LETTER FROM THE EDITOR

The first two articles in this issue of THE MAGAZINE are statistical in nature. First, James Ciecka and Gary Skoog consider a fallacy about the present value of a life annuity, touching on contributions by Christiaan Huygens, Jan de Witt, and Edmond Halley. Then, Brett and David Hemenway answer the question: Why people die before they expect to? They explain how moments before a person's death, that person still has a positive life expectancy.

The next four articles all use geometry. Samuel Moreno and Esther García-Caballero place a regular hexagon in the complex plane to relate two equalities. Richard Mabry uses dissection to show how connecting specific points on a regular hexagon yields another one that is one-thirteenth the size of the original. Mabry's article includes an online supplement with additional visual proofs of his result. Günhan Caglayan computes the area of a regular dodecagon in two ways, one via a dissection, to yield the cotangent of 15 degrees. If successive rotations of a convex polyhedron result in a continuous periodic pattern in the place, then the polyhedron is referred to as a stamper. Jin Akiyama and Kiyoko Matsunaga prove that a polyhedron is a stamper if and only if it is an isotetrahedron.

The next batch of articles involves discrete mathematics. Alexander Yong uses combinatorics to explain linguist Joseph Greenberg's classification of languages and corrects Greenberg's 1957 mistake in combinatorial enumeration. Samuel Erickson, Adam Goyt, and Josiah Reiswig ask a Fibonacci-like question about aphids, instead of rabbits, because aphids reproduce both sexually and asexually. They determine recurrence relations and generating functions for the number of aphids under different circumstances. Jaime Gaspar provides a proof without words using trapezoids to compute triangular numbers. Using elementary means, B. J. Gardner constructs fields of order p^2 for various primes, including $p < 29$, as subrings of 2×2 matrices over $\mathbb{Z}_p$.

In the final two articles of this issue, Dan Ştefan Marinescu and Mihai Monea use convexity to connect the Chebyshev, Jensen, and Hermite-Hadamard inequalities and Christopher Bernhardt uses Markov chains in his proof of Perron's 1907 theorem on positive matrices and their eigenvalues/eigenvectors. Perron's theorem and Markov chains are two of my favorite topics.

As with previous June issues, Brendan Sullivan provides a MathFest inspired crossword. There is also another Partiti puzzle, new Reviews, and Problems and Solutions. Interspersed there is a novel proof of the divergence of the harmonic series.

This issue marks the last one with María Luisa Pérez-Seguí serving as an assistant editor for the Problems section. Although she started serving formally at the beginning of 2017, she had served on an ad hoc basis in 2016. We thank her for her service and commitment. Problems Editor Eduardo Dueñez informs me that there has been a dearth of problem proposals recently. If you have an idea for a problem for the Problems section, consider submitting it. Information on how to do so appears at the bottom of the first page of the Problems section.

Michael A. Jones, Editor

ARTICLES

Life Expectancies and Annuities: A Modern Look at an Old Fallacy

JAMES E. CIECKA
DePaul University
Chicago, Illinois 60604
jciecka@depaul.edu

GARY R. SKOOG
Legal Econometrics
Glenview, Illinois 60025
DePaul University
Chicago, Illinois 60604
gskoog@umich.edu

Suppose you own a life annuity (i.e., an annuity that pays as long as you live but stops paying upon your death). If your life annuity pays $60,000 annually, then how much should you receive if you were to "cash out" your annuity for a lump sum of money? It seems reasonable and intuitive that the fair price would be the present value of $60,000 paid annually for a period of time equal to your life expectancy (see, for example, Hacking [5]). However, perhaps the oldest inequality in mathematical economics states that a life annuity is worth less than an annuity paid with certainty to life expectancy if the interest rate is positive (hereafter referred to as the "inequality"). We use the term "fallacy" when this inequality is treated as equality. Beyond historical interest, the inequality continues to have contemporary importance; examples include personal financial planning and valuation of pensions and life care plans that frequently arise in civil litigation. Judges, attorneys, and economists must deal with such situations but some fall into the fallacy because of lack of appropriate mathematical training.

The next section of the paper describes the historical development of the fallacy and the inequality. Concepts are defined in discrete time as they developed historically, but we use modern notation. The inequality and fallacy are of historical interest because they arose during the infancy of the development of probability theory, the concepts of life expectancy and the expected present value of a life annuity being among the very first nongambling applications of mathematical statistics. We do not prove, but merely state, the inequality in discrete time. However, using insights from discrete time, we prove the inequality in continuous time, and interpret the difference between the terms of the inequality.

Historical development of the fallacy and inequality

The concept of life expectancy dates to correspondence between Christiaan Huygens and his younger brother Lodewijk [10]. In modern terminology, they essentially defined life expectancy as the expected value of a random variable measuring additional

Math. Mag. **91** (2018) 163–170. doi:10.1080/0025570X.2018.1439256
MSC: Primary 60–03, Secondary 01A45; 91B30

years of life using Christiaan's formulation of expected value first advanced in 1657. Lodewijk initiated the fraternal correspondence on August 22, 1669. As of that date, Lodewijk already had made life expectancy calculations using his brother's definition of expected value and Graunt's [4] mortality data for London residents.

Let $F_x(t) = \Pr(T_x \leq t)$ denote the discrete distribution function for a life age x failing before or at $T_x = t$ years beyond age x; the corresponding survivor function is $S_x(t) = 1 - F_x(t) = \Pr(T_x > t)$. Letting ω denote the youngest age at which there are no survivors, $F_x(0) = 0$, $F_x(\omega - x) = 1$, $S_x(0) = 1$, $S_x(\omega - x) = 0$. Lodewijk Huygens essentially defined life expectancy for a person age x as

$$e_x \equiv \sum_{t=0}^{\omega-x-1} (t + .5)[S_x(t) - S_x(t+1)]. \tag{1}$$

This definition is remarkably modern because we think of T_x as an additional-years-of-life random variable with probability for $T_x = t$ given by the change in the probability of survival (i.e., the probability of dying) between ages $x + t$ and $x + t + 1$ for a life age x. (See Ciecka [2], for more details about Lodewijk Huygens calculations and Graunt's mortality data.) With a little bit of algebra, life expectancy in equation (1) can be shown to be equivalent to the simpler expression

$$e_x = .5 + \sum_{t=0}^{\omega-x-1} S_x(t+1). \tag{2}$$

However, equation (2) only sums survival probabilities, and characteristics of additional years of life, other than the mean, are not easily calculated.

In his letter to Christiaan, Lodewijk refers to his life expectancy calculations as "useful for the compositions of life pensions." In a letter dated November 28, 1669, Christiaan made a clear distinction between life expectancy and median additional years of life. Christiaan thought life expectancy important to "set [a] life pension" and the median for "wagers." The term "wagers" is clear because Christiaan was referring to the "age to which there is equal likelihood [to] survive or not survive," but Christiaan agreed with Lodewijk that life expectancy is the important concept for a life pension. The Huygens' correspondence in 1669 marks the invention of the concept of life expectancy and the simultaneous appearance of the fallacy. It is true that the Huygens brothers did not calculate life annuities; rather, they were solely concerned with the additional-years-of-life random variable and its characteristics. However, they strayed, perhaps inadvertently, into life annuities when they spoke of a pension – a leading example of a life annuity – and identified their newly invented life expectancy concept as the appropriate number of years to "set [a] life pension." That is, had they calculated a pension, they would have used their new concept of life expectancy to value it. The fallacy could not have been corrected until 1671 at the earliest when de Witt, also using the distribution of deaths, was the first person to give a correct definition of the expected present value of a life annuity [3, 7, 8, 9].

De Witt's definition of the expected present value of a life annuity plays off the Huygens brothers' definition of life expectancy. Define an annuity with fixed term of t years at annual interest rate i as

$$a_{\overline{t|}} \equiv \sum_{j=1}^{t} (1+i)^{-j}. \tag{3}$$

De Witt's definition of the expected present value of a life annuity for a life age x is

$$a_x \equiv \sum_{t=0}^{\omega-x-1} a_{\overline{t}|}[S_x(t) - S_x(t+1)], \tag{4}$$

where we define $a_{\overline{0}|} = 0$; that is, no annuity payment is received if death occurs within one year after age x. De Witt replaced the additional-years-of-life random variable T_x in Huygens' definition of life expectancy in equation (1) with an annuity-certain random variable$a_{\overline{T_x}|}$ in equation (4). This marks the first time that interest rates and mortality probabilities were incorporated into a single formula and predates Edmond Halley's seminal paper on life annuities by 22 years. It is natural that de Witt developed his concept of a life annuity from the Huygens brothers' formulation of life expectancy. All were Dutch, were contemporaries, and de Witt and Christiaan Huygens had the same mathematics teacher and mentor.

Halley [**6**], unaware of de Witt's work, developed a view of a life annuity in a different, but mathematically equivalent, manner. Today, Halley's method is used far more commonly than de Witt's concept; but de Witt's formulation is more amenable to calculating distributional characteristics such as variance, skewness, and kurtosis of a life annuity. Halley used the simpler definition

$$a_x = \sum_{t=1}^{\omega-x-1} (1+i)^{-t} S_x(t), \tag{5}$$

which is equivalent to de Witt's definition after equation (4) is simplified by replacing $a_{\overline{t}|}$ using equation (3).

Following King [**11**], we define the present value of a fixed annuity to life expectancy as

$$a_{\overline{e_x}|} \equiv \sum_{j=1}^{n} (1+i)^{-j} + \gamma(1+i)^{-(n+1)} \tag{6}$$

where $e_x = n + \gamma$ with n being a whole number and γ a proper fraction. In 1887 King showed that

$$a_x \leq a_{\overline{e_x}|} \tag{7}$$

with the equality holding when $i = 0$ and the strict inequality holding when $i > 0$. King's proof, in discrete time, was the first proof of the inequality. Although the fallacy is well known today, at times writers still make the mistake of equating the expected present value of a life annuity to an annuity certain with term equal to life expectancy as Hacking [**5**] has done. We suspect a few sources for the persistence of the fallacy. First, thinking about future years lived as a random variable was not common in actuarial science training until the mid-1980s. Once we view additional years of life as a random variable, there is less tendency to think in terms of a specific numerical value such as life expectancy. Second, it is natural to simplify the complex situation of keeping track of several outcomes by the device of substituting the mean in their place. Since the concept of life expectancy is commonly used, it might be incorrectly applied as a shortcut when valuing a life annuity.

Figure 1 illustrates the magnitude of the fallacy for 50-year-old men and women for interest rates $0 \leq i \leq 1$ using current US mortality data [**1**]. The figure shows the difference in the two terms in the inequality expressed as a fraction of the expected present value of a life annuity [i.e., the difference between formula (6) and formula (5) divided by formula (5)]. The maximum relative difference is about 7.4% at an annual interest rate of 7.0% for men and the maximum relative difference of 5.6% at an annual interest rate of 6.0% for women. Consider a life care plan for health related purchases, which frequently arises in civil litigation, requiring a fixed yearly expenditure to the end of a person's life. At $i = .02$, the cost of such a plan is 5% greater for 50-year-old men and 4% greater for women using formula (6) than using formula (5).

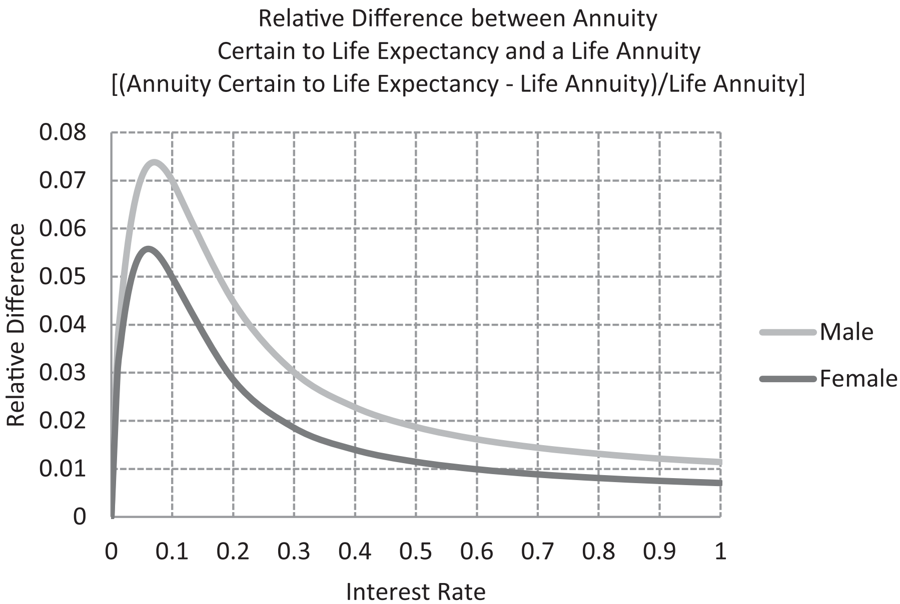


Figure 1 Relative difference between an annuity certain to life expectancy and life annuity for 50-year-old men and women.

The inequality in continuous time

Let δ denote the force of interest defined as $\delta = \ln(1+i)$, i.e., δ is the rate on a continuously compounded investment that results in the same return as i paid at the end of a year. A temporary continuous annuity paying at the annual rate of 1 discounted at δ for n years has a present value

$$\bar{a}_{\overline{n}|} \equiv \int_0^n e^{-\delta t} dt = \delta^{-1}(1 - e^{-\delta n}), \tag{8}$$

implying

$$1 = \delta \bar{a}_{\overline{n}|} + e^{-\delta n} = \delta \bar{a}_{\overline{n}|} + (1+i)^{-n}. \tag{9}$$

Now, let $F_x(t) = \Pr(T_x \le t)$ denote the continuous distribution function for a life age x failing before or at time $T_x = t$ years of life beyond age x with probability density function $f_x(t) = \frac{dF_x(t)}{dt}$; the corresponding continuous survivor function is $S_x(t) = 1 - F_x(t) = \Pr(T_x > t)$.

Define the expected present value of a continuous life annuity as

$$\bar{a}_x \equiv \int_0^\infty \bar{a}_{\overline{t}|} f_x(t) dt \quad \big[\text{de Witt-type definition but in continuous time}\big], \tag{10}$$

or equivalently using equation (8) and integrating equation (10) by parts

$$\bar{a}_x = \int_0^\infty e^{-\delta t} S_x(t)dt \quad \big[\text{Halley-type definition but in continuous time}\big]. \tag{11}$$

Continuous life time expectancy is defined as $\bar{e}_x \equiv \int_{t=0}^{\infty} S_x(t)dt$, a continuous annuity to life expectancy is $\bar{a}_{\overline{\bar{e}_x}|} \equiv \int_0^{\bar{e}_x} e^{-\delta t}dt$, and the present value of continuous insurance paying 1 at the instant of death is defined as $\bar{A}_x \equiv \int_0^\infty e^{-\delta t} f_x(t)dt$.

Consider an individual who purchases a continuous life annuity paying the annual rate δ while alive and 1 at the instant of his death. The cost will be $\delta\bar{a}_x + \bar{A}_x$. This amount is 1 as shown in equation (12) by using equation (9) and the definitions of a continuous life annuity and continuous insurance:

$$\delta\bar{a}_x + \bar{A}_x = \int_0^\infty [\delta\bar{a}_{\overline{t}|} + e^{-\delta t}]f_x(t)dt = \int_0^\infty [1]f_x(t) = 1. \tag{12}$$

This result, of course, is not new. It is the continuous counterpart of the equation $ia_x + (1 + i)A_x = 1$, where a_x is the value of a discrete life annuity paying i, and A_x is the present value of insurance paying 1 at the end of year of death, see Promislow [**12**]. The equation forms the basis for taxation in life estate and inheritance matters in which ia_x measures life estate income and $1 - ia_x$ is the remainder that has insurance value $(1 + i)A_x$.

We can now state the inequality in continuous time and specify the difference between the terms of the inequality.

Proposition (Life Annuity-Life Expectancy). *For a life age x and for force of interest $\delta > 0$, the expected present value of a continuous life annuity is less than the value of a continuous annuity to life expectancy. That is, $\bar{a}_x < \bar{a}_{\overline{\bar{e}_x}|}$. Furthermore, the difference between the terms in the inequality is related to continuous insurance by the equation $\bar{a}_{\overline{\bar{e}_x}|} - \bar{a}_x = [\bar{A}_x - e^{-\delta\bar{e}_x}]/\delta$.*

Proof. Years not lived between age x and $\bar{e}_x$ are

$$\int_0^{\bar{e}_x} [1 - S_x(t)]dt = \bar{e}_x - \int_0^{\bar{e}_x} S_x(t)dt = \int_0^\infty S_x(t)dt - \int_0^{\bar{e}_x} S_x(t)dt = \int_{\bar{e}_x}^\infty S_x(t)dt. \tag{13}$$

The integrand $[1 - S_x(t)]$ in equation (13) captures time not lived between age x and $\bar{e}_x$. Years not lived between age x and $\bar{e}_x$ also equals years lived after life expectancy as shown in the right-most integral in equation (13).

The value of a continuous life annuity can be broken into two pieces, as

$$\bar{a}_x = \int_0^{\bar{e}_x} e^{-\delta t} S_x(t)dt + \int_{\bar{e}_x}^\infty e^{-\delta t} S_x(t)dt.$$

Rewriting, we get

$$\bar{a}_x = \int_0^{\bar{e}_x} e^{-\delta t}\,\{S_x(t) + [1 - S_x(t)]\}\,dt - \int_0^{\bar{e}_x} e^{-\delta t}[1 - S_x(t)]dt + \int_{\bar{e}_x}^{\infty} e^{-\delta t} S_x(t)dt$$

$$= \int_0^{\bar{e}_x} e^{-\delta t}dt - \int_0^{\bar{e}_x} e^{-\delta t}[1 - S_x(t)]dt + \int_{\bar{e}_x}^{\infty} e^{-\delta t} S_x(t)dt$$

$$= \bar{a}_{\overline{\bar{e}_x}|} - \int_0^{\bar{e}_x} e^{-\delta t}[1 - S_x(t)]dt + \int_{\bar{e}_x}^{\infty} e^{-\delta t} S_x(t)dt$$

We have, after putting $\bar{a}_{\overline{\bar{e}_x}|}$ on the left side of the equation immediately above,

$$\bar{a}_x - \bar{a}_{\overline{\bar{e}_x}|} = -\int_0^{\bar{e}_x} e^{-\delta t}[1 - S_x(t)]dt + \int_{\bar{e}_x}^{\infty} e^{-\delta t} S_x(t)dt < 0. \tag{14}$$

The two terms on the right side of equation (14) distribute the same probability as the two extreme terms in equation (13), but the first term on the right side of equation (14) is less heavily discounted than the second term. Therefore, the right side of equation (14) is negative. This proves the inequality $\bar{a}_x < \bar{a}_{\overline{\bar{e}_x}|}$ because $\bar{a}_x - \bar{a}_{\overline{\bar{e}_x}|} < 0$.

From equations (12) and (9), we have

$$1 = \delta\bar{a}_x + \bar{A}_x = \delta\bar{a}_{\overline{\bar{e}_x}|} + e^{-\delta\bar{e}_x}, \text{ and thus}$$
$$\bar{a}_{\overline{\bar{e}_x}|} - \bar{a}_x = [\bar{A}_x - e^{-\delta\bar{e}_x}]/\delta. \qquad \square$$

We believe that the foregoing proof of the inequality and the magnitude of the difference in the terms of the inequality are new in continuous time. However, the proof uses King's [**11**] idea, but in continuous time, and Halley's definition of a life annuity in continuous time. King, in the first proof of the inequality, says that discounting has greater affect in a life annuity than an annuity certain with term equal to life expectancy. In explaining his proof, King says:

> The life annuity is, on average, equivalent to an annuity-certain of which the payment at the end of the first year is ${}_1p_x$, and at the end of the second year ${}_2p_x$, &c., and the total payments of which aggregate [to]${}_1p_x + {}_2p_x + c. + {}_kp_x = n + \gamma$. The annuity-certain is an annuity consisting of n payments of 1 each, and 1 payment of γ, and the total payments of which, therefore, also aggregate [to] $n + \gamma$. In each case, therefore, the total payment is on average the same; but in the case of the life annuity, the payments are spread over a longer period of time, and consequently are more affected by discount and have a smaller present value.

In the foregoing statement, King used the notation ${}_tp_x$ where we have used $S_x(t)$, $k = \omega - x$, and King's $n + \gamma$ has two interpretations – it is the curtate life expectancy with n denoting the integer number of years in life expectancy and γ denoting the fractional part of life expectancy and $n + \gamma$ also is the average of payments made in a life annuity. The essential idea used in our proof in continuous time is captured by King's observation that life annuity payments are "spread over a longer period of time, and consequently are more affected by discount."

We used de Witt's definition of the expected present value of a life annuity to get the exact difference $\bar{a}_{\overline{\bar{e}_x}|} - \bar{a}_x$ between the two terms in the inequality that depends on δ and the difference between continuous insurance $\bar{A}_x$ payable at the instant of death and insurance paying at the end of life expectancy $e^{-\delta \bar{e}_x}$. However, this difference in insurance values must be small, because the scaling factor δ^{-1} is large.

Although a straightforward application of Jensen's inequality could have been used to prove the inequality in either discrete or continuous time, it does not help us understand the historical context of the inequality nor does it provide an interpretation of the difference between the terms of the inequality. Our proof provides this interpretation using continuous insurance, and our previous discussion provided the historical context of the fallacy and inequality.

The fallacy still occurs today, even though it originated more than 300 years ago concurrent with early contributions to mathematical statistics by Christiaan and Lodewijk Huygens, Jan de Witt, and Edmond Halley. Current examples of the fallacy are pension plans and life care plans when valued using life expectancy (a mean) as a substitute for a set of more complex survival probabilities occurring over time. A similar fallacy arises when the expected present value of a person's future earnings is calculated using worklife expectancy (i.e., the mean of future time in the labor force) rather than using age-specific probabilities of being in the labor force. Indeed, errors or fallacies could occur whenever a single value like a mean is used in place of a wider set of outcomes.

Acknowledgment The authors wish to thank two anonymous referees, the editor, and members of the editorial board for many comments and suggestions, which led to a much improved paper.

REFERENCES

[1] Arias, E. (2014). *United States Life Tables, 2010. National Vital Statistics, 63, no. 7*. Hyattsville, MD: U.S. Department of Health and Human Services, Center for Disease Control and Prevention.

[2] Ciecka, J. E. (2011). The first probability based calculations of life expectancies, joint life expectancies, and median additional years of life. *J. Legal Econ*. 17:47–58.

[3] De Witt, J. (1671). *Value of Life Annuities in Proportion to Redeemable Annuities*. The Hague, the Netherlands: Report issued by Grand Pensionary de Witt to the States General of Holland and West Friesland. Published in Dutch with English translation in Hendricks 1852 and 1853.

[4] Graunt, J. (1662). *Natural and Political Observations Mentioned in the Following Index, and Made Upon the Bills of Mortality*. London, England: John Martin.

[5] Hacking, I. (1975). *The Emergence of Probability*. Cambridge, UK: Cambridge University Press.

[6] Halley, E. (1693). An estimate of the degrees of mortality of mankind, drawn from the curious tables of the births and funerals at the city of Breslaw; with an attempt to ascertain the price of annuities upon lives. *Philos Trans*. 17:596–610. doi:10.1098/rstl.1693.0007.

[7] Hendricks, F. (1852). Contributions of the history of insurance and the theory of life annuities, with a restoration of the Grand Pensionary de Witt's treatise on life annuities. *Assur Mag*. 2:121–150 doi:10.1017/S2046164X00000235.

[8] Hendricks, F. (1852). Contributions of the history of insurance and the theory of life annuities, with a restoration of the Grand Pensionary de Witt's treatise on life annuities (concluded from No. VI). *Assur Mag*. 2:222–258. doi:10.1017/S2046164X00000405.

[9] Hendricks, F. (1853). Contributions of the history of insurance and the theory of life annuities. *Assur Mag*. 3:93–120. doi:10.1017/S2046165800020815.

[10] Huygens, C. (1669). "The Correspondence of Huygens Concerning the Bills of Mortality of John Graunt." Extracted from Volume V of the *Oeuvres Completes* of Christiaan Huygens. For English translation, `http://cerebro.xu.edu/math/Sources/Huygens/sources/cor_mort.pdf`.

[11] King, G. (1887). *Life Contingencies*. London: Charles & Edwin Layton.

[12] Promislow, S. D. (2006). *Fundamentals of Actuarial Mathematics*. Hoboken, NJ: John Wiley & Sons.

Summary. In continuous time, we prove that the expected present value of a life annuity is less than the value of an annuity certain with term equal to life expectancy using insights from George King's proof in discrete time from the late 19th century. The concept and estimation of life expectancy and a life annuity were among the first applications of mathematical statistics using the emerging doctrine of chances in the second half of the 17th century. The paper is placed in the historical context of the rich and extraordinarily important contributions of Christiaan and Lodewijk Huygens, Jan de Witt, and Edmond Halley more than 300 years ago. A fallacy occurs when expected present value of a life annuity is thought to be equal to the value of an annuity certain with term equal to life expectancy. The fallacy appears today when pension plans and life care plans are evaluated using life expectancy (a mean) as a substitute for a set of more complex outcomes occurring over time.

JAMES E. CIECKA (MR Author ID: 1239016) is a professor of economics at DePaul University and past chairman of the department. He is the executive co-editor of the *Journal of Forensic Economics* which specializes in research applying economics theory and econometrics to litigation related matters. He received his Ph.D. in economics from Purdue University. He used to play a lot of golf; but most walking time is now spent with Newton – a delightful wheaten terrier who, contrary to golf, does not spoil a good walk.

GARY R. SKOOG (MR Author ID: 206116) heads his own litigation-related consulting firm, Legal Econometrics, and is an adjunct professor in the economics department at DePaul University. He is the past president of the National Association of Forensic Economics. He received his Ph.D. in economics from the University of Minnesota. He enjoys and is a serious student of baseball, has contributed to its sabermetrics literature, and greatly enjoys University of Michigan football.

Partiti Puzzle Solution

12 156	10 28	16 367	12 129	13 67	18 2349
19 379	4 4	5 5	12 48	5 5	9 18
15 258	7 16	12 39	8 17	14 239	10 46
7 34	7 7	10 28	11 56	8 8	12 57
6 6	5 5	13 49	10 37	10 19	2 2
23 12479	11 38	7 16	10 28	9 45	24 3678

Math. Mag. **91** (2018) 170. doi:10.1080/0025570X.2018.1446598

Why People Die Before They Expect To

BRETT HEMENWAY
University of Pennsylvania
Philadelphia, PA 19104
fbrett@cis.upenn.edu

DAVID HEMENWAY
Harvard University
Cambridge, MA 02138
hemenway@hsph.harvard.edu

The Social Security Administration (SSA) provides a handy tool [2] to help people plan for their retirement. You enter your date of birth and the SSA estimates the expected number of years remaining in your life. The tool makes a very simple calculation, it uses historical life-expectancy data to compute the conditional expectation of lifespan conditioned on having reached your current age. Thus if lifespans are modeled by a random variable, X, and you are T years old, your expected lifespan is $E(X|X > T)$.

The SSA's tool is designed to help you plan for retirement—but there is a problem. To see this, note that if you were born today, your life expectancy would be 83 years, but at age 83, your (conditional) life expectancy has increased to 90, while at 90 your (conditional) life expectancy has jumped to 95. This is just an illustration of the basic mathematical fact that the conditional additional expected lifespan is always positive.

Actually, according to the SSA's tool, your conditional additional expected lifespan is merely nonnegative rather than strictly positive. The SSA puts a hard-cap on lifespan at 120 years of age, so if you are currently 120 years old your conditional additional expected lifespan is 0. If you have not reached 120 years old, your conditional expected lifespan (according to the SSA) is strictly positive. And it is a mathematical fact, however, that for any distribution of lifespans your conditional additional expected lifespan is always nonnegative.

But this fact has an interesting consequence, at whatever age you die, you should have "expected" to live longer (well, to be precise, you should have expected to live not shorter). Mathematically, this corresponds to the fact that at the moment of your death, your conditional expected additional lifespan is positive. In other words, everyone dies before they expect to.

It makes complete sense for those of us living to recalibrate our life expectancy each year. For example, if we never recalibrated, at age 85, our life expectancy will be minus 2 years! But by recalibrating, it means that all of us will die before we reach our (conditional) life expectancy.

Using the life-expectancy tables provided by the CDC [1] it is easy to calculate each person's conditional additional expected lifespan at time of death. The answer is surprisingly large: 11.9 years! On average, moments before death you should "expect" to live 11.9 more years. In fact, recalculating can lead to extremely odd and counterintuitive results. There are natural distributions (outlined below) where the average expected lifespan at the moment of death is much larger than the expected lifespan at the moment of birth!

Consider a very simple distribution where half the population dies at 50 years old, and the remaining half dies at age 100. The expected lifespan is clearly 75 years old.

Math. Mag. **91** (2018) 171–181. doi:10.1080/0025570X.2018.1421388
MSC: Primary 62N05
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

Those unfortunates who die at age 50 expected another 25 years of life. The lucky ones who make it to age 51 now recalculate and find their conditional expected lifespan is 100 (which is exactly right). Thus half the population dies 25 years before their expectation, while the other half dies exactly at expectation, leading to an average overestimate of 12.5 years (which is one quarter of the unconditioned expected lifespan). Note that if there were no reconditioning, half the population would die 25 years "early," while the other half would die 25 years "late," leaving an average overestimate of 0.

