Europ. J. Combinatorics (1984) **5**, 313–321

On a Recurrence involving Stirling Numbers

TAMÁS LENGYEL

We count the number $Z(n)$ of (not necessarily maximal) chains from 0 to 1 in the partition lattice of an n-set. This function satisfies the recurrence $Z(n)=\sum_{k=1}^{n-1} S(n, k)Z(k)$ where $S(n, k)$ denotes the Stirling numbers of the second kind. We find the asymptotic order of magnitude of $Z(n)$.

1. INTRODUCTION

The partitions of an n-set form the lattice $\mathrm{Eq}(n)$ with minimal element $\{\{1\}, \{2\}, \ldots, \{n\}\}$ and maximal element $\{\{1, 2, \ldots, n\}\}$.

Let $Z(n)$ denote the number of those chains of $\mathrm{Eq}(n)$ which contain these two extreme elements. It is easy to see, that $Z(n)$ satisfies the recurrence

$$Z(n)=\sum_{k=1}^{n-1} S(n, k)Z(k)$$

where $S(n, k)$ denotes the Stirling numbers of the second kind.

Our aim is to find the asymptotic order of magnitude of $Z(n)$. This computation seems worth making since the partition lattice is the first natural lattice without the structure of binomial poset, and hence does not succumb to the techniques of Doubilet, Rota and Stanley [5] and Bender [3]. In this paper, we prove the following result.

THEOREM 1.1. *There exist positive constants C_1 and C_2 such that*

$$C_1 \leqslant Z(n)/f(n) \leqslant C_2$$

where $f(n)=(n!)^2(2\log 2)^{-n} n^{-1-(\log 2)/3}$, (log stands for the natural logarithm).

In a subsequent paper we shall give a fairly general convergence criterion [2], an application of which will prove

THEOREM 1.2 *The following limit exists*

$$\lim_{n\to\infty} Z(n)/f(n)=C$$

where C is a positive constant [2].

REMARK 1. Although we do not obtain any explicit estimate of the constant C, numerical evidence appears to suggest that $C\approx 1\cdot 1$. Below we tabulate the values of $Z(n)$ and of $Z(n)/f(n)$ for $n\leqslant 16$ (with 4-digit precision):

n	1	2	3	4	5	6	7	8
$Z(n)$	1	1	4	32	436	9012	$2\cdot 628\times 10^5$	$1\cdot 027\times 10^7$
$Z(n)/f(n)$	1·386	1·128	1·145	1·131	1·124	1·120	1·117	1·115

n	9	10	11	12	13	14	15	16
$Z(n)$	$5\cdot 183\times 10^8$	$3\cdot 280\times 10^{10}$	$2\cdot 543\times 10^{12}$	$2\cdot 371\times 10^{14}$	$2\cdot 617\times 10^{16}$	$3\cdot 376\times 10^{18}$	$5\cdot 030\times 10^{20}$	$8\cdot 575\times 10^{22}$
$Z(n)/f(n)$	1·113	1·111	1·110	1·110	1·109	1·108	1·107	1·107

0195-6698/84/040313+09 \$02.00/0

A list of the tree representations of partition-chains for small vaules of n is given in Figure 1. (The numbers in parentheses in the figure indicate multiplicities arising from different numberings of the end vertices.)

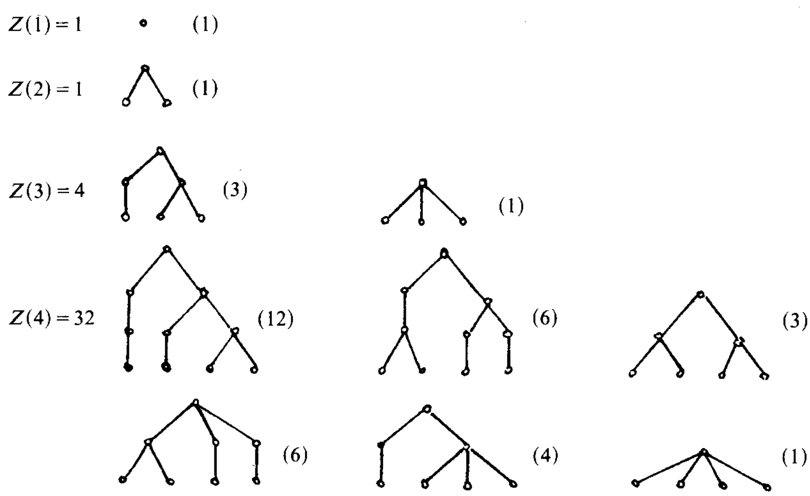

FIGURE 1.

