Jumping Champions

Andrew Odlyzko, Michael Rubinstein, and Marek Wolf

CONTENTS

The asymptotic frequency with which pairs of primes below x differ by some fixed integer is understood heuristically, although not rigorously, through the Hardy–Littlewood k-tuple conjecture. Less is known about the differences of consecutive primes. For all x between 1000 and 10^{12}, the most common difference between consecutive primes is 6. We present heuristic and empirical evidence that 6 continues as the most common difference (jumping champion) up to about $x = 1.7427 \cdot 10^{35}$, where it is replaced by 30. In turn, 30 is eventually displaced by 210, which is then displaced by 2310, and so on. Our heuristic arguments are based on a quantitative form of the Hardy–Littlewood conjecture. The technical difficulties in dealing with consecutive primes are formidable enough that even that strong conjecture does not suffice to produce a rigorous proof about the behavior of jumping champions.

1. INTRODUCTION

An integer D is called a *jumping champion* if it is the most frequently occurring difference between consecutive primes $\leq x$ for some x. Occasionally there are several jumping champions. Since the initial primes are 2, 3, 5, 7, 11, the jumping champions are: 1 for $x = 3$; 1 and 2 for $x = 5$; 2 for $x = 7$; and 2 for $x = 11$. (It is clear that we only need to consider prime values of x.)

Jumping champions for various x up to around 1000 are presented in Table 1. Initially 2 and 4 dominate, with 2 showing up more frequently than 4, and 6 showing up only a few times. However, at $x = 563$, $D = 6$ takes over as jumping champion, and except for $x = 941$, where it shares the leadership with $D = 4$, it is the only champion at least up to $x = 10^{12}$. One might therefore be led to conclude that 6 should remain the jumping champion out to infinity. However, this appears to be another of the many number theoretic functions where the initial behavior is misleading. We will present heuristics that suggest that 6 does not remain a jumping champion forever.

1058-6458/1999 $0.50 per page

x	Champion(s)	x	Champion(s)
5	1 2	421	2 6
7	2	431	2 6
11	2	433	2
⋮	⋮	439	2 6
97	2	443	2 6
101	2 4	449	6
103	2	457	6
107	2 4	461	6
109	2	463	2 6
113	2 4	467	2 4 6
127	2 4	479	2 4 6
131	4	487	2 4 6
137	4	491	4
139	2 4	⋮	⋮
149	2 4	541	4
151	2	547	4 6
157	2	557	4 6
163	2	563	6
167	2 4	⋮	⋮
173	2 4	937	6
179	2 4 6	941	4 6
181	2	947	6
⋮	⋮	953	6
373	2	967	6
379	2 6	971	6
383	2 6	977	6
389	6	983	6
397	6	⋮	⋮
401	6	$1.7427 \cdot 10^{35}$	? 30 ?
409	6	⋮	⋮
419	6	10^{425}	? 210 ?

TABLE 1. Jumping champions for small x.

Conjecture 1. The jumping champions are 4 and the primorials 2, 6, 30, 210, 2310, ...

The heuristics (see Section 2) suggest that 6 is the jumping champion up to about $x = 1.7427 \cdot 10^{35}$, where 30 becomes the jumping champion. (Harley [1994], stimulated by a report on an early phase of our research, has independently computed this number as the point of transition between 6 and 30.) In turn, 30 is displaced as jumping champion by 210 around $x = 10^{425}$. This is substantiated by numerical experimentation (see the end of Section 2 and Table 3). It is likely that in the transition zones, the two contenders in all cases trade places as jumping champions, but we have neither the computing power to verify this numerically nor the theoretical tools to prove it. Although Conjecture 1 is very simple and elegant, it is surprisingly deep.

The heuristics we develop are based on the famous Hardy–Littlewood k-tuple conjecture. Recall that the twin prime conjecture says that there exist infinitely many primes p such that $p+2$ is also a prime. On the other hand, there is only a single prime p such that p, $p+2$, and $p+4$ are all primes, since at least one of these three integers is divisible by 3. The Hardy–Littlewood k-tuple conjecture [1922] is that unless there is a trivial divisibility condition that stops $p, p+a_1, \dots, p+a_k$ from consisting of primes infinitely often, then such prime tuples will occur, and will even occur with a certain asympotic density that is easy to compute in terms of the a_i. While there is a general belief that the k-tuple conjecture is true, it remains unproved.

There seems to be little hope of making any progress towards a proof of Conjecture 1 without assuming at least a quantitative form of the k-tuple conjecture. However, as we will show, even assuming the strongest form of that conjecture that seems reasonable in view of our knowledge of prime numbers, we are still left with formidable obstacles that prevent us from obtaining a complete proof of Conjecture 1. Still, in investigating jumping champions, we are led to some nice combinatorics related to the coefficients in the k-tuple conjecture.

A strong form of the k-tuple conjecture leads to an explicit asymptotic formula for the frequency with which an integer D appears as the difference of consecutive primes $\leq x$. This formula has some interesting arithmetical properties, and it leads to the "irregularly regular" behavior shown in Figure 1. Brent [1974] was the first to suggest this formula and gave an algorithm for computing certain coefficients that arise in it.

A conjecture that follows from Conjecture 1, but should be considerably easier to prove, and might conceivably be provable unconditionally, is the following.

Conjecture 2. The jumping champions tend to infinity. Furthermore, any fixed prime p divides all sufficiently large jumping champions.

The first part of Conjecture 2 was proved by Erdős and Straus [1980] under the assumption of a quantitative form of the k-tuple conjecture.

