5.7 Lengyel’s Constant

5.7.1 Stirling Partition Numbers

Let \( S \) be a set with \( n \) elements. The set of all subsets of \( S \) has \( 2^n \) elements. By a partition of \( S \) we mean a disjoint set of nonempty subsets (called blocks) whose union is \( S \). The set of partitions of \( S \) that possess exactly \( k \) blocks has \( S_{n,k} \) elements, where \( S_{n,k} \) is a
Stirling number of the second kind. The set of all partitions of \( S \) has \( B_n \) elements, where \( B_n \) is a Bell number:

\[
B_n = \sum_{k=0}^{n} S_{n,k} = \frac{1}{e} \sum_{j=0}^{\infty} \frac{j^n}{j!} = \frac{d^n}{dx^n} \exp(e^x - 1) \bigg|_{x=0}.
\]

For example, \( S_{1,1} = 1 \), \( S_{4,2} = 7 \), \( S_{4,3} = 6 \), \( S_{4,4} = 1 \), and \( B_4 = 15 \). More generally, \( S_{n,1} = 1 \), \( S_{n,2} = 2^{n-1} - 1 \), and \( S_{n,3} = \frac{1}{2}(3^{n-1} + 1) - 2^{n-1} \). The following recurrences are helpful [1–4]:

\[
S_{n,0} = \begin{cases} 
1 & \text{if } n = 0, \\
0 & \text{if } n \geq 1,
\end{cases} \quad S_{n,k} = kS_{n-1,k} + S_{n-1,k-1} \quad \text{if } n \geq k \geq 1,
\]

\[
B_0 = 1, \quad B_n = \sum_{k=0}^{n} \binom{n-1}{k} B_k.
\]

and corresponding asymptotics are discussed in [5–9].

5.7.2 Chains in the Subset Lattice of \( S \)

If \( U \) and \( V \) are subsets of \( S \), write \( U \subset V \) if \( U \) is a proper subset of \( V \). This endows the set of all subsets of \( S \) with a partial ordering; in fact, it is a lattice with maximum element \( S \) and minimum element \( \emptyset \). The number of chains \( \emptyset = U_0 \subset U_1 \subset \cdots \subset U_{k-1} \subset U_k = S \) of length \( k \) is \( k!S_{n,k} \). Hence the number of all chains from \( \emptyset \) to \( S \) is [1, 6, 10]

\[
\sum_{k=0}^{n} k!S_{n,k} = \sum_{j=0}^{\infty} \frac{j^n}{2j+1} = \frac{1}{2} \Li_{-n} \left( \frac{1}{2} \right) = \frac{d^n}{dx^n} \frac{1}{2 - e^x} \bigg|_{x=0} \sim \frac{n!}{2} \left( \frac{1}{\ln(2)} \right)^{n+1},
\]

where \( \Li_n(x) \) is the polylogarithm function. Wilf [10] marveled at how accurate this asymptotic approximation is.

If we further insist that the chains are maximal, equivalently, that additional proper insertions are impossible, then the number of such chains is \( n! \). A general technique due to Doubilet, Rota & Stanley [11], involving what are called incidence algebras, can be used to obtain the two aforesaid results, as well as to enumerate chains within more complicated posets [12].

As an aside, we give a deeper application of incidence algebras: to enumerating chains of linear subspaces within finite vector spaces [6]. Define the \( q \)-binomial coefficient and \( q \)-factorial by

\[
\binom{n}{k}_q = \frac{\prod_{j=1}^{n}(q^j - 1)}{\prod_{j=1}^{k}(q^j - 1) \cdot \prod_{j=1}^{n-k}(q^j - 1)},
\]