This overestimate of lifespan occurs in this case because people are continuously recalculating their conditional expected lifespan. It is this recalculation that causes the problem. If, at birth, you predicted each person would live to the median lifespan, exactly half the population would exceed that lifespan, while the other half would "die young." On the other hand, if you allow each person to continuously recalculate their median lifespan conditioned on their current age, essentially no one will ever reach their median lifespan because the conditional median keeps moving ever further away.

This type of overestimate that occurs as a result of reconditioning can be easily seen in a simple game. Suppose you roll a fair six-sided die until a one appears, at which point you stop. The number of rolls is distributed according to the geometric distribution, and it is straightforward to calculate that the expected number of rolls before the game ends is 6. Now suppose you roll the die and your first roll is four. How many more rolls do you expect to make before the game ends? Since each roll is independent of the last, you should expect to have six more rolls until the game ends, thus the conditional expected length of the game is now seven. If your next roll is a three, you should still expect six more rolls until the game ends. In fact, the memoryless property of the die ensures that at any point, you should expect six more rolls before the game ends. In particular, at moments before you roll a one to end the game, you should still expect six more rolls before the game ends.

Why does all this matter? Because people are trying to find easy, sensible decision rules to help them make reasonable decisions about their life. What should actual actuaries and mathematicians be telling them? Currently they are being told something about their life expectancy—conditioned on their current age. But how should they act on that? Most people do not realize that none of them will reach their life expectancy—if they keep recalibrating, which makes sense for them to do.

Lifespan calculation occurs in a variety of settings, and overestimating lifespan can cause people to make poor planning decisions. How long do you have until you have to replace your car, your boiler, or the roof on your house? Deciding when to start saving for a replacement for this type of big-ticket item can be difficult, but continuously recalculating the conditional expected lifespan will likely lead you to be unprepared when the item actually fails. Even for the simple six-sided die game described above, suppose you are trying to plan your next activity. You figure it will take you a bit of time to decide on what game to play next, so you create a decision-rule that says you will start thinking about that when you have only three rolls left. If you follow this strategy, you will never start planning for the end of this game because at every point in time, you still expect to go six more rounds. If you want to be more confident that you will be prepared when the game ends, you should not recalculate.

It is interesting to compare the average lifespan with the average number of years remaining when death occurs. Can the latter ever be more than the former? You bet it can. Consider a simple steady-state world where 80% of those born die at age 1, 10% die at age 12, and 10% die at age 100. The average lifespan at birth is 12. What is the average lifespan remaining at death? For 80% of the population it is 11 years, for 10% of the population, which continues to recalibrate each year it is 44 years. The average lifespan remaining at death is 13.2 years (The 10% dying at age 100 have zero expected years of life remaining because in the example there is 100% certainty they will die at exactly age 100).

In what follows, we use mathematics to show—among other things—that: (1) for the exponential distribution (e.g., the six-sided die game described above) the average lifespan equals the average lifespan remaining at time of death; (2) when everyone dies at three distinct times, the average number of years remaining can be greater than the expected lifespan (see example above) but at most double the average lifespan at birth; when everyone dies at four distinct times, the average number of years remaining at death can be at most triple the average lifespan; (3) the Weibull distribution provides a natural lifespan distribution where the average number of years remaining at death can be arbitrarily large compared to the expected lifespan; and (4) a Gompertz-Makeham distribution for the current USA lifespan provides a very good estimate of the actual average number of years remaining when death occurs for the population.

Survival analysis

Although recalculating conditional expected years remaining will always lead to an overestimate of remaining lifespan, the exact overestimate is highly dependent on the underlying distribution.

Throughout this note, we will use the term overestimate instead of the more accurate term "not underestimate." Although it is mathematically possible for this overestimate to be zero (i.e., not be a true overestimate), this can only occur when the distribution function is a point mass, i.e., there is a fixed age at which everyone dies and there is no deviation. Eliminating this uninteresting (and usually unrealistic) case allows us to use the simpler and more intuitive term "overestimate" or "positive" instead of the cumbersome "not underestimate" or "nonnegative."

Below, we derive a formula for this overestimate, and apply this formula to a collection of distributions that are commonly used to model lifespan in a variety of settings. First, we review some of the basic mathematics that underlie the theory of survival analysis, and we give a simple characterization of the average number of years remaining at time of death (Lemma 1).

Let X be a random variable denoting the time of death. Let $f(x)$ denote the probability density function (pdf) of X, and $F(x) = \int_0^x f(t)dt$ denote the cumulative distribution function (cdf) of X. In survival analysis, the functions $S(x) \stackrel{\text{def}}{=} 1 - F(x)$ measures the probability of being alive at time x, and is termed the "survival function."

For a given distribution, we will be interested in the following question: on average, what is the conditional expected lifespan at the moment of death? Lemma 1 gives a simple formula for computing this average overestimate.

Lemma 1. *If X is a random variable denoting time of death, with pdf $f(\cdot)$ and survival function $S(\cdot)$, then the average expected lifespan at the time of death is*

$$\int_0^\infty f(t)E(X - t|X > t)dt$$

which is equivalent to

$$-\int_0^\infty xf(x)[1 + \log S(x)]dx \tag{1}$$

provided the latter integral exists. Since $0 \le S(x) \le 1$, $\log S(x) < 0$.

Proof. Upon reaching age t, the expected number of years remaining is then

$$E(X-t|X>t) = E(X|X>t) - t = \frac{1}{\Pr(X>t)}\int_t^\infty xf(x)dx - t$$
$$= \frac{1}{S(t)}\int_t^\infty xf(x)dx - t.$$

This means that the average remaining lifespan at time of death is

$$\int_0^\infty f(t)E(X-t|X>t)dt = \int_0^\infty f(t)\left[\frac{1}{S(t)}\int_t^\infty xf(x)dx - t\right]dt$$
$$= \int_0^\infty \frac{f(t)}{S(t)}\int_t^\infty xf(x)dxdx - E(X) = \int_0^\infty xf(x)\int_0^x \frac{f(t)}{S(t)}dtdx - E(X).$$

Now, the term $\frac{f(t)}{1-F(t)}$ is denoted the "hazard function," and is equivalently defined as

$$\lambda(t) \stackrel{\text{def}}{=} \lim_{\Delta t\to 0}\frac{\Pr[t \le X < t+\Delta t|X \ge t]}{\Delta t} = \frac{f(t)}{S(t)}.$$

Notice that

$$\frac{d}{dt}\log S(t) = \frac{d}{dt}\log(1-F(t)) = \frac{1}{1-F(t)}\frac{d}{dt}(1-F(t)) = \frac{-f(t)}{1-F(t)} = \frac{-f(t)}{S(t)}.$$

Thus, the hazard function can be rewritten as $\lambda(t) = -\frac{d}{dt}\log S(t)$, which yields

$$\int_0^\infty xf(x)\int_0^x \frac{f(t)}{S(t)}dtdx - E(X) = \int_0^\infty xf(x)\int_0^x -\frac{d}{dt}\log S(t)dtdx - E(X)$$
$$= \int_0^\infty -xf(x)\log S(x)dx - E(X).$$

Since $E(X) = \int_0^\infty xf(x)dx$, the above integral becomes

$$-\int_0^\infty xf(x)[1+\log S(x)]dx.$$

■

Bounds for discrete distributions

Upon noticing that the average lifespan at time of death is positive, it is natural to ask how large this overestimate can be.

The simplest distribution to understand is one supported on two points. Some fraction of the people die at time t_1, and the rest die at time t_2. Let p_1 denote the fraction of people who die at time t_1, and p_2 the fraction of those who die at time t_2. Then the average lifespan is $p_1t_1 + p_2t_2$. The people who die at time t_1, will have their lifespan overestimated by $p_1t_1 + p_2t_2 - t_1$, while the people who die at time t_2 will predict their moment of death perfectly. Thus the average overestimate is

$$p_1(p_1t_1 + p_2t_2 - t_1) = p_1(E - t_1),$$

where E is the global expectation. Because $p_1 \le 1$, and $t_1 \ge 0$, this overestimate will always be upper bounded by the (unconditional) expectation. The overestimate can be arbitrarily close to the expectation. To see this, let $p_1 = 1 - \frac{1}{n}$, $t_1 = \frac{1}{n}$, and $t_2 = n^2$. Then the expectation is $(1-\frac{1}{n})\frac{1}{n} + \frac{1}{n}n^2 = n + \frac{1}{n} - \frac{1}{n^2}$. Everyone who dies at time t_1, will have their lifespan overestimated by $E - t_1 = E - \frac{1}{n}$. Thus the average overestimate is $(1 - \frac{1}{n})(E - \frac{1}{n}) \to E$.

Is it always true that the overestimate is bounded by the expectation? In fact it is not. For distributions supported on a finite number of points, Lemma 2 gives an upper bound on the overestimate. We will see later that for continuous distributions, the overestimate can be unbounded (Lemma 3).

Lemma 2. *If X is a discrete random variable with distribution supported on n points, then the estimated number of years remaining at time of death is at most $(n-1)E(X)$.*

Proof. Suppose X is a discrete random variable distributed on n positive points $a_1 < \cdots < a_n$ such that $\Pr[X = a_i] = p_i$, where $0 < p_i < 1$, for $i = 1, \ldots, n$, where $0 < p_i < 1$, for $i = 1, \ldots, n$. When $n = 2$, the expected overestimate is $p_1(E(X) - a_1) + p_2(a_2 - a_2) = p_1(E(X) - a_1) \leq E(X)$.

Assume that for a discrete positive random variable Y distributed on $n - 1$ points, the average overestimate is at most $(n-2)E(Y)$. In general, the average overestimate is

$$\sum_{i=1}^{n} p_i \left(\frac{\sum_{j \geq i} a_i p_i}{\sum_{j \geq i} p_i} - a_i \right).$$

For a random variable Y and any $q > 0$, we have

$$\begin{aligned}
\sum_{i=1}^{n} p_i \left(\frac{\sum_{j \geq i} a_i p_i}{\sum_{j \geq i} p_i} - a_i \right) &= \sum_{i=1}^{n-1} p_i \left(\frac{\sum_{j \geq i} a_i p_i}{\sum_{j \geq i} p_i} - a_i \right) \\
&= \sum_{i=1}^{n-1} \frac{p_i}{q} \left(\frac{\sum_{j=i}^{n} a_i q \frac{p_i}{q}}{\sum_{j=i}^{n} \frac{p_i}{q}} - a_i q \right) \\
&= \sum_{i=1}^{n-1} \frac{p_i}{q} \left(\frac{\sum_{j=i}^{n-1} a_i q \frac{p_i}{q}}{\sum_{j=i}^{n} \frac{p_i}{q}} - a_i q \right) + \sum_{i=1}^{n-1} p_i \frac{a_n p_n}{\sum_{j=i}^{n} p_i} \\
&\leq \sum_{i=1}^{n-1} \frac{p_i}{q} \left(\frac{\sum_{j=i}^{n-1} a_i q \frac{p_i}{q}}{\sum_{j=i}^{n-1} \frac{p_i}{q}} - a_i q \right) + \sum_{i=1}^{n-1} p_i \frac{a_n p_n}{\sum_{j=i}^{n} p_i} \\
&\leq \sum_{i=1}^{n-1} \frac{p_i}{q} \left(\frac{\sum_{j=i}^{n-1} a_i q \frac{p_i}{q}}{\sum_{j=i}^{n-1} \frac{p_i}{q}} - a_i q \right) + (n-1) a_n p_n.
\end{aligned}$$

Now, let $q = 1 - p_n$, and let $Y = qX | X \neq a_n$, thus $\Pr[Y = qa_i] = \frac{p_i}{1-p_n} = \frac{p_i}{q}$ for all $i \neq n$. Since Y is supported on $n - 1$ points, by the induction hypothesis, we have

$$\begin{aligned}
\sum_{i=1}^{n-1} \frac{p_i}{q} \left(\frac{\sum_{j=i}^{n-1} a_i q \frac{p_i}{q}}{\sum_{j=i}^{n-1} \frac{p_i}{q}} - a_i q \right) &\leq (n-2)E(Y) \\
&= (n-2) \sum_{i=1}^{n-1} \frac{p_i}{q} \cdot G a_i = (n-2) \sum_{i=1}^{n-1} p_i a_i.
\end{aligned}$$

Thus, we have

$$\sum_{i=1}^{n} p_i \left(\frac{\sum_{j\geq i} a_i p_i}{\sum_{j\geq i} p_i} - a_i \right) \leq (n-2) \sum_{i=1}^{n-1} p_i a_i + (n-1) p_n a_n \leq (n-1) E(X). \qquad \blacksquare$$

Overestimates for specific distributions

Exponential distribution The exponential distribution is the simplest to analyze. The exponential distribution has a pdf $f(t) = \lambda e^{-\lambda t}, t \geq 0, \lambda > 0$, and thus has constant hazard rate $\lambda(t) = \lambda$. The fact that the exponential distribution has a constant hazard rate makes it "memoryless," meaning that your expected life conditioned on reaching a certain age is identical to the unconditioned expectation. Thus, $E(X|X > t) = E(X) = \frac{1}{\lambda}$, which means that at every moment your expected life remaining is exactly $E(X)$. In particular, at the moment of death, your expected remaining life is $E(X) = \frac{1}{\lambda}$. This means that the average overestimate is also $\frac{1}{\lambda}$.

To verify this, we note that for the exponential distribution

$$S(t) = e^{-\lambda t} \text{ and } \lambda(t) = \lambda.$$

In this case, equation (1) becomes

$$-\int_0^\infty x f(x)[1 + \log S(x)]dx = -\int_0^\infty x f(x)[1 - \lambda x]dx = -E(X) + \lambda E(X^2)$$
$$= -\frac{1}{\lambda} + \lambda \frac{2}{\lambda^2} = \frac{1}{\lambda} = E(X).$$

Weibull distribution The Weibull distribution generalizes the exponential distribution by adding a second "shape" parameter, k. When $k = 1$, the Weibull distribution is the exponential distribution with parameter $\frac{1}{\lambda}$ and constant failure rate λ. The mean of a Weibull distribution is $\frac{1}{\lambda}\Gamma(1 + \frac{1}{k})$. Hence, over time, the failure rate increases if $k > 1$, decreases when $k < 1$, and is constant when $k = 1$.

For the Weibull distribution, the survival and hazard rate functions are

$$S(t) = e^{-(\lambda t)^k} \quad \text{and} \quad \lambda(t) = k\lambda(\lambda t)^{k-1}.$$

In this case, equation (1) becomes

$$\begin{aligned} -\int_0^\infty x f(x)[1 + \log S(x)]dx &= -\int_0^\infty x f(x)[1 - (\lambda x)^k]dx \\ &= -E(X) + \lambda^k E(X^{k+1}) \\ &= \frac{1}{\lambda}\Gamma\left(2 + \frac{1}{k}\right) - \frac{1}{\lambda}\Gamma\left(1 + \frac{1}{k}\right) \end{aligned} \tag{2}$$

Note that when $k = 1$, the Weibull distribution is the exponential distribution with parameter $1/\lambda$, thus equation (2) becomes

$$\frac{1}{\lambda}(\Gamma(3) - \Gamma(2)) = \frac{1}{\lambda},$$

which coincides with the overestimate for the exponential distribution calculated in the last section.

The Weibull distribution also provides an example where the overestimates can be arbitrarily large (relative to the population mean). Contrast this with the case for finite distributions, where the overestimate is bounded by $(n - 1)E(X)$ (Lemma 2).

Lemma 3. *For a random variable X following Weibull distribution with parameters λ and $k = \frac{1}{n}$, the average overestimate of lifespan at time of death is $nE(X)$.*

Proof. Let X be distributed according to a Weibull distribution with parameters λ and $k = \frac{1}{n}$ for some positive integer n. Then the average overestimate of lifespan at time of death satisfies

$$\begin{aligned}\int_0^\infty f(t)\, E(X - t|X > t)dt &= \frac{1}{\lambda}\left(\Gamma\left(2 + \frac{1}{k}\right) - \Gamma\left(1 + \frac{1}{k}\right)\right)\\ &= \frac{1}{\lambda}\left(\Gamma(n+2) - \Gamma(n+1)\right) = \frac{1}{\lambda}((n+1)! - n!)\\ &= n\frac{1}{\lambda}n! = n\frac{1}{\lambda}\Gamma(n+1) = n\frac{1}{\lambda}\Gamma\left(1 + \frac{1}{k}\right) = nE(X).\end{aligned}$$

■

Another way to say this is that the average number of years remaining at time of death can be arbitrarily large relative to the average number of years remaining at birth. Formally this is stated in the following corollary.

Corollary. *For any $c > 0$, there exists a distribution where the average overestimate of lifespan is at least c times the unconditional expected lifespan, i.e.,*

$$\int_0^\infty f(t)E(X - t|X > t)dt > cE(X).$$

This is counterintuitive: the average expected lifespan at the moment of death can be much larger than the average expected lifespan at birth! Actually, Lemma 3 says something even stronger, the average *overestimate* of lifespan at time of death can be much larger than your expected lifespan at birth!

This type of overestimate occurs when there is high infant mortality. As $k \to 0$, the Weibull distribution becomes concentrated near 0 (see Figure 1).

Gompertz-Makeham distribution The Gompertz-Makeham distribution is commonly used to model lifespan. The pdf, cdf, survival, and hazard functions of this distribution are, respectively,

$$\begin{aligned}f(x) &= (\alpha e^{\beta x} + \lambda)\exp\left(-\lambda x - \frac{\alpha}{\beta}\left(e^{\beta x} - 1\right)\right),\\ F(x) &= 1 - \exp\left(-\lambda x - \frac{\alpha}{\beta}\left(e^{\beta x} - 1\right)\right),\\ S(x) &= \exp\left(-\lambda x - \frac{\alpha}{\beta}\left(e^{\beta x} - 1\right)\right), \quad \text{and}\\ \lambda(x) &= -\frac{d}{dx}\log S(x) = \alpha e^{\beta x} + \lambda,\end{aligned}$$

where $x \geq 0$ and α, β, and λ are nonnegative constants. For the Gompertz-Makeham distribution, the mean is

$$\frac{1}{\lambda}\left(1 - e^{\frac{\alpha}{\beta}}\left(\frac{\alpha}{\beta}\right)^{\frac{\lambda}{\beta}}\Gamma\left(1 - \frac{\lambda}{\beta}, \frac{\alpha}{\beta}\right)\right).$$

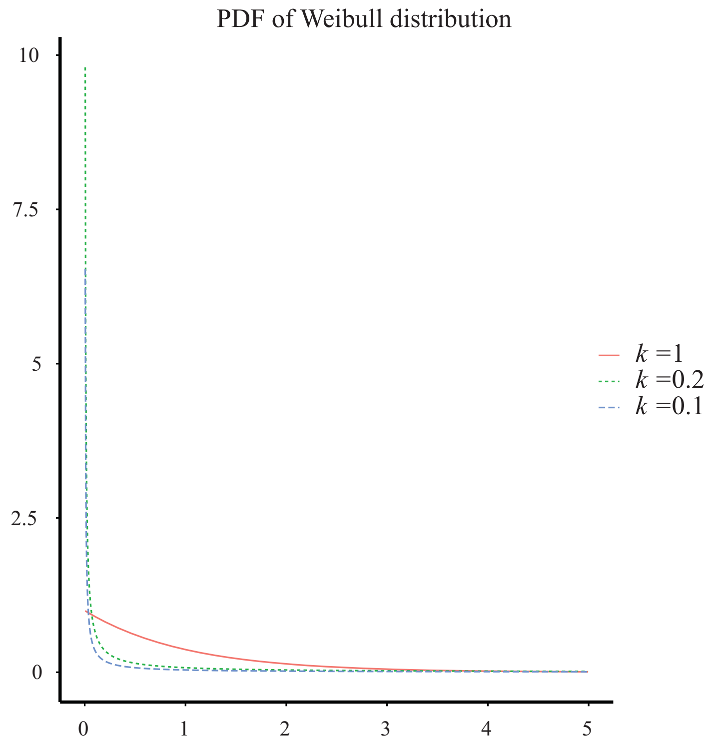


Figure 1. As the Weibull distribution concentrates near zero, the average expected lifespan at time of death increases (relative to the population mean).

The average expected lifespan at the moment of death is then given by

$$-\int_0^\infty xf(x)[1+\log S(x)]dx = -\int_0^\infty xf(x)\left[1-\lambda x-\frac{\alpha}{\beta}\left(e^{\beta x}-1\right)\right]dx$$
$$= -\left(1+\frac{\alpha}{\beta}\right)E(X)+\lambda E(X^2)+\frac{\alpha}{\beta}\int_0^\infty xf(x)e^{\beta x}dx.$$

The Gompertz-Makeham distribution provides a fairly accurate model of the actual population information provided by the CDC (See Figure 2). In Figure 2, the fitted parameters for the Gompertz-Makeham distribution are

$$\alpha = 9.661407\cdot 10^{-6}, \beta = .1073342, \text{ and } \lambda = .001218835.$$

This gives a mean age of 77.43, and the average expected lifespan at the moment of death of 12.28 years. This coincides well with the empirical distribution, which has mean 79.1 and average expected lifespan at the moment of death of 11.9 years.

Averaging predictions over a lifespan

The expected lifespan is an unbiased estimator of an individual's true lifespan, and similarly the conditional expected lifespan is an unbiased estimator of the actual conditional lifespan. We only run into trouble when we restrict our attention to the moment of death. At the moment of death, this prediction is an overestimate, and this causes the "paradox" that everyone dies before they expect. Although your expected remaining lifespan will always be an overestimate at time of death, it may be an underestimate

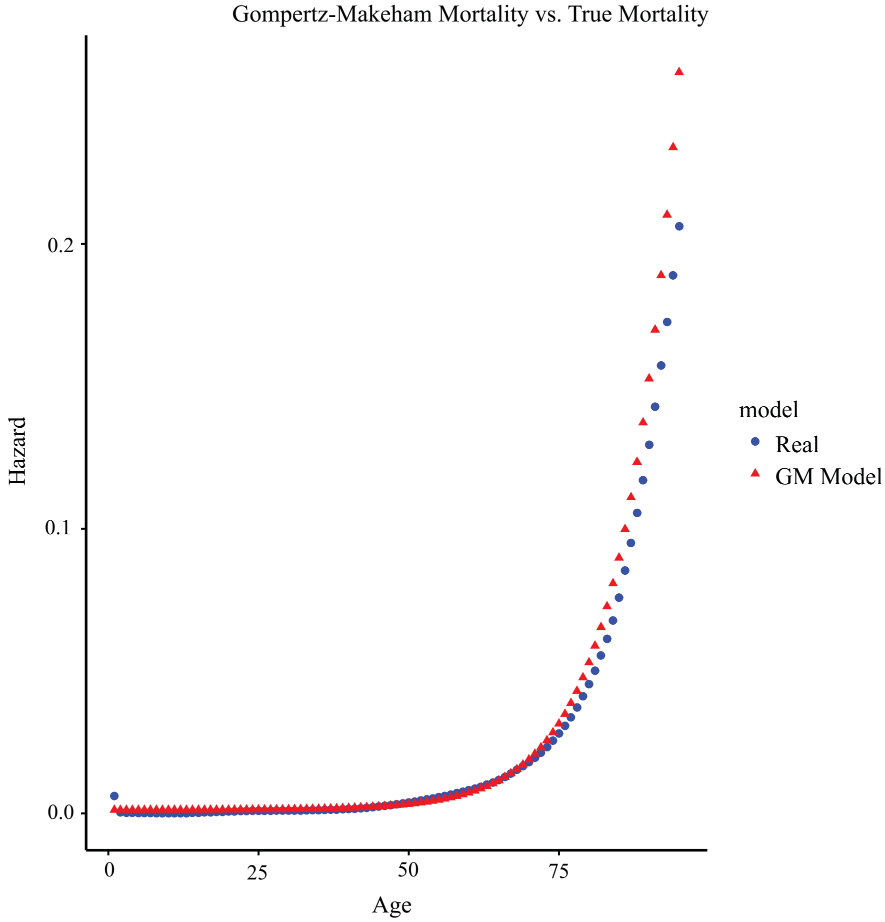


Figure 2. Fitting the Gompertz-Makeham distribution to the US mortality rates provided by the CDC.

when you are young. Instead of focusing on the expected remaining lifespan at the moment of death, we can examine these predictions at every moment along your life, and this removes the paradox.

The actuarial tables make a prediction for your lifespan at every moment of your life (conditioned on having lived to that moment). Lemma 4 shows that if we average these predictions over your lifespan, and then over the population, we obtain an accurate predictor of the population mean.

Lemma 4. *The conditional expected lifespan is an unbiased estimator of the true lifespan, when averaged first over the individual then over the population.*

Proof. Suppose you live to age T, then the average prediction error is

$$\frac{1}{T}\int_0^T (E(x|x>t)-T)\,dt = \frac{1}{T}\int_0^T \frac{1}{S(t)}\int_t^\infty xf(x)dxdt - T.$$

Now, integrating this against the distribution function, we have

$$\int_0^\infty f(T)\int_0^T (E(X|X>t)-T)\,dtdT$$

$$
\begin{aligned}
&= \int_0^\infty f(T) \int_0^T E(X|X > t)dtdT - \int_0^\infty f(T) \int_0^T TdtdT \\
&= \int_0^\infty f(T) \int_0^T E(X|X > t)dtdT - \int_0^\infty T^2 f(T)dT \\
&= \int_0^\infty f(T) \int_0^T E(X|X > t)dtdT - E(X^2) \\
&= \int_0^\infty \int_t^\infty f(T)E(X|X > t)dTdt - E(X^2) \\
&= \int_0^\infty E(X|X > t) \int_t^\infty f(T)dTdt - E(X^2) \\
&= \int_0^\infty E(X|X > t)S(t)dt - E(X^2) \\
&= \int_0^\infty \frac{1}{S(t)} \int_t^\infty xf(x)dxS(t)dt - E(X^2) \\
&= \int_0^\infty \int_t^\infty xf(x)dxdt - E(X^2) = \int_0^\infty \int_0^x xf(x)dtdx - E(X^2) \\
&= \int_0^\infty x^2 f(x)dx - E(X^2) = E(X^2) - E(X^2) = 0.
\end{aligned}
$$

■

Lemma 4 means that if a person averages all predictions received over their lifetime, then averages this over the population, they will find that this is an unbiased predictor of lifetime. Comparing Lemmas 1 and 4, we see that by restricting our attention to the moment of death, we turn an unbiased estimator into a heavily biased one.

Conclusion

The problem for normal people in the real world remains. We want to plan for when we are going to die and no one except experts in mathematics can calculate (or wants to calculate) all the predictions we receive over our entire lifespan, nor would that information be of much help in making plans. We are left with this paradoxical situation: It makes sense to recalculate your likelihood of death, but that means when you die you will die before your time. Yet if you correctly assume that you will die before your time and plan accordingly, this will mean that you often will make plans earlier than you might like, and these plans might have to change.

Although the SSA's data indicate that the average person dies 11.9 years before they expect, this effect is mitigated somewhat if we only consider those above a certain age. Since most people do not consider their own mortality until they reach a certain age, maybe we should focus on those who are likely to calculate their expected lifespan at all. If we restrict our attention to those above 40, the average overestimate is 10.4 years, for those above 50, the average overestimate is 9.7 years, and for those above 60, the average overestimate is at 8.7 years, and by 70 it is down to 7.2 years.

Fortunately, in the real world, the death of our possessions and ourselves rarely comes without any warning. It is not typically a bolt out of the blue. Instead our cars and houses and appliances start having problems—warning signs of impending doom. We are not like the wonderful one-hoss shay where all is fine and then everything goes at once. By the time we die, while our age cohort may have 12 years life expectancy, our cohort of equally sick or injured contemporaries may have only about 12 weeks left. Mathematically, this is just a simple statement that we should not recalculate our lifespan based solely on our current age, but instead we should take other warning signs into account.

Still, for all of us, it is useful to understand for planning purposes, if we correctly recalibrate, we will all die before we expect. We should try to prepare for that certain contingency.

REFERENCES

[1] Centers for Disease Control. (2010). Life table for the total population: United States. ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Publications/NVSR/63_07/Table01.xlsx.

[2] Social Security Administration. Retirement & survivors benefits: Life expectancy calculator. socialsecurity.gov/oact/population/longevity.html.

Summary. In the USA, the average life expectancy at birth is 83 years. For the average person, however, moments before their death, their remaining life expectancy is not zero, but instead is closer to 12 years. This article explains how the recalculating your conditional life expectancy can lead to large overestimates of how much time you have remaining. We examine this phenomenon for a number of natural survival distributions, and show how the expected lifespan at the moment of death can be significantly larger than the expected lifespan at birth.

BRETT HEMENWAY (MR Author ID: 863162, ORCID 0000-0002-4344-7406) obtained a Ph.D. in mathematics from UCLA in 2010. Currently, he is a research assistant professor in the Department of Computer and Information Science at the University of Pennsylvania. Dr. Hemenway's research focuses on coding theory, cryptography, and mathematical modeling. After employing a father-son team of contractors on a home renovation project, Dr. Hemenway has been actively seeking collaborations with his father. (Dr. Hemenway's sons are not yet 2 years old, and are thus currently unsuited for professional collaborations.)

