cos(\cos(2x))}{2}dx - \int_0^{\frac{\pi}{2}} \frac{\sin(1)\sin(\cos(2x))}{2}dx \\
&= \frac{\pi}{4} - \frac{\cos(1)}{2}\int_0^{\frac{\pi}{2}} \cos(\cos(2x))dx = \frac{\pi}{4} - \frac{\cos(1)}{4}\int_0^{\pi} \cos(\cos(x))dx \\
&= \frac{\pi}{4} - \frac{\cos(1)\pi}{4} J_0(1).
\end{aligned}$$

Exercise 310. Prove the following identity:

$$\sum_{n\geq 0} \frac{(2n)!}{n!^3} = e^2 I_0(2).$$

Proof. By exploiting the integral representation for central binomial coefficients we have

$$\begin{aligned}
\sum_{n\geq 0} \frac{(2n)!}{n!^3} &= {}_1F_1\left(\tfrac{1}{2};1;4\right) = \frac{2}{\pi}\int_0^{\pi/2} \frac{(2\cos\theta)^{2n}}{n!}\,d\theta = \frac{2}{\pi}\int_0^{\pi/2} e^{4\cos^2\theta}\,d\theta \\
&= \frac{2e^2}{\pi}\int_0^{\pi/2} e^{2\cos(2\theta)}\,d\theta = \frac{e^2}{\pi}\int_0^{\pi} e^{2\cos\varphi}\,d\varphi = \frac{2e^2}{\pi}\int_0^{\pi/2} \cosh(2\cos\varphi)\,d\varphi \\
&= \frac{2e^2}{\pi}\sum_{m\geq 0}\frac{1}{(2m)!}\int_0^{\pi/2} (2\cos\varphi)^{2m}\,d\varphi = e^2\sum_{m\geq 0}\frac{1}{m!^2} = e^2\,{}_1F_2(1;1,1;1) = e^2 I_0(2).
\end{aligned}$$

□

Exercise 311. Investigate about the series

$$S = \sum_{n\geq 0} \frac{(2n)!}{n!^4}$$

and hypergeometric identities provided by such a series.

Proof. Through the series representation for $I_0(2z)$ and the integral representation for central binomial coefficients we have

$$S = {}_1F_2\left(\tfrac{1}{2};1,1;4\right) = \frac{2}{\pi}\int_0^{\pi/2} I_0(4\cos\theta)\,d\theta = \frac{2}{\pi}\int_0^1 \frac{I_0(4x)}{\sqrt{1-x^2}}\,dx$$

and the RHS can be written as the following double series:

$$\frac{2}{\pi}\sum_{m,n\geq 0}\frac{\binom{2m}{m}4^n}{4^m n!^2(2m+2n+1)} = \sum_{n\geq 0}\frac{\binom{2n}{n}}{n!^2} = \sum_{m,n\geq 0}\frac{1}{m!^2 n!^2} = I_0(2)^2 = {}_1F_2(1;1,1;1)^2.$$

Actually S is just one of the coefficients of the Fourier cosine series of $I_0(4\cos\theta)$:

$$I_0(4\cos\theta) = I_0(2)^2 + \sum_{m\geq 1} c_m\cos(2m\theta),\qquad c_m = \frac{1}{\pi}\int_0^{2\pi}\sum_{n\geq 0}\frac{4^n(\cos\theta)^{2n}}{n!^2}\cos(2m\theta)\,d\theta.$$

Using both $\cos\theta = \text{Re}(e^{i\theta}) = \frac{e^{i\theta}+e^{-i\theta}}{2}$ and the binomial theorem we have

$$c_m = 2\sum_{n\geq 0}\frac{\binom{2n}{n+m}}{n!^2} = 2I_m(2)^2$$

and the pointwise convergence of the involved Fourier series leads to

$$I_0(4) = I_0(2)^2 + 2\sum_{m\geq 1} I_m(2)^2,$$

i.e. to an instance of the duplication formula for the I_0 function, which can be seen as a direct consequence of Parseval's identity, too. Since the Fourier coefficients of $I_0(4\cos\theta)$ depend on I_m^2, Parseval's identity provides a curious identity about the series of $I_m(2)^4$.

$$I_0(4z)^2 = \sum_{n\geq 0}\frac{4^n(2n)!}{n!^4}z^{2n}$$

gives

$$\int_0^{\pi/2} I_0(4\cos\theta)^2\,d\theta = \frac{\pi}{2}\sum_{n\geq 0}\frac{(2n)!^2}{n!^6} = \frac{\pi}{2}\cdot{}_2F_3\left(\tfrac{1}{2},\tfrac{1}{2};1,1,1;16\right)$$

hence by the orthogonality relations

$$\sum_{n\geq 0}\frac{(2n)!^2}{n!^6} = {}_2F_3\left(\tfrac{1}{2},\tfrac{1}{2};1,1,1;16\right) = I_0(2)^4 + 2\sum_{m\geq 1} I_m(2)^4.$$

□

Some series from Ramanujan and some generalizations. We have

$$\left(\sqrt{x+1}-1\right)^4 = -4\sum_{m\geq 3}\left[\binom{1/2}{m-1}+2\binom{1/2}{m}\right]x^m$$

$$\left(\sqrt{1+\frac{1}{x}}-\sqrt{\frac{1}{x}}\right)^4 = -4\sum_{m\geq 1}\left[\binom{1/2}{m+1}+2\binom{1/2}{m+2}\right]x^m$$

hence:

$$\begin{aligned}\sum_{n\geq 1}\left(\sqrt{n+1}-\sqrt{n}\right)^4 &= -4\sum_{m\geq 1}\left[\binom{1/2}{m+1}+2\binom{1/2}{m+2}\right]\zeta(m)\\ &= -4\int_0^{+\infty}\frac{dx}{e^x-1}\sum_{m\geq 1}\left[\binom{1/2}{m+1}+2\binom{1/2}{m+2}\right]\frac{x^{m-1}}{(m-1)!}\\ &= \int_0^{+\infty}\frac{2e^{-x/2}\left(x\,I_0\left(\frac{x}{2}\right)-4\,I_1\left(\frac{x}{2}\right)\right)}{x^2(e^x-1)}\,dx\\ &= 2\int_0^{+\infty}\frac{I_2(x)\,dx}{xe^x(e^x-1)(e^x+1)}\end{aligned}$$

where the last integrand function is pretty close to $\frac{1}{16}e^{-2x}$, from which it follows that the original series is pretty close to $\frac{1}{16}$. The following integral representation for the Bessel function I_2

$$I_2(x) = \frac{x^2}{3\pi}\int_0^{\pi} \exp\left(x\cos\theta\right)\sin^4(\theta)\,d\theta$$

leads to:

$$\begin{aligned}\sum_{n\geq 1}\left(\sqrt{n+1}-\sqrt{n}\right)^4 &= \frac{1}{6\pi}\int_0^{\pi}\psi'\left(\tfrac{3-\cos\theta}{2}\right)\sin^4(\theta)\,d\theta \\ &= -1+\frac{1}{6\pi}\int_0^{\pi}\psi'\left(\tfrac{1-\cos\theta}{2}\right)\sin^4(\theta)\,d\theta \\ &= -1+\frac{\pi}{6}\int_0^{\pi/2}\frac{\sin^4(\theta)}{\sin^2\left(\pi\sin^2\frac{\theta}{2}\right)}\,d\theta \\ &= -1+\frac{\pi}{3}\int_0^{\pi/4}\frac{\sin^4(2\theta)\,d\theta}{\sin^2\left(\pi\sin^2\theta\right)}\end{aligned}$$

by the reflection formula for the trigamma function. The blue integral can be written in the more symmetric form

$$\boxed{\sum_{n\geq 0}\left(\sqrt{n+1}-\sqrt{n}\right)^4 = \frac{4\pi}{3}\int_0^1\frac{x^{3/2}(1-x)^{3/2}}{\sin^2(\pi x)}\,dx = \frac{\pi}{6}\int_0^1\frac{(1-x^2)^{3/2}\,dx}{\cos^2\frac{\pi x}{2}}.}$$

Let us tackle the case $s=3$ with a similar approach. We have

$$(\sqrt{x+1}-1)^3 = \sum_{m\geq 3}\left[4\binom{1/2}{m}+\binom{1/2}{m-1}\right]x^m$$

$$\left(\sqrt{1+\frac{1}{x}}-\sqrt{\frac{1}{x}}\right)^3 = \sum_{m\geq 3}\left[4\binom{1/2}{m}+\binom{1/2}{m-1}\right]x^{m-3/2}$$

$$\begin{aligned}\sum_{n\geq 1}\left(\sqrt{n+1}-\sqrt{n}\right)^3 &= \sum_{m\geq 3}\left[4\binom{1/2}{m}+\binom{1/2}{m-1}\right]\zeta\left(m-\frac{3}{2}\right) \\ &= \sum_{m\geq 3}\left[4\binom{1/2}{m}+\binom{1/2}{m-1}\right]\frac{1}{(m-\frac{5}{2})!}\int_0^{+\infty}\frac{x^{m-5/2}}{e^x-1}\,dx \\ &= \int_0^{+\infty}\frac{3e^{-x}\left(2-2e^x+x+e^x x\right)}{2\sqrt{\pi}x^{5/2}(e^x-1)}\,dx \\ &= \frac{3}{\sqrt{\pi}}\int_0^{+\infty}\left[\frac{1}{2x^{3/2}e^x}-\frac{1}{x^{5/2}e^x}+\frac{1}{x^{5/2}e^x(e^x-1)}\right]dx\end{aligned}$$

hence

$$\boxed{\sum_{n\geq 0}\left(\sqrt{n+1}-\sqrt{n}\right)^3 = \tfrac{3}{2\pi}\,\zeta\left(\tfrac{3}{2}\right)}$$

just follows from integration by parts, Frullani's Theorem and the integral representation for the ζ function. The same approach allows an explicit evaluation in terms of the ζ function for any odd value of s. For instance:

$$\sum_{n\geq 0}\left(\sqrt{n+1}-\sqrt{n}\right)^5 = \tfrac{15}{2\pi^2}\,\zeta\left(\tfrac{5}{2}\right),\qquad \sum_{n\geq 0}\left(\sqrt{n+1}-\sqrt{n}\right)^7 = \tfrac{7}{2\pi}\,\zeta\left(\tfrac{3}{2}\right)-\tfrac{105}{2\pi^3}\,\zeta\left(\tfrac{7}{2}\right)$$

$$\sum_{n\geq 0}\left(\sqrt{n+1}-\sqrt{n}\right)^9 = \tfrac{90}{2\pi^2}\zeta\left(\tfrac{5}{2}\right)-\tfrac{945}{2\pi^4}\zeta\left(\tfrac{9}{2}\right).$$

Relations between Hermite polynomials and Bessel functions Starting with the generating function

$$e^{2xt-t^2} = \sum_{n\geq 0} H_n(x)\frac{t^n}{n!}$$

then replacing t with $te^{i\theta}$ we have

$$\exp\left[2xte^{i\theta} - t^2 e^{2i\theta}\right] = \sum_{n\geq 0} H_n(x)e^{ni\theta}\frac{t^n}{n!}$$

and by Parseval's identity

$$\int_{-\pi}^{\pi} \exp\left[4xt\cos\theta - 2t^2\cos(2\theta)\right]\,d\theta = 2\pi\sum_{n\geq 0} H_n(x)^2\frac{t^{2n}}{n!^2}.$$

Now we can multiply both sides of (3) by $e^{-x^2}e^{kix}$ and apply $\int_{\mathbb{R}}(\ldots)\,dx$ to get

$$\sqrt{\pi}e^{-k^2/4}\int_{-\pi}^{\pi} e^{2t^2-2ikt\cos\theta}\,d\theta = 2\pi\sum_{n\geq 0}\left(\int_{\mathbb{R}} H_n(x)^2 e^{-x^2}e^{kix}\,dx\right)\frac{t^{2n}}{n!^2}$$

which simplifies into

$$\sqrt{\pi}e^{-k^2/4}e^{2t^2}\,J_0(2kt) = \sum_{n\geq 0}\left(\int_{\mathbb{R}} H_n(x)^2 e^{-x^2}e^{kix}\,dx\right)\frac{t^{2n}}{n!^2}$$

and the wanted integral can be recovered from the Cauchy product between the Taylor series of e^{2t^2} and the Taylor series of $J_0(2kt)$:

$$\left(\int_{\mathbb{R}} H_n(x)^2 e^{-x^2}e^{kix}\,dx\right) = \sqrt{\pi}e^{-k^2/4}n!^2\cdot[t^{2n}]\sum_{a,b\geq 0}\frac{2^a(-1)^b k^{2b}t^{2a+2b}}{a!b!^2}$$

such that:

$$\boxed{\int_{\mathbb{R}} H_n(x)^2 e^{-x^2}e^{kix}\,dx = \sqrt{\pi}e^{-k^2/4}n!^2\sum_{b=0}^{n}\frac{2^{n-b}(-1)^b k^{2b}}{(n-b)!b!^2}.}$$

One could place the right hand side of the last equation into the more known form of

$$\sqrt{\pi}\,2^n\,n!\,e^{-k^2/4}\,L_n\left(\frac{k^2}{2}\right),$$

where $L_n(x)$ are the Laguerre polynomials.

12.1 The Gauss circle problem

Disclaimer: most of the contents of this subsection are freely adapted from the books, notes and articles of T. Jameson, A. Ivic, J. Bell and G.N. Watson. Let $N(R)$ be the number of lattice points in the region $\{(x,y) : x^2+y^2\leq R^2\}$ and let $r_2(N)$ be the following arithmetic function:

$$r_2(N) = \left|\{(x,y)\in\mathbb{Z}^2 : x^2+y^2 = N\}\right|.$$

A double-counting argument easily leads to the following identity:

$$N(R) = \sum_{x=-R}^{R}\left(1+2\left\lfloor\sqrt{R^2-x^2}\right\rfloor\right) = \sum_{N=0}^{R^2} r_2(N)$$

and a reasonable claim is that $N(R)$, for sufficiently large values of R, is close to the area of a circle with radius R, i.e. πR^2. Since $\mathbb{Z}[i]$ (the ring of Gaussian integers) is an Euclidean domain, $r_2(N)$ only depends on the prime factors of N. Such numbers of representations can be shown to be four times a multiplicative function, since $\mathbb{Z}[i]$ has four invertible elements:

$$r_2(N) = 4(\chi_4 * 1)(N) = 4\sum_{d|N}\chi_4(d), \qquad \chi_4(d) = \begin{cases} 1 & \text{if} \quad d \equiv 1 \pmod 4, \\ -1 & \text{if} \quad d \equiv 3 \pmod 4, \\ 0 & \text{if} \quad d \equiv 0 \pmod 2. \end{cases}$$

The algebra of Dirichlet series hence ensures

$$L(r_2, s) \stackrel{\text{def}}{=} \sum_{N\geq 1}\frac{r_2(N)}{N^s} = 4\zeta(s)L(\chi_4, s), \qquad L(\chi_4, s) = \sum_{n\geq 0}\frac{(-1)^{n+1}}{(2n+1)^s}$$

for any $s \in \mathbb{C}$ such that $\text{Re}(s) > 1$. We may notice that

$$L(\chi_4, 1) = \sum_{n\geq 1}\frac{(-1)^{n+1}}{2n+1} = \int_0^1 \frac{dx}{1+x^2} = \frac{\pi}{4}$$

and that summation by parts grants

$$\zeta(s) = \frac{1}{s-1} + \gamma + O(s-1)$$

as $s \to 1^+$. Summation by parts also gives

$$L(r_2, s) = \sum_{n\geq 1} N(\sqrt{n})\left[\frac{1}{n^s} - \frac{1}{(n+1)^s}\right] = \sum_{n\geq 1}\frac{N(\sqrt{n})}{n}\left[\frac{1}{n+1} + (s-1)\cdot O\left(\frac{\log n}{n^2}\right)\right]$$

hence by assuming that $\frac{N(R)}{R^2}$ is convergent, its limit has to be $4L(\chi_4, 1) = \pi$. This can be shown through a simple geometric argument. Let us consider $U(R)$ as the union of the squares centered at the lattice points of $\{(x, y) : x^2 + y^2 \leq R^2\}$, all of them having unit side length. $U(R)$ is contained in a circle having radius $R + \frac{\sqrt{2}}{2}$. Conversely, $U\left(R - \frac{\sqrt{2}}{2}\right)$ is contained in a circle with radius R. It follows that

$$\pi\left(R - \frac{\sqrt{2}}{2}\right)^2 \leq N(R) \leq \pi\left(R + \frac{\sqrt{2}}{2}\right)^2$$

hence

Theorem 312 (Gauss)**.** For any $\varepsilon > 0$,

$$\left|N(R) - \pi R^2\right| \leq (\sqrt{2} + \varepsilon)R = O(R)$$

as $R \to +\infty$.

Soon after Gauss' work, mathematicians wondered about the optimality of the bound $O(R)$ for the difference between $N(R)$ and the area of the circle with radius R. Since the Dirichlet L-function $L(\chi_4, s)$ has an infinitude of zeroes in the strip $0 \leq \text{Re}(s) \leq 1$, it can be shown (as done by Hardy) that the bound for the error term cannot be improved beyond $O(R^{1/2})$. The purpose of the final part of this section is to prove that it *can* be improved as follows:

Theorem 313 (Sierpinski,Voronoi)**.** As $R \to +\infty$,

$$\left|N(R) - \pi R^2\right| = O(R^{2/3}).$$

The classical proofs revolve around a few key ingredients. The first fact is that the sum

$$\sum_{x=-R}^{R}\left(1+2\sqrt{R^2-x^2}\right)$$

can be estimated through the trapezoid method or Simpson's rule, hence the problem is reduced to an accurate estimation of

$$\sum_{x=-R}^{R}\left\{\sqrt{R^2-x^2}\right\}$$

where $\{z\}$ stands for the fractional part of z. The Fourier sine series of $\rho(z)=\frac{1}{2}-\{z\}$ is well-known, the discrete Fourier transform of $\{\sqrt{R^2-x^2}\}$ is related to Bessel functions of the first kind, with a known asymptotic behavior. In order to partially compensate the erratic behaviour of $r_2(n)$ over the integers, actual bounds for $N(R)$ are produced by considering an averaged version of $r_2(n)$ over suitably short intervals. Let $N'(R)$ denote the number of pairs of integers (m,n) satisfying $m^2+n^2\leq R^2, m>0, n\geq 0$ and let $M=\lfloor\frac{R}{\sqrt{2}}\rfloor$, $f(x)=\sqrt{R^2-x^2}-x$. Then

$$N'(R)=M+\lfloor R\rfloor+2\sum_{m=1}^{M}\lfloor f(m)\rfloor=\lfloor R\rfloor+2\sum_{m=1}^{M}f(m)+2\sum_{m=1}^{M}\rho(f(m))$$

and integration by parts leads to the fundamental formula

$$N'(R)=\frac{\pi R^2}{4}+2\sum_{m=1}^{M}\rho(f(m))+O(1).$$

If we define $P(x)$ as $N'(\sqrt{x})-\frac{\pi x}{4}$, from the fact that $N'(\sqrt{x})$ is increasing we get

$$P(X)\leq P(X+y)+\frac{\pi y}{4},\qquad P(X)\geq P(X-y)-\frac{\pi y}{4},$$

hence by integration and the triangle inequality we have

$$|P(X)|\leq\frac{1}{Y}\max\left(\left|\int_{X-Y}^{X}P(x)\,dx\right|,\left|\int_{X}^{X+Y}P(x)\,dx\right|\right)+\frac{\pi Y}{8}$$

yielding a bound for $P(X)$, given a bound for the average of $P(x)$ over short intervals. In the following manipulations we will assume $Y\ll\sqrt{X}$, such that $x=X+O(Y)$ ensures $\sqrt{x}=\sqrt{X}+O(1)$. We have:

$$P(x)=2\sum_{m\leq\sqrt{X/2}}\rho(\sqrt{x-m^2})+O(1),$$

$$\int_{X-Y}^{X+Y}P(x)\,dx=2\sum_{m\leq\sqrt{X/2}}\int_{X-Y}^{X+Y}\rho(\sqrt{x-m^2})\,dx+O(Y).$$

By using the Fourier series of the fractional part provided by Bernoulli polynomials we obtain

$$\int_{X-Y}^{X+Y}P(x)\,dx=O(\sqrt{X})+\frac{2}{\pi^2}\operatorname{Re}\sum_{h\geq 1}\frac{1}{h^2}\sum_{m\leq\sqrt{X/2}}\sqrt{X-m^2}\left[e\left(\sqrt{X-Y-m^2}\right)-e\left(\sqrt{X+Y-m^2}\right)\right]\tag{1}$$

where $e(z)$ is the common shorthand notation for $e^{2\pi iz}$. The original problem is now converted into the approximated evaluation of exponential sums. Such task can be accomplished by recalling two classical results due to Kusmin, Landau and Van der Corput.

Theorem 314 (Kusmin-Landau)**.** Let $f_A,\ldots,f_B$ be real numbers (where A,B are integers with $A<B$) and let $\delta_n=f_n-f_{n-1}$. Suppose that δ_n is a weakly monotonic function of n. Suppose also that there is some integer K and some $\delta\in\left(0,\frac{1}{2}\right]$ such that $\delta_n\in[K+\delta,K+1-\delta]$ for all n. Then

$$S\stackrel{\text{def}}{=}\left|\sum_{n=A}^{B}e(f_n)\right|\leq\frac{2}{\sin(\pi\delta)}.$$

If f_n is decreasing we may negate all the f_ns, which simply turns S into $\overline{S}$, which has the same size as S. Then we may assume without loss of generality that δ_n is non-decreasing. We may also assume that $K = 0$, by replacing f_n by $f_n - Kn$. The proof of the Kusmin-Landau theorem relies on a clever trick to express $e(f_n)$ as a difference, then on the application of summation by parts. We begin by writing (for $A < n \leq B$)

$$e(f_n) - e(f_{n-1}) = e(f_n)(1 - e(-\delta_n))$$

then setting

$$g_n = \frac{1}{1 - e(-\delta_n)} = \frac{1}{2} - \frac{i}{2}\cot(\pi\delta_n)$$

such that

$$e(f_n) = g_n e(f_n) - g_n e(f_{n-1})$$

and

$$|1 - g_n| = |\overline{g}_n| = |g_n| = \frac{1}{2\sin(\pi\delta_n)}.$$

The expression on the RHS is positive because from our assumptions we have $0 < \delta_n < 1$.
By using summation by parts we get

$$S = (1 - g_{A+1})e(f_A) + g_B e(f_B) + \sum_{n=A+1}^{B-1}(g_n - g_{n-1})e(f_n)$$

and the claim now follows from the triangle inequality:

$$|S| \leq |1 - g_{A+1}| + |g_B| + \sum_{n=A+1}^{B-1}|g_n - g_{n-1}| \leq \frac{1}{\sin(\pi\delta_{A+1})} + \frac{1}{\sin(\pi\delta_B)} \leq \frac{2}{\sin(\pi\delta)}.$$

Suppose we have a differentiable function $f(x)$ with $f'(x)$ being monotonic and fulfilling $f'(x) \in [K + \delta, K + 1 - \delta]$ for all $x \in [A, B]$. Then with $f_n = f(n)$ we have $\delta_n = f(n) - f(n-1) = \int_{n-1}^{n} f'(x)\,dx$ which clearly satisfies the conditions for the applicability of the Kusmin-Landau theorem. Since we also have $\sin(\pi\delta) \geq 2\delta$ for $\delta \in \left(0, \frac{1}{2}\right]$, we may write

$$\left|\sum_{n=A}^{B} e(f(n))\right| \leq \frac{1}{\delta}.$$

Also, if $A = B$ then this sum contains only one term and it has modulus 1, so that this bound holds trivially (since $\frac{1}{\delta} \geq 2$). Suppose now that we have a *twice* differentiable $f(x)$ defined for $x \in [a, b]$, such that $0 < \lambda \leq f''(x) \leq h\lambda$. Here we have $a < b$ and a, b need not be integers. This implies that $h \geq 1$ and that $f'(x)$ is strictly increasing. From the Kusmin-Landau theorem we will derive a bound for

$$S \stackrel{\text{def}}{=} \sum_{a \leq n \leq b} e(f(n)).$$

Let $f'(a) = \alpha$ and $f'(b) = \beta$. For a free parameter $\delta \in \left(0, \frac{1}{2}\right]$ we partition the interval $[\alpha, \beta]$ into sub-intervals of the form $I_n = (n - \delta, n + \delta)$ (containing the reals close to the integer n), and of the form $J_n = [n + \delta, n + 1 - \delta]$ (containing reals at least δ-apart from the integers). Accordingly, we split the sum S into subsums over ranges for x corresponding to these ranges for $f'(x)$. The condition that $I_n \subseteq [\alpha, \beta]$ is equivalent to the condition $n \in (\alpha - \delta, \beta + \delta)$, and the number of ns satisfying this is $\leq \beta - \alpha + 2$. Similarly, the condition that $J_n \subseteq [\alpha, \beta]$ is equivalent to the condition $n \in (\alpha - 1 + \delta, \beta - \delta)$ and the number of ns satisfying this is $\leq \beta - \alpha + 2$ as well. By invoking the Kusmin-Landau theorem we have

$$|S| \leq (\beta - \alpha + 2)\left(\frac{1}{\delta} + \frac{2\delta}{\lambda} + 1\right)$$

and this expression is minimized by choosing $\delta = \sqrt{\lambda/2}$, leading to:

Theorem 315 (Van der Corput).
$$|S| \leq (\beta - \alpha + 2)\left(\sqrt{\frac{8}{\lambda}} + 1\right).$$

The square root appearing in the RHS is crucial in encoding the cancellations in the involved exponential sum. When dealing with weighted exponential sums, we may observe that

$$\begin{aligned} \left|\sum_{a<n\leq b} g(n)c_n\right| &= \left|\int_{a^+}^{b^+} g(x)\, d\sum_{a<n\leq x} c_n\right| \\ &= \left|g(b)\sum_{a<n\leq b} c_n - \int_a^b g'(x)\sum_{a<n\leq x} c_n\, dx\right| \\ &\leq \left(|g(b)| + \int_a^b |g'(x)|\, dx\right)\max_{a<x\leq b}\left|\sum_{a<n\leq x} c_n\right| \end{aligned}$$

which is very practical in inserting (or removing, depending on which way around we read things) a smooth slowly varying weight $g(x)$. In the current case Van der Corput's theorem leads to

$$\sum_{m\leq t} e\left(h\sqrt{X-m^2}\right) \ll h\left(\frac{h}{\sqrt{X}}\right)^{-1/2} = X^{1/4}h^{1/2}$$

for $t \leq \sqrt{X/2} + O(1)$, then summation by parts grants

$$\sum_{m\leq t} \sqrt{X-m^2}\, e(h\sqrt{X-m^2}) \ll X^{3/4}h^{1/2}$$

and
$$\sum_{m\leq\sqrt{X/2}} \sqrt{X-m^2}\left[e(h\sqrt{X-Y-m^2}) - e(h\sqrt{X+Y-m^2})\right] \ll X^{3/4}h^{1/2}\cdot\sqrt{X}\cdot\frac{hY}{X} = X^{1/4}h^{3/2}Y.$$

Recalling (1) now we state

$$\begin{aligned} \int_{X-Y}^{X+Y} P(x)\, dx &\ll \sum_{h\leq X^{1/2}Y} YX^{1/4}h^{-1/2} + \sum_{h>X^{1/2}Y} X^{3/4}h^{-3/2} \\ &\ll YX^{1/4}\left(\frac{X^{1/2}}{Y}\right)^{1/2} + X^{3/4}\left(\frac{X^{1/2}}{Y}\right)^{-1/2} \\ &\ll (YX)^{1/2} \end{aligned}$$

and the mean-to-max trick gives us

$$P(X) \ll Y + \left(\frac{X}{Y}\right)^{1/2}$$

which is clearly minimized by taking $Y = X^{1/3}$, finally giving

$$P(X) \ll X^{1/3}.$$

It should not take too much imagination to envisage that there is nothing spectacularly important about us restricting to a circle in all of this, so that (using mainly just a little notational gameplay) we can get similar results for simple closed curves satisfying certain conditions (like twice differentiability and radius of curvature bounded above and below). We have had formulae with square root signs all over the place coming from Pythagoras, but in essence we are just using linearising tricks. More usually a proof is given using the Poisson summation formula and bounds for the resulting "exponential integrals": this is exactly the second approach we are going to outline.

