Nov. 27, 2010

Estimating the Number of Goldbach Partitions

An alternative approach

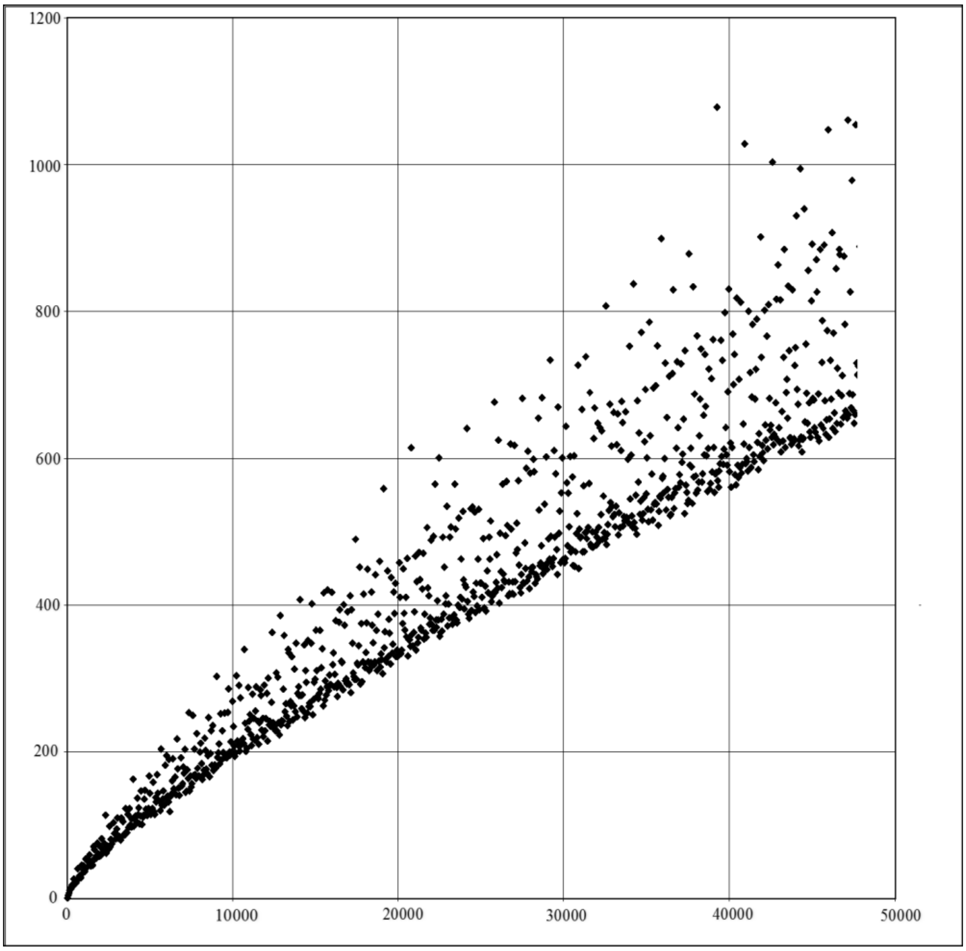
Mathieu Marchal

Abstract

The sum oft two odd primes is an even number. The inverse proposition (any even number $n \geq 6$ can be written in at least one way as a sum of two odd primes) has been the subject of a correspondence between Goldbach and Euler, in 1742. It has never been proved, nor disproved.

As n grows larger, the number of possible partitions, $g(n)$, increases in an irregular way, depending heavily on n's divisors. A plot of $g(n)$ reveals a comet-like pattern as shown on **Figure 1**.

Fig. 1 Plot of $g(n)$ for values $n = 6 + m \cdot 48$; $(1 \leq m \leq 1000)$



Asymptotic formulas, describing $g(n)$'s dependency on *n*'s divisors, have been conjectured by a number of authors. Most formulas contain the factors, which had been identified by Sylvester as far as in 1871.

In this paper, we present a new approach, yielding fairly accurate results, which are compared with those obtained by known formulas.