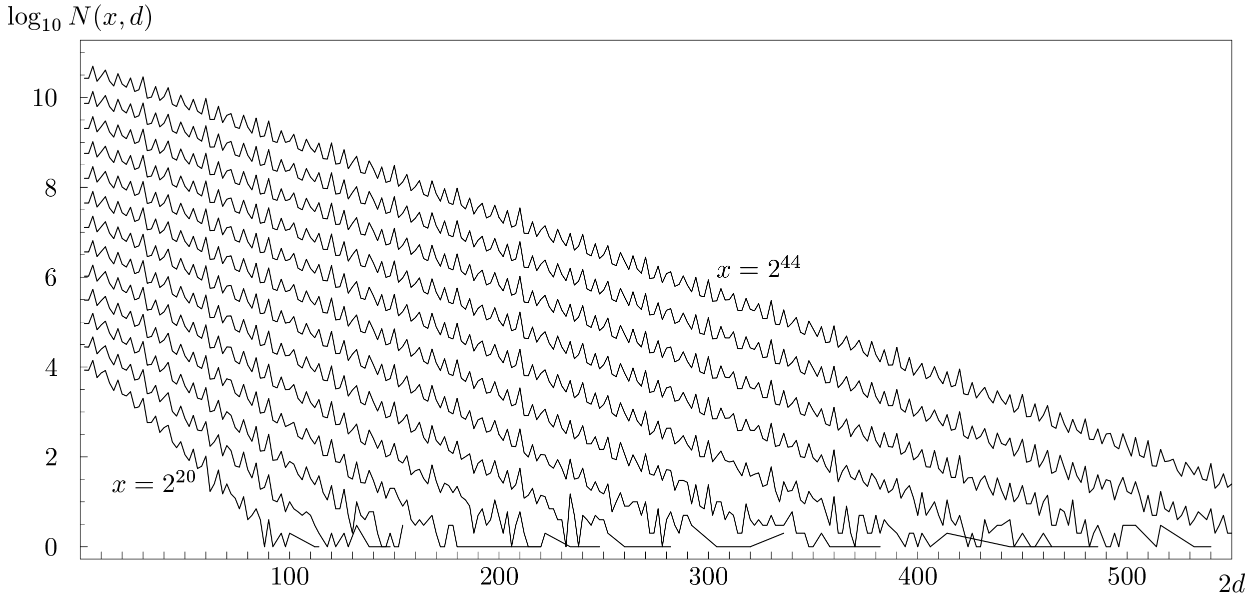

FIGURE 1. Log plot of $\log N(x,d)$, the number of jumps of length $2d$ between consecutive primes up to x, for $x = 2^{20}, 2^{22}, \ldots, 2^{44}$. Integrating (2–8) by parts and taking logarithms we see that, for fixed x, the graph should follow a straight line with small pertubations of size $\log A_{d,1}$ (see text); this can be seen clearly in the figure. The prominent bump at $2d = 210$ reflects the relatively large size of $A_{105,1}$.

As far as we are aware, the first question about the behavior of jumping champions was raised by Harry Nelson [1978–79] — though the term jumping champion itself was introduced by John Horton Conway in 1993. Erdős and Straus, motivated by Nelson's note, proved, under the assumption of a form of the k-tuple conjecture, that jumping champions for x tend to infinity with x. They also raised the question of the rate at which champions tend to infinity. We answer this question in our note, assuming stronger conjectures, as Erdős and Straus suggested might have to be done. Evidence suggests that the size of the champion jumps from

$$(1 + o(1)) \log x / (\log\log x)^2$$

to

$$(1 + o(1)) \log x / (\log\log x)$$

at the transition point, and then, as x increases, slowly goes down again to

$$(1 + o(1)) \log x / (\log\log x)^2.$$

Jumping champions have been thought about independently several times since the work of Erdős and Straus. We were led to look at them by Conway. Mealy and Leech have also asked about their behavior [Guy 1994].

2. THE HEURISTICS

The k-tuple Conjecture

Let $0 < m_1 < m_2 < \cdots < m_k$. The k-tuple conjecture predicts that the number of primes $p \le x$ such that $p + 2m_1$, $p + 2m_2$, …, $p + 2m_k$ are all prime is

$$P(x; m_1, m_2, \ldots, m_k) \sim C(m_1, m_2, \ldots, m_k) \int_2^x \frac{dt}{\log^{k+1} t}, \qquad (2\text{–}1)$$

where

$$C(m_1, m_2, \ldots, m_k) = 2^k \prod_q \frac{(1 - w(q; m_1, m_2, \ldots, m_k)/q)}{(1 - 1/q)^{k+1}}. \qquad (2\text{–}2)$$

In Equation (2–2), q runs over all odd primes, and $w(q; m_1, m_2, \ldots, m_k)$ denotes the number of distinct residues of $0, m_1, m_2, \ldots, m_k \bmod q$. Note that if $k = 1$ then

$$C(m) = 2 \prod_q \frac{q(q-2)}{(q-1)^2} \prod_{q \mid m} \frac{(q-1)}{(q-2)} \qquad (2\text{–}3)$$

depends only on the odd primes dividing m, and $C(m_1) = C(m_2)$ if and only if m_1 and m_2 have the same odd prime factors (possibly raised to different powers).

For a discussion on the k-tuple conjecture and references to numerical computations in its support, see [Halberstam and Richert 1974, Introduction].

Brent [1974; 1975] was apparently the first one to study the size of the error term in the k-tuple conjecture. Hardy and Littlewood did not make any predictions about its size, although the standard arguments that assume random cancellation of various terms suggest it should be about $\sqrt{x}$ for each k-tuple. Brent's computations [1975, Table 4] support this suggestion for tuples $p, p+2$ where we find a remainder with roughly half as many digits as the main term. See also the comment following (2–7).

The Heuristics

Let $N(x,d)$ be the number of primes $p \le x$ such that $p+2d$ is the smallest prime $> p$. By inclusion-exclusion we have, for each $K = 0, 1, \ldots,$

$$N(x,d) \le \sum_{k=0}^{2K} (-1)^k \sum_{0<m_1<\cdots<m_k<d} P(x; m_1, \ldots, m_k, d) \tag{2–4}$$

and

$$N(x,d) \ge \sum_{k=0}^{2K+1} (-1)^k \sum_{0<m_1<\cdots<m_k<d} P(x; m_1, \ldots, m_k, d) \tag{2–5}$$

(here the $k=0$ term is $P(x;d)$). So it is natural to compare $N(x,d)$ with

$$\int_2^x \sum_{k=1}^{M} \frac{A_{d,k}}{\log^{k+1} t} dt \tag{2–6}$$

where M is a positive integer and

$$A_{d,k} = (-1)^{k+1} \sum_{0<m_1<\cdots<m_{k-1}<d} C(m_1, \ldots, m_{k-1}, d) \tag{2–7}$$

(here $A_{d,1} = C(d)$).

Computations of Brent [1974] indicate that taking all the terms in (2–6) (that is, choosing M so that $A_{d,M+1} = 0$) approximates $N(x,d)$ to within $O(x^{1/2})$. This can be seen in [Brent 1974, Table 2], which shows an agreement between theoretical approximation and reality to roughly half the decimal places.