DAVID HEMENWAY, Ph.D. (MR Author ID: 1201712) is a Professor of Health Policy and Director of the Harvard Injury Control Research Center. He is an economist and Professor of Health Policy at the Harvard School of Public Health. His previous article in Mathematics Magazine (Why Your Classes are Larger Than Average) was published when his wonderful son (above) was 1 year old.

Hidden in the Shape of a Regular Hexagon

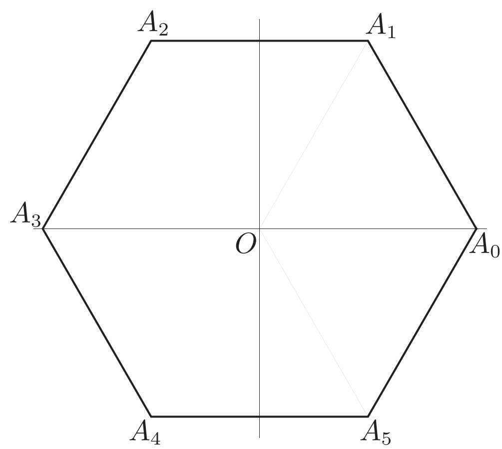
SAMUEL G. MORENO
ESTHER M. GARCÍA–CABALLERO
Universidad de Jaén
23071 Jaén, Spain
samuel@ujaen.es
emgarcia@ujaen.es

The motivation of this note is to explore a generalization of a problem posed in [**1**], and our invitation to a visual proof (in Exercise below) comes after reading [**2**].

Theorem. *If* $x + \frac{1}{x} = 1$ *and* $m \in \mathbb{Z}$, *then*

$$x^{6m+\ell} + \frac{1}{x^{6m+\ell}} = \begin{cases} (-1)^{\ell} \cdot 2 & \text{if} \quad \ell \in \{0, 3\}, \\ (-1)^{\ell-1} & \text{if} \quad \ell \in \{1, 2, 4, 5\}. \end{cases}$$

Proof. In the complex plane, a regular hexagon inscribed in a unit-radius circle has vertices $A_n \in \mathbb{C}$ as shown in the figure below. Multiplication (division) by A_1 corresponds to a $(\pi/3)$-turn counter-clockwise (resp. clockwise), and thus $A_n = (A_1)^n$ for $n \in \mathbb{Z}$, with $A_6 = A_0$, $A_7 = A_1$, $A_{-1} = A_5$ and in general $A_{6m+i} = A_i$ for all $m, i \in \mathbb{Z}$. The parallelogram law (for adding complex numbers) gives us $A_1 + 1/A_1 = A_1 + A_{-1} = A_1 + A_5 = A_0 = 1$, from which A_1 is a solution of $x + 1/x = 1$ (the other solution being $1/A_1 = A_5$).



Using again the parallelogram law we get $A_1^2 + 1/A_1^2 = A_2 + A_4 = A_3 = -1$ (case $\ell = 2$) and $A_1^3 + 1/A_1^3 = A_3 + A_3 = -2$ (case $\ell = 3$). After computing the remaining cases to complete the cycle, the result follows. ∎

Exercise. Find a proof without words of the following fact: If $x + 1/x$ is a positive integer, say n, then $x^2 + 1/x^2 = n^2 - 2$ and $x^3 + 1/x^3 = n^3 - 3n$.

Acknowledgment The authors thank the referees for their thorough reports and for fine and judicious suggestions that helped improve the presentation of this note.

Math. Mag. **91** (2018) 182–183. doi:10.1080/0025570X.2018.1426345
MSC: Primary: 51M04, Secondary: 97D50

REFERENCES

[1] Nahin, P. J. (2015). *Inside Interesting Integrals*. (Pages xi, in Preface, and 369). New York: Springer.
[2] Stack Exchange Inc. (2018). Mathematics. Questions. math.stackexchange.com/questions/1043823/x1-x-an-integer-implies-xn1-xn-an-integer.

Summary. Without actually computing the value of x in $x + 1/x = 1$ we obtain the values of $x^n + 1/x^n$, n being any integer. To this end, we use elementary facts associated with the shape of a regular hexagon.

SAMUEL G. MORENO (MR Author ID: 721728, ORCID 0000-0003-2233-2092) graduated from Universidad Autónoma de Madrid (Spain), received a M.S. from Universidad Nacional de San Luis (Argentina) under the direction of Felipe Zó, and did his doctoral studies at Universidad de Jaén (Southern Spain). He enjoys spending time with his family, practicing all kind of sports, and enjoying the sun and beach when time permits.

ESTHER M. GARCÍA–CABALLERO (MR Author ID: 661661, ORCID 0000-0001-5889-3145) received her Ph.D. in Mathematics from Universidad de Granada (Spain). She loves her family, her cat *Piruleta*, mathematics, and her kitchen for cooking baked goods (though, depending on the day, not necessarily in that order).

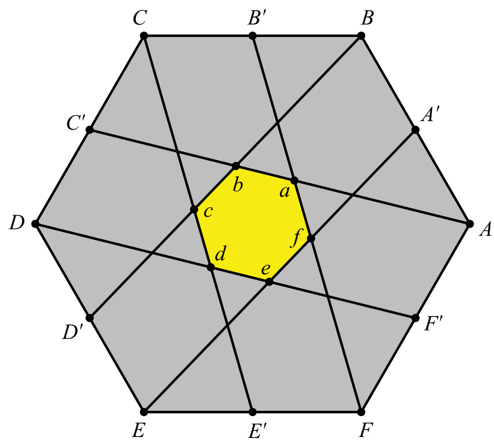
N	O	D	U	H	■	C	U	S	P	■	C	A	L	■
P	R	O	V	O	■	O	R	A	L	■	O	N	A	T
R	E	B	U	S	■	I	S	L	E	■	L	I	P	O
■	■	■	L	A	U	R	A	T	A	A	L	M	A	N
■	L	O	A	N	S	■	■	■	S	N	E	E	Z	E
E	U	G	E	N	I	A	C	H	E	N	G	■	■	■
S	N	L	■	A	N	N	O	Y	■	S	E	D	A	N
S	E	E	K	■	G	A	B	E	S	■	S	I	D	E
A	S	S	A	M	■	C	O	N	E	S	■	S	D	O
■	■	■	P	A	M	E	L	A	G	O	R	K	I	N
C	H	A	R	R	O	■	■	■	A	M	A	S	S	■
P	E	T	E	R	W	I	N	K	L	E	R	■	■	■
T	R	E	K	■	G	R	I	N	■	H	E	M	A	N
N	O	A	A	■	L	I	N	E	■	O	L	I	V	E
■	N	M	R	■	I	S	A	W	■	W	Y	N	N	E

Math. Mag. **91** (2018) 183. doi:10.1080/0025570X.2018.1456231

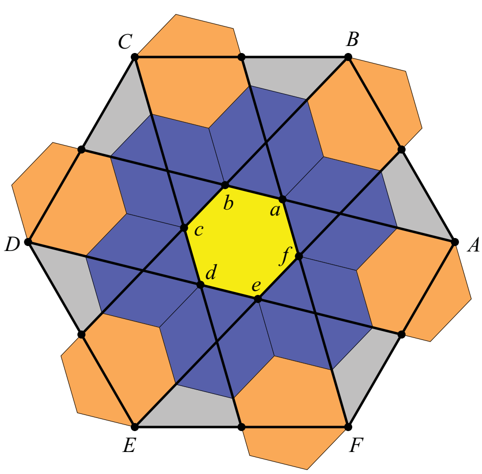
Proof Without Words: One-Thirteenth of a Hexagon

RICK MABRY
Louisiana State University Shreveport
Shreveport, LA 71115
Richard.Mabry@LSUS.edu

Let $[ABCDEF]$ be a regular hexagon. Let A' denote the midpoint of $[AB]$ et cyclica. Form the six segments $[AC']$, etc. Then a regular hexagon $[abcdef]$ is formed, where $a = [AC'] \cap [FB']$, etc., with area equal to one-thirteenth the area of $[ABCDEF]$.



Proof.



Math. Mag. **91** (2018) 184–185. doi:10.1080/0025570X.2018.1445378 © Mathematical Association of America
MSC: Primary 51M04
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.
Supplemental data for this article can be accessed on the publisher's website.

For those who like words with their proofs without words, several detailed proofs are included in an online supplement. Also included there is a generalization (with similar visual proofs) to an infinite family of related cross-cut hexagons, in which a certain collection of “Hogben’s central polygonal numbers” (of which 13 is a member) occur as the area ratios of outer to inner hexagons.

Summary. We give a visual proof that when a regular hexagon is cross-cut by six lines, each through a vertex and an opposite midpoint, then the regular hexagon bounded by these lines has one-thirteenth the area of the original hexagon.

RICK MABRY (MR Author ID: 117420, ORCID 0000-0002-2479-6416) is an emeritus professor at LSU Shreveport, where he held his only academic position for a very short quarter century. He now lives in Bavaria, where he still tries to make the occasional cut.

Partiti Puzzle

12	10	16	12	13	18
19	4	5	12	5	9
15	7	12	8	14	10
7	7	10	11	8	12
6	5	13	10	10	2
23	11	7	10	9	24

How to play. In each cell, place one or more distinct integers from 1 to 9 so that they sum to the value in the top left corner. No integer can be used more than once in horizontally, vertically, or diagonally adjacent cells. For an introduction to the Partiti Puzzle, see [Caicedo, A. E., Shelton, B. (2018). Of puzzles and partitions: Introducing Partiti. *Math. Mag.* 91(1): 20–23.]

This solution is on page 170.

—contributed by Lai Van Duc Thinh,
Vietnam; fibona2cis@gmail.com

Math. Mag. **91** (2018) 185. doi:10.1080/0025570X.2018.1445930

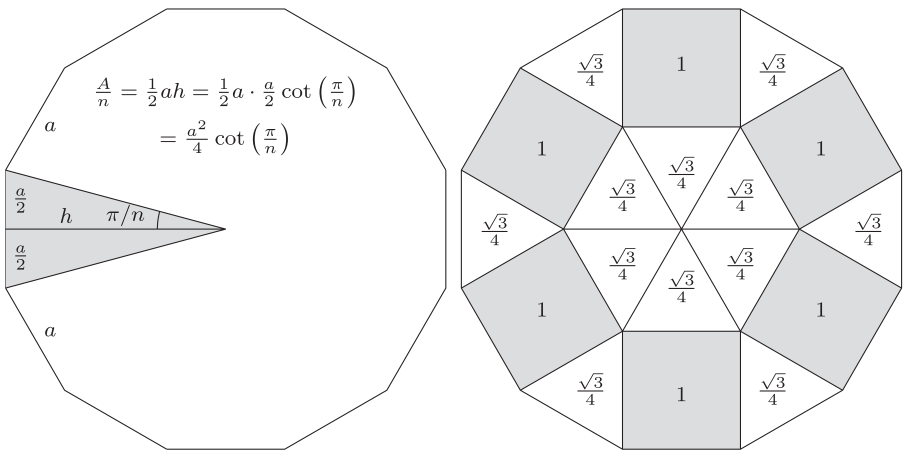
Proof Without Words: Regular Dodecagon and Cotangent of 15 Degrees

GÜNHAN CAGLAYAN
New Jersey City University
Jersey City, NJ 07305
gcaglayan@njcu.edu

The following proposition is proved by dissecting a regular dodecagon in two ways.

Proposition. $\cot(\pi/12) = 2 + \sqrt{3}$

Proof. Let $n = 12$ and $a = 1$.



$$A = \frac{na^2}{4} \cdot \cot\left(\frac{\pi}{12}\right) = \frac{12}{4} \cdot \cot\left(\frac{\pi}{12}\right) \qquad A = 6 + 12 \cdot \frac{\sqrt{3}}{4} = 6 + 3\sqrt{3}$$

$$3\cot(\pi/12) = 6 + 3\sqrt{3} \Rightarrow \cot(\pi/12) = 2 + \sqrt{3}$$

■

Summary. A regular dodecagon of unit side length partitioned into 12 regular triangles and six squares of unit side length to demonstrate the value of the cotangent of 15 deg.

GÜNHAN CAGLAYAN (MR Author ID: 1116420) teaches mathematics at New Jersey City University. His main interests are mathematical visualization and student learning through modeling and visualization.

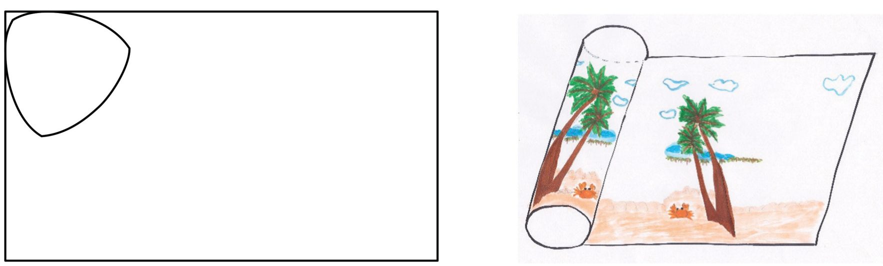
Math. Mag. **91** (2018) 186. doi:10.1080/0025570X.2018.1447186 © Mathematical Association of America
MSC: Primary 51M04, Secondary 97G40; 97G60

Only Isotetrahedra Can Be Stampers

JIN AKIYAMA
KIYOKO MATSUNAGA
Tokyo University of Science
Tokyo 162-8601, Japan
ja@jin-akiyama.com
matsunaga@mathlab-jp.com

"RULO" is a popular robot vacuum cleaner in Japan with a Reuleaux triangle-like shape. It is popular because of the convenience it brings to a household. One only needs to turn it on before leaving the house; when one gets back home, it has cleaned the whole floor including the corners (Figure 1(a)). Middle-aged Japanese have become accustomed to greeting their rulos with "*Tadaima* (I am home)" as if they are living things. Moreover, rulos are said to be kinder to the house owners than the latter's own dogs. Hence, they are taken care of like household pets by some.

One may also imaginatively think of this robot vacuum cleaner as a painter creating a masterpiece on the floor—filling each gap with colors and patterns. It is similar to a cylindrical rotary printing press in which images to be printed out are carved in the cylinder. Such a machine can create a periodic pattern without any gaps or overlaps as it continues to rotate clockwise or counterclockwise (Figure 1(b)).



(a) (b)

Figure 1 (a) RULO. (b) Rotary printing press.

Stampers

Let us now combine the ideas of that rotary printing press and a rulo, and consider embossing different patterns on each of the faces of a cube C. Brush each pattern on each face of C with ink. Then place it on paper and stamp the pattern of the base of C. Repeat this process through a succession of rotations of C around one of the edges of its square base. Even if you do it around one vertex of C, there will be some overlapping of patterns (Figure 2).

What if you do the same procedure with a regular tetrahedron T?

Emboss different patterns on each of the four faces of T, for example 1, 2, 3, and 4 (Figure 3(a)); then ink each of the embossed patterns before stamping it onto the paper. Rotate T around any of the edges of the base of T. No matter which direction you

Math. Mag. **91** (2018) 187–191. doi:10.1080/0025570X.2018.1447718 © Mathematical Association of America
MSC: Primary 52C20, Secondary 05B45; 52A10
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

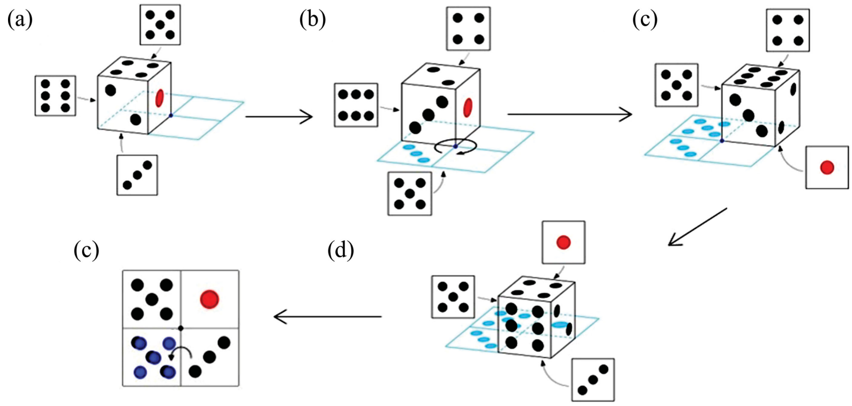


Figure 2 Stamping by rotation around a vertex of base C.

rotate T, it creates a continuous periodic pattern W that covers the plane without overlaps or any gaps (Figure 3(b)). Note that a pair of adjacent faces (for example, faces 3 and 2) of a regular tetrahedron T that share an edge (the edge ab) must touch each other in the identical orientation in the tiling W. Furthermore, around each vertex $v \in \{a, b, c, d\}$ in W a trio of faces (for example, 1, 2, and 3) of T that share the vertex v (the vertex b) must appear twice by turn and two of the same-stamped-faces are point-symmetric with regard to v. These properties uniquely determine the tiling W once an initial position is stamped by T on the plane, in spite of which route T may be rolled along. (This may be verified by mathematical induction with regard to n, where n is the number of rollings of T from the initial position).

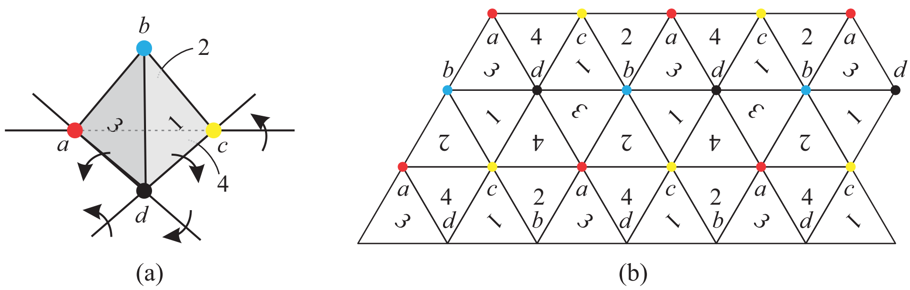


Figure 3 (a) The initial position of a stamper: The base is 4 of T. (b) Stamping a surface by rotating a tetrahedron T.

We call a convex polyhedron P a **rolling polyhedral stamper** (or **stamper** for short) if a succession of rotations of P around one of the edges of its base polygon can produce a continuous periodic pattern, containing each design embossed in each of the faces of P, that covers the plane without overlaps or any gaps.

A regular tetrahedron is a stamper but a cube is not a stamper as shown in Figure 2. What kinds of polyhedra are stampers other than a regular tetrahedron?

The main theorem

There are many strategies to solve problems [2]. One of them is to narrow down candidates for solutions by using necessary conditions and to check the sufficiency of each of the candidates. Such a strategy may be used to solve the stamper problem.

Let v be a vertex of a polyhedron P, and let $S(v)$ be the sum of the planar angles converging at v. We first show a necessary condition for a convex polyhedron to be a stamper.

Proposition 1. *A necessary condition for P to be a stamper is that for each vertex v of P, $S(v)$ is a divisor of $360°$, and it is less than $360°$.*

Proof. When a stamper is rotated along its edges in any direction around one fixed vertex, it is necessary that no overlapping occurs and that the printed designs and their orientations are properly preserved. Therefore, $S(v)$ at each vertex v of P is a divisor of $360°$, and is less than $360°$. ■

Proposition 2. *Let P be a stamper. Then, any inner angle of each of its face is less than $90°$.*

Proof. Suppose that there exists an inner angle α at a vertex v which is contained in some face F of P. We claim that $S(v) > 2\alpha$ since the faces around v do not form a part of a solid if $S(v) \leq 2\alpha$. If an inner angle α of v in F is greater than or equal to $90°$, then $S(v) > 2\alpha \geq 2 \times 90° = 180°$. This inequality violates the condition of Proposition 1. Therefore $\alpha < 90°$, completing the proof. ■

Proposition 3. *There is no stamper with n-gonal faces for $n \geq 4$.*

Proof. Any n-gon ($n \geq 4$) has at least one angle which is greater than or equal to $\frac{(n-2)\times 180°}{n} \geq 90°$. According to Proposition 2, no polyhedron with an n-gonal face ($n \geq 4$) is a stamper. ■

The next proposition follows at once from Propositions 2 and 3.

Proposition 4. *Every face of a stamper is an acute triangle.*

Proposition 5. *The number of faces of a polyhedron with all triangular faces is even.*

Proof. Let E, V, and F be the number of edges, vertices, and faces of P, respectively. A polyhedron with all triangular faces satisfies $3F = 2E$. Therefore, F must be even. ■

From Propositions 4 and 5, we have the following proposition.

Proposition 6. *For a polyhedron P to be a stamper, it is necessary that the boundary of P consists of $2n$ acute triangular faces, where $n \geq 2$.*

Proposition 7. *Every stamper has $n + 2$ vertices, $3n$ edges, and $2n$ triangular faces where $n \geq 2$.*

Proof. Recall that Euler's formula relates V, E, and F for convex polyhedra by $V - E + F = 2$. By substituting $F = 2n$, $E = 3F/2 = 3n$ into Euler's formula, it follows that

$$V = 3n - 2n + 2 = n + 2.$$

■

We need one more piece of terminology to state our main result. A tetrahedron T is called an **isotetrahedron** if all faces of T are congruent. Notice that there are infinitely many non-similar **isotetrahedra**.

Theorem. *A polyhedron is a stamper if and only if it is an isotetrahedron.*

Proof. Let us check which polyhedra with $2n$ acute triangular faces (for $n = 2, 3, 4, \ldots$) can be stampers.

Case 1: $n = 2$
Any tetrahedron T with four vertices v_1, v_2, v_3, and v_4 other than an isotetrahedron has at least one vertex

$$v_i \ (1 \le i \le 4) \text{ such that } S(v_i) > 180^\circ \text{ since } \sum_{i=1}^{4} S(v_i) = 4 \times 180^\circ.$$

Therefore, T cannot be a stamper from Proposition 2.

If T is an isotetrahedron, the equalities $S(v_1) = S(v_2) = S(v_3) = S(v_4) = 180^\circ$ hold (Figure 4), which satisfies the necessary condition from Proposition 1 for T to be a stamper. Furthermore, we can easily check that T is a stamper (Figure 5).

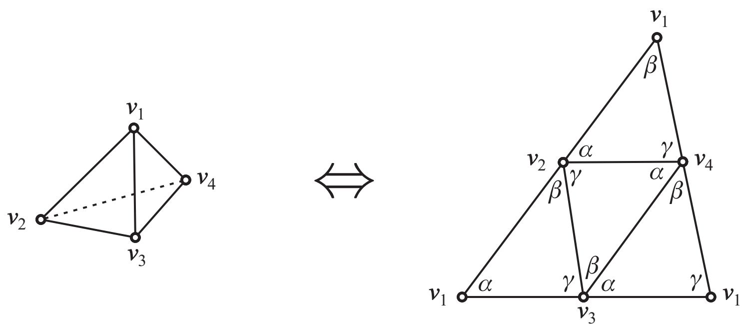


Figure 4 Isotetrahedron T with $S(v_1) = S(v_2) = S(v_3) = S(v_4) = 180^\circ$.

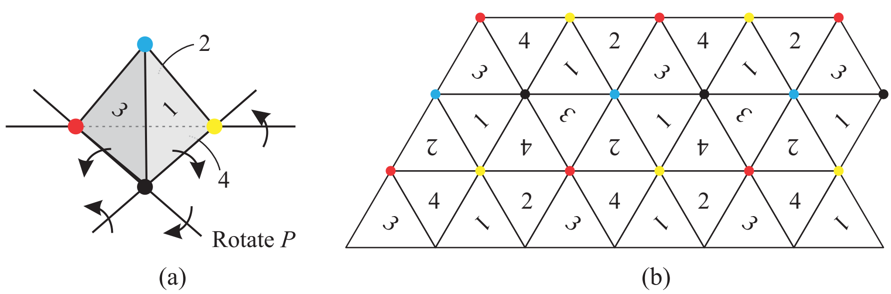


Figure 5 (a) The initial position of a stamper P: the base is 4. (b) Stamping a surface by rotating P.

Case 2: $n \ge 3$
A polyhedron with $2n$ triangular faces has $n + 2$ vertices $v_1, v_2, \ldots, v_{n+2}$ and $3n$ edges. Since $\sum_{i=1}^{n+2} S(v_i) = 180^\circ \times 2n$, there exists at least one vertex v_i such that

$$S(v_i) \ge \frac{180^\circ \times 2n}{n+2} > 180^\circ, \ (n \ge 3).$$

According to Proposition 1, any polyhedron with $2n$ triangular faces ($n \ge 3$) cannot be a stamper. Therefore, only isotetrahedra are stampers. ■

Remark. It is proved in [1] that copies of any net of an isotetrahedron tile the plane. Thus, embossing cutting lines which makes a net of an isotetrahedron T on the four faces of T, then dipping T into ink and repeatedly rolling T around any edge of the base of T onto the floor can do wonders. Such a stamper will produce various wall paper patterns on the floor like what one can see on the wall of the Alhambra Palace.

Acknowledgments We would like to thank the referees and editor for their many helpful suggestions.

REFERENCES

[1] Akiyama, J. (2007). Tile-makers and semi-tile-makers. *Amer. Math. Monthly* 114: 602–609.
[2] Larson, L. C. (1983). *Problem-Solving Through Problems*. New York, NY: Springer.

Summary. We call a convex polyhedron P a (rolling polyhedral) stamper if a succession of rotations of P around one of the edges of its base polygon can produce a continuous periodic pattern, containing each design embossed in each of the faces of P, that covers the plane without overlaps or any gaps. We prove that a polyhedron is a stamper if and only if it is an isotetrahedron.

JIN AKIYAMA (MR Author ID: 24050) studied math at Tokyo University of Science, Sophia University, and the University of Michigan. He is the founding editor of the journal *Graphs and Combinatorics*, and now a distinguished professor at his alma mater. He has authored and co-authored many books, including *Factors and Factorizations of Graphs* (jointly with Mikio Kano, Lect. Notes Math., Springer, 2011) and *A Days Adventure in Math Wonderland* (jointly with Mari-Jo Ruiz, World Scientific, 2008).

KIYOKO MATSUNAGA (MR Author ID: 1139080) is a science writer for mathematics. She studied math at Waseda University. She has contributed not only through books including *Treks into Intuitive Geometry* (jointing with Jin Akiyama, Springer, 2015), but also for NHK TV programs on mathematics, including *Math, I Like It* for elementary students, *Math Time Travel* for junior and senior high school students, and *The Joy of Math* and *Math Wonderland* for the broader public.

The Joseph Greenberg Problem: Combinatorics and Comparative Linguistics

ALEXANDER YONG
University of Illinois at Urbana-Champaign
Urbana, IL 61801
ayong@uiuc.edu

In 1957, the prominent linguist Joseph H. Greenberg (1915–2001) proposed the method of *mass comparison* (also known as *multilateral comparison*) for determining genetic relatedness between languages [**6**]. I first learned about his work through the PBS/BBC documentary "In search of the first language"; a transcript (quoted below) is found at http://www.pbs.org/wgbh/nova/transcripts/2120glang.html. At the time of writing, it is also available on Youtube.

Allowing Greenberg to summarize his idea:

> "I usually had preliminary notebooks in which I took those elements of a language, which, on the whole, we know are the most stable over time. These are things like the personal pronouns, particularly first and second person, names for the parts of the human body... I would look at a very large number of languages in regard to these matters, and I did find that they fell into quite obvious groupings."

This "presorting" technique is controversial, see, e.g., [**10**, **12**, **15**]. Its use [**9**] to see that 650 languages of North and South America fall into three families (Eskimo-Aleut, Na-Dene, and Amerind) is hotly debated. For example, James Matisoff argued:

> "Eyeballing data is prescientific, or nonscientific. There are so many ways you can be led astray, because very often, words look as if they have some connection, and they have no historical connection whatsoever."

Greenberg [**7**] offered a combinatorial argument for mass comparison. In particular, he questioned whether hypothesis testing of language groupings by extensive bilateral comparisons of languages is even feasible. Specifically, he noted that for 8 languages, there are 4,140 possible classifications to test, whereas for 25 languages there is the astronomical 4,749,027,089,305,918,018 number of possibilities.

Actually, Greenberg's critique of comprehensive relationship testing contains a straw man. In fact, many language relationships can be quickly ruled out, greatly reducing the search space. Nevertheless, the exactness of his enumeration for 25 languages makes a strong visual impression. However, in the interest of accuracy, we point out that in fact it is *slightly* erroneous: the correct enumeration is 4,638,590,332,229,999,353.

Combinatorialists will recognize the numbers Greenberg wanted to be the Bell numbers $B(n)$. Nowadays, a quick lookup at the On-Line Encyclopedia of Integer sequences (http://oeis.org) detects the discrepancy. However, since his count may be of some historical interest, we elaborate upon the correction, and how one comes to notice it.

Actually, a main purpose for this elaboration is pedagogical. The author introduces the "Joseph Greenberg problem" to introductory classes in combinatorics: *Compute*

Math. Mag. **91** (2018) 192–197. doi:10.1080/0025570X.2018.1444897
MSC: Primary 05A15, Secondary 11B73; 91F20

Greenberg's stated numbers. Thus, the discussion might be of interest to the combinatorics instructor, or to the reader who is not already versed on the topic.

