WORKING NOTES: BOUNDS ON FOURIER COEFFICIENTS ARISING IN BENFORD'S LAW IN IMAGES

C. T. ABDALLAH, G. L. HEILEMAN, T. HILL, S. J. MILLER, AND F. PÉREZ-GONZÁLEZ

ABSTRACT. In studying applications of Benford's Law to images, the size of certain coefficients of Fourier expansions play a central role. In particular, we need to bound

$$b_n \ = \ \prod_{k=0}^{\infty}\left(1+\frac{\alpha_n^2}{(ck+1)^2}\right)^{-1}, \quad \alpha_n \ = \ \frac{2\pi n}{\log 10}.$$

If $\frac{c+1}{\alpha_n} < 1$ then below we show that

$$b_n \ \le \ \frac{\log^2 10}{\log^2 10+(2\pi n)^2}\cdot\exp\left(\frac{c+1}{c}\right)\cdot\exp\left(-\frac{\pi^2 n}{c\log 10}\right).$$

Better bounds are obtainable with only slightly more work. It is worth noting that we have lots of decay; the coefficients are falling off at least as fast as $1/n^2$ as well as exponentially with n. We show

$$\sum_{n=2}^{\infty} b_n \quad \le \quad g(c),$$

where

$$g(c) \ = \ 0.0421603\cdot\exp\left(\frac{c+1}{c}\right)\cdot\exp\left(-\frac{2\pi^2}{c\log 10}\right)\cdot\left[1-\exp\left(-\frac{2\pi^2}{c\log 10}\right)\right]^{-1/2}.$$

In particular, for $c = 1$ we obtain a bound of $5.24\cdot 10^{-5}$, and for $c = 2$ a bound of .00232.

1. ESTIMATES FOR b_n

In previous work, the modulus squared of the Fourier coefficients

$$b_n \ = \ |a_n|^2 \ = \ \prod_{k=0}^{\infty}\left(1+\frac{\alpha_n^2}{(ck+1)^2}\right)^{-1}, \quad \alpha_n \ = \ \frac{2\pi n}{\log 10} \tag{1.1}$$

were encountered. It is desirable to obtain estimates on how rapidly b_n decays with n. We provide some below, concentrating on ease of exposition rather than best possible (though of course the bounds can be improved a bit if needed).

Date: January 1, 2008.
2000 *Mathematics Subject Classification.* 11K06, 60A10 (primary), (secondary).
Key words and phrases. Benford's Law, Images.

We have

$$\begin{aligned} b_n &= \prod_{k=0}^{\infty} \left(\frac{(ck+1)^2 + \alpha_n^2}{(ck+1)^2} \right)^{-1} \\ &= \prod_{k=0}^{\infty} \left(\frac{(ck+1)^2}{(ck+1)^2 + \alpha_n^2} \right) \\ &= \frac{1}{1+\alpha_n^2} \prod_{k=1}^{\infty} \left(1 - \frac{\alpha_n^2}{(ck+1)^2 + \alpha_n^2} \right) \\ \log b_n &= -\log(1+\alpha_n^2) + \sum_{k=1}^{\infty} \log\left(1 - \frac{\alpha_n^2}{(ck+1)^2 + \alpha_n^2} \right). \end{aligned} \tag{1.2}$$

Using

$$\log(1-u) \ = \ -\sum_{\ell=1}^{\infty} \frac{u^\ell}{\ell}, \tag{1.3}$$

we find

$$\log b_n \ = \ -\log(1+\alpha_n^2) - \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \frac{1}{\ell} \left(\frac{\alpha_n^2}{(ck+1)^2 + \alpha_n^2} \right)^{\ell}. \tag{1.4}$$

Thus we obtain an upper bound for $\log b_n$ (and hence an upper bound for b_n) by only keeping the $\ell = 1$ term above (each summand is positive, but is hit by a negative sign). Thus

$$\log b_n \ \leq \ -\log(1+\alpha_n^2) - \sum_{k=1}^{\infty} \frac{\alpha_n^2}{(ck+1)^2 + \alpha_n^2}. \tag{1.5}$$