REMARK 2. We remark that the number of maximal chains in $\mathrm{Eq}(n)$ is easily seen to be $n!(n-1)!/2^{n-1}$. The number $Z(n)$ exceeds this by an exponentially large factor.

REMARK 3. Some similarly looking tree-enumeration problems were considered by Schröder [9] in 1870, see Comtet [4, pp. 165, 223]. Since these results have no direct relevance to our problem, we do not discuss them here. The author wishes to thank the referee for pointing out this reference as well as for other valuable comments.

2. PRELIMINARIES

The Stirling number of the second kind, $S(n, k)$ is defined to be the number of ways of partitioning a set of n elements into k non-empty subsets. Hence

$$S(n, 1) = S(n, n) = 1$$

and

$$S(n, k) = S(n-1, k-1) + kS(n-1, k).$$

LEMMA 2.1. The following recurrence holds for $Z(n)$

$$Z(n) = \sum_{k=1}^{n-1} S(n, k)Z(k), \qquad n \geqslant 2. \tag{2.1}$$

PROOF. The proof is clear.

From this recurrence one can easily derive the functional equation

$$2Z(x) = Z(e^x - 1) + x$$

for the (divergent) exponential generating function

$$Z(x) = \sum_{n=1}^{\infty} Z(n)\frac{x^n}{n!}.$$

This appealing equation has led us nowhere and thus we have been left with the less sophisticated approach that follows below.

Since the dominant terms on the right side of expression (2.1) will be those with k close to n, we shall replace k by $n-k$.

It will be conveinent to consider the quantity

$$Z^*(n)=\frac{Z(n)2^n}{(n!)^2}$$

instead of the rapidly growing $Z(n)$. The function $Z^*(n)$ satisfies the recurrence

$$Z^*(n)=\sum_{k=1}^{n-1} a(n,k)Z^*(n-k), \tag{2.2}$$

where

$$a(n,k)=\frac{S(n,n-k)2^k}{[n]_k^2} \tag{2.3}$$

(We use $[n]_k$ to denote $k!\binom{n}{k}$.)

We shall see that $a(n,k)\sim 1/k!$ for $1\leqslant k<n^{1/3-\varepsilon}$ (ε is an arbitrarily small positive real) and the contribution of the terms $k\geqslant n^{1/5}$ to the right side of expression (2.2) is negligible. In fact this is true even for $k>C'\log n$.

To obtain the asymptotic order of magnitude of $Z^*(n)$, a finer estimate of $a(n,k)$ will be necessary (Lemma 3.1′) which includes the first error term and thus reduces the error to $k^6\mathrm{O}(1/n^2)$.

The next idea is that we explicitly define a function $y(n)$ which nearly satisfies the recurrence (2.2).

Let

$$y(n)=(\log 2)^{-n}n^{-1-(\log 2)/3}.$$

LEMMA 2.2.

$$\sum_{k=1}^{n-1} a(n,k)y(n-k)=y(n)(1+\mathrm{O}(1/n^2)).$$

This result will suffice to guarantee that $Z^*(n)$ and $y(n)$ are asymptotically proportional in view of the following general observation which has been pointed out to me by L. Babai.

LEMMA 2.3. *Let $x(n)$ and $y(n)$ be sequences of positive real numbers. Suppose that $x(n)$ satisfies the recurrence*

$$x(n)=\sum_{k=1}^{n-1} c(n,k)x(n-k), \tag{2.4}$$

where $c(n,k)\geqslant 0$ for $1\leqslant k\leqslant n-1$. Suppose furthermore that for $n\geqslant N_1$

$$(1-\varepsilon_n)y(n)\leqslant\sum_{k=1}^{n-1} c(n,k)y(n-k)\leqslant(1+\varepsilon_n)y(n), \tag{2.5}$$

where $0\leqslant\varepsilon_n<1$. Then for every $n\geqslant N_1$ we have

$$\frac{A}{\prod_{j=N_1}^{n}(1+\varepsilon_j)}\leqslant\frac{y(n)}{x(n)}\leqslant\frac{B}{\prod_{j=N_1}^{n}(1-\varepsilon_j)}, \tag{2.6}$$

where

$$A = \min_{1 \le j \le N_1 - 1} \frac{y(j)}{x(j)} \quad \text{and} \quad B = \max_{1 \le j \le N_1 - 1} \frac{y(j)}{x(j)}.$$

