A Simple Heuristic Proof of Hardy and Littlewood's Conjecture B

Michael Rubinstein

Hardy and Littlewood have conjectured (see [3]) that *there are infinitely many prime pairs* $(p, p+m)$ *for every even m. If* $\pi_m(x)$ *is the number of pairs less than x, then*

$$\pi_m(x) \sim 2C_2 \frac{x}{(\log x)^2} \prod_{\substack{p>2 \\ p \mid m}} \frac{p-1}{p-2} \tag{*}$$

where $C_2 = \prod_{p>2}(1 - 1/(p-1)^2)$.

If $m = 2$ then we call pairs of primes $(p, p+2)$ twin primes. Thus, in particular, (*) implies that

$$\pi_2(x) \sim 2C_2 \frac{x}{(\log x)^2}. \tag{**}$$

In this paper, I present a short heuristic proof of (**) and generalize to give a heuristic proof of (*).

For other heuristic proofs consult [2] and [4].

"*Proof*" We begin by considering the following combinatorial problem: Let $A = \{1, 2, 3, \dots, a\}$ and choose from A a subset B consisting of b elements. Now, say we choose another subset B' from A consisting of b' elements. The elements in B' are chosen randomly (by randomly I mean that every element of A has an equal chance of being picked). Then, the expected number of elements in $B \cap B'$ is equal to $b' \cdot (b/a) = b'b/a$. Since this trivial result is used throughout this article, we refer to it as (1).

We also reference Dirchlet's Theorem (see [1] for a proof of Dirichlet's Theorem) which states that given an arithmetical progression $ak + b$, where $(a, b) = 1$, then

$$\sum_{\substack{p \in ak+b \\ p \le x}} 1 \sim \frac{x}{\phi(a) \log x}. \tag{2}$$

We now give a heuristic proof of (**).

Consider the following 2 pairs of arithmetical progressions:

$$(2k+0), (2k+2)$$
$$(2k+1), (2k+3)$$

where $k = 0, 1, 2, 3 \dots$.

All twin primes will fall in the second pair $(2k+1), (2k+3)$. Therefore, the question of how many twin primes exist up to a certain x is equivalent to finding out how many values of k make $2k+1$ and $2k+3$ simultaneously prime.

Now, assume that any value of k is equally likely to make $2k + 1$ prime, and any value of k is equally likely to make $2k + 3$ prime (this assumption is incorrect and will be improved on shortly). Further, assume that the primes belonging to $2k + 1$ behave independently from the primes belonging to $2k + 3$. Thus, using (1) and (2), we have that

$$\pi_2(x) \sim \frac{\left(\dfrac{x}{\phi(2)\log x}\right)^2}{\dfrac{x}{2}} = 2\frac{x}{(\log x)^2}.$$

In the above,

$$b \sim b' \sim \frac{x}{\phi(2)\log x}, \quad \text{and} \quad a = \left[\frac{x}{2}\right] \sim \frac{x}{2}.$$

The $x/(\log x)^2$ part reflects the conjecture, but the constant 2 does not. The assumption that any value of k is equally likely to make $2k + 1$ prime, and that any value of k is equally likely to make $2k + 3$ prime is incorrect. For, if $k \equiv 1 \bmod 3$, then $2k + 1$ is composite ($k > 1$), and if $k \equiv 0 \bmod 3$, then $2k + 3$ is composite ($k > 0$). And so, we improve on the above result by considering pairs of arithmetical progressions mod 6 rather than pairs of arithmetical progressions mod 2. Thus, consider the following 6 pairs of arithmetical progressions:

$$(6k + 0), (6k + 2)$$
$$(6k + 1), (6k + 3)$$
$$(6k + 2), (6k + 4)$$
$$(6k + 3), (6k + 5)$$
$$(6k + 4), (6k + 6)$$
$$(6k + 5), (6k + 7)$$

We can eliminate all pairs except for $(6k + 5), (6k + 7)$ as the others cannot contribute anything to $\pi_2(x)$ (except when $k = 0$). Thus by (1) and (2), we have that,

$$\pi_2(x) \sim \frac{\left(\dfrac{x}{\phi(2 \cdot 3)\log x}\right)^2}{\dfrac{x}{2 \cdot 3}} = \frac{3}{2}\frac{x}{(\log x)^2}.$$

Somewhat better.