While Greenberg used Bell numbers to support mass comparison, the same combinatorics does not support the plausibility of the resulting Americas classification. The probability, under the uniform distribution, of a classification of 650 languages having less than 100 families is near zero, but almost 100% for those in the range [120, 150]. Similarly, for 1000 languages, which seems like the upper bound for described languages of the Americas, the range is [170, 210]. These ranges bound, and are close to the more generally accepted viewpoint that there are between 150 and 180 families (see, e.g., [**2**], and references therein). They therefore suggest a theoretical interpretation of the observed range as, roughly, coming from the uniform distribution.

Also, most random classifications with this number of families and languages have a moderate number (9 to 19) of *language isolates*. This is somewhat consistent with the number of isolates/unclassified languages in the actual consensus classification (we seem to underestimate the number of isolates/unclassified languages by a factor of 5–10).

We contribute this combinatorial analysis to the above debate within comparative linguistics.

Multiset counting, Stirling and Bell numbers, and generating series

The 8 languages case Greenberg's calculation of 4,140 for the number of ways to classify 8 languages into families is correct. I find it helpful for the student to first determine the number using multiset counting.

Suppose that the (Native American) languages to be classified are

Alutiiq, Eyak, Cup'ik, Naukan, Mahican, Inupiaq, Tlinqit, Kalallisut

There are 22 partitions of 8. Each partition describes language family sizes. For example, in the case $5 + 2 + 1$, classifications correspond to arrangements of the multiset A, A, A, A, A, B, B, C. Hence, the arrangement A, B, A, A, C, A, B, A encodes

$$A \leftrightarrow \{\text{Alutiiq, Cup'ik, Naukan, Inupiaq, Kalallisut}\},$$
$$B \leftrightarrow \{\text{Eyak, Tlinqit}\},\ C \leftrightarrow \{\text{Mahican}\}$$

Thus, there are $\binom{8}{5\ 2\ 1} = \frac{8!}{5!2!1!}$ such arrangements.

For the partition $4 + 2 + 2$, we rearrange A, A, A, A, B, B, C, C and both

$$A, B, B, A, C, C, A, A \quad \text{and} \quad A, C, C, A, B, B, A, A$$

encode the same language groupings. As there are two repeated parts one corrects for this overcount by dividing by $\frac{1}{2!}$ so that the number of distinct groupings is $\frac{1}{2!}\binom{8}{4\ 2\ 2}$.

Similarly, for the partition $2 + 2 + 2 + 1 + 1$ the count is $\frac{1}{3!}\frac{1}{2!}\binom{8}{2\ 2\ 2\ 1\ 1}$. The reader will find it not too laborious to compute all 22 numbers of this kind, add them all up, and thus recover Greenberg's stated number. Perhaps Greenberg did somesuch calculation as a check.

The 25 languages case The method just used for the 8 language case becomes unpalatable for 25 languages because now there are 1,958 partitions. It is logical to discuss now standard combinatorics, found in textbooks such as [**1**].

Suppose $\{a_n\}_{n\in\mathbb{Z}_{\geq 0}}$ is a sequence of answers to a counting problem (such as the Joseph Greenberg problem for n languages). The uninitiated might find it surprising that one can ever get traction on a problem by considering the infinitely many such problems, and then rephrasing the question as the coefficient of $\frac{x^n}{n!}$ in the exponential generating series

$$\sum_{n=0}^{\infty} a_n \frac{x^n}{n!} = a_0 + a_1 \frac{x^1}{1!} + a_2 \frac{x^2}{2!} + a_2 \frac{x^3}{3!} + \cdots .$$

Let $S(n, k)$ be the number of ways to split n languages into exactly k language families; this is known as the Stirling number. One has the exponential generating series identity

$$\sum_{n=k}^{\infty} k! S(n, k) \frac{x^n}{n!} = (e^x - 1)^k. \tag{1}$$

This can be obtained by the product rule for exponential generating series: given

$$f(x) = \sum_{k=0}^{\infty} f_k \frac{x^k}{k!}, \quad \text{and} \quad g(x) = \sum_{\ell=0}^{\infty} g_\ell \frac{x^\ell}{\ell!},$$

the coefficient of $\frac{x^n}{n!}$ in $h(x) = f(x)g(x)$ is $\sum_{k+\ell=n} \binom{n}{k} f_k g_\ell$. This is interpreted as counting two (ordered) boxes worth of combinatorial objects:

(I) the first of the type enumerated by $f(x)$, and
(II) the second of g's type;
(III) with the labels distributed to the boxes in all possible ways.

Note that $e^x - 1$ is the series for unordered collections of languages. Thus, $(e^x - 1)^k$ is the series for classifications into k ordered families (which is a $k!$ factor overcount of $S(n, k)$).

Expanding the right-hand side of (1) by the binomial theorem, one deduces

$$S(n, k) = \frac{1}{k!} \sum_{i=0}^{k} (-1)^i \binom{k}{i} (k - i)^n, \tag{2}$$

where $S(n, 0) = 0$ (for $n \neq 0$). One can lament that while this expression is explicit and nonrecursive, it is not manifestly nonnegative or even integral, even though it computes $S(n, k) \in \mathbb{Z}_{\geq 0}$.

Greenberg's problem for n languages is computed by the Bell number defined by $B(n) = \sum_{k=0}^{n} S(n, k)$. By (2) we have a non-recursive expression for $B(n)$. However, computing $B(25)$ by hand this way is impractical.

Actually, Greenberg's footnote on p. 43 of [7] shows he knew the recurrence

$$B(n + 1) = \sum_{k=0}^{n} \binom{n}{k} B(k), \tag{3}$$

presumably used in his computation. The classroom combinatorial proof of (3) is as follows. Arbitrarily order the $n + 1$ languages first through last. Think about the number, $0 \leq k \leq n$ of other languages *not* in the same family as the last language. For each k, the number of choices is $\binom{n}{k}$. These k languages not in the same family as the last one has $B(k)$ many classifications. Hence, there are $\binom{n}{k} B(k)$ such classifications altogether.

Since the cases are disjoint, we sum each separate enumeration to obtain $B(n+1)$. Still, executing the calculation for $B(25)$ in 1957 would have been a task*.

Now, since $(e^x - 1)^k/k!$ is the generating series for $S(n, k)$, summing over all disjoint cases k of the number of families,

$$\sum_{k=0}^{\infty} \frac{(e^x-1)^k}{k!} = e^{e^x-1} \tag{4}$$

is the generating series for $B(n)$.

Often, this is where the classroom or textbook analysis stops, with the computation treated as moot. However, carrying it out, using, e.g., `Maple`, the instructor can make clear the speed one calculates the answer. For example, in `Maple` one computes it as follows:

```
> 25!*coeftayl(exp(exp(x)-1),x^25);
                                        4638590332229999353
```

In the interpretation for the Joseph Greenberg problem, this is a small surprise.

Back-of-the-envelope calculations

An advantage of generating series is that they are adaptable to related enumerations. While Greenberg used $B(25)$ to support mass comparison, similar numerics do not seem to be consistent with a key consequence, his work on languages of the Americas [9].

A simple plausibility test and interpreting the consensus range Greenberg [9] classified 650 indigenous languages of North and South America into three families*. This has been criticized for, e.g., not providing sufficient statistical evidence for claimed commonalities between languages. Specific to this situation is the use of a disputed method (mass comparison) and a large disagreement (two orders of magnitude) among linguists as to the number of families. Side-stepping well-established lines of debate, imagine a restart using combinatorics. Roughly, how many language families should there be?

Consider each classification in an unbiased way. By summing over disjoint cases, as in the derivation of (4), the generating series for language families with between a and b families is $\sum_{k=a}^{b} \frac{(e^x-1)^k}{k!}$. Using this, the probability a random classification on 650 languages has at most three families is $0.238 \times 10^{-843}\%$. (This is comparable to the probability of randomly finding a prechosen atom from the observable universe correctly, ten times in succession.) This hardly disproves Greenberg's classification—but it does quantify how improbable it is from the baseline.

The probability that the number of families are in the ranges [50, 110], [111, 120], [121, 130], [131, 140] and [141, 150] are 0.0000565%, 0.56%, 37.1%, 58.8% and 3.5% (rounded), respectively. Hence, almost all of the density is in the range [121, 150]. Thus, we would naïvely guess (taking into account margin of error) that there are a few hundred families. This is in decent agreement with the stated consensus of 150–180 families. For 1,000 languages, which seems to me to be the upper end of the number of described indigenous languages of the Americas, the model predicts the range [170, 210].

* Greenberg cites [14] that cites [5] (which has values of $B(n)$ for $n \leq 20$). By 1962, [13] gave values for $n \leq 74$ and cited [11] which had values for $n \leq 50$. Hence $B(25)$ was known at least to experts, if not widely available, by the time of [7]. Anyway, it seems likely he just computed it himself. Later, in [9] he discusses the Bell numbers again. He writes that, using a computer program, $B(20) \approx 5.172 \times 10^{10}$ (the correct value is $\approx 5.172 \times 10^{13}$); he does not restate $B(25)$ but refers the reader to [7].

* The 650 figure is found, e.g., in [3, p. 105].

These bounds on the consensus range suggest that the current “observed” range of [150, 180] families is consistent with the uniform distribution. That is, the naïve model gives a theoretical interpretation of the observed range.

Language isolates A *language isolate* is a language family consisting of only one language. Euskara, the ancestral language of the Basque people, is one such example.

If all (modern) human languages originate from a single source, then one should consider, as Greenberg does, language classifications with no isolates. Indeed, in his Americas classification, Greenberg placed many generally regarded isolates (Yuchi, Chitimacha, Tunica, among others) into his Amerind superfamily.

The exponential generating series for a box having at least two elements is

$$e^x - x - 1 = \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots . \tag{5}$$

Therefore, by the same reasoning as the derivation of the generating series (4) for $B(n)$, the generating series for classifications, where each family has at least two languages (i.e., no language isolates) is

$$\sum_{k=0}^{\infty} \frac{(e^x - x - 1)^k}{k!} = e^{e^x - x - 1}.$$

Thus, suppose we consider classifications, with no language isolates, for eight languages. The number of possibilities is the coefficient of $\frac{x^8}{8!}$ in e^{e^x-x-1} (which is 715). While this is not small, it says that the probability of a random language grouping having no language isolate is $\frac{715}{4140}$ or approximately 17%. For 25 languages, the probability is about 8.75%, whereas for 200 languages it is about 1.93%. In the case of $n = 500$ (about the number of languages analyzed in [**8**]), it is 0.927%. For 650 languages (roughly the number studied in [**9**]) it is 0.747%. In other words, most random language family configurations have a language isolate. This is not supportive of Greenberg’s hypothesis.

Suppose we are interested in the number of classifications of languages with f families and i isolates (for some fixed f and i). Imagine two boxes, the first consists of i unordered elements (the language isolates); the generating series is $\frac{x^i}{i!}$. The second room consists of collections of $f - i$ families, where each family has at least two languages (i.e., no language isolates); the generating series is $\frac{(e^x-x-1)^{f-i}}{(f-i)!}$; cf. (5). Therefore, by the product rule, the number of classifications on n languages with f families and i isolates is the coefficient of $\frac{x^n}{n!}$ in $\frac{x^i}{i!}\frac{(e^x-x-1)^{f-i}}{(f-i)!}$.

For $n = 650$, if $f = 150$, the expected number of isolates is about 9 and nearly 5% have more than 14 isolates in the right tail.

If instead $f = 180$ (the upper range of the consensus number of families), the expected number of isolates is roughly 19.5 with nearly 5% having more than 26 isolates in the right tail.

Thus, the couple of dozen isolates/unclassified languages identified in the actual classification seems somewhat high (or our estimate seems somewhat low). This would predict that a number of current isolates/unclassified languages should amalgamate, reducing the total number of families down towards 150. We refrain from a “just-so” argument for why the number of isolates seems higher than the model predicts.

What about Africa? About 1,500 languages ([**3**, p. 104]) of Africa were classified by Greenberg [**8**], using mass comparison, into four language families. Our naïve tests also reject this conclusion, estimating instead a few hundred language families.

Yet Greenberg’s classification is generally regarded as a success! However, in [**16**, p. 561] this success is qualified: “for the majority of Africa’s best documented languages." In *loc. cit.* it is noted that there is a documentation problem for less popular languages, and the total number of languages in Africa varies substantially: from 2,058 by one count, to 1,441 in another. Moreover, the review [**4**] argues that there are 19 language families, given present knowledge. Also, D. Ringe suggested to us that Africa is special because of 2,000 years of Bantu expansion that wiped out many language families. In any case, the situation is more complicated than first supposed. At this time, we merely conclude that language dispersal in Africa does not follow the same model as that for the Americas.

Concluding, the topic provides a source of classroom discussion/debate, and has room for further analysis and experimentation. The author used the same estimation methodology of this article to critique the famous bibliometric h-index; see [**17**]. Can the reader think of another study of similar type?

Acknowledgments We thank Donald Ringe for a helpful and encouraging email correspondence (and his appearance in the documentary), from which I learned much of the comparative linguistics background I have presented (any errors are my own). We also thank Eugene Lerman, Oliver Pechenik (especially for suggesting a criticism of [**9**]), Anh Yong, the Fall 2013 Math 580 class at UIUC, and the anonymous referees. The author was supported by an NSF grant.

REFERENCES

[1] Brualdi, R. (2010). *Introductory Combinatorics*. Upper Sadle River, NJ: Pearson Prentice Hall.

[2] Campbell, L. (1997). *American Indian Languages: The Historical Linguistics of Native America*. New York: Oxford University Press.

[3] Chapman, S., Routledge, C. (ed.), (2005). *Key Thinkers in Linguistics and the Philosophy of Language*. Oxford: Oxford University Press.

[4] Dimmendaal, G. (2008). Language ecology and linguistic diversity on the African continent. *Lang. Linguist. Compass* 2(5): 840–858.

[5] Epstein, L. F. (1939). A function related to the series for exp(exp$x$). *J. Math. Phys.* 18: 153–173.

[6] Greenberg, J. H. (1957). The nature and uses of linguistic typologies. *Int. J. Amer. Linguist.* 23(2): 68–77.

[7] ——— (1957). Genetic Relationship Between Languages. in "Essays in Linguistics", Chicago: University of Chicago Press, chapter 3, 35–45.

[8] ——— (1963). The Languages of Africa. *Int. J. Amer. Linguist.* 29, 1, part 2.

[9] ——— (1987). *Language in the Americas*, Stanford University Press.

[10] ——— (1993). Observations Concerning Ringe's 'Calculating the Factor of Chance in Language Comparison'. *Proc. Am. Philos. Soc.* 137(1): 79–90.

[11] Gupta, H. (1950). *Tables of Distribution*. East Punjab University, Research Bulletin, Vol. 2.

[12] Kessler, B. (2001). *The Significance of Word Lists*. Stanford: CSLI publications.

[13] Levine, J., Dalton, R. E. (1962). Minimum Periods, Modulo p, of First-Order bell exponential integers. *Math. Comput.* 16(80): 416–423.

[14] Ore, O. (1942). Theory of equivalence relations. *Duke Math. J.* 9: 573–627.

[15] Ringe, D. (1992). On calculating the factor of chance in language comparison. *Trans. Am. Philos. Soc.* 82(Part 1).

[16] Sands, B. (2009). Africa's Linguistic Diversity. *Languag. Linguist. Compass* 3(2): 559–580.

[17] Yong, A. (2014). Critique of Hirsch's citation index: A combinatorial Fermi problem. *Notice. AMS.* 61(9): 1040–1050.

Summary. We correct a 1957 combinatorial enumeration by the linguist J. Greenberg. The desired count, the Bell number $B(25)$, supported using his mass comparison method for language classification. In 1987, he used this method to classify indigenous languages of the Americas into three families. Actually, the same combinatorics provides a back-of-the-envelope estimate for the number of families. This suggests that alternative classifications with over a hundred families possess the right order of magnitude.

ALEXANDER YONG (MR Author ID: 693675) is a professor of mathematics at the University of Illinois at Urbana-Champaign. He enjoys watching documentaries on a loop; this article is one consequence. Great amusement came from communicating math to scholars appearing in "In search of the first language" and its BBC version "Before Babel." His area of research is combinatorics, usually in connection to algebraic geometry and representation theory. He and his wife Anh have two children, David and Elizabeth. He can often be found losing to his son in chess, despite his training on the mean streets of Toronto.

Aphid Sequences: Turning Fibonacci Numbers Inside Out

SAMUEL ERICKSON
ADAM M. GOYT
Minnesota State University Moorhead
Moorhead, MN 56563
ericksosam@mnstate.edu
goytadam@mnstate.edu

JOSIAH REISWIG
University of South Carolina
Columbia, SC 29208
jreiswig@math.sc.edu

As he rested in his garden after a long day of writing, Fibonacci [11] noticed a pair of rabbits peaking their heads out from behind the cabbage. Instantly, another pair of baby rabbits appeared, and as if no maturation time was needed, a new pair of baby rabbits appeared. Suddenly, Fibonacci watched as hundreds of rabbits stampeded out of his cabbage patch.

Fibonacci awoke panting, "It was just a dream. There is no way a baby rabbit could begin reproducing before it was even born." Or is there?

Luckily, Fibonacci was probably not aware of aphids. Aphids are one of the most destructive pests on farms and have a curious, although biologically common, reproductive pattern. They lay eggs in the fall, which lay dormant until the spring. In the spring the eggs hatch and until the next fall aphids reproduce asexually using parthenogenetic reproduction. That is, they give live birth to a clone of themselves. Some aphids will produce 20 to 40 generations during a summer. What is curious about parthenogenetic reproduction is that before an aphid is born it can be incubating its own offspring [4].

These telescoping generations introduce a twist on Fibonacci numbers. In the rabbit definition of Fibonacci numbers we did not have pregnant embryonic rabbits. Once born, the pair of rabbits had to undergo a maturation period before being able to create their own offspring. Thus, the initial offspring of a pair of rabbits takes longer to produce than any successive offspring, leading to the well-known recurrence relation $F_0 = 1$, $F_1 = 1$, and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$, where F_n is the number of rabbits in the nth month, i.e., the nth Fibonacci number. With aphids, the opposite is true. The aphid produces its first daughter sooner than any subsequent daughters. So we pose Fibonacci's original question but for aphids. How many aphids are there after n time periods?

In the next section, we develop the case where aphids are immortal by producing a simple recurrence relation and an associated generating function. Then, we introduce mortality and determine a generating function for these. In the final section, we determine bounds on the growth rates of the sequences.

Math. Mag. **91** (2018) 198–205. doi:10.1080/0025570X.2018.1445924
MSC: Primary 05A15, Secondary 11B39

Immortal aphids

Aphids produce offspring in two ways. The first is by laying eggs. This happens once per year, in the fall before the aphids die off for the season. During the summer season, aphids only reproduce parthenogenetically. For the sake of simplicity, we will focus on this period of parthenogenetic reproduction and let the time periods be days. Let A_n be the number of aphids alive after n days have passed. Also, let I_n be the number of aphids born on the nth day. Assume that an aphid gives birth to its first child after m days (the immature birth rate), and it gives birth to each subsequent child every M days (the mature birth rate). We will do an example to help make this clear.

Consider the case when $m = 1$ and $M = 2$, and we begin with a mature aphid, that is, an aphid whose next birth will not be her first. This gives us that $I_0 = 0$ and $A_0 = 1$. Now this aphid does not give birth the next day, so we have that $I_1 = 0$ and $A_1 = 1$. The first birth takes place on day 2, so we have $I_2 = 1$ and $A_2 = 2$. On the third day the immature aphid will give birth and will become mature. The mature aphid will not give birth. This gives us that $I_3 = 1$ and $A_3 = 3$. The first five terms of $\{I_n\}$ and $\{A_n\}$ appear in Table 1.

TABLE 1: The case where $m = 1$ and $M = 2$

Day n	Immature I_n	Mature $A_n - I_n$	Total A_n
0	0	1	1
1	0	1	1
2	1	1	2
3	1	2	3
4	2	3	5
5	3	5	8

These look like the familiar Fibonacci numbers, and they are. What is different is that the reproduction pattern of the aphids has turned the Fibonacci numbers inside out as now we have immature aphids giving birth every day and mature ones giving birth every 2 days, which is exactly the opposite of what happened with the Fibonacci numbers.

We also observe that the sequence $\{I_n\}$ looks like the Fibonacci numbers. In this case we have that $A_n = A_{n-1} + A_{n-2}$ for $n \geq 2$ with $A_0 = 1, A_1 = 1$, and $I_n = I_{n-1} + I_{n-2}$ for $n \geq 3$ with $I_0 = I_1 = 0$ and $I_2 = 1$. This is just a specific case of the general recurrence relation that we prove in Theorem 1.

Theorem 1. *Assume an aphid gives birth to its first child after m days and gives birth to each subsequent child every M days after. If A_n is the total number of aphids after n days and I_n is the total number of aphids born on day n we have the following:*

$$I_n = I_{n-m} + I_{n-M} \text{ for } n \geq M + 1 \text{ and } A_n = A_{n-m} + A_{n-M} \text{ for } n \geq M,$$

where $A_0 = A_1 = \cdots = A_{M-1} = 1$, $I_0 = I_1 = \cdots = I_{M-1} = 0$, *and* $I_M = 1$.

Proof. We begin by proving the statements about the number of immature aphids, I_n. The initial conditions hold since there are no births until the Mth day when there is one aphid born.

The number of immature aphids on day n is the number of aphids born on day n. Now, each aphid born is either the first daughter of its mother or not. If it is the first daughter of its mother, then the mother was immature m days ago, which gives us the first term. If the aphid is not the first daughter of its mother then its mother gave birth M days ago, which means that this aphid has a sister that is M days older than it. This gives us a one-to-one correspondence between the immature aphids on day n that are not the first daughter and the immature aphids on day $n - M$. Thus, $I_n = I_{n-m} + I_{n-M}$.

Now, we discuss the statements about the total number of aphids. The initial conditions hold as there is only one aphid until that aphid gives birth, which does not happen until the Mth day.

Assume now that $n \geq M$. Since the process started with a mature aphid and any aphid alive on day n was either the initial mature aphid or was born sometime in the last n days, we have $A_n = I_0 + I_1 + \cdots + I_n + 1$. There are no immature aphids until the Mth day, so we have $A_n = I_M + I_{M+1} + \cdots + I_n + 1$. Using the recurrence relation for I_n gives us that

$$\begin{aligned}
A_n &= I_M + I_{M+1} + I_{M+2} + \cdots + I_n + 1 \\
&= I_M + (I_{M+1-m} + I_{M+1-M}) + (I_{M+2-m} + I_{M+2-M}) + \cdots \\
&\qquad + (I_{n-m} + I_{n-M}) + 1 \\
&= (I_{M+1-m} + I_{M+2-m} + \cdots + I_{n-m} + 1) \\
&\qquad + (I_{M+1-M} + I_{M+2-M} + \cdots + I_{n-m} + I_M)\,.
\end{aligned}$$

Since $I_M = 1$, we have that $A_n = A_{n-m} + A_{n-M}$. ■

This is a very general Fibonacci type recurrence relation. In fact, it correlates with many generalizations studied in the past. Table 2 gives a list of some previously studied versions and their number in the Online Encyclopedia of Integer Sequences [12].

TABLE 2: Special cases of aphid sequences

m	M	Recurrence relation	Name of sequence if one exists	OEIS number
1	1	$A_n = 2A_{n-1}$	Powers of 2	A000079
1	2	$A_n = A_{n-1} + A_{n-2}$	Fibonacci sequence	A000045
1	4	$A_n = A_{n-1} + A_{n-4}$	Narayana's cow sequence	A000930
1	5	$A_n = A_{n-1} + A_{n-5}$		A003520
1	6	$A_n = A_{n-1} + A_{n-6}$		A005708
1	7	$A_n = A_{n-1} + A_{n-7}$		A005709
1	8	$A_n = A_{n-1} + A_{n-8}$		A005710
2	3	$A_n = A_{n-2} + A_{n-3}$	Padovan sequence	A000931
2	4	$A_n = A_{n-2} + A_{n-4}$	Fibonacci numbers repeated	A103609
2	5	$A_n = A_{n-2} + A_{n-5}$		A005686
2	6	$A_n = A_{n-2} + A_{n-6}$	Narayana's cows repeated	A108104
3	4	$A_n = A_{n-3} + A_{n-4}$		A079398
3	5	$A_n = A_{n-3} + A_{n-5}$		A226503
4	5	$A_n = A_{n-4} + A_{n-5}$		A103372

As you might expect, starting with an immature aphid only affects the initial conditions. To handle this case, we focus on the recurrence relation for the immature aphids.

The total number of aphids on day n is just the sum of the aphids that were born on or before day n.

Theorem 2. *Suppose that an aphid gives birth to its first child after m days and gives birth to each subsequent child every M days after that. Then the number of immature aphids, I_n, on day n is given by*

$$I_n = \begin{cases} I_{n-m} & \text{if } m \leq n \leq M \\ I_{n-m} + I_{n-M} & \text{if } M+1 \leq n \end{cases},$$

where $I_0 = 1$ and $I_n = 0$ for $1 \leq n \leq m-1$. Also, for $n \geq 0$,

$$A_n = \sum_{j=0}^{n} I_j.$$

Proof. We begin with the recurrence relation for the immature aphids. The initial conditions arise because we begin with an immature aphid and no births take place during the first $m-1$ days. The first recurrence follows because no mature aphids could have given birth yet. The second recurrence comes from the fact that mature aphids are now able to produce offspring.

Now, as in the proof of Theorem 1 the total number of aphids on day n is the sum of the aphids born on or before day n. Since we began with an immature aphid, this gives us the desired result. ■

We will now give the generating function for $\{A_n\}$ in the case where we begin with a mature aphid. The case where we begin with an immature aphid will be addressed in the next section. Letting $A(x) = \sum_{n=0}^{\infty} A_n x^n$ and applying standard generating function techniques to the recurrence relation from Theorem 1, we have that

$$A(x) = \frac{1-x^m}{(1-x^m-x^M)(1-x)}.$$

Mortal aphids

The idea of letting the rabbits die in the traditional Fibonacci sequence was first addressed by Brother U. Alfred [**1**, **2**]. Unfortunately, there was a mistake in his work and Hoggatt and Lind [**7**] gave a correct solution to this problem. Since then others have taken on dying rabbits [**5**, **10**].

Let us assume now that the aphids have a lifespan of d days. Again, let A_n be the number of aphids alive on day n and I_n be the number of aphids born on day n. It turns out that in this case it is easier to begin with an immature aphid, so each sequence will begin $I_0 = A_0 = 1$. Table 3 gives the number of aphids by age for the first 5 days for $m = 1$, $M = 2$, and $d = 5$.

One important observation here is that a mature aphid will give birth on day 5 because $m = 1$ and $M = 2$. This causes an issue because aphids are supposed to die on day 5 as well. We assume that the aphid gives birth and then dies on day 5, so its offspring is alive on day 5, but it is not. This is why the total in the last row is 12.

In the following theorem, we determine a recurrence relation for immature aphids in the case where aphids are mortal.

Theorem 3. *Suppose that an aphid gives birth to its first child after m days and gives birth to each subsequent child every M days after that. Also, suppose that an aphid*

TABLE 3: The number of aphids by age for the first 5 days when $m = 1$, $M = 2$, and $d = 5$

Day n	Immature I_n	1 day old	2 days old	3 days old	4 days old	Total A_n
0	1	0	0	0	0	1
1	1	1	0	0	0	2
2	1	1	1	0	0	3
3	2	1	1	1	0	5
4	3	2	1	1	1	8
5	5	3	2	1	1	12

dies on its dth day of life. Then the number of immature aphids, I_n, on day n is given by

$$I_n = \begin{cases} I_{n-m} & \text{if } m \le n \le M \\ I_{n-m} + I_{n-M} & \text{if } M+1 \le n \le m+kM-1\,, \\ I_{n-m} + I_{n-M} - I_{n-m-kM} & \text{if } n \ge n-m-kM \end{cases}$$

where $k = \left\lfloor \frac{d-m}{M} \right\rfloor + 1$, $I_0 = 1$, and $I_n = 0$ for $1 \le n \le m-1$.

Proof. The initial conditions and the first two recurrence relations are the same as those in Theorem 2. The first two terms of the final recurrence relation have the same meaning as the recurrence relation in Theorem 2. The last term of this recurrence relation removes the immature aphids that would have given birth on day n, but are now dead. ■

We find the total number of aphids alive on day n by summing the number of aphids born in the last d days. If we assume that $I_n = 0$ if $n < 0$ then we have the following theorem.