The following corollary is immediate:

COROLLARY 2.4. *Under the conditions of Lemma* 2.3, *if* $\sum_{j=N_1}^{\infty} \varepsilon_j < \infty$ *then there exist positive reals* C_1 *and* C_2 *such that for every* n

$$C_1 \le x(n)/y(n) \le C_2.$$

This corollary together with Lemma 2.2 suffices for the estimate $C_1 f(n) \le Z(n) \le C_2 f(n)$ ($0 < C_1 < C_2$; C_1, C_2 constants) thus yielding the asymptotic order of magnitude of $Z(n)$. In Section 3 we prove the required estimates for Stirling numbers. Lemmas 2.2 and 2.3 will be proved in Section 4.

3. Estimates for Stirling Numbers

Hsu [7] found the asymptotic expansion of $S(n+k, n)$ for fixed k in the form

$$S(n+k, n) = \frac{n^{2k}}{2^k k!}\left(1 + \frac{f_1(k)}{n} + \cdots + \frac{f_t(k)}{n^t} + O(n^{-t-1})\right)$$

where $f_i(k)$ is a polynomial of degree $2i$. Hsu asserts

$$f_1(k) = (2k^2 + k)/3, \qquad f_2(k) = (4k^4 - k^2 - 3k)/18.$$

For $t = 1$ his formula reduces (with some change of notation) to

$$S(n, n-k) = \frac{n^{2k}}{2^k k!}\left\{1 - \frac{k(4k-1)}{3n} + O\left(\frac{1}{n^2}\right)\right\}$$
$$= \frac{[n]_k^2}{2^k k!}\left\{1 - \frac{k(k+2)}{3n} + O\left(\frac{1}{n^2}\right)\right\}.$$

We shall prove the validity of this estimate over a large range of values of k.

LEMMA 3.1. *There is an absolute constant* C'' *such that for* $k^3 < n/16$

$$\left| S(n, n-k) - \frac{[n]_k^2}{k! 2^k}\left\{1 - \frac{k(k+2)}{3n}\right\}\right| < C'' \cdot \frac{[n]_k^2}{k! 2^k} \cdot \frac{k^6}{n^2}. \tag{3.1}$$

Equivalently, we obtain an approximation for $a(n, k)$:

LEMMA 3.1′. *For* $k^3 < n/16$,

$$\left| a(n, k) - \frac{1}{k!}\left\{1 - \frac{k(k+2)}{3n}\right\}\right| < C'' \cdot \frac{1}{k!} \cdot \frac{k^6}{n^2}. \tag{3.2}$$

The following lemma takes care of those values of k not covered by Lemma 3.1.

LEMMA 3.2. *For each* k *in the interval* $3 \log n / \log\log n < k \le n - 1$, *we have*

$$S(n, n-k) < \frac{[n]_k^2}{n^2 \cdot 2^k}, \tag{3.3}$$

when n *is large enough.*

Equivalently, we obtain an approximation for $a(n, k)$:

LEMMA 3.2′. *For each k in the interval $3\log n/\log\log n < k < n-1$, we have*

$$a(n, k) < \frac{1}{n^2}.$$

PROOF OF LEMMA 3.1. $S(n, n-k)$ counts the $(n-k)$-partitions of an n-set. Such a partition consists of l_1 singletons, l_2 pairs, etc., l_i classes of size i. The automorphism group of such a partition has order

$$F(l_1, \ldots, l_n) = l_1!(1!)^{l_1} l_2!(2!)^{l_2} \cdots l_n!(n!)^{l_n}.$$

For $k, L \geq 0$ let $\mathscr{L}(k, L)$ denote the set of those n-tuples $(l_1, \ldots, l_n)$ of nonnegative integers satisfying