Now, the sum in (2–7) runs over $\binom{d-1}{k-1}$ terms and it would not be unreasonable to guess that $A_{d,k}$ grows nicely with this binomial coefficient. In fact, we show in Theorem 3 (Section 3) that for k fixed,

$$A_{d,k+1} \sim (-1)^k A_{d,1} \frac{(2d)^k}{k!} \quad \text{as } d \to \infty.$$

This suggests, in conjuction with (2–6), that, for d large,

$$N(x,d) \sim A_{d,1} \int_2^x \frac{\exp(-2d/\log t)}{\log^2 t} dt \tag{2–8}$$

should approximate well the number of gaps of size $2d$ up to height x. However, not only does d have to be large for this to be a good approximation, but x has to be large compared to d, and this restricts the range in which we may use (2–8).

The presence of the $A_{d,1}$ factor in (2–8) indicates that, in order to make $N(x,d)$ huge, it is preferable for d to have many small prime factors. On the other hand, the $\exp(-2d/\log t)$ term in the integrand tells us that amongst all d that produce the same value for $A_{d,1}$, the smallest one wins. More precisely, let

$$\begin{aligned} 2d_1 &= 2^{a_0} p_1^{a_1} \cdots p_j^{a_j}, \\ 2d_2 &= 2p_1 \cdots p_j, \\ 2d_3 &= 2\cdot 3 \cdots q_j, \end{aligned}$$

where $a_i \ge 1$, the p_i's are odd primes, and q_j is the j-th odd prime ($q_1 = 3$, $q_2 = 5$, and so on). Note that $d_3 \le d_2 \le d_1$.

Formula (2–8) tells us that, for d_3 large enough, we should expect

$$N(x, d_2) \ge N(x, d_1)$$

(because $A_{d_2,1} = A_{d_1,1}$ but $d_2 \le d_1$), and $N(x,d_3) \ge N(x,d_2)$ (because $A_{d_3,1} \ge A_{d_2,1}$ and $d_3 < d_2$). So we see that primorials are favored.

Furthermore, integrating Formula (2–8) by parts, we find that $N(x, 3\cdots q_{j+1})$ should begin to overtake $N(x, 3\cdots q_j)$ *roughly* when

$$\frac{q_{j+1}-1}{q_{j+1}-2} \exp\frac{-2\cdot 3\cdots q_{j+1}}{\log x} > \exp\frac{-2\cdot 3\cdots q_j}{\log x},$$

that is, roughly when

$$x > \exp(2\cdot 3\cdots q_j (q_{j+1}-1)(q_{j+1}-2)).$$

These considerations justify Conjecture 1, at least for sufficiently large gaps (and very large x). For smaller d, rather than using (2–8), we could use the first few terms of (2–6) to study $N(x,d)$.

For example, $A_{1,1} = A_{2,1}$, and $A_{2,2} = 0$ (since there are no triplets of primes $p, p+2, p+4$ other than $3, 5, 7$). Hence both $N(x,1)$ and $N(x,2)$ should be very close to

$$A_{1,1} \int_2^x \frac{dt}{\log^2 t}.$$

This explains why 4 also appears as a champion.

We can also determine roughly when 30 will take over from 6 as champion, and when 210 will first beat 30. Using the coefficients from [Brent 1974] to compute (2–6) with all the terms ($M = 2$ when $2d = 6$ and $M = 8$ when $2d = 30$), we find that 30 should take over as champion roughly at $x = 1.7427 \cdot 10^{35}$. Further, taking $M = 4$ terms in (2–6), predicts that 210 will first begin to beat 30 sometime in the interval $10^{425} < x < 10^{426}$. Numerical experimentation substantiates these claims. We used Maple's probable prime function to test intervals of length 10^7. If all the probable primes that this function produced for us are indeed prime, then in the interval $[10^{30}, 10^{30} + 10^7]$ there are 5278 gaps of size 6, and 5060 gaps of size 30, whereas in the interval $[10^{40}, 10^{40} + 10^7]$ there are 3120 gaps of size 6 and 3209 gaps of size 30. (Note that even if some of the probable primes we found are not prime, it is extremely likely there are few of them, so the statistics we produce would not be noticeably affected.) Further, in the intervals $[10^{400}, 10^{400} + 10^7]$ we find that gaps of size 30 and 210 show up 50 and 33 times, respectively, and 26 and 34 times in the interval $[10^{450}, 10^{450} + 10^7]$. These last results are only roughly indicative of true behavior, since sample sizes are so small. In fact, in our data for 10^{450}, 198 appears to be the champion, as it shows up as a gap of consecutive primes 40 times!

Section 3 is devoted to studying the coefficients $A_{d,k}$ that appear in (2–6).

3. THE COEFFICIENTS $A_{d,k}$

We turn now to the problem of estimating the coefficients $A_{d,k}$ that appear in (2–6). In this section we use the O-notation: $a = O(b)$ is equivalent to $|a| \le K\,|b|$ for some constant K, and $a = O_c(b)$ is equivalent to $|a| \le K(c)\,|b|$ for some $K(c)$.

We can prove (unconditionally)

Theorem 3. *Let $1 \le k \le c \log\log d$, where c is a constant. Then*

$$A_{d,k+1} = -A_{d,k}\frac{2d}{k}\left(1 + O_c(k/\log\log d)\right) \qquad (3\text{–}1)$$

Remark. Numerical data suggest (see Figure 2) that the $1 + O_c(k/\log\log d)$ in this formula can be replaced by $1 + O(k \log d/d)$ or even $1 + o(k \log d/d)$.