Theorem 4. *For $n \ge 0$, $A_n = \sum_{j=n-d+1}^{n} I_j$.*

We will use these two results to determine the generating functions for $\{A_n\}$ and $\{I_n\}$. Let $I(x) = \sum_{n=0}^{\infty} I_n x^n$. We use the third recurrence relation in Theorem 3 to start. So, we have that

$$\sum_{n=m+kM}^{\infty} I_n x^n = \sum_{n=m+kM}^{\infty} I_{n-m} x^n + \sum_{n=m+kM}^{\infty} I_{n-M} x^n - \sum_{n=m+kM} I_{n-m-kM} x^n.$$

This means that

$$I(x) - \sum_{j=0}^{m+kM-1} I_j x^j =$$
$$x^m \left(I(x) - \sum_{j=0}^{kM-1} I_j x^j \right) + x^M \left(I(x) - \sum_{j=0}^{m+(k-1)M-1} I_j x^j \right) - x^{m+kM} I(x).$$

Finally, we find that

$$(1 - x^m - x^M + x^{m+kM})I(x) = \sum_{j=0}^{m+kM-1} I_j x^j - \sum_{j=m}^{m+kM-1} I_{j-m} x^j - \sum_{j=M}^{m+kM-1} I_{j-M} x^j.$$

Now, the right-hand side can be reorganized into three sums, so we have

$$(1 - x^m - x^M + x^{m+kM})I(x) = \\ \sum_{j=0}^{m-1} I_j x^j + \sum_{j=m}^{M-1} (I_j - I_{j-m}) x^j + \sum_{j=M}^{m+kM-1} (I_j - I_{j-m} - I_{j-M}) x^j.$$

By Theorem 3, the first sum is 1, the second sum is 0, and the final sum is x^M. This gives us the generating function

$$I(x) = \frac{1 - x^M}{1 - x^m - x^M + x^{m+kM}}.$$

Now, let $A(x) = \sum_{n=0}^{\infty} A_n x^n$ be the generating function for the total number of aphids. By Theorem 4 we have that

$$A(x) = I(x) \cdot \frac{1 - x^{d+1}}{1 - x}.$$

Growing aphids

The generating function for the Fibonacci numbers is well known to be $F(x) = 1/(1 - x - x^2)$, and can be used to tell us that $F_n \sim \varphi^n$, where φ is the golden ratio. The generating functions for aphids are very similar to this, so we wanted to know what sorts of growth rates they could achieve.

We will use the generating function $A(x)$ to determine the growth rate of A_n. Since the aphid sequences consist of positive integers, the series representation of $A(x)$ has positive coefficients. Also, $A(x)$, when thought of as a complex function, is analytic at zero. This allows us to use Pringsheim's theorem [**6**], which says that if a complex function that is representable at the origin by a series expansion that has nonnegative coefficients has radius of convergence R then the point $z = R$ is a singularity. The consequence is that $A_n \sim R^n$, where $1/R$ is the smallest singularity of $A(x)$ on the positive real axis. Thus, we need only determine the smallest real positive singularity of $A(x)$.

Now, $A(x) = I(x)(1 - x^{d+1})/(1 - x)$. Since $1 - x$ is a factor of $1 - x^{d+1}$, we can focus on $I(x)$ to find the growth rate of $\{A_n\}$. Since

$$I(x) = \frac{1 - x^M}{1 - x^m - x^M - x^{m+kM}} = \frac{1}{1 - x^m - x^{m+M} - \cdots - x^{m+(k-1)M}},$$

we only need to concern ourselves with finding the smallest positive real zero of the denominator. Let

$$q(x) = 1 - x^m - x^{m+M} - \cdots - x^{m+(k-1)M},$$

then

$$q'(x) = -mx^{m-1} - (m + M)x^{m+M-1} - \cdots - (m + (k - 1)M)x^{m+(k-1)M-1},$$

which is negative for any positive value of x. Thus, we may assume that the denominator is decreasing on $[0, \infty)$. Furthermore, $q(0) = 1$ and $q(1) < 0$, so there is exactly one zero between 0 and 1. In fact, $q(1/2) > 0$, so we may conclude that the positive real zero of $q(x)$ lies between $1/2$ and 1. This tells us that for any values of m, M, and d, with $d \geq M > m$, we have that $A_n \sim \kappa^n$, where $1 \leq \kappa \leq 2$.

Can it be that $\kappa = 1$ or $\kappa = 2$ for any sequences? The answer to the first is yes and to the second is only if we let our aphids be immortal. Furthermore, we can also kill off a family of aphids if we let $d < m$. These are the results of the following theorem.

Theorem 5. *Suppose that an aphid gives birth to its first daughter after m days and gives birth to each subsequent child every M days after that. Also, suppose that an aphid dies after d days. If A_n is the total number of aphids on day n then*

(a) $\{A_n\}$ *will eventually be 0 if* $d < m$,

(b) $A_n \sim 1$ *if* $m \leq d < m + M$, *and*

(c) $A_n \sim \kappa^n$ *for some* $1 < \kappa < 2$ *if* $d \geq m + M$.

Proof. If no aphid is ever allowed to give birth then eventually they will all die off. This proves (a).

To prove part (b), we observe that if $d < m + M$ then each aphid will only produce one offspring in its lifetime. So for $m = 1$ there will never be more than $d + 1$ aphids alive on any one day and for $m > 1$ there will never be more than d/m aphids on any one day. Thus the sequence $\{A_n\}$ acts like a constant sequence. Notice that we can make $\{A_n\}$ eventually the sequence $\{\kappa\}$ for any positive integer κ in this case, by letting $m = 1$ and $d = \kappa - 1$.

We proved part (c) above. Notice that $\kappa = 2$ when $m = 1$ and $M = 1$ and the aphids are immortal. See Table 2 for this example. ■

The same result can be proved for the growth rate in the immortal case. In this case, we have the generating function from earlier

$$A(x) = \frac{1 - x^m}{(1 - x^m - x^M)(1 - x)}.$$

Since $1 - x$ is a factor of $1 - x^m$, we need only consider the first positive real zero of the polynomial $1 - x^m - x^M$, which is really finding the positive zero of the denominator of $I(x)$ when $k = 2$. Thus, we have the following theorem.

Theorem 6. *Suppose that an aphid gives birth to its first daughter after m days and gives birth to each subsequent child every M days after that. If A_n is the total number of aphids on day n then $A_n \sim \kappa^n$, where $1 < \kappa \leq 2$.*

Theorems 5 and 6 state that there is some κ between 1 and 2 where $A_n \sim \kappa^n$. An interesting question is what are these κ? As we mentioned before $\kappa = 2$ in the case when $m = 1$ and $M = 1$. Also, $\kappa = \varphi$, the golden ratio, when $m = 1$ and $M = 2$ since this is the growth rate of the Fibonacci numbers.

One obvious question we can ask is whether given a specific value of $1 < \kappa \leq 2$, can we find m and M so that $A_n \sim \kappa^n$? Unfortunately, the answer is no. Let us focus on the zeros of $1 - x^m - x^M$. The case for mortal aphids is similar. Just as in the mortal case, this function is decreasing on $[0, \infty)$, and it has a unique zero in the interval $[\frac{1}{2}, 1)$. Assume that ζ is the zero of $1 - x^m - x^M$. If $s > m$ or $t > M$, then ζ cannot be the zero of $1 - x^s - x^t$. In fact, $1 - \zeta^s - \zeta^t > 0$, if $s > m$ or $t > M$. This means that if $\kappa_{m,M}$ is the growth rate of the aphid sequence $\{A_n\}$ with immature birth rate m and mature birth rate M then $\kappa_{m,M} < \kappa_{s,t}$ whenever $s > m$ or $t > M$.

One conclusion that we can immediately draw from this is that there is no aphid sequence with growth rate κ^n, where $\varphi < \kappa < 2$. Thus, we rephrase the question we asked in the previous paragraph. For which values of κ can we find an aphid sequence with the growth rate κ^n? Certainly κ must be algebraic, so that eliminates uncountably many possibilities, but there are also some algebraic values that cannot be growth rates. We do not have any good answers to this question at this time.

Acknowledgments We would like to thank Dr. Thomas Gulya for many interesting conversations.

REFERENCES

[1] Alfred, B. U. (1963). Exploring Fibonacci numbers. *Fibonacci Quart.* 1(1): 57–63.
[2] Alfred, B. U. (1963). Dying rabbit problem revived. *Fibonacci Quart.* 1(4): 53–56.
[3] Benjamin, A. T., Quinn, J. J. (2003). *Proofs That Really Count. The Art of Combinatorial Proof*. Washington, DC: Mathematical Association of America.
[4] Dixon, A. F. G. (1976). *Biology of Aphids*. London, UK: Hodder.
[5] Feng, J. (2011). Some new remarks about the dying rabbit problem. *Fibonacci Quart.* 49(2): 171–176.
[6] Flajolet, P., Sedgewick, R. (2009). *Analytic Combinatorics*. Cambridge: Cambridge University Press.
[7] Hoggatt, V. E. Jr., Lind, D. A. (1965). The dying rabbit problem. *Fibonacci Quart.* 7(5): 482–487.
[8] Miles, E. P. Jr (1960). Generalized Fibonacci numbers and associated matrices. *Am. Math. Monthly.* 67: 745–752.
[9] Munarini, E. (1997). A combinatorial interpretation of the generalized Fibonacci numbers. *Adv. Appl. Math.* 19(3): 306–318.
[10] Oller-Marcén, A. M. (2009). The dying rabbit problem revisited. *Integers* 9(2): 129–138.
[11] Pisano, L. (2002). *Fibonacci's Liber Abaci: A Translation into Modern English of Leonardo Pisano's Book of Calculation* (Sigler, L. E., trans.). New York, NY: Springer-Verlag.
[12] Sloane, N. (2012). The encyclopedia of integer sequences. Available at: `http://oeis.org`.

Summary. Fibonacci numbers are well known to enumerate pairs of rabbits assuming liberally interpreted reproductive patterns and immortality. We turn these numbers inside out by considering the reproductive behavior of aphids, whose offspring can begin to reproduce before they are even born. We study the sequences that arise both from allowing aphids to be immortal and from assuming a given lifespan, tying together many different generalizations of Fibonacci numbers. We determine recurrence relations and generating functions for them. We then use the generating functions to determine bounds on the growth rates of these sequences.

SAMUEL ERICKSON (MR Author ID: 1190727) is a software developer in Fargo-Moorhead Area. His interests include machine learning, data science, and statistics.

ADAM M. GOYT (MR Author ID: 841007) is a professor of mathematics at Minnesota State University Moorhead. His areas of research are algebraic and enumerative combinatorics and history of mathematics.

JOSIAH REISWIG (MR Author ID: 1170827) is a mathematics graduate student at the University of South Carolina. His interests include extremal graph theory, extremal set theory, and racquetball.

Proof Without Words: Using Trapezoids to Compute Triangular Numbers

JAIME GASPAR
University of Kent
Canterbury CT2 7NF, UK
Centro de Matemática e Aplicações (CMA), FCT, UNL
Quinta da Torre, 2829-516 Caparica, Portugal
mail@jaimegaspar.com

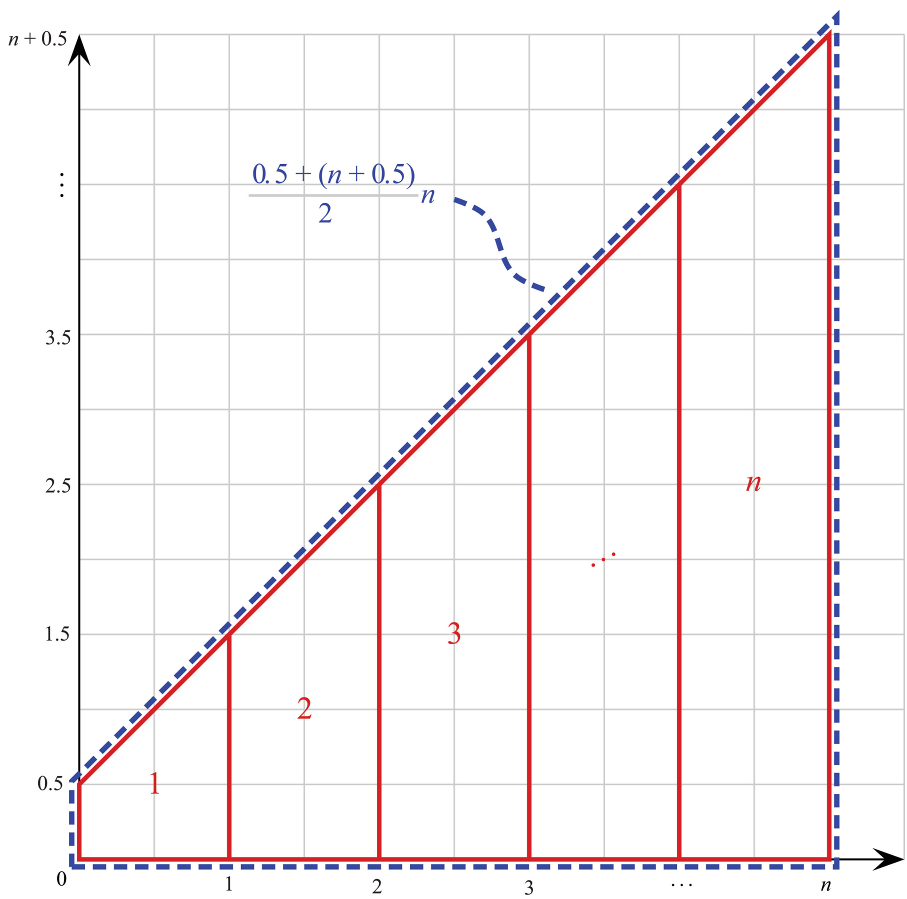


Figure 1

Remark. An anonymous referee (essentially) suggested to transform this proof without words with a figure into the following proof without words with symbols:

$$\int_{k-1}^{k} (x + 1/2)\,\mathrm{d}x = k \quad\Rightarrow\quad 1 + 2 + 3 + \cdots + n = \int_0^n (x + 1/2)\,\mathrm{d}x = \frac{n(n+1)}{2}.$$

Math. Mag. **91** (2018) 206–207. doi:10.1080/0025570X.2018.1451172
MSC: Primary 05A19
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/umma.

Remark. There is a classic proof shown in Figures 2 and 3 [2, p. 69]. An anonymous referee noticed a relation illustrated in Figure 4 between the proof from Figure 1 and the proof from Figures 2 and 3.

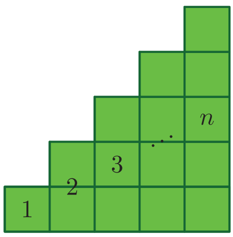


Figure 2

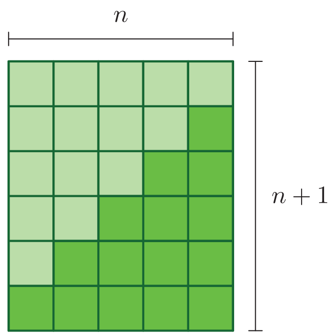


Figure 3

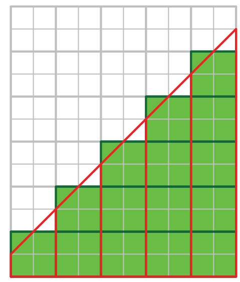


Figure 4

Acknowledgments A figure similar to Figure 1 was previously published elsewhere by the author [1, p. 42]. Funding at the time of writing: Martí Franquès Research Fellowship Programme grant number 2013PMF-PIPF-24 of the Universitat Rovira i Virgili. Funding at the time of submission: Research Postgraduate Scholarship from the Engineering and Physical Sciences Research Council / School of Computing, University of Kent. Affiliation at the time of writing: Universitat Rovira i Virgili, Department of Computer Engineering and Mathematics, Avinguda Països Catalans 26, E-43007 Tarragona, Catalonia; Centro de Matemática e Aplicações (CMA), FCT, UNL.

REFERENCES

[1] Gaspar, J. (2016). Mathematical Candies. *Bulletin of the Portuguese Mathematical Society* (in Portuguese). Special Issue, Proceedings of the 2014 National Meeting of the Portuguese Mathematical Society (in Portuguese). pp. 41–44.

[2] Nelsen, R. B. (1993). *Proofs Without Words: Exercises in Visual Thinking*, 3rd ed. Washington DC: The Mathematical Association of America.

Summary. Giving a closed-form expression for the triangular numbers is perhaps the result most often proved without words. We use trapezoids to present a proof without words that the nth triangular number is $n(n+1)/2$.

JAIME GASPAR (MR Author ID: 875672, ORCID 0000-0001-5810-9293) studies, researches in, and teaches mathematics and computer science; he has been at the University of Lisbon in Portugal, Technical University of Darmstadt in Germany, National Institute for Research in Computer Science and Automation in France, and Rovira i Virgili University in Spain; now he is at the University of Kent in the United Kingdom.

Fossicking for Finite Fields

B. J. GARDNER
University of Tasmania
Hobart, Australia
Barry.Gardner@utas.edu.au

Students generally become aware of the existence of finite fields in a first course in abstract algebra, but the only examples at that stage tend to be the fields $\mathbb{Z}_p$ of integers modulo a prime p. These students must wait until a later course for a full treatment of finite fields and many will never see this. It seems desirable to be able to give some more examples of finite fields in the first course, based on concepts which are familiar at that stage. We shall show how to find, for many primes p including the small ones, a field with p^2 elements which is a subring of the ring $M_2(\mathbb{Z}_p)$ of 2×2 matrices over the field $\mathbb{Z}_p$.

We shall use integer symbols to denote elements of the various $\mathbb{Z}_p$ without being fastidious about any relationship between the integers and p, and if, for instance, in an expression involving $\mathbb{Z}_5$, 6 is replaced by 1, the new expression will be joined to the old by $=$, rather than by $\equiv \pmod 5$.

Fields

Recall that the real 2×2 matrices of the form $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ form a field isomorphic to the complex numbers *via* the correspondence

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \mapsto a + bi;$$

so, in particular,

$$\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} \mapsto a \text{ for every } a \text{ and } \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mapsto i.$$

We shall first try to imitate this method of constructing a field by starting with $\mathbb{Z}_p$ and examining the subring

$$C(p) = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} : a, b \in \mathbb{Z}_p \right\}$$

of $M_2(\mathbb{Z}_p)$. It is easy to show that $C(p)$ *is* a subring, and moreover a commutative one, but this will be shown below anyway. Since a and b can take p possible values, $C(p)$ has precisely p^2 elements. Perhaps not surprisingly, $C(p)$ is not a field if $\mathbb{Z}_p$ contains a square root of -1. Here are the details.

Lemma. *$C(p)$ is a field if and only if $\mathbb{Z}_p$ does not contain a square root of -1.*

Proof. The determinant $det\left(\begin{bmatrix} a & b \\ -b & a \end{bmatrix}\right)$ of $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ is $a^2 + b^2$. For $a = 0$ or $b = 0$ this is zero if and only if $a = b = 0$. For $a, b \neq 0$ we have $a^2 + b^2 = 0$ if and only if

Math. Mag. **91** (2018) 208–212. doi:10.1080/0025570X.2018.1445377 © Mathematical Association of America
MSC: Primary 12E30

$(\frac{a}{b})^2 = -1$, so if there is no square root of -1 in $\mathbb{Z}_p$, each nonzero element of $C(p)$ is invertible. Then since

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix}^{-1} = \frac{1}{a^2+b^2}\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \in C(p),$$

it follows that $C(p)$ is a field. On the other hand, if there is an element $t \in \mathbb{Z}_p$ with $t^2 = -1$, then $det\left(\begin{bmatrix} t & 1 \\ -1 & t \end{bmatrix}\right) = t^2 + 1 = 0$, so $C(p)$ is not a field. ■

It happens that $\mathbb{Z}_p$ contains a square root of -1 if and only if $p \equiv 1 \pmod 4$ or $p = 2$. (See [**1**, p. 69 (Theorem 82)] for odd primes; in $\mathbb{Z}_2$ we have $1^2 = 1 = -1$.) Thus, we have the following proposition.

Proposition. *The following are equivalent for p a prime.*

1. *$C(p)$ is a field.*
2. *$\mathbb{Z}_p$ contains no square root of -1.*
3. *$p \equiv 3 \pmod 4$.*

Example 1. The field $C(3)$ consists of the following nine matrices (where $2 = -1$ and $1 = -2$):

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix},$$
$$\begin{bmatrix} 2 & 1 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 2 & 2 \\ 1 & 2 \end{bmatrix}.$$

For a field of order p^2 when $\mathbb{Z}_p$ *does* contain a square root of -1, we shall have to use a different method. There is another way of looking at $C(p)$ which is helpful in suggesting other methods. Let I denote the 2×2 identity matrix over $\mathbb{Z}_p$ and $u = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ (so that $u^2 = -I$). Then I and u are in $C(p)$ and, for all $a, b \in \mathbb{Z}_p$, we have

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} = aI + bu.$$

As I and u are linearly independent over $\mathbb{Z}_p$, $C(p)$ is a two-dimensional $\mathbb{Z}_p$-vector space and therefore has p^2 elements. Since for all $a, b, c, d \in \mathbb{Z}_p$ we have

$$\begin{aligned}(aI + bu)(cI + du) &= acI^2 + adIu + bcuI + bdu^2 \\ &= (ac - bd)I + (ad + bc)u = (cI + du)(aI + bu),\end{aligned}$$

where the resemblance to complex number multiplication is not accidental, $C(p)$ is a commutative subring with identity of $M_2(\mathbb{Z}_p)$. This is perhaps easier to imitate, using another matrix in place of u, than the procedure we used initially.

Thus, with a view to finding more fields, we shall proceed similarly, using the matrix v which we now introduce. *For the moment, p is an odd prime and all matrices have entries in* $\mathbb{Z}_p$. We will treat $\mathbb{Z}_2$ separately.

Let $v = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$. Then v and I are linearly independent. Let

$$D(p) = \{aI + bv : a, b \in \mathbb{Z}_p\}.$$

This too is a two-dimensional vector space over $\mathbb{Z}_p$. We have

$$v^2 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = I + v.$$

Thus $v^2 \in D(p)$, so $D(p)$ is closed under multiplication, and as with $C(p)$, we see that $D(p)$ is a commutative subring with identity of $M_2(\mathbb{Z}_p)$. Let us see when (if ever) it is a field.

A typical element of $D(p)$ has the form

$$x\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} x & b \\ b & x+b \end{bmatrix}$$

and its determinant is $x(x+b) - b^2 = x^2 + xb - b^2$. If $x = 0$ or $b = 0$, then the determinant is zero if and only if $x = b = 0$, so we will consider $x, b \neq 0$. Purely formally from $x^2 + xb - b^2 = 0$, we get

$$x = \frac{-b \pm \sqrt{b^2 - 4(-b^2)}}{2} = \frac{-b \pm \sqrt{5b^2}}{2} = \frac{-b \pm b\sqrt{5}}{2}.$$

Whatever can be said about the intermediate expressions, the final one will represent an element of $\mathbb{Z}_p$ as long as we can assign a meaning to "$\sqrt{5}$," as p is odd and hence 2 has an inverse in $\mathbb{Z}_p$.

Lemma. *$D(p)$ is a field if and only if $\mathbb{Z}_p$ contains no square root of* 5.

Proof. If $s^2 = 5$ for some $s \in \mathbb{Z}_p$, let $\alpha = \frac{-b+bs}{2}$. Then $s \neq 1$ so $b \neq bs$ and α is a nonzero element of $\mathbb{Z}_p$, and we have

$$\begin{aligned}\alpha^2 + \alpha b - b^2 &= \frac{b^2 - 2b^2 s + b^2 s^2}{4} + \frac{-b+bs}{2}b - b^2 \\ &= \frac{b^2 + b^2 s^2 - 2b^2 - 4b^2}{4} = \frac{b^2 + 5b^2 - 2b^2 - 4b^2}{4} = 0.\end{aligned}$$

Thus $det\left(\begin{bmatrix} \alpha & b \\ b & \alpha + b \end{bmatrix}\right) = 0$, though $\alpha, b \neq 0$, so $D(p)$ is not a field if $\mathbb{Z}_p$ contains a square root of 5.

On the other hand, it is easy to show that if a matrix in $D(p)$ is invertible, then its inverse is in $D(p)$ (cf. $C(p)$), so if $D(p)$ is not a field, then $x^2 + xb - b^2 = 0$ for some nonzero $x, b \in \mathbb{Z}_p$, and then, manipulating our formal quadratic solution, we get $\sqrt{5} = \pm(2x+b)/b$. The right-hand side represents elements of $\mathbb{Z}_p$ and these are indeed square roots of 5:

$$\left(\frac{2x+b}{b}\right)^2 = \frac{4x^2 + 4xb + b^2}{b^2} = \frac{4(x^2 + xb - b^2) + 5b^2}{b^2} = \frac{5b^2}{b^2} = 5.$$

∎

If $p \neq 2, 5$ then $\mathbb{Z}_p$ contains a square root of 5 if and only if $p \equiv 1$ or $9 \pmod{10}$. This is proved explicitly in [**1**, pp. 76 (Theorem 97) and 78]. Thus, e.g., $\mathbb{Z}_{13}$ and $\mathbb{Z}_{47}$ contain no square root of 5, so $D(13)$ and $D(47)$ are fields.

It is not usual to number 0 among the "squares modulo a prime" (*quadratic residues*). But $0^2 \equiv 5 \pmod 5$, so we need to look separately at $D(5)$. Substituting 0 for $\sqrt{5}$ in our formal quadratic solution we get $x = -b/2$. Taking $x = -1$, $b = 2$, for instance, we get

$$det\left(\begin{bmatrix} -1 & 2 \\ 2 & -1+2 \end{bmatrix}\right) = det\left(\begin{bmatrix} -1 & 2 \\ 2 & 1 \end{bmatrix}\right) = -1 - 4 = -5 = 0.$$

Hence $D(5)$ is not a field.

And where did v come from? It actually fell out serendipitously when the author needed to characterize the one-generator subrings of $M_2(\mathbb{Z}_2)$ for another investigation. Actually, though, because its entries give it such a simple arithmetic, it is just about **the** natural matrix to try in an attempt to imitate what has been done with u.

We still do not know about $D(2)$; it is not covered by [**1**, p. 76 (Theorem 97)] as this result deals only with odd primes. Since $5 \equiv 1 \pmod 2$, $\mathbb{Z}_2$ actually has a square root of 5, but as often happens, 2 is "different." As $D(2)$ is so small, we can easily write down all the values of $a^2 + ab - b^2$, i.e., of $a + ab + b$, for $a, b \in \mathbb{Z}_2$. Since

$$0 + 0 \cdot 0 + 0 = 0; \quad 0 + 0 \cdot 1 + 1 = 1 = 1 + 1 \cdot 0 + 0; \quad 1 + 1 \cdot 1 + 1 = 1,$$

all nonzero matrices in $D(2)$ are invertible, so $D(2)$ is a field.

Summarizing all this we have the following result, since an odd prime must be congruent to 1,3,7, or 9 (mod 10).

Proposition. *The following conditions are equivalent.*

1. *$D(p)$ is a field;*
2. *$p = 2$ or $\mathbb{Z}_p$ contains no square root of 5.*
3. *$p = 2$ or $p \equiv 3$ or $7 \pmod{10}$.*

By using $C(p)$ or $D(p)$ as convenient and appropriate we can now get fields with p^2 elements for many values of p:

$C(p)$ for $p = 3, 7, 11, 19, 23, 31, 43, 47, 59, 67, 71, \ldots$ and
$D(p)$ for $p = 2, 3, 7, 13, 17, 23, 37, 43, 47, 53, 67, \ldots$.

The ones missing out are 5, 29, 41, 61, 73, 89, 101, ...

This prompts the question: What is a small prime like 5 doing in a list like this? We would like to have examples of fields with p^2 elements for the small primes (including 5), so we shall use another matrix in place of u, v to find a field of order 25.

For a prime p, let $w = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \in M_2(\mathbb{Z}_p)$. (Why *this* matrix? It is a bit like v, it is not in $D(\mathbb{Z}_5)$, its entries are small, and it works.) Then

$$w^2 = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 2 & 4 \end{bmatrix} = I + 2w,$$

so as in the previous cases, $\{aI + bw : a, b \in \mathbb{Z}_p\}$, which we will call $E(p)$, is a commutative subring and a two-dimensional subspace of $M_2(\mathbb{Z}_p)$. For all $a, b \in \mathbb{Z}_p$ we have

$$aI + bw = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix} + \begin{bmatrix} 0 & b \\ b & 2b \end{bmatrix} = \begin{bmatrix} a & b \\ b & a+2b \end{bmatrix}$$

and $det\left(\begin{bmatrix} a & b \\ b & a+2b \end{bmatrix}\right) = a(a+2b) - b^2 = a^2 + 2ab - b^2$. We examine the values of this expression for $p = 5$.

If $a, b \in \mathbb{Z}_5$ and a or $b = 0$, then $a^2 + 2ab - b^2 = 0$ if and only if $a = b = 0$, so assume $a, b \neq 0$. If a and b have the same square, then $a^2 + 2ab - b^2 = 2ab \neq 0$. If a and b have different squares, then as the nonzero squares are 1 and 4, we have $a^2 - b^2 = 2$ or 3 and (as $\{a, b\} = \{1, 2\}, \{1, 3\}, \{2, 4\}$ or $\{3, 4\}$) $2ab = 1$ or 4, so $a^2 + 2ab - b^2 \neq 0$. As invertible elements of $E(5)$ have their inverses in $E(5)$, we now have what we need.

