On the number of ways of writing t as a product of factorials

Daniel M. Kane

December 31, 2005

Abstract

Let $\mathbb{N}_0$ denote the set of non-negative integers. In this paper we prove that

$$\limsup_{t\to\infty} \left|\left\{(n,m) \in \mathbb{N}_0^2 : n!m! = t\right\}\right| = 6.$$

1 Introduction

Let $\mathbb{N}_0$ denote the set of non-negative integers. In this paper we will prove that

$$\limsup_{t\to\infty} \left|\left\{(n,m) \in \mathbb{N}_0^2 : n!m! = t\right\}\right| = 6.$$

We use three techniques to prove this result. First, it is not difficult to generate an infinite set of t each of which has at least 6 representations as a product of factorials thus establishing the lower bound. We then use considerations of the number of times two divides t in order to show that all of the solutions must be near each other. Lastly we use some analytic techniques analogous to those in [1].

The following three conjectures also seem likely.

Conjecture 1.

$$\max \left|\left\{(n,m) \in \mathbb{N}_0^2 : n!m! = t\right\}\right| = 6.$$

Conjecture 2.

$$\limsup_{t\to\infty} \left|\left\{(n,m) \in \mathbb{N}^2 : n!m! = t\right\}\right| = 4.$$

Conjecture 3.

$$\max \left|\left\{(n,m) \in \mathbb{N}^2 : n!m! = t\right\}\right| = 4.$$

It is true that conjecture 3 would imply both other conjectures, and that any of these conjectures is stronger than our main theorem.

2 The lower bound

Notice that for any integer $n > 2$, we have that

$(n!)! = 0! \cdot (n!)! = 1! \cdot (n!)! = n! \cdot (n! - 1)! = (n! - 1)! \cdot n! = (n!)! \cdot 1! = (n!)! \cdot 0!.$

Therefore, we have that

$$\limsup_{t\to\infty} \left|\left\{(n,m) \in \mathbb{N}_0^2 : n!m! = t\right\}\right| \geq 6.$$

3 The first technique

For a positive integer n, let $e(n)$ denote the largest k so that 2^k divides n. Notice that

$$\begin{aligned} e(n!) &= \sum_{i=1}^{\lfloor \log_2(n) \rfloor} \left\lfloor \frac{n}{2^i} \right\rfloor \\ &= \sum_{i=1}^{\lfloor \log_2(n) \rfloor} \frac{n}{2^i} - \sum_{i=1}^{\lfloor \log_2(n) \rfloor} O(1) \\ &= n + O(\log n). \end{aligned}$$

Therefore we have that if $n!m! = t$, then $e(n!) + e(m!) = e(t)$, and therefore, $n + m + O(\log n + \log m) = e(t)$. Since $(n/2)^{n/2} < n! < t$, we have that $n < \log t$ for sufficiently large t. Therefore, for sufficiently large t, $n + m + O(\log\log t) = e(t)$. Hence if $n_1!m_1! = n_2!m_2! = t$, then $n_1 + m_1 = n_2 + m_2 + O(\log\log t)$. This fact provides an elementary proof that for fixed t the number of solutions to $n!m! = t$ is $O(\log\log t)$ because by convexity of the log of the factorial function, at most two solutions to $n!m! = t$ have a given sum of $n + m$, and this sum cannot vary by more than $O(\log\log t)$.

4 The second technique

Our second technique is similar to that used in [1]. We begin with the following lemma:

Lemma 1. *If $F(x) : \mathbb{R} \to \mathbb{R}$ is an infinitely differentiable function and if $F(x) = 0$ for $x = x_1, x_2, ..., x_{n+1}$ (where $x_1 < x_2 < ... < x_{n+1}$), then $F^{(n)}(y) = 0$ for some $y \in (x_1, x_{n+1})$.*

Proof. We proceed by induction on n. The case of $n = 1$ is Rolle's Theorem. Given the statement of Lemma 2.1 for $n - 1$, if there exists such an F with $n + 1$ zeroes, $x_1 < x_2 < ... < x_{n+1}$, then by Rolle's theorem, there exist points $y_i \in (x_i, x_{i+1})$ $(1 \leq i \leq n)$ so that $F'(y_i) = 0$. Then since F' has at least n roots, by the induction hypothesis there exists a y with $x_1 < y_1 < y < y_n < x_{n+1}$, and $F^{(n)}(y) = (F')^{(n-1)}(y) = 0$. □

We now state a lemma that helps us to count the number of integer points on smooth curves.