The Sonine-Schafheitlin integrals ensure that

$$\frac{2^{s-p-1}\Gamma(s/2)}{\Gamma(p+1-s/2)} \qquad \left(0<\operatorname{Re}(s)=\sigma<p+\tfrac{3}{2}\right)$$

is the Mellin transform of $x^{-p}J_p(x)$. This brings to the table the Perelli S-class of Dirichlet series.
If a_n is a arithmetic function and the associated Dirichlet series

$$L(a,s)=\sum_{n\geq 1}\frac{a_n}{n^s} \qquad \operatorname{Re}(s)>1$$

has an analytic continuation to $\mathbb{C}$ whose only possible pole is at $s=1$, and there are $A,a_1,\ldots,a_g$ such that

$$\gamma(s)=A^s\prod_{j=1}^{g}\Gamma(a_j s)$$

fulfills the reflection formula $\gamma(s)L(a,s)=\gamma(1-s)L(a,1-s)$, then by assuming $f\in\mathcal{S}(\mathbb{R})$ and defining

$$\mathscr{M}(f)(s)=\int_0^{+\infty}x^{s-1}f(x)\,dx,\qquad H(x)=\frac{1}{2\pi i}\int_{\operatorname{Re}(s)=\frac{3}{2}}\frac{\gamma(s)}{\gamma(1-s)}x^{-s}\,ds,\qquad g(x)=\int_0^{+\infty}f(y)H(xy)\,dy$$

we have:

$$\sum_{n\geq 1}a_n f(n)=f(0)L(a,0)+\operatorname*{Res}_{s=1}\mathscr{M}(f)(s)L(a,s)+\sum_{n\geq 1}a_n g(n).$$

Since f is a Schwartz function its Mellin transform $\mathscr{M}(f)$ is holomorphic on $\operatorname{Re}(s)>0$. Over such region integration by parts ensures

$$\mathscr{M}(f)(s)=-\frac{1}{s}\mathscr{M}(f')(s+1),$$

hence $\mathscr{M}(f)$ has an analytic continuation to $\mathbb{C}$ possibly with poles at $0,-1,-2,\ldots$.
Denoting $\mathscr{M}(f)$ as F, the Mellin inversion formula grants

$$\begin{aligned}\sum_{n\geq 1}a_n f(n) &= \sum_{n\geq 1}\frac{a_n}{2\pi i}\int_{\operatorname{Re}(s)=\frac{3}{2}}n^{-s}F(s)\,ds\\ &= \frac{1}{2\pi i}\int_{\operatorname{Re}(s)=\frac{3}{2}}F(s)L(a,s)\,ds.\end{aligned}$$

The only possible pole of $L(a,s)$ is at $s=1$. From $\mathscr{M}(f)(s)=-\frac{1}{s}\mathscr{M}(f')(s+1)$ the only possible pole of $F(s)$ in the half-plane $\operatorname{Re}(s)>-1$ is at $s=0$, and the residue of $F(s)$ at $s=0$ is

$$-\mathscr{M}(f')(1)=-\int_0^{+\infty}f'(x)\,dx=f(0),$$

so the residue of $F(s)L(a,s)$ at $s=0$ is $f(0)L(a,0)$. By the residue theorem, taking as given that $F(s)L(a,s)\to 0$ uniformly in $-\frac{1}{2}\leq\operatorname{Re}(s)\leq\frac{3}{2}$ as $|\operatorname{Im}(s)|\to\infty$, we have

$$\sum_{n\geq 1}a_n f(n)=f(0)L(a,0)+\operatorname*{Res}_{s=1}F(s)L(a,s)+\frac{1}{2\pi i}\int_{\operatorname{Re}(s)=-\frac{1}{2}}F(s)L(a,s)\,ds$$

by shifting the integration line. Now we may exploit the reflection formula for $L(a,s)$, introducing $G(s)=F(s)\frac{\gamma(s)}{\gamma(1-s)}$. The last integral in the previous line turns into $\int_{\operatorname{Re}(s)=\frac{3}{2}}G(s)L(a,s)\,ds$, and it is tedious but straightforward to check that

$$\frac{1}{2\pi i}\int_{\operatorname{Re}(s)=\frac{3}{2}}x^{-s}G(s)\,ds=g(x),$$

proving the Lemma through the Mellin inversion formula. Since $L(r_2, s)$ belongs to the Perelli S-class and

$$L(\chi_4, s) = 2^{1-2s}\frac{\pi^{(s+1)/2}\Gamma\left(\frac{2-s}{2}\right)}{\pi^{(2-s)/2}\Gamma\left(\frac{1+s}{2}\right)}L(\chi_4, 1-s)$$

holds as a consequence of the Poisson summation formula, we have the following *resummation formula* for r_2:

Theorem 316 (Voronoi)**.**

$$\sum_{a\leq n\leq b}{}^{'} r_2(n)f(n) = \pi\int_a^b f(x)\,dx + \sum_{n\geq 1} r_2(n)\int_a^b f(x)J_0(2\pi\sqrt{xn})\,dx$$

where $f(x)$ is a suitably smooth function and $\sum'$ denotes that at $n = a$ or $n = b$ the summand is to be halved if a or b is an integer.

[11] In view of the non-negativity of $r_2(n)$ we have

$$\sum_{n\geq 1} f_-(n)r_2(n) \leq \sum_{X<n\leq 2X} r_2(n) \leq \sum_{n\geq 1} f_+(n)r_2(n)$$

where f_- is a smooth, non-negative function supported in $[X, 2X]$ such that $f(x) = 1$ for $x \in [X+G, 2X-G]$ ($X^\varepsilon \leq G \leq \sqrt{X}$), while similarly f_+ is supported in $[X-G, 2X+G]$ and satisfies $f(x) = 1$ for $x \in [X, 2X]$. If henceforth we denote by $f(x)$ either $f_-(x)$ or $f_+(x)$, then $f^{(r)}(x) \ll_r G^{-r}$ ($r = 0, 1, 2, \ldots$) and by the Voronoi summation formula

$$\sum_{n\geq 1} f(n)r_2(n) = \pi X + O(G) + \sum_{n\geq 1} r_2(n)\int_{X-G}^{2X+G} f(x)J_0(2\pi\sqrt{xn})\,dx.$$

From the theory of Bessel functions we recall the identity $\frac{d}{dz}\left[z^\nu J_\nu(z)\right] = z^\nu J_{\nu-1}(z)$ and the bound $J_\nu(z) \ll \frac{1}{\sqrt{z}}$ for $z \to +\infty$. It follows that for small values of n ($n \leq Y$) the integral appearing in the RHS of the last line is $\ll X^{1/4}n^{-3/4}$, and by invoking $\sum_{n\leq x} r_2(n) \ll x$ and summation by parts we have

$$\sum_{n\leq Y} r_2(n)X^{1/4}n^{-3/4} \ll (XY)^{1/4}.$$

By using the recurrence relation for J'_ν, performing two integration by parts and noting that the support of f'' has measure $\ll G$, we obtain that

$$\begin{aligned}\sum_{n>Y} r_2(n)\int_{X-G}^{2X+G} f(x)J_0(2\pi\sqrt{xn})\,dx &= \sum_{n>Y}\frac{r_2(n)}{\pi^2 n}\int_{X-G}^{2X+G} f''(x)xJ_2(2\pi\sqrt{xn})\,dx \\ &\ll \sum_{n>Y} r_2(n)n^{-5/4}G^{-1}X^{3/4} \ll X^{3/4}G^{-1}Y^{-1/4}.\end{aligned}$$

By combining the bounds obtained for small/large values of n we get

$$\sum_{n\geq 1} f(n)r_2(n) = \pi X + O(G) + O((XY)^{1/4}) + O(X^{3/4}G^{-1}Y^{-1/4})$$

which simplifies into

$$\sum_{X<n\leq 2X} r_2(n) = \pi X + O(X^{1/3})$$

by choosing $G = Y = X^{1/3}$. Replacing X with $2^{-j}R^2$ and summing over $j \geq 1$,

$$\left|N(R) - \pi R^2\right| = O(R^{2/3})$$

[11] A *resummation formula* exists every time a Dirichlet L-series has a suitably structured reflection formula, **and vice-versa**.

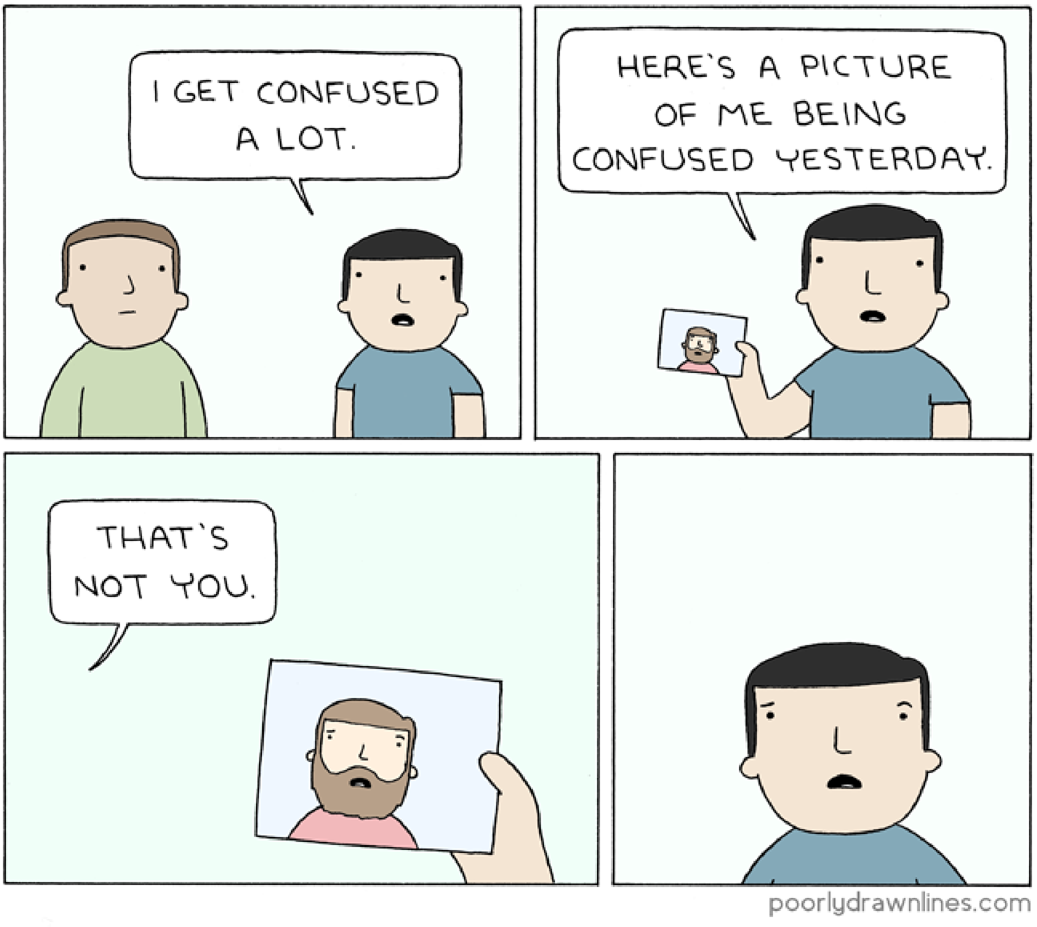
is finally proved. Various authors have exploited the oscillations of J_0 and J_1 to refine the previous bounds. For instance Huxley has proved (in 2003) that the exponent $\frac{2}{3}$ can be replaced with the slightly smaller $\frac{131}{208}$. In the opposite direction, Hardy has shown in 1925 that $\left|N(R) - \pi R^2\right|$ is as large as $R^{1/2}$ infinitely often, by exploiting the reflection formula

$$\sum_{n\geq 0} \frac{r_2(n)}{\sqrt{n+a}} e^{-2\pi\sqrt{(n+a)b}} = \sum_{n\geq 0} \frac{r_2(n)}{\sqrt{n+b}} e^{-2\pi\sqrt{(n+b)a}}$$

due to Ramanujan. The techniques outlined in this section can be applied to the *divisor problem* too, concerning the number of lattice points in the first quadrant under a rectangular hyperbola. We have

$$D(x) = {\sum_{n\leq x}}' d(n) = x(\log x + 2\gamma - 1) + \Delta(x)$$

where $\Delta(x) = O(x^{1/2})$ follows from simple geometric arguments. Voronoi's summation formula for $d(n)$ involves Bessel functions of the second kind and it allows to prove that $\Delta(x) = O(x^{1/3}\log x)$. In the opposite direction, $\Delta(x)$ is as large as $x^{1/4}\log x$ infinitely often.



13 Dilworth, Erdos-Szekeres, Brouwer and Borsuk-Ulam's Theorems

Definition 317. Let $E \subset \mathbb{R}^n$. We define the **convex hull** of E as:

$$\text{Hull}(E) = \{\lambda\, e_1 + (1-\lambda)\, e_2 : \lambda \in [0,1],\ e_1, e_2 \in E\}.$$

It is simple to check that such set is the smallest (with respect to $\subseteq$) convex subset of $\mathbb{R}^n$ containing E.

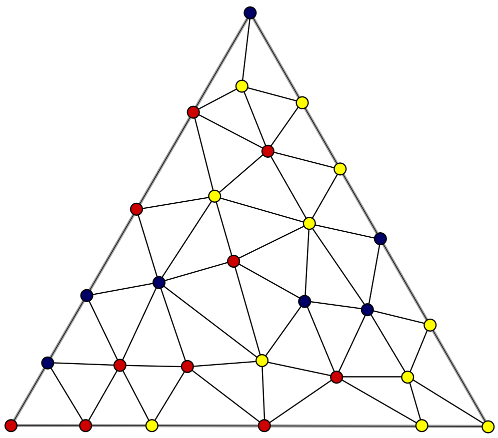
Definition 318. Let x, y, z be three distinct points in $\mathbb{R}^2$. The set $\text{Hull}(\{x,y,z\})$ is a *triangle* having vertices x, y, z, and the segments $\text{Hull}(\{x,y\}), \text{Hull}(\{x,z\}), \text{Hull}(\{y,z\})$ are its *sides*.

Definition 319. Given a triangle $T \subset \mathbb{R}^2$, a set of triangles $\{T_1, \ldots, T_n\}$ is a **triangulation** for T if it fulfills the following constraints:

- $T = \bigcup_{j=1}^{n} T_j$;
- if $i \neq j$ and T_i intersects T_j, the intersection $T_i \cap T_j$ is made by a single vertex, or by a whole side, shared by T_i and T_j.

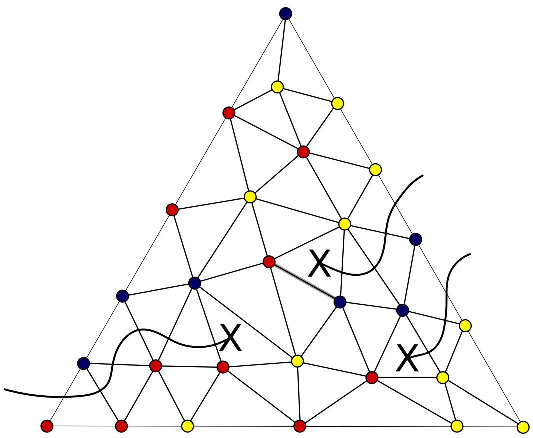
With such assumptions any vertex or side of a T_j is said to be a vertex or side of the triangulation.

Definition 320. Given a triangle T having vertices x_1, x_2, x_3 and a triangulation $S = \{T_1, \ldots, T_n\}$ of it, a 3-*coloring* of the vertices of S is a function f from the vertices of S in the set $\{1,2,3\}$. We say that a 3-coloring of the vertices of S is a *Sperner coloring* if $f(x_j) = j$ and for any vertex z of the triangulation belonging to the side $\text{Hull}(\{x_i, x_j\})$ of T we have $f(z) = f(x_i)$ or $f(z) = f(x_j)$.



An example of Sperner coloring.

Theorem 321 (Sperner)**.** In every Sperner coloring there is an element of the triangulation whose vertices have three distinct colors.



Paths ending in a triangle whose vertices have three distinct colors.

Proof. We may first notice that the number of sides of the triangulation, belonging to $\text{Hull}(\{x_1, x_2\})$ and having endpoints of distinct colors, is odd - we say that such sides are *boundary sides*. Assume, for the moment, that every element of the triangulation has two vertices with the same color. In such a case, every side with distinct-colored endpoints and colors in $\{1,2\}$ shares a vertex with a different side fulfilling the same constraints. Thus it is possible to travel across such sides by starting at a boundary side and ending at a different boundary side. However the number of boundary sides with colors in $\{1,2\}$ is odd, hence at least one path has to "get stuck" at some point, by reaching a triangle with distinct-colored vertices. □

Theorem 322 (Brouwer)**.** If $T \subseteq \mathbb{R}^2$ is a triangle and $f : T \to T$ is a continuous function, f has a fixed point in T, i.e.:

$$\exists\, x \in T : \quad f(x) = x.$$

Before starting an actual proof it is practical to introduce the concepts of *trilinear coordinates* and *mesh* of a triangulation.

Definition 323. Given a triangle $T \subseteq \mathbb{R}^2$ with vertices A, B, C and a point $x \in T$, let us denote

$$d_A(x) = d(x, BC), \qquad d_B(x) = d(x, AC), \qquad d_C(x) = d(x, AB),$$

where the distance of a point from a segment is defined as usual:

$$d(x, AB) = \min_{\lambda \in \mathbb{R}} d(x, \lambda A + (1 - \lambda) B).$$

We say that the *trilinear coordinates* of x are given by:

$$\left(\frac{d_A(x)}{d_A(x) + d_B(x) + d_C(x)}, \frac{d_B(x)}{d_A(x) + d_B(x) + d_C(x)}, \frac{d_C(x)}{d_A(x) + d_B(x) + d_C(x)} \right).$$

A point belonging to T has non-negative trilinear coordinates, whose sum equals 1.

By trilinear coordinates, any continuous function from T in T can be interpreted as a continuous function from $D \subseteq \mathbb{R}^3$ in itself, where:

$$D = \{(x, y, z) \in \mathbb{R}^3 : \ x, y, z \geq 0, \ x + y + z = 1\}.$$

Definition 324. The *mesh* of a triangulation is the maximum side length for the elements of such triangulation.

Definition 325. Given two triangulations S and S', we say that S' *extends* or *refines* S if every element of S is given by the union of some elements in S'.

Proof. Let S be a triangulation of D with mesh ε. Assuming f has no fixed points, we may give a color to every vertex (a, b, c) in S according to $(a', b', c') = f((a, b, c))$:

- if $a' < a$, we give to (a, b, c) the color 1;
- if $a' \geq a$ and $b' < b$, we give to (a, b, c) the color 2;
- if $a' \geq a$ and $b' \geq b$, then $c' < c$, and we give to (a, b, c) the color 3.

We leave to the reader to check this coloring is a Sperner coloring. It follows there is an element $T_j \in S$ whose vertices have three distinct colors. The triangulation S can be extended to a triangulation S' having mesh $\frac{\varepsilon}{2}$-fine: by following the above instructions for assigning colors and repeating the same construction multiple times, we may reach the wanted conclusion through the Bolzano-Weierstrass Theorem. In T_j there is a point t that is the common limit of three distinct sequences of points: the first sequence made by points with color 1, the second sequence made by points with color 2 and the third sequence made by points with color 3. However f is uniformly continuous over D, hence for any $x \in D$ there is a neighbourhood of x with at most two colors. This leads to a contradiction, proving that f surely has a fixed point in D. □

Corollary 326 (Brouwer)**.** If D is a closed disk in $\mathbb{R}^2$ and $f : D \to D$ is a continuous function, f has a fixed point in D.

Proof. Let $T \in \mathbb{R}^2$ be a triangle and let us assume the existence of a continuous and invertible map $g : T \to D$ with a continuous inverse function. Such assumptions grant

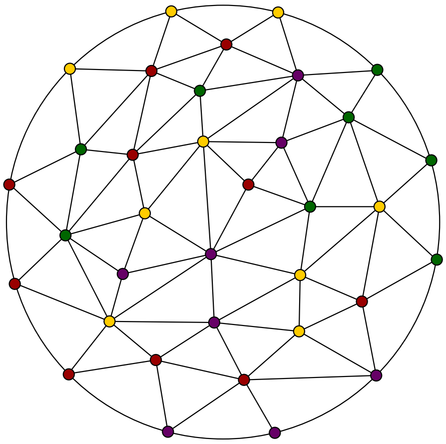
$$(g^{-1} \circ f \circ g) : T \to T$$

has a fixed point $x \in T$, from which $g(x) \in D$ is a fixed point for f.
To prove the existence of the previous map g is an exercise left to the reader. □

Theorem 327 (Tucker)**.** Let $P_1, \ldots, P_{2n} \in \partial B(0,1) \subseteq \mathbb{R}^2$ such that $-P_j = P_{n+j}$ for any $j \in [1,n]$. Let S be a triangulation of $E = \text{Hull}(\{P_1, \ldots, P_{2n}\})$ and f a 4-coloring of the vertices of S, with colors given by $\{-2,-1,+1,+2\}$, such that:

$$\forall j \in [1,n], \quad f(P_j) = -f(P_{n+j}).$$

With such assumptions, a side of S has opposite-colored endpoints.



A Tucker coloring: red is opposite to green and yellow is opposite to purple.

Proof. If the sides forming ∂E never have opposite-colored endpoints, on the boundary of E there are an odd number of segments whose endpoints have colors -1 and $+2$. Assuming that in $S \setminus \partial E$ we do not have segments with opposite-colored endpoints, every element of S with a side colored by -1 and $+2$ has another side fulfilling the same constraints: travelling across such sides by starting from the exterior of E we get that some path necessarily stops at the interior of E. □

Corollary 328. Every continuous function $f : D = \bar{B}(0,1) \to \mathbb{R}^2$ such that

$$\forall\, x \in \partial D, \quad f(x) = -f(-x)$$

has a zero in D.

Assuming that f is non-vanishing on D, the function $g : D \to S^1$ defined by

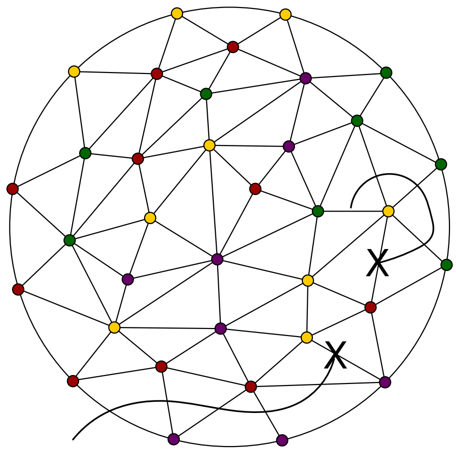
$$g(x) = \frac{f(x) - f(-x)}{\|f(x) - f(-x)\|}$$

is continuous and fulfills the constraint

$$\forall\, x \in \partial D, \quad g(x) = -g(-x).$$

It follows we may give to every point of $(a,b) \in D$ a color in the set $\{-2,-1,+1,+2\}$ according to $(c,d) = g((a,b))$:

- if $d \geq 0$ and $c > 0$, we give to (a,b) the color $+1$;
- if $d > 0$ and $c \leq 0$, we give to (a,b) the color $+2$;
- if $d < 0$ and $c \leq 0$, we give to (a,b) the color -1;
- if $d \leq 0$ and $c > 0$, we give to (a,b) the color -2.



The proofs of Sperner's Theorem and Tucker's Theorem are *very* similar.

Considering $2n$ points on ∂D pairwise symmetric with respect to the origin and a triangulation with mesh ε of their convex hull, the coloring given by the above algorithm fulfills the hypothesis of Tucker's Theorem. It follows we have two distinct points $z, w \in D$ such that

$$\|z - w\| \leq \varepsilon, \qquad \|g(z) - g(w)\| \geq \sqrt{2},$$

leading to a contradiction by the uniform continuity of g on D. Thus f has a fixed point in D.

