the squares which can be written as a sum of two positive squares are those with a prime factor congruent to 1 mod 4. We see that among positive squares, $(2^k)^2$ and $(5 \times 2^k)^2$ are the only ones which cannot be written as a sum of three positive squares and that 1 and 9 are the only ones which cannot be written as a sum of four positive squares.

Combining these conditions, we learn that with the exception of $(5 \times 2^k)^2$, a square can be written as sums of 2, 3, and 4 positive squares if and only if it has at least one prime factor congruent to 1 mod 4. Moreover such a square n can be written as a sum of k positive squares for all k from 1 to $n - 14$.

The first few squares meeting the combined conditions are 169, 225, 289, 625, 676, 841, 900. Going out a little farther we find $n = 1\,000\,002\,000\,001 = (101 \times 9901)^2$ with 101 being a prime congruent to 1 mod 4. So this square can be written as a sum of k positive squares for all k from 1 to 1 000 001 999 987, making 169's run of 155 look not so special after all.

REFERENCES

1. Leonard Eugene Dickson, *History of the Theory of Numbers, Vol. II: Diophantine Analysis*, Dover, New York, 2005.
2. Owen Fraser and Basil Gordon, On representing a square as the sum of three squares, *Amer. Math. Monthly* **76** (1969) 922–923. doi:10.2307/2317949
3. E. Grosswald, A. Calloway, and J. Calloway, The representation of integers by three positive squares, *Proc. Amer. Math. Soc.* **10** (1959) 451–455. doi:10.2307/2032865
4. Adolf Hurwitz, *Mathematische Werke. Bd. II: Zahlentheorie, Algebra und Geometrie*, Birkhäuser Verlag, Basel, 1963.
5. Kelly Jackson, Francis Masat, and Robert Mitchell, Extensions of a sums-of-squares problem, *Math. Mag.* **66** (1993) 41–43.
6. Carlos J. Moreno and Samuel S. Wagstaff, Jr., *Sums of Squares of Integers*, Chapman & Hall/CRC, Boca Raton, FL, 2006.
7. Gordon Pall, On sums of squares, *Amer. Math. Monthly* **40** (1933) 10–18. doi:10.2307/2301257
8. Don Redmond, *Number Theory: An Introduction*, Marcel Dekker, New York, 1996.

Summary This note shows that with the exception of $(5 \times 2^k)^2$, an integer square can be written as sums of 2, 3, and 4 positive squares if and only if it has at least one prime factor congruent to 1 mod 4. Moreover such a square n can be written as a sum of k positive squares for all k from 1 to $n - 14$. The question of when a non-square can be written as a sum of k positive squares is also examined.

How Fast Will We Lose?

RON HIRSHON
College of Staten Island
Staten Island, NY 10314
rhirshon@aol.com

Two players X and Y play a gambling game. They start with bankrolls of x and y dollars respectively, where x and y are positive integers and $(x, y) \neq (1, 1)$. They repeatedly flip a coin, which may be a fair or unfair coin. When heads appears, X wins and receives one dollar from Y; when tails appears, X loses and pays one dollar to Y.

Math. Mag. **83** (2010) 213–218. doi:10.4169/002557010X494869.

The game continues until one player runs out of money. Let L be the event that X loses the match; that is, that it is X who ends the game with a zero balance.

We assume that the flips are independent. We write p for the probability that X wins a given flip, and we always write q for $1-p$. Then the probability that X loses is

$$\mathbf{Pr}(L) = q^x \frac{p^y - q^y}{p^{x+y} - q^{x+y}} \quad (p \neq q); \quad \mathbf{Pr}(L) = \frac{y}{x+y} \quad \left(p = q = \frac{1}{2}\right). \tag{1}$$

This is a well-known formula. Our gambling game is called "gambler's ruin," and can also be described as a random walk on the integers with two absorbing barriers. A classical reference is Feller [**1**], chapters III and XIV; see especially equations (3.4) and (3.5) in section XIV.3. The theory goes back over 300 years, and early investigators include Huygens, DeMoivre, Monmart, and two Bernoullis. A good source, both for history and results, is Takács [**4**]. Formula (1) is also used in [**3**], for which this paper is a sequel.