$$\sum_{i=1}^{n} il_i = n, \tag{3.4}$$

$$\sum_{i=2}^{n} (i-1)l_i = k, \tag{3.5}$$

$$\sum_{i=3}^{n} (i-2)l_i = L. \tag{3.6}$$

Clearly, $\mathscr{L}(k, L) = \varnothing$ unless $L \leq k$, and $l_{L+3} = \cdots = l_n = 0$ for any $(l_1, \ldots, l_n) \in \mathscr{L}(k, L)$. Now, it is easy to see that

$$S(n, n-k) = \sum_{(l_1,\ldots,l_n)} \frac{1}{l_1!\ldots l_n!}\binom{n}{1^{l_1}, \ldots, n^{l_n}},$$

where the summation extends over all n-tuples satisfying equations (3.4) and (3.5); the parentheses contain a multinomial coefficient. Grouping those n-tuples $(l_1, \ldots, l_n)$ with $\sum_{i=3}^{n} (i-2)l_i = L$ together, we obtain

$$S(n, n-k) = \sum_{L=0}^{k} \left([n]_{2k-L} \sum_{\mathscr{L}(k,L)} \frac{l_1!}{F(l_1, \ldots, l_n)}\right).$$

We shall prove that for $k^3 = \mathrm{o}(n)$ the terms on the right side decrease rapidly with increasing L. In fact for $L \geq 2$, they will contribute only to the error term in (3.1). For $L = 0$ we have $\mathscr{L}(k, 0) = \{(n-2k, k, 0, \ldots, 0)\}$. The corresponding term is

$$\frac{[n]_{2k}}{k!2^k} = \frac{[n]_k^2}{k!2^k}\left(1 - \frac{k^2}{n} + k^4 \mathrm{O}\left(\frac{1}{n^2}\right)\right).$$

For $L = 1$ we have $\mathscr{L}(k, 1) = \{(n-2k+1, k-2, 1, 0, \ldots, 0)\}$. Hence the second term is

$$\frac{[n]_{2k-1}}{(k-2)!2^{k-2}3!} = \frac{[n]_k^2}{k!2^k}\left(\frac{2k(k-1)}{3n} + k^4 \mathrm{O}\left(\frac{1}{n^2}\right)\right).$$

To prove (3.1), we have to give an upper bound for the remaining terms.

Let us observe that

$$k - L = \sum_{i=2}^{n} l_i \leq l_2 + L,$$

hence

$$l_2 \geq k - 2L,$$

and therefore, for $k \geq 2L$,

$$F(l_1, \ldots, l_n) \geq l_1!(k-2L)!2^{k-2L}.$$

Our next claim is that

$$|\mathscr{L}(k, L)| \leqslant (2k)^L.$$

As a matter of fact, $|\mathscr{L}(k, L)|$ is less than the number of nonnegative integral solutions of $L = x_1 + \cdots + x_k$ (by (3.5) and (3.6)), and the latter number is equal to

$$\binom{L+k-1}{k-1} = \binom{L+k-1}{L} \leqslant \binom{2k}{L} \leqslant (2k)^L.$$

Therefore

$$\begin{aligned}\sum_{L=2}^{k} [n]_{2k-L} \sum_{\mathscr{L}(k,L)} \frac{l_1!}{F(l_1, \ldots, l_n)} &\leqslant \sum_{L=2}^{k} [n]_{2k-L} \frac{(2k)^L}{(k-2L)!2^{k-2L}} \\ &< \frac{[n]_k^2}{k!2^k} \sum_{L=2}^{k} \frac{1}{n^L} k^{2L} 2^{2L} (2k)^L \\ &= \frac{[n]_k^2}{k!2^k} \sum_{L=2}^{k} \left(\frac{8k^3}{n}\right)^L \\ &< \frac{[n]_k^2}{k!2^k} \cdot \frac{128k^6}{n^2}\end{aligned}$$

as $n \to \infty$. (For $k < 2L$, the expression $(k-2L)!$ appearing twice in the above proof should in both cases be interpreted to mean 1.)