Proposition. *$E(5)$ is a field (with 25 elements).*

We have introduced $E(p)$ only to plug a hole by providing a field of order 5^2. We leave it as an exercise for the reader to show that in general $E(p)$ is a field if and only if $\mathbb{Z}_p$ contains no square root of 2. For a more general investigation, take the subring of $M_2(\mathbb{Z}_p)$ generated by a single matrix and see if it is a field or not. If it is a field, what can be deduced about the existence of square roots in $\mathbb{Z}_p$ or the nature of p? If it is not a field, what does it look like? In general, what can its dimension be? All of these questions are easier to answer for $M_2(\mathbb{Z}_2)$, where the calculations are easier and there are fewer matrices.

In a second abstract algebra course, finite fields can be described more systematically and perhaps classified completely, as notions like ideals and factor rings of polynomial rings have become familiar to students. In a first course such notions are not treated in depth if they are mentioned at all. On the other hand, elementary matrix theory should be familiar (at least over the reals, and the extension to other rings is easy). Using only the most elementary matrix theory, we have obtained examples of fields with p^2 elements for many primes p, including all those less than 29. Moreover, our methods should suggest accessible investigations whereby students can produce further examples.

Acknowledgment The author thanks the referees for a stimulating dialogue.

REFERENCES

[1] Hardy, G. H., Wright, E. M. (1979). *An Introduction to the Theory of Numbers*. 5th ed. Oxford, UK: Oxford University Press.

Summary. Fields of order p^2 for various primes p, including all $p < 29$, are exhibited by elementary means as subrings of the 2×2 matrix ring over $\mathbb{Z}_p$.

B. J. GARDNER (MR Author ID: 71400) “retired” from the University of Tasmania, Hobart, Australia in 2014. He is now an Honorary Research Associate of that university, but lives in a small town of about 500 souls on the shores of Bass Strait. His publications (including two books) are mainly in algebra, principally radical theory and ring theory.

From Chebyshev to Jensen and Hermite-Hadamard

DAN ŞTEFAN MARINESCU
National College "Iancu de Hunedoara"
Hunedoara, Romania
marinescuds@gmail.com

MIHAI MONEA
National College "Decebal"
Deva, Romania
mihaimonea@yahoo.com

It has been said that the relation which truly governs mathematics is that of inequality, equality being a special case. In this note, we connect three classical inequalities, namely, the Chebyshev, Jensen, and Hermite-Hadamard inequalities, by showing an unexpected relationship between Chebyshev's inequality and convexity. We are motivated by Mercer's original proof in [4] of Jensen's inequality for a function with positive second derivative.

Chebyshev's order inequality

> *The most obviously important named inequalities are those of Hölder and Minkowski, but the watershed paper, in my estimation, is the paper of Chebyshev.*

This opinion is due to Fink (see [2]). In this paper, we explore the discrete analog of the integral inequality that is the main result of [1]. We recall this result, often called Chebyshev's order inequality (see, for example, [6, p. 76]), with a simple proof.

Chebyshev inequality. Let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be real numbers such that $a_1 \geq a_2 \geq \cdots \geq a_n$ and $b_1 \leq b_2 \leq \cdots \leq b_n$. Then, for any $w_1, w_2, \ldots, w_n \in [0, 1]$ such that $\sum_{k=1}^{n} w_k = 1$, we have

$$\sum_{i=1}^{n} w_i a_i b_i \leq \left(\sum_{i=1}^{n} w_i a_i\right) \cdot \left(\sum_{i=1}^{n} w_i b_i\right). \tag{1}$$

Proof. The case $n = 1$ is clear. For $n \geq 2$, denote $A = \sum_{i=1}^{n} w_i a_i$. There exists $k \in \{1, 2, \ldots, n-1\}$ such that $a_k \geq A \geq a_{k+1}$. Then $(a_i - A)(b_i - b_k) \leq 0$, for any $i \in \{1, 2, \ldots, n\}$. We obtain $a_i b_i + A b_k \leq a_i b_k + A b_i$. By multiplying this relation with w_i and summing from 1 to n, we obtain

$$\sum_{i=1}^{n} w_i a_i b_i + A b_k \sum_{i=1}^{n} w_i \leq b_k \sum_{i=1}^{n} w_i a_i + A \sum_{i=1}^{n} w_i b_i.$$

Since $\sum_{i=1}^{n} w_i = 1$, the following inequality is equivalent to Equation (1).

$$\sum_{i=1}^{n} w_i a_i b_i + A b_k \leq A b_k + A \sum_{i=1}^{n} w_i b_i.$$

■

Math. Mag. **91** (2018) 213–217. doi:10.1080/0025570X.2018.1445376 © Mathematical Association of America
MSC: Primary 26D15

The proof shows that we obtain equality when $b_i = b_k$ or $a_i = A$, for all $i \in \{1, 2, \ldots n\}$. This leads to $a_1 = a_2 = \cdots = a_n$ or $b_1 = b_2 = \cdots = b_n$.

Convexity and some related results

Let I denote a real interval and I^o its interior. Let $f : I \to \mathbb{R}$ be a convex function, so that

$$f(tx + (1-t)y) \leq t \cdot f(x) + (1-t) \cdot f(y),$$

for any $x, y \in I$, and $t \in [0, 1]$. If the reverse inequality holds for any $x, y \in I$, and $t \in [0, 1]$, the function is called concave. Convex functions are continuous on I^o. Moreover, if a convex function is differentiable on I^o, then its derivative is nondecreasing on I^o (for example, see Theorem 1.3.3 from [5]). An important tool for our proof is represented by the next result (see assertion (iv) of Proposition 1.1.7 from [5]).

Proposition 1. *Let $f : I \to \mathbb{R}$ be a convex function. For any $x, y \in I$ with $x \leq y$, the function $g(t) = f((1-t)x + ty)$, defined for any $t \in [0, 1]$, is convex.*

Proof. By writing the first 1 below as $1 = \alpha + (1-\alpha)$, we see that

$$\begin{aligned}&(1 - \alpha u - (1-\alpha)\,v)\,x + (\alpha u + (1-\alpha)\,v)\,y\\&\quad = \alpha\,((1-u)\,x + uy) + (1-\alpha)\,((1-v)\,x + vy)\,.\end{aligned}$$

Taking $u, v \in [0, 1]$, and $\alpha \in [0, 1]$ and using this identity, we have

$$\begin{aligned}g\,(\alpha u + (1-\alpha)\,v) &= f\,((1 - \alpha u - (1-\alpha)\,v)\,x + (\alpha u + (1-\alpha)\,v)\,y)\\&= f\,(\alpha\,((1-u)\,x + uy) + (1-\alpha)\,((1-v)\,x + vy))\\&\leq \alpha f\,((1-u)\,x + uy) + (1-\alpha)\,f\,((1-v)\,x + vy)\\&= \alpha g\,(u) + (1-\alpha)\,g\,(v)\,,\end{aligned}$$

which proves that the function g is convex. ■

In this paper, we relate Chebyshev's inequality to Jensen's inequality and the Hermite-Hadamard inequality (see [5, pp. 122 and 139]). For completeness, we recall these inequalities below. While Chebyshev's inequality involves real numbers, the other two inequalities involve convex functions.

Jensen inequality. Let $f : I \to \mathbb{R}$ be a convex function. Let $n \in \mathbb{N}$ and $x_1, x_2, \ldots, x_n \in I$. Let $w_1, w_2, \ldots, w_n \in [0, 1]$ such that $\sum_{k=1}^{n} w_k = 1$. Then

$$f\left(\sum_{k=1}^{n} w_k x_k\right) \leq \sum_{k=1}^{n} w_k f(x_k).$$

Hermite inequality. Let $f : I \to \mathbb{R}$ be a convex function. Then, for any $a, b \in I$ with $a \leq b$, we have

$$f\left(\frac{a+b}{2}\right) \leq \frac{1}{b-a}\int_a^b f(x)dx \leq \frac{f(a)+f(b)}{2}.$$

From Chebyshev's inequality to Jensen's and the Hermite-Hadamard inequalities

Let $f : I \to \mathbb{R}$ be a convex function, differentiable on I^o. Let $x_1, x_2, \ldots, x_n \in I$ and $w_1, w_2, \ldots, w_n \in [0, 1]$ such that $\sum_{k=1}^n w_k = 1$. Let $A = \sum_{k=1}^n w_k x_k$. We introduce the function $F : [0, 1] \to \mathbb{R}$ defined, for any $t \in [0, 1]$, by

$$F(t) = \sum_{k=1}^n w_k f((1-t)x_k + tA).$$

The next proposition follows from the relationship between a differentiable function having nonnegative derivative on an interval and being nonincreasing on that interval.

Proposition 2. *The function F is convex and nonincreasing.*

Proof. Without loss of generality, we can assume that $x_1 \le x_2 \le \cdots \le x_n$. First, we observe that if $x_1 = x_n$, then the function F is constant and the proof is clear. Thus, we can assume that $x_1 < x_n$, and hence $x_1 < A < x_n$. The function F is convex since by Proposition 1 it is a linear combination of convex functions. Therefore, it is continuous on $(0, 1)$. Since $F(1) = f(A)$, it follows that F is continuous at $t = 1$.

On the other hand, F is differentiable on $(0, 1)$ since

$$F'(t) = \sum_{k=1}^n w_k (A - x_k) f'((1-t)x_k + tA),$$

for any $t \in (0, 1)$. From $x_1 \le x_2 \le \cdots \le x_n$, we obtain $A - x_1 \ge A - x_2 \ge \cdots \ge A - x_n$. Then $f'((1-t)x_1 + tA) \le f'((1-t)x_2 + tA) \le \cdots \le f'((1-t)x_n + tA)$. By using Chebyshev's inequality, we obtain

$$\begin{aligned} F'(t) &= \sum_{k=1}^n w_k (A - x_k) f'((1-t)x_k + tA) \\ &\le \left(\sum_{k=1}^n w_k (A - x_k)\right)\left(\sum_{k=1}^n w_k f'((1-t)x_k + tA)\right) \\ &= \left(A - \sum_{k=1}^n w_k x_k\right)\left(\sum_{k=1}^n w_k f'((1-t)x_k + tA)\right) \\ &= 0. \end{aligned}$$

We conclude that F is nonincreasing on $(0, 1)$. By the continuity of F, it follows that F is nonincreasing on $(0, 1]$.

Now, let $x \in (0, 1)$. Then $x = x \cdot 1 + (1-x) \cdot 0$ and $F(x) \le xF(0) + (1-x)F(1)$. Since $F(1) \le F(x)$, we obtain

$$F(x) \le xF(0) + (1-x)F(x),$$

which is equivalent to $xF(x) \le xF(0)$. We obtain $F(x) \le F(0)$ and the proof is complete. ■

Since $F(0) \ge F(1)$, it follows that

$$\sum_{k=1}^n w_k f(x_k) \ge f\left(\sum_{k=1}^n w_k x_k\right),$$

which is Jensen's inequality.

On the other hand, $F(1) \le F(t) \le F(0)$, for any $t \in [0, 1]$. If $n = 2$ and $x_1 < x_2$, we obtain

$$f(w_1x_1 + w_2x_2) \le \sum_{k=1}^{2} w_k f((1-t)x_k + tA) \le w_1 f(x_1) + w_2 f(x_2).$$

If we choose $w_1 = w_2 = \frac{1}{2}$ and integrate the previous inequality with respect to t over $[0, 1]$, then

$$f(\frac{x_1 + x_2}{2}) \le \frac{1}{2}\sum_{k=1}^{2}\int_0^1 f((1-t)x_k + tA)dt \le \frac{f(x_1) + f(x_2)}{2}.$$

Further, we have

$$\begin{aligned}\int_0^1 f((1-t)x_1 + tA)dt &= \int_0^1 f\left((1-t)x_1 + t\,\frac{x_1 + x_2}{2}\right)dt \\ &= \int_0^1 f\left(x_1 + \frac{x_2 - x_1}{2}\,t\right)dt.\end{aligned}$$

By using the substitution $s = x_1 + \frac{x_2 - x_1}{2}t$, we find that

$$\int_0^1 f\left(x_1 + \frac{x_2 - x_1}{2}\,t\right)dt = \frac{2}{x_2 - x_1}\int_{x_1}^{\frac{x_1+x_2}{2}} f(s)ds.$$

Interchanging x_1 and x_2, we have

$$\begin{aligned}\int_0^1 f\left(x_2 - \frac{x_2 - x_1}{2}\,t\right)dt &= -\frac{2}{x_2 - x_1}\int_{x_2}^{\frac{x_1+x_2}{2}} f(s)ds \\ &= \frac{2}{x_2 - x_1}\int_{\frac{x_1+x_2}{2}}^{x_2} f(s)ds.\end{aligned}$$

Then

$$\frac{1}{2}\sum_{k=1}^{2}\int_0^1 f((1-t)x_k + tA)dt = \frac{1}{x_2 - x_1}\int_{x_1}^{x_2} f(s)ds,$$

and hence

$$f\left(\frac{x_1 + x_2}{2}\right) \le \frac{1}{x_2 - x_1}\int_{x_1}^{x_2} f(s)ds \le \frac{f(x_1) + f(x_2)}{2},$$

which is the Hermite-Hadamard inequality.

We conclude this paper with a final remark. The condition that the function f is differentiable on I^o is unnecessary. Any convex function $f : I \to \mathbb{R}$ has finite left and right derivatives at each point of I^o. In particular, both f'_- and f'_+ are nondecreasing on I^o (for example, see [**5**, Theorem 1.3.3]) and the monotonicity of a function can be studied only with left or right derivatives (see [**3**, Corollary 1.4.3]). In fact, any continuous function $f : I \to \mathbb{R}$ is nonincreasing if the left derivative $f'_-(x_0)$ exists and is negative, for every $x_0 \in I^o$. Therefore, we obtain the same conclusion if we replace f' with f'_- in the proof of Proposition 2.

REFERENCES

[1] Chebyshev, P. L. (1882). Sur les expressions approximatives des intégrales définies par les autres prises entre les mêmes limites. *Proc. Math. Soc. Kharkov.* 2: 93–98 (Russian), translated in Oeuvres, 2 (1907), 716–719.
[2] Fink, A. M. (2000). An essay on the history of inequalities. *J. Math. Anal. Appl.* 249: 118–134.
[3] Flett, T. M. (1980). *Differential Analysis*. Cambridge, UK: Cambridge University Press.
[4] Mercer, A. McD. (2010). Short proofs of Jensen's and Levinson's inequalities. *Math. Gazette.* 94: 492–495.
[5] Niculescu, C. P., Persson, L.-E. (2006). *Convex Functions and Their Applications: A Contemporary Approach*. New York, NY: Springer.
[6] Steele, J. M. (2004). *The Cauchy–Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. Cambridge, UK: Cambridge University Press.

Summary. In this note, we connect three classical inequalities, namely, the Chebyshev, Jensen, and Hermite-Hadamard inequalities, by showing an unexpected relationship between Chebyshev's inequality and convexity.

DAN ŞTEFAN MARINESCU (MR Author ID: 676895) is a teacher of National College Iancu de Hunedoara, Hunedoara, Romania. He has a Ph.D. in mathematics, obtained from the University Babeş-Bolyai of Cluj Napoca. He is a member of the Romanian Central National Olympiad Committee.

MIHAI MONEA (MR Author ID: 897493) is a teacher of National College Decebal, Deva, Romania. He is Ph.D. student of the University Politehnica from Bucharest. He is a member of the Romanian Central National Olympiad Committee.

The Generalized Harmonic Series Diverges by the AM-GM Inequality

A nice survey article on proofs of the harmonic series diverging is [1] and an unpublished addendum is [2]. The proof presented here only uses the AM-GM inequality. Note that without the AM-GM inequality the proof follows since $\frac{1}{2n+1} + \frac{1}{2n+2} > \frac{2}{2n+2}$. If the harmonic series converges, then

$$\begin{aligned}\sum_{n=0}^{\infty} \frac{1}{n+1} &= \sum_{n=0}^{\infty} \left(\frac{1}{2n+1} + \frac{1}{2n+2} \right) \\ &\geq \sum_{n=0}^{\infty} \frac{2}{\sqrt{(2n+1)(2n+2)}} \quad \text{(by the AM-GM inequality)} \\ &> \sum_{n=0}^{\infty} \frac{2}{2n+2} = \sum_{n=0}^{\infty} \frac{1}{n+1}.\end{aligned}$$

—Submitted by Ángel Plaza (MR Author ID: 350023),
University of Las Palmas de Gran Canaria, Spain

REFERENCES

[1] Kifowit, S. J., Stamps, T. A. (2006). The harmonic series diverges again and again. *AMATYC Rev.* 27: 31–43.

[2] Kifowit, S. J. (2017). More proofs of divergence of the harmonic series. Available at http://stevekifowit.com/pubs/harm2.pdf.

Math. Mag. **91** (2018) 217. doi:10.1080/0025570X.2018.1446649
MSC: Primary 40A05

Powers of Positive Matrices

CHRIS BERNHARDT
Fairfield University
Fairfield, CT 06824
cbernhardt@fairfield.edu

In 1907, Oskar Perron proved a beautiful theorem about powers of positive matrices. This theorem is not just of theoretical interest, but has, as we shall see, a number of important applications. Surprisingly, this result is often omitted from courses and texts in undergraduate linear algebra. The aim of this paper is to give an elementary proof of the result. But first, we begin with an example to help guide our intuition.

A matrix with real entries is called positive if all its entries are positive. We will look at powers of a positive matrix and show a remarkable result concerning limits as the power gets larger. To illustrate we begin with an example. Consider

$$M = \begin{bmatrix} 1 & 2 \\ 8 & 1 \end{bmatrix}.$$

Looking at a few powers gives

$$M^3 = \begin{bmatrix} 49 & 38 \\ 152 & 49 \end{bmatrix}, M^4 = \begin{bmatrix} 353 & 136 \\ 544 & 353 \end{bmatrix}, M^5 = \begin{bmatrix} 1441 & 842 \\ 3368 & 1441 \end{bmatrix}.$$

There is not an obvious pattern, apart from the diagonal entries being equal, but it is clear that the entries are getting larger quickly. Examining each of the four entries for larger and larger powers shows that they have an exponential growth rate with base of 5. If we look at $\frac{1}{5^n}M^n$ instead of M^n, we obtain

$$\begin{bmatrix} .392 & .304 \\ 1.216 & .392 \end{bmatrix}, \begin{bmatrix} .5648 & .2176 \\ .8704 & .5648 \end{bmatrix}, \begin{bmatrix} .46112 & .26944 \\ 1.07776 & .46112 \end{bmatrix},$$

for $n = 3, 4,$ and 5, respectively. It seems possible that each of the entries of $\frac{1}{5^n}M^n$ might approach a limit as n goes to infinity. Plugging in $n = 20$ gives, to five decimal places,

$$\frac{1}{5^{20}}M^{20} \approx \begin{bmatrix} .50002 & .24999 \\ .99996 & .50002 \end{bmatrix}.$$

A natural conjecture is that

$$\lim_{n\to\infty} \frac{1}{5^n}M^n = \begin{bmatrix} .5 & .25 \\ 1 & .5 \end{bmatrix}.$$

It turns out that this conjecture is correct, but this leaves the questions of why 5 is the exponential growth rate and where the entries in the limiting matrix come from. The answers all have to do with eigenvalues and eigenvectors.

It is straightforward to check that $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ is a right eigenvector with eigenvalue 5 and that $[\,2 \quad 1\,]$ is a left eigenvector with eigenvalue 5. Scalar multiple of eigenvectors are also eigenvectors. We will pick the left and right eigenvectors so that their dot

Math. Mag. **91** (2018) 218–227. doi:10.1080/0025570X.2018.1446615 © Mathematical Association of America
MSC: 15A24

product is 1. We take $\mathbf{R} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as a right eigenvector and $\mathbf{L} = [\,1/2 \quad 1/4\,]$ as a left eigenvector.

We can think of vectors as being special cases of matrices and then thinking of them as matrices multiply them together. For our example,

$$\mathbf{LR} = \begin{bmatrix} 1/2 & 1/4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \end{bmatrix}.$$

This is not surprising. We are essentially calculating the dot product and we chose the vectors to have a dot product of 1. However, reversing the order of multiplication to obtain what is often called the *outer* or *tensor* product gives

$$\mathbf{RL} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1/2 & 1/4 \end{bmatrix} = \begin{bmatrix} .5 & .25 \\ 1 & .5 \end{bmatrix}.$$

Notice that this matrix is the limiting matrix we found above. We are now ready to state our theorem about powers of a positive matrix.

Theorem 1 (Perron [8, 9]). *Any finite positive square matrix A has the following properties:*

1. *A has a positive eigenvalue, λ, such that if μ is any other eigenvalue, then $\lambda > |\mu|$.*
2. *A has a positive right eigenvector, associated with λ, with one-dimensional eigenspace.*
3. *A has a positive left eigenvector, associated with λ, with one-dimensional eigenspace.*
4. *Let $\mathbf{R}$ and $\mathbf{L}$ denote right and left eigenvectors corresponding to λ with $\mathbf{LR} = 1$. Let $H = \mathbf{RL}$. Then*

$$\lim_{n\to\infty} \frac{1}{\lambda^n} A^n = H.$$

This beautiful theorem deserves be more widely known than it is. Oskar Perron proved this in 1907. He needed the result to prove something else he was working on and published it as a lemma in [8], but immediately realized that the result was important in its own right and published it as the centerpiece in [9]. Over the years there have been many proofs of Perron's results and there have been many applications. C. R. MacCluer has written a paper [4] that gives both a number of proofs and applications.

Most of these proofs use sophisticated ideas. We will give an elementary argument, borrowing ideas from previous proofs, and arrange them in such a way that each step is explained. The resulting proof is much longer than many, but we do not use any results that are not seen at the undergraduate level.

The first step is to consider a much more well-known case. This is when the matrix is stochastic.

Powers of stochastic matrices

A stochastic matrix is a matrix that has real entries between 0 and 1 such that the entries in each row sum to one. This means that the entries in each row can be considered as probabilities. The entries in stochastic matrices are allowed to be 0, so they are nonnegative matrices, but not necessarily positive ones.

These matrices are often used in studying processes where an experiment is repeated, but the outcome of a given experiment affects the outcome of the next one. There are a finite number of possible outcomes, often called *states*, for the experiment. The probability of obtaining a certain outcome the next time an experiment is performed depends solely on the current state. Processes where the next outcome depends solely on the current state are called *Markov chains*, named after the Russian mathematician Andrey Markov.

Markov chains have an enormous number of practical applications. We will give one simple example to illustrate—Markov's analysis of Pushkin's *Eugene Onegin*.

Markov was interested in the proportions of vowels and consonants in Russian literature. In particular, what proportion of vowels and consonants followed a vowel and what proportions of vowels and consonants followed a consonant? To study this, he took 20,000 consecutive letters from the text of *Eugene Onegin* and did the calculation. He published his results in a paper [5] in 1913.

The results he obtained can be represented by the stochastic matrix

$$S = \begin{array}{c} \\ \text{Vow} \\ \text{Con} \end{array} \begin{array}{c} \begin{array}{cc} \text{Vow} & \text{Con} \end{array} \\ \begin{bmatrix} .128 & .872 \\ .663 & .337 \end{bmatrix} \end{array}.$$

So, for example, the probability of obtaining another vowel is .128 if we have just read a vowel and .663 if we have just read a consonant.

Squaring the matrix gives, to three decimal places,

$$S^2 \approx \begin{array}{c} \\ \text{Vow} \\ \text{Con} \end{array} \begin{array}{c} \begin{array}{cc} \text{Vow} & \text{Con} \end{array} \\ \begin{bmatrix} .595 & .405 \\ .308 & .692 \end{bmatrix} \end{array}.$$

This matrix tells us about what happens two letters further on. If we have just read a vowel, the probability that we will read a vowel two letters afterward is .595.

As we raise the matrix to higher and higher powers it approaches a limit. The limit matrix, again with entries rounded to three decimal places, is

$$\begin{array}{c} \\ \text{Vow} \\ \text{Con} \end{array} \begin{array}{c} \begin{array}{cc} \text{Vow} & \text{Con} \end{array} \\ \begin{bmatrix} .432 & .568 \\ .432 & .568 \end{bmatrix} \end{array}.$$

Notice that the entries in each of the columns are the same. This agrees with Markov's calculation that in the total sample the proportion of vowels is .432 and the proportion of consonants is .568.

Later, Claude Shannon used an analysis of English language using Markov chains as the basic example to illustrate the ideas of *entropy* and *information* in his landmark paper *A Mathematical Theory of Communication* [10]—this is the paper that forms the foundation for information theory.

The reader may have noticed that not all stochastic matrices have a limit as the power gets larger. For example, the matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ keeps oscillating between itself and $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ as the powers increase. However, if the stochastic matrix is positive there is a limit.

Theorem 2. *Let S be a positive, stochastic matrix, then* $\lim_{n\to\infty} S^n = T$ *exists. The matrix T is stochastic and the entries in any given column are all equal.*

There are many proofs of this widely known and widely applied theorem. The proof below is given by Grinstead and Snell [2].

Proof. Let S be a positive stochastic matrix. Let $\mathbf{1}$ denote the column vector with each entry equal to 1. Since S is stochastic, $\mathbf{1}$ is an eigenvector with eigenvalue 1. For every n, S^n will be positive and will have $\mathbf{1}$ as an eigenvector with eigenvalue 1. So S^n is stochastic for every n.

Let $\mathbf{v}$ be a column vector of the appropriate size with nonnegative entries that sum to 1. Each entry of the vector $S\mathbf{v}$ is a weighted average of the entries of $\mathbf{v}$. This means that the maximum of the entries of $S\mathbf{v}$ is less than or equal to the maximum of the entries of $\mathbf{v}$ and that the minimum of the entries of $S\mathbf{v}$ is greater than or equal to the minimum of the entries of $\mathbf{v}$. We examine this idea in more detail.

Let M_0 denote the maximum of the entries of $\mathbf{v}$ and M_i denote the maximum of the entries of $S^i\mathbf{v}$. Then we have

$$M_0 \geq M_1 \geq \cdots \geq M_n \geq \cdots \geq 0.$$

So $\lim_{n\to\infty} M_n$ exists.

Let m_0 denote the minimum of the entries of $\mathbf{v}$ and m_i denote the minimum of the entries of $S^i\mathbf{v}$. Then we have

$$m_0 \leq m_1 \leq \cdots \leq m_n \leq \cdots \leq 1.$$

So $\lim_{n\to\infty} m_n$ exists.

Let s denote the smallest entry of S. Recall that M_1 is a weighted average of the elements of $\mathbf{v}$. Then $M_1 \leq sm_0 + (1-s)M_0$ and $m_1 \geq sM_0 + (1-s)m_0$. Combining these two inequalities, we obtain $M_1 - m_1 \leq (1-2s)(M_0 - m_0)$. Induction yields $M_n - m_n \leq (1-2s)^n(M_0 - m_0)$. Since $0 < s \leq 1$, we know that $0 \leq 1 - 2s < 1$, and consequently $\lim\limits_{n\to\infty}(M_n - m_n) = 0$. This means that $\lim\limits_{n\to\infty} S^n\mathbf{v}$ exists and that all of its entries are equal.

Let $\mathbf{e}_i$ denote the column vector with a 1 in the ith entry and zeros elsewhere. Then by the argument above $\lim\limits_{n\to\infty} S^n\mathbf{e}_i$ exists and all of its entries are equal. This will be the ith column of $\lim\limits_{n\to\infty} S^n$. Consequently, $\lim\limits_{n\to\infty} S^n$ exists and each of its columns has all of its entries equal. ■

Now that we know the limit of powers of positive stochastic matrices, we return to the study of powers of positive matrices. We will use our result as a tool to help us in this more general case. The idea is, given any positive matrix A, we want to find a square matrix V, a positive stochastic matrix S and a number λ such that

$$\frac{1}{\lambda}A = VSV^{-1}.$$

Then we will have

$$\frac{1}{\lambda^n}A^n = VS^nV^{-1},$$

for any nonnegative integer n. Then we can use our result about limits of stochastic matrices to find the limit. The question is how do we find V, S, and λ?

Some results on powers of Jordan blocks

Positive stochastic matrices form a very special subset of positive matrices. Given a stochastic matrix, we can see immediately that it has a positive eigenvalue, namely, 1, and a positive eigenvector, $\mathbf{1}$. We want to show that any positive matrix also has a positive eigenvalue and a positive eigenvector. This is not obvious and requires some work. We will do this in the next two sections. To prove these facts we will use the Jordan normal form. (A good reference for matrix theory is [6] by C. Meyer. This is a rare book at the undergraduate level that also includes Perron's results and the more general Perron-Fröbenius theory. Much of what follows in the next two sections is based on this book.)

The aim of this section is to begin our study by proving two results concerning square positive matrices. The first is that any such matrix must have a nonzero eigenvalue. The second is that if the spectral radius of the matrix A is less than one, then A^n approaches the zero matrix as n goes to infinity. (The *spectral radius* of a square matrix M is the maximum element of $\{|\lambda| : \lambda \text{ is an eigenvalue of } M\}$.)