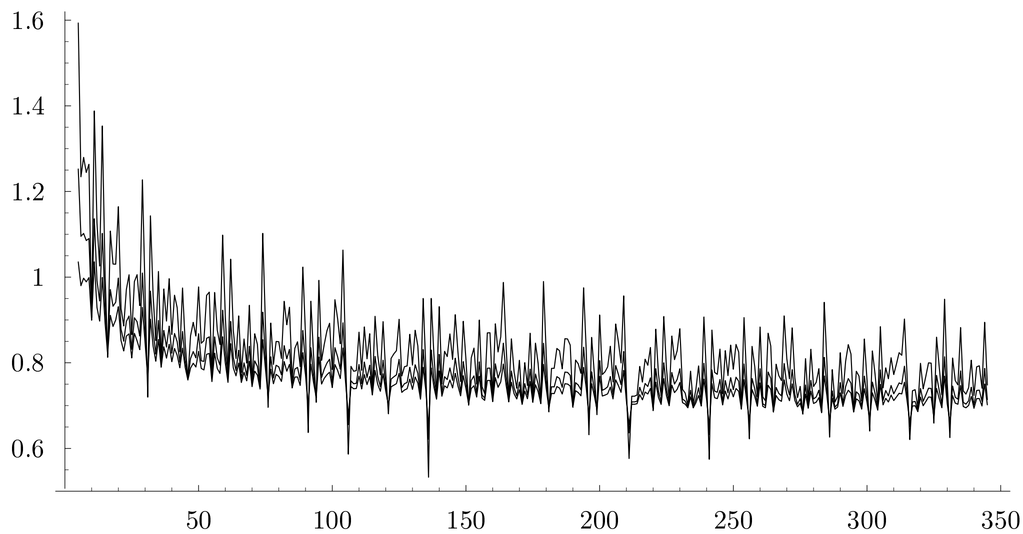

FIGURE 2. A figure substantiating the remark that follows (3–1). It shows (superimposed on one another) the graphs of d versus $\left(\frac{1}{k} + \frac{1}{2d}\frac{A_{d,k+1}}{A_{d,k}}\right)\frac{d}{\log d}$, for $k = 1, 2, 3$. According to the remark, these graphs should all be bounded. The picture suggests that they not only are bounded, but, for fixed d, deviate from some constant value by an amount proportional to $1/k$.

Proof. First observe that if $A_{d,k}=0$ then $A_{d,k+1}=0$ and the theorem holds trivially. ($A_{d,k}=0$ implies that all p, $p+2m_1$, $\dots$, $p+2m_{k-1}$, $p+2d$ tuples are ruled out. Hence, so are all the p, $p+2m_1$, $\dots$, $p+2m_k$, $p+2d$ tuples, because each one contains (many) p, $p+2m_1$, $\dots$, $p+2m_{k-1}$, $p+2d$ subtuples). Therefore, assume $A_{d,k}\neq 0$. To shorten the notation, we set $w_{r,s}(q;d)=w(q;m_r,m_{r+1},\dots,m_s,d)$. From (2–2) and (2–7) we have

$$\frac{A_{d,k+1}}{A_{d,k}}=\frac{-2\sum_{\substack{0<m_1<\cdots\\<m_k<d}}\prod_q\bigl(1-w_{1,k}(q;d)/q\bigr)}{\sum_{\substack{0<m_1<\cdots\\<m_{k-1}<d}}\prod_q\bigl(1-w_{1,k-1}(q;d)/q\bigr)(1-1/q)}.$$

(If $k=1$, the sum in the denominator has a single term, $\prod_q\bigl(1-w(q;d)/q\bigr)(1-1/q)$.) If $q>d$, we have $w_{1,k}(q;d)=k+2$ and $w_{1,k-1}(q;d)=k+1$. So the quotient above is

$$\frac{A_{d,k+1}}{A_{d,k}}=-2P_1P_2, \tag{3–2}$$

with

$$P_1=\frac{\sum_{\substack{0<m_1<\cdots\\<m_k<d}}\prod_{q\le d}\bigl(1-w_{1,k}(q;d)/q\bigr)}{\sum_{\substack{0<m_1<\cdots\\<m_{k-1}<d}}\prod_{q\le d}\bigl(1-w_{1,k-1}(q;d)/q\bigr)(1-1/q)} \tag{3–3}$$

and

$$P_2=\prod_{q>d}\frac{(1-(k+2)/q)}{(1-1/q)(1-(k+1)/q)}. \tag{3–4}$$

P_2 poses little difficulty and is easily estimated by using the Taylor series for $\log(1-x)$,

$$P_2=\exp\biggl(-\sum_{m=2}^{\infty}\sum_{q>d}\frac{1}{m}\biggl(\Bigl(\frac{k+2}{q}\Bigr)^m-\Bigl(\frac{k+1}{q}\Bigr)^m-\frac{1}{q^m}\biggr)\biggr), \tag{3–5}$$

for $k+2\le d$.

Now $0<(k+2)^m-(k+1)^m-1<m(k+2)^{m-1}$, for $m\ge 2$, which can be seen by writing

$$\begin{aligned}(k+2)^m&-(k+1)^m\\&=(k+2)^{m-1}+(k+2)^{m-2}(k+1)+\cdots+(k+1)^{m-1}.\end{aligned}$$

Hence

$$1>P_2>\exp\biggl(-\sum_{m=2}^{\infty}(k+2)^{m-1}\sum_{q>d}\frac{1}{q^m}\biggr).$$

But

$$\sum_{q>d}\frac{1}{q^m}<\sum_{n=d+1}^{\infty}\frac{1}{n^m}<\int_d^{\infty}\frac{dt}{t^m}=\frac{1}{(m-1)}\frac{1}{d^{m-1}},$$

so

$$\begin{aligned}1>P_2&>\exp\biggl(-\sum_{m=2}^{\infty}\frac{1}{(m-1)}\frac{(k+2)^{m-1}}{d^{m-1}}\biggr)\\&=1-\frac{k+2}{d},\end{aligned}$$

for $k+2<d$, which is to say

$$P_2=1+O(k/d), \tag{3–6}$$

for $k+2<d$. In fact, a better estimate is not hard to establish. Since (3–6) contributes less than the error claimed in the theorem, we omit the proof and simply state that

$$P_2=1-\frac{k}{d\log d}+O\left(\frac{1}{d\log d}+\frac{k}{d\log^2 d}\right), \tag{3–7}$$

for $k<d/2$.

Next, consider P_1. On scrutinizing (3–3), we see that each term in the denominator may be matched with terms in the numerator. We write

$$P_1=\frac{1}{k}\frac{\sum_{\substack{0<m_1<\cdots\\<m_{k-1}<d}}\sum_{\substack{0<m_0<d\\m_0\neq m_1,\dots,m_{k-1}}}\prod_{q\le d}\bigl(1-w_{0,k-1}(q;d)/q\bigr)}{\sum_{\substack{0<m_1<\cdots\\m_{k-1}<d}}\prod_{q\le d}\,(1-w_{1,k-1}(q;d)/q)\,(1-1/q)} \tag{3–8}$$

and claim that each inner sum in the numerator is approximately d times its corresponding term in the denominator. More precisely, we show that, for $k\le c\log\log d$, with c a constant,

$$\begin{aligned}\sum_{\substack{0<m_0<d\\m_0\neq m_1,\dots,m_{k-1}}}&\prod_{q\le d}\bigl(1-w_{0,k-1}(q;d)/q\bigr)\\&=d(1+O_c(k/\log\log d))\\&\quad\times\prod_{q\le d}\bigl(1-w_{1,k-1}(q;d)/q\bigr)(1-1/q).\end{aligned} \tag{3–9}$$

The theorem would then follow on combining (3–9) with (3–8), (3–6), and (3–2).

