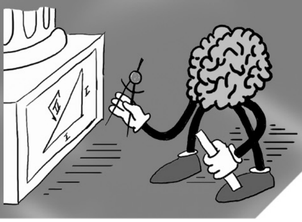

Trifectas in geometric progression

Gerry Myerson*

Abstract

The trifecta in the 2007 Melbourne Cup was the numbers 6–12–24, a geometric progression. How many trifectas in geometric progression are there in an n–horse race?

Introduction

Sydney radio presenter (and first class honours Mathematics graduate) Adam Spencer noted that the winning trifecta in the 2007 Melbourne Cup was the numbers 6, 12 and 24, a geometric progression. He asked his audience: 'How many trifectas in geometric progression are there with 24 horses?'

'Trifecta' is, of course, a technical term, referring to choosing the horses that finish first, second and third, in that order. The order is important — if you bet on the 6-12-24 trifecta, and the horses come in 24-12-6, you don't win. The three numbers have to be integers, but the common ratio does not. Thus, for example, 4-6-9 is a trifecta in geometric progression.

With these conventions in place, it is not hard to draw up a systematic list of all the qualifying trifectas, and to find that there are 24 of them. What if, instead of horses numbered 1 to 24, there are horses numbered 1 to n? How many trifectas in geometric progression are there then?

The Spencer function and its relatives

Let the Spencer function, $S(n)$, be the number of trifectas in geometric progression in a race with horses numbered 1 to n. Estimating $S(n)$ is a nice problem we can use as a vehicle to illustrate some simple techniques of analytic number theory. We will prove the following theorem.

Theorem 1. $S(n) = \frac{6}{\pi^2} n \log n + O(n)$.

The appearance of π in this formula is undoubtedly related to the circular portion of the racetrack.

Received 19 November 2007; accepted for publication 17 December 2007.
*Mathematics, Macquarie University, NSW 2109. E–mail: gerry@maths.mq.edu.au

We introduce some relatives of the Spencer function that have already appeared in the literature. Let $f(n)$ be the number of triples of integers (a,b,n) in geometric progression with $1 \le a \le b \le n$. Let $F(n)$ be the summatory function for $f(n)$, that is, $F(n) = \sum_{m=1}^{n} f(m)$. Then $F(n)$ differs from $S(n)$ in two ways; it counts the progressions (a,a,a), which $S(n)$ does not, and it omits the progressions with $a > b > n$, which $S(n)$ includes. Thus, the two functions are related by

$$S(n) = 2F(n) - 2n\,. \tag{1}$$

The values of the function $f(n)$ for $n = 1, 2, \ldots, 12$ are $1, 1, 1, 2, 1, 1, 1, 2, 3, 1, 1, 2$. We can search for this sequence in [2], a resource not available to Hardy and the other giants of analytic number theory of years gone by. We find that it agrees with sequence A000188, which is given as the number of solutions to $x^2 \equiv 0 \pmod{n}$. We can show that this is our $f(n)$. If (a,b,n) are in geometric progression, then $an = b^2$, so $x = b$ is a solution of $x^2 \equiv 0 \pmod{n}$. Conversely, if $b^2 \equiv 0 \pmod{n}$ with $1 \le b \le n$, then $an = b^2$ for some a, $a \le b \le n$, and (a,b,n) is a geometric progression.

An integer is said to be *squarefree* if it is not divisible by any square number other than 1. Every positive integer n can be written in exactly one way as $n = d^2e$, where e is squarefree. A000188 is also described as the sequence whose nth term is the square root of the largest square dividing n; that is, the nth term is d, where $n = d^2e$ and e is squarefree. Let us show that this, too, is our $f(n)$. If (a,b,n) is a three-term geometric progression, and $n = d^2e$, then for an to be a square we need $a = c^2e$ for some c, and then $b = cde$. In fact, c can be any of the numbers $1, 2, \ldots, d$, so $f(n) = d$.