The Hardy-Littlewood conjecture

The work of Brun(1915), Stäckel (1916-1918), Hardy and Littlewood (1922), among others, confirmed a formula, which, essentially, is equivalent to the one which Sylvester had arrived at:

$$g(n) \sim C \cdot \frac{n}{\log^2(n)} \cdot \prod_{\substack{3<p<\sqrt{n} \\ p/n}} \frac{p-1}{p-2}, \qquad (1.1)$$

$g(n)$ representing the count of partitions into prime numbers q and $n-q$, with $q \le n/2$, and the product being taken over the odd prime divisors of n.

The approach taken by British mathematicians Hardy and Littlewood $[1]$ is interesting. Dealing with the general problem of partitioning numbers into primes, they had developed their *circle method*, which enabled them to prove, that *every sufficiently large even number is the sum of four odd prime numbers, and every sufficiently large odd number the sum of three*. If, in the case of two primes, the method failed to yield a proof, it nevertheless helped to support Sylvesters formula. Moreover, reviewing the different values, which had been proposed for the constant factor C (Sylvester had taken $C = 0.5$), they found a strong argument for it being equal to the “twin primes constant”:

$$C_{HL} = \prod_{3 \le p \le \infty} \frac{p \cdot (p-2)}{(p-1)^2} = 0.6601618158..., \qquad (1.2)$$

An alternative formula

From the $\pi(n)$ prime numbers, which are inferior to n, we sieve out those, which do not lend themselves to a Goldbach partition. This approach implies multiplying $\pi(n)$ by a factor :

$$g_{alt}(n) \sim \frac{1}{2} \cdot \pi(n) \cdot \prod_{\substack{3<p<\sqrt{n} \\ p \times n}} \frac{p-2}{p-1}, \qquad (2a)$$

the product being taken over the prime numbers p *which are primes relative to* n :
$n \ne 0 \pmod p$.

To support this formula we use Merten's 3rd theorem:

$\lim_{n\to\infty} \prod_{2\le p\le n} \frac{p-1}{p} \cdot \log(n) = e^{-\gamma}$, with the Euler-Mascheroni constant $\gamma = 0.5772156649...$,

from which we deduce the asymptotic equivalences

$$\frac{1}{2} \cdot e^{\gamma} \cdot \prod_{2\le p<\sqrt{n}} \frac{p-1}{p} \sim \frac{1}{\log(n)}, \quad \text{and} \quad \frac{\pi(n)}{n} \sim \frac{1}{2} \cdot e^{\gamma} \cdot \prod_{2\le p<\sqrt{n}} \frac{p-1}{p} \qquad (2.1)$$

Numbers q ($\sqrt{n} < q < n$) which lend themselves to a Goldbach Partition must satisfy the two conditions:

$n \ne 0 \pmod{p}$ and $n \ne q \pmod{p}$ for all prime numbers p ($2 \le p < \sqrt{n}$) (2.2.1)

To estimate their density, we have to modify the formula (2.1) to :

$$\frac{G(n)}{n} \sim \frac{1}{2} \cdot e^{\gamma} \cdot \prod_{\substack{2\le p<\sqrt{n} \\ p/n}} \frac{p-1}{p} \cdot \prod_{\substack{3\le p<\sqrt{n} \\ p\times n}} \frac{p-2}{p} \qquad (2.2.2)$$

This can be written in the form :

$$G(n) \sim \pi(n) \cdot \prod_{\substack{2\le p<\sqrt{n} \\ p/n}} (p-1) \cdot \prod_{\substack{3\le p<\sqrt{n} \\ p\times n}} (p-2) \Big/ \prod_{3\le p<\sqrt{n}} (p-1) \qquad (2.2.3)$$

Simplifying the quotient, and counting partitions into prime numbers q and $n-q$, with $q \le n/2$, we obtain the formula (2a).