Lemma 2. *Let $f : \mathbb{R} \to \mathbb{R}$ be a C^k function. Suppose that for $x \in (a, b)$, that*

$$0 < \left| \frac{1}{k!} \frac{\partial^k}{\partial x^k} f(x) \right| < \alpha.$$

Then if we have $a < x_0 < x_1 < ... < x_k < b$ where $x_i \in \mathbb{Z}$ and $f(x_i) \in \mathbb{Z}$ for all $0 \leq i \leq k$, then $x_k - x_0 \geq \alpha^{\frac{-2}{k(k+1)}}$.

Proof. Let

$$g(x) = \sum_{i=0}^{k} f(x_i) \prod_{\substack{0 \leq j \leq k \\ i \neq j}} \frac{x - x_j}{x_i - x_j}$$

be the polynomial of degree k that interpolates f at the x_i. Let $h(x) = f(x) - g(x)$. Then $h(x_i) = 0$. Hence by Lemma 1 we have that for some $a < x_0 < y < x_k < b$ that $\frac{\partial^k}{\partial x^k} h(y) = 0$. Or that

$$\left(\frac{1}{k!} \frac{\partial^k}{\partial x^k} f(x) \right)_{x=y} = \left(\frac{1}{k!} \frac{\partial^k}{\partial x^k} g(x) \right)_{x=y} = \sum_{i=0}^{k} f(x_i) \prod_{\substack{0 \leq j \leq k \\ j \neq i}} \frac{1}{x_i - x_j}.$$

Therefore,

$$s = \left(\frac{1}{k!} \frac{\partial^k}{\partial x^k} f(x) \right)_{x=y}$$

is an integer multiple of $M = \prod_{0 \leq i < j \leq k} \frac{1}{x_j - x_i}$. Therefore, either $s = 0$ or else $|s| \geq M$. But by assumption, $0 < |s| < \alpha$. Therefore, $\alpha \geq |s| \geq M$. Hence $\alpha \geq (x_k - x_0)^{\frac{-k(k+1)}{2}}$, and hence we have that $\alpha^{\frac{-2}{k(k+1)}} \leq x_k - x_0$ as desired. □

We will also make use of a generalization of Stirling's formula which states that:

$$\log(\Gamma(z + 1)) = (z + \frac{1}{2}) \log(z) - z + \frac{1}{2} \log(2\pi) + O(z^{-1})$$

uniformly for $\Re(z) > 0$. This follows readily from the $m = 2$ case of

$$\begin{aligned} \log \Gamma(z + 1) \quad &= \quad \frac{1}{2} \log(2\pi) + \left(z + \frac{1}{2} \right) \log(z) - z \\ &+ \sum_{j=1}^{m} \frac{B_{2j}}{(2j - 1)(2j) z^{2j-1}} - \frac{1}{2m} \int_0^\infty \frac{B_{2m}(x - [x])}{(x + z)^{2m}} dx. \end{aligned}$$

where B_{2j} and B_{2m} are the Bernoulli numbers and Bernoulli polynomials (see [2]).

5 The Strategy

We have yet to prove that for sufficiently large t

$$\left|\left\{(n,m)\in\mathbb{N}_0^2 : n!m! = t\right\}\right| \leq 6.$$

It is sufficient to show that $\left|\left\{(n,m)\in\mathbb{N}_0^2 : n\geq m, n!m! = t\right\}\right| \leq 3$ for all sufficiently large t. We will split solutions of this form into three overlapping cases:
1. $m < \exp(2\sqrt{\log\log t})$
2. $m > \exp(\sqrt{\log\log t}), n-m > (\log t)^{25/36}$
3. $n-m < (\log t)^{26/36}$

Furthermore, we will show by our results from sections 3 and 4, that for all sufficiently large t, that all integer solutions to $n!m! = t$ lie in one of these regions.