The site [2] gives three formulas for $f(n)$:

(1) $f(n) = \sum_{r^2 \mid n} \phi(r)$, where ϕ is the Euler phi-function ($\phi(n)$ is the number of positive integers relatively prime to and not exceeding n),
(2) f is multiplicative (that is, if m and n are relatively prime, then $f(mn) = f(m)f(n)$), and, for p prime, $f(p^e) = p^{[e/2]}$ (here and below, $[x]$ is the greatest integer not exceeding x),
(3) For real part of s exceeding 1, $\sum_{n=1}^{\infty} f(n)n^{-s} = \zeta(2s-1)\zeta(s)/\zeta(2s)$, where $\zeta(s) = \sum_{n=1}^{\infty} n^{-s}$ is the Riemann zeta-function.

Let us establish these formulas. First, if $n = d^2e$, as above, then $r^2 \mid n$ is equivalent to $r^2 \mid d^2$, and so to $r \mid d$. Thus $\sum_{r^2 \mid n} \phi(r) = \sum_{r \mid d} \phi(r)$, and it is a standard fact of elementary number theory that the second sum here is d.

Second, if $m = r^2s$, and $n = u^2v$, with m and n relatively prime and s and v squarefree, then $mn = (ru)^2sv$ and sv is squarefree, so $f(mn) = ru = f(m)f(n)$. Also, p^e is either $(p^{e/2})^2$ or $(p^{[e/2]})^2p$, depending on whether e is even or odd, so in either case $f(p^e) = p^{[e/2]}$.

Finally,

$$\begin{aligned}\frac{1-p^{-2s}}{(1-p^{-2s+1})(1-p^{-s})} &= (1+p^{-s})(1+p^{1-2s}+p^{2-4s}+p^{3-6s}+\cdots) \\ &= (1+\frac{1}{p^s}+\frac{p}{p^{2s}}+\frac{p}{p^{3s}}+\frac{p^2}{p^{4s}}+\frac{p^2}{p^{5s}}+\cdots) \\ &= \sum_{e=1}^{\infty}\frac{p^{[e/2]}}{p^{es}} \\ &= \sum_{e=1}^{\infty}\frac{f(p^e)}{(p^e)^s}\,. \end{aligned} \tag{2}$$

Now, Euler showed that $\zeta(s)=\prod_p(1-p^{-s})^{-1}$, so

$$\frac{\zeta(2s-1)\zeta(s)}{\zeta(2s)} = \prod_p \frac{1-p^{-2s}}{(1-p^{-2s+1})(1-p^{-s})} = \prod_p\sum_{e=1}^{\infty}\frac{f(p^e)}{(p^e)^s} = \sum_{n=1}^{\infty}\frac{f(n)}{n^s}\,, \tag{3}$$

where the last equation is a consequence of the unique factorization theorem and the multiplicativity of f. The manipulations of infinite series and products are justified by absolute convergence for real part of s exceeding 1.

Proof of Theorem 1

Finch and Sebah [1] applied the method of Selberg and Delange to (3) to establish $F(n)\sim 3\pi^{-2}n\log n$. Indeed, that paper applies the Selberg–Delange method to a large collection of similar problems, thus showcasing the versatility of the method. However, if one is only interested in estimating F, the method is overkill, and we will have no more to say about it here.

As explained earlier, every three-term geometric progression with no term exceeding n can be written uniquely in the form (a^2e, abe, b^2e) with e squarefree, $a^2e\le n$, and $b^2e\le n$. Thus, if $P(n)$ is the number of three-term geometric progressions with no term exceeding n, then $P(n)$ is given by

$$P(n) = \sum_{\substack{e\le n\\ e \text{ squarefree}}}\ \sum_{a\le\sqrt{n/e}}\ \sum_{b\le\sqrt{n/e}} 1\,. \tag{4}$$

It is related to our earlier functions by $P(n)=2F(n)-n=S(n)+n$. We get

$$\begin{aligned}P(n) &= \sum\nolimits^*\left[\sqrt{n/e}\right]^2 \\ &= \sum\nolimits^*\left(\sqrt{n/e}+O(1)\right)^2 \\ &= n\sum\nolimits^*\frac{1}{e}+O\Big(\sqrt{n}\sum\nolimits^*\frac{1}{\sqrt{e}}\Big)+O\Big(\sum\nolimits^* 1\Big)\,,\end{aligned} \tag{5}$$

where all the sums, $\sum^*$, are on squarefree e up to n. Trivially, $\sum_e 1\le n$, so the third term on the right is $O(n)$. The second sum on the right is less than $\sum_{e=1}^{n}e^{-1/2}$, which is $O(\sqrt{n})$ by comparison with the integral, so the second term on the right is also $O(n)$.