We argue (somewhat) crudely below; a better estimate is obtainable by using the Euler-MacLaurin formula. We have

$$\sum_{k=1}^{\infty} \frac{\alpha_n^2}{(ck+1)^2 + \alpha_n^2} \ \geq \ \int_{x=1}^{\infty} \frac{\alpha_n^2 dx}{(cx+1)^2 + \alpha_n^2}; \tag{1.6}$$

the reason for this is that the integrand is monotonically decreasing, and the sum is basically the upper sum approximation. Because of the minus sign, we thus increase the bound on $\log b_n$ if we replace the sum by this integral (a better estimate would be to pull off the first term and then use the integral switch, but let us see what this yields). We find

$$\begin{aligned} \log b_n &\leq -\log(1+\alpha_n^2) - \int_{x=1}^{\infty} \frac{\alpha_n^2 dx}{(cx+1)^2 + \alpha_n^2} \\ &= -\log(1+\alpha_n^2) - \int_{x=1}^{\infty} \frac{dx}{1 + \left(\frac{cx+1}{\alpha_n}\right)^2}. \end{aligned} \tag{1.7}$$

We change variables. Let $u = (cx+1)/\alpha_n$ (so $dx = \frac{\alpha_n}{c} du$). Thus

$$\log b_n \ \leq \ -\log(1+\alpha_n^2) - \frac{\alpha_n}{c} \int_{\frac{c+1}{\alpha_n}}^{\infty} \frac{du}{1+u^2}. \tag{1.8}$$

As the anti-derivative of $(1+u^2)^{-1}$ is $\arctan(u)$ and $\arctan(\infty) = \pi/2$, we find

$$\begin{aligned}\log b_n &\leq -\log(1+\alpha_n^2) - \frac{\alpha_n}{c}\left(\arctan(\infty) - \arctan\left(\frac{c+1}{\alpha_n}\right)\right)\\ &= -\log(1+\alpha_n^2) - \frac{\alpha_n \pi}{2c} + \frac{\alpha_n}{c}\arctan\left(\frac{c+1}{\alpha_n}\right). \qquad (1.9)\end{aligned}$$

Thus we are left with estimating the remaining arc-tangent. The Taylor series of arc-tangent is

$$\arctan(x) = \sum_{\ell=0}^{\infty} \frac{(-1)^\ell x^{2\ell+1}}{2\ell+1}. \qquad (1.10)$$

WE ASSUME FROM NOW ON THAT $\frac{c+1}{\alpha_n} < 1$. As we are only concerned with $n \geq 2$, this is a weak condition, and holds whenever $c < 4.45$. Using $0 < x < 1$ implies $0 < \arctan(x) < x$ (this follows because we have an alternating sum of terms which decrease in absolute value), we have

$$\log b_n \leq -\log(1+\alpha_n^2) - \frac{\alpha_n \pi}{2c} + \frac{c+1}{c}. \qquad (1.11)$$

Exponentiating yields

$$b_n \leq \frac{1}{1+\alpha_n^2} \cdot \exp\left(-\frac{\alpha_n \pi}{2c}\right) \cdot \exp\left(\frac{c+1}{c}\right). \qquad (1.12)$$

While we could obtain better estimates with a little more work, already this estimate is nice. Plugging in the value for α_n yields

$$b_n \leq \frac{\log^2 10}{\log^2 10 + (2\pi n)^2} \cdot \exp\left(\frac{c+1}{c}\right) \cdot \exp\left(-\frac{\pi^2 n}{c \log 10}\right). \qquad (1.13)$$

From this it should be easy to get great estimates on $\sum_n b_n$. We have *plenty* of decay to exploit. One natural way to proceed is to use the Cauchy-Schwartz inequality (although perhaps it is better to use a Holder-type inequality and optimize the exponent). Below we assume $c_n, \gamma_n \geq 0$, as this is what happens in our applications. The Cauchy-Schwarz inequality states that

$$\left|\sum_{n=2}^{\infty} c_n \gamma_n\right| \leq \sqrt{\sum_{n=2}^{\infty} c_n^2}\sqrt{\sum_{n=2}^{\infty} \gamma_n^2}. \qquad (1.14)$$