Define the function $f:\mathbb{R}\to\mathbb{R}$ implicitly by $\Gamma(f(x)+1)\Gamma(x+1) = t$. It is clear that

$$f'(x) = -\frac{g(x)}{g(f(x))}$$

where

$$g(x) = \frac{\partial}{\partial x}\log\Gamma(x+1) = \log(x) + O(x^{-1})$$

(by our strong form of Stirling's formula). So

$$f'(x) = -\frac{\log x + O(1)}{\log(f(x))}.$$

So if we have two pairs of solutions $(n_1,m_1),(n_2,m_2)$ to $n\geq m, n!m! = t$, where $m_2 > m_1$, then

$$O(\log\log t) > n_2 + m_2 - (n_1+m_1) = \int_{m_1}^{m_2} 1 + f'(x)dx.$$

We need to show that if there are solutions with m too big for region 1, there are none with m too small for region 2, and that if there are solutions with m too small for region 3, there are none with m too big for region 2.

We can show the first of these by verifying that for sufficiently large t

$$\begin{aligned}\int_{\exp(\sqrt{\log\log t})}^{\exp(2\sqrt{\log\log t})} 1 + f'(x)dx \quad &> \\ \int_{\exp(\sqrt{\log\log t})}^{\exp(2\sqrt{\log\log t})} \frac{\log t - 4\sqrt{\log\log t} + O(1)}{\log t}dx \quad &= \\ \exp(2\sqrt{\log\log t}) + O(1) \quad &\gg \quad \log\log t.\end{aligned}$$

This shows that for sufficiently large t, it is impossible to have two solutions, one of which has m too small to be in region 2, and the other of which has m too large to be in region 1 since this would imply that $n_1+m_1-(n_2+m_2) \gg \log\log t$.

Now if x_1 and x_2 are the numbers so that $f(x_2) - x_2 = (\log t)^{25/36}$ and $f(x_1) - x_1 = (\log t)^{26/36}$, we notice that since the log of the gamma-function is convex that $x_2 - x_1 > \frac{1}{3}(\log t)^{26/36}$. Thus we verify the second of these statements by noticing that

$$\begin{aligned}
\int_{x_1}^{x_2} 1 + f'(x)dx &> \\
\frac{1}{3}(\log t)^{26/36}(1 + f'(x_2)) &> \\
\frac{1}{3}(\log t)^{26/36}\frac{\log(f(x_2)/x_2) + O(x_2^{-1})}{\log(f(x_2))} &> \\
\Omega\left((\log t)^{26/36}\frac{(f(x_2) - x_2)/x_2}{\log(f(x_2))}\right) &= \\
\Omega\left(\frac{(\log t)^{15/36}}{\log\log t}\right) &\gg \\
\log\log t.&
\end{aligned}$$

Recall that $a(t) = \Omega(b(t))$ means that there exists a constant $c > 0$ so that for all sufficiently large t, $a(t) > cb(t)$, and that $a(t) = \Theta(b(t))$ means that there exist $c_1 > 0$ and $c_2 > 0$ so that for all sufficiently large t, $c_1 a(t) > b(t) > c_2 a(t)$.

In section 6, we will cover the case where there are solutions in the first region. In section 7, we will cover the case where there are solutions in the second region. In section 8, we will cover the case where there are solutions in the third region.

6 The First Region

In this section, we will prove that for sufficiently large t, that there are at most 2 solutions to $n!m! = t$ with $0 < m \leq \exp(\sqrt{\log\log t})$.