To prove (3–9), break up $\prod_{q\le d}$ into two pieces. Let $d\ge 15$ satisfy

$$3\cdot 5\cdots q_a\le d<3\cdot 5\cdots q_{a+1}, \tag{3–10}$$

and write

$$\prod_{q\le d}=\prod_{q\le q_{a-1}}\prod_{q_a\le q\le d}. \tag{3–11}$$

By the Prime Number Theorem,

$$q_a \sim \log d. \tag{3–12}$$

Now, if the right-hand side of (3–9) is zero (this happens if $w_{1,k-1}(q;d) = q$ for some $q \le d$) then so is the left-hand side (since then $w_{0,k-1}(q;d)$ also equals q), and (3–9) is trivially true. So, assume that this isn't the case and consider

$$\sum_{\substack{0<m_0<d \\ m_0 \neq m_1,\ldots,m_{k-1}}} \prod_{q \le q_{a-1}} \prod_{q_a \le q \le d} f_q(m_0,\ldots,m_{k-1},d), \tag{3–13}$$

where

$$f_q(m_0,\ldots,m_{k-1},d) = \frac{(1-w_{0,k-1}(q;d)/q)}{(1-w_{1,k-1}(q;d)/q)\,(1-1/q)}.$$

To simplify things, (3–13) may be written as

$$\sum_{m_0=1}^{d} \prod_{q \le q_{a-1}} \prod_{q_a \le q \le d} f_q(m_0,\ldots,m_{k-1},d) - k \prod_{q \le d} \frac{1}{1-1/q}.$$

The second term above is $O(k \log d)$ (in fact, by a theorem of Mertens [Ingham 1932], it contributes $\sim -\frac{k}{2} e^{\gamma} \log d$) and will be overshadowed by the first term. So, let

$$S = \sum_{m_0=1}^{d} \prod_{q \le q_{a-1}} \prod_{q_a \le q \le d} f_q(m_0,\ldots,m_{k-1},d). \tag{3–14}$$

Our goal is to show $S = d(1+O(k/\log\log d))$. We first estimate the contribution from $\prod_{q_a \le q \le d}$. The value of $w_{0,k-1}(q;d)$ is

$$w_{0,k-1}(q;d) = \begin{cases} w_{1,k-1}(q;d) & \text{if } q \mid m_0(m_1-m_0)\cdots(m_{k-1}-m_0)(d-m_0), \\ w_{1,k-1}(q;d)+1 & \text{otherwise.} \end{cases}$$

For most q (when k is small compared to d) the latter holds. In fact, let

$$L = \#\left\{ \begin{matrix} q \text{ such that } q_a \le q \le d \text{ and } q \text{ divides} \\ m_0(m_1-m_0)\cdots(m_{k-1}-m_0)(d-m_0) \end{matrix} \right\}. \tag{3–15}$$

Now, $m_0(m_1-m_0)\cdots(m_{k-1}-m_0)(d-m_0) < d^{k+1}$, so $q_a^L < d^{k+1}$. Hence, from (3–12),

$$L = O\left(\frac{k \log d}{\log\log d}\right). \tag{3–16}$$

But

$$\prod_{q_a \le q \le d} \frac{1-(k+2)/q}{(1-1/q)(1-(k+1)/q)} \le \prod_{q_a \le q \le d} f_q \le \frac{1}{(1-1/q_a)^L}.$$

The left-hand side above is roughly of the same form as (3–4), and by (3–6), it is $1+O(k/q_a) = 1+O(k/\log d)$, (so long as $k < (q_a - 2) \sim \log d$). Meanwhile,

$$\frac{1}{(1-1/q_a)^L} = e^{O(L/q_A)} = e^{O(k/\log\log d)} = 1 + O_c\left(k/\log\log d\right),$$

assuming $k \le c \log\log d$, with c a constant. Therefore, pulling out $\prod_{q_a \le q \le d} f_q$ from (3–14), we get

$$S = (1 + O_c\left(k/\log\log d\right)) \times \sum_{m_0=1}^{d} \prod_{q \le q_{a-1}} f_q(m_0,\ldots,m_{k-1},d), \tag{3–17}$$

for $k \le c \log\log d$. Next, write

$$d = \alpha(3\cdot 5 \cdots q_{a-1}) + \beta = \alpha Q + \beta,$$

where, by (3–10), $\alpha, \beta \in \mathbb{Z}$, $\alpha \ge q_a$, and $0 \le \beta < 3\cdot 5 \cdots q_{a-1}$. Then break up the sum over m_0:

$$\sum_{m_0=1}^{d} = \sum_{m_0=1}^{\alpha Q} + \sum_{\alpha Q+1}^{d}.$$

The contribution of the second sum on the right is $O(\beta \log\log d)$, as can be seen from

$$\prod_{q \le q_{a-1}} f_q \le \prod_{q \le q_{a-1}} 1/(1-1/q).$$

But $\beta < d/q_a = O(d/\log d)$, so the contribution to (3–17) from this sum is $O(d \log\log d/\log d)$. To complete our proof we show that

$$\sum_{m_0=1}^{\alpha Q} \prod_{q \le q_{a-1}} f_q(m_0,\ldots,m_{k-1},d) = \alpha Q = d(1+O(1/\log d)). \tag{3–18}$$

This in combination with all our other estimates will establish the theorem.