Recall that for any square matrix M, we can find a matrix P such that $P^{-1}MP = J$, where J is in Jordan normal form. The matrix J has Jordan blocks on the main diagonal and has zero entries elsewhere; more specifically, J has the following form:

$$J = \begin{bmatrix} J_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & J_d \end{bmatrix},$$

where, for each eigenvalue λ_i of M, the block J_i has the form

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & \mathbf{0} & \\ & & \ddots & \ddots & \\ & \mathbf{0} & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}.$$

We will look at powers of a Jordan block for which $|\lambda_i| < 1$. First, we consider the case when $\lambda_i = 0$. Here

$$J_i = \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & \mathbf{0} & \\ & & \ddots & \ddots & \\ & \mathbf{0} & & 0 & 1 \\ & & & & 0 \end{bmatrix}.$$

If J_i has size $d \times d$ then it is straightforward to check that J_i^d is the zero matrix. A matrix with the property that some power is the zero matrix is called *nilpotent*. So J_i is nilpotent. It is also clear that if J has only blocks that look like J_i then it will also be nilpotent. From this we can deduce that if all the eigenvalues of A are zero then A must be nilpotent. However, if A is a positive matrix, then no power of A can give us the zero matrix. This gives the following lemma.

Lemma 1. *If A is a square positive matrix, then it has at least one nonzero eigenvalue.*

We now look at powers of J_i for which $|\lambda_i| < 1$. Write

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & \mathbf{0} & \\ & & \ddots & \ddots & \\ & \mathbf{0} & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix} = \lambda_i I + \begin{bmatrix} 0 & 1 & & & \\ & 0 & 1 & \mathbf{0} & \\ & & \ddots & \ddots & \\ & \mathbf{0} & & 0 & 1 \\ & & & & 0 \end{bmatrix} = \lambda_i I + B.$$

The binomial theorem gives us

$$J_i^n = (\lambda_i I + B)^n = \sum_{k=0}^{n} \binom{n}{k} \lambda^{n-k} B^k.$$

If J_i has size $d \times d$ and $n > d$, we know from above, that B is nilpotent and that

$$J_i^n = \sum_{k=0}^{d} \binom{n}{k} \lambda^{n-k} B^k.$$

Now

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k!} \le \frac{n^k}{k!}.$$

Since the entries of B^k are either 0 or 1 each entry of $\binom{n}{k}\lambda^{n-k}B^k$ is between 0 and

$$\frac{n^k \lambda^{n-k}}{k!} = \frac{n^k \lambda^n}{k! \lambda^k}.$$

Either by repeatedly using L'Hôpital's rule or by noting that one term grows polynomially and one decays exponentially, we have $\lim_{n\to\infty} n^k \lambda^n = 0$. Hence, in the limit each entry approaches zero, so $\lim_{n\to\infty} J_i^n$ exists and equals the zero matrix.

The next lemma follows from the argument above.

Lemma 2. *Suppose that M is a square $m \times m$ (not necessarily positive, not necessarily real) matrix with spectral radius less than 1, then*

$$\lim_{n\to\infty} M^n = 0_{m\times m}.$$

A positive matrix has a positive eigenvalue and a positive eigenvector

First, we introduce some notation. For vectors of the same size, we will write $\mathbf{v}_1 \le \mathbf{v}_2$ if for each i the ith entry of $\mathbf{v}_1$ is less than or equal to ith entry of $\mathbf{v}_2$. We write $\mathbf{v}_1 < \mathbf{v}_2$ if for each i the ith entry of $\mathbf{v}_1$ is less than ith entry of $\mathbf{v}_2$.

Given a vector $\mathbf{v}$ that has complex entries, we let $|\mathbf{v}|$ denote the vector that has as ith entry the modulus of the ith entry of $\mathbf{v}$. We use the same notation for matrices, so if M is a matrix that has complex entries we let $|M|$ denote the matrix that has each entry of M replaced by its modulus.

We illustrate the notation with a simple example. Let $A = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. Then, $|A\mathbf{v}| = \begin{bmatrix} |c_1 v_1 + c_2 v_2| \\ |c_3 v_1 + c_4 v_2| \end{bmatrix}$ and $|A||\mathbf{v}| = \begin{bmatrix} |c_1||v_1| + |c_2||v_2| \\ |c_3||v_1| + |c_4||v_2| \end{bmatrix}$.

The triangle inequality then gives $|A\mathbf{v}| \le |A||\mathbf{v}|$. This argument holds in the general case and we obtain for any matrix M and vector $\mathbf{v}$ that

$$|M\mathbf{v}| \le |M||\mathbf{v}|.$$

In what follows, we will need the real number 0, the zero vector $\mathbf{0}$, and the zero matrix $0_{m\times m}$. The context should make clear which zero object is meant, but the notation for each is also distinct.

With these preliminaries out of the way we will state and prove the major result of this section.

Theorem 3. *Let A be a square, positive matrix with spectral radius ρ. Then*

1. *ρ is an eigenvalue and has a positive eigenvector, and*
2. *if $\mu \ne \rho$ is an eigenvalue, then $|\mu| < \rho$.*

Proof. Lemma 1 tells us that A has a nonzero eigenvalue, so $\rho > 0$. Let λ be an eigenvalue with $|\lambda| = \rho$ and let $\mathbf{v}$ be an eigenvector associated to λ. We will show that $|\lambda|$ is an eigenvalue and that $|\mathbf{v}|$ is an eigenvector.

First note

$$|\lambda||\mathbf{v}| = |\lambda\mathbf{v}| = |A\mathbf{v}| \le |A||\mathbf{v}| = A|\mathbf{v}|.$$

We will suppose for a contradiction that $|\lambda||\mathbf{v}| \ne A|\mathbf{v}|$. Then $A|\mathbf{v}| - |\lambda||\mathbf{v}| \ge \mathbf{0}$ and $A|\mathbf{v}| - |\lambda||\mathbf{v}| \ne \mathbf{0}$. Note that this does not imply that $A|\mathbf{v}| - |\lambda||\mathbf{v}| > \mathbf{0}$ because it could have some zero entries, but $A|\mathbf{v}| - |\lambda||\mathbf{v}|$ must have at least one positive entry and, since A is positive, we do know that $A(A|\mathbf{v}| - |\lambda||\mathbf{v}|) > \mathbf{0}$. So we have

$$A(A|\mathbf{v}|) > |\lambda|A|\mathbf{v}|.$$

To simplify the notation, we will let $A|\mathbf{v}| = \mathbf{w}$. The inequality above becomes $A\mathbf{w} > |\lambda|\mathbf{w}$.

We can find a small real $\epsilon > 0$ such that

$$A\mathbf{w} > |\lambda|(1+\epsilon)\mathbf{w}, \text{ or equivalently } \frac{1}{|\lambda|(1+\epsilon)}A\mathbf{w} > \mathbf{w}.$$

Again to simplify notation we let $\frac{1}{|\lambda|(1+\epsilon)}A = B$. Notice that B is also a positive matrix and we know that $B\mathbf{w} > \mathbf{w}$. These two facts imply that $B(B\mathbf{w}) > B\mathbf{w}$. Inductively, we obtain

$$\cdots > B^3\mathbf{w} > B^2\mathbf{w} > B\mathbf{w} > \mathbf{w} \ge \mathbf{0}.$$

From this we can deduce that $\lim\limits_{n\to\infty} B^n\mathbf{w} \ne \mathbf{0}$. However, the spectral radius of B is $\frac{1}{|\lambda|(1+\epsilon)}|\lambda| = \frac{1}{1+\epsilon} < 1$. So, by Lemma 2, we must have $\lim_{n\to\infty} B^n = 0_{m\times m}$ and so

$$\lim_{n\to\infty} B^n\mathbf{w} = \mathbf{0}.$$

We have now derived a contradiction and so the assumption that $|\lambda||\mathbf{v}| \ne A|\mathbf{v}|$ is false. We conclude that $|\lambda| = \rho$ is an eigenvalue and $|\mathbf{v}|$ is an eigenvector. It is clear that $|\mathbf{v}| \ge \mathbf{0}$, but since A is positive it must be the case that $A|\mathbf{v}| > \mathbf{0}$. Since $A|\mathbf{v}| = \rho|\mathbf{v}|$ we deduce that $|\mathbf{v}|$ is positive.

To complete the proof we need to show that λ has to actually equal ρ. First note that $|A\mathbf{v}| = |A||\mathbf{v}| = A|\mathbf{v}|$. We know from the triangle inequality that this can only happen if the entries of $\mathbf{v}$ are scalar multiples of one another or, equivalently, if $\mathbf{v} = k\mathbf{u}$ for some real vector $\mathbf{u}$. Then $|A\mathbf{u}| = A|\mathbf{u}|$ tells us that all the components of $\mathbf{u}$ are all nonpositive or all nonnegative. Since $A\mathbf{u} = \lambda\mathbf{u}$, the eigenvalue λ must be positive. So $\lambda = \rho$. ■

Proof of Perron's theorem

Proof. Given any positive matrix A, we now know that it has a positive eigenvalue λ equal to the spectral radius and that it has a positive right eigenvector. We can use the same argument to show that it also has a positive left eigenvector. We let $\mathbf{R}$ and $\mathbf{L}$ denote right and left eigenvectors with $\mathbf{LR} = 1$. We let

$$A = \begin{bmatrix} a_{1,1} & \dots & a_{1,m} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \dots & a_{m,m} \end{bmatrix}, \ \mathbf{L} = \begin{bmatrix} w_1 & \dots & w_m \end{bmatrix}, \text{ and } \mathbf{R} = \begin{bmatrix} v_1 \\ \vdots \\ v_m \end{bmatrix}.$$

We now introduce two diagonal matrices that come from $\mathbf{R}$,

$$V = \begin{bmatrix} v_1 & & 0 \\ & \ddots & \\ 0 & & v_m \end{bmatrix} \text{ and } V^{-1} = \begin{bmatrix} 1/v_1 & & 0 \\ & \ddots & \\ 0 & & 1/v_m \end{bmatrix},$$

and define S by $S = \frac{1}{\lambda}V^{-1}AV$. Clearly S is a positive matrix. As before, we let $\mathbf{1}$ denote the column vector with each entry equal to 1. Then

$$S\mathbf{1} = \frac{1}{\lambda}V^{-1}AV\mathbf{1} = \frac{1}{\lambda}V^{-1}A\mathbf{R} = \frac{1}{\lambda}V^{-1}\lambda\mathbf{R} = V^{-1}\mathbf{R} = \mathbf{1},$$

which implies that S is a positive stochastic matrix. We can now use our stochastic matrix theorem to deduce that

$$\lim_{n\to\infty} S^n = \lim_{n\to\infty} \frac{1}{\lambda^n}V^{-1}A^nV = T,$$

or equivalently

$$\lim_{n\to\infty} \frac{1}{\lambda^n}A^n = VTV^{-1}.$$

We know that T is a stochastic matrix and the entries in any given column are equal. Let

$$T = \begin{bmatrix} t_1 & t_2 & \dots & t_m \\ t_1 & t_2 & \dots & t_m \\ \vdots & \vdots & \ddots & \vdots \\ t_1 & t_2 & \dots & t_m \end{bmatrix}.$$

We now multiply VTV^{-1} on the left by $\mathbf{L}$. Since, for all n, $\mathbf{L}$ is a left eigenvector with eigenvalue 1 of $\frac{1}{\lambda^n}A^n$, we have $\mathbf{L}VTV^{-1} = \mathbf{L}$.

Now $\mathbf{L}V$ is a row vector with ith entry equal to $w_i v_i$. When we multiply $\mathbf{L}VT$, we get a row vector with ith entry equal to

$$t_i \sum_{j=1}^{m} w_j v_j = t_i.$$

Finally, multiplying on the right by V^{-1} gives us a row vector with ith entry equal to t_i/v_i. Since $\mathbf{L}$ is an eigenvector with eigenvalue 1, we know the ith entry equals w_i. So for all i, $t_i/v_i = w_i$, or equivalently $t_i = w_i v_i$.

It is now clear that the jth column of TV^{-1} has each entry equal to w_j. Consequently, the entry of VTV^{-1} in the ith row and jth column is $v_i w_j$. So

$$\lim_{n\to\infty} \frac{1}{\lambda^n} A^n = \mathbf{RL}.$$

The only thing left to prove is that the left and right eigenspaces associated with λ are one-dimensional. This follows from the fact that the limit matrix is the outer product of two vectors. Let $\mathbf{w}$ any column vector. Then $\mathbf{RLw} = \mathbf{R}(\mathbf{Lw})$, but $\mathbf{Lw}$ is just a constant. So $\mathbf{RL}$ projects all vectors onto scalar multiples of $\mathbf{R}$. Similarly, multiplying row vectors on the right by $\mathbf{RL}$ sends the vectors to scalar multiples of $\mathbf{L}$. ■

Concluding comments

As we noted in the introduction, Oskar Perron was the first to prove this theorem. Actually he proved a slightly stronger result. A square matrix is called *primitive* if some power of it is positive. Perron's theorem is true if the hypothesis that the matrix is positive is replaced with it being primitive. Slight modifications have to be made to our proofs to deal with primitive matrices that are not positive, but these are fairly easy to do.

Perron also showed that the spectral radius was not just *geometrically simple*, but *algebraically simple*. An eigenvalue is *geometrically simple* if the associated eigenspace is one-dimensional. An eigenvalue is *algebraically simple* if it is a simple, not repeated, root of the characteristic polynomial. If an eigenvalue is algebraically simple, then it must be geometrically simple, but the converse is not true. As an example consider the following matrices:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}.$$

For A, the eigenvalue 1 is neither geometrically nor algebraically simple. For B, the eigenvalue 1 is geometrically, but not algebraically simple. For C, the eigenvalue 1 is both geometrically and algebraically simple.

We have shown that the spectral radius is a geometrically simple eigenvalue. Perron's result is stronger, but needs another argument. We do not give this here because our main aim was to get to the result about the limits of powers of positive matrices and the fact that the spectral radius is algebraically simple is not needed to prove this.

Soon after Perron published his results, Georg Fröbenius extended the arguments to nonnegative matrices [**1**]. There are simple proofs of the results of Fröbenius that use directed graphs. Two good books that give these proofs are [**3**] and [**11**].

Nowadays the study of nonnegative matrices is called Perron-Fröbenius theory. This theory has many applications. Population models in biology, models of markets in economics, applications to thermodynamics in physics have all used this theory.

(See [4] for a more complete listing.) One example that has impact on us daily is the Google PageRank algorithm. Here the underlying matrix has size $n \times n$ where n is the number of webpages and the entries count the links between pages. This has been described as the world's largest matrix computation, see [7], with n being about 2.7 billion when the paper was written in 2002. There are many more pages today!

Another important application is in the study of dynamical systems, especially in the sub-discipline of symbolic dynamics. The book [11] by Sternberg introduces various areas of dynamical systems that contain Perron-Fröbenius theory. It uses only elementary ideas and is aimed at undergraduates. The book [3] by Lind and Marcus is somewhat more specialized, but is also accessible to undergraduates. They are both recommended for readers who would like to learn more about this subject.

REFERENCES

[1] Fröbenius, G. (1912). Über matrizen aus nicht negativen elementen. *Sitzungsberichte Preussische Akademie der Wissenschaft*, Berlin, pp. 456–477.

[2] Grinstead, C. M., Snell, J. L. (1997). *Introduction to Probability*. Providence, RI: American Mathematical Society. Available at: `https://math.dartmouth.edu/prob/prob/prob.pdf`

[3] Lind, D., Marcus, B. (1995). *An Introduction to Symbolic Dynamics and Coding*. Cambridge: Cambridge University Press.

[4] MacCluer, C. R. (2000). The many proofs and applications of Perron's theorem. *SIAM Rev.* 42(3): 487–498.

[5] Markov, A. A. (1913). Primer statisticheskogo issledovaniya nad tekstom Evgeniya Onegina, illyustriruyuschij svyaz ispytanij v cep. *Izv. Akad. Nauk., SPb, VI Seriya* 7(93): 153–162.

[6] Meyer, C. (2000). *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM).

[7] Moler, C. (2002). The world's largest matrix computation. *Matlab News Notes*. 10: 12–13.

[8] Perron, O. (1907). Grundlagen für eine theorie des Jacobischen kettenbruchalgorithmus. *Math. Ann.*, 64: 11–76.

[9] Perron, O. (1907). Zur theorie der matrices. *Math. Ann.* 64: 248–263.

[10] Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27(3): 379–423.

[11] Sternberg, S. (2010). *Dynamical Systems*. Mineola, NY: Dover Publications, Inc.

Summary. In 1907, Oskar Perron proved a beautiful theorem about powers of positive matrices. This theorem is not just of theoretical interest, but has a number of important applications. We explain what the theorem says and give an elementary proof.

CHRIS BERNHARDT (MR Author ID: 35645) received his Ph.D. in mathematics from the University of Warwick in 1980. After completing his Ph.D., he moved to the United States joining the faculty at Fairfield University in 1987. He has been there since. His research has been in one-dimensional dynamics, but he often gets distracted by beautiful results in other areas of mathematics.

ACROSS

1. "Obviously!"
6. Point of non-differentiability, sometimes
10. Baseball legend Ripken
13. Location of Brigham Young University
14. A 66-Across may be this style
15. "The news is ____ 5 o'clock."
17. Picture puzzle
18. Emerald ____
19. Certain surgery, for short
20. * She (also known as "mathgrrl") will give the MAA Chan Stanek Lecture for Students about learning from failure in 3D printing
23. Interest may be charged on them
25. Nothing to ____ at
26. * She (also a concert pianist) will give an MAA Invited Address about inclusion and exclusion in mathematics (but not the combinatorial kind)
30. NBC comedy show
31. Pester
32. Four-door car
36. Look for
38. Baseball player and manager Kapler, and actor and poker player Kaplan
40. Geometry postulate: ____ Angle Side
41. Black tea named for a region in northeast India
43. $z^2 = x^2 + y^2$ (Cartesian) and $\phi = \frac{\pi}{2}$ (spherical) are examples of these
45. NASA mission watching the Sun since 2010
46. * She (of Bucknell Univ.) will give the AWM-MAA Etta Zuber Falconer Lecture about ellipses
49. 1969 western film starring Elvis Presley (and his only film in which he didn't sing)
52. Conglomerate
53. * He (of Dartmouth Coll.) will give the Pi Mu Epsilon J. Sutherland Frame Lecture on large random systems
57. Arduous journey
58. Jack-o'-lantern feature
59. Herculean main character of a comic book and TV series from the 1980s
63. Scientific agcy. under the Dept. of Commerce that issues weather warnings
64. Graph of $f(x) = ax + b$
65. Martini garnish
66. Spectroscopy technique: Abbr.
67. Lyric from 1993 Ace of Base hit: "____ the sign"
68. Arthur, inventor of the crossword puzzle (!)

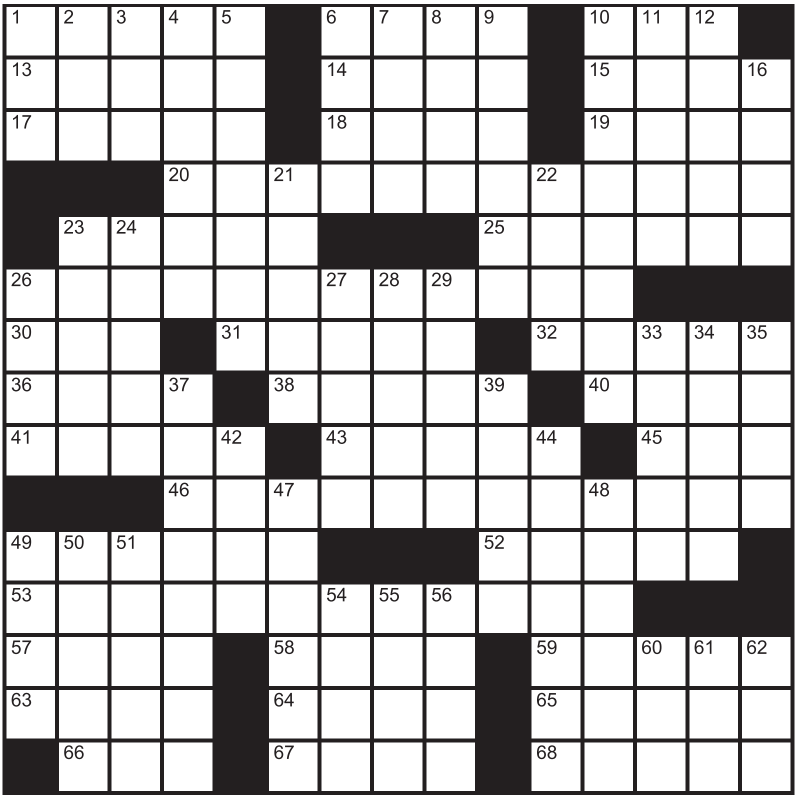
DOWN

1. "Fresh Air" airer
2. Miner's find
3. Piece of info on a govt. form, usually
4. Latinate plural of the flesh that hangs in the back of your throat
5. Biblical cry for divine help: "____; blessed is the one who comes in the name of the Lord."
6. Fiber from the outer husk of a coconut
7. ____ Major (constellation)
8. NaCl
9. "The magic word"
10. Where many attendees of Math Fest teach or take classes
11. Japanese cartoon art
12. Capital of Bolivia
16. Specific pitch
21. Taking advantage of
22. Landers and others
23. Crescent shapes formed by two intersecting circles
24. Makes eyes at
26. US law that replaced NCLB (No Child Left Behind) in 2015
27. Have ____ up your sleeve
28. Common business-oriented language (this practically gives it away)
29. Shenzi, Banzai, or Ed in "The Lion King"
33. Planar regions bounded by circles
34. Capital of Ethiopia: ____ Ababa
35. Element #10, a noble gas
37. 6174 is ____'s constant, named for the Indian mathematician who discovered its bizarre property
39. Mathematician Irving credited with introducing the term "$C^\star$-algebra", or mathematician Graeme who was President of the London Mathematical Society from 2011 to 2013
42. Guitarist Johnny of The Smiths
44. By one manner or another
47. Protagonist of "The Jungle Book"
48. Not often at all
49. Hook and Crunch hold this naval title: Abbr.
50. Long-legged, long-necked bird
51. Mr. T's group
54. Part of the eye that controls the size of the pupil
55. Girl: Spanish
56. Understood
60. Least element of a set, for short
61. Stormy Daniels has won several awards given out by this org.
62. French word used before one's maiden name

Math. Mag. **91** (2018) 228–229. doi:10.1080/0025570X.2018.1456157

MathFest 2018

BRENDAN SULLIVAN
Emmanuel College
Boston, MA 02115
sullivanb@emmanuel.edu



Clues start at left, on page 228. The Solution is on page 183.

This crossword puzzle and solution are open access and available for download on Taylor & Francis' website for THE MAGAZINE (https://maa.tandfonline.com).

Crossword Puzzle Creators

If you are interested in submitting a mathematically themed crossword puzzle for possible inclusion in MATHEMATICS MAGAZINE, please contact the editor at mathmag@maa.org.

PROBLEMS

Proposals

To be considered for publication, solutions should be received by November 1, 2018.

2046. *Proposed by Ioan Băetu, Botoşani, Romania.*

For integers m, n such that $1 \leq m < n$, let S_n be the group of all permutations of $\{1, 2, \ldots, n\}$, let F be the set of permutations $\sigma \in S_n$ such that $\sigma(m) < \sigma(m+1) < \cdots < \sigma(n)$, and let T be the set of transpositions in F. Prove that there exists a unique subgroup G of S_n such that $T \subset G \subset F$.

2047. *Proposed by George Stoica, Saint John, New Brunswick, Canada.*

Let (a_n) be a sequence of nonzero real numbers such that

$$\lim_{n\to\infty} n\left(\left|\frac{a_n}{a_{n+1}}\right| - 1\right)$$

exists and is strictly positive. Prove or disprove: The sequence (a_n) is necessarily convergent.

2048. *Proposed by Julien Sorel, Piatra Neamt, PNI, Romania.*

Three points A, B, C are chosen uniformly at random in the three-quarter disk

$$\mathcal{Q} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1, \text{ and either } x \leq 0 \text{ or } y \leq 0\}$$

obtained by removing the first quadrant of the unit disk. What is the probability that the origin $O = (0, 0)$ lies inside $\triangle ABC$?

Math. Mag. **91** (2018) 230–238. doi:10.1080/0025570X.2018.1456271 © Mathematical Association of America

We invite readers to submit original problems appealing to students and teachers of advanced undergraduate mathematics. Proposals must always be accompanied by a solution and any relevant bibliographical information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. Submitted problems should not be under consideration for publication elsewhere.

Proposals and solutions should be written in a style appropriate for this MAGAZINE.

Authors of proposals and solutions should send their contributions using the Magazine's submissions system hosted at http://mathematicsmagazine.submittable.com. *More detailed instructions are available there. We encourage submissions in PDF format, ideally accompanied by LaTeX source. General inquiries to the editors should be sent to* mathmagproblems@maa.org.

2049. *Proposed by Scott Duke Kominers, Harvard University, Cambridge, MA.*

Show that any finite set of squares in the plane (possibly of different sizes and not necessarily disjoint) has a subset consisting of non-overlapping squares that together cover at least 7% of the area covered by the full set.

2050. *Proposed by Sung Soo Kim, Hanyang University, Ansan, Korea.*

Find the number of sequences $a_1, a_2, \ldots, a_9$ in $\{1, 2, 3\}$ such that

(i) $a_1 = a_2 = 1$, and
(ii) the nine pairs $(a_1, a_2), (a_2, a_3), \ldots, (a_8, a_9), (a_9, a_1)$ are the same as the nine pairs $(1, 1), (1, 2), \ldots, (3, 2), (3, 3)$ in some order.

Quickies

1081. *Proposed by Lokman Gökçe, Adana, Turkey.*

Let $\triangle ABC$ be an obtuse "golden triangle" with angles $\angle BAC = 108^\circ$, $\angle BCA = \angle CBA = 36^\circ$. Let R and r be the circumradius and inradius of $\triangle ABC$, respectively, and let I be its incenter. Show that

$$AI + BI + CI = 3R - 2r, \qquad \text{and}$$
$$AI^2 + BI^2 + CI^2 = 4R^2 - 6Rr.$$

1082. *Proposed by Michel Bataille, Rouen, France.*

Let f be a nonnegative, continuous function on $[0, 1]$. Prove that

$$\frac{1}{6}\int_0^1 f(x)\,dx + \frac{1}{2}\int_0^1 (f(x))^5\,dx \geq \frac{1}{3}\int_0^1 (f(x))^4\,dx + \frac{1}{5}\int_0^1 (f(x))^2\,dx.$$

Solutions

When an operation on primes commutes **June 2017**

2021. *Proposed by Mihai Caragiu, Ohio Northern University, Ada, OH.*

For an integer $n > 1$, let $\Pi(n)$ be the greatest prime factor of n. Consider a binary operation '$*$', on the set $\mathcal{P} = \{2, 3, 5, 7, \ldots\}$ of all primes, defined by $p * q = \Pi(2p + q)$ for all primes p, q. Find all distinct primes p, q such that $p * q = q * p$.

Solution by Abhay Goel (student), Kalamazoo College, Kalamazoo, MI.
We have $2 * 5 = \Pi(9) = 3 = \Pi(12) = 5 * 2$ and $2 * 23 = \Pi(27) = 3 = \Pi(48) = 23 * 2$, so we have solutions $\{p, q\} = \{2, 5\}$ or $\{2, 23\}$. We prove that there are no others.

Assume that distinct primes p, q satisfy $p * q = q * p = r$, say. Thus, r is a prime, $r \mid 2p + q$, and $r \mid 2q + p$. Hence, $r \mid 3(p + q) = (2p + q) + (2q + p)$. Since r is prime, either $r = 3$ or $r \mid p + q$. However, if $r \mid p + q$, then $r \mid p = (2p + q) - (p + q)$ and $r \mid q = (2q + p) - (p + q)$; since p, q are different primes, they cannot have a common prime factor r. Thus, we must have $r = 3$. Since r is the largest prime divisor of each

$2p+q$ and $2q+p$, both these numbers must be of the form 2^a3^b for some integer exponents $a \geq 0$, $b \geq 1$. We prove that $p=2$ or $q=2$. Indeed, if we had $p \neq 2$ and $q \neq 2$, then p and q, and hence $2p+q$ and $2q+p$, would both be odd; thus,

$$2p+q=3^b \quad \text{and} \quad 2q+p=3^c, \qquad \text{for some integers } b, c \geq 1.$$

Solving for p, q we obtain,

$$p = 2\cdot 3^{b-1} - 3^{c-1} \quad \text{and} \quad q = 2\cdot 3^{c-1} - 3^{b-1}.$$

If we had $b>1$ and $c>1$ then the expressions above show that $3 \mid p$ and $3 \mid q$, contradicting the assumption that p and q are distinct primes; therefore, $b=1$ or $c=1$. Now, if $b=1$, then $p=2-3^{c-1}<2$; if $c=1$, then $q=2-3^{b-1}<2$, a contradiction. Thus, we conclude that $p=2$ or $q=2$. Without loss of generality, assume $p=2$ and $q>2$. Since $2p+q=4+q$ is odd,

$$4+q=3^b \quad \text{and} \quad 2q+2=2^a3^c \qquad \text{for some integers } a, b, c \geq 1.$$

It follows that

$$3(3^{b-1}-1) = 3^b - 3 = q+1 = 2^{-1}(2q+2) = 2^{a-1}3^c.$$

Clearly, $3^c \leq 2^{a-1}3^c < 3^b$, so we must have $b > c \geq 1$. Thus, the left-hand side of the equation above is divisible by 3 but not by 3^2, while the right-hand side is divisible by 3^c. Hence, we must have $c=1$, so the equation above gives

$$3^{b-1}-1 = 2^{a-1} \qquad \text{for integers } a, b \geq 1.$$

The solutions to this Diophantine equation are well known, namely $(a-1, b-1) = (1,1)$ or $(3,2)$ (a proof is given below for completeness). Thus, $b=2$ or $b=3$, giving $q=3^2-4=5$ or $q=3^3-4=23$ (and $p=2$), and so proving that the only solutions to the problem are $\{2, 5\}$ and $\{2, 23\}$. Now we prove:

If $2^n+1=3^m$ for nonnegative integers n, m, then either $n=1$ and $m=1$, or else $n=3$ and $m=2$.