Notice that

$$\begin{aligned}
e\left(\frac{(n+x)!}{n!}\right) &= e((n+x)!) - e(n!) \\
&= \sum_{i=1}^{\infty} \left\lfloor \frac{n+x}{2^i} \right\rfloor - \left\lfloor \frac{n}{2^i} \right\rfloor \\
&= \sum_{i=1}^{\lfloor \log x \rfloor} \frac{x}{2^i} + O(1) + \max_{n<c\leq n+x} e(c) - \log x \\
&= x + \max_{n<c\leq n+x} e(c) + O(\log x).
\end{aligned}$$

Therefore, if we have any two such solutions, $n_1!m_1! = n_2!m_2! = t$, with $n_1 > n_2$ then $e\left(\frac{n_1!}{n_2!}\right) = e\left(\frac{m_2!}{m_1!}\right)$. Therefore,

$$n_1 - n_2 + O(\log(n_1 - n_2)) + \max_{n_1<c\leq n_2} e(c) \geq m_2 - m_1 + O(\log(m_2 - m_1)).$$

Which implies that

$$\max_{n_1 < c \le n_2} e(c) > (m_2 - m_1) - (n_1 - n_2) + O(\log(m_2 - m_1)).$$

Notice that if $n_1!m_1! = n_2!m_2!$, then $\frac{n_1!}{n_2!} = \frac{m_2!}{m_1!}$, and therefore, $m_2 - m_1 > (n_1 - n_2) \cdot \frac{\log n_2}{\log m_2}$. Now, $n_2 \log n_2 > \log t - 2\sqrt{\log\log t}\exp(2\sqrt{\log\log t})$. Therefore, $n_2 = \Omega(\frac{\log t}{\log\log t})$, so $\log n_2 = \Omega(\log\log t)$. Hence $m_2 - m_1 = \Omega(\sqrt{\log\log t})(n_1 - n_2)$. Therefore, we have that

$$\begin{aligned}
\max_{n_1 < c \le n_2} e(c) &> (m_2 - m_1)(1 + O((\log\log t)^{-1/2})) + O(\log(m_2 - m_1)) \\
&= \Omega(m_2 - m_1) \\
&= \Omega(\sqrt{\log\log t}).
\end{aligned}$$

Therefore, if we have three solutions in region 1, $(n_1, m_1), (n_2, m_2), (n_3, m_3)$ with $0 < m_1 < m_2 < m_3$, then we have that there exist $n_3 < c_1 \le n_2 < c_2 \le n_1$ with $e(c_i) = \Omega(\sqrt{\log\log t})$. Therefore, since $\min(e(x), e(y)) \le e(x - y)$, we have that $e(c_2 - c_1) = \Omega(\sqrt{\log\log t})$. Therefore, $n_1 - n_3 > c_2 - c_1 > \exp(\Omega(\sqrt{\log\log t}))$. But we notice that this and previous inequalities imply that

$$m_3 + n_3 - (m_1 + n_1) = \Omega(\sqrt{\log\log t})\exp(\Omega(\sqrt{\log\log t})).$$

Since this cannot be $O(\log\log t)$, we have that for sufficiently large t, there are at most 2 solutions with $m \ne 0$ in region 1. Hence there are at most 3 solutions in region 1.

7 The Second Region

In this section we will show that there are at most 2 solutions with $m > \exp(\sqrt{\log\log t})$ and $n - m > (\log t)^{25/36}$. Recall that $f : \mathbb{R} \to \mathbb{R}$ so that $\Gamma(f(x) + 1)\Gamma(x + 1) = t$. This is defined in the range we are interested in, because the gamma-function is increasing. Let $g(x) = \log\Gamma(x + 1)$. So $g(f(x)) + g(x) = \log t$. Differentiating implicitly, we get that

$$f'(x) = -\frac{g'(x)}{g'(f(x))}.$$

Therefore,

$$\begin{aligned}
f''(x) &= -\frac{g''(x)}{g'(f(x))} + f'(x)\frac{g'(x)g''(f(x))}{(g'(f(x)))^2} \\
&= -\frac{g''(x)(g'(f(x)))^2 + (g'(x))^2 g''(f(x))}{(g'(f(x)))^3}.
\end{aligned}$$

