Computation of Stirling numbers and generalizations

S. Ilie
Dept. Mathematics,
Ryerson University
silvana@ryerson.ca

D. J. Jeffrey
Dept. Applied Mathematics
Western University Canada
djeffrey@uwo.ca

R. M. Corless
Dept. Applied Mathematics
Western University Canada
rcorless@uwo.ca

X. Zhang
Dept. Applied Mathematics
Western University Canada
xzhan695@uwo.ca

***Abstract*—We consider the computation of Stirling numbers and generalizations for positive and negative arguments. We describe computational schemes for Stirling Partition and Stirling Cycle numbers, and for their generalizations to associated Stirling numbers. The schemes use recurrence relations and are more efficient than the current method used in MAPLE for cycle numbers, which is based on an algebraic expansion. We also point out that the proposed generalization of Stirling numbers due to Flajolet and Prodinger changes the evaluation of Stirling partition numbers for negative arguments. They are no longer zero, but become rational numbers.**

I. Introduction

Among the remarkable sequences of numbers with important combinatorial significance one can count the sequences of Stirling numbers. These numbers were first introduced by James Stirling in [13] to express the connection between the ordinary powers and the factorial powers. In older literature, they are called Stirling numbers of the first kind and second kind. A modern notation for these numbers follows Knuth's suggestions [10]. He proposed the notations

$$\left\{ {n \atop m} \right\}_r \quad \text{and} \quad \left[{n \atop m} \right]_r$$

for respectively Stirling Partition numbers (Stirling numbers of the second kind) and Stirling Cycle numbers (Stirling numbers of the first kind). These notations in the special case $r = 1$ were first suggested by Karamata in [8].

The definitions used in this paper give all numbers as nonnegative, again following Knuth, in contrast to earlier definitions [1], [2]. The names reflect the combinatorial significance of the numbers, and the notations are inspired by the similar notation for the binomial coefficients.

We describe a new implementation of Stirling cycle numbers in MAPLE, which is faster than the existing implementation in the `combinat` package. The existing implementation does not use recurrence relations, but expands a polynomial. The current MAPLE functions have names `stirling2`, for Stirling partition numbers and `stirling1`, for signed Stirling cycle numbers.

Following a challenge by Knuth, several authors suggested generalizations of Stirling numbers to non-integral arguments [12], [5]. The widely accepted generalization [5] allows Stirling numbers to be extended to complex arguments. The proposal leaves the value of $\left\{ {0 \atop 0} \right\}$ undecided, and we discuss possible values here.

II. Definitions and Properties for $r = 1$

We begin with the case $r = 1$, which has been the traditional meaning given to Stirling numbers.

Definition II.1. The Stirling partition number $\left\{ {n \atop k} \right\}$ is the number of ways to partition a set of n objects into k nonempty subsets.

Definition II.2. The Stirling cycle number $\left[{n \atop k} \right]$ is the number of permutations of n objects having k cycles.

Stirling numbers satisfy the recurrence relations

$$\left\{ {n+1 \atop k} \right\} = k \left\{ {n \atop k} \right\} + \left\{ {n \atop k-1} \right\} ; \tag{1}$$

$$\left[{n+1 \atop k} \right] = n \left[{n \atop k} \right] + \left[{n \atop k-1} \right] ; \tag{2}$$

similar to the one satisfied by the binomial coefficients

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1} .$$

The identities (1) and (2) hold for all integers n and k (positive or negative). The boundary conditions

$$\left\{ {k \atop k} \right\} = \left[{k \atop k} \right] = 1 \quad \text{for} \quad k > 0 ;$$

$$\left\{ {n \atop 0} \right\} = \left[{n \atop 0} \right] = \delta_{0n} \quad \text{for} \quad n > 0 ;$$

where δ_{0n} is the Kronecker delta, lead to unique solutions for all $k; n$ integers. The Stirling cycle numbers and the Stirling partition numbers are connected by the remarkable law of duality

$$\left\{ {n \atop k} \right\} = \left[{-k \atop -n} \right] ;$$

which is valid for all $k; n$ integers. Other special values are

$$\left\{ {n \atop 1} \right\} = 1 ; \quad \text{and} \quad \left[{n \atop 1} \right] = (n-1)! . \tag{3}$$

For $n; k > 1$, it is possible to write Stirling partition numbers as a sum over binomial coefficients:

$$\left\{ {n \atop k} \right\} = \frac{1}{k!} \sum_{j=1}^{k} \binom{k}{j} (-1)^{k-j} j^n . \tag{4}$$