\[
[n!]_q = (1 + q)(1 + q + q^2) \cdots (1 + q + \cdots + q^{n-1}).
\]
where \( q > 1 \). Note the special case in the limit as \( q \to 1^+ \). Consider the \( n \)-dimensional vector space \( \mathbb{F}_q^n \) over the finite field \( \mathbb{F}_q \), where \( q \) is a prime power [12–16]. The number of \( k \)-dimensional linear subspaces of \( \mathbb{F}_q^n \) is \( \binom{n}{k}_q \), and the total number of linear subspaces of \( \mathbb{F}_q^n \) is asymptotically \( c_e q^{n^2/4} \) if \( n \) is even and \( c_o q^{n^2/4} \) if \( n \) is odd, where [17, 18]

\[
c_e = \frac{\sum_{k=0}^{\infty} q^{-k^2}}{\prod_{j=1}^{\infty} (1 - q^{-j})}, \quad c_o = \frac{\sum_{k=0}^{\infty} q^{-(k+\frac{1}{2})^2}}{\prod_{j=1}^{\infty} (1 - q^{-j})}.
\]

We give a recurrence for the number \( \chi_n \) of chains of proper subspaces (again, ordered by inclusion):

\[
\chi_1 = 1, \quad \chi_n = 1 + \sum_{k=1}^{n-1} \binom{n}{k}_q \chi_k \quad \text{for } n \geq 2.
\]

For the asymptotics, it follows that [6, 17]

\[
\chi_n \sim \frac{1}{\zeta_q'(r)r^n} \left( \frac{1}{r} \right)^n \prod_{j=1}^{n} (q^j - 1) = \frac{A}{r^n} (q - 1)(q^2 - 1)(q^3 - 1) \cdots (q^n - 1),
\]

where \( \zeta_q(x) \) is the zeta function for the poset of subspaces:

\[
\zeta_q(x) = \sum_{k=1}^{\infty} \frac{x^k}{(q - 1)(q^2 - 1)(q^3 - 1) \cdots (q^k - 1)}
\]

and \( r > 0 \) is the unique solution of the equation \( \zeta_q(r) = 1 \). In particular, when \( q = 2 \), we have \( c_e = 7.3719688014 \ldots \), \( c_o = 7.371949407 \ldots \), and

\[
\chi_n \sim \frac{A}{r^n} \cdot Q \cdot 2^{\frac{n(n+1)}{2}},
\]

where \( r = 0.7759021363 \ldots \), \( A = 0.8008134543 \ldots \), and

\[
Q = \prod_{k=1}^{\infty} \left( 1 - \frac{1}{2^k} \right) = 0.2887880950 \ldots
\]

is one of the digital search tree constants [5.14]. If we further insist that the chains are maximal, then the number of such chains is \( [n!]_q \).

### 5.7.3 Chains in the Partition Lattice of \( S \)

We have discussed chains in the poset of subsets of the set \( S \). There is, however, another poset associated naturally with \( S \) that is less familiar and more difficult to study: the poset of partitions of \( S \). Here is the partial ordering: Assuming \( P \) and \( Q \) are two partitions of \( S \), then \( P < Q \) if \( P \neq Q \) and if \( p \in P \) implies that \( p \) is a subset of \( q \) for some \( q \in Q \). In other words, \( P \) is a refinement of \( Q \) in the sense that each of its blocks fits within a block of \( Q \). For arbitrary \( n \), the poset is, in fact, a lattice with minimum element \( m = \{1\}, \{2\}, \ldots, \{n\} \) and maximum element \( M = \{1, 2, \ldots, n\} \).
What is the number of chains \( m < P_1 < M \) in the partition lattice of the set \( \{1, 2, 3\} \)? In the case \( n = 3 \), there is only one chain for \( k = 1 \), specifically, \( m < M \). For \( k = 2 \), there are three such chains as pictured in Figure 5.10.