By differentiating our strong form of Stirling's formula, we find that for $f(x) > x$

$$f''(x) = -\frac{(\log f(x))^2 x^{-1} + (\log x)^2 (f(x))^{-1}(1 + O(x^{-1}))}{((\log f(x)) + O(f(x)^{-1}))^3}.$$

Therefore, for all sufficiently large t, for $x > \exp(\sqrt{\log\log t})$ and $f(x) - x > (\log t)^{25/36}$ we have that

$$0 < \left|\frac{1}{2}f''(x)\right| < O\left(\frac{1}{x(\log f(x))}\right).$$

Assume for sake of contradiction that we have three solutions to $n!m! = t$ in region 2 ($n_i!m_i! = t$ for $1 \leq i \leq 3$ where $m_i < m_{i+1}$). Then we have that m_i is an integer and that $f(m_i)$ is an integer. Since between m_1 and m_3 we have that

$$0 < \left|\frac{1}{2}f''(x)\right| < O\left(\frac{1}{m_1}\right).$$

Therefore, by Lemma 2, $m_3 - m_1 > \Omega(m_1^{1/3})$. But we also have that $(m_3 + n_3) - (m_1 + n_1) = O(\log\log t)$. Therefore

$$\int_{m_1}^{m_1+\Omega(m_1^{1/3})} \frac{(\log f(x)) - \log x}{\log f(x)} dx = O(\log\log t).$$

Now we have that for x in the range we are concerned with that

$$\frac{(\log f(x)) - \log x}{\log f(x)} = \Omega\left(\frac{(f(x) - x)/x}{\log f(x)}\right) = \Omega((\log t)^{-11/36}).$$

Therefore, it must be that $m_1 = O((\log\log t)^3 (\log t)^{11/12})$. But in this range, the integrand we are concerned with is at least $\frac{1}{12} + o(1)$. Therefore, it must be that $m_1 = O((\log\log t)^3)$ which does not hold. Therefore, for sufficiently large t, there are at most 2 solutions in region 2.

8 Region Three

In this section we will show that there are at most 3 solutions in region 3 for sufficiently large t. This proof depends on the fact that if n and m are integers, then so are $n + m$ and $(n - m)^2$ and applications of Lemma 2 and results from section 3.

Suppose that $\Gamma((a + \sqrt{x} + 2)/2)\Gamma((a - \sqrt{x} + 2)/2) = t$, where $x = O(a^{3/2})$. Then we have by our strong form of Stirling's formula that

$$\begin{aligned}
\log t &= \frac{a + \sqrt{x} + 1}{2}\log\left(\frac{a + \sqrt{x}}{2}\right) + \frac{a - \sqrt{x} + 1}{2}\log\left(\frac{a - \sqrt{x}}{2}\right) \\
&\quad -a + \log 2\pi + O(a^{-1}) \\
&= (a + 1)\log(a/2) - a + \log(2\pi) \\
&\quad + \frac{a + \sqrt{x} + 1}{2}\log\left(1 + \frac{\sqrt{x}}{a}\right) + \frac{a - \sqrt{x} + 1}{2}\log\left(1 - \frac{\sqrt{x}}{a}\right) + O(a^{-1}) \\
&= (a + 1)\log(a/2) - a + \log(2\pi) + \frac{x}{a} + O(a^{-1}).
\end{aligned}$$

Therefore, we have that

$$x = a \log\left(\frac{t}{2\pi}\right) - a(a+1)\log\left(\frac{a}{2}\right) + a^2 + O(1).$$

Let a_0 be the positive real value so that $(\Gamma((a_0/2)+1))^2 = t$. So then we have that

$$a_0 \log\left(\frac{t}{2\pi}\right) - a_0(a_0+1)\log\left(\frac{a_0}{2}\right) + a_0^2 = O(1).$$

It is also true that $a_0 = O(\log t)$. If we pick an a so that

$$a \log\left(\frac{t}{2\pi}\right) - a(a+1)\log\left(\frac{a}{2}\right) + a^2 + O(1) < (\log t)^{16/11}.$$