For us,

$$c_n = \frac{\log^2 10}{\log^2 10 + (2\pi n)^2}, \quad \gamma_n = \exp\left(-\frac{\pi^2 n}{c \log 10}\right). \qquad (1.15)$$

Lemma 1.1. *Notation as above, we have*

$$\begin{aligned}\sum_{n=2}^{\infty} c_n^2 &\leq \frac{\log 10}{8\pi}\left(\pi - 2\arctan\left(\frac{2\pi}{\log 10}\right) - \sin\left(2\arctan\left(\frac{2\pi}{\log 10}\right)\right)\right)\\ &\approx 0.00177749\\ \sum_{n=2}^{\infty} \gamma_n^2 &= \exp\left(-\frac{4\pi^2}{c \log 10}\right) \cdot \left[1 - \exp\left(-\frac{2\pi^2}{c\log 10}\right)\right]^{-1}. \qquad (1.16)\end{aligned}$$

Proof. We have

$$\begin{aligned}
\sum_{n=2}^{\infty} c_n^2 &= \sum_{n=2}^{\infty}\left(\frac{1}{1+\left(\frac{2\pi n}{\log 10}\right)^2}\right)^2 \\
&= \left(1+\left(\frac{2\cdot 2\pi}{\log 10}\right)^2\right)^{-2} + \sum_{n=3}^{\infty}\left(\frac{1}{1+\left(\frac{2\pi n}{\log 10}\right)^2}\right)^2 \\
&\le \left(1+\left(\frac{2\cdot 2\pi}{\log 10}\right)^2\right)^{-2} + \int_2^{\infty}\left(1+\left(\frac{2\pi x}{\log 10}\right)^2\right)^{-2} dx \qquad (1.17)
\end{aligned}$$

(as the integrand is monotonically decreasing, thus we only increase the integral by starting at 2 instead of 3). We first change variables by letting $y = 2\pi x/\log 10$, and find

$$\sum_{n=3}^{\infty} c_n^2 \le \frac{\log 10}{2\pi}\int_{4\pi/\log 10}^{\infty}\frac{dy}{(1+y^2)^2}. \qquad (1.18)$$

We now change variables by letting $y = \tan\theta$, so $dy = \sec^2\theta d\theta$:

$$\begin{aligned}
\sum_{n=3}^{\infty} c_n^2 &\le \frac{\log 10}{2\pi}\int_{\arctan(4\pi/\log 10)}^{\pi/2}\frac{\sec^2\theta d\theta}{(1+\tan^2\theta)^2} \\
&= \frac{\log 10}{2\pi}\int_{\arctan(4\pi/\log 10)}^{\pi/2}\frac{\sec^2\theta d\theta}{\sec^4\theta} \\
&= \frac{\log 10}{2\pi}\int_{\arctan(4\pi/\log 10)}^{\pi/2}\cos^2\theta d\theta \\
&= \frac{\log 10}{2\pi}\left[\frac{\theta}{2}+\frac{\sin(2\theta)}{4}\right]_{\theta=\arctan(4\pi/\log 10)}^{\pi/2} \\
&= \frac{\log 10}{8\pi}\left(\pi - 2\arctan\left(\frac{4\pi}{\log 10}\right) - \sin\left(2\arctan\left(\frac{4\pi}{\log 10}\right)\right)\right) \\
&\approx 0.00072228. \qquad (1.19)
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_{n=2}^{\infty} c_n^2 &\le \left(1+\left(\frac{4\pi}{\log 10}\right)^2\right)^{-2} \\
&\quad + \frac{\log 10}{8\pi}\left(\pi - 2\arctan\left(\frac{4\pi}{\log 10}\right) - \sin\left(2\arctan\left(\frac{4\pi}{\log 10}\right)\right)\right) \\
&\approx 0.00177749. \qquad (1.20)
\end{aligned}$$

If instead we use Mathematica to numerically evaluate the sum, we get approximately 0.00140459 (we could easily get closer to this result by keeping more terms). In particular, we see our estimation is pretty good (we are off by about 26%).