However, once again not every value of k is equally likely to make $6k + 5$ prime, and not every value of k is equally likely to make $6k + 7$ prime. For, if $k \equiv 0 \bmod 5$, then $6k + 5$ is composite ($k > 0$), and if $k \equiv 3 \bmod 5$ then $6k + 7$ is composite. And so, we improve on the above by considering pairs of arithmetical progressions mod 30 instead of pairs of arithmetical progressions mod 6. Thus, consider the following 30 pairs of arithmetical progressions:

$$(30k + 0), (30k + 2)$$
$$(30k + 1), (30k + 3)$$

$$(30k + 29), (30k + 31)$$

The only three pairs that contribute to $\pi_2(x)$ are $(30k+11), (30k+13)$, $(30k+17), (30k+19)$, $(30k+29), (30k+31)$ (with the obvious few finite exceptions such as 3, 5 when $k=0$).

Now, by (1) and (2), each pair contributes

$$\frac{\left(\dfrac{x}{\phi(2\cdot3\cdot5)\log x}\right)^2}{\dfrac{x}{2\cdot3\cdot5}} = \frac{2\cdot3\cdot5x}{(1\cdot2\cdot4\log x)^2} = \frac{15}{32}\frac{x}{(\log x)^2}$$

to $\pi_2(x)$. Thus

$$\pi_2(x) \sim 3\left(\frac{15}{32}\frac{x}{(\log x)^2}\right).$$

Once again, not all values of k are equally likely to make $30k+b$ prime. And so we continue in the above fashion indefinitely, and find that

$$\pi_2(x) \sim \lim_{k\to\infty}\phi_2(2\cdot3\cdot5\cdots p_k)\left(\frac{\left(\dfrac{x}{\phi(2\cdot3\cdot5\cdots p_k)\log x}\right)^2}{\dfrac{x}{2\cdot3\cdot5\cdots p_k}}\right)$$

$$= \lim_{k\to\infty}\frac{\phi_2(2\cdot3\cdot5\cdots p_k)(2\cdot3\cdot5\cdots p_k)}{(\phi(2\cdot3\cdot5\cdots p_k))^2}\frac{x}{(\log x)^2},$$

where $\phi_2(n)$ denotes the number of pairs $c, c+2$ such that $(n,c)=(n,c+2)=1$, and $0\le c\le n-1$ (i.e. the number of pairs of arithmetical progressions $(nk+c), (nk+c+2)$ that contain infinitely many pairs of primes).

The reader will object that we are unjustified in passing to the limit since (2) is no longer valid when a (in (2)) is infinite. We would have to fix this problem by letting k (in the above) be a function of x (so that $p_k \ll x$), and use a more precise version of (2). However, since the method being described is heuristic, we omit the details, and proceed.

I claim that $\phi_2(2\cdot3\cdot5\cdot7\cdots p_k)=(3-2)(5-2)(7-2)(11-2)\cdots(p_k-2)$. This will be proven shortly. Assuming this for the moment, we see that the above is simply

$$\pi_2(x) \sim 2\prod_{p>2}\frac{(p-2)p}{(p-1)^2}\frac{x}{(\log x)^2} = 2\prod_{p>2}\left(1-\frac{1}{(p-1)^2}\right)\frac{x}{(\log x)^2}.$$

We have arrived at (**).

It remains to be shown that $\phi_2(2\cdot3\cdot5\cdots p_k)=(3-2)(5-2)\dots(p_k-2)$. In fact, I prove that.

Lemma 1. *Given $n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_j^{\alpha_j}$ where the p_i's are arbitrary primes and $\alpha_i\ge1$, then*

$$\phi_2(n)=p_1^{\alpha_1-1}p_2^{\alpha_2-1}\cdots p_j^{\alpha_j-1}\prod_{p_i>2}(p_i-2).$$

Note the similarity to Euler's totient function.

Proof: Two steps are required:

a) Show that it is true for p^α, p a prime. That,

$$\phi_2(p^\alpha) = p^{\alpha-1}\cdot(p-2) \qquad (\text{if } p > 2)$$
$$= p^{\alpha-1} \qquad (\text{if } p = 2)$$

b) Show that $\phi_2(ab) = \phi_2(a)\phi_2(b)$ if $(a,b) = 1$.

Proof of a): If $p > 2$ then, of the integers $0, 1, 2, 3, 4 \dots, p^\alpha - 1$ there are exactly $p^{\alpha-1}$ integers that are not relatively prime to p^α (namely $0, p, 2p, 3p, 4p, \dots, (p^{\alpha-1} - 1)p$). Each of these $p^{\alpha-1}$ integers kills two arithmetical progressions and so $\phi_2(p^\alpha) = p^\alpha - 2p^{\alpha-1} = p^{\alpha-1}(p-2)$.

If $p = 2$, then each multiple of p only kills one pair of arithmetical progressions (since we must not double count) and so $\phi_2(p^\alpha) = p^\alpha - p^{\alpha-1} = p^{\alpha-1}$ (since $p = 2$).