Obviously, the formulas (2a) and

$$g'_{alt}(n) \sim \frac{1}{2} \cdot \frac{n}{\log(n)} \cdot \prod_{\substack{3< p<\sqrt{n} \\ p\times n}} \frac{p-2}{p-1} \quad \text{, and} \qquad (2b)$$

$$g''_{alt} \sim \frac{n}{4} \cdot e^{\gamma} \cdot \prod_{\substack{2\le p<\sqrt{n} \\ p/n}} \frac{p-1}{p} \cdot \prod_{\substack{3\le p<\sqrt{n} \\ p\times n}} \frac{p-2}{p} \qquad (2c)$$

are equivalent, as are the formulas (1.1, 1.2), and (2) :

$$\frac{g_{alt}}{g_{HL}} = \left[\frac{\pi(n)}{n/\log(n)}\right] \cdot \left[\frac{1}{1/\log(n)} \cdot \frac{e^{\gamma}}{2} \cdot \prod_{2\le p<\sqrt{n}} \frac{(p-1)}{p}\right] \cdot \left[\prod_{\sqrt{n}<p\le\infty} \frac{(p-1)^2}{p\cdot(p-2)}\right] \sim 1,$$

each of the three factors in square brackets tending towards unity.

Results

A preliminary numerical verification did show, that the results calculated with (2a) are not necessarily more accurate than those computed with (2b) or (2c), which are both about equally precise.

On **Table 1** and **Fig. 2**, the values for $g_{alt}(n)$, which have been calculated with the formula (2c) are compared with $g(n)$, as well with those calculated with the « corrected » Hardy-Littlewood formula*)

$$g_{SW}(n) \sim C_{HL} \cdot \frac{n}{(\log^2 n - 2 \cdot \log n)} \cdot \prod_{\substack{3 \le p \le \sqrt{n} \\ p / n}} \frac{p-1}{p-2} \tag{1.3}$$

Discussion

For its simplicity, the formula $g_{alt}(n)$ adequately describes the behaviour of $g(n)$, though it does not quite attain the degree of precision granted by $g_{SW}(n)$. The latter seems to be a remarkably accurate formula, indeed. After testing several formulas up to $n \le 5 \cdot 10^8$, Richstein $[3]$ finds it "a bit surprising" , that $g_{SW}(n)$ yields better results, on the whole, than other, more elaborate ones (with the exception of Selmer's formulas).

Acknowledgment

The author is indebted to Greg Bernhardt for signalling an error contained in a preliminary version of this paper.

---

*) In their paper $[1]$, Hardy and Littlewood stated, that, in the formula (1.1), the term $\log^2(n)$ of the denominator *is certainly in error wrong to an order* $\log(n)$ and, as had been suggested by Shah and Wilson, *must be replaced by* $\log^2(n) - 2 \cdot \log(n)$, a substitution, which is asymptotically irrelevant, but *essential for the purpose of verification within the limits of calculation*.

References

[1] G.H.Hardy, J.E. Littlewood, *Some problems of 'partitio numerorum'; III: On the expression of a number as a sum of primes*, Acta Math. **44** (1922), 32 – 39

[2] Wang Yuan (Ed.), *Goldbach Conjecture*, 1984, World Scientific Publishing, Singapore, ISBN 9971-966-08-5, and ISBN 9971-966-09-3 pbk (a collection of selected papers, containing a photographic reproduction of [1])

[3] Richstein J. ; Bosma Wieb (Editor*); Computing the number of Goldbach partitions up to 5.10*^{8}, Conference : Algorithmic number theory. International symposium, 4, Leiden, NLD, 2000-07-02 , ISSN : 0302-9743 ISBN : 3-540-67695-3