We now turn to the sum of γ_n^2:

$$\begin{aligned}
\sum_{n=2}^{\infty} \gamma_n^2 &= \sum_{n=2}^{\infty} \exp\left(-\frac{2\pi^2 n}{c\log 10}\right) \\
&= \exp\left(-\frac{4\pi^2}{c\log 10}\right) \sum_{\ell=0}^{\infty} \left(\exp\left(-\frac{2\pi^2}{c\log 10}\right)\right)^n \\
&= \exp\left(-\frac{4\pi^2}{c\log 10}\right) \cdot \left[1 - \exp\left(-\frac{2\pi^2}{c\log 10}\right)\right]^{-1}. \qquad (1.21)
\end{aligned}$$

While our bound depends on c, we can evaluate it exactly as it is a geometric series. □

Lemma 1.2. *We have*

$$\sum_{n=2}^{\infty} b_n \ \leq\ 0.0421603 \cdot \exp\left(\frac{c+1}{c}\right) \cdot \exp\left(-\frac{2\pi^2}{c\log 10}\right) \cdot \left[1 - \exp\left(-\frac{2\pi^2}{c\log 10}\right)\right]^{-1/2}. \qquad (1.22)$$

Proof. This follows immediately from the previous lemma and the Cauchy-Schwartz inequality. □

We can now give some good estimates on $\sum_{n\geq 2} b_n$ for various values of c. Instead of using our exact bound of about $.001777$ for the sum of c_n^2 we instead use the numerical bound of about $.001405$. We have

$$\sum_{n=2}^{\infty} b_n \quad \leq \quad g(c), \qquad (1.23)$$

where

$$g(c) \ =\ 0.0421603 \cdot \exp\left(\frac{c+1}{c}\right) \cdot \exp\left(-\frac{2\pi^2}{c\log 10}\right) \cdot \left[1 - \exp\left(-\frac{2\pi^2}{c\log 10}\right)\right]^{-1/2}. \qquad (1.24)$$

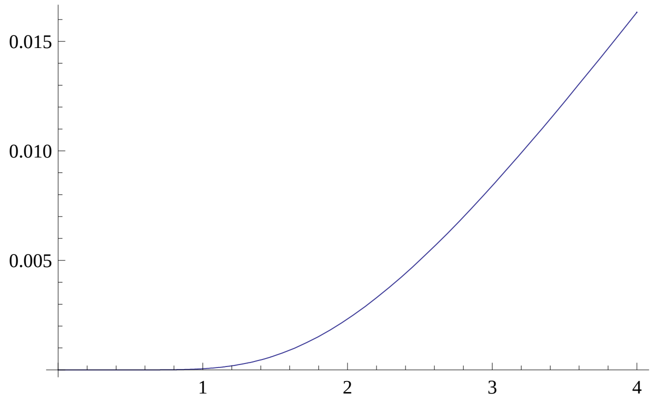
Thus

c	$g(c)$
0.50	$2.70 \cdot 10^{-8}$
0.75	$4.20 \cdot 10^{-6}$
1.00	$5.24 \cdot 10^{-5}$
1.25	$2.38 \cdot 10^{-4}$
1.50	$6.55 \cdot 10^{-4}$

For larger values of c we find

c	$g(c)$
1.75	.00135
2.00	.00232
2.25	.00356
2.50	.00501
2.75	.00664

Graphically, we find



E-mail address: chaouki@ece.unm.edu

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING, UNIVERSITY OF NEW MEXICO, ALBUQUERQUE, NM 87131

E-mail address: heileman@ece.unm.edu

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING, UNIVERSITY OF NEW MEXICO, ALBUQUERQUE, NM 87131

E-mail address: hill@math.gatech.edu

E-mail address: sjmiller@math.brown.edu

DEPARTMENT OF MATHEMATICS, BROWN UNIVERSITY, PROVIDENCE, RI 02912

E-mail address: fperez@gts.tsc.uvigo.es

DEPARTAMENTO DE TEORÍA DE LA SEÑAL Y LAS COMUNICACIONES, UNIVERSIDAD DE VIGO, SPAIN