Proof of Lemma 3.2′. For any n and k, clearly $S(n, k) \leqslant k^n/k!$. Using this inequality for $n-k$ in the place of k, a simple calculation shows that

$$a(n, k) < 2^k \mathrm{e}^n/[n]_k.$$

For $n/2 \leqslant k \leqslant n-1$ the stated inequality $a(n, k) < 1/n^2$ readily follows. We have

$$\frac{a(n, k+1)}{a(n, k)} = \frac{S(n, n-k-1)}{S(n, n-k)} \frac{2}{(n-k)^2}.$$

It is known that the sequence $S(n, 1), S(n, 2), \ldots, S(n, n)$ is logarithmically concave [6], i.e.

$$\frac{S(n, k)}{S(n, k+1)} \geqslant \frac{S(n, k-1)}{S(n, k)}.$$

By repeated application of this inequality we find for any $m \leqslant k \leqslant n-1$

$$\frac{a(n, k+1)}{a(n, k)} \leqslant \frac{2}{(n-k)^2} \frac{S(n, n-m)}{S(n, n-m+1)}. \tag{3.7}$$

Now if $m^5 < n$ we may apply (3.1) to both the numerator and the denominator on the right side of (3.7):

$$\frac{S(n, n-m)}{S(n, n-m+1)} = \frac{(n-m+1)^2}{2m}(1 + \mathrm{o}(1)). \tag{3.8}$$

If in addition $k < n/2$ we obtain from (3.7), (3.8) that

$$\frac{a(n, k+1)}{a(n, k)} < 2\frac{2n^2}{\left(\frac{n}{2}\right)^2 2m} = \frac{8}{m},$$

as n is large enough.

From inequality (3.2) we obtain for $8 \leq m < k < n/2$ and $m^5 < n$ that

$$a(n, k) < a(n, m) \leq \frac{2}{m!}. \tag{3.9}$$

Let, finally, $m = [3 \log n/\log\log n]$. We conclude that $a(n, k) < 2/m! < 1/n^2$ (if n is sufficiently large).

4. Proofs of some Lemmas

Lemma 4.1. *Let f and i be arbitrary fixed reals and let q be such that $|q| < 1$. Then for $m = o(n)$ and $m! > n^2$ we have*

$$\sum_{k=1}^{m} \frac{q^k}{k!}\left(1 - \frac{k+i}{n}\right)^f = e^q - 1 - \frac{f(q+i)(e^q-1) + fq}{n} + O_{f,i,q}\left(\frac{1}{n^2}\right), \tag{4.1}$$

as $n \to \infty$.

Proof. We can expand the left side of equation (4.1) as follows:

$$\begin{aligned}
&\sum_{k=1}^{m} \frac{q^k}{k!}\left\{1 - \frac{f(k+i)}{n} + \frac{k^2}{n^2} O_{f,i,q}(1)\right\} \\
&= \sum_{k=1}^{m} \left\{\frac{q^k}{k!} - \frac{f}{n} q \frac{q^{k-1}}{(k-1)!} - \frac{fi}{n}\frac{q^k}{k!}\right\} \\
&\quad + O_{f,i,q}\left(\frac{1}{n^2}\right) \sum_{k=1}^{m} \frac{q^k}{k!} k^2 \\
&= e^q - 1 - \frac{fq}{n} e^q - \frac{fi}{n}(e^q - 1) + O_{f,i,q}\left(\frac{1}{n^2}\right) \\
&= e^q - 1 - \frac{f(q+i)(e^q-1) + fq}{n} + O_{f,i,q}\left(\frac{1}{n^2}\right),
\end{aligned}$$

as $n \to \infty$. (We used the condition $m! > n^2$ in the next to last equality.)

Proof of Lemma 2.2. Setting $m = [n^{1/5}]$, $f = -1 - (\log 2)/3$ and $q = \log 2$ we obtain by Lemma 4.1, that for sufficiently large n

$$\begin{aligned}
&\sum_{k=1}^{m} \frac{1}{k!}\left\{1 - \frac{k(k+2)}{3n}\right\} y(n-k) \\
&= y(n) \sum_{k=1}^{m} \frac{(\log 2)^k}{k!}\left\{1 - \frac{k(k-1)}{3n} - \frac{3k}{3n}\right\}\left(1 - \frac{k}{n}\right)^f \\
&= y(n)\left\{1 - \frac{1}{n}(2f \log 2 + \tfrac{2}{3}(\log 2)^2 + 2 \cdot \log 2) + O\left(\frac{1}{n^2}\right)\right\} \\
&= y(n)\left\{1 + O\left(\frac{1}{n^2}\right)\right\}.
\end{aligned} \tag{4.2}$$