In this paper, we study the probability of the event L_n that X loses in *exactly n* flips. DeMoivre calculated this probability in 1718, but his formula was quite complicated; see [**4**], equations (13) and (12). Our goal is to give a simple method for finding these probabilities. As explained in the last section of [**3**], this will involve the parallel goal of counting the number $c_n = c_n(x, y)$ of different sequences of H and T of length n that lead to losing in exactly n flips. Also, given that X loses, we determine the expected time it will take to lose.

For X to go broke, X must lose x more coin flips than X wins. Thus, for some integer $k \geq 0$, the sequence consists of $x+k$ tails and k heads. The probability of each such sequence is $q^{x+k}p^k$, and the number of such sequences is c_{x+2k}. Thus $\mathbf{Pr}(L_{x+2k}) = c_{x+2k}q^{x+k}p^k$. If n is not of the form $x+2k$, then $\mathbf{Pr}(L_n) = 0$. Therefore

$$\mathbf{Pr}(L) = \sum_{k=0}^{\infty} c_{x+2k}q^{x+k}p^k. \tag{2}$$

As noted in [**3**], the numbers c_{x+2k} are the coefficients for the power series of a certain rational function $g = g_{x,y}$. This means that g is a *generating function* for the sequence $\{c_{x+2k}\}$, $k = 0, 1, 2, \ldots$.

First we rewrite equation (1) using $S_n = p^{n-1} + p^{n-2}q + \cdots + pq^{n-2} + q^{n-1}$, which is positive for $0 \leq p \leq 1$. Observe that

$$S_n = \frac{p^n - q^n}{p - q} \quad (p \neq q) \quad \text{and} \quad S_n = \frac{n}{2^{n-1}} \quad \left(p = q = \frac{1}{2}\right). \tag{3}$$

It follows that

$$\mathbf{Pr}(L) = q^x \frac{S_y}{S_{x+y}} \quad \text{for} \quad 0 \leq p \leq 1; \tag{4}$$

to see this, for $p \neq q$ divide the numerator and denominator in (1) by $p - q$, and for $p = \frac{1}{2}$, note that

$$q^x \frac{S_y}{S_{x+y}} = \left(\frac{1}{2}\right)^x \frac{y}{2^{y-1}} \cdot \frac{2^{x+y-1}}{x+y} = \frac{y}{x+y}.$$

LEMMA. *The expression S_n may be expressed as a polynomial in $u = pq$ with integer coefficients.*

Proof. For $n = 1$ or $n = 2$, (3) reduces to 1 so that $S_1 = S_2 = 1$. Since

$$p^{n+1} - q^{n+1} = p^n p - q^n q = p^n(1-q) - q^n(1-p)$$
$$= (p^n - q^n) - pq(p^{n-1} - q^{n-1}),$$

for $p \neq q$ and $n \geq 2$ we have from (3) that

$$S_{n+1} = S_n - pqS_{n-1} = S_n - uS_{n-1}. \tag{5}$$

This identity also holds for $p = \frac{1}{2}$, which can be verified directly or by using a continuity argument. The lemma follows by induction. ■

Iterating (5), we obtain the sample calculations summarized in TABLE 1.