Theorem 329 (Borsuk-Ulam). If $f : S^2 \to \mathbb{R}^2$ is a continuous function, there exist two antipodal points on S^2 where the function f attains the same value.

Proof. Assume, by contradiction, that for any $x \in S^2 = \partial B(0,1) \subseteq \mathbb{R}^3$ we have $f(x) \neq f(-x)$. There exists a continuous and invertible function g from $S^2 \cap \{z \geq 0\}$ in $D^2 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$, which coincides with the identity function when restricted to $S^1 = S^2 \cap \{z = 0\}$:

$$g\left(\left(x, y, \sqrt{1-(x^2+y^2)}\right)\right) = (x,y).$$

As a consequence, there exists $h : D^2 \to \mathbb{R}^2$, defined by:

$$h(x) = f(g^{-1}(x)) - f(-g^{-1}(x)),$$

which is continuous, non-vanishing and fulfilling $h(x) = -h(-x)$ on the points of $\partial D^2 = S^1$. However, such a function cannot exist by the previous result. □

Corollary 330. At every moment, there exist two antipodal points on the surface of Earth having the same temperature and the same pressure.

Theorem 331 (Lusternik-Schnirelmann). If C_1, C_2, C_3 are three closed sets covering S^2, at least one of them contains a couple of antipodal points.

Proof. Since the functions $d_i : S^2 \to \mathbb{R}$, with $i \in \{1,2,3\}$, defined by:

$$d_i(x) = \min_{y \in C_i} \|x - y\|$$

are continuous, the function $f : S^2 \to \mathbb{R}^2$ definined by:

$$f(x) = (d_1(x), d_2(x))$$

is continuous as well. Due to Borsuk-Ulam's Theorem there exists some $x \in S^2$ such that:

$$(d_1(x), d_2(x)) = f(x) = y = f(-x) = (d_1(-x), d_2(-x)).$$

If y has its first or second component equal to zero, then both x and $-x$ belong to C_1 or C_2, respectively. If y does not have any component equal to zero, then neither x or $-x$ belong to C_1 or C_2, hence they necessarily belong to C_3, since:

$$S^2 = \bigcup_{i=1}^{3} C_i.$$

□

Exercise 332. Prove that Lusternik-Schnirelmann's Theorem is *equivalent* to Borsuk-Ulam's Theorem.

Surprisingly, these quite deep results in Topology have interesting applications in Combinatorics. If we denote by $[n]$ the set $\{1, 2, \ldots, n\}$ and by $\binom{[n]}{k}$ the subsets of $[n]$ with k elements, the **Kneser graph** $\mathrm{KG}(n,k)$ is the graph admitting $\binom{[n]}{k}$ as vertex set and having edges only between nodes associated with disjoint subsets. Given a finite, undirected graph G, its **chromatic number** $\chi(G)$ is defined as the minimum number of colors required to give a color to every vertex of G, in such a way that neighbours have distinct colors.

Theorem 333 (Kneser's Conjecture / Lovasz Theorem).

$$\chi(\mathrm{KG}(n,k)) = n - 2k + 2.$$

It is not difficult to prove that $\chi(\mathrm{KG}(n,k)) \leq n-2k+2$, for instance by giving to any vertex $v \in V$ associated with a subset F the color defined by $\min(F \cup \{n-2k+2\})$. With such assumptions, if F_1 and F_2 have the same color $i < (n-2k+2)$, then $i \in F_1 \cap F_2$, hence $(F_1, F_2) \in E$. On the other hand, if both F_1 and F_2 have color $(n-2k+2)$, both of them are subsets of $\{n-2k+2, n-2k+2, \ldots, n\}$, which is a set with cardinality $2k-1$. As a consequence, F_1 and F_2 have a non-empty intersection. Laszlo Lovasz proved Kneser's conjecture in 1978 by using the Borsuk-Ulam theorem, and his proof has been greatly simplified in 2002 by J. Greene (just a college student at the time). Le us set $d = n-2k+1$: we wish to prove that $\chi(\mathrm{KG}(n,k)) > d$. Let $X \subseteq S^d$ be a set of n points *in general position* (i.e. such that at most d points of X belong to a hyperplane through the origin) and let us consider X and $[n]$ as the same object. Assuming to have a valid (in the chromatic number sense) coloring of $\mathrm{KG}(n,k)$ using d colors, let us define, for any $x \in S^d$, $H(x)$ as the open hemisphere made by points y such that $\langle x, y\rangle > 0$. Additionally, let us define a covering $A_1, A_2, \ldots, A_{d+1}$ of S^d as follows: for any $i = 1, 2, \ldots, d$ we let

$$A_i \stackrel{\text{def}}{=} \left\{ x : \exists F \in \binom{X}{k} \text{ with color } i \text{ such that } F \subseteq H(x) \right\}$$

and define A_{d+1} as $S^d \setminus \bigcup_{i=1}^d A_i$. Invoking the Lusternik-Schnirelmann Theorem, there exists some $x \in S^d$ such that both x and $-x$ belong to A_i for some $i \in [d+1]$. If $i \leq d$, $H(x)$ and $H(-x)$ contain, respectively, sets F_1 and F_2 of the same color, i, but since $H(x)$ and $H(-x)$ are disjoint, F_1 and F_2 are disjoint, hence they cannot have the same color, since there is an edge between them. If $i = d+1$, then both x and $-x$ belong to A_{d+1}, hence $H(x)$ contains at most $(k-1)$ points from X (otherwise $H(x)$ would contain some F with color $j \leq d$ and x would belong to A_j and not to A_{d+1}). In a similar way, $H(-x)$ contains at most $(k-1)$ points from X, hence $S^d \setminus (H(x) \cup H(-x))$, a subset of the hyperplane $\langle x, y\rangle = 0$, contains at least $n-2k+2 = d+1$ punti, which contradicts the previous "*general position*" assupmtion. We may notice that in order to plane n points on S^n in general position it is enough to follow a probabilistic approach: a random configuration (with respect to the uniform probability distribution on S^n) is almost surely in general position. A deterministic approach is to exploit the **moment curve** $t \to (1, t, \ldots, t^d)$: it is simple to prove by Ruffini's rule that d distinct points on the curve, together with the origin, never lie on the same hyperplane. By normalizing these d points one gets d points in general position on S^{d-1}.

Corollary 334. There are triangle-free graphs with an arbitrarily large chromatic number.

This results is usually proved by invoking **Mycielski**'s construction: let $G = ([n], E)$ be a triangle-free graph and let us consider the graph having vertex set $\{1, 2, \ldots, 2n, 2n+1\}$ and the following set of edges:

- an edge between j and k if (j,k) is an edge in G;
- an edge between $n+\tau$ and k if (τ, k) is an edge in G;
- an edge between $2n+1$ and $n+\tau$ for any $\tau \in [n]$.

This is a triangle-free graph and its chromatic number equals $\chi(G)+1$.
Kneser graphs provide an alternative construction: if $n < 3k$, $\mathrm{KG}(n,k)$ is triangle-free, hence

$$\mathrm{KG}(3M-4, M-1)$$

for any $M \geq 3$, is a triangle-free graph with chromatic number M by Lovasz' Theorem.
Kneser graphs play a crucial role in another remarkable result:

Theorem 335 (Erdös-Ko-Rado)**.** If A is a family of subsets with cardinality r of $\{1, 2, \ldots, n\}$, with $n \geq 2r$, and every couple of elements of A has a non empty intersection, then:

$$|A| \leq \binom{n-1}{r-1}.$$

In terms of Kneser graphs, the Erdös-Ko-Rado Theorem states that $\binom{n-1}{r-1} = \alpha(\mathrm{KG}(n,k))$ is the maximum number of elements in a **independent subset** of $\mathrm{KG}(n,k)$, i.e. a subset of the vertex set in which there are no edges from the original graph. The proof we are going to outile is due to Katona (1972). Let us assume to arrange the elements of $\{1, 2, \ldots, n\}$ in a cyclic ordering, and to consider the intervals with length r in this ordering. For instance, assuming $n = 8$ and $r = 3$, the cyclic ordering $(3, 1, 5, 4, 2, 7, 6, 8)$ produces the intervals:

$$(3, 1, 5),\ (1, 5, 4),\ (5, 4, 2),\ (4, 2, 7),\ (2, 7, 6),\ (7, 6, 8),\ (6, 8, 3),\ (8, 3, 1).$$

However it is not possible that all these intervals belong to A, since many of them are disjoint. The key observation of Katona is that at most r of these intervals may belong to A, disregarding the considered cyclic ordering. Indeed, if $(a_1, a_2, \ldots, a_r)$ is an interval belonging to A, any other interval of the same cyclic ordering belonging to A has to **separate** a_i and a_{i+1} for some i, i.e. it has to contain exactly one of these two elements. The two intervals separating these two elements are disjoint, hence at most one of them may belong to A. As a consequence, the number of intervals belonging to A is at most 1 plus the number of "separable couples", i.e. $\leq 1 + (r-1) = r$. Now we may count the number of couples (S, C) where S is a set in A and C is a cyclic ordering for which S is an interval. On one hand, for any S it is possible to generate C by choosing the $r!$ permutations of S and the $(n-r)!$ permutations of the remaining elements. In particular the number of the previous (S, C) couples is $|A|r!(n-r)!$. On the other hand, there are $(n-1)!$ cyclic orderings, and any of them produces at most r intervals belonging to A, hence the number of the previous (S, C) couples is at most $r(n-1)!$. By *double counting*:

$$|A|r!(n-r)! \leq r(n-1)! \quad \longleftrightarrow \quad |A| \leq \binom{n-1}{r-1}.$$

Theorem 336 (Lubell-Yamamoto-Meshalkin)**.** Let $\mathcal{A}$ be a family of subsets of $\{1, 2, \ldots, n\}$, with the property that two distinct elements $A_1, A_2 \in \mathcal{A}$ never fulfill $A_1 \subseteq A_2$ or $A_2 \subseteq A_1$. By denoting as τ_k the number of elements in $\mathcal{A}$ with cardinality k, we have:

$$\sum_{k=0}^{n} \frac{\tau_k}{\binom{n}{k}} \leq 1.$$

The proof of this remarkable result is very similar to the outlined proof of the Erdös-Ko-Rado Theorem. A family of subsets fulfilling the hypothesis is also said to be a *Sperner family*, and it is an **antichain** in the partial ordering of the subsets of $\{1, 2, \ldots, n\}$ induced by $\subseteq$. If we consider a random ordering $(x_1, x_2, \ldots, x_n)$ of $[n]$ and the sets $C_0 = \emptyset, \ldots, C_k = \{x_1, \ldots, x_k\}$, every couple of elements of $\mathcal{C} = \{C_0, C_1, \ldots, C_n\}$ fulfills $\subseteq$ or $\supseteq$, hence the expected value of $|\mathcal{C} \cap \mathcal{A}|$ certainly is ≤ 1. On the other hand, for any $A \in \mathcal{A}$ the probability that A belongs to the random **chain** $\mathcal{C}$ is exactly $\binom{n}{|A|}^{-1}$, since $\mathcal{C}$ contains just one set with cardinality $|A|$, and the probability is uniform over any element of $\mathcal{A}$. In particular:

$$1 \geq \mathbb{E}[|\mathcal{A} \cap \mathcal{C}|] = \sum_{A \in \mathcal{A}} \mathbb{P}[A \in \mathcal{C}] = \sum_{A \in \mathcal{A}} \frac{1}{\binom{n}{|A|}} = \sum_{k=0}^{n} \frac{\tau_k}{\binom{n}{k}}.$$

Corollary 337 (Sperner)**.** Given the set $\{1, 2, \dots, n\}$, the Sperner family with the greatest number of elements is the one made by the sets with cardinality $\lfloor n/2 \rfloor$.

Hall's Theorem is also known as Hall's *marriage* Theorem, since it is usually presented through this formulation [12]: *let us suppose to have n women and n men, and that every woman has her personal set of* potential partners *among the previous n men. Which conditions ensure it is possible to declare them all* wife and husband*, without marrying any woman with someone not in her list of candidates?* Hall's Theorem states that necessary and sufficient conditions are provided by having no subset of women with a joint list of candidates that is too small.

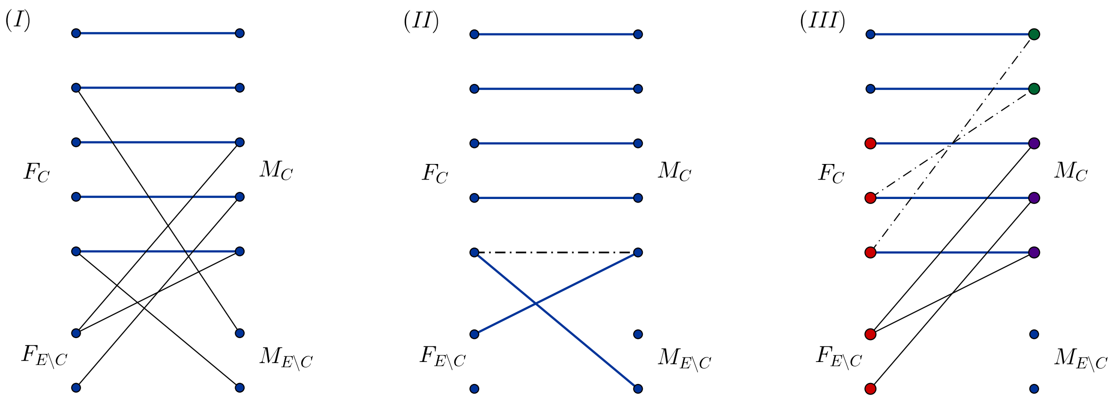
Definition 338. We say that a graph $G = (V, E)$ admits a **perfect matching** if there exists $E' \subseteq E$ such that the elements of E' do not have any common vertex and every vertex $v \in V$ is an endpoint for some edge in E'.

Theorem 339 (Hall's marriage Theorem, 1935)**.** Let us assume that $G = (V, E)$ is a balanced bipartite graph, i.e. $V = F \cup M$, $F \cap M = \emptyset$, $|F| = |M|$ and every edge in G joins a vertex in F with a vertex in M. G has a perfect matching if and only if for any non-empty subset $F' \subseteq F$, the number of neighbours of F' in M is $\geq |F'|$.

Proof. Part of the claim is trivial: if a perfect matching exists, clearly every non-empty subset $F' \subset F$ has a neighbourhood in M whose cardinality is $\geq |F'|$.
The converse implication is not trivial, but it is a consequence of Dilworth's Theorem. We are going to approach the proof of Hall's Theorem through a **maximality** argument. Let us assume $|F| = |M| = n$ and that the hypothesis of Hall's Theorem are fulfilled, but no perfect matching exists. We may consider $C \subseteq E$ that is a *coupling*, i.e. a set of disjoint edges, with the maximum cardinality, then denote through $F_C, F_{E\setminus C}$, respectively, the sets of elements of F reached / not reached by C. In a similar way we may define M_C and $M_{E\setminus C}$. We may notice that every element of E has an endpoint in C, otherwise C would not be maximal. There are at least $|F_{E\setminus C}|$ edges of G with an endpoint in $F_{E\setminus C}$, and for all of them the other enpoint has to lie in M_C, again by maximality of C. In particular $|M_C| \geq |F_{E\setminus C}|$ which implies $|M_C| = |F_C| \geq \frac{n}{2}$. On the other hand, all the edges of C with an endpoint in F_C have the other endpoint in M_C. In order that F fulfills the given hypothesis, there have to be at least $|M_{E\setminus C}|$ edges of G joining F_C and $M_{E\setminus C}$. If $f \in F_C$ and $m \in M_C$ are joined by an edge in C and, additionally, f is joined with an element of $M_{E\setminus C}$ by $e_1 \in E$ and m is joined with an element of $F_{E\setminus C}$ by $e_2 \in E$, by removing from C the edge (m, f) and inserting the edges e_1, e_2 we get a coupling with a greater cardinality, violating the maximality of C. Since that cannot happen, we may define M'_C as the neighbourhood of $F_{E\setminus C}$ in M_C and consider that C induces a bijection between M'_C and a subset of F_C. Let us denote such subset as F'_C. If $F_{E\setminus C}$ is non-empty and $F_{E\setminus C} \cup F'_C$ meets the hypothesis, the neighbourhood of F'_C in M_C contains at least one element not belonging to M'_C. If we denote as M''_C the last neighbourhood and repeat the same construction, we ultimately get that there are no edges of E between F_C and $M_{E\setminus C}$, but in such a case F does not meet the hypothesis. It follows that $F_{E\setminus C}$ is empty and C is a perfect matching. □

[12] Truth to be told, the *very* usual formulation has the roles of women and men exchanged. It does not really matter... or does it?

A summary of the key steps in the proof of Hall's Theorem.
The elements of C are blue, the elements of $E \setminus C$ are black.



- (I) : C induces a bijection between F_C and M_C, but there have to be some edge of E joining $F_{E\setminus C}$ and M_C, together with some edge of E joining $M_{E\setminus C}$ and F_C;
- (II) : the depicted situation cannot happen, since otherwise we could remove an edge from C and insert two edges, getting a larger coupling;
- (III) : on the other hand the neighbourhood of the red points cannot be made by purple points only, since otherwise the set of red points would violate the hypothesis.

Exercise 340. Exploiting Hall's Theorem, prove that if we remove two opposite-colored squares from a $2n \times 2n$ chessboard, the remaining part can be tiled by dominoes like ◻◻ or its 90°-rotated version.

Exercise 341 (Birkhoff)**.** We say that a $n \times n$ matrix is **doubly stochastic** if its entries are non-negative and the sum of the entries along every row or column equals one. Prove that any doubly stochastic matrix can be written as a convex combination of permutation matrices.

Theorem 342 (Erdös-Szekeres)**.** From each sequence of $mn+1$ real numbers it is possible to extract a weakly increasing subsequence with $n+1$ terms, or a weakly decreasing subsequence with $m+1$ terms.

Proof. The proof we are going to outline proceeds in a algorithmic fashion: let us assume that $a_1, a_2, \ldots, a_{mn+1}$ is the main sequence and that we have n queues in a post office, initially empty, where to insert the elements of the main sequence one by one. For any $i \in \{1, \ldots, \}$ we act as follows: a_i visits the first queue and if it is empty, or the value of a_i is greater than the value of the last entry of the queue, a_i takes place in such a queue, otherwise it goes to the next queue and behaves in the same way. If, at some point, some element a of the original sequence is not able to take place in any queue, the subsequence given by the last entries of each queue, together with a, provides a weakly decreasing subsequence with $n+1$ terms. Conversely, if every element of the original sequence is able to take place somewhere, at the end of the process there will be $mn+1$ entries in n queues, hence at least a queue with $m+1$ entries by the Dirichlet box principle, with such queue providing an increasing subsequence with $m+1$ terms. □

The Erdös-Szekeres Theorem can be exploited to outline a "quantitative" proof of the Bolzano-Weierstrass Theorem. If $\{a_n\}_{n\geq 1}$ is a sequence of real numbers belonging to the interval $[a,b]$, the Erdös-Szekeres Theorem allows us to extract a weakly monotonic subsequence. Since every bounded and weakly monotonic sequence is convergent to its infimum or supremum, every bounded sequence has a convergent subsequence.
The Erdös-Szekeres Theorem has been the key for proving the following statement, too:

Theorem 343 (Happy Ending Problem[13])**.** Given some positive integer N, in every set of distinct points in $\mathbb{R}^2$ with a sufficiently large cardinality there are the vertices of a convex N-agon.

By denoting as $f(N)$ the minimum cardinality granting the existence of a convex N-agon, up to 2015 the sharpest bound for $f(N)$ was

$$\forall N\geq 7,\qquad f(N)\leq\binom{2N-5}{N-2}+1=O\left(\frac{4^N}{\sqrt{N}}\right),$$

but in 2016 Suk and Tardos presented an article (currently under revision) where they claimed $f(N)=2^{N+O(\sqrt{N\log N})}$, i.e. an almost-optimal bound, since it is not difficult to prove that $f(N)\geq 1+2^{N-2}$. The Erdös-Szekeres can also be exploited to prove the following statement (even if "the usual way" is to do just the opposite):

Theorem 344 (Dilworth-Mirsky)**.** Given a partial ordering on a finite set, the maximum length of a chain equals the minimum number of chains in which the original set can be partitioned.

The last part of this section is dedicated to a famous application of the **combinatorial compactness** principle: in some cases it is possible to prove statements for finite sets by deducing them from statements involving infinite sets.

Theorem 345 (Ramsey, 1930)**.** For any $n\geq 1$ there exists an integer $R(n)$ such that, given any 2-coloring of the edges of $K_{R(n)}$, there exists a monochromatic subgraph K_n.

We are going to prove such claim by starting with its infinite version:

Theorem 346 (Ramsey, infinite version)**.** Let $k,c\geq 1$ be integers, and let V be an infinite set. Given any coloring of $\binom{V}{k}$ with c colors, there exists an infinite subset $U\subseteq V$ such that $\binom{U}{k}$ is monochromatic.

Proof. We proceed by induction on k: if $k=1$ the claim is trivial.
So we may assume that $k\geq 2$ and $\binom{V}{k}$ is colored through c colors, i.e. there exists a map $\chi:\binom{V}{k}\to C$ where $|C|=c$. We wish to construct an infinite sequence $V_0\supset V_1\supset V_2\supset\ldots$ of infinite subsets of V, together with an infinite sequence of elements $x_0,x_1,x_2,\ldots$ such that, for any i:

- (I) $x_i\in V_i$ and $V_{i+1}\subseteq V_i\setminus\{x_i\}$;
- (II) the k-subsets of the form $\{x_i\}\cup Y$, as Y ranges over V_{i+1}, all have the same color $c(i)$.

[13] Erdös gave such a name to this problem since it someway contributed to making George Szekeres and Esther Klein a married couple.

We start by setting $V_0 = V$ and we pick some $x_0 \in V$. Let us assume to already have $V_0, V_1, \ldots, V_i$ and $x_0, x_1, \ldots, x_i$. We want to construct a couple (V_{i+1}, x_{i+1}): for such a purpose we introduce an auxiliary coloring χ_i, which assigns to any $Y \in \binom{V_i \setminus \{x_i\}}{k-1}$ the color of the k-subset $\{x_i\} \cup Y$ in the original coloring χ, i.e.:

$$\chi_i : \binom{V_i \setminus \{x_i\}}{k-1} \to C$$
$$Y \to \chi(\{x_i\} \cup Y).$$

This construction produces a c-coloring χ_i of $\binom{V_i \setminus \{x_i\}}{k-1}$. Since $V_i \setminus \{x_i\}$ is infinite, the inductive hypothesis ensures the existence of an infinite subset V_{i+1} which is $(k-1)$-monochromatic. Additionally we get that, with respect to the original coloring χ, all the k-subsets of $V_{i+1} \cup \{x_i\}$ containing the element x_i have the same color, which we denote as $c(i)$. Let x_{i+1} be any element of V_{i+1}. The sequences produced by our algorithm clearly fulfill (I) and (II). Since c is finite, the Dirichlet box principle ensures the existence of an infinite subset $A \subset \mathbb{N}$ such that for any $i \in A$ the color $c(i)$ is always the same. By letting $U = \{x_i : i \in A\}$ we get an infinite subset of V that is k-monochromatic as wanted. ☐

If we restrict our attention to graphs (i.e. to the $k = 2$ case), the Theorem just proved can be also stated in the following way: *let G be a graph with an infinite vertex set V, such that for any infinite subset $X \subset V$ there is an edge of V whose endpoints belong to X. Then G contains an infinite complete subgraph.* In order to prove the finite version, it is enough to invoke the following Lemma, which can be informally stated as *in a infinite tree there is an infinite branch.*

Lemma 347 (König's infinity Lemma)**.** Let $G = (V, E)$ a graph with an infinite vertex set. Let $V = \bigcup_{i \geq 0} V_i$, where $V_0, V_1, \ldots$ are finite, non-empty and disjoint sets. Let us assume there is a map $f : V \setminus V_0 \to V$ such that for any $v \in V_i$ with $i \geq 1$, $f(v) \in V_{i-1}$ and $\{v, f(v)\} \in E$. Then there exists an infinite sequence $v_0, v_1, v_2, \ldots$ such that $v_i \in V_i$ and $v_i = f(v_{i+1})$ for any $i \geq 0$.

Proof. Let us denote through $\mathcal{C}$ the set of "f -paths", i.e. the set of paths of the form

$$v, f(v), f^2(v), \ldots, f^{s(v)}(v),$$

terminating at some element $f^{s(v)}(v) \in V_0$. Any $v \in V$ is a vertex for an element of $\mathcal{C}$ and every f-path is finite, hence $\mathcal{C}$ has to contain an infinite number of paths. Since V_0 is a finite set, there exists an infinite subset $\mathcal{C}_0 \subseteq \mathcal{C}$ made by paths terminating at some $v_0 \in V_0$. Since V_1 is a finite set, there exists at least one vertex $v_1 \in V_1$ such that the subset $\mathcal{C}_1 \subseteq \mathcal{C}_0$, made by paths through v_1, is infinite. By repeating the same argument we easily get an infinite sequence $v_0, v_1, \ldots$ meeting the wanted constraints. ☐

We are ready to deduce the finite version of Ramsey's Theorem from its infinite version.

For any $r \in \mathbb{N}^+$, let us denote as I_r the set $\{1, 2, \ldots, r\}$. Let us assume the existence of a triple (k, c, n) such that, for any integer $r \geq k$, there exists a c-coloring $\gamma_r : \binom{I_r}{k} \to |C|$, with $|C| = c$, of the k-subsets of I_r, such that no n-subset of I_r is k-monochromatic (i.e. for any $X \subseteq I_r$ with cardinality n, there exist Y and Y' in $\binom{X}{k}$ such that $\gamma_r(Y) \neq \gamma_r(Y')$).
Denoting through V_r the set of this "faulty" c-colorings γ_r of I_r, every V_r is finite and non-empty. Let us set $V = \bigcup_{r \geq k} V_r$. For $r \geq k$ and $\gamma \in V_{r+1}$, let us define $f(\gamma)$ as the restriction of γ to I_r. $f(\gamma)$ then is an element of V_r (the restriction of a faulty coloring still is a faulty coloring). By considering $G = (V, E)$, where the edge set is given by $E = \{\{\gamma, f(\gamma)\} : \gamma \in V \setminus V_k\}$, we are in the hypothesis of König's infinity Lemma. It follows the existence of a coloring of $\binom{\mathbb{N}}{k}$ which is faulty when restricted to any I_r. This contradicts the statement of Ramsey's theorem in its infinite version.