One can easily see by Lemma 3.2′ that

$$\begin{aligned}
\sum_{k=m+1}^{n-1} a(n, k) y(n-k) &\leq \frac{1}{n^2} \sum_{k=m+1}^{n-1} y(n-k) \\
&\leq \frac{y(n)}{n^2} n (\log 2)^{m+1} n^{-f} \\
&= y(n) O\left(\frac{1}{n^2}\right).
\end{aligned}$$

Now, we have

$$\sum_{k=1}^{n-1} a(n,k)y(n-k)=\sum_{k=1}^{m} a(n,k)y(n-k)+\sum_{k=m+1}^{n-1} a(n,k)y(n-k).$$

We have just proved that the second term here is $y(n)O(1/n^2)$. By Lemma 3.1′, the first term differs from the left hand side of equation (4.2) only by the term $y(n)O(1/n^2)$.

PROOF OF LEMMA 2.3. Let us define the positive numbers γ_n and δ_n for $n \geq N_1$ by

$$\delta_n=\frac{\delta_{n-1}}{1+\varepsilon_n}, \qquad \gamma_n=\frac{\gamma_{n-1}}{1-\varepsilon_n}.$$

Furthermore we set

$$\delta_1=\delta_2=\cdots=\delta_{N_1-1}=A,$$
$$\gamma_1=\gamma_2=\cdots=\gamma_{N_1-1}=B.$$

Clearly, the γ are increasing and the δ are decreasing. Now expression (2.6) can be rewritten as

$$\delta_n \leq \frac{y(n)}{x(n)} \leq \gamma_n, \qquad n=1,2,\dots. \tag{4.3}$$

We prove expression (4.3) by induction on n. The assertion holds for $n=1,2,\dots,N_1-1$ by the definition of A and B. For $n \geq N_1$ we obtain, using the inductive hypothesis, that

$$\begin{aligned}\frac{\gamma_{n-1}}{\gamma_n}y(n)=(1-\varepsilon_n)y(n) &\leq \sum_{k=1}^{n-1} c(n,k)y(n-k)\\ &\leq \sum_{k=1}^{n-1} c(n,k)x(n-k)\gamma_{n-k}\\ &\leq \gamma_{n-1}\sum_{k=1}^{n-1} c(n,k)x(n-k)=\gamma_{n-1}x(n)\end{aligned}$$

and

$$\begin{aligned}\frac{\delta_{n-1}}{\delta_n}y(n)=(1+\varepsilon_n)y(n) &\geq \sum_{k=1}^{n-1} c(n,k)y(n-k)\\ &\geq \sum_{k=1}^{n-1} c(n,k)x(n-k)\delta_{n-k}\\ &\geq \delta_{n-1}\sum_{k=1}^{n-1} c(n,k)x(n-k)=\delta_{n-1}x(n).\end{aligned}$$

In view of the remark made after Corollary 2.4, the proof of Theorem 1.1 is now complete.

ACKNOWLEDGEMENT

I wish to express my gratitude to L. Babai who made many valuable remarks and suggestions.

References

[1] L. Babai, private communication.

[2] L. Babai and T. Lengyel, A convergence criterion for recurrent sequences with application to the partition lattice (in preparation).

[3] E. A. Bender, Asymptotic methods in enumeration, *SIAM Review* **16** (1974), 485–515.

[4] L. Comtet, *Advanced Combinatorics*, D. Reidel, Dordrecht, 1974.

[5] P. Doubilet, G.-C. Rota and R. Stanley, On the foundations of combinatorial theory, VI: the idea of generating function, *Proc. 6th Berkeley Symp. on Math. Stat. and Prob., vol II*, University of California Press, Berkeley, 1970, 267–318.,

[6] L. H. Harper, Stirling behavior is asymptotically normal, *Ann. Math.. Statist.*, **38** (1967), 410–414.

[7] L. C. Hsu, Note on an asymptotic expansion of the n-th difference of zero, *Ann. Math. Statist* **19** (1948), 273–277.

[8] G.-C. Rota, On the foundations of combinatorial theory, I. Theory of Möbius functions, Z. *Warscheinlichkeits-theorie und Verw. Gebiete* **2** (1964), 340-368.

[9] D. Schröder, Vier combinatorische Probleme, *Z. für M. Phys.* **15** (1870), 361-376.

T. Lengyel

Computer and Automation Institute of the Hungarian Academy of Sciences,
Kende u. 13-17, Budapest, Hungary, H-1111