To handle the first sum on the right, we introduce the Möbius function. We define it on prime powers by $\mu(1) = 1$, $\mu(p) = -1$, and $\mu(p^r) = 0$ for $r \geq 2$. Then we extend it to all positive integers n by insisting that it be multiplicative. A crucial fact about the Möbius function is that $\sum_{d|n} \mu(d)$ is 1 if $n = 1$ and is 0 otherwise. For if $n \neq 1$ and p is any prime dividing n, then to each squarefree divisor d not divisible by p there corresponds the squarefree divisor pd divisible by p, and $\mu(d) + \mu(pd) = 0$.

From this it follows that $\sum_{d^2|n} \mu(d)$ is 1 if n is squarefree and is 0 otherwise. For if n is squarefree then the only integer d such that $d^2 \mid n$ is $d = 1$, while if $n = r^2 s$ with $r > 1$ and s squarefree then as we saw earlier d^2 divides n if and only if d divides r, and $\sum_{d|r} \mu(d) = 0$. Therefore, we can write

$$A = \sum_{\substack{e \leq n \\ e \text{ squarefree}}} \frac{1}{e} = \sum_{e=1}^{n} \sum_{r^2|e} \frac{\mu(r)}{e} = \sum_{r \leq \sqrt{n}} \sum_{q \leq n/r^2} \frac{\mu(r)}{qr^2} = \sum_{r \leq \sqrt{n}} \frac{\mu(r)}{r^2} \sum_{q \leq n/r^2} \frac{1}{q}, \quad (6)$$

where we have changed the order of summation and introduced the new variable $q = e/r^2$. As is well known, $\sum_{q \leq x} q^{-1} = \log x + O(1)$, so

$$\begin{aligned} A &= \sum_{r \leq \sqrt{n}} \frac{\mu(r)}{r^2} \left(\log \frac{n}{r^2} + O(1) \right) \\ &= \log n \sum_{r \leq \sqrt{n}} \frac{\mu(r)}{r^2} - 2 \sum_{r \leq \sqrt{n}} \frac{\mu(r) \log r}{r^2} + O\left(\sum_{r \leq \sqrt{n}} \frac{|\mu(r)|}{r^2} \right). \end{aligned} \quad (7)$$

Now $|\mu(r)| \leq 1$, and $\sum_1^\infty r^{-2} \log r$ converges, so the second and third terms in the last expression are $O(1)$. For the first term, we need

$$\sum_{r=1}^{\infty} \frac{\mu(r)}{r^2} = \frac{6}{\pi^2}. \quad (8)$$

This can be proved as follows:

$$\sum_{r=1}^{\infty} \frac{\mu(r)}{r^2} \sum_{s=1}^{\infty} \frac{1}{s^2} = \sum_{n=1}^{\infty} \frac{\sum_{d|n} \mu(d)}{n^2} = 1 \quad (9)$$

and, of course, $\sum_{s=1}^{\infty} s^{-2} = \pi^2/6$. Moreover, $|\sum_{r>\sqrt{n}} r^{-2}\mu(r)| < \sum_{r>\sqrt{n}} r^{-2} = O(n^{-1/2})$, so $A = (6/\pi^2) \log n + O(1)$. It follows that $P(n) = (6/\pi^2) n \log n + O(n)$, which proves Theorem 1.