TABLE 1

S_3	S_4	S_5	S_6
$1-u$	$1-2u$	$1-3u+u^2$	$1-4u+3u^2$

Set the expressions in (2) and (4) for $\mathbf{Pr}(L)$ equal and cancel q^x from both sides of the resulting identity. Setting $u = pq$, we obtain the identity

$$\sum_{k=0}^{\infty} c_{x+2k}u^k = \frac{S_y}{S_{x+y}}. \tag{6}$$

We write $g(u) = g_{x,y}(u)$ for the rational function $\frac{S_y}{S_{x+y}}$. From (6) and (4), we have

$$g(u) = \sum_{k=0}^{\infty} c_{x+2k}u^k \quad \text{and} \quad \mathbf{Pr}(L) = q^x g(u). \tag{7}$$

We call g the loss function of X for the parameters x and y (in the variable u), and we call the coefficients of the Maclaurin expansion in (7) the loss sequence of X for these parameters. To repeat, the first term in the loss sequence is always $c_x = 1$.

THEOREM. *Given the loss sequence c_{x+2k}, we have*

$$\mathbf{Pr}(L_{x+2k}) = c_{x+2k}q^{x+k}p^k \quad \textit{for integers} \quad k \geq 0. \tag{8}$$

Similarly, there is a win sequence d_{y+2k} for X, based on Y's loss function $g_{y,x}$, so that X's probability of winning in exactly $y + 2k$ steps is $d_{y+2k}p^kq^{y+k}$.

Note that, in the beginning, we had equation (2) but we did not know the coefficients. Equation (7) gets the power series to represent a rational function g. Now by direct means, we can obtain the rational function, then its power series, and then easily read off as many coefficients as we like. This is valid because of the uniqueness theorem for power series: If two power series agree on an interval, then their coefficients are equal.

To illustrate the Theorem, see TABLE 2. For example, from the $(x, y) = (4, 2)$ line, we conclude that $\mathbf{Pr}(L_4) = q^4$, $\mathbf{Pr}(L_6) = 4q^5p$, $\mathbf{Pr}(L_8) = 13q^6p^2$, $\mathbf{Pr}(L_{10}) = 40q^7p^3$, etc. Note also that the number of ways of losing in 22 flips is 29,524.

The loss functions in TABLE 2 were obtained using equation (6) and the results in TABLE 1. Most of the loss sequences in TABLE 2 can be verified by rewriting the

TABLE 2

x	y	Loss Function $g(u)$	Loss Sequence—first ten terms
1	2	$1/(1-u)$	1, 1, 1, 1, 1, 1, 1, 1, 1, 1
1	3	$(1-u)/(1-2u)$	$1, 1, 2, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8$
1	4	$(1-2u)/(1-3u+u^2)$	1, 1, 2, 5, 13, 34, 89, 233, 610, 1597
2	4	$(1-2u)/(1-4u+3u^2)$	1, 2, 5, 14, 41, 122, 365, 1094, 3281, 9842
5	1	$1/(1-4u+3u^2)$	1, 4, 13, 40, 121, 364, 1093, 3280, 9841, 29524
4	2	$1/(1-4u+3u^2)$	1, 4, 13, 40, 121, 364, 1093, 3280, 9841, 29524
3	3	$1/(1-3u)$	$1, 3, 3^2, 3^3, 3^4, 3^5, 3^6, 3^7, 3^8, 3^9$

loss function using partial fractions and then using the expansion $\frac{1}{1-w} = \sum_{k=0}^{\infty} w^k$. For example, for $(x, y) = (1, 3)$, we obtain

$$g(u) = \frac{1-u}{1-2u} = 1 + \frac{u}{1-2u} = 1 + \sum_{k=1}^{\infty} 2^{k-1} u^k,$$

which explains the powers of 2 in the loss sequence. The relationship $g_{2,4}(u) = ug_{4,2}(u) + \frac{1}{1-u}$ explains why the loss sequences in lines (4, 2) and (2, 4) look similar.