To prove (3–18), break up the range of summation $m_0 = 1,\ldots,\alpha Q$ into blocks of length Q (there are

α such blocks). Each block contributes the same amount to (3–18) because

$$\prod_{q \le q_{a-1}} f_q(m_0, \ldots, m_{k-1}, d)$$

depends only on the values modulo Q of its arguements. Next, we show by induction on a that

$$\sum_{m_0=1}^{q_1 \cdots q_{a-1}} \prod_{q \le q_{a-1}} f_q(m_0, \ldots, m_{k-1}, d) = Q. \tag{3–19}$$

If $a - 1 = 1$, our sum is

$$\sum_{m_0=1}^{q_1} f_{q_1}(m_0, \ldots, m_{k-1}, d) \tag{3–20}$$

We find that (3–20) sums to

$$w_{1,k-1}(q_1; d) \frac{1}{1 - 1/q_1} + (q_1 - w_{1,k-1}(q_1; d)) \frac{1 - (w_{1,k-1}(q_1; d) + 1)/q}{(1 - w_{1,k-1}(q_1; d)/q_1)(1 - 1/q_1)},$$

which equals q_1. Now say that (3–19) has been proved for $a - 1$ and consider the a case

$$\sum_{m_0=1}^{q_1 \cdots q_a} \prod_{q \le q_a} f_q(m_0, \ldots, m_{k-1}, d).$$

Group the m_0's according to their values modulo q_a

$$\sum_{n_0=1}^{q_a} \sum_{n=0}^{q_1 \cdots q_{a-1} - 1} \prod_{q \le q_a} f_q(nq_a + n_0, m_1, \ldots, m_{k-1}, d).$$

Now, because f_{q_a} only depends on its values modulo q_a, the expression above is

$$\sum_{n_0=1}^{q_a} f_{q_a}(n_0, m_1, \ldots, m_{k-1}, d) \times \sum_{n=0}^{q_1 \cdots q_{a-1} - 1} \prod_{q \le q_{a-1}} f_q(nq_a + n_0, m_1, \ldots, m_{k-1}, d).$$

As n runs from 0 to $q_1 \cdots q_{a-1} - 1$, the value of $nq_a + n_0$ runs over the complete set of residues modulo $q_1 \cdots q_{a-1}$, since q_a is relatively prime to $q_1 \cdots q_{a-1}$. Hence the inner sum is, by our induction hypothesis, equal to $q_1 \cdots q_{a-1}$, so the expression above equals

$$q_1 \cdots q_{a-1} \sum_{n_0=1}^{q_a} f_{q_a}(n_0, m_1, \ldots, m_{k-1}, d) = q_1 \cdots q_{a-1} q_a = Q. \qquad \square$$

Remarks. Gallagher [1976] studied the combinatorics of a related problem, essentially that of the asymptotics of the sum $\sum_{d \le M} A_{d,k}$. His method can be adapted for our problem, albeit with messier combinatorics. The remainder term obtained grows very quickly with k (though for small k, his method provides a stronger result). On the other hand, Theorem 3 can be used, along with Corollary 5 below and summation by parts, to obtain the asymptotics of $\sum_{d \le M} A_{d,k}$ (though, they are not needed for the champions problem).

To establish Corollary 5 we first give a general counting formula which is useful for averaging certain types of products. As usual $\{x\} = x - \lfloor x \rfloor$ will denote the fractional part of x, and empty products are taken to be 1.

Theorem 4. *Let S be a set of pairwise relatively prime positive integers, and let f be a complex valued function on this set. Then*

$$\sum_{d=1}^{M} \prod_{\substack{a \mid d \\ a \in S}} f(a) = M \prod_{\substack{a \le M \\ a \in S}} \Bigl(1 + \frac{1}{a}(f(a) - 1)\Bigr) - \sum_{\sigma} \Bigl\{ \frac{M}{\prod_{a \in \sigma} a} \Bigr\} \prod_{a \in \sigma} (f(a) - 1),$$

where σ ranges over all finite non-empty subsets of S whose elements are at most M.

This formula can be derived using an inclusion-exclusion argument as in the sieve of Eratosthenes.

Corollary 5.

$$\sum_{d=1}^{M} A_{d,1} = 2M \prod_{q > M} \frac{q(q-2)}{(q-1)^2} - A_{1,1} \sum_{i=1}^{\pi(M)-1} \sum_{\substack{q_1 < \cdots \\ < q_i \le M}} \Bigl\{ \frac{M}{q_1 \cdots q_i} \Bigr\} \frac{1}{(q_1 - 2) \cdots (q_i - 2)}.$$

This implies

$$\sum_{d=1}^{M} A_{d,1} = 2M + O(\log M).$$

Proof. The first part of the corollary follows from Theorem 4 and formulas (2–7) and (2–3).

The second part follows by noting that

$$\prod_{q > M} \frac{q(q-2)}{(q-1)^2} = 1 + O(M^{-1}),$$

and

$$0 \le \sum_{i=1}^{\pi(M)-1} \sum_{\substack{q_1<\cdots\\<q_i\le M}} \left\{ \frac{M}{q_1\cdots q_i} \right\} \frac{1}{(q_1-2)\cdots(q_i-2)}$$

$$< \prod_{q\le M} \left(1+\frac{1}{q-2}\right) = O(\log M). \qquad \square$$

This corollary was also proved in [Bombieri and Davenport 1966, p. 10], but with $O(\log^2 M)$ instead of $O(\log M)$ for the remainder, and, with the stronger remainder, in [Montgomery 1971, Lemma 17.4].

4. ADDITIONAL DATA

Table 2 compares two estimates for $N(x,d)$ at $x = 10^{12}$. The first estimate was computed using (2–6) with $M = 4$. The second was computed using (2–8). The table shows that the higher terms in (2–6) are important for estimating $N(x,d)$ if d is allowed to grow (notice that the middle column gives a good approximation roughly up to $d = 18$). Brent [1974] has observed this fact. His computations also show that taking all the terms in (2–6) gives numbers that agree very well with $N(x,d)$. This is what (2–8) attempts to do (in closed form). However, d needs to be large for (2–8) to be a good approximation and x has to be large compared to d (though, even for small d and x not too huge, the table reveals that (2–8) gives a decent, uniform approximation to $N(x,d)$).

Figure 3 plots $N(x,d)\log^2(x)/x$ versus x, for various values of d, with emphasis on the primorials $2d = 6$ and $2d = 30$. (The $\log^2(x)/x$ factor was included to stabilize the curves.) The picture shows 6 dominating as champion for $x > 941$, presumably until roughly $x = 1.7427\cdot 10^{35}$.