Let \( Z_n \) denote the number of all chains from \( m \) to \( M \) of any length; clearly \( Z_1 = Z_2 = 1 \) and, by the foregoing, \( Z_3 = 4 \). We have the recurrence

\[
Z_n = \sum_{k=1}^{n-1} S_{n,k} Z_k
\]

and exponential generating function

\[
Z(x) = \sum_{n=1}^{\infty} \frac{Z_n}{n!} x^n, \quad 2Z(x) = x + Z(e^x - 1),
\]

but techniques of Doubilet, Rota & Stanley and Bender do not apply here to give asymptotic estimates of \( Z_n \). The partition lattice is the first natural lattice without the structure of a binomial lattice, which implies that well-known generating function techniques are no longer helpful.

Lengyel [19] formulated a different approach to prove that the quotient

\[
r_n = \frac{Z_n}{(n!)^2(2 \ln(2))^{-n} n^{-1 - \ln(2)/3}}
\]

must be bounded between two positive constants as \( n \to \infty \). He presented numerical evidence suggesting that \( r_n \) tends to a unique value. Babai & Lengyel [20] then proved a fairly general convergence criterion that enabled them to conclude that \( \lambda = \lim_{n \to \infty} r_n \) exists and \( \lambda = 1.09 \ldots \). The analysis in [19] involves intricate estimates of the Stirling numbers; in [20], the focus is on nearly convex linear recurrences with finite retardation and active predecessors.

In an ambitious undertaking, Flajolet & Salvy [21] computed \( \Lambda = 1.0986858055 \ldots \). Their approach is based on (complex fractional) analytic iterates of \( \exp(x) - 1 \) and much more, but unfortunately their paper is presently incomplete. See [5.8] for related discussion of the Takuchi-Prellberg constant.

By way of contrast, the number of maximal chains is given exactly by \( n!(n - 1)!/2^{n-1} \) and Lengyel [19] observed that \( Z_n \) exceeds this by an exponentially large factor.
5.7.4 Random Chains

Van Cutsem & Ycart [22] examined random chains in both the subset and partition lattices. It is remarkable that a common framework exists for studying these and that, in a certain sense, the limiting distributions of both types of chains are identical. We mention only one consequence: If \( \kappa_n = k/n \) is the normalized length of the random chain, then

\[
\lim_{n \to \infty} E(\kappa_n) = \frac{1}{2 \ln(2)} = 0.7213475204 \ldots
\]

and a corresponding Central Limit Theorem also holds.

5.8 Takeuchi–Prellberg Constant

In 1978, Takeuchi defined a triply recursive function [1, 2]

\[ t(x, y, z) = \begin{cases} 
  t(t(x - 1, y, z), t(y - 1, z, x), t(z - 1, x, y)) & \text{if } x \leq y, \\
  t(t(x - 1, y, z), t(y - 1, z, x), t(z - 1, x, y)) & \text{otherwise}
\end{cases} \]

that is useful for benchmark testing of programming languages. The value of \(t(x, y, z)\) is of no practical significance; in fact, McCarthy [1, 2] observed that the function can be described more simply as

\[ t(x, y, z) = \begin{cases} 
  y & \text{if } x \leq y, \\
  z & \text{if } y \leq z, \\
  x & \text{otherwise},
\end{cases} \]

The interesting quantity is not \(t(x, y, z)\), but rather \(T(x, y, z)\), defined to be the number of times the \(otherwise\) clause is invoked in the recursion. We assume that the program is memoryless in the sense that previously computed results are not available at any time in the future. Knuth [1, 3] studied the Takeuchi numbers \(T_n = T(n, 0, n+1)\):

\[ T_0 = 0, \quad T_1 = 1, \quad T_2 = 4, \quad T_3 = 14, \quad T_4 = 53, \quad T_5 = 223, \ldots \]

and deduced that

\[ e^{n \ln(n) - n \ln(\ln(n)) - n} < T_n < e^{n \ln(n) - n + \ln(n)} \]

for all sufficiently large \(n\). He asked for more precise asymptotic information about the growth of \(T_n\).