Some remarkable results in Infinite Combinatorics.

Theorem 348 (Schur)**.** For any coloring of $\mathbb{N}$ with a finite number of colors, there exist infinite triples of the form $(a, b, a+b)$ which are monochromatic.

Theorem 349 (Van Der Waerden)**.** For any coloring of $\mathbb{N}$ with a finite number of colors, there exist monochromatic arithmetic progressions of arbitrary length.

Theorem 350 (Hales-Jewett)**.** For any coloring of $\mathbb{Z}^d$ with a finite number of colors, there exist infinite monochromatic sets of the form $P_1 \times P_2 \times \ldots \times P_d$, in which every P_k is an arithmetic progression in $\mathbb{Z}$.

Theorem 351 (Szemeredi)**.** The claim of Van Der Waerden's Theorem still holds if $\mathbb{N}$ is replaced by a subset of $\mathbb{N}$ with positive asymptotic density.

Theorem 352 (Green-Tao)**.** The set of prime numbers contains arithmetic progressions with arbitrary length.

Exercise 353. Let us consider a graph G having $\mathbb{N}$ as its vertex set and edges joining a and b iff $|a-b|$ is a number of the form $m!$ for some $m \in \mathbb{N}^+$. Prove that the chromatic number of G equals 4.

Exercise 354 (Sarközy)**.** Let us consider a graph G having $\mathbb{N}$ as its vertex set and edges joining a and b iff $|a-b|$ is a number of the form m^2 for some $m \in \mathbb{N}^+$. Prove that the chromatic number of G is $+\infty$. Hint: use a refined version of Van Der Corput's trick to show that in any subset of $\mathbb{N}$ with positive asymptotic density, there are two distinct elements whose difference is a square.

14 Continued fractions and elements of Diophantine Approximation

A (ordinary) **continued fraction** is an object of the following form [14]:

$$a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \ldots}},$$

where a_0 is a non-negative integer and every a_j with $j \geq 1$ is a positive integer.
For the sake of brevity and readability, the following compact notations are used very often:

$$[a_0; a_1, a_2, \ldots, a_n] \quad \text{or} \quad a_0 + \frac{1}{a_0+}\,\frac{1}{a_0+}\ldots\frac{1}{a_n}.$$

The numbers a_j are said *terms* or *partial quotients* of the continued fraction, while the following rational numbers

$$q_0 = a_0, \quad q_1 = [a_0; a_1], \quad q_2 = [a_0; a_1, a_2], \quad \ldots$$

are said *convergents* of the continued fraction. We may notice that:

$$[a_0; a_1, \ldots, a_n, 1] = [a_0; a_1, \ldots, a_n + 1],$$

hence from now on we will assume that the last term of a ordinary continued fraction (associated with an element of $\mathbb{Q}^+ \setminus \mathbb{N}$) is never 1, in order to preserve the unicity of the representation.

Theorem 355. Every rational number $q \in \mathbb{Q}^+$ has a unique continued fraction representation.

We may simply consider the orbit of q with respect to the map $\varphi : x \mapsto \frac{1}{x - \lfloor x \rfloor}$.
For instance, by starting at $q = \frac{24}{7}$ we have:

$$\frac{24}{7} = 3 + \frac{3}{7}, \qquad \frac{7}{3} = 2 + \frac{1}{3},$$

from which:

$$\frac{24}{7} = 3 + \frac{1}{2 + \frac{1}{3}} = [3; 2, 3].$$

We may notice that the fractional parts involved in the process are rational numbers < 1 and their numerators form a strictly decreasing sequence: if at some point we are dealing with a fractional part $\frac{a}{b}$ with $a < b$, at the next step the numerator of the new fractional part is $b - \left\lfloor \frac{b}{a} \right\rfloor a < a$. It follows that at some point the involved fractional part has a unit numerator and the previous algorithm stops, producing a continued fraction representation for q.

Lemma 356. If two continued fractions with $n+1$ terms $[a_0; a_1, \ldots, a_n] = q_a$ and $[b_0; b_1, \ldots, b_n] = q_b$ share the first n terms and $a_n > b_n$, then $q_a > q_b$ or $q_a < q_b$ according to the parity of n.

The result can be simply proved by induction on n: the base case $n = 0$ is trivial and the inequality

$$[c_0; c_1, \ldots, c_{n-1}, c] < [c_0; c_1, \ldots, c_{n-1}, d]$$

is equivalent, up to substracting c_0 from both terms, to the inequality:

$$[0; c_1, \ldots, c_{n-1}, c] < [0; c_1, \ldots, c_{n-1}, d],$$

or, by considering the reciprocal of both the RHS and the LHS, to the inequality:

$$[c_1; \ldots, c_{n-1}, c] > [c_1; \ldots, c_{n-1}, d],$$

[14] Temporarily, we will assume to deal with a finite number of nested ratios. We will see soon that this assumption can be dropped, and dropping such assumption leads to an algorithmic way for representing *every* element of $\mathbb{R}^+$ as a continued fraction.

whose truth follows from the inductive hypothesis. In general, two continued fractions which differ by at least one term represent two distinct rational numbers. Let us assume to have:

$$[a_0; a_1, \ldots, a_n, b_0, \ldots, b_m] = [a_0; a_1, \ldots, a_n, c_0, \ldots, c_r],$$

with $b_0 \neq c_0$. By subtracting a suitable integer number from both sides and reciprocating $n+1$ times we get:

$$[b_0; \ldots, b_m] = [c_0; \ldots, c_r]$$

which leads to a contradiction, since by applying $x \mapsto \lfloor x \rfloor$ to both sides of the last supposed identity we get two distinct integer numbers. This proves that any positive rational number can be represented in a unique way as a continued fraction. We may also notice that the process leading to the construction of the terms of the continued fraction of $\frac{a}{b}$, where a and b are coprime positive integers, is analogous to the application of the extended Euclidean algorithm for computing $\gcd(a, b)$. Since a and b are coprime by assumption, $\gcd(a, b) = 1$ and the Euclidean algorithm stops in a finite number of steps, producing a finite continued fraction representing $\frac{a}{b}$.

Theorem 357. Let h_m and k_m, respectively, the numerator and denominator of the m-th convergent of the continued fraction $[a_0; a_1, \ldots, a_n]$. For any integer $x \geq 1$ we have:

$$[a_0; a_1, \ldots, a_n, x] = \frac{x\, h_n + h_{n-1}}{x\, k_n + k_{n-1}}.$$

Proof. We may prove the statement by induction on n: the base case $n = 2$ is simple to check. Let p_m and q_m be the numerator and denominator of the m-th convergent of $[a_1; \ldots, a_n, x]$. We have:

$$[a_0; a_1, \ldots, a_n, x] = a_0 + \frac{q_n}{p_n} = \frac{a_0 p_n + q_n}{p_n},$$

but the inductive hypothesis ensures $p_n = x \cdot p_{n-1} + p_{n-2}$ and $q_n = x \cdot q_{n-1} + q_{n-2}$, from which:

$$[a_0; a_1, \ldots, a_n, x] = \frac{x(a_0 p_{n-1} + q_{n-1}) + (a_0 p_{n-2} + q_{n-2})}{x k_n + k_{n-1}} = \frac{x\, h_n + h_{n-1}}{x\, k_n + k_{n-1}}.$$

□

Theorem 358. If $\frac{h_m}{k_m}$ and $\frac{h_{m+1}}{k_{m+1}}$ are consecutive convergents of the same continued fraction,

$$h_{m+1}\, k_m - h_m\, k_{m+1} = (-1)^m.$$

Proof. Let $A_m = h_{m+1}\, k_m - h_m\, k_{m+1}$. By the previous Theorem,

$$h_{m+1} = a_m h_m + h_{m-1}, \qquad k_{m+1} = a_m k_m + k_{m-1},$$

from which $A(m) = -A(m-1)$ immediately follows. □

In particular, continued fraction can be exploited for computing the multiplicative inverse of any $a \in \mathbb{F}_p^*$:

Exercise 359. Compute the inverse of 35 modulus 113 through continued fractions.

Proof. We have that:

$$\frac{113}{35} = [3;4,2,1,2], \qquad [3;4,2,1] = \frac{42}{13}$$

hence the absolute value of the difference between $\frac{113}{35}$ and $\frac{42}{13}$ equals $\frac{1}{35\cdot 13}$. In particular:

$$113 \cdot 13 - 42 \cdot 35 = -1,$$

hence the inverse of 35 in $\mathbb{Z}/(113\mathbb{Z})^*$ is 42. □

Corollary 360. If the continued fraction of $q \in \mathbb{Q}^+$ has n terms, by denoting as k_m the denominator of the m-th convergent we have:

$$q = a_0 - \sum_{j=1}^{n} \frac{(-1)^j}{k_{j-1}k_j}.$$

Proof. Due to the previous Theorem we have:

$$q = \frac{h_n}{k_n} = a_0 + \sum_{j=1}^{n}\left(\frac{h_j}{k_j} - \frac{h_{j-1}}{k_{j-1}}\right) = a_0 + \sum_{j=1}^{n} \frac{A(j-1)}{k_{j-1}k_j}.$$

□

Corollary 361. If the continued fraction of $q \in \mathbb{Q}^+$ has at least $k+2$ terms and $\frac{p_k}{q_k}$ is its k-th convergent, we have:

$$\frac{1}{q_k(q_k + q_{k+1})} \leq \left|q - \frac{p_k}{q_k}\right| < \frac{1}{q_k\, q_{k+1}}.$$

Proof. As a consequence of the Corollary (360) q certainly is between $\frac{p_k}{q_k}$ and $\frac{p_{k+1}}{q_{k+1}}$, hence:

$$\left|q - \frac{p_k}{q_k}\right| < \left|\frac{p_{k+1}}{q_{k+1}} - \frac{p_k}{q_k}\right| = \frac{1}{q_{k+1}q_k},$$

where the last identity follows from the Theorem (358). On the other hand, if a, b, c, d are four positive integers, the ratio $\frac{a+c}{b+d}$ certainly is between $\frac{a}{b}$ and $\frac{c}{d}$, hence the sequence

$$\frac{p_k}{q_k}, \frac{p_k + p_{k+1}}{q_k + q_{k+1}}, \frac{p_k + 2p_{k+1}}{q_k + 2q_{k+1}}, \ldots, \frac{p_k + a_k\, p_{k+1}}{q_k + a_k\, q_{k+1}} = \frac{p_{k+2}}{q_{k+2}}$$

is monotonic. Additionally, $\frac{p_k}{q_k}$ and $\frac{p_{k+2}}{q_{k+2}}$ lie on the same side with respect to q, hence:

$$\left|q - \frac{p_k}{q_k}\right| \geq \left|\frac{p_k + p_{k+1}}{q_k + q_{k+1}} - \frac{p_k}{q_k}\right| = \frac{1}{q_k(q_k + q_{k+1})},$$

where the last identity is again a consequence of the Theorem (358). □

Corollary 362. If the continued fraction of $\alpha \in \mathbb{Q}^+$ has at least $n+2$ terms and $\frac{p_k}{q_k}$ is its k-th convergent, we have:

$$|q_k\alpha - p_k| > |q_{k+1}\alpha - p_{k+1}|.$$

Proof. The following inequality holds as a consequence of the Corollary (361):

$$|q_k\alpha - p_k| \geq \frac{1}{q_k + q_{k+1}} \geq \frac{1}{q_{k+2}} > |q_{k+1}\alpha - p_{k+1}|.$$

□

Remark. Continued fractions provide a bijection between $[1,+\infty)$ and the set of (finite or infinite) sequences of positive integers. This leads to a very short proof of the non-countability of $\mathbb{R}$:

$$|\mathbb{R}| \geq |[1,+\infty)| \geq \left|\mathbb{N}^{\mathbb{N}}\right| \geq \left|2^{\mathbb{N}}\right| > |\mathbb{N}|.$$

Despite not being very common, to use continued fractions to *define* the set of real numbers is a perfectly viable alternative approach to using Dedekind cuts or the completion of a metric space, since continued fractions encode the order structure of $\mathbb{R}$ in a very efficient way.

Theorem 363 (Hurwitz)**.** If we consider two consecutive convergents of the continued fraction of $\alpha \in (\mathbb{R}\setminus\mathbb{Q})^+$, at least one of them fulfills:

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{2\,q^2}.$$

Proof. Let us assume, by contradiction, that both $\frac{p_k}{q_k}$ and $\frac{p_{k+1}}{q_{k+1}}$ violate the previous inequality. Such assumptions lead to:

$$\frac{1}{q_k q_{k+1}} = \left|\frac{p_k}{q_k} - \frac{p_{k+1}}{q_{k+1}}\right| = \left|\alpha - \frac{p_k}{q_k}\right| + \left|\alpha - \frac{p_{k+1}}{q_{k+1}}\right| \geq \frac{1}{2q_k^2} + \frac{1}{2q_{k+1}^2},$$

from which:

$$2q_k q_{k+1} \geq q_k^2 + q_{k+1}^2,$$

or $0 \geq (q_{k+1}-q_k)^2$, implying $k=0$ and $q_k = 1$: in such a case, however, $\left|\alpha - \frac{p_0}{q_0}\right| < \frac{1}{2}$. □

Theorem 364 (Hurwitz)**.** If we consider three consecutive convergent of the continued fraction of $\alpha \in (\mathbb{R}\setminus\mathbb{Q})^+$, at least one of them fulfills:

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{\sqrt{5}\,q^2}.$$

Proof. Let us assume that the consecutive convergents $\frac{p_k}{q_k}, \frac{p_{k+1}}{q_{k+1}}, \frac{p_{k+2}}{q_{k+2}}$ all violate the wanted inequality. In such a case:

$$\frac{1}{q_k q_{k+1}} > \frac{1}{\sqrt{5}q_k^2} + \frac{1}{\sqrt{5}q_{k+1}^2}, \qquad \frac{1}{q_{k+1} q_{k+2}} > \frac{1}{\sqrt{5}q_{k+1}^2} + \frac{1}{\sqrt{5}q_{k+2}^2},$$

hence by setting $\lambda_1 = \frac{q_{k+1}}{q_k}$ and $\lambda_2 = \frac{q_{k+2}}{q_{k+1}}$ we have:

$$\lambda_i^2 - \sqrt{5}\lambda_i + 1 < 0,$$

from which it follows that $\lambda_i \in \left(\frac{\sqrt{5}-1}{2}, \frac{\sqrt{5}+1}{2}\right)$. This leads to:

$$\lambda_2 = \frac{q_{k+2}}{q_{k+1}} = a_{k+1} + \frac{1}{\lambda_1} > 1 + \frac{2}{\sqrt{5}+1} = \frac{\sqrt{5}+1}{2},$$

i.e. to a contradiction. □

Theorem 365 (Hurwitz duplication formula)**.**

$$2\,[0,2a+1,b,c,\ldots] = [0,a,1,1+2\,[0,b-1,c,\ldots]]\,,$$
$$2\,[0,2a,b,c,\ldots] = [0,a,2\,[b,c,\ldots]]\,.$$

Proof. Let $x = [b, c, \ldots]$. In such a case the LHS of the first equality is given by

$$2 \cdot \frac{1}{(2a+1) + \frac{1}{x}} = \frac{2x}{(2a+1)x + 1}$$

and the RHS of the first equality is given by

$$\frac{1}{a + \frac{1}{1+\frac{1}{1+\frac{2}{x-1}}}} = \frac{1}{a + \frac{1}{1+\frac{x-1}{x+1}}} = \frac{1}{a + \frac{x+1}{2x}} = \frac{2x}{2ax + (x+1)}.$$

Similarly, in order to prove the second equality it is enough to check that:

$$2 \cdot \frac{1}{2a + \frac{1}{x}} = \frac{1}{a + \frac{1}{2x}}.$$

□

Hurwitz duplication formula leads to a simple derivation of the continued fraction of e, related to the interplay between continued fractions and Riccati differential equations. Let $f_0(x) = \tanh(x)$ and

$$f_{n+1}(x) = \frac{1}{f_n(x)} - \frac{2n+1}{x}.$$

The Taylor series of $f_0(x)$ at the origin is $x - \frac{x^3}{3} + \ldots$, so $\frac{1}{f_0(x)}$ has a simple pole with residue 1 at the origin and $f_1(x) = \frac{1}{f_0(x)} - \frac{1}{x}$ is an analytic function in a neighbourhood of the origin. It is not difficult to prove by induction that $f_n(x)$ fulfills the differential equation $f_n'(x) = 1 - f_n(x)^2 - \frac{2n}{x} f_n(x)$ and it is an analytic function in a neighbourhood of the origin. Assuming the fact that $f_n(x)$ is uniformly bounded over n, the continued fraction representation

$$\tanh(x) = \left[0, \tfrac{1}{x}, \tfrac{3}{x}, \tfrac{5}{x}, \tfrac{7}{x}, \ldots\right]$$

(due to Lambert) follows, and a careful study reveals it is convergent over the whole $\mathbb{C}^*$.
By evaluating both sides of the last identity at $x = \frac{1}{2}$ we get

$$\begin{aligned}
\frac{e-1}{e+1} &= [0; 2, 6, 10, 14, \ldots] \\
\frac{e+1}{e-1} &= [2; 6, 10, 14, 18, \ldots] \\
\frac{2}{e-1} &= [1; 6, 10, 14, 18, \ldots] \\
\frac{e-1}{2} &= [0; 1, 6, 10, 14, \ldots] \\
e &= 1 + 2 \cdot [0; 1, 6, 10, 14, \ldots]
\end{aligned}$$

and the continued fraction for e can be deduced from Hurwitz duplication formula.

Exercise 366. By exploiting Lambert's continued fraction prove that for any $n \in \mathbb{Z} \setminus \{0\}$ we have $e^n \notin \mathbb{Q}$, then show that for any $q \in \mathbb{Q}^+ \setminus 1$ we have $\log q \notin \mathbb{Q}$.

Theorem 367 (Lagrange)**.** If α is an irrational positive numbers and two coprime integers p, q fulfill:

$$\left|\alpha - \frac{p}{q}\right| \leq \frac{1}{2q^2},$$

then $\frac{p}{q}$ is a convergent of the continued fraction of α.

Proof. Let $[a_0; a_1, \ldots, a_n]$ be the continued fraction of $\frac{p}{q}$: such assumption grants $p = p_n$ and $q = q_n$. By setting

$$\beta = \frac{p_{n-1} - \alpha q_{n-1}}{\alpha q_n - p_n}$$

we have:

$$\alpha = \frac{\beta p_n + p_{n-1}}{\beta q_n + q_{n-1}}, \tag{2}$$

and additionally:

$$\left|\alpha - \frac{p}{q}\right| = \frac{1}{q_n(\beta q_n + q_{n-1})}.$$

The RHS is $\leq \frac{1}{2q_n^2}$, from which:

$$\beta \geq 2 - \frac{q_{n-1}}{q_n} > 1.$$

If we represent $\beta = [a_{n+1}; \ldots]$ and exploit the equation (2) we have that the continued fraction of α is given by the concatenation of the continued fraction of $\frac{p}{q}$ and the continued fraction of β, i.e.:

$$\alpha = [a_0; a_1, \ldots, a_n, a_{n+1}, \ldots],$$

which proves that $\frac{p}{q}$ is a convergent of α. □

Lagrange's Theorem implies that the convergents of a continued fraction are in some sense the *best* rational approximations for the represented positive real number. Such remark leads to an interesting approach for proving the irrationality of some positive real numbers:

Corollary 368. If $\alpha \in \mathbb{R}^+$ and there exists a sequence $\left\{\frac{p_n}{q_n}\right\}_{n\geq 1}$ of rational numbers fulfilling

$$\forall n \geq 1, \qquad |q_n\alpha - p_n| \leq \frac{1}{2q_n},$$

then $\alpha \in \mathbb{R} \setminus \mathbb{Q}$.

Informally speaking, if a real number is (in the previous sense) *too close* to the set of rational numbers, then it is certainly irrational (*kind of counterintuitive, isn't it?*). A possible proof of the irrationality of some constant is so given by finding an infinite number of *good enough* rational approximations. Remarkably, in many cases such sequences of accurate approximations can be found by exploiting peculariar integrals and/or orthogonal polynomials: this idea can be traced back to Beuker ($\approx$ 1970) and it has been studied and extended by many others, like Hadjicostas, Rivoal, Sorokin, Hata, Viola, Rhin, Nesterenko and Zudilin. Let us see a famous instance of such technique:

Lemma 369. The following approximations are pretty accurate:

$$e \approx \frac{19}{7}, \qquad \pi \approx \frac{22}{7}.$$

Proof. On the interval $(0, 1)$, the positive $x(1 - x)$ is positive and never exceeds $\frac{1}{4}$.
In particular, the following integrals

$$\int_0^1 x^2(1-x)^2 e^{-x}\,dx = 14e - 38, \qquad \int_0^1 \frac{x^4(1-x)^4}{1+x^2}\,dx = \frac{22}{7} - \pi$$

are positive but respectively bounded by $\frac{1}{16}$ and $\frac{1}{256}$. In a similar way, $\pi^2 \approx \frac{493}{50}$ follows from:

$$\frac{1}{120} \geq \int_0^1 \frac{x^4(1-x)^2}{1+x}(-\log x)\,dx = \frac{493}{150} - \frac{\pi^2}{3}.$$

Even better approximations can be obtained by replacing the weights $x^\alpha(1-x)^\beta$ with shifted Legendre polynomials. For instance:

$$\int_0^1 P_4(2x-1)e^{-x}\,dx = 1001 - \frac{2721}{e} \approx 4\cdot 10^{-5}.$$

In the last case the magnitude of the involved integral is a bit more difficult to estimate, but that can be done through Rodrigues' formula, for instance. □

Quadratic surds also play an important role in the theory of continued fractions. We wish to investigate the structure of the continued fraction of $\sqrt{D}$, where $D\in\mathbb{N}^+$ is not a square. Let us consider, for instance, the $D=19$ case:

$$\sqrt{19} = [4;2,1,3,1,2,8,2,1,3,1,2,8,2,1,3,1,2,8,\ldots] = [4;\overline{2,1,3,1,2,8}].$$

The seeming periodicity is not accidental:

Theorem 370 (Lagrange)**.** If $D\in\mathbb{N}^+$ is not a square,

$$\sqrt{D} = [a_0;\overline{a_1,a_2,\ldots,a_2,a_1,2a_0}].$$

For the sake of simplicity we are going to prove the claim only in the $D=19$ case, but the reader can easily extend the following argument and prove Lagrange's Theorem in full generality. For starters, we have that $\sqrt{19}-4$ and $-\sqrt{19}-4$ are the only roots of the polynomial x^2+8x-3. In particular, $\frac{1}{\sqrt{19}-4}=\frac{4+\sqrt{19}}{3}$ is a root of the polynomial $3x^2-8x-1$ ed and its integer part equals 2. It follows that $\gamma=\frac{1}{\sqrt{19}-4}-2$ is a root of the polynomial $3x^2+4x-5$ and $\frac{1}{\gamma}$ is a root of the reciprocal polynomial $5x^2-4x-3$. Thus the construction of the continued fraction of $\sqrt{19}$ is associated with the sequence of polynomials

$$\begin{gathered}3x^2-8x-1\\5x^2-4x-3\\2x^2-6x-5\\5x^2-6x-2\\3x^2-4x-5\\x^2-8x-3\end{gathered}$$

all having integer coefficients and discriminant $4\cdot 19=76$. There is a finite number of such polynomials and the transitions between a polynomial and the next one are fixed. In particular this sort of "Euclidean algorithm for quadratic forms" allows us to write $\alpha=\sqrt{19}-4$ as a fractional linear function of the conjugated root $-\sqrt{19}-4$, and the continued fraction of $\sqrt{19}$ turns out to have a structure that is both periodic and almost-palindromic, since the map associating a root of an irreducible quadratic polynomial with its algebraic conjugate is an involution.
By studying the behaviour of $p_n^2-Dq_n^2$ as $\frac{p_n}{q_n}$ ranges among the convergents of $\sqrt{D}$ we also have the following results about **Pell's equations**[15]:

Theorem 371. If $D\in\mathbb{N}^+$ is not a square, the Diophantine equation $x^2-Dy^2=1$ has an infinite number of solutions. Assuming $\sqrt{D}$ has the following continued fraction representation:

$$\sqrt{D} = [a_0;\overline{a_1,a_2,\ldots a_{k-1},a_k,a_{k-1},\ldots,a_2,a_1,2a_0}],$$

the fundamental solution is given by

$$\frac{x}{y} = [a_0;a_1,a_2,\ldots,a_2,a_1].$$

[15] A renowned case of di *mathematical misattribution*: the outlined results about the Diophantine equation $x^2-Dy^2=\pm1$ are due to Brounckner, but an erroneous quotation from Euler led them to being attributed to Pell forever since.

