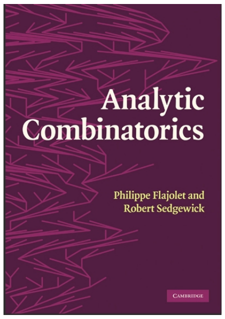
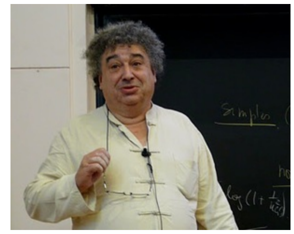


"Analytic combinatorics aims to enable precise quantitative predictions of the properties of large combinatorial structures. ..."



Amazon As pdf See also Philippe Flajolet's lectures and courses

"... The theory has emerged over recent decades as essential both for the analysis of algorithms and for the scientific models in many disciplines, including probability theory, statistical physics, computational biology and information theory. With a careful combination of symbolic enumeration methods and complex analysis, drawing heavily on generating functions, results of sweeping generality emerge that can be applied to fundamental structures such as permutations, sequences, strings, walks, trees, graphs and maps."
Foreword to "Analytic Combinatorics", Flajolet-Sedgewick 2009, Cambridge University Press.

Philippe Flajolet (1948-2011) laid the foundations of Analytic Combinatorics and extensively developed the methods and techniques used in this field.

Examples

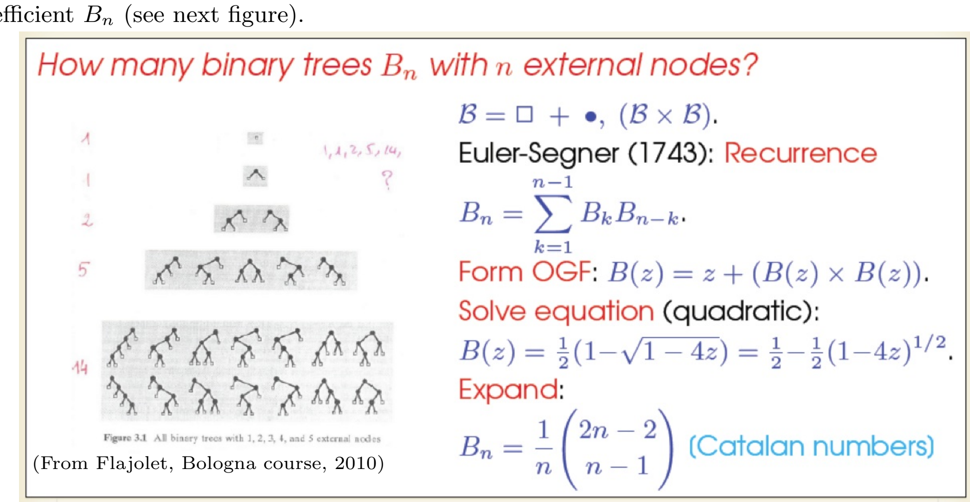
Binary trees. If you ask to a five or six years old child to draw binary trees with 1, 2, 3, 4, and 5 external nodes, and ask him about how many (different) ones there are, he will tell you the sequence (provided he or she does not get tired)

$$1, 1, 2, 5, 14 \ldots$$

Counting is also natural for mathematicians. Considering the sequence (B_n) enumerating binary trees and its OGF (ordinary generating function) $B(z)$, we have

$$(B_n) = (B_1, B_2, B_3, B_4, \ldots) = (1, 1, 2, 5, 14, \ldots) \quad \text{and} \quad B(z) = \sum_{n\geq 1} B_n z^n.$$

Now, if there are more than one external node in a binary tree, removing the root gives two subtrees that are equivalent (from a counting point of view) to any binary tree: there is a recursive decomposition that translates to a functional equation verified by the generating function $B(z)$, from which it is possible to extract the n-th Taylor coefficient B_n (see next figure).



(From Flajolet, Bologna course, 2010)

The example of binary trees is typical of the process of Analytic Combinatorics which works as follows.

1. Construct a symbolic equation on the combinatorial classes occurring in your problem (in the case of binary tree, these are the class $\mathcal{B}$ and the class $\square$ representing a leaf with OGF z).