Second proof of Theorem 1

To illustrate a few more techniques, here is a second proof, starting from the formula $f(n) = \sum_{r^2|n} \phi(r)$ given earlier. We get

$$\begin{aligned} F(n) &= \sum_{m=1}^{n} \sum_{r^2|m} \phi(r) \\ &= \sum_{r\le\sqrt{n}} \sum_{q\le n/r^2} \phi(r) \\ &= \sum_{r\le\sqrt{n}} \left[\frac{n}{r^2}\right] \phi(r) \\ &= \sum_{r\le\sqrt{n}} \left(\frac{n}{r^2} + O(1)\right) \phi(r) \\ &= n \sum_{r\le\sqrt{n}} \frac{\phi(r)}{r^2} + O\left(\sum_{r\le\sqrt{n}} \phi(r)\right). \end{aligned} \tag{10}$$

As in the first proof, we changed the order of summation and introduced the new variable $q = m/r^2$. Trivially, $\phi(r) \le r$, so the second sum in the last expression is $O(n)$. All we need now is an estimate for $\sum_{r\le\sqrt{n}} r^{-2}\phi(r)$.

If p is prime, then $\phi(p^e) = p^e - p^{e-1} = p^e(1 - p^{-1})$, since the only numbers not exceeding p^e and not relatively prime to it are the multiples of p, of which there are p^{e-1}. Also, ϕ is multiplicative — a proof can be found in any introductory number theory text — so $\phi(r) = r \prod_{p|r}(1 - p^{-1}) = r \sum_{d|r} d^{-1}\mu(d)$. We have

$$\sum_{r\le\sqrt{n}} \frac{\phi(r)}{r^2} = \sum_{r\le\sqrt{n}} \sum_{d|r} \frac{\mu(d)}{dr} = \sum_{d\le\sqrt{n}} \frac{\mu(d)}{d^2} \sum_{q\le\sqrt{n}/d} \frac{1}{q}, \tag{11}$$

where yet again we have interchanged the order of summation and introduced a new variable q given by $q = r/d$. The last expression can be processed the same way as the last expression in (6), leading to $\sum_{r\le\sqrt{n}} r^{-2}\phi(r) = (3/\pi^2)\log n + O(1)$. So $F(n) = (3/\pi^2)n\log n + O(n)$, and $S(n) = (6/\pi^2)n\log n + O(n)$, as was to be proved.

Numerics

Table 1 gives, for $n = 1000, 2000, \ldots, 10000$, the values of $S(n)$, $(6/\pi^2)n\log n$ (rounded to the nearest integer), and $E(n) = S(n) - (6/\pi^2)n\log n + 0.86n$ (again, rounded). The constant 0.86 was chosen by eye to make $E(n)$ consistently small for the given values of n; the methods of this paper are too crude to shed any light on the error term. We computed these values of $S(n)$ from the table of values of $f(n)$ in [2].

Table 1.

n	1000	2000	3000	4000	5000	6000	7000	8000	9000	10000
$S(n)$	3344	7496	11996	16748	21544	26552	31780	36872	42068	47296
$(6/\pi^2)n\log n$	4199	9242	14602	20169	25889	31732	37677	43708	49816	55992
$E(n)$	5	–26	–26	19	–45	–20	123	44	–8	–96

Envoi

We encourage the reader to solve the considerably simpler problem of computing the number of trifectas in arithmetic progression.

Addendum

Steven Finch and Pascal Sebah inform me by email that the method of Selberg and Delange can be used to show that $S(n) = (6/\pi^2)n\log n + \big(C + o(1)\big)n$ with $C = -2 + (18\gamma - 6)\pi^{-2} - 72\zeta'(2)\pi^{-4} = -.86222249\ldots$. Here, γ is Euler's constant, defined as $\lim_{n\to\infty}(\sum_{m=1}^{n} m^{-1} - \log n)$.

References

[1] Finch, S. and Sebah, P. (2006). Squares and cubes modulo n. http://xxx.adelaide.edu.au/abs/math/0604465 (accessed 3 June 2008).
[2] Sloane, N.J.A. (2007). The on-line encyclopedia of integer sequences. http://www.research.att.com/~njas/sequences/ (accessed 3 June 2008).