Starting with Knuth’s recursive formula for the Takeuchi numbers

\[ T_{n+1} = \sum_{k=0}^{n} \left[ \binom{n+k}{n} - \binom{n+k}{n+1} \right] T_{n-k} + \sum_{k=1}^{n-1} \binom{2k}{k} \frac{1}{k+1} \]

and the somewhat related Bell numbers [5, 7]

\[ B_{n+1} = \sum_{k=0}^{n} \binom{n}{k} B_{n-k}, \quad B_0 = 1, \quad B_1 = 1, \quad B_2 = 2, \quad B_3 = 5, \quad B_4 = 15, \quad B_5 = 52, \ldots \]

Prellberg [4] observed that the following limit exists:

\[ c = \lim_{n \to \infty} \frac{T_n}{B_n \exp \left( \frac{1}{2} W_n^2 \right)} = 2.2394331040 \ldots, \]

where \(W_n \exp(W_n) = n\) are special values of the Lambert \(W\) function [6, 11].

Since both the Bell numbers and the \(W\) function are well understood, this provides an answer to Knuth’s question. The underlying theory is still under development, but
Prellberg's numerical evidence is persuasive. Recent theoretical work [5] relates the constant \( c \) to an associated functional equation,

\[
T(z) = \sum_{n=0}^{\infty} T_n z^n, \quad T(z) = \frac{T(z - z^2)}{z} - \frac{1}{(1-z)(1-z+z^2)},
\]

in a manner parallel to how Lengyel's constant [5.7] is obtained.


### 5.9 Pólya's Random Walk Constants

Let \( L \) denote the \( d \)-dimensional cubic lattice whose vertices are precisely all integer points in \( d \)-dimensional space. A **walk** \( \omega \) on \( L \), beginning at the origin, is an infinite sequence of vertices \( \omega_0, \omega_1, \omega_2, \omega_3, \ldots \) with \( \omega_0 = 0 \) and \( |\omega_{j+1} - \omega_j| = 1 \) for all \( j \). Assume that the walk is random and symmetric in the sense that, at each time step, all \( 2d \) directions of possible travel have equal probability. What is the likelihood that \( \omega_n = 0 \) for some \( n > 0 \)? That is, what is the **return probability** \( p_d \)?

Pólya [1–4] proved the remarkable fact that \( p_1 = p_2 = 1 \) but \( p_d < 1 \) for \( d > 2 \). McCrea & Whipple [5], Watson [6], Domb [7] and Glasser & Zucker [8] each contributed facets of the following evaluations of \( p_3 = 1 - 1/m_3 = 0.3405373295 \ldots \), where the expected number \( m_3 \) of returns to the origin, plus one, is

\[
m_3 = \frac{3}{(2\pi)^3} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \frac{1}{3 - \cos(\theta) - \cos(\phi) - \cos(\psi)} \, d\theta \, d\phi \, d\psi
\]

\[
= \frac{12}{\pi^2} \left( 18 + 12\sqrt{2} - 10\sqrt{3} - 7\sqrt{6} \right) K \left[ (2 - \sqrt{3})(\sqrt{3} - \sqrt{2}) \right]^2
\]

\[
= 3 \left( 18 + 12\sqrt{2} - 10\sqrt{3} - 7\sqrt{6} \right) \left[ 1 + 2 \sum_{k=1}^{\infty} \exp(-\sqrt{6}\pi k^2) \right]^4
\]

\[
= \frac{\sqrt{6}}{32\pi^3} \Gamma \left( \frac{1}{24} \right) \Gamma \left( \frac{5}{24} \right) \Gamma \left( \frac{7}{24} \right) \Gamma \left( \frac{11}{24} \right) = 1.5163860591 \ldots
\]

Hence the **escape probability** for a random walk on the three-dimensional cubic lattice is \( 1 - p_3 = 0.6594626704 \ldots \). In these expressions, \( K \) denotes the complete elliptic integral of the first kind [1.4.6] and \( \Gamma \) denotes the gamma function [1.5.4]. Return and escape probabilities can also be computed for the body-centered or face-centered cubic