2. Translate the symbolic equation into a functional equation on generating functions.

3. Extract the Taylor coefficient of interest; asymptotically, this is often done by complex analysis and Cauchy integrals or variants of these.

The counting is much more general than univariate counting as we see next.

Cycles in permutations. The *cycle construction* puts in equivalence classes sequences taken up to a circular shift; considering the permutations of the symmetric group $\mathfrak{S}_4$ of size 4!, we have

$$1234 \equiv 2341 \equiv 3412 \equiv 4123,\ 1243 \equiv 2341 \equiv \ldots,\ 1324 \equiv \ldots,\ 1342 \equiv \ldots,\ 1423 \equiv \ldots,\ 1432 \equiv \ldots.$$

If $C_n = n!/n$ is the number of classes of the symmetric group $\mathfrak{S}_n$ quotiented by the cycle construction, the corresponding *exponential* generating function verifies

$$C(z) = \sum_{n\geq 0} \frac{C_n z^n}{n!} = \sum_{n\geq 0} \frac{z^n}{n} = \log\left(\frac{1}{1-z}\right).$$

Considering any permutation, we can decompose it as a set of cycles, as seen in the following example

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 \\ 11 & 12 & 13 & 17 & 10 & 15 & 14 & 9 & 3 & 4 & 6 & 2 & 7 & 8 & 1 & 5 & 16 \end{pmatrix},$$

one of the cycle being $4 \to 17 \to 16 \to 5 \to 10 \to 4$.

If $\mathcal{C}$ is a generic cycle, and $\mathcal{P}$ a generic permutation, the decompositions is written symbolically as

$$\mathcal{P} = \{\epsilon\} + \mathcal{C} + (\mathcal{C} \star \mathcal{C}) + (\mathcal{C} \star \mathcal{C} \star \mathcal{C}) + \ldots \qquad \text{(Permutation = Set of Cycles)}.$$

As *Set* $\rightsquigarrow$ exp and *Cycle* $\rightsquigarrow$ log, using again exponential generating functions that count labelled objects, and moreover a variable u that counts the number of cycles, we have (being very sketchy)

$$P(z,u) = \sum_{\substack{n\geq 0\\ u\leq n}} \begin{bmatrix} n \\ k \end{bmatrix} u^k z^n = 1 + uC(z) + \frac{1}{2!}u^2C^2(z) + \frac{1}{3!}u^3C^3(z) + \cdots = \exp\left(u \log\left(\frac{1}{1-z}\right)\right) = (1-z)^{-u},$$

where $\begin{bmatrix} n \\ k \end{bmatrix}$ is the Stirling cycle number that counts the number of permutations of size n with k cycles.

We obtain by the binary theorem

$$[z^n](1-z)^{-u} = \sum_{k\leq n} \begin{bmatrix} n \\ k \end{bmatrix} u^k = u(u+1)(u+2)\ldots(u+n-1),$$

and, by logarithmic differentiation, the expected number of cycles $\mu_n = \sum_k \frac{k}{n!} \begin{bmatrix} n \\ k \end{bmatrix}$ in a random permutation of size n is the n-th harmonic number,

$$\mu_n = H_n \equiv 1 + \frac{1}{2} + \cdots + \frac{1}{n} \qquad (\rightsquigarrow \mu_{100} \equiv H_{100} = 5.18738).$$

Second moment follows easily, and an asymptotic method known as *quasi-powers theorem* leads to a limiting Gaussian law. (There are equivalent probabilistic approaches.)

What can you learn from Analytic Combinatorics?

The projected courses will aim providing a thorough introduction to Flajolet-Sedgewick book "Analytic Combinatorics"; an additional course will be related to the Boltzmann random generation of objects. If you are a *mathematician* or a *physicist*, you cannot avoid being touched by the beauty of symbolic structures and by relatively simple mathematical concepts that lead to deep results with "real life" applications. If you are a *computer scientist* you will learn evaluating combinatorial structures that have algorithmic counterparts; *i.e* the (generalized) birthday paradox provides an analysis of collisions in data hashing.