The sequence for $(x, y) = (1, 4)$ in TABLE 2 no doubt looks familiar. In fact, it is $1, f_1, f_3, f_5, \ldots$ where f_n is the Fibonacci sequence ($f_1 = f_2 = 1$, $f_3 = 2$, $f_4 = 3$, $f_5 = 5, \ldots$). To see this, we note that

$$f_1 + f_2 z + f_3 z^2 + f_4 z^3 + f_5 z^4 + \cdots = \frac{1}{1-z-z^2};$$

see, for example, formulas (6.116) and (6.117) in [**2**]. Also

$$f_1 - f_2 z + f_3 z^2 - f_4 z^3 + f_5 z^4 - \cdots = \frac{1}{1+z-z^2}.$$

Adding, we obtain

$$f_1 + f_3 z^2 + f_5 z^4 + \cdots = \frac{1-z^2}{(1-z^2)^2 - z^2}.$$

Replacing z^2 by u, we get for $0 < u < 1$,

$$f_1 + f_3 u + f_5 u^2 + \cdots = \frac{1-u}{1-3u+u^2},$$

and so

$$1 + f_1 u + f_3 u^2 + f_5 u^3 + \cdots = 1 + u \cdot \frac{1-u}{1-3u+u^2} = \frac{1-2u}{1-3u+u^2}.$$

The rational function on the right is the loss function $g(u)$ in TABLE 2 for $(x, y) = (1, 4)$, and we now see why the corresponding loss sequence consists of Fibonacci numbers.

We return to the power series in (7). $\mathbf{Pr}(L)$ is defined for all p between 0 and 1 inclusive. Hence, from (7), $\sum_{k=0}^{\infty} c_{x+2k} u^k$ converges for $p = \frac{1}{2}$ or $u = \frac{1}{4}$, so the radius of convergence R of the Maclaurin series of any loss function, with $(x, y) \neq (1, 1)$, obeys $\frac{1}{4} \le R < 1$. If we set $u = \frac{1}{4}$ and $p = q = \frac{1}{2}$ in (7), and if we use the second equation of (1), we obtain the following useful relation for the loss sequence:

$$\sum_{k=0}^{\infty} c_{x+2k} \left(\frac{1}{4}\right)^k = g\left(\frac{1}{4}\right) = 2^x \mathbf{Pr}(L) = \frac{2^x y}{x+y}. \tag{9}$$

This shows that, given the value of x and the loss sequence of X, the value of y is uniquely determined. As an example, suppose the loss sequence is one whose general term is 3^k, $k \ge 0$. If $x = 3$, then by using (9), y is determined by the equations $\sum_{k=0}^{\infty} (\frac{3}{4})^k = 4 = \frac{8y}{3+y}$, which has unique solution $y = 3$.

Different pairs (x, y) may yield the same loss function. For example, $(x, y) = (n, 1)$ yields the same loss function as $(x, y) = (n-1, 2)$. In each case, the common loss function is $1/S_{n+1}$. However, one can never find three distinct pairs (x, y) that have the same loss function. To see this, note that if for $n > 1$, we arrange the powers of u in the expansion of S_n in ascending order as in Table 1, the first two terms in this expansion will be

$$1 - (n-2)u. \tag{10}$$

This is easily proved by induction using the defining relation $S_{n+1} = S_n - uS_{n-1}$. Now suppose that two pairs (x, y) and (x^*, y^*) yield the same loss function, so that $S_y/S_{x+y} = S_{y^*}/S_{x^*+y^*}$ and

$$S_y S_{x^*+y^*} = S_{x+y} S_{y^*}. \tag{11}$$

First suppose that both y and y^* are greater than 1. Performing the multiplications of polynomials in (11) and using (10), we see that the start of the calculation gives

$$[1 - (y-2)u][1 - (x^* + y^* - 2)u] = [1 - (x + y - 2)u][1 - (y^* - 2)u].$$

Equating coefficients of u, we find that $x = x^*$. But then $y = y^*$ by the statement following equation (9).