This can be extended to cycle numbers by using the identity

$$(-1)^{n+k}\left[\begin{matrix}n\\k\end{matrix}\right] = \sum_{h=0}^{n-k}(-1)^h\binom{n-1+h}{n-k+h}\binom{2n-k}{n+h}\left\{\begin{matrix}n-k+h\\h\end{matrix}\right\}. \quad (5)$$

The original definitions used by Stirling are important. Using the 'falling' and 'rising' notation of Knuth, we can write:

$$\begin{aligned} z^{\underline{n}} &:= z(z-1)(z-2)\ldots(z-n+1) \\ &= \sum_k \left[\begin{matrix}n\\k\end{matrix}\right](-1)^{n-k}z^k. \quad (6) \\ z^{\overline{n}} &:= z(z+1)(z+2)\ldots(z+n-1) \\ &= \sum_k \left[\begin{matrix}n\\k\end{matrix}\right] z^k. \quad (7) \end{aligned}$$

III. Associated Stirling numbers

For extended discussions of associated Stirling numbers, we refer to [2], [7].

Definition III.1. The number $\left\{\begin{matrix}n\\m\end{matrix}\right\}_{\geq r}$ gives the number of partitions of a set of size n into m subsets, each subset having a cardinality $\geq r$.

Definition III.2. The number $\left[\begin{matrix}n\\m\end{matrix}\right]_{\geq r}$ gives the number of permutations of n objects into m cycles, each cycle having a cardinality $\geq r$.

```
stirling1:= (n,k)->
    coeff(mul(z-m,m=0..n-1),z,k);
```

If fact the implementation in Maple works a little differently from this. For any given n, the function computes all $n \quad 1$ coefficients of the polynomial, or equivalently all ${n \atop k}$ for fixed n and $0 \quad k \quad n$, and stores them in cache memory. Thus after one Stirling number is requested, subsequent requests for another number with the same n, but differing k can be returned immediately.

We compare computation of Stirling cycle numbers computed by four methods.

1) equation (5).
2) Maple using (6).
3) equation (15).
4) equation (2).

A. Stirling cycle numbers by recurrence for $0 \quad k \quad n$

To calculate ${n \atop k}$, the program computes only those elements which are needed for its evaluation, according to the recurrence relation. Thus, in coordinates $(N; K)$, these are the pairs inside the parallelogram delimited by the lines $N = K$, $N = K + k$, $K = 0$, and $K = k$. Notice that, when k is close to 0 or n, the number of pairs computed becomes of order n, because the area of the parallelogram is of order n. (By 'of order n' we mean that the number of pairs will be approximately linear in n, with only weak dependence on k.)

When k approaches n=2, more pairs are computed, and so the algorithm is slower. The worst case is when k is closest to n=2. Then our algorithm computes a number of pairs of order n^2, because the area of the parallelogram is of order n^2. (The number of pairs will be approximately quadratic in n with a weak dependence on ($k \quad n$=2).) Our algorithm is always better than the algorithm existing in MAPLE, even in the case $k = bn$=2c. MAPLE's algorithm is always of order n^2 because the number of arithmetic operations that are required to compute the coefficients of the polynomial (6)) (of order n^2 arithmetic operations are done in order to compute, so when k is farther from n=2, the difference between our algorithm and MAPLE's increases. The computation works bottom-up, starting with the known boundary values for ${n \atop n}$ and ${n \atop 1}$ and considers two cases depending on whether $n > 2k$ or not. In either case, three subcases are considered. For example, when $n > 2k$, the algorithm computes as follows:

- the pairs $[i; j]$, $1 \quad i \quad k$, $1 \quad j \quad i$,
- the pairs $[i; j]$, $k + 1 < i < n \quad k + 1$, $1 \quad j \quad k$, and
- the pairs $[i; j]$, $n \quad k + 2 \quad i \quad n$, $i + k \quad n \quad j \quad k$.

On the other hand, if $n \quad 2k$, we compute as follows:

- the pairs $[i; j]$, $1 \quad i \quad n \quad k$, $1 \quad j \quad i$,
- the pairs $[i; j]$, $n \quad k + 1 < i < k$, $i \quad n + k \quad j \quad i + n \quad k \quad 1$, and
- the pairs $[i; j]$, $k + 1 \quad i \quad n$, $i + k \quad n \quad j \quad k$.