Table 1 Calculated results $g_{alt}(n)$ and $g_{SW}(n)$ compared with $g(n)$

n			*g(n)*	$g_{alt}(n)$	$g_{SW}(n)$	$g_{alt}(n)$	$g_{SW}(n)$	$(g_{alt}-g_{SW})/g_{SW}$
						Relative	Error	
2*3*5	=	30	3	4	11	33,33%	266,67%	-63,64%
2*3*5*7	=	210	19	15	25	-21,05%	31,58%	-40,00%
2*3*5*7*11	=	2310	114	93	122	-18,42%	7,02%	-23,77%
2*3*5*7*11*13	=	30030	905	788	898	-12,93%	-0,77%	-12,25%
2*3*5*7*11*13*17	=	510510	9493	8969	9521	-5,52%	0,29%	-5,80%
2*3*5*7*11*17*19*23*29	=	497668710*)	3977551*)	4006949	3970818	0,74%	-0,17%	0,91%
		1000	28	18	26	-35,71%	-7,14%	-30,77%
		10000	127	114	133	-10,24%	4,72%	-14,29%
		100000	810	731	804	-9,75%	-0,74%	-9,08%
		1000000	5402	5139	5392	-4,87%	-0,19%	-4,69%
		10000000	38807	37974	38681	-2,15%	-0,32%	-1,83%
		100000000	291400	290577	290999	-0,28%	-0,14%	-0,15%
2^{10}	=	1024	22	14	20	-36,36%	-9,09%	-30,00%
2^{11}	=	2048	25	24	32	-4,00%	28,00%	-25,00%
2^{12}	=	4096	53	42	51	-20,75%	-3,77%	-17,65%
2^{13}	=	8192	76	71	86	-6,58%	13,16%	-17,44%
2^{14}	=	16384	151	125	145	-17,22%	-3,97%	-13,79%
2^{15}	=	32768	244	217	248	-11,07%	1,64%	-12,50%
2^{16}	=	65536	435	388	429	-10,80%	-1,38%	-9,56%
2^{17}	=	131072	749	689	751	-8,01%	0,27%	-8,26%
2^{18}	=	262144	1314	1228	1324	-6,54%	0,76%	-7,25%
2^{19}	=	524288	2367	2220	2353	-6,21%	-0,59%	-5,65%
2^{20}	=	1048576	3957	4010	4209	1,34%	6,37%	-4,73%
2^{21}	=	2097152		7302	7575			-3,60%
2^{22}	=	4194304		13305	13705			-2,92%
2^{23}	=	8388608		24340	24914			-2,31%
2^{24}	=	16777216		44750	45491			-1,63%
2^{25}	=	33554432		82614	83393			-0,93%
2^{26}	=	67108864		152885	153433			-0,36%
2^{27}	=	134217728		283717	283247			0,17%
2^{28}	=	268435456		527313	524511			0,53%
2^{29}	=	536870912		983427	974065			0,96%

*)For $6 \le n \le 5 \cdot 10^8$, the maximal value assumed by $g(n)$ (Richstein[3])

Fig. 2a Calculated results $g_{alt}(n)$ compared with $g(n)$

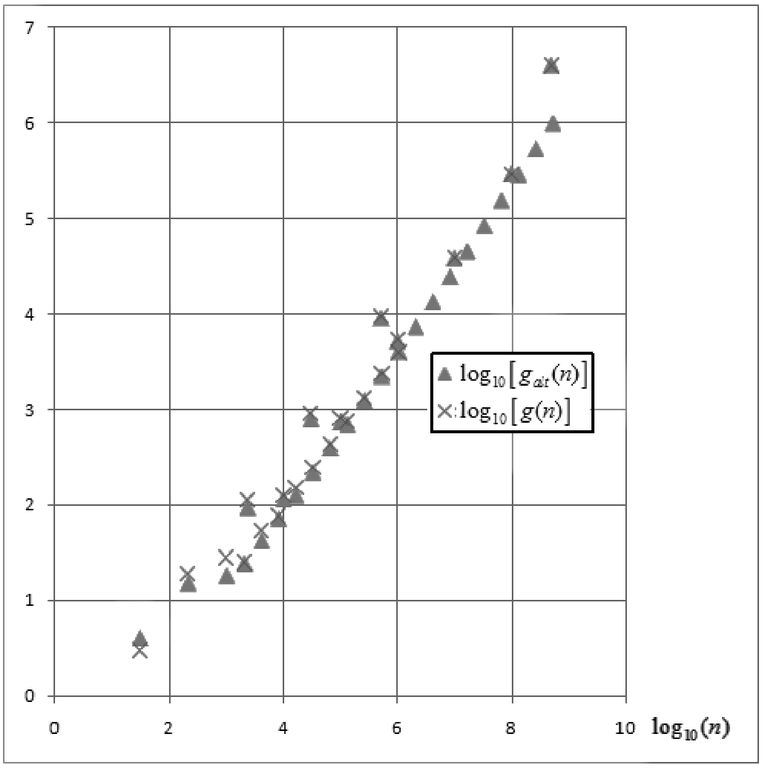


Fig. 2b Calculated results $g_{alt}(n)$ compared with $g_{SW}(n)$