If $\sqrt{D}$ has the following continued fraction representation:

$$\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots a_{k-1}, a_k, a_k, a_{k-1}, \ldots, a_2, a_1, 2a_0}],$$

the fundamental solution is given by

$$\frac{x}{y} = [a_0; a_1, a_2, \ldots, a_2, a_1, 2a_0, a_1, a_2, \ldots, a_2, a_1].$$

Theorem 372. If $D \in \mathbb{N}^+$ is not a square, and $\sqrt{D}$ has the following continued fraction representation:

$$\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots a_{k-1}, a_k, a_k, a_{k-1}, \ldots, a_2, a_1, 2a_0}],$$

then the Diophantine equation $x^2 - Dy^2 = -1$ has no solution. Conversely, if

$$\sqrt{D} = [a_0; \overline{a_1, a_2, \ldots a_{k-1}, a_k, a_{k-1}, \ldots, a_2, a_1, 2a_0}],$$

the Diophantine equation $x^2 - Dy^2 = -1$ has an infinite number of solutions and the fundamental solution is given by:

$$\frac{x}{y} = [a_0; a_1, a_2, \ldots a_{k-1}, a_k, a_{k-1}, \ldots, a_2, a_1].$$

In the $D = 19$ case, for instance, the Diophantine equation $x^2 - 19y^2 = -1$ has no solutions, but the equation $x^2 - 19y^2 = 1$ has the fundamental solution $(170, 39)$. In greater generality, Lagrange's Theorem proves that, if $D \in \mathbb{N}^+$ is not a square, the ring $\mathbb{Z}[\sqrt{D}]$ has an infinite number of invertible elements.

The theory of continued fractions is also deeply related to **Padé approximants** and the theory of **moments**.

Theorem 373 (Brounckner)**.**

$$\pi = \cfrac{4}{1 + \cfrac{1^2}{2 + \cfrac{3^2}{2 + \cfrac{5^2}{2 + \cfrac{7^2}{\cdots}}}}}$$

Proof. Let us set:

$$I_n = \int_0^1 \frac{x^{2n}}{1+x^2}\,dx.$$

We have $I_0 = \frac{\pi}{4}$, $I_1 = 1 - \frac{\pi}{4}$ and

$$I_n + I_{n+1} = \int_0^1 x^{2n}\,dx = \frac{1}{2n+1},$$

from which it follows that:

$$\frac{I_n + I_{n+1}}{I_{n+1} + I_{n+2}} = \frac{2n+3}{2n+1}$$

and by setting $r_n = \frac{I_{n+1}}{I_n}$ we get:

$$\frac{1 + 1/r_n}{r_{n+1} + 1} = \frac{2n+3}{2n+1}, \qquad r_n = \frac{2n+1}{2 + (2n+3)r_{n+1}},$$

hence:

$$r_0 = \frac{1}{2+3r_1} = \cfrac{1}{2+\cfrac{3^2}{2+5r_2}} = \cfrac{1}{2+\cfrac{3^2}{2+\cfrac{5^2}{2+7r_3}}} = \ldots$$

and the claim follows from noticing that $r_0 = \frac{4}{\pi} - 1$ and $I_n = O\left(\frac{1}{2n+1}\right)$. □

Theorem 374 (Euler)**.**

$$e = [2; 1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, \ldots]$$

Proof. The following proof is due to Cohn. For any real polynomial r having degree k we have:

$$\int_0^1 r(t)e^{xt}\,dx = \frac{q(x)e^x - p(x)}{x^{k+1}}$$

where $p(x)$ and $q(x)$ are polynomials with degree $\leq k$:

$$p(x) = r(0)x^k - r'(0)x^{k-1} + r''(0)x^{k-2} - \ldots \qquad q(x) = r(1)x^k - r'(1)x^{k-1} + r''(1)x^{k-2} - \ldots$$

By considering $r(t) = r_{m,n}(t) = t^n(t-1)^m$ we have that:

$$\int_0^1 r_{m,n}(t)e^t\,dt = \int_0^1 t^n(t-1)^m\,dt = q(1)e - p(1)$$

where $p(1), q(1)$ are integers and the absolute value of the integral does not exceed $4^{-\min(m,n)}$.
In particular, $\frac{p(1)}{q(1)}$ is a good rational approximations of e. If we define

$$A_n = \frac{1}{n!}\int_0^1 t^n(t-1)^n e^t\,dt, \qquad B_n = \frac{1}{n!}\int_0^1 t^{n+1}(t-1)^n e^t\,dt, \qquad C_n = \frac{1}{n!}\int_0^1 t^n(t-1)^{n+1}e^t\,dt$$

we may easily check that the following identities hold:

$$A_n = -B_{n-1} - C_{n-1}, \qquad B_n = -2nA_{n-1} + C_{n-1}, \qquad C_n = B_n - A_n,$$

hence by denoting as $\frac{p_n}{q_n}$ the n-th convergent of the continued fraction of

$$[1, 0, 1, 1, 2, 1, 1, 4, 1, 1, 6, 1, \ldots],$$

we get that:

$$A_n = q_{3n}e - p_{3n}, \qquad B_n = p_{3n+1} - q_{3n+1}e, \qquad C_n = p_{3n+2} - q_{3n+2}e$$

and the claim immediately follows. □

We may notice that the rational approximations of e deriving from Beuker's integral

$$\int_0^1 t^n(1-t)^n e^{-t}\,dt$$

belong to the set of the *best rational approximations*, hence to the set of convergents of the continued fraction of e.
The following Corollary is a straightforward consequence:

Corollary 375.

$$e \notin \mathbb{Q}.$$

Truth to be told, continued fractions are not strictly needed to prove the irrationality of e.
Euler himself provided a simpler, alternative proof: let us assume $e = \frac{p}{q}$ for some couple (p, q) of coprime positive integers. With such assumptions:

$$\frac{q}{p} = \frac{1}{e} = \sum_{n\geq 0} \frac{(-1)^n}{n!}$$

and $\frac{p!}{e}$ is an integer. That leads to a contradiction, since:

$$\frac{p!}{e} = \sum_{n\leq p} \frac{(-1)^n p!}{n!} + \sum_{n>p} \frac{(-1)^n}{n(n-1)\cdot\ldots\cdot(p+1)}$$

is the sum between an integer and a real number whose absolute value is $\leq \frac{1}{p+1}$. However it is important to remark that Cohn's proof of the irrationality of e can be suitably modified in order to prove something way more subtle, i.e. that e is a **trascendental** number.

Theorem 376 (Hermite, 1873)**.** If $p(x)$ is a non-constant polynomial with rational coefficients, $p(e) \neq 0$.

Proof. Let us assume, by contradiction, to have:

$$a_0 + a_1 e + a_2 e^2 + \ldots + a_n e^n = 0$$

with $a_0, \ldots, a_n \in \mathbb{Z}$ and $a_0 \neq 0$. For any polynomial $f(t)$ we have:

$$e^x \int_0^x f(t)e^{-t}\,dt = e^x F(0) - F(x), \qquad F(x) = \sum_{k\geq 0} f^{(k)}(x).$$

If we evaluate both sides of the last identity at $x = k = 0, 1, \ldots, n$, multiply both sides by a_k and sum such contributions, we get:

$$\sum_{k=0}^{n} a_k e^k \int_0^k f(t)e^{-t}\,dt = F(0)\sum_{k=0}^{n} a_k e^k - \sum_{k=0}^{n} a_k F(k) = -\sum_{k=0}^{n} a_k F(k)$$

where we still have the chance to choose $f(t)$ in a suitable way. If the RHS of the last identity is a non-zero integer, but the LHS is a real number whose absolute value is less than one, we reach the wanted contradiction.
The most difficult part of the current proof is to check that by picking

$$f(t) = \frac{t^{p-1}}{(p-1)!}(t-1)^p (t-2)^p \cdot\ldots\cdot (t-n)^p$$

we meet the previous constraints as soon as p is a large enough prime number. By defining

$$M = \max_{t\in[0,n]} |t(t-1)\cdot\ldots\cdot(t-n)|$$

we get that:

$$\left|\sum_{k=0}^{n} a_k e^k \int_0^k f(t)e^{-t}\,dt\right| \leq \frac{n e^n M^p}{(p-1)!} \sum_{k=0}^{n} |a_k|\,,$$

hence by choosing large enough p the LHS can be made arbitrarily close to zero, since neither M or n depend on p. Let us assume $p > n$ and $p > |a_0|$. Since $F(k)$ equals the sum of f and all its derivatives at k, and $k = 1, 2, \ldots, n$ are roots with multiplicity p for f, $a_k F(k)$ is an integer number for any $k \in \{1, 2, \ldots, n\}$. Additionally, it is a multiple of p. About $F(0)$ we have that it is an integer, but since[16]

$$f^{(p-1)}(0) = \left[(-1)^n n!\right]^p,$$

we have $F(0) \equiv \pm 1 \pmod{p}$, hence $\sum_{k=0}^{n} a_k F(k)$ is a *non-zero* integer as wanted. □

[16] "If $n \not\equiv 0 \pmod{p}$, then $n \neq 0$" is a trivial statement with a huge number of non-trivial consequences in Mathematics.

The technique shown for proving first the irrationality, then the trascendence of e, has a close analogue for π, too.

Theorem 377 (Lambert, 1761).
$$\pi \notin \mathbb{Q}.$$

Proof. If we set $I_n = \int_{-1}^{1}(1-x^2)^n \cos(\alpha x)\,dx$, the integration by parts formula leads to:

$$\alpha^2 I_n = 2n(2n-1)I_{n-1} - 4n(n-1)I_{n-2}$$

hence, in particular:

$$\alpha^{2n+1} I_n = n!\,[P_n(\sin\alpha) + Q_n(\cos\alpha)]$$

where P_n, Q_n are polynomials with integer coefficients and degree $< 2n+1$.
Let us set $\alpha = \frac{\pi}{2}$ and suppose we have $\frac{\pi}{2} = \frac{b}{a}$ with a, b being coprime positive integers. Such assumptions grant that

$$J_n \stackrel{\text{def}}{=} \frac{b^{2n+1} I_n}{n!}$$

is an integer. However $0 < I_n \leq \int_{-1}^{1}\cos(\pi x/2)\,dx = \frac{4}{\pi}$ for any n, hence

$$\lim_{n\to+\infty} J_n = 0,$$

but we cannot have $J_n = 0$ for any $n \in \mathbb{N}$, and the irrationality of π follows. □

Theorem 378 (Lindemann, 1882). If $p(x)$ is a non-constant polynomial with rational coefficients, $p(\pi) \neq 0$.

Proof. If we assume that π is an algebraic number over $\mathbb{Q}$, the complex number $i\pi$ is algebraic as well. Let us assume that $\theta_1(x)$ is the minimal polynomial of $i\pi$ over $\mathbb{Q}$, having roots $\alpha_1, \ldots, \alpha_n$. As a consequence of De Moivre's formula,

$$(e^{\alpha_1}+1)(e^{\alpha_2}+1)\cdot\ldots\cdot(e^{\alpha_n}+1) = 0.$$

With such assumptions all the numbers of the form $\alpha_j + \alpha_k$ (with $j \neq k$) are roots of a polynomial $\theta_2(x) \in \mathbb{Q}[x]$, all the numbers of the form $\alpha_i + \alpha_j + \alpha_k$ (with i, j, k being three distinct elements of $\{1, \ldots, n\}$) are roots of a polynomial $\theta_3(x) \in \mathbb{Q}[x]$ and so on. Let us set:

$$\theta(x) = \theta_1(x)\cdot\theta_2(x)\cdot\ldots\cdot\theta_n(x) = x^m\left(c_0 x^r + c_1 x^{r-1} + \ldots + c_r\right).$$

We have $c_r \in \mathbb{Q}\backslash\{0\}$ and the roots of $\theta(x)$ are given by the sum of the elements of any non-empty subset of $\{\alpha_1, \ldots, \alpha_n\}$. If we denote as $\beta_1, \ldots, \beta_r$ the roots of $c_0 x^r + c_1 x^{r-1} + \ldots + c_r$, De Moivre's identity leads us to:

$$\sum_{k=1}^{r} e^{\beta_k} + K = 0$$

with K being a positive integer. Let us consider

$$f(x) \stackrel{\text{def}}{=} c_0^s x^{p-1}\frac{\theta(x)^p}{(p-1)!},\qquad s \stackrel{\text{def}}{=} rp-1,\qquad F(x) \stackrel{\text{def}}{=} f(x) + f'(x) + \ldots + f^{(s+p)}(x)$$

where p is a prime number large enough. We have:

$$KF(0) + \sum_{j=1}^{r} F(\beta_j) = -\sum_{j=1}^{r}\beta_j \int_0^1 e^{(1-\lambda)\beta_j} f(\lambda\beta_j)\,d\lambda$$

and we may reach the wanted conclusion by proceeding like in the proof of the trascendence of e. □

Lindemann's Theorem puts a final word on the problem of *squaring the circle*, establishing its impossibility. It also has an interesting generalization:

Theorem 379 (Lindemann-Weierstrass)**.** If $\alpha_1, \ldots, \alpha_n$ are linearly independent algebraic numbers over $\mathbb{Q}$, the following quantities

$$e^{\alpha_1},\ e^{\alpha_2},\ \ldots\ e^{\alpha_n}$$

are **algebraically independent** over $\mathbb{Q}$, i.e. there is no multivariate polynomial $p \in \mathbb{Q}[x_1, \ldots, x_n]$ such that

$$p(e^{\alpha_1}, \ldots, e^{\alpha_n}) = 0,$$

with the only trivial exception of the polynomial which constantly equals zero.

Corollary 380. $\sin(1)$ is a trascendental number over $\mathbb{Q}$. Indeed, if we assume that it is algebraic we get that

$$i \sin(1) + \sqrt{1 - \sin^2(1)} = e^i,$$

is algebraic as well, but that contradicts the Lindemann-Weierstrass Theorem.

Corollary 381. For any natural number $n \geq 2$, $\log(n)$ is a trascendental number over $\mathbb{Q}$. Assuming it is algebraic, the identity $e^{\log n} = n$ together with the Lindemann-Weierstrass Theorem lead to a contradiction.

Corollary 382.

$$\arctan \frac{1}{2} = \frac{i}{2} \cdot \log \frac{2-i}{2+i}$$

is a trascendental number over $\mathbb{Q}$.

Definition 383. In Number Theory, a Liouville number is an irrational number x with the property that, for every positive integer n, there exist integers p and q with $q > 1$ and such that

$$0 < \left| x - \frac{p}{q} \right| < \frac{1}{q^n}.$$

Theorem 384 (Liouville, 1844)**.** Every Liouville number is a trascendental number over $\mathbb{Q}$.

Proof. Is enough to show that if α is an irrational number which is the root of a polynomial f of degree $n > 0$ with integer coefficients, then there exists a real number $A > 0$ such that, for all integers p, q, with $q > 0$,

$$\left| \alpha - \frac{p}{q} \right| > \frac{A}{q^n}.$$

Let M be the maximum value of $|f'(x)|$ over the interval $[\alpha - 1, \alpha + 1]$. Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be the distinct roots of f which differ from α. Select some value $A > 0$ satisfying

$$A < \min \left(1, \frac{1}{M}, |\alpha - \alpha_1|, |\alpha - \alpha_2|, \ldots, |\alpha - \alpha_m| \right).$$

Now assume that there exist some integers p, q contradicting the lemma. Then

$$\left| \alpha - \frac{p}{q} \right| \leq \frac{A}{q^n} \leq A < \min \left(1, \frac{1}{M}, |\alpha - \alpha_1|, |\alpha - \alpha_2|, \ldots, |\alpha - \alpha_m| \right).$$

Then $\frac{p}{q}$ is in the interval $[\alpha - 1, \alpha + 1]$ and $\frac{p}{q}$ is not in $\{\alpha_1, \alpha_2, \ldots, \alpha_m\}$, so $\frac{p}{q}$ is not a root of f and there is no root of f between α and $\frac{p}{q}$. By the mean value theorem, there exists an x_0 between $\frac{p}{q}$ and α such that

$$f(\alpha) - f\left(\tfrac{p}{q}\right) = \left(\alpha - \tfrac{p}{q}\right) f'(x_0).$$

Since α is a root of f but $\frac{p}{q}$ is not, we see that $|f'(x_0)| > 0$ and we can rearrange:

$$\left|\alpha - \frac{p}{q}\right| = \frac{\left|f(\alpha) - f\left(\frac{p}{q}\right)\right|}{|f'(x_0)|} = \left|\frac{f\left(\frac{p}{q}\right)}{f'(x_0)}\right|$$

Now, f is of the form $\sum_{i=0}^{n} c_i x^i$ where each c_i is an integer, so we can express $\left|f\left(\frac{p}{q}\right)\right|$ as

$$\left|f\left(\frac{p}{q}\right)\right| = \frac{1}{q^n}\left|\sum_{i=0}^{n} c_i p^i q^{n-i}\right| \geq \frac{1}{q^n},$$

the last inequality holding because $\frac{p}{q}$ is not a root of f and the c_i are integers. Thus we have that $\left|f\left(\frac{p}{q}\right)\right| \geq \frac{1}{q^n}$. Since $|f'(x_0)| \leq M$ by the definition of M, and $\frac{1}{M} > A$ by the definition of A, we have that

$$\left|\alpha - \frac{p}{q}\right| = \left|\frac{f\left(\frac{p}{q}\right)}{f'(x_0)}\right| \geq \frac{1}{Mq^n} > \frac{A}{q^n} \geq \left|\alpha - \frac{p}{q}\right|$$

which is a contradiction; therefore, no such p, q exist, proving the Lemma. As a consequence of this Lemma, let x be a Liouville number; if x is algebraic there exists some integer n and some positive real A such that for all p, q

$$\left|x - \frac{p}{q}\right| > \frac{A}{q^n}.$$

Let r be a positive integer such that $\frac{1}{2^r} \leq A$. If we let $m = r + n$, then, since x is a Liouville number, there exists integers $a, b > 1$ such that

$$\left|x - \frac{a}{b}\right| < \frac{1}{b^m} = \frac{1}{b^{r+n}} \leq \frac{1}{2^r b^n} \leq \frac{A}{b^n}$$

which contradicts the Lemma; therefore x is not algebraic over $\mathbb{Q}$. □

Corollary 385. Both

$$\sum_{n\geq 0} \frac{1}{2^{n!}} \quad \text{and} \quad \sum_{n\geq 0} \frac{1}{10^{n!}}$$

are trascendental numbers.

A cornerstone in the Diophantine Approximation Theory is given by the following result, proved by combining the shown techniques (of an analytic and arithmetical nature at the same time) with a Lemma in Linear Algebra due to Siegel:

Theorem 386 (Gelfond-Schneider)**.** If a, b are algebraic numbers over $\mathbb{Q}$, $a \notin \{0, 1\}$ and b is an irrational number, then

$$a^b \text{ is trascendental over } \mathbb{Q}.$$

Just like the Lindemann-Weierstrass Theorem, the Gelfond-Schneider Theorem has plenty of remarkable consequences:

- $2^{\sqrt{2}}$ is a trascendental number. Assuming the opposite $(2^{\sqrt{2}})^{\sqrt{2}} = 2^2 = 4$ would be trascendental, but it clearly is not;
- e^{π} and $e^{-\pi}$ are trascendental numbers, since $e^{-\pi} = (e^{\pi i})^i = (-1)^i$.

Exercise 387. Prove that if $x \in \left(\frac{1}{3}, \frac{2}{3}\right) \cap \mathbb{Q}$,

$$\frac{\log(1-x)}{\log x}$$

is a trascendental number over $\mathbb{Q}$.

Despite the abundance of deep results, there still are many open problems in Diophantine Approximation. For instance it is a widespread opinion (supported by the computation of many terms in the involved continued fractions) that all the following numbers are irrational:

$$\zeta(5), \qquad \Gamma\left(\frac{1}{5}\right), \qquad \gamma, \qquad \pi+e$$

but no one (up to 2017) has been able to actually prove the irrationality of any of them.

Exercise 388. By considering the sequence of integrals given by

$$I_n(\alpha) = \int_0^{\pi} \cos(nx) e^{\alpha\cos(x)}\,dx,$$

prove that the following identity holds:

$$[1;2,3,4,5,6,\ldots] = \left(\sum_{m\geq 0}\frac{1}{m!^2}\right)\cdot\left(\sum_{m\geq 0}\frac{1}{m!(m+1)!}\right)^{-1}.$$

Exercise 389. By considering the continued fraction of $\frac{\log 10}{\log 2}$, prove that the leading digits in the decimal representation of 2^m are 999 for an infinite number of $m\in\mathbb{N}$.

Exercise 390. By setting $\alpha_0 = 2, \beta_0 = 1$ and

$$\alpha_{n+1} = \alpha_n^2 + 3\beta_n^2, \qquad \beta_{n+1} = 2\alpha_n\beta_n$$

(sequence generated by Newton's method or the Babylonian algorithm), prove that the ratios $\frac{\alpha_n}{\beta_n}$ are convergents of the continued fraction of $\sqrt{3}$ for any $n\geq 1$.

Exercise 391. Considering that the function $g(z) = \frac{\sinh z}{z}$ is a solution of the differential equation $(zg)'' = (zg)$, deduce from

$$\frac{g'}{g} = \frac{1}{\frac{2}{z}+\frac{g''}{g'}}$$

that for any x in a neighbourhood of the origin we have:

$$\tanh(x) = \cfrac{x}{1+\cfrac{x^2}{3+\cfrac{x^2}{5+\cfrac{x^2}{7+\ldots}}}}$$

hence $\sqrt{2}\,\frac{e^{\sqrt{2}}-1}{e^{\sqrt{2}}+1}$ is an irrational number.

Exercise 392 (A continued fraction representation for the error function)**.**
Prove that for any $z > 1$ the following identity holds:

$$\int_{z}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}\,dx = \frac{e^{-z^2/2}}{\sqrt{2\pi}}\cdot\cfrac{1}{z+\cfrac{1}{z+\cfrac{2}{z+\cfrac{3}{\ldots}}}}$$

Notice that the LHS is the probability that a normal random variable with distribution $N(0,1)$ takes values in $(z,+\infty)$. In particular, the above identity is an accurate **tail inequality** for the normal distribution.

On the irrationality of $\zeta(3)$ and $\zeta(2)$.
We are going to provide a sketch of Frits Beukers' approach for proving $\zeta(3)\notin\mathbb{Q}$.
By Fubini's Theorem and termwise integration,

$$\int_0^1\int_0^1\frac{-\log(xy)}{1-xy}x^r y^r\,dx\,dy = 2\left(\zeta(3)-\sum_{k=1}^{r}\frac{1}{k^3}\right)$$

so by denoting as d_m the least common multiple of $1,2,\ldots,m$ we have that

$$\int_0^1\int_0^1\frac{-\log(xy)}{1-xy}x^r y^s\,dx\,dy$$

is a rational number whose reduced denominator is a divisor of $d^3_{\max(r,s)}$. By considering

$$I_n = \int_0^1\int_0^1\frac{-\log(xy)}{1-xy}\tilde{P}_n(x)\tilde{P}_n(y)\,dx\,dy$$

where $\tilde{P}_n$ is a shifted Legendre polynomial, we have that I_n is the sum between a positive integral multiple of $\zeta(3)$ and a rational number whose reduced denominator is a divisor of d_n^3. In particular $I_n = \frac{A_n+B_n\zeta(3)}{d_n^3}$ with $A_n, B_n\in\mathbb{Z}$. Since $\frac{-\log(xy)}{1-xy} = \int_0^1\frac{dz}{1-(1-xy)z}$, by exploiting Rodrigues formula and suitable substitutions it is not difficult to show that I_n equals the following triple integral:

$$I_n = \int_0^1\int_0^1\int_0^1 u^n(1-u)^n v^n(1-v)^n w^n(1-w)^n\frac{du\,dv\,dw}{(1-(1-uv)w)^{n+1}}$$

fulfilling a simple inequality:

$$0 < I_n < \frac{1}{27^n}\int_0^1\int_0^1\int_0^1\frac{du\,dv\,dw}{(1-(1-uv)w)^{n+1}} = \frac{2\,\zeta(3)}{27^n}.$$

Now the claim essentially follows from $e<3$. By the Prime Number Theorem we have that $\log d_m = m + o(1)$, hence for any n large enough we have $d_n \leq (2.8)^n$. Assuming $\zeta(3)=\frac{a}{b}$ we get

$$0 < |bA_n + aB_n| \leq 2b\,\zeta(3)\frac{d_n^3}{27^n} < 2b\,\zeta(3)(0.9)^n$$

leading to a contradiction as soon as $(0.9)^n < \frac{1}{2b\,\zeta(3)}$.

Exercise 393. Show that the argument above can be modified (actually simplified) to establish the irrationality of $\zeta(2)=\frac{\pi^2}{6}$. Use the identity

$$\int_0^1\int_0^1\frac{\tilde{P}_n(x)(1-y)^n}{1-xy}\,dx\,dy = (-1)^n\int_0^1\int_0^1\frac{x^n(1-x)^n y^n(1-y)^n}{(1-xy)^{n+1}}\,dx\,dy.$$

For a simpler proof of the irrationality of π^2, one may consider the function

$$f_n(x) = \frac{x^n(1-x)^n}{n!} = \sum_{m=n}^{2n} c_m x^m$$

and assuming that $\pi^2 = \frac{a}{b}$ holds, letting

$$G(x) = b^n \left[\pi^{2n} f(x) - \pi^{2n-2} f''(x) + \pi^{2n-4} f^{(4)}(x) - \ldots + (-1)^n f^{(2n)}(x)\right].$$

We have that $f(0) = 0$ and $f^{(m)}(0) = 0$ if $m < n$ or $m > 2n$. But, if $n \leq m \leq 2n$, then

$$f^{(m)}(0) = \frac{m!}{n!} c_m$$

is an integer. Therefore $f(x)$ and all its derivatives take integral values at $x = 0$; since $f(x) = f(1-x)$ the same is true at $x = 1$, so that $G(0)$ and $G(1)$ are integers. We have:

$$\begin{aligned}\frac{d}{dx}\left[G'(x)\sin(\pi x) - \pi G(x)\cos(\pi x)\right] &= \left[G''(x) + \pi^2 G(x)\right]\sin(\pi x)\\ &= b^n \pi^{2n+2} f(x)\sin(\pi x)\\ &= \pi^2 a^n f(x)\sin(\pi x),\end{aligned}$$

hence

$$\pi \int_0^1 a^n f(x)\sin(\pi x)\,dx = \left[\frac{G'(x)\sin(\pi x)}{\pi} - G(x)\cos(\pi x)\right]_0^1 = G(0) + G(1) \in \mathbb{Z},$$

however:

$$0 < \pi \int_0^1 a^n f(x)\sin(\pi x)\,dx < \frac{\pi a^n}{n! 4^n} < 1$$

for any n large enough, leading to a contradiction.