The two cases are illustrated in figures 1 and 2. For the transition case $n = 2k$, a third case could be added, but the efficiency gains would be negligible. Our algorithm uses only a vector of length k to store all pairs computed. We can do that because ${n \atop k}$ depends only on the pairs in the level below, that is, with the first component equal to $n \quad 1$.

.)

TABLE I

TIMINGS IN SECONDS OF COMPUTATIONS OF S

If these are to hold true when n ceases to be integral, then substituting $n = 0$ in these equations gives a contradiction. One of them must take precedence for the determination of $\genfrac{[}{]}{0pt}{}{0}{0}$.

The integral definition (17) shows the same behaviour. It has been shown that [14]

$$\lim_{x \to 0} \frac{\genfrac{[}{]}{0pt}{}{x=n}{}}{x} = n .$$

Thus the origin is a singular point and the value there is a matter of convention.

B. Consequences of different assumptions

We consider here the consequences of retaining the recurrence relation (18), and allowing different conventions for $\genfrac{[}{]}{0pt}{}{0}{0}$. Under the definition $\genfrac{[}{]}{0pt}{}{0}{0} = 1$, we have seen $\genfrac{[}{]}{0pt}{}{0}{1} = 0$. This in turn implies $\genfrac{[}{]}{0pt}{}{0}{k} = 0$ for all $k > 0$, and by further extension $\genfrac{[}{]}{0pt}{}{n}{k} = 0$ for all $k > 0, n < 0$.

In contrast, under the definition $\genfrac{[}{]}{0pt}{}{0}{0} = 0$, we have from the recurrence relation

$$\genfrac{[}{]}{0pt}{}{0}{1} = \genfrac{[}{]}{0pt}{}{1}{1} - \genfrac{[}{]}{0pt}{}{0}{0} = 1 .$$

From this we have $2\genfrac{[}{]}{0pt}{}{0}{2} = \genfrac{[}{]}{0pt}{}{1}{2} - \genfrac{[}{]}{0pt}{}{0}{1} = -1$. Thus by induction, it is easy to establish

$$\genfrac{[}{]}{0pt}{}{0}{k} = \frac{(-1)^{k+1}}{k!} ; \qquad \text{for} \quad k > 0 .$$

To proceed further, we notice that the recurrence relation for partition numbers can be written

$$k \genfrac{[}{]}{0pt}{}{n}{k} + \genfrac{[}{]}{0pt}{}{n}{k-1} = \genfrac{[}{]}{0pt}{}{n+1}{k} .$$

If n is regarded as fixed and $\genfrac{[}{]}{0pt}{}{n+1}{k}$ is regarded as known, then the recurrence relation can be solved. We therefore state the lemma

Lemma. The recurrence relation $mS_m + S_{m-1} = g_m$ has the solution

$$S_m = \frac{(-1)^m}{m!} \left[S_0 + \sum_{j=1}^{m} (-1)^j (j-1)! \, g_j \right] .$$

Proof. Direct substitution in the recurrence relation. □

We now apply this to

$$k \genfrac{[}{]}{0pt}{}{1}{k} + \genfrac{[}{]}{0pt}{}{1}{k-1} = \genfrac{[}{]}{0pt}{}{0}{k} = \frac{(-1)^{k+1}}{k!}$$

giving

$$\begin{aligned} \genfrac{[}{]}{0pt}{}{1}{k} &= \frac{(-1)^k}{k!} \genfrac{[}{]}{0pt}{}{1}{0} + \frac{(-1)^{k+1}}{k!} \sum_{j=1}^{k} \frac{1}{j} \\ &= \frac{(-1)^k}{k!} \genfrac{[}{]}{0pt}{}{1}{0} + \frac{(-1)^{k+1}}{k!} H_k ; \end{aligned}$$

where H_m is a harmonic number. It seems natural to continue with $\genfrac{[}{]}{0pt}{}{1}{0} = 0$, giving

$$\genfrac{[}{]}{0pt}{}{1}{k} = \frac{(-1)^{k+1}}{k!} H_k .$$

VIII. Conclusions

We have considered two computational problems in this paper. The first problem is a more efficient algorithm for evaluating the known Stirling Cycle numbers with positive arguments. For negative arguments there are no commonly accepted definitions for either cycl-7.307 Td [(x=n)]TJ 5.4e accepte0e2yeitheta132y