Table 3 on page 117 shows the number $g_{2d}(x)$ of gaps of size $2 \le 2d \le 240$ presumed to lie in the intervals $[x, x+10^7]$, for $x = 10^{30}$, 10^{40}, 10^{400}, 10^{450}. These values of x are chosen to illustrate the predicted transitions. At $x = 10^{30}$, $g_6(10^{30}) = 5278$ dominates $g_{30}(10^{30}) = 5060$, but $g_{30}(10^{40}) = 3209$ beats $g_6(10^{40}) = 3120$. Next, $g_{30}(10^{400}) = 50$ and $g_{210}(10^{400}) = 33$, but $g_{30}(10^{450}) = 26$ and $g_{210}(10^{450}) = 34$. These numbers are consistent with our predictions that 30 begins to beat 6 as champion near $x = 10^{35}$, and that 210 first beats 30 near $x = 10^{425}$. Note, however, that, at $x = 10^{450}$, the apparent champion seems to be $2d = 198$, which shows up 40 times! Such are the dangers of working with small samples.

REFERENCES

[Bombieri and Davenport 1966] E. Bombieri and H. Davenport, "Small differences between prime numbers", *Proc. Roy. Soc. Ser. A* **293** (1966), 1–18.

[Brent 1974] R. P. Brent, "The distribution of small gaps between successive primes", *Math. Comp.* **28** (1974), 315–324.

[Brent 1975] R. P. Brent, "Irregularities in the distribution of primes and twin primes", *Math. Comp.* **29** (1975), 43–56. Correction in **30** (1976), 198.

[Erdős and Straus 1980] P. Erdős and E. G. Straus, "Remarks on the differences between consecutive primes", *Elem. Math.* **35**:5 (1980), 115–118.

[Gallagher 1976] P. X. Gallagher, "On the distribution of primes in short intervals", *Mathematika* **23**:1 (1976), 4–9. Corrigendum in **28**:1 (1981), 86.

[Guy 1994] R. K. Guy, *Unsolved problems in number theory*, Second ed., Springer, New York, 1994.

[Halberstam and Richert 1974] H. Halberstam and H.-E. Richert, *Sieve methods*, London Mathematical Society Monographs **4**, Academic Press, New York, 1974.

[Hardy and Littlewood 1922] G. H. Hardy and J. E. Littlewood, "Some problems of 'partitio numerorum', III: On the expression of a number as a sum of primes", *Acta Math.* **44** (1922), 1–70. Reprinted as pp. 561–630 in *Collected Papers of G. H. Hardy*, vol. 1, Oxford Univ. Press, 1966.

[Harley 1994] R. Harley, "Some estimates by Richard Brent applied to the "high jumpers" problem", 1994. See http://pauillac.inria.fr/~harley/wnt.html.

[Ingham 1932] A. E. Ingham, *The distribution of prime numbers*, Cambridge tracts in mathematics and mathematical physics **30**, Cambridge Univ. Press, Cambridge, 1932. Reprinted by Stechert-Hafner, New York, 1964; Cambridge Univ. Press, 1990.

[Montgomery 1971] H. L. Montgomery, *Topics in multiplicative number theory*, Lecture Notes in Math. **227**, Springer, 1971.

[Nelson 1978–79] H. Nelson, "Problem 654", *J. Rec. Math.* **11** (1978–79), 231.

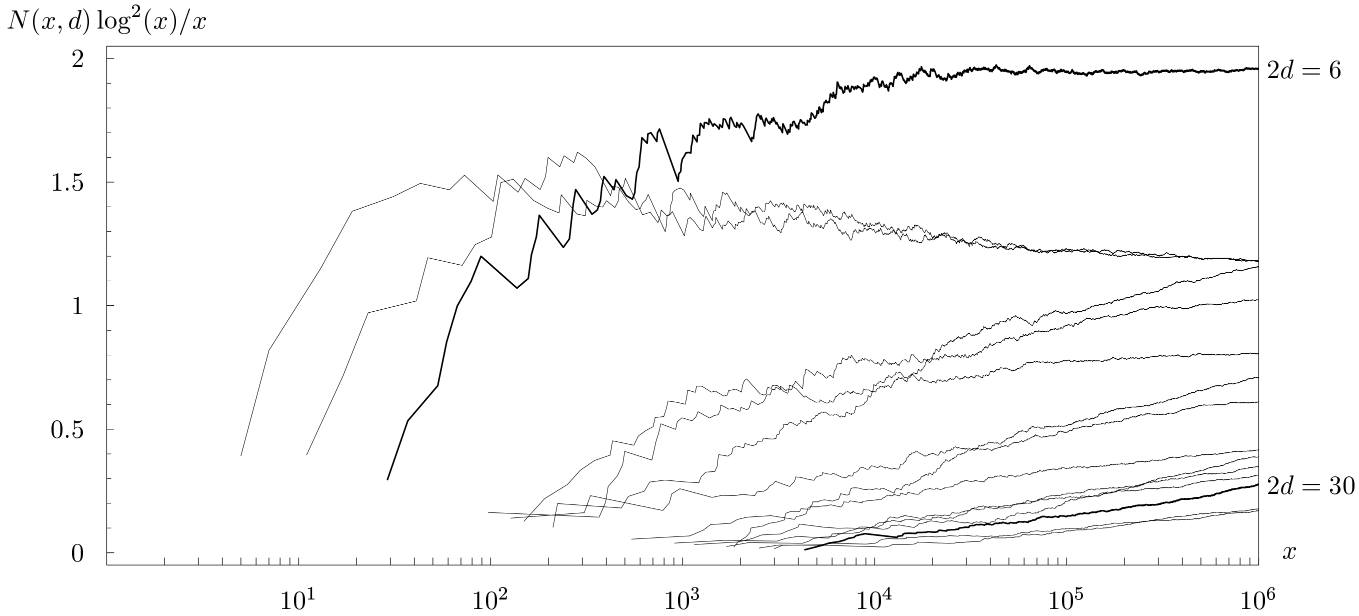

FIGURE 3. $N(x,d)\log^2(x)/x$ as a function of x, for $2d = 2, 4, \dots, 30$ (in order or increasing x-coordinate of starting point: $2d$ = 2, 4, 6, 8, 14, 10, 12, 18, 20, 22, 24, 16, 26, 28, 30; in order of decreasing y-coordinate along the right-hand edge: $2d$ = 6, 2, 4, 12, 10, 8, 18, 14, 16, 24, 20, 22, 30, 28, 26). The two curves in bold represent $2d = 6$ and $2d = 30$.