15 Symmetric functions and elements of Analytic Combinatorics

The key idea in Analytic Combinatorics is to transfer all the information contained in a combinatorial problem in the sequence of coefficients of a power series, in order to deduce identities (or inequalities) from the geometric and analytic behavior of the function built that way. Such trick is possible due to the following Lemma:

Lemma 394 (Cauchy, Liouville)**.** If $\{a_n\}_{n\geq 0}$ is a sequence of complex numbers fulfilling $|a_n| \leq C \cdot M^n$ with $C > 0$ and $M > 1$,

$$f(z) \stackrel{\text{def}}{=} \sum_{n\geq 0} a_n z^n$$

is a holomorphic function in the region $\|z\| < \frac{1}{M}$ and we have:

$$a_n = \frac{1}{2\pi}\int_0^{2\pi} f(e^{i\theta}) e^{-ni\theta}\,d\theta = \frac{1}{2\pi i}\oint_{\|z\|=\frac{1}{2M}} \frac{f(z)}{z^{n+1}}\,dz.$$

Exercise 395 (Frobenius coin problem)**.** Let $R(N)$ be the number of ways for paying an integer price N with coins whose values are $1, 2$ or 3. Find the asymptotic behavior of $R(N)$ as $N \to +\infty$.

Proof. It is not difficult to check that $R(n)$ is given by the coefficient of x^n in the following product:

$$\begin{aligned}\left(1+x+x^2+x^3+\ldots\right)\left(1+x^2+x^4+x^6+\ldots\right)\left(1+x^3+x^6+x^9+\ldots\right) &= \frac{1}{1-x}\cdot\frac{1}{1-x^2}\cdot\frac{1}{1-x^3}\\ &= \frac{1}{(1-x)^3(1+x)(1+x+x^2)}\stackrel{\text{def}}{=} f(x).\end{aligned}$$

$f(x)$ is a meromorphic function with a triple pole at $x=1$ and simple poles at $x\in\{-1,\omega,\omega^2\}$, with $\omega=\exp\frac{2\pi i}{3}$. We may consider the *partial fraction decomposition* of $f(z)$:

$$(\clubsuit)\qquad f(z)=\frac{\mathbf{A}}{(1-z)^3}+\frac{B}{(1-z)^2}+\frac{C}{(1-z)}+\frac{D}{(1+z)}+\frac{E}{(\omega-z)}+\frac{F}{(\omega^2-z)}.$$

We have that the leading term of the asymptotic expansion of $R(n)$ is simply given by $\mathbf{A}$.
We may notice that[17] many coefficients of $(\clubsuit)$ can be readily computed through simple limits:

$$\begin{aligned}\mathbf{A} &= \lim_{z\to 1} f(z)(1-z)^3=\lim_{z\to 1}\frac{1}{(1+z)(1+z+z^2)}=\frac{1}{6}\\ D &= \lim_{z\to -1} f(z)(1+z)=\lim_{z\to -1}\frac{1}{(1-z)^3(1+z+z^2)}=\frac{1}{8}\\ E &= \lim_{z\to\omega} f(z)(\omega-z)=\lim_{z\to\omega}\frac{-1}{(1-z)^3(1+z)(\omega+z)}=-\frac{\omega}{9}\\ F &= \overline{E}=-\frac{\omega^2}{9}.\end{aligned}$$

The same holds for B and C:

$$\begin{aligned}B &= -\lim_{z\to 1}\frac{d}{dz}(1-z)^3 f(z)=\frac{1}{4}\\ C &= -\frac{1}{2}\lim_{z\to 1}\frac{d^2}{dz^2}(1-z)^3 f(z)=-\frac{17}{72}.\end{aligned}$$

The equality $C+D+E+F=0$ is not accidental: $f(z)$ behaves like $\frac{1}{z^6}$ for $\|z\|\to+\infty$, hence the sum of its residues has to be zero. The Taylor series of $\frac{1}{(1-x)^{k+1}}$ at the origin can be computed through the *hockey stick identity* and the *stars and bars* combinatorial argument, or by repeated differentiation:

$$\frac{1}{(1-x)^{k+1}}=\sum_{n\geq 0}\binom{n+k}{k}x^n,$$

hence $(\clubsuit)$ implies:

$$R(N)=A\binom{N+2}{2}+B\binom{N+1}{1}+\left[C+D(-1)^N+E\omega^{2N+2}+F\omega^{N+1}\right]$$

where the term in square brackets has an absolute value not exceeding $\frac{7}{12}$: in particular the simple poles of $f(z)$ do not give a significant contribution to the asymptotic behaviour of $R(N)$, while the triple pole at $z=1$ gives that $R(N)$ has a quadratic growth and fixes the coefficients $A=\frac{1}{6}$ and $B=\frac{1}{4}$: by exploiting the fact that $R(N)$ is certainly an integer, we get that:

$$R(N)\text{ is the closest integer to }\frac{(N+3)^2}{12}.$$

□

Exercise 396. Let $R(N)$ by the number of ways for paying an integer price $N\in\mathbb{N}^+$ through coins whose value is $1,2,5$ or 10. Prove that:

$$R(N)=\frac{N^2(N+27)}{600}+O(N).$$

[17] In general, the residue Theorem provides an alternative way for computing, through suitable limits, the coefficients of a partial fraction decomposition. Such approach is equivalent to solving a system of linear equations.

We solved an instance of the *Frobenius coin problem* through the manipulation of a *generating function* with singularities along the unit circle. In similar circumstances, a pole with order k at $\exp\left(2\pi i\frac{p}{q}\right)$ has an influence on the asymptotic behaviour of a_N that depends in the first place by the magnitude of k, in the second place by the magnitude of q. This observation led to the birth of Hardy, Littlewood and Ramanujan's **circle method**. Let us assume that a_N represents the number of ways for writing N as the sum of τ elements from some $A \subseteq \mathbb{N}$: in such a case,

$$f(x) = \sum_{n\geq 0} a_n x^n = \left(\sum_{a\in A} x^a\right)^{\tau} = g_A(x)^{\tau}$$

where g_A is a holomorphic function on the interior of the unit disk. It is possible to partition the unit circle through the set $\mathfrak{M}$ of **major arcs**, made by points of the form $\exp(2\pi i\theta)$, for some $\theta \in [0,1)$ close enough to a rational number with a small denominator, and the set $\mathfrak{m}$ of **minor arcs**, given by da $S^1 \setminus \mathfrak{M}$. The adjectives *major* and *minor* are related to the fact that major arcs usually give the most significant contribution to the asymptotic behaviour of a_N, despite the fact their measure can be very small or even negligible with respect to the measure of $\mathfrak{m}$. The behaviour of f at $\mathfrak{M}$ depends on particular exponential sums, while the behaviour of f at $\mathfrak{m}$ can be usually estimated through Cauchy-Schwarz, Hölder's or similar inequalities. The circle method has been a very effective technique for tackling a vast amount of problems in Analytic Number Theory having the following form:

Prove that any $n \in \mathbb{N}$ large enough can be written as the sum of (at most) τ elements from A.

but it often involved highly non-trivial strategies for producing tight bounds for the involved exponential sums, choosing the parameters defining $\mathfrak{m}$ and $\mathfrak{M}$ in a practical way, controlling the contribution provided by minor arcs. For instance it is often easier to deal with "weighted representations", i.e. to apply the original circle method to

$$f_\omega(x) = \sum_{n\geq 1} \omega_n x^n = \left(\sum_{a\in A} \omega(a) x^a\right)^{\tau}$$

where the weight-function $\omega(a)$ is suitably chosen according to the structure of A. Just to mention a technical difficulty, **Vinogradov's inequality** estimates f_ω in a neighbourhood of $\mathfrak{M}$ when A is the set of prime numbers, $\omega(n) = \log(n)$ and $\tau \geq 3$. It led its author to one of the greatest achievement of the circle method, namely the proof of the *ternary* Goldbach conjecture:

Theorem 397 (ternary Goldbach conjecture / Vinogradov's Theorem)**.**
Every odd and large enough natural number can be written as the sum of three primes.

but due to intrinsic limitations of the circle method it looks unlikely it might lead to a proof of the (binary) Goldbach conjecture, i.e. the statement *every even and large enough natural number is the sum of two primes.* Other remarkable achievements are related to **Waring's problem**:

Theorem 398. Every natural number large enough is the sum of $g(k)$ k-th powers.

and the behaviour of the partition function $p(n)$:

Theorem 399 (Ramanujan)**.** Let $p(N)$ be the number of representations of N as $n_1 + n_2 + \ldots + n_k$, where $n_1 \geq n_2 \geq \ldots \geq n_k$ and k are positive natural numbers. We have:

$$p(N) \sim \frac{1}{4N\sqrt{3}} \exp\left(\pi\sqrt{\frac{2N}{3}}\right).$$

As an excellent reference on the subject, we mention again Terence Tao's **blog**.
We now consider the following problem:

Exercise 400. Let $z_1, z_2, \ldots, z_n$ be distinct complex numbers and let us denote as $V = V(z_1, z_2, \ldots, z_n)$ the matrix

$$V = \begin{pmatrix} z_1^{n-1} & z_1^{n-2} & \ldots & 1 \\ z_2^{n-1} & z_2^{n-2} & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots \\ z_n^{n-1} & z_n^{n-1} & \ldots & 1 \end{pmatrix}.$$

Prove that V is invertible and compute its inverse matrix.

Proof. We may notice that solving the problem

$$V \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$$

is equivalent to solving the **interpolation** problem

$$\begin{cases} p(z_1) &= w_1 \\ p(z_2) &= w_2 \\ \ldots &= \ldots \\ p(z_n) &= w_n \end{cases}$$

where $p(z) = a_1 z^{n-1} + a_2 z^{n-2} + \ldots + a_n$. The Lagrange interpolating polynomial

$$q(z) = \sum_{r=1}^{n} w_r \prod_{k \neq r} \frac{z - z_k}{z_r - z_k}$$

provides a solution to such problem, and such solution is unique: assuming two distinct polynomials with degree $\leq (n-1)$ attain the value w_1 at $z = z_1$, the value w_2 at $z = z_2$, ... , the value w_n at $z = z_n$, their difference is a non-zero polynomial with degree $\leq (n-1)$ and at least n roots: a contradiction by Ruffini's rule. It follows that the existence of the Lagrange basis implies the non-singularity of the **Vandermonde** matrix V. Let us consider the case given by $w_1 = 1$ and $w_2 = \ldots = w_n = 0$. Due to Cramer's rule[18] we have that:

$$a_1 = \frac{1}{\det V} \det \begin{pmatrix} 1 & z_1^{n-2} & \ldots & 1 \\ 0 & z_2^{n-2} & \ldots & 1 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & z_n^{n-1} & \ldots & 1 \end{pmatrix} = \prod_{k \neq 1} \frac{1}{z_1 - z_k}$$

and by a Laplace expansion and induction on n we immediately get:

$$\det V(z_1, \ldots, z_n) = \prod_{1 \leq j < k \leq n} (z_j - z_k).$$

As an alternative, we may consider that $\det V(z_1, \ldots, z_n)$ is a (multivariate) polynomial with degree $\binom{n}{2}$ that vanishes every time $z_j = z_k$ for some $j \neq k$. That proves the above identity up to a multiplicative constant, where such constant can be found by an explicit evaluation of $\det V(z_1, \ldots, z_n)$ at $z_j = \exp\left(\frac{2\pi i j}{n}\right)$: in such a case the product between V and V^H is a diagonal matrix, as a consequence of the discrete Fourier transform.

[18] Although the first person proving and using such identity has been MacLaurin.

At last, it is simple to prove that the element appearing at the μ-th row and ν-th column of V^{-1} is given by:

$$\frac{\sigma_{\mu-1}\left(\{z_1, z_2, \ldots, z_n\} \setminus \{z_\nu\}\right)}{\prod_{k\neq\nu}(z_\nu - z_k)}$$

where $\sigma_0 = 1$ and $\sigma_\eta\left(\{u_1, \ldots, u_m\}\right)$, for $\eta \geq 1$, is the η-th elementary symmetric function of $u_1, u_2, \ldots, u_m$. □

The situation above is not the only case in which the inverse of a structured matrix still is a structured matrix. Another classical example is related to Stirling numbers of the first and second kind, where the convolution identity

$$\sum_{j=0}^{n} \begin{bmatrix} n \\ j \end{bmatrix} \begin{Bmatrix} j \\ k \end{Bmatrix} = \delta(n, k)$$

can be deduced by manipulating well-known generating functions, or just by noticing that both $\{1, x, x^2, \ldots, x^r\}$ and $\{1, x, x(x-1), \ldots, x(x-1) \cdot \ldots (x-r+1)\}$ provide a basis for the vector space of polynomials with degree $\leq r$.

The **Newton-Girard formulas** also appear in this context. Given n complex variables $z_1, \ldots, z_n$, let us denote as e_k the k-th elementary symmetric function of such variables, corresponding to the following sums of products:

$$e_0 = 1, \qquad e_1 = z_1 + z_2 + \ldots + z_n, \qquad e_2 = \sum_{i<j} z_i z_j, \qquad \ldots \qquad e_n = z_1 \cdot z_2 \cdot \ldots \cdot z_n.$$

Let us denote as p_k the sum of the k-th powers of such variables:

$$p_0 = n, \qquad p_k = \sum_{j=1}^{n} z_j^k.$$

Both $\{e_0, e_1, \ldots, e_n\}$ and $\{p_0, p_1, \ldots, p_n\}$ provide a basis for the ring of the symmetric functions in n variables, in particular:

$$\left\{\begin{array}{rcl} e_1 & = & p_1 \\ 2e_2 & = & e_1 p_1 - p_2 \\ 3e_3 & = & e_2 p_1 - e_1 p_2 + p_3 \\ 4e_4 & = & e_3 p_1 - e_2 p_2 + e_1 p_3 - p_4 \\ \ldots & = & \ldots \end{array}\right. \qquad \left\{\begin{array}{rcl} p_1 & = & e_1 \\ p_2 & = & e_1 p_1 - 2e_2 \\ p_3 & = & e_1 p_2 - e_2 p_1 + 3e_3 \\ p_4 & = & e_1 p_3 - e_2 p_2 + e_3 p_1 - 4e_4 \\ \ldots & = & \ldots \end{array}\right.$$

A slick proof of such identities through creative telescoping is due to Mead: for any $k > 0$, let $r(i)$ be the sum of all distinct monomials with degree k, given by the product of the $(k-i)$-th power of a variable and other i distinct variables. We have:

$$p_i e_{k-i} = r(i) + r(i+1), \qquad 1 < i < k$$

and trivially $p_k e_0 = p_k = r(k)$ and $p_1 e_{k-1} = k e_k + r(2)$, so by adding such identies with alternating signs we immediately get Newton's formulas. The classical proof goes as follows: by introducing

$$f(t) = \prod_{h=1}^{n}(1 - z_h t) = \sum_{h=0}^{n}(-1)^h e_h t^h$$

we are able to write $f'(t)$ as the product between $f(t)$ and its logarithmic derivative:

$$f'(t) = \sum_{h=1}^{n}(-1)^h h e_h t^{h-1} = \left(\sum_{h=0}^{n}(-1)^h e_h t^h\right) \cdot \left(-\sum_{h=1}^{n}\frac{z_h}{1 - z_h t}\right) = \left(\sum_{h=0}^{n}(-1)^h e_h t^h\right) \cdot \left(-\sum_{m\geq 0} p_{m+1} t^m\right)$$

so by comparing the coefficients of t^{k-1} in both sides:

$$(-1)^k k e_k = \sum_{j=1}^{k}(-1)^{k-j-1} p_j e_{k-j}.$$

In general:

$$\exp\left(-\sum_{m\geq 1}\frac{p_m}{m}x^m\right) = \sum_{r\geq 0}(-1)^r e_r x^r$$

and in order to derive the elementary symmetric functions from the power sums it is enough to apply an exponential map (EXP), while the opposite can be achieved by applying a logarithmic map (LOG). Morever it is possible to write the Newton-Girard formulas in a matrix form, in such a way that, by Cramer's rule, p_m can be written as the ratio of the determinants of two matrices with entries in $\{e_0, e_1, \ldots, e_n\}$, and vice versa.

Exercise 401. Compute the solutions (α, β, γ) of the following system of polynomial equations:

$$\begin{cases} \alpha+\beta+\gamma &= 2\\ \alpha^2+\beta^2+\gamma^2 &= 6\\ \alpha^3+\beta^3+\gamma^3 &= 8\end{cases}$$

Proof. It is enough to find the elementary symmetric functions e_1, e_2, e_3 of α, β, γ to get a cubic polynomial vanishing at α, β, γ. Since

$$\exp\left(-2x-3x^2-\frac{8}{3}x^3\right) = 1-2x-x^2+2x^3+O(x^4)$$

we have that α, β, γ are the roots of

$$x^3+2x^2-x-2$$

hence $\{\alpha, \beta, \gamma\} = \{-1, 1, 2\}$. □

It might be interesting to study the conditions ensuring that a problem like the above one has non-negative, real solutions. For sure, it is pretty simple to devise a necessary condition: if, for instance, $p_{k-1}p_{k+1} < p_k^2$, the Cauchy-Schwarz inequality is violated. A similar criterion for the elementary symmetric functions is the following one:

Theorem 402 (Newton's inequality)**.** If $1 = e_0, e_1, e_2, \ldots, e_n$ are the elementary symmetric functions associated with n real variables, by setting

$$S_k = \frac{e_k}{\binom{n}{k}}$$

(S_k is so the "mean contribution" provided by a monomial appearing in e_k) we have:

$$S_{k-1}S_{k+1} \leq S_k^2$$

for any $k \in \{1, 2, \ldots, n-1\}$.

Corollary 403 (MacLaurin's inequality)**.** If $1 = e_0, e_1, e_2, \ldots, e_n$ are the elementary symmetric functions associate with n non-negative real variables, by setting $S_k = e_k\binom{n}{k}^{-1}$ we have:

$$S_1 \geq \sqrt{S_2} \geq \sqrt[3]{S_3} \geq \ldots \geq \sqrt[n]{S_n}.$$

We may notice that MacLaurin's inequality gives many intermediate terms between the geometric mean $\sqrt[n]{S_n}$ and the arithmetic mean S_1.

Exercise 404. $p(x) \in \mathbb{Q}[x]$ is an irreducible polynomial over $\mathbb{Q}$ with degree 4, vanishing at $\alpha, \beta, \gamma, \delta$. Design an algorithm that manipulates the coefficients of $p(x)$ and returns a sixth-degree polynomial $q(x) \in \mathbb{Q}[x]$ vanishing at

$$\alpha\beta, \quad \alpha\gamma, \quad \alpha\delta, \quad \beta\gamma, \quad \beta\delta, \quad \gamma\delta.$$

Exercise 405. Given a $n \times n$ matrix A with real entries, such that

$$\text{Tr}(A) = \text{Tr}(A^2) = \ldots = \text{Tr}(A^n) = 0,$$

prove that A is nilpotent, i.e. $A^k = 0$ for some $k \in \mathbb{N}$.

Exercise 406. Prove that the number of partitions of n into **odd** parts equals the number of partitions of n into **distinct** parts.

Proof. The generating function for the partitions into odd parts is given by:

$$\frac{1}{1-x} \cdot \frac{1}{1-x^3} \cdot \frac{1}{1-x^5} \cdot \ldots = \prod_{k \geq 0} \frac{1}{1-x^{2k+1}},$$

while the generating function for the partitions into distinct parts is given by:

$$(1+x)(1+x^2)(1+x^3) \cdot \ldots = \prod_{k \geq 1}(1+x^k) = \prod_{k \geq 1} \frac{1-x^{2k}}{1-x^k} = \frac{\prod_{k \text{ even}}(1-x^k)}{\prod_k (1-x^k)} = \prod_{k \text{ odd}} \frac{1}{(1-x^k)}.$$

We invite the reader to find a purely combinatorial proof in terms of Ferrers diagrams / Young tableaux and bijective maps. □

Exercise 407. If we pay 100 dollars by using only 1, 2 or 3 dollars bills in one of the possible $R(100)$ ways, what is the expected number of the used bills?

Proof. Let us consider the two-variables generating function

$$\begin{aligned} f(x,y) &= \left(1+yx+y^2x^2+y^3x^3+\ldots\right)\left(1+yx^2+y^2x^4+y^3x^6+\ldots\right)\left(1+yx^3+y^2x^6+y^3x^9+\ldots\right) \\ &= \frac{1}{(1-xy)(1-x^2y)(1-x^3y)}. \end{aligned}$$

The coefficient of x^{100} is a polynomial in the y variables which encodes the information concerning how many bills we used to pay the given price. For instance, if such polynomial were

$$q(y) = 27y^{30} + 51y^{32} + 87y^{36} + \ldots$$

that would mean we had 27 ways for paying the given amount through 30 bills, 51 ways for paying the given amount through 32 bills, 87 ways for paying the given amount through 36 bills etcetera. In particular the expected number of bills is given by $\frac{q'(1)}{q(1)} = \frac{d}{dy}\log q(y)\Big|_{y=1}$, i.e. by the ratio between the coefficient of x^{100} in

$$\frac{x+3x^2+4x^3+3x^4}{(1-x)^4(1+x)^2(1+x+x^2)^2}$$

and the coefficient of x^{100} in

$$\frac{1}{(1-x)^3(1+x)(1+x+x^2)}.$$

Keeping track of the contributions provided by the singularities at $x=1$ only, we get that the answer is a number close to

$$\frac{11}{36}\binom{103}{3}\cdot\left(\frac{1}{6}\binom{102}{2}\right)^{-1}\approx\frac{11}{18}\cdot 100$$

that is a pretty accurate approximation of the exact answer $\frac{q'}{q}(1)=\frac{3195}{52}=61.4423\ldots$ □

Exercise 408 (Coupon collector's problem)**.** In order to complete an album, N distinct stickers are needed. Such stickers are sold in packets containing just one sticker, and every sticker has the same probability to be inside some packet. What is the average number of packets we need to buy to complete the album?

Proof. If we assume that our album already contains n stickers, a packet just bought contains a sticker we already have with probability $\frac{n}{N}$ and a new sticker with probability $\frac{N-n}{N}$. In particular the average number of packets we need to buy to acquire a new sticker is $\frac{N}{N-n}$, and the average number of packets we need to buy to complete the whole album is:

$$\frac{N}{N}+\frac{N}{N-1}+\frac{N}{N-2}+\ldots+\frac{N}{1}=NH_N=N\log N+\gamma N+\frac{1}{2}+o(1).$$

□

Due to **Cantelli's inequality**, the probability that in order to complete the album with need to but more than NH_N+cN packets is $\leq\frac{\pi^2}{12c^2}$. Erdös and Rényi proved that as $N\to+\infty$, such probability converges to

$$1-e^{-e^{-c}}.$$

Exercise 409. What is the probability that a random element of $\sigma\in S_{12}$ is an involution, i.e. a map such that $\sigma=\sigma^{-1}$?

Proof. An element of S_n is an involution if and only if its decomposition into disjoint cycles is made by fixed points and transpositions only. Let us denote as L_n the number of elements of S_n that decompose through cycles having lengths $\in L=\{l_1,\ldots,l_k\}$. We have:

$$\sum_{n\geq 0}\frac{L_n}{n!}z^n=\exp\left(\sum_{l_i\in L}\frac{z^{l_i}}{l_i}\right),$$

hence in our case the wanted probability is given by the coefficient of z^{12} in $\exp\left(z+\frac{z^2}{2}\right)$, i.e.:

$$\sum_{k=0}^{6}\frac{1}{2^k k!}\cdot\frac{1}{(12-2k)!}=\frac{1}{2^6\cdot 6!}\left(1+\sum_{k=1}^{6}\binom{6}{k}\frac{1}{(2k-1)!!}\right)=0.000292591924536368 98\ldots$$

A simpler upper bound is provided by:

$$\frac{1}{2^6 6!}\left(1+\sum_{k\geq 1}\frac{6^k}{k!(2k-1)!!}\right)=\frac{\cosh(2\sqrt{3})}{2^6\cdot 6!}\leq 0.000347.$$

□

Exercise 410 (Bernstein's limit)**.** Prove that:

$$\lim_{n\to+\infty} e^{-n}\sum_{k=0}^{n}\frac{n^k}{k!}=\frac{1}{2}.$$

Proof. Due to the Taylor formula with integral remainder,

$$\sum_{k=0}^{n}\frac{n^k}{n!}=e^n-\frac{1}{n!}\int_0^n (n-t)^n e^t\,dt$$

hence the claim is equivalent to:

$$\lim_{n\to+\infty}\frac{n^{n+1}}{n!e^n}\int_0^1\left((1-t)e^t\right)^n\,dt=\frac{1}{2}.$$

On the interval $(0,1)$ the inequality $(1-t)e^t\leq\sqrt{1-t^2}$ holds, hence:

$$\int_0^1\left((1-t)e^t\right)^n\,dt\leq\frac{1}{2}\int_0^1(1-t)^{n/2}t^{-1/2}\,dt=\frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(1+\frac{n}{2}\right)}{2\,\Gamma\left(\frac{3}{2}+\frac{n}{2}\right)}$$

and by Stirling's inequality, if the wanted limit exists, it certainly is $\leq\frac{1}{2}$. It follows that we just need to prove that the approximation $(1-t)e^t\approx\sqrt{1-t^2}$ is accurate enough to provide the exact main term of the asymptotic expansion. Since on the interval $(0,1)$ we have:

$$0\leq\sqrt{1-t^2}-(1-t)e^t\leq\frac{t^3}{3},$$

it follows that:

$$0\leq\int_0^1(1-t^2)^{n/2}\,dt-\int_0^1\left((1-t)e^t\right)^n\,dt\leq n\int_0^1\frac{t^3}{3}(1-t^2)^{n/2-1}\,dt=\frac{2}{3n}+O\left(\frac{1}{n^2}\right)$$

and the proof is complete. □

The statement just proved is usually is usually approached by applying the strong law of large numbers to the Poisson distribution.