Now suppose that $y = 1$, so we are investigating the case when $(x, 1)$ and (x^*, y^*) yield the same loss function. Then $S_1/S_{x+1} = S_{y^*}/S_{x^*+y^*}$ and $S_{x^*+y^*} = S_{x+1}S_{y^*}$. The same analysis as in the last paragraph leads to $x^* = x$ if $y^* = 1$, and $x^* = x - 1$ if $y^* > 1$. We are left with the case that $(x, 1)$ and $(x-1, y^*)$ give the same loss function. For $y^* = 2$ we already observed, prior to equation (10), that this happens. In general, there cannot be three such pairs $(x, 1)$, $(x-1, y^*)$, $(x-1, y^{**})$ because, as we noted after equation (9), the y value is uniquely determined by the x value and the loss sequence. Thus only one y value can go with $x - 1$ and $y^* = y^{**}$.

Finally, here are two questions that come to mind.

QUESTION 1. Can an infinite number of the loss functions have a common root?

QUESTION 2. Our main ideas are actually “probability free” in their definition. Can one give, in a manner as simple as ours, a method of determining the loss function for any (x, y) without referring to the probability result (1)?

Average time to lose As promised, we compute the expected time it will take to lose, given that we lose. If T represents the number of flips before losing, then we want the conditional expectation $E(T|L)$ and this equals

$$\frac{1}{\mathbf{Pr}(L)} \sum_{k=0}^{\infty} (x+2k) c_{x+2k} q^{x+k} p^k = \frac{xq^x g(u) + 2q^x u g'(u)}{q^x g(u)} = x + 2u\frac{g'(u)}{g(u)}.$$

For $(x, y) = (4, 2)$, we have $g'(u)/g(u) = \frac{4-6u}{3u^2-4u+1}$, so the expected number of flips is

$$4 + 2pq \cdot \frac{4 - 6pq}{3p^2q^2 - 4pq + 1}.$$

For $p = q = \frac{1}{2}$ and $u = \frac{1}{4}$, the expected time to lose is $32/3$.

Acknowledgment The author would like to express his thanks to Emeric Deutsch for reading several versions of this paper and for general advice. Special thanks are due to Ken Ross, Associate Editor, for a great deal of improvement of this paper, mathematically, historically, and stylistically.

REFERENCES

1. W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed., John Wiley, 1968.
2. R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, Addison-Wesley, 1989.
3. R. Hirshon and R. De Simone, An offer you can't refuse, *Mathematics Magazine* **81** (2008) 146–152.
4. L. Takács, On the classical ruin problems, *J. American Statistical Association* **64** (1969) 889–906. doi:10.2307/2283470

Summary In a version of gambler's ruin, players start with x and y dollars respectively, and flip coins for one dollar per flip until one player runs out of money. This is a random walk with two absorbing barriers. We consider the number of ways for the first player to lose on the nth flip, for $n = x, x + 2, \dots$. We use probabilistic arguments to construct generating functions for these quantities along with explicit methods for computing them. This paper builds on the paper by Hirshon and De Simone, *Mathematics Magazine* **81** (2008) 146–152.

More Polynomial Root Squeezing

CHRISTOPHER FRAYER
University of Wisconsin–Platteville
Platteville, WI 53818-3099
frayerc@uwplatt.edu

Suppose you're looking at the graph of a polynomial $y = p(x)$ in a java applet, with blue dots on the x-axis indicating the polynomial's roots, and red dots on the x-axis showing the positions of the critical points. Let's assume that all the roots are real and that you grab the blue dots and move them around on the x-axis. As you do this, what happens to the red dots?

This is a fair question because the roots determine the polynomial up to a constant multiple, and they determine the critical points exactly. For simplicity (and without loss of generality) we will only consider monic polynomials (that is, polynomials with leading coefficient 1).

If you move all the blue points (roots) the same amount, the whole graph just translates, and all the red dots simply move along for the ride. If you move all the roots in the same direction but by different amounts, it seems reasonable that the critical points all move in that same direction. This is in fact true, according to the Polynomial Root Dragging Theorem (see [**1**], [**3**]). But suppose you take two roots and symmetrically squeeze them closer to each other, something we call polynomial root squeezing. Then

Math. Mag. **83** (2010) 218–221. doi:10.4169/002557010X494878.