d	$N(10^{12},d)$	(2–6), $M=4$	(2–8)	d	$N(10^{12},d)$	(2–6), $M=4$	(2–8)
1	1870585221	1870559866.	1734571973.	26	299020127	19357608.	287761502.
2	1870585458	1870559866.	1608489045.	27	511589763	−117485659.	489342519.
3	3435528229	3435458600.	2983176210.	28	276101593	−190236598.	272337270.
4	1573331564	1573293311.	1383199071.	29	238482555	−159446866.	218306665.
5	2052293026	2052377278.	1710267841.	30	521616486	−872270696.	520705710.
6	2753597777	2753698149.	2379035785.	31	173395125	−542475987.	187370709.
7	1556469349	1556538305.	1323739864.	32	174696822	−466395227.	168010801.
8	1202533145	1202481778.	1023002316.	33	337881160	−1472349367.	346327794.
9	2246576317	2246300116.	1897433561.	34	144475047	−901708546.	154203810.
10	1298682892	1297504207.	1173113388.	35	209257685	−1446734637.	214563934.
11	1105634145	1104842257.	906625819.	36	225244356	−2345640221.	248794573.
12	1754011594	1748689938.	1513472556.	37	112410088	−1279821387.	118692508.
13	866077378	860228350.	765617165.	38	103953673	−1562442677.	113342851.
14	946685406	940272873.	781065469.	39	202872036	−3480363786.	216657899.
15	1803413614	1768917778.	1609765148.	40	109107891	−2536053455.	122824166.
16	596278790	571983719.	559868265.	41	79287666	−2097549341.	87646234.
17	629634308	602935653.	553874113.	42	169541709	−5569989899.	190259148.
18	1069300358	994461819.	963192792.	43	63992940	−2740157702.	75335519.
19	520188423	469051756.	472946539.	44	67022921	−3106662564.	75804586.
20	626694626	549365467.	552378496.	45	141957467	−8653244845.	168777258.
21	979052296	757589403.	922195739.	46	49878328	−3851360864.	61511925.
22	414087760	277381704.	395992947.	47	46375798	−3982359526.	55682088.
23	366906343	217998577.	346302520.	48	83989444	−8412724248.	101068993.
24	651790197	305395231.	613209321.	49	45681754	−5553974513.	56258792.
25	386726111	71637118.	379182356.	50	48416676	−6460114606.	57992596.

TABLE 2. A comparison of two estimates for $N(10^{12},d)$.

$2d$	$g_{2d}(10^k)$ for $k =$				$2d$	$g_{2d}(10^k)$ for $k =$				$2d$	$g_{2d}(10^k)$ for $k =$			
	30	40	400	450		30	40	400	450		30	40	400	450
2	2769	1539	25	11	82	932	654	12	12	162	509	582	27	16
4	2772	1473	20	29	84	1982	1582	44	20	164	264	297	16	13
6	5278	3120	32	26	86	882	674	18	12	166	252	247	11	15
8	2630	1520	17	13	88	835	652	12	8	168	619	622	44	24
10	3462	1998	15	19	90	2119	1664	49	29	170	328	389	8	8
12	5016	2761	37	28	92	769	634	13	11	172	247	235	13	12
14	2900	1644	19	18	94	813	609	12	12	174	466	550	28	25
16	2392	1397	20	13	96	1452	1101	19	22	176	242	239	17	11
18	4578	2681	27	17	98	804	648	25	14	178	225	233	9	15
20	2866	1760	23	11	100	916	706	27	16	180	526	641	27	26
22	2450	1460	14	22	102	1392	1118	27	18	182	255	273	20	16
24	4305	2544	25	24	104	692	559	10	9	184	205	211	10	7
26	2241	1315	23	12	106	672	532	21	9	186	372	386	29	19
28	2410	1472	21	10	108	1207	1047	33	15	188	180	206	11	10
30	5060	3209	50	26	110	884	705	18	21	190	240	279	13	16
32	1828	1217	17	13	112	707	631	15	19	192	342	413	22	16
34	1938	1257	18	9	114	1145	967	24	15	194	161	186	11	10
36	3518	2268	19	22	116	567	432	16	14	196	215	243	13	9
38	1758	1129	17	15	118	512	471	22	4	198	323	423	33	40
40	2260	1397	20	19	120	1285	1162	40	30	200	207	234	13	16
42	3718	2536	25	24	122	447	439	17	12	202	130	154	7	6
44	1798	1124	13	13	124	463	436	13	5	204	305	354	36	21
46	1655	1066	6	14	126	1051	1011	28	22	206	151	170	12	5
48	2919	1974	32	21	128	466	408	8	6	208	152	190	11	16
50	1968	1255	18	12	130	647	595	14	17	210	438	512	33	34
52	1475	1068	19	9	132	892	831	25	23	212	112	159	8	7
54	2748	1826	23	21	134	380	367	11	10	214	121	155	14	12
56	1557	1051	18	14	136	406	361	10	15	216	212	301	27	25
58	1312	924	11	11	138	765	802	18	16	218	99	149	10	6
60	3305	2269	38	30	140	598	543	19	17	220	173	208	24	14
62	1270	825	15	8	142	369	345	10	8	222	222	300	14	24
64	1214	863	13	8	144	662	664	32	16	224	139	156	14	14
66	2588	1739	29	17	146	333	318	9	9	226	113	131	11	11
68	1107	816	8	14	148	336	361	15	15	228	216	292	28	27
70	1658	1231	21	18	150	876	833	37	29	230	129	169	19	11
72	2008	1456	25	25	152	311	332	15	15	232	95	103	15	4
74	984	785	14	13	154	398	418	20	13	234	184	251	20	20
76	1036	777	13	8	156	650	629	26	23	236	87	116	14	15
78	2130	1588	28	25	158	286	302	10	12	238	97	160	14	10
80	1238	940	25	15	160	364	369	17	12	240	211	276	33	23

TABLE 3. Gap histograms. The columns show the presumed value of $g_{2d}(x) = N(x+10^7, d) - N(x, d)$ (as calculated by Maple's probabilistic prime function), for values of $2d$ in the range $2 \le 2d \le 240$.

Andrew Odlyzko, AT&T Labs – Research, 180 Park Avenue, Florham Park, NJ 07932, United States (amo@research.att.com)

Michael Rubinstein, Department of Mathematics, University of Texas, Austin, TX, 78712, United States (miker@math.utexas.edu)

Marek Wolf, Institute of Theoretical Physics, University of Wroclaw, pl. M. Brona 9, Wroclaw, PL-50-204, Poland (mwolf@proton.ift.uni.wroc.pl)

Received February 2, 1998; accepted July 6, 1998