Exercise 411. Prove that for any $x\in(0,1)$ the following identity holds:

$$\frac{x}{1-x}=\sum_{k\geq 0}\frac{2^k x^{2^k}}{1+x^{2^k}}.$$

Exercise 412 (Erdös)**.** $S=\{a_1,\ldots,a_n\}$ is a set of positive natural number with the property that disjoint subsets of S have different sums of their elements. Prove that:

$$\sum_{k\geq 1}\frac{1}{a_k}<2.$$

The circle method for locating the roots of polynomials.

Lemma 413. If z is a complex number whose real part is $\neq 0$ and $f(z) = \frac{1}{2}\left(z + \frac{1}{z}\right)$, the sequence defined by

$$z_0 = z, \qquad z_{n+1} = f(z_n)$$

is rapidly convergent to $\text{Sign}(\text{Re}\, z)$.

Lemma 414. If $h(z)$ is an entire function and $p(z)$ is a monic polynomial with roots $\zeta_1, \ldots, \zeta_r$ (counted according to their multiplicity) in the region $\|z\| < 1$,

$$\sum_{k=1}^{r} h(\zeta_k) = \frac{1}{2\pi i} \oint_{\|z\|=1} \frac{p'(z)}{p(z)}\, h(z)\, dz.$$

These preliminary results allow us to design an algorithm for locating the roots of a polynomial through a *divide et impera* approach. Given a non-constant $p(z) \in \mathbb{C}[z]$, it is simple to apply a translation to the z variable in such a way that $p(z)$ has about the same number of zeroes in the right halfplane H_+ and in the left halfplane H_- (for instance by performing substitution in such a way that $[z^{\partial p - 1}]p(z) = 0$). With such assumptions we have:

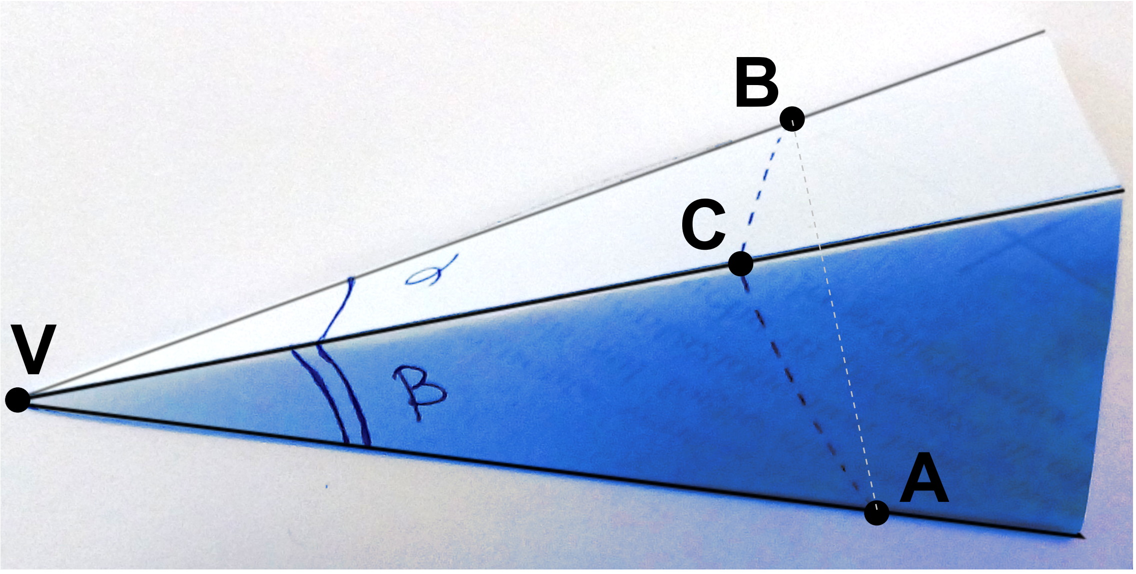
$$p(z) = p_+(z) \cdot p_-(z), \qquad p_+(z) = \prod_{\zeta_k \in H_+} (z - \zeta_k), \qquad p_-(z) = \prod_{\zeta_k \in H_-} (z - \zeta_k),$$

and by applying (multiple times if needed) the transformation f of the Lemma 413 we may also assume without loss of generality that all the roots of $p_-(z)$ lie in the region $\|z + 1\| \leq \frac{1}{2}$ and all the roots of $p_+(z)$ lie in the region $\|z - 1\| \leq \frac{1}{2}$. Let us assume that $\zeta_1, \ldots, \zeta_m$ are the roots of $p_+(z)$. In order to find the coefficients of p_+ it is enough to compute the elementary symmetric functions $e_1, e_2, \ldots, e_m$ of $\zeta_1, \zeta_2, \ldots, \zeta_m$, where the values of $e_1, e_2, \ldots, e_m$ can be deduced from the values $p_1, p_2, \ldots, p_m$ by the Newton-Girard formulas. Additionally, as a consequence of the Lemma 414:

$$p_h = \frac{1}{2\pi i} \oint_{\|z-1\|=1} \frac{p'(z)}{p(z)}\, z^h\, dz$$

hence $p_1, \ldots, p_m$ can be approximated by applying quadrature formulas to m integrals of regular functions, depending on p and p'. From the approximated values of $\tilde{p}_1, \ldots, \tilde{p}_m$ it is simple to find $\tilde{e}_1, \ldots, \tilde{e}_m$, i.e. the coefficients of $\tilde{p}_+$. Such approximation of p_+ can be improved by applying a variant of Newton's method, leading to approximations for the coefficients of p_+ and p_- with an arbitrary degree of precision. At this point it is enough to repeat the same steps till decomposing $p(z)$ in linear factors and finding the whole set of roots.

16 Spherical Trigonometry



Exercise 415. We know that $VV_AV_BV_C$ is a tetrahedron (i.e. a pyramid with a triangular base) in the Euclidean space and we know the amplitudes of

$$\alpha = \widehat{V_BVV_C}, \quad \beta = \widehat{V_AVV_C}, \quad \gamma = \widehat{V_AVV_B}.$$

How is it possible to find the angles between two faces meeting at V?

Proof. Let us consider three points A, B, C on the edges from V in such a way that $VC = 1$ and both the triangles VCA and VCB have a right angle at C. With such assumptions the angle between the faces VV_AV_C and VV_BV_C is exactly the angle $\widehat{ACB}$. We have that $VA = \frac{1}{\cos\beta}$ and $VB = \frac{1}{\cos\alpha}$, as well as $CA = \tan\beta$ and $CB = \tan\alpha$. Then the length of AB can be found by applying the cosine Theorem to the triangle AVB or to the triangle ACB:

$$AB^2 = AV^2 + BV^2 - 2\,AV \cdot BV \cos\gamma = AC^2 + BC^2 - 2\,AC \cdot BC \cos\widehat{ACB}.$$

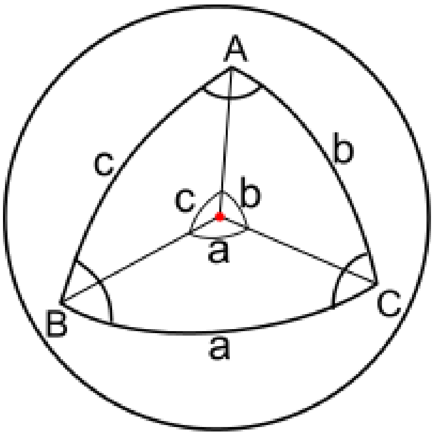
Due to such identities we have:

$$\frac{1}{\cos^2\beta} + \frac{1}{\cos^2\alpha} - \frac{\sin^2\beta}{\cos^2\beta} - \frac{\sin^2\alpha}{\cos^2\alpha} = \frac{2}{\cos\alpha\cos\beta}\cos\gamma - \frac{2\sin\alpha\sin\beta}{\cos\alpha\cos\beta}\cos\widehat{ACB}$$

where the last equality can be simplified as follows:

$$\sin\alpha\sin\beta\cos\widehat{ACB} = \cos\gamma - \cos\alpha\cos\beta.$$

We may get the angle between the faces sharing the VV_A edge or the VV_V edge in a similar way, by applying a suitable permutation to the variables A, B, C. □



Let us assume that A, B, C are three distinct points on a sphere centered at O with unit radius. Geodesics (paths of minimum length) between two of the previous points are given by arcs of maximal circles (sections delimited by the intersection between the sphere and a plane through its center), whose lengths equal the amplitudes of the angles

$$a = \widehat{BOC}, \quad b = \widehat{AOC}, \quad c = \widehat{AOB}.$$

The smallest region of the sphere delimited by such geodesics is known as **spherical triangle**. Let us denote through A, B, C also the angles between such geodesics on the sphere: in particular A is the angle between the planes BAO and CAO. We may notice two maximal circles always intersect at two antipodal points on the sphere:
in particular, by naming as **segments** the geodesics and as **lines** their prolongations, i.e. the intersections between the sphere and the planes through its center, we have that by taking as a **distance** the geodetic distance the **spherical geometry** meets the first four Euclid's axioms, **but not the fifth**: given a line r and a point P not belonging to such line, it does not exist any line through P that is parallel to r. A pretty straightforward consequence of this "violation" is that theorems about the amplitude of "external angles" in a triangle cease to hold, hence the sum of the angle of a spherical triangle, $A+B+C$, **is not** costant. However, it is possible to approach the study of spherical trigonometry like elementary trigonometry is usually approached: we are going to study the relations between sides and angles in spherical triangles (producing analogues of the sine and cosine Theorems) and how to write the area in terms of three fundamental elements (three sides, two sides and an angle, one side and two angles). The previous exercise provides an excellent starting point:

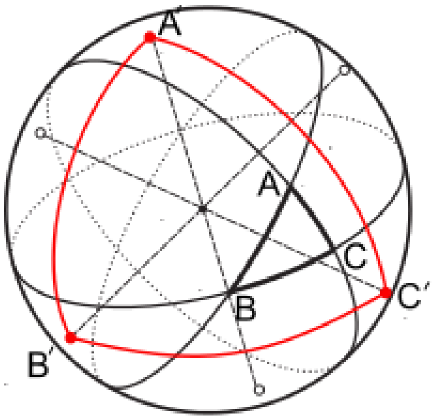
Theorem 416 (From sides to angles)**.** According to the notation just introduced, in a spherical triangle on a unit sphere we have:

$$\cos A = \frac{\cos a - \cos b \cos c}{\sin b \sin c}, \quad \cos B = \frac{\cos b - \cos a \cos c}{\sin a \sin c}, \quad \cos C = \frac{\cos c - \cos a \cos b}{\sin a \sin b}.$$

In spherical geometry the concepts of **symmetry** and **duality** have a great relevance.
They lead, for instance, to the following result:

Theorem 417 (From angles to sides)**.** According to the notation just introduced, in a spherical triangle on a unit sphere we have:

$$\cos a = \frac{\cos A + \cos B \cos C}{\sin B \sin C}, \quad \cos b = \frac{\cos B + \cos A \cos C}{\sin A \sin C}, \quad \cos c = \frac{\cos C + \cos A \cos B}{\sin A \sin B}.$$



Proof. Given a spherical triangle with vertices at A, B, C, we may define its **dual** triangle $A'B'C'$ as follows: we consider the maximal circle given by B, C, O and the perpendicular through O to the plane π_{BC} containing such circle. This line meets the sphere at two antipodal points: if we denote as A' the intersection lying on the same side of π_{BC} with respect to A, we may proceed in a similar way for defining B' and C'.
It is simple to notice that the angle between OA' and OB' and the angle between the planes ACO and BCO are supplementary, hence the amplitude of the former angle is $\pi - C$. Morever in the spherical triangle $A'B'C'$ the angles C' and $\widehat{AOB}$ are supplementary, hence the amplitude of the former angle is $\pi - c$.
Given the relations allowing us to find the angles of a spherical triangle ABC from its sides, the inverse relations (allowing us to write the sides as functions of the angles) can by simply deduced by replacing c with $\pi - C$, C with $\pi - c$ and so on. Since $\cos(\pi - \theta) = -\cos\theta$ and $\sin(\pi - \theta) = \sin\theta$, the claim immediately follows. This brilliant idea is due to the Persian mathematician Abu Nasr Mansur, ≈1000 a.C. □

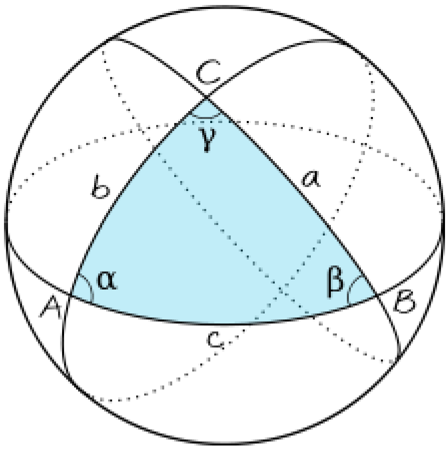
We have so far collected the equivalent versions of the sine and cosine Theorems for planar triangles.
It is straightforward to state the Pythagorean Theorem for spherical triangles:

Theorem 418 (Spherical Pythagorean Theorem)**.** If a spherical triangle ABC lies on the surface of a unit sphere and fulfills $C = \frac{\pi}{2}$, we have:

$$\cos(c) = \cos(a)\cos(b)$$

due to $\cos C = 0$.

Corollary 419. If a spherical triangle with $C = \frac{\pi}{2}$ lies on the surface of a sphere with radius R we have $\cos\frac{c}{R} = \cos\frac{a}{R}\cos\frac{b}{R}$; since for angles close to zero we have $\cos\theta \approx 1 - \frac{\theta^2}{2}$ the Pythagorean Theorem for planar triangles ($c^2 = a^2 + b^2$) can be seen as a limit case (for $R \to +\infty$) of the spherical Pythagorean Theorem.



Leu us consider a spherical triangle ABC on a unit sphere, having all its angles A, B, C between 0 and $\frac{\pi}{2}$. We may denote as π_{AB} the plane through A, B, O and define in a similar way π_{BC} and π_{AC}: the planes π_{AC} and π_{AB} meet at A and at the antipode of A. It follows that by prolongating the sides of ABC we get an antipodal spherical triangle. The surface area of the spherical wedge delimited by the planes π_{AB} and π_{AC} is clearly given by the product between $\frac{2A}{2\pi}$ and the surface area of the sphere, namely 4π. If we consider the planes $\pi_{AB}, \pi_{AC}, \pi_{BC}$ pairwise and sum the surface areas of the delimited spherical wedges, we get the surface area of the sphere plus four times the surface area of the spherical triangle ABC:
$4A + 4B + 4C = 4\pi + 4[ABC]$.

In particular, by introducing the **spherical excess** (also known as *angular excess*) of ABC as $E = A + B + C - \pi$, we have:

Theorem 420 (Area of a spherical triangle)**.** The spherical excess E of a spherical triangle is always positive and the area of a spherical triangle is simply given by ER^2, where R is the sphere radius.

E can be computed from the side lengths by a spherical analogue of Heron's formula:

Theorem 421 (l'Huilier). If ABC is a spherical triangle on a unit sphere with spherical excess E and semiperimeter s, we have:

$$\tan\frac{E}{4} = \sqrt{\tan\frac{s}{2}\tan\frac{s-a}{2}\tan\frac{s-b}{2}\tan\frac{s-c}{2}}.$$

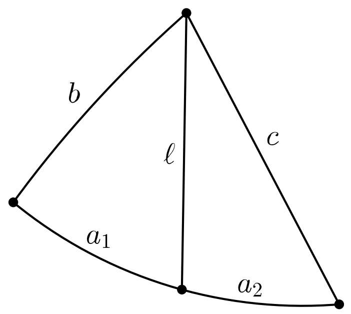
Remark: the triangle inequality $a+b>c$ still holds for spherical triangles, since sides still are geodesics. The motto "*in order to go from Paris to London, it is not very practical to go through Moskow*" applies to plane geometry and to spherical geometry just as well.

Theorem 422 (Sine Theorem for spherical triangles). If ABC is a spherical triangle on a unit sphere centered at O, we have:

$$\frac{\sin A}{\sin a} = \frac{\sin B}{\sin b} = \frac{\sin C}{\sin c} = \frac{6\,\text{Vol}(OABC)}{\sin a \sin b \sin c}.$$

Proof. Just like in the Euclidean case, it is enough to compare two different expressions for the measure of a geometric object (in the Eucliden case, a triangle, in the spherical case, a tetrahedron). As an alternative one may use the identity $\sin^2\theta = 1-\cos^2\theta$ to deduce the above statement from the formulas for $\cos a$ and $\cos A$. □

Analogies with the Euclidean case go further: we may define a (spherical) **cevian** as a segment (i.e. a geodesic) joining a vertex of a triangle with some point on the opposite side. In the Euclidean case the instrument allowing us to compute the length of a cevian is Stewart's Theorem, which can be deduced from the cosine Theorem by noticing two supplementary angles sharing the foot of such cevian as common vertex. The same approach works in the spherical case, too:

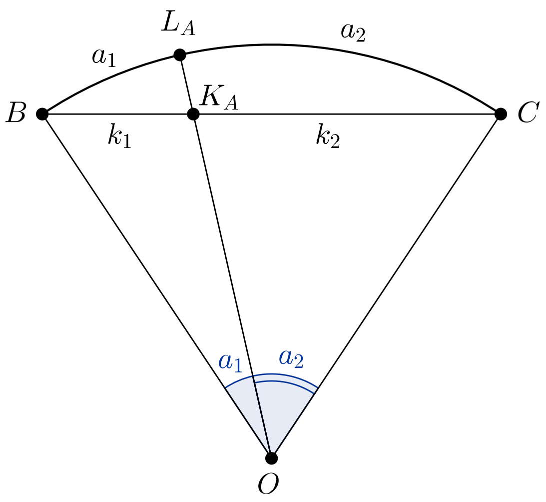


Theorem 423 (Stewart's Theorem, spherical version). The diagram represents a spherical triangle on a unit sphere and a cevian with length ℓ with its foot on the a side. In such configuration we have:

$$\cos\ell = \frac{\sin a_1 \cos c + \sin a_2 \cos b}{\sin a}.$$

Let us assume that $a_1, a_2, b_1, b_2, c_1, c_2$ (in this order) are the lengths of the segments cut by three cevians on the triangle sides. It is pretty natural to wonder if the concurrence of those cevians is equivalent to some algebraic identity involving the previous lengths.

By denoting as π_{ABC} the plane given by A, B, C, every segment with endpoints F, G, belonging to the interior of ABC, can be associated with a straight segment on the plane π_{ABC}: it is enough to consider the intersection between π_{ABC} and the plane triangle FGO. Three spherical cevians are concurrent if and only if their associated segments on the plane π_{ABC} are concurrent, so the spherical version of Ceva's Theorem can be deduced from applying the usual version of Ceva's Theorem to the associated configuration in the plane π_{ABC}. In order to do that, of course, we need to compute the lengths of a few straight segments.



Let us denote as L_A the foot of the spherical cevian through A and let us consider the plane containing B, L_A, C and O. By denoting as K_A the intersection between the straight segments OL_A and BC on such plane we have that K_A is the foot of the associated cevian. We just need to compute the ratio $\frac{k_1}{k_2} = \frac{BK_A}{K_AC}$.

We may notice that BK_AO and CK_AO have the same heigth with respect to the BC base. If we assume that the angles $\widehat{BOK_A} = \widehat{BOL_A}$ and $\widehat{COK_A} = \widehat{COL_A}$ have amplitudes a_1 and a_2, we get:

$$\frac{k_1}{k_2} = \frac{[BK_AO]}{[CK_AO]} = \frac{\sin a_1}{\sin a_2}$$

from which it is simple to deduce the following statement.

Theorem 424 (Ceva's Theorem, spherical version)**.** Three spherical cevians are concurrent if and only if

$$\frac{\sin a_1}{\sin a_2} \cdot \frac{\sin b_1}{\sin b_2} \cdot \frac{\sin c_1}{\sin c_2} = 1.$$

Corollary 425 (Existence of the spherical centroid)**.** The spherical cevians joining the vertices of a spherical triangle with the midpoints of the opposite sides are concurrent. If we denote as G the centroid of the plane triangle ABC, the common intersection of the previous spherical cevians lies on the ray OG, with O being the center of the sphere.

Corollary 426 (Existence of the spherical Gergonne point)**.** Given a spherical triangle ABC, let us denote as D, E, F the tangency points of the inscribed circle on the sides a, b, c. The spherical cevians AD, BE, CF are concurrent.

Let us denote as $A_1, A_2, B_1, B_2, C_1, C_2$ (in this order) the angles cut by three concurrent spherical cevians ($A_1 + A_2 = A$ and so on). Let ℓ be the length of the spherical cevian through A and let D and $\pi - D$ the angles adjacent to the foot of such cevian. By the spherical version of the sine Theorem we have::

$$\frac{\sin D}{\sin b} = \frac{\sin C}{\sin \ell} = \frac{\sin A_2}{\sin a_2}, \qquad \frac{\sin D}{\sin c} = \frac{\sin B}{\sin \ell} = \frac{\sin A_1}{\sin a_1}$$

hence the spherical version of Ceva's Theorem has the following equivalent form:

Theorem 427 (Spherical Ceva Theorem, equivalent form)**.** Three spherical cevians are concurrent if and only if

$$\frac{\sin A_1}{\sin A_2} \cdot \frac{\sin B_1}{\sin B_2} \cdot \frac{\sin C_1}{\sin C_2} = 1.$$

Corollary 428 (Existence of the spherical incenter)**.** The spherical cevians that divide the angles of A, B, C in halves are concurrent. This also follows from the first form of Ceva's Theorem for spherical triangles: if Γ is a circle on the surface of a sphere, B, C are two distinct points on Γ and both the geodesics AB and AC are tangent to Γ, then AB and AC have the same lengths. This is straightforward to prove by symmetry, just like in the Euclidean case.

Corollary 429. Just like in the Euclidean case, the two forms of Ceva's Theorem allow to define the concepts of **isogonal conjugate** and **isotomic conjugate** for spherical triangles, too.

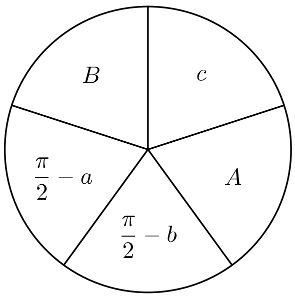
Exercise 430. By suitably modifying the proof of the spherical version of Ceva's Theorem, state and prove the spherical versions of Ptolemy and Van Obel's Theorems.

We now introduce some useful Lemmas on right spherical triangles and some bisection formulas. These results will later allow us to prove the existence of the spherical orthocenter, to find the radius r of the incirle, to find the radius R of the circumcircle and much more.

Lemma 431. Let ABC be a spherical triangle with $C = \frac{\pi}{2}$. We have:

$$\cos A = \frac{\tan b}{\tan c}, \qquad \tan A = \frac{\tan a}{\sin b}.$$

Proof. Due to the spherical versions of the sine, cosine and Pythagorean Theorems, the following identities hold:



$$\begin{aligned}\cos A &= \frac{\cos a - \cos b \cos c}{\sin b \sin c} = \frac{\cos a - \cos^2 b \cos a}{\sin b \sin c} \\ &= \frac{\cos a \sin b}{\sin c} = \frac{\cos c \sin b}{\cos b \sin c} = \frac{\tan b}{\tan c}, \\ \tan A &= \frac{\sin A \tan c}{\tan b} = \frac{\sin a \tan c}{\tan b \sin c} \\ &= \frac{\sin a \cos b}{\sin b \cos c} = \frac{\sin a}{\sin b \cos a} = \frac{\tan a}{\sin b}.\end{aligned}$$

As a by-product we have just proved many other interesting identities, which can be recalled by a slick mnemonic trick. The diagram represents **Napier's pentagon**: the cosine of some angle equals the product of the cotangents of its neighbours, or the product of the sines of the two opposite angles. □

There are similar identities for the dual of a right spherical triangle, i.e. a triangle with a side equal to one fourth of a maximal circle, also known as **quadrantal triangle**. We recall that in order to apply Mansur's duality it is enough to replace the angle w with the angle $\pi - W$ and the angle W with the angle $\pi - w$. If we denote as s the semiperimeter of a spherical triangle, $s = \frac{a+b+c}{2}$, we have the following result:

Lemma 432 (bisection formulas for angles)**.** In a spherical triangle ABC we have:

$$\sin \frac{A}{2} = \sqrt{\frac{\sin(s-b)\sin(s-c)}{\sin b \sin c}}, \qquad \cos \frac{A}{2} = \sqrt{\frac{\sin(s)\sin(s-a)}{\sin b \sin c}}.$$

Proof. By the addition formulas for the sine and cosine function and the spherical version of the cosine Theorem we have:

$$2\sin(s-c)\sin(s-b) = \cos(b-c) - \cos(a) = \sin b \sin c(1 - \cos A),$$

$$2\sin(s)\sin(s-a) = \cos(a) - \cos(b+c) = \sin b \sin c(1 + \cos A).$$

The claim follows from the fact that in a spherical triangle the amplitude of each angle is $< \pi$, hence both $\sin \frac{A}{2}$ and $\cos \frac{A}{2}$ are positive. By applying Mansur's duality we also have bisection formulas for sides: we may notice that such duality essentially maps the semiperimeter s into the semi-excess $E/2$ and vice versa.

Lemma 433 (bisection formulas for sides)**.** In a spherical triangle ABC we have:

$$\sin \frac{a}{2} = \sqrt{\frac{\sin \frac{E}{2} \sin\left(A - \frac{E}{2}\right)}{\sin B \sin C}}, \qquad \cos \frac{a}{2} = \sqrt{\frac{\sin\left(B - \frac{E}{2}\right)\sin\left(C - \frac{E}{2}\right)}{\sin B \sin C}}.$$

We may notice that the sine Theorem is a consequence of such bisection formulas:

$$\frac{\sin A}{\sin a} = \frac{2\sin\frac{A}{2}\cos\frac{A}{2}}{\sin a} = \frac{2n}{\sin a \sin b \sin c}$$

where the **staudtian** $n = \sqrt{\sin(s)\sin(s-a)\sin(s-b)\sin(s-c)}$ bears a striking resemblance to Heron's formula. We have already proved the staudtian is just the volume of the tetrahedron $OABC$, up to a fixed multiplicative constant. The bisection formulas for sides also reveal an interesting relation between the staudtian and the spherical excess.