Then there must be a unique complex x with $|x| \leq (\log t)^{16/11}$ so that $\Gamma((a+\sqrt{x}+2)/2)\Gamma((a-\sqrt{x}+2)/2) = t$. Since the derivative of

$$a \log\left(\frac{t}{2\pi}\right) - a(a+1)\log\left(\frac{a}{2}\right) + a^2$$

is

$$\log\left(\frac{t}{2\pi}\right) - (2a+1)\log\left(\frac{a}{2}\right) + a - 1$$

and since its second derivative is $O(\log a)$, this should hold as long as $|a - a_0| < O((\log t)^{9/20})$. Furthermore, for a in this range, x attains all values with $|x| \leq (\log t)^{13/9}$. This allows use to define an analytic function $h(a)$ defined on $|a - a_0| < O((\log t)^{29/20})$ so that

$$\Gamma\left(\frac{a+\sqrt{h(a)}}{2}+1\right)\Gamma\left(\frac{a-\sqrt{h(a)}}{2}+1\right) = t.$$

Furthermore, $h(a)$ attains all values of absolute value at most $(\log t)^{13/9}$ when $|a-a_0| = O(\log t)^{5/9}$. Additionally, we have that

$$h(a) = a \log\left(\frac{t}{2\pi}\right) - a(a+1)\log\left(\frac{a}{2}\right) + a^2 + O(1).$$

Since the $O(1)$ is uniform in the region stated, its third derivative when $|a-a_0| = O(\log t)^{13/29}$ (notice that $4/9 < 13/29 < 9/20$) can be written using Cauchy's Integral formula as

$$\int_C O((z-a)^{-4})dz,$$

where C is a contour that traverses a circle centered at a of radius $\Omega((\log t)^{9/20})$ once in the counter-clockwise direction. This is $O((\log t)^{-27/20})$. Therefore, when $|h(a)| < (\log t)^{13/9}$ we have that

$$\begin{aligned} h'''(a) &= \frac{\partial^3}{\partial a^3}\left(a \log\left(\frac{t}{2\pi}\right) - a(a+1)\log\left(\frac{a}{2}\right) + a^2\right) + O((\log t)^{-27/20}) \\ &= -\frac{1}{a} + O((\log t)^{-27/20}). \end{aligned}$$

Since it is clear by Stirling's formula that $a_0 = \Theta\left(\frac{\log t}{\log\log t}\right)$ we have that for sufficiently large t, when $|h(a)| \leq (\log t)^{13/9}$ then $0 < |h'''(a)| < O\left(\frac{\log\log t}{\log t}\right)$.

Now suppose for sake of contradiction that (n_i, m_i) are distinct region 3 solutions for $1 \leq i \leq 4$. Then $n_i + m_i \in \mathbb{Z}$ and $h(n_i + m_i) = (n_i - m_i)^2 \in \mathbb{Z}$. Furthermore, $|h(n_i + m_i)| \leq (\log t)^{13/9}$. Then since in the range between the $n_i + m_i$ we have that $0 < |h'''(a)| < O\left(\frac{\log\log t}{\log t}\right)$, Lemma 2 implies that the difference between the largest and smallest of the $n_i + m_i$ is at least

$$\Omega\left(\left(\frac{\log t}{\log\log t}\right)^{1/6}\right).$$

Since this is larger than $O(\log\log t)$, we have from results in section 3, that for sufficiently large t, this is impossible.

Hence there are at most three solutions in region 3 for sufficiently large t.

9 Conclusions

Hence we have proved our result that

$$\limsup_{t\to\infty} \left|\left\{(n,m) \in \mathbb{N}_0^2 : n!m! = t\right\}\right| = 6.$$

Notice that all of our statements about there being at most 6 solutions for "sufficiently large t" can be made effective, although this was not done in this paper. I do not believe that the effective bound that is achieved would be small enough to allow for a reasonable proof that there are at most 6 solutions for any t, at least without further insight.

References

[1] D. Kane, *New bounds on the number of representations of t as a binomial coefficient*, Integers 4 (2004).

[2] Hans Rademacher, *Topics in Analytic Number Theory* Springer-Verlag, Berlin-Heidelberg-New York 1970.