Proof of b): Partition the integers $0, 1, 2, 3 \dots, ab - 1$ as follows:

$$\underset{A1}{0, 1, \dots, a-1}, \quad \underset{A2}{a, \dots, 2a-1}, \quad \underset{A3}{2a, \dots, 3a-1}, \quad \underset{Ab}{\dots (b-1)a, \dots, ba-1}.$$

Now, in each Ai, there are $\phi_2(a)$ pairs $c, c+2$, where $(i-1)a \le c \le ia - 1$, such that $(a,c) = (a, c+2) = 1$. Furthermore, each pair appears b times modulo a (i.e. once in each Ai). We need to find how many of these pairs are relatively prime to b. Well, for any pair $c, c+2$, where $0 \le c \le a - 1$, we may list the b times that it appears modulo a:

$$c, c+2,\ a+c, a+c+2, \dots, (b-1)a + c, (b-1)a + c + 2. \tag{3}$$

If we examine these pairs modulo b, we see that, since $(a,b) = 1$, they run through all pairs $i, i+2 \bmod b$. And so, of the pairs in (3), $\phi_2(b)$ are relatively prime to b. Thus, the total number of pairs relatively prime to a and b is equal to $\phi_2(a)\phi_2(b)$.

The above method that I've described above for twin primes may be generalized to work for (*). In fact, an identical argument is used, the only difference being in our counting function $\phi_2(n)$. I leave out the details (as they are more or less identical) but summarize the results below. The reader should fill in the details while taking a shower.

Given m is an even integer then

$$\pi_m(x) \sim \lim_{k\to\infty} \phi_m(2\cdot 3\cdot 5 \cdots p_k)\left(\frac{\left(\dfrac{x}{\phi(2\cdot 3\cdot 5\cdots p_k)\log x}\right)^2}{\dfrac{x}{2\cdot 3\cdot 5\cdots p_k}}\right)$$

$$= \lim_{k\to\infty} \frac{\phi_m(2\cdot 3\cdot 5\cdots p_k)(2\cdot 3\cdot 5\cdots p_k)x}{(\phi(2\cdot 3\cdot 5\cdots p_k))^2(\log x)^2},$$

where $\phi_m(n)$ denotes the number of pairs $c, c+m$ such that $(n,c) = (n, c+m) = 1$, and $0 \le c \le n-1$ (i.e. the number of pairs of arithmetical progressions $(nk+c), (nk+c+m)$ that contain infinitely many primes).

In a similar manner to the proof of lemma 1, it is easy to establish that

$$\begin{aligned}\phi_m(p^\alpha) &= p^{\alpha-1}\cdot(p-2) \qquad (\text{if } (p,m)=1,\ p \text{ a prime})\\ &= p^{\alpha-1}(p-1) \qquad (\text{if } p|m,\ p \text{ a prime})\end{aligned}$$

and that

$$\phi_m(ab) = \phi_m(a)\phi_m(b) \quad \text{if } (a,b)=1.$$

Thus, in particular we have,

$$\begin{aligned}\phi_m(2\cdot 3\cdot 5\cdots p_k) &= \prod_{\substack{p\le p_k\\ (p,m)=1}}(p-2)\prod_{\substack{p\le p_k\\ p|m}}(p-1)\\ &= \prod_{2<p\le p_k}(p-2)\prod_{\substack{2<p\le p_k\\ p|m}}\frac{p-1}{p-2}.\end{aligned}$$

And so,

$$\begin{aligned}\pi_m(x) &\sim 2\left(\prod_{p>2}\frac{(p-2)p}{(p-1)^2}\right)\left(\prod_{\substack{p>2\\ p|m}}\frac{p-1}{p-2}\right)\frac{x}{(\log x)^2}\\ &= 2C_2\frac{x}{(\log x)^2}\prod_{\substack{p>2\\ p|m}}\frac{p-1}{p-2}.\end{aligned}$$

We have arrived at (*).

CONCLUSION. The beauty of the heuristic method that I have described in this article lies in its simplicity, and in the fact that it gives the conjectured value. It presents the reader with a somewhat convincing reason as to why the conjecture ought to be true. It should also be remarked that the same strategy can be applied to similar problems (such as pairs of primes that differ linearly and not just by a constant).

REFERENCES

1. Apostol, *An Introduction to Analytic Number Theory*, 1976, Springer-Verlag, New York, 146–156.
2. Cherwell, *Quarterly Journal of Mathematics* (Oxford), 17 (1946) 46–62.
3. Hardy and Littlewood, *Acta Math.*, 44 (1923) 1–70.
4. Polya, *American Math. Monthly*, 66 (1959) 375–384.

68 Banstead Rd.
Montreal West, Quebec
Canada, H4X 1P2
miker@phoenix.princeton.edu