Theorem 434 (Cagnoli)**.**

$$\sin\frac{E}{2} = \frac{n}{2\cos\frac{a}{2}\cos\frac{b}{2}\cos\frac{c}{2}}.$$

Proof. By the bisection formulas for sides we have:

$$\sin\frac{E}{2} = \frac{\sin\frac{b}{2}\sin\frac{c}{2}\sin A}{\cos\frac{a}{2}}$$

and by invoking the sine Theorem we may conclude that $\sin A = \frac{2n}{\sin b\sin c} = \frac{n}{2\sin\frac{b}{2}\sin\frac{c}{2}\cos\frac{b}{2}\cos\frac{c}{2}}$. □

Exercise 435. A regular icosahedron (regular polyhedron with 20 triangular faces) is inscribed in a unit sphere. Find the length of its edges.

Proof. Let us denote as O the center of the circumscribed sphere and let us consider the vertices A, B, C of some face. We may consider the spherical triangle ABC given by the intersections of the planes OAB, OAC, OBC with the surface of the sphere. ABC is an equilateral spherical triangle and its area equals $\frac{1}{20}$ times the surface area of the sphere, hence $E = \frac{4\pi}{20} = \frac{\pi}{5}$ and $A = B = C = \frac{\pi+E}{3} = \frac{2\pi}{5}$. By denoting as ℓ the edge length, the bisections formulas for sides allow us to state:

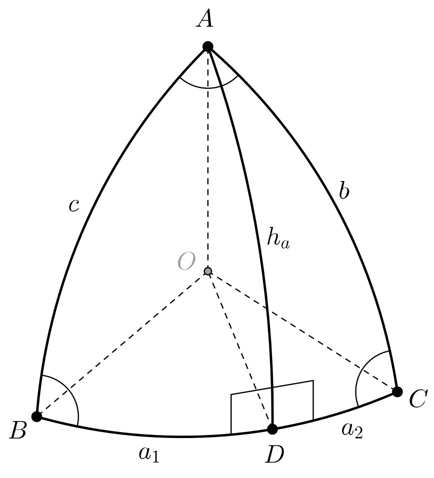
$$\ell = 2\sin\frac{a}{2} = 2\sqrt{\frac{\sin\frac{E}{2}\sin\left(A-\frac{E}{2}\right)}{\sin B\sin C}} = \frac{2}{\sin\frac{\pi}{5}}\sqrt{\sin\frac{\pi}{10}\sin\frac{3\pi}{10}} = \sqrt{2-\frac{2}{\sqrt{5}}}.$$

We also have:

$$\text{Vol}(OABC) = \frac{n}{3} = \frac{\sin b\sin c\sin A}{3} = \frac{(1-\cos^2 a)\sin A}{3}$$

and from $\cos a = \frac{\cos A - \cos^2 A}{\sin^2 A}$ it is simple to find the volume of the icosahedron, $20\cdot\text{Vol}(OABC)$.

Exercise 436. Find the edge length and the volume of a regular dodecahedron (polyhedron with 12 pentagonal faces) inscribed in a unit sphere.



Exercise 437 (Existence of the spherical orthocenter)**.** Prove that in any spherical triangle the altitudes (geodetics through some vertex, orthogonal to the opposite side) are concurrent.

Proof. Let us denote as h_a the altitude through A and as D its foot on a. By setting $BD = a_1$ and $DC = a_2$ we have that both ABD and ACD are right triangles. Due to the identities provided by Napier's pentagon:

$$\sin(a_1) = \frac{\tan h_a}{\tan B}, \qquad \sin(a_2) = \frac{\tan h_a}{\tan C}, \qquad \frac{\sin a_1}{\sin a_2} = \frac{\tan C}{\tan B}$$

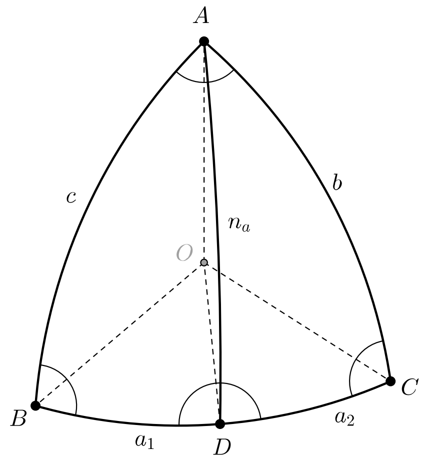
so the claim immediately follows from the spherical version of Ceva's Theorem.

About the spherical circumcenter, its existence is quite trivial: the plane π_{ABC} meets the surface of the sphere at the circumscribed circle of ABC. On a sphere only a point Q and its antipode have the property that the geodetic distances QA, QB, QB are equal; additionally, the perpendicular to a side through its midpoint, just like in the Euclidean case, is the locus of points equating two geodetic distances. It follows that the spherical circumcenter lies on each spherical perpendicular bisector, and the line joining the spherical circumcenter with the circumcenter of the plane triangle ABC goes through the center of the sphere.
We are ready to find algebraic expressions for the inradius and circumradius of a spherical triangle.

- **Inradius** r. The incenter is a point lying on each angle bisector, both in the Euclidean and in the spherical case. If we draw the geodesics joining the incenter with its orthogonal projections on the triangle sides, then the geodesics joining the incenter with the vertices of our spherical triangle ABC, such triangle is partitioned into six right triangles having a leg with length r and the other leg having length $\in \{s-a, s-b, s-c\}$. By exploiting the identities provided by Napier's pentagon we have $\tan\frac{A}{2} = \frac{\tan r}{\sin(s-a)}$; by applying the bisection formulas for angles we have:
$$\tan r = \sqrt{\frac{\sin(s-a)\sin(s-b)\sin(s-c)}{\sin s}} = \frac{n}{\sin s}.$$
- **Circumradius** R. The circumcenter is a point lying on each perpendicular bisector, both in the Euclidean and in the spherical case. If we draw the geodesics joining the circumcenter with the vertices of our triangle ABC and the midpoints of its sides, such triangle is partitioned into six right triangles having hypotenuse R and a leg with length $\in \{\frac{a}{2}, \frac{b}{2}, \frac{c}{2}\}$. By considering the dual version of the previous argument we get:
$$\cot R = \sqrt{\frac{\sin\left(A-\frac{E}{2}\right)\sin\left(B-\frac{E}{2}\right)\sin\left(C-\frac{E}{2}\right)}{\sin\frac{E}{2}}} = \frac{n}{2\sin\frac{a}{2}\sin\frac{b}{2}\sin\frac{c}{2}}.$$

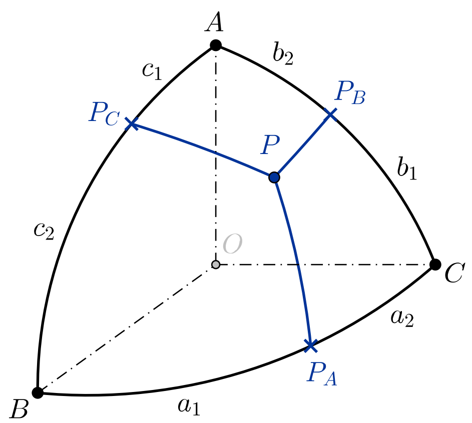
Corollary 438. Euler's inequality for plane triangles, $R \geq 2r$, has the following spherical analogue:
$$\frac{\tan R}{\tan r} \geq 2.$$

We will now introduce a couple of results due to Steiner.



Exercise 439 (Existence of the pseudocentroid)**.** In a spherical triangle we say that a cevian is a **pseudomedian** if it divides the triangle in two regions with the same area. Prove that the pseudomedians are concurrent.

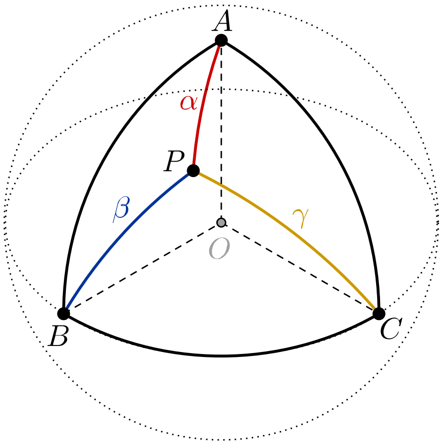
Proof. Let us assume that n_a is the pseudomedian through A, with its foot at D. Since the area is additive the spherical excess is additive as well, hence both the triangles ABD and ACD have an excess equal to $\frac{E}{2}$. By the bisection formulas for sides:
$$\sin\frac{n_a}{2} = \sqrt{\frac{\sin\frac{E}{4}\sin\left(B-\frac{E}{4}\right)}{\sin\widehat{BAD}\sin\widehat{ADB}}} = \sqrt{\frac{\sin\frac{E}{4}\sin\left(C-\frac{E}{4}\right)}{\sin\widehat{CAD}\sin\widehat{ADC}}}$$
but $\widehat{ADB}$ and $\widehat{ADC}$ are supplementary angles, hence $\frac{\sin\widehat{BAD}}{\sin\widehat{CAD}} = \frac{\sin\left(B-\frac{E}{4}\right)}{\sin\left(C-\frac{E}{4}\right)}$ and the claim follows from the second form of Ceva's Theorem for spherical triangles. □



Exercise 440. Let P be a point in a spherical triangle ABC and let P_A, P_B, P_C be its projections on the sides. Assuming P_A, P_B, P_C split the sides of ABC in portions having lengths $a_1, a_2, b_1, b_2, c_1, c_2$ (in this order), prove that:
$$\cos a_1 \cos b_1 \cos c_1 = \cos a_2 \cos b_2 \cos c_2.$$

Proof. By joining P with the vertices A, B, C the spherical triangle ABC is partitioned into six right triangles. Napier's pentagon provides the identity
$$\frac{\cos PA}{\cos PB} = \frac{\cos c_1}{\cos c_2}$$
immediately proving the claim. □

Exercise 441. A smooth surface $S \subset \mathbb{R}^3$ has the following property: for any plane π such that $\pi \cap S$ is non-empty, $\pi \cap S$ is either a point or a circle. Prove that S is a sphere.



Exercise 442. ABC is a spherical triangle on a unit sphere, such that $A = B = C = \frac{\pi}{2}$. P is a point inside such triangle and α, β, γ are the lengths of the geodesics PA, PB, PC. Prove that:

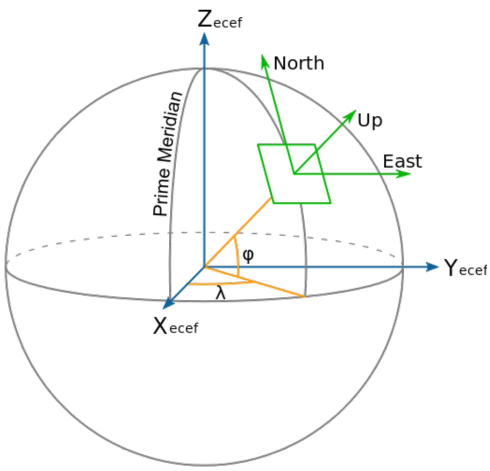
$$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1.$$

Proof. The spherical triangles PAB, PBC, PCA all have a side equal to one fourth of a maximal circle. By denoting as $A_1, A_2, B_1, B_2, C_1, C_2$ (in this order) the angles split by the geodesics PA, PB, PC, we have that A_1 and A_2 are complementary angles. By the dual version of Napier's identities:

$$\cos\beta = \sin\alpha\cos A_1, \qquad \cos\gamma = \sin\alpha\cos A_2$$

hence $\cos^2\beta + \cos^2\gamma = \sin^2\alpha$ and the claim trivially follows. □

It is pretty obvious that spherical trigonometry plays a crucial role in geodesy and astronomy. The usual system of coordinates used (for instance, by the GPS) for describing the location of a point on the surface of Earth (that we assume to be perfectly spherical) is based on two angles, a **latitude** φ, having values between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$, and a **longitude** λ, having values between $-\pi$ and π. The prime meridian through Greenwich gives the set of points whose longitude is zero and the equator gives the set of points whose latitude is zero.



The diagram represents the standard reference system, where the acronym *ECEF* stands for *Earth Centred Earth Fixed*. In order to be able to apply the results outlined in this section, it is essential to understand how to compute the distance d between two points on the surface of Earth whose coordinates are given by (φ_1, λ_1) and (φ_2, λ_2). By denoting as R_T the Earth radius, we have the following relation:

Theorem 443 (haversine formula)**.**

$$\operatorname{hav}\frac{d}{R_T} = \operatorname{hav}(\varphi_2 - \varphi_1) + \cos\varphi_1\cos\varphi_2\operatorname{hav}(\lambda_2 - \lambda_1)$$

where $\operatorname{hav}\theta = \frac{1-\cos\theta}{2} = \sin^2\frac{\theta}{2}$ is the *haversine* function.

Proof. Let us denote as A and B the points we are dealing with, and let us consider a point C on the sphere such that the spherical triangle ABC fulfills $\widehat{ACB} = \frac{\pi}{2}$. With such assumptions a only depends on a difference of latitudes and b only depends on a difference of longitudes, hence by suitably rearranging the terms of the identity $\cos c = \cos a\cos b + \sin a\sin b\cos C$ the claim easily follows. □

Exercise 444. ABC is a spherical triangle on a unit sphere and $a = b = \frac{3}{2}, c = 1$. Denoting as D the midpoint of BC, estimate the difference between the areas of the spherical triangles ABD and ACD.

Sketch of proof: we may compute the length of the median AD through Stewart's Theorem and use it for computing the staudtians of ABD and ACD. Due to Cagnoli's Theorem, both the spherical excess of ABD and the spherical excess of ACD are given by twice the arccosine of the sine of a known quantity. As an alternative, one may employ l'Huilier's Theorem to write those spherical excesses as four times an arctangent. □

We finish this section by mentioning an astonishing and very recent result, extending both Euler's Theorem on the collinearity of O, G, H and Feuerbach's Theorem on the property of the nine-point-circle to the spherical case.

We say that in a spherical triangle ABC a cevian AD is a **pseudoaltitude** if it fulfills

$$\widehat{BDA} = \widehat{DAB} + B - \frac{E}{2}, \qquad \widehat{CDA} = \widehat{CAD} + C - \frac{E}{2}$$

with E being the spherical excess of ABC. Using the spherical version of Ceva's Theorem it is not difficult to prove that the pseudoaltitudes concur, at a point (not surprisingly) known as **pseudoorthocenter**. We have already proved Steiner's Theorem about the concurrency of the pseudomedians at the **pseudocentroid**.

Theorem 445 (Akopyan)**.** In a spherical triangle the circumcenter, the pseudocentroid and the pseudoorthocenter lie on the same geodesic. Additionally, the feet of the pseudoaltitudes and the midpoints of the sides belong to the same circle, which is tangent to the inscribed circle and to every ex-inscribed circle.

Sketch of proof. It is not difficult to locate the positions, on the triangle sides, of the feet of the cevians through the circumcenter, the pseudocentroid and the pseudoorthocenter. Then it is possible to employ the projective map given by the central projection from the surface of the sphere to the plane π_{ABC} containing the vertices of ABC. The first part of the statement follows from Van Obel's and Ceva's Theorems. The existence of a spherical analogue of the nine points circle can be proved with a similar technique, i.e. by showing the concurrence of three perpendicular bisectors on π_{ABC}. The last part of the statement is the most difficult to prove, but that can be done by invoking the generalization of Ptolemy's theorem known as Casey's Theorem. □

The material composing this section, up to minor changes due to the author, mainly comes from three sources:

1. John Casey, **A Treatise in Spherical Trigonometry**, Hodges, Figgis & co, Dublin 1889;
2. Ren Guo, Estonia Black, Caleb Smith, **Strengthened Euler's inequality in spherical and hyperbolic geometries**, ArXiv, 17/04/2017;
3. Arseniy V.Akopian, **On some classical constructions extended to hyperbolic geometry**, ArXiv, 11/05/2011.

We invite the reader to have a look at them in order to acquire a broader view on the subject. Casey's treatise, for instance, is an impressive and very detailed work on spherical trigonometry, with the appearance of being a complete compendium. However the research of new results in spherical geometry is far from being over: Akopyan's Theorem, proved more than a century after Casey's treatise, is exemplary, and many interesting questions just arise from picking a non-trivial result in Euclidean Geometry and wondering if it admits a spherical analogue.

Students interested in having one or more references among the following ones can contact me at the address jacopo.daurizio@gmail.com.

References

[1] Aigner, Ziegler, *Proofs from The Book*

[2] Ahlfors, *Complex Analysis*

[3] N.Alon, *The Combinatorial Nullstellensatz*

[4] K.Ball, *An Elementary Introduction to Modern Convex Geometry*

[5] Boros, Moll, *Irresistible Integrals*

[6] Borwein, Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*

[7] N.G. de Bruijn, *Asymptotic Methods in Analysis*

[8] N.L. Carothers, *A Short Course on Approximation Theory*

[9] C. Casolo, *Corso di Teoria dei Grafi*

[10] K. Chandrasekharan, *Introduction to Analytic Number Theory*

[11] Flajolet, Sedgewick, *Analytic Combinatorics*

[12] Flajolet, Salvy, *Euler Sums and Contour Integral Representations*

[13] D.M. Kane, *An elementary derivation of the asymptotics of partition functions*

[14] Konrad, Knopp, *Theory and Applications of Infinite Series*

[15] Ireland, Rosen, *A Classical Introduction to Modern Number Theory*

[16] Mitrinovic, Pecaric, Fink, *Classical and New Inequalities in Analysis*

[17] P. Nahin, *Inside Interesting Integrals*

[18] Petrushev, Popov, *Rational Approximations of Real Functions*

[19] M. Steele, *The Cauchy-Schwarz Master Class*

[20] Elias M. Stein, *Fourier Analyis - an introduction*

[21] G. Szegö, *On a inequality of P.Turán concerning Legendre polynomials*

[22] A. Treibergs, *Mixed Area and the Isoperimetric Inequality*

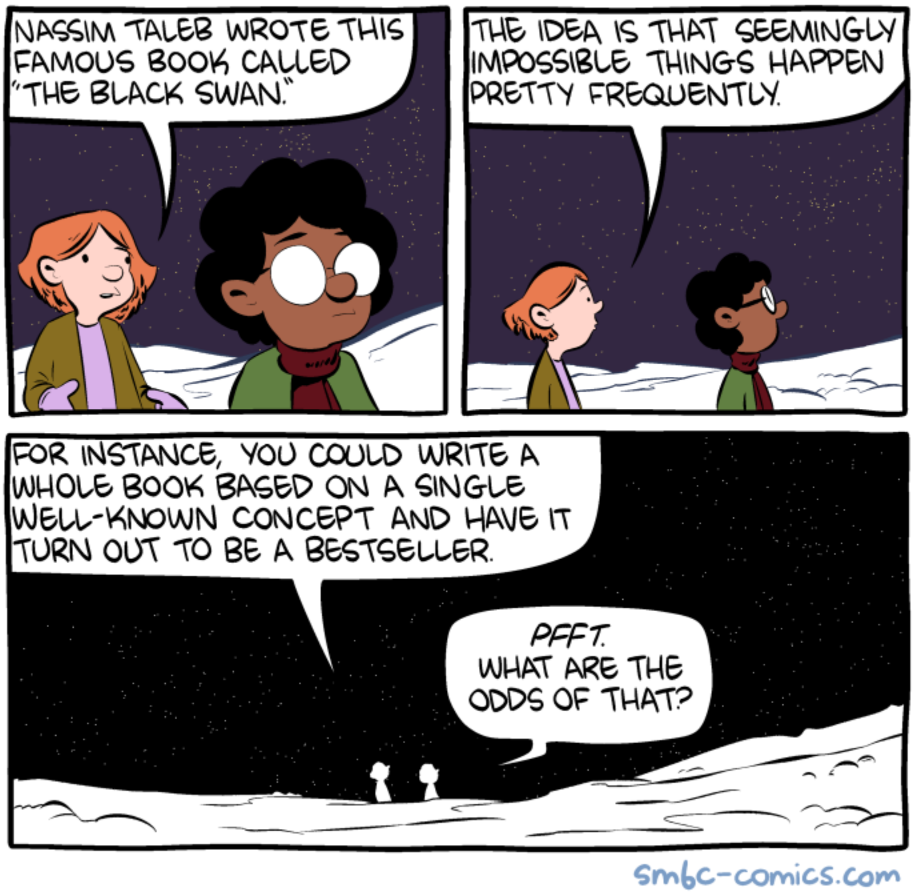
[23] J.H. van Lint, R.M. Wilson, *A course in Combinatorics*

[24] Don Zagier, *The Dilogarithm function*

[25] Whittaker and Watson, *A Course of Modern Analysis*

[26] Herbert S. Wilf, *Generatingfunctionology*

Postscript.

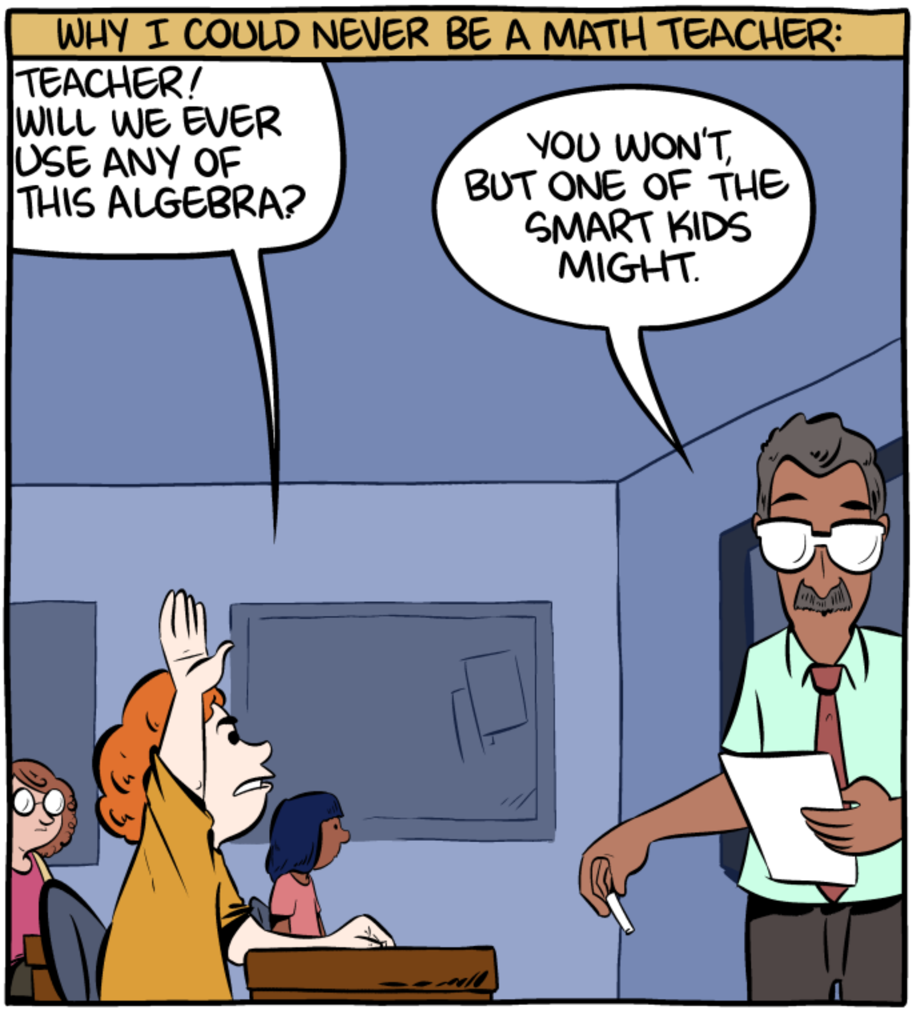


The author is well-aware of the following flaws in these course notes:

- A debatable ordering of sections, such that if often happens that in section X something belonging to section $X+\tau$ is used;
- A degree of malice in not presenting the minimal hypothesis allowing some manipulations, or a greater degree of malice in not presenting hypothesis at all, and just using expressions like "*as soon as everything makes sense*" or surrogates;
- To have declared this course as a problem solving *omnibus* in Calculus, Arithmetics, Algebra, Geometry and Combinatorics, while this course is *especially* about Calculus and Combinatorics;
- To have included many exercises without worked solutions in a section they *do not* rightfully belong, according to the usual strategies for solving them. I did this on purpose, so I guess I reach the maximum degree of malice here.

I have several faults and some excuses. For starters, the spirit of these pages. F.Iandoli left a comment about them: "*It is a very personal piece of work, starting from the title itself. To do* Superior Mathematics from an Elementary point of view *is a good summary of the author's way for approaching Mathematics - we have a difficult problem. Well, let's bring it to its knees, by throwing the right amount of stones.* Second point, the safeguard mechanism I activated in the introduction: this chaotic collection of deep results and dirty tricks is not meant to be an institutional course, but just as a compendium for young mathematicians, where to find interesting ideas in a concise form, interesting exercises for getting better at problem solving and a (not so) small *toolbox enlargement kit.* Third point, the author really loves *not to take* the usual way, and these notes are a bit of a challenge both for their writer (*How could I explain in the best possible way why these results are really important?*) and their readers (*Here it is, in a single line, a very difficult problem. If you reached this point, you can solve it, so solve it.*) Anyway these pages will be improved

through time, maybe through their readers' contributions, too. I am really grateful to Professor *Massimo Caboara* to let me build this strange experiment in Mathematics education. My deep gratitude also goes to *Greta Malaspina*[19] for her several helpful comments and her collaboration in translating these course notes from Italian to English. I am also greatly indebted to many users of Math.Stackexchange that contributed in some way to these notes: in no particular order, *achillehui*, *robjohn*, *Mike Spivey*, *Olivier Oloa*, *Mark Viola*, *Sangchul Lee*, *Vladimir Reshetnikov*, *Ron Gordon*, *Tolaso J. Kos* and *Zaid Alyafeai*.



Jacopo D'Aurizio (aka **Jack**)
jacopo.daurizio@gmail.com
http://www.matemate.it

[19] I will always regret she left me.