Clearly, these are the all the solutions with $n \leq 3$, so it only remains to show that there are no solutions with $n \geq 4$. If $n \geq 4$, then 2^n is divisible by $2^4=16$, so we have

$$3^m = 2^n + 1 \equiv 1 \pmod{16}.$$

The smallest proper power of 3 satisfying the above congruence is $3^4 = 81 \equiv 1 \pmod{16}$, hence $4 \mid m$. However, we also have $3^4 = 81 \equiv 1 \pmod 5$, hence $3^m \equiv 1 \pmod 5$, so

$$2^n = 3^m - 1 \equiv 1 - 1 = 0 \pmod 5.$$

Since no power of 2 is a multiple of 5, this contradiction shows that there are no solutions with $n \geq 4$, completing the proof.

Also solved by Michel Bataille (France), Brian Beasley, Bruce Burdick, Robert Calcaterra, John Christopher, Joseph DiMuro, Wenwen Du & Paul Peck, James Duemmel, Dmitry Fleischman, Michael Goldenberg & Mark Kaplan, Russell Gordon, Graham Lord, Rick Mabry, Missouri State University Problem Solving Group, Michael Reid, Celia Schacht, Nicholas Singer, Skidmore College Problem Group, John Smith, David Stone & John Hawkins, Joseph Walsh, Edward White & Roberta White, and the proposer. There were two incomplete or incorrect solutions.

Poorly distributed irrational geometric sequences (modulo 1) **June 2017**

2022. *Proposed by Mihàly Bencze, Bucharest, Romania.*

Given an irrational number $\beta > 1$, show that there exists a number $\alpha \in (1, 2)$, such that

$$0 < \{\alpha\beta^n\} < \frac{1}{\beta - 1} \qquad \text{for all } n \in \mathbb{N},$$

where $\{x\}$ denotes the fractional part of x.

Solution by Northwestern University Math Problem Solving Group, Northwestern University, Evanston, IL.

Let $\beta > 1$ be irrational. If $1 < \beta \leq 2$, then $1/(\beta - 1) \geq 1 > \{\alpha\beta^n\}$ for all n and all α, so we only need to choose α ensuring that the inequality $\{\alpha\beta^n\} > 0$ holds, i.e., that $\alpha\beta^n$ is not an integer, for any $n \in \mathbb{N}$. In case β is algebraic, choose α to be any transcendental number in $(1, 2)$; if β is transcendental, let α be any irrational algebraic number in $(1, 2)$. Then, $\alpha\beta^n$ is transcendental (or, when $n = 0$, at least irrational) and hence not an integer for arbitrary $n \in \mathbb{N}$, completing the proof when $\beta \leq 2$.

Henceforth, assume $\beta > 2$. Define a sequence $(d_n)_{n\geq 0}$ of integers recursively by:

$$d_0 = 1, \qquad \text{and} \qquad d_n = \lceil \beta d_{n-1} \rceil \quad \text{for } n \geq 1,$$

where $\lceil x \rceil$ denotes the least integer no less than x. Since d_0 and β are positive, it is clear that $d_n \geq 1$ for all n. Let $\alpha_n = d_n\beta^{-n}$. Since d_n is a positive integer and β is irrational, βd_n is not an integer, so $d_n = \lceil \beta d_{n-1} \rceil > \beta d_{n-1}$ for all $n \geq 1$. Thus, $\alpha_n = d_n\beta^{-n} > d_{n-1}\beta^{-(n-1)} = \alpha_{n-1}$ for all $n \geq 1$, so $(\alpha_n)_{n\geq 0}$ is strictly increasing. Next, we have, for $n \geq 1$:

$$\begin{aligned} 0 < \alpha_n - \alpha_{n-1} &= d_n\beta^{-n} - d_{n-1}\beta^{-(n-1)} = (d_n - \beta d_{n-1})\beta^{-n} \\ &= (\lceil \beta d_{n-1} \rceil - \beta d_{n-1})\beta^{-n} < \beta^{-n}, \end{aligned}$$

since $\lceil x \rceil - x < 1$ for all x. Therefore, the sequence (α_n) is strictly increasing, and we have

$$1 = \alpha_0 < \alpha_n = \alpha_0 + \sum_{i=1}^{n}(\alpha_i - \alpha_{i-1}) < 1 + \sum_{i=1}^{n} \beta^{-n} = \frac{\beta}{\beta - 1}$$

for all $n \geq 1$. By the monotone sequence theorem, the sequence (α_n) has a limit $\alpha = \lim_{n\to\infty} \alpha_n$ such that $1 < \alpha \leq \beta/(\beta - 1) < 2$ (by the assumption $\beta > 2$), hence $\alpha \in (1, 2)$, and moreover $\alpha_n < \alpha$ for all n since (α_n) is strictly increasing; furthermore,

$$0 < \alpha - \alpha_n = \sum_{i=n+1}^{\infty} (\alpha_i - \alpha_{i-1}) < \sum_{i=n+1}^{\infty} \beta^{-i} = \frac{\beta^{-n}}{\beta - 1},$$

hence $0 < \alpha\beta^n - d_n = \beta^n(\alpha - \alpha_n) < 1/(\beta - 1) < 1$ (since $\beta > 2$). It follows that $\{\alpha\beta^n\} = \alpha\beta^n - d_n \in (0, 1/(\beta - 1))$ for all n, concluding the proof.

Editor's Note. Celia Schacht and George Stoica independently communicated that the property stated in the problem was proved by R. Tijdeman in 1972. (R. Tijdeman, Note on Mahler's 3/2-problem, *Det Kongelige Norske Videnskabers Selskabs Skrifter* **16** (1972) 1–4.) Celia Schacht further pointed out that A. Dubickas has improved upon Tijdeman's result in recent years, and also remarked that the numbers α satisfying the stated property for a fixed irrational $\beta > 2$ form a zero-measure set of exceptions to Weyl's equidistribution property (H. Weyl, Über die Gleichverteilung von Zahlen modulo Eins, *Mathematische Annalen* **77** (1916) 313–352): For any irrational $\beta > 1$ and almost every $\alpha \in \mathbb{R}$, the fractional parts $\{\alpha\beta^n\}$ $(n = 0, 1, \ldots)$ are uniformly distributed in $[0, 1)$ (in particular, they are dense in $[0, 1)$).

Also solved by Robert Calcaterra, Souvik Dey (India), Soo Young Kim (South Korea), Reiner Martin (Germany), Michael Reid, Celia Schacht, Nicholas C. Singer, George Stoica (Canada), Enrique Treviño, and the proposer. There was one incomplete or incorrect solution.

Expressing natural numbers in the form $a + a + b + c$ **June 2017**

2023. *Proposed by Mircea Merca, Craiova, Romania.*

For every natural number n, let $f(n)$ be the number of representations of n in the form

$$n = a + a + b + c$$

where a, b, c are distinct positive integers such that $b < c$. Show that there are infinitely many values of n such that $f(n + 1) < f(n)$.

Solution by Michael Reid, University of Central Florida, Orlando, FL.
For a positive integer m, let $g(m)$ denote the number of expressions $m = b + c$ with b, c positive integers and $b < c$. For each integer b in the range $1 \le b < m/2$ there is exactly one such expression, so $g(m) = \lfloor (m - 1)/2 \rfloor$. Next, let $\widetilde{f}(n)$ be the number of expressions $n = a + a + b + c$ with a, b, c positive integers and $b < c$. For each a in the range $1 \le a < n/2$ there are exactly $g(n - 2a)$ such expressions, hence

$$\begin{aligned}\widetilde{f}(n) = \sum_{a=1}^{\lfloor \frac{n-1}{2} \rfloor} g(n - 2a) &= \sum_{a=1}^{\lfloor \frac{n-1}{2} \rfloor} \left\lfloor \frac{n - 2a - 1}{2} \right\rfloor = \sum_{a=1}^{\lfloor \frac{n-1}{2} \rfloor} \left(\left\lfloor \frac{n-1}{2} \right\rfloor - a \right) \\ &= \frac{1}{2} \left\lfloor \frac{n-1}{2} \right\rfloor^2 - \frac{1}{2} \left\lfloor \frac{n-1}{2} \right\rfloor = \frac{1}{2} \left\lfloor \frac{n-1}{2} \right\rfloor \left\lfloor \frac{n-3}{2} \right\rfloor .\end{aligned}$$

Finally, to compute $f(n)$, we subtract from $\widetilde{f}(n)$ the number of expressions $n = a + a + b + c$ where either $a = b$ or $a = c$. For each $a \ge 1$ with $n - 2a > a$, there is exactly one such expression to exclude (namely $n = a + a + a + (n - 3a)$ if $a < n - 3a$, $n = a + a + (n - 3a) + a$ if $n - 3a < a$), except when $n - 3a = a$, in which case there are none to exclude. Thus, $f(n) = \widetilde{f}(n) - \lfloor (n - 1)/3 \rfloor + \delta(n)$, where $\delta(n) = 1$ if 4 divides n; otherwise, $\delta(n) = 0$. This gives the explicit formula

$$f(n) = \frac{1}{2} \left\lfloor \frac{n-1}{2} \right\rfloor \left\lfloor \frac{n-3}{2} \right\rfloor - \left\lfloor \frac{n-1}{3} \right\rfloor + \delta(n).$$

For any positive integer n of the form $12k + 9$ (k a nonnegative integer), we have:

$$\begin{aligned} f(n) &= f(12k + 9) = \tfrac{1}{2}(6k + 4)(6k + 3) - (4k + 2) + 0 = 18k^2 + 17k + 4, \\ f(n + 1) &= f(12k + 10) = \tfrac{1}{2}(6k + 4)(6k + 3) - (4k + 3) + 0 = 18k^2 + 17k + 3. \end{aligned}$$

This provides infinitely many n satisfying $f(n + 1) = f(n) - 1 < f(n)$.

Also solved by Dawson Bolus, Robert Calcaterra, Dmitry Fleischman, Graham Lord, Missouri State University Problem Solving Group, Celia Schacht, Nicholas Singer, David Stone, Brendan Sullivan, Enrique Treviño, Edward White, and the proposer.

Multiple binary expansions of 1 to a fractional base **June 2017**

2024. *Proposed by George Stoica, Saint John, New Brunswick, Canada.*

A *binary expansion* is an expression of the form

$$0.d_1d_2d_3\ldots d_i\ldots$$

where each numeral (digit) d_i is either 0 or 1 ($i = 1, 2, 3, \ldots$). Given a real number $\beta > 1$ (called the *base*) the *base-β value* of the binary expansion above is

$$(0.d_1d_2d_3\ldots)_\beta = \sum_{i=1}^{\infty} \frac{d_i}{\beta^i}.$$

(i) If $1 < \beta < 2$, show that some binary expansion has base-β value equal to 1; in fact, if $\beta \leq \phi$, where $\phi = (1 + \sqrt{5})/2$ is the golden ratio, then there are infinitely many such expansions.

(ii) Find all binary expansions with value 1 when $\beta = \phi$.

Editor's Note. Due to an editorial error, part (*i*) of the original published problem (*Math. Mag.* **90** (2017) 231–238) asked for a proof that infinitely many binary expansions have base-β value equal to 1 for any $\beta \in (1, 2)$. This, however, is not true: The Komornik–Loreti constant $\gamma \doteq 1.7872356\ldots$ has the property that the value 1 corresponds to a unique base-γ binary expansion, but to multiple binary expansions to any base $\beta \in (1, \gamma)$. (V. Komornik and P. Loreti, Unique Developments in Non-Integer Bases, *Amer. Math. Monthly* **105** (1998) 636–639.) The critical base γ is the unique positive number such that

$$1 = (0.d_1d_2\ldots)_\gamma = (0.1101001100101101\ldots)_\gamma$$

where $d_1, d_2, \ldots$ is the Thue–Morse sequence recursively defined by $d_0 = 0$, $d_{2n} = d_n$, and $d_{2n+1} = 1 - d_n$ for $n = 0, 1, 2, \ldots$. The solution below only proves that infinitely many expansions with value 1 exist for $\beta \leq \phi \doteq 1.618\ldots$. We apologize for this mistake.

Solution by the editors.
(*i*) Fix $\beta \in (1, 2)$. As customary, we identify a binary expression $\mathbf{d} = 0.d_1d_2\ldots$ with its value $(\mathbf{d}) = (0.d_1d_2\ldots)_\beta$. A terminating binary expansion $0.d_1d_2\ldots d_n$ means $0.d_1\ldots d_n000\ldots$, while $0.d_1d_2\ldots d_k\overline{1}$ means $0.d_1d_2\ldots d_k111\ldots$ (an expansion with a tail of unit digits). By $\mathbf{d}_{\leq n}$ (resp., $\mathbf{d}_{<n}$) we mean the expansion $0.d_1\ldots d_n$ (resp., $0.d_1\ldots d_{n-1}$). When we speak about the digits $d_1, d_2, \ldots$ of an expansion, we ignore the leading "0."; in particular, $\mathbf{d}_{\leq 0} = \mathbf{d}_{<1} = 0.$ is a valid expansion regarded as having no digits and value zero. Let $B = 1/(\beta - 1) = 0.\overline{1}$. Note that $B > 1$ since $1 < \beta < 2$. We prove that any value $\upsilon \in [0, B]$ has a binary expansion $\mathbf{d}$. Recursively define the digits d_i ($i = 1, 2, \ldots$) by letting

$$d_i = \begin{cases} 0, & \text{if } (\mathbf{d}_{<i}) + \beta^{-i} > \upsilon, \\ 1, & \text{if } (\mathbf{d}_{<i}) + \beta^{-i} \leq \upsilon. \end{cases} \tag{1}$$

Straightforward induction shows that $0 \leq \upsilon - (\mathbf{d}_{\leq k}) \leq \beta^{-k}B$ for $k = 0, 1, 2, \ldots$; moreover, if $d_k = 0$, the stronger bound $\upsilon - (\mathbf{d}_{\leq k}) \leq \beta^{-k}$ holds (in particular, $d_k = 0$

cannot be followed by an infinite sequence of unit digits). Clearly, $\beta^{-k} \to 0$ as $k \to \infty$ (since $\beta > 1$), so the equality $v = 0.d_1 d_2 \ldots$ follows, proving the existence of an expansion with given value $v \in [0, B]$.

We call an expansion $\mathbf{d}$ chosen according to (1) *greedy* since at every step the digit 1 is chosen if at all possible. Obviously, any expansion (whether greedy or not) with value $v \in (0, B)$ must have both zero and unit digits.

Next, assume $\beta \leq \phi = (1 + \sqrt{5})/2$. We have $(1 + \beta^{-1})(\beta - 1) = \beta - \beta^{-1} \leq \phi - \phi^{-1} = 1$, hence $1 + \beta^{-1} \leq 1/(\beta - 1) = B$. Consider the base-$\beta$ greedy expansion $\mathbf{d} = 0.d_1 d_2 \ldots$ with value $v = 1$. Since $0 < 1 < B$, the expansion $\mathbf{d}$ has both zero and unit digits. Since $\mathbf{d}$ is greedy, it does not end with a tail of unit digits, so there is a position k such that $d_k = 1$ and $d_{k+1} = 0$. By construction of the greedy expansion, we have $(\mathbf{d}_{\leq k}) = (\mathbf{d}_{<k}) + \beta^{-k}$ and $(\mathbf{d}_{<k}) + \beta^{-k} \leq 1 < (\mathbf{d}_{<k}) + \beta^{-k} + \beta^{-(k+1)}$. Let $v' = \beta^k[1 - (\mathbf{d}_{<k})]$; thus, we have $1 \leq v' < 1 + \beta^{-1} \leq B$. (Note the strict inequality.) The value v' has a greedy expansion $\mathbf{e} = 0.e_1 e_2 \ldots$ with $e_1 = 1$ (since $v' \geq 1 > \beta^{-1}$), giving an expansion

$$1 = (\mathbf{d}') = 0.d_1' d_2' \ldots = 0.d_1 d_2 \ldots d_{k-1} 01 e_2 e_3 \ldots$$

different from the greedy expansion $\mathbf{d}$. This construction may be iterated: the greedy expansion $\mathbf{e}$ of $v' \in [1, B)$ has both zero and unit digits, hence for some position k' we have $e_{k'} = 1$ and $e_{k'+1} = 0$, from which we obtain different expansion $\mathbf{e}' = 0.e_1 e_2 \ldots e_{k'-1} 01 f_2 f_3 \ldots$ with value v', and hence a new expansion

$$1 = (\mathbf{d}'') = 0.d_1 \ldots d_{k-1} 01 e_2 \ldots e_{k'-1} 01 f_2 f_3 \ldots.$$

Successively repeating this procedure, infinitely many different expansions $\mathbf{d}, \mathbf{d}', \mathbf{d}'', \ldots$ with value 1 are obtained, provided $\beta \leq \phi$.

(*ii*) For the base $\beta = \phi$, we have $1 = \phi^{-1} + \phi^{-2}$, hence $\phi^{-i} = \phi^{-(i+1)} + \phi^{-(i+2)}$. Hence, from any (nonzero) terminating expansion one can obtain another by replacing the trailing digit 1 by the digits 011 (i.e., "100" becomes "011"). Starting with the greedy expansion

$$\mathbf{a} = 0.11 = \phi^{-1} + \phi^{-2} = 1,$$

we obtain the following expansions with value 1:

$$\mathbf{a}' = 0.1011, \quad \mathbf{a}'' = 0.101011, \quad \ldots, \quad \mathbf{a}^{(n)} = 0.1010\ldots1011, \quad \ldots,$$

where in $\mathbf{a}^{(n)}$ there are n pairs "10" before the trailing "11" (thus, $\mathbf{a}$ above is $\mathbf{a}^{(0)}$). (These are precisely the expansions obtained by application of the iterative procedure in the solution to part (*i*), starting with $\mathbf{a} = 0.11$.) We also have

$$\mathbf{b} = 0.0\overline{1} = 0.01111\ldots = \phi^{-2} + \phi^{-3} + \cdots = \frac{\phi^{-2}}{1 - \phi^{-1}} = 1.$$

Performing the substitution "100" for "011" as above, we obtain the following expansions with value 1:

$$\mathbf{b}' = 0.100\overline{1}, \quad \mathbf{b}'' = 0.10100\overline{1}, \quad \ldots, \quad \mathbf{b}^{(n)} = 0.1010\ldots100\overline{1}, \quad \ldots,$$

where in $\mathbf{b}^{(n)}$ there are n pairs "10" before the trailing "$0\overline{1}$" (thus, $\mathbf{b}^{(0)}$ is $\mathbf{b}$). Also,

$$\mathbf{c} = 0.\overline{10} = 0.101010101010\ldots = \sum_{j=0}^{\infty} \phi^{-(2j+1)} = \frac{\phi^{-1}}{1 - \phi^{-2}} = 1.$$

We prove that there are no other base-ϕ expansions with value 1. Indeed, let $\mathbf{d} = 0.d_1d_2\ldots$ be any base-ϕ expansion with value 1. If $\mathbf{d}$ differs from $\mathbf{c}$, let the first difference occur at position k. We prove that $\mathbf{d}$ is $\mathbf{a}^{(l-1)}$ if $k = 2l$ is even, $\mathbf{d}$ is $\mathbf{b}^{(l-1)}$ if $k = 2l - 1$ is odd ($l = 1, 2, \ldots$).

If $k = 1$, then $d_1 \neq c_1 = 1$, so $d_1 = 0$; hence, $1 = (\mathbf{d}) \leq 0.0\overline{1} = 1$. Since equality holds, $\mathbf{d}$ must be the expansion $\mathbf{b}$. If $k = 2l - 1$ for some $l > 1$, then $c_k = 1$ and $d_k = 0$, so

$$0.d_kd_{k+1}\ldots = \phi^{k-1}[(\mathbf{d}) - (\mathbf{d}_{<k})] = \phi^{k-1}[(\mathbf{c}) - (\mathbf{c}_{<k})] = 0.\overline{10} = (\mathbf{c}) = 1.$$

As shown above, $0.d_kd_{k+1}\ldots$ is the expansion $\mathbf{b}^{(0)}$, so $\mathbf{d}$ is the expansion $\mathbf{b}^{(l-1)}$. If $k = 2l$ for some $l \geq 1$, then $c_k = 0$ and $d_k = 1$, so

$$0.d_kd_{k+1}\ldots = \phi^{k-1}[(\mathbf{d}) - (\mathbf{d}_{<k})] = \phi^{k-1}[(\mathbf{c}) - (\mathbf{c}_{<k})] = 0.\overline{01} = \phi^{-1} = 0.1.$$

Since $d_k = 1$, we must have $0 = d_{k+1} = d_{k+2} = \ldots$, so $\mathbf{d}$ is the expansion $\mathbf{a}^{(l-1)}$.

Also solved by Ram Dubey, Dmitry Fleischman, Enrique Treviño, and the proposer.

Sign fluctuations of weighted partial sums of a sequence **June 2017**

2025. *Proposed by Valerian Nita, Sterling Heights, MI.*

Let n be a positive integer and let $x_1, x_2, \ldots, x_n$ and $a_1, a_2, \ldots, a_n$ be real numbers such that $\sum_{k=1}^{n} x_k = 0$ and $0 < a_1 < a_2 < \ldots < a_n$. Define $s_1, s_2, \ldots, s_n$ by $s_k = \sum_{j=1}^{k} a_jx_j$ for $k = 1, 2, 3, \ldots, n$. If there is at least one nonzero number among $x_1, x_2, \ldots, x_n$, prove that there is at least one positive and at least one negative number among $s_1, s_2, \ldots, s_n$.

Solution by Michel Bataille, Rouen, France.

There is nothing to prove if $n = 1$, because in this case $x_1 = \sum_{k=1}^{1} x_k = 0$ by hypothesis, so it is not possible for the number x_1 to be nonzero. Henceforth assume $n \geq 2$. Without loss of generality we may assume x_1 is nonzero, because the truth of the statement for $x_m, x_{m+1}, \ldots, x_n$ and $a_m, \ldots, a_n$ implies its truth for $0, \ldots, 0, x_m, \ldots, x_n$ and $a_1, a_2, \ldots, a_n$, and moreover at least one of $x_1, x_2, \ldots, x_n$ is nonzero by assumption. By changing the sign of all numbers $x_1, x_2, \ldots, x_n$ if necessary, we may further assume $x_1 > 0$. Thus, $s_1 = a_1x_1$ is positive, so it remains only to show that one of the numbers $s_2, \ldots, s_n$ is negative. For $k = 1, 2, \ldots, n$, let $X_k = \sum_{j=1}^{k} x_j$. We have, for $2 \leq k \leq n$:

$$\begin{aligned} s_k &= a_1X_1 + a_2(X_2 - X_1) + a_3(X_3 - X_2) + \cdots + a_k(X_k - X_{k-1}) \\ &= a_kX_k + \sum_{j=1}^{k-1}(a_j - a_{j+1})X_j. \end{aligned}$$

Note that the coefficients $a_j - a_{j+1}$ above are all negative by hypothesis. We have $X_1 = x_1 > 0$ and $X_n = \sum_{k=1}^{n} x_k = 0$, so there exists $m \in \{2, \ldots, n\}$ such that $X_1, \ldots, X_{m-1}$ are positive and X_m is nonpositive. Thus, we have

$$s_m = a_mX_m + \sum_{j=1}^{m-1}(a_j - a_{j+1})X_j < 0$$

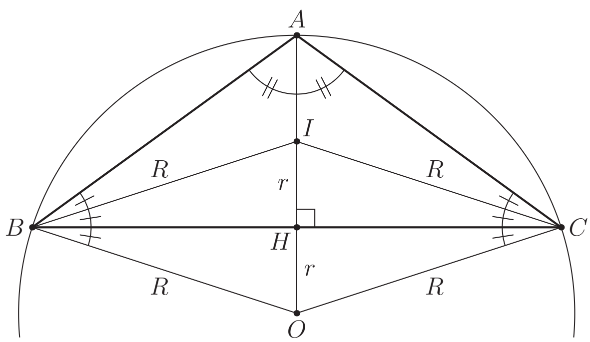
since the term $a_m X_m$ is nonpositive and the terms $(a_j - a_{j+1})X_j$ are negative (there is at least one of the latter, namely $(a_1 - a_2)X_1 < 0$), concluding the proof.

Also solved by Paul Budney, Robert Calcaterra, Richard Daquila, James Duemmel, Dmitry Fleischman, Eugene A. Herman, Dain Kim (South Korea), Miguel Lerma, ONU-SOLVE Group, Edward White, and the proposer.

Answers

Solutions to the Quickies from page 230.

A1081.



Let O be the circumcenter of $\triangle ABC$. Since $OA = OB = OC = R$ and $\triangle ABC$ is isosceles with $AB = AC$, the segments $\overline{AO}$ and $\overline{BC}$ are perpendicular and intersect at the midpoint H of $\overline{BC}$; moreover, $\overline{AH}$ bisects $\angle BAC$, so I lies on $\overline{OA}$ and $IH = r$. It follows that $\angle OBA = \angle OAB = \frac{1}{2}\angle BAC = 54^\circ$, hence $\angle OBC = \angle OBA - \angle CBA = 54^\circ - 36^\circ = 18^\circ$. Since $\overline{BI}$ bisects $\angle CBA$ we also have $\angle CBI = \angle IBA = \frac{1}{2}\angle CBA = 18^\circ = \angle OBC$, and similarly $\angle BCI = \angle OCB = 18^\circ$, so $OBIC$ must be a rhombus, with sides of length $OB = R$ and with $OH = IH = r$. It follows that $AI = OA - OI = R - 2r$, so $AI + BI + CI = (R - 2r) + R + R = 3R - 2r$.

Next, we have $OI^2 = R^2 - 2Rr$ by Euler's theorem. Since $OI = 2r$:

$$\begin{aligned} AI^2 + BI^2 + CI^2 &= (R-2r)^2 + R^2 + R^2 = 3R^2 - 4Rr + (2r)^2 \\ &= 3R^2 - 4Rr + (R^2 - 2Rr) = 4R^2 - 6Rr. \end{aligned}$$

A1082. The left-hand side of the stated inequality is equal to

$$\begin{aligned} L &= \left(\int_0^1 y^5\,dy\right)\left(\int_0^1 f(x)\,dx\right) + \left(\int_0^1 y\,dy\right)\left(\int_0^1 (f(x))^5\,dx\right) \\ &= \int_0^1 \int_0^1 \left(y^5 f(x) + y(f(x))^5\right)dx\,dy. \end{aligned}$$

Now, for $a, b \geq 0$, the inequality $a^5b + ab^5 \geq a^2b^4 + a^4b^2$ holds (since $a^5b + ab^5 - a^2b^4 - a^4b^2 = ab(a-b)^2(a^2 + ab + b^2) \geq 0$). Since f is nonnegative, it follows by integration that $L \geq R$, where

$$\begin{aligned} R &= \int_0^1 \int_0^1 \left(y^2(f(x))^4 + y^4(f(x))^2\right)dx\,dy \\ &= \left(\int_0^1 y^2\,dy\right)\left(\int_0^1 (f(x))^4\,dx\right) + \left(\int_0^1 y^4\,dy\right)\left(\int_0^1 (f(x))^2\,dx\right) \end{aligned}$$

is equal to the right-hand side of the stated inequality, completing the proof.

REVIEWS

PAUL J. CAMPBELL, *Editor*
Beloit College

Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles, books, and other materials are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of mathematics literature. Readers are invited to suggest items for review to the editors.

Chang, Kenneth, Influential mathematician wins honor akin to Nobel: Robert P. Langlands is awarded the Abel Prize, a top math honor, *New York Times* (20 March 2018), https://www.nytimes.com/2018/03/20/science/robert-langlands-abel-prize-mathematics.html.

Robert Langlands, mathematical visionary, wins the Abel Prize, https://www.quantamagazine.org/robert-langlands-mathematical-visionary-wins-the-abel-prize-20180320/.

Weisstein, Eric W., Langlands Program. From MathWorld—A Wolfram Web Resource. http://mathworld.wolfram.com/LanglandsProgram.html.

The Abel Prize, awarded annually by the Norwegian Academy of Science and Letters, has gone to Robert Langlands, originator of the conjectures known as the Langlands program. As Weisstein puts it, that program is a "grand unified theory of mathematics"; like the grand unified theory in physics, it is unproven. The Langlands program is a major ongoing project in mathematical research that suggests connections between reciprocity laws in number theory and periodic functions in analysis. Langlands conjectured that transformations underlying reciprocity laws can be represented by matrices; proof of a special case led to the proof of Fermat's Last Theorem. Langlands's conjectures have been proved for all local fields (e.g., $\mathbb{R}$ and p-adic fields) and all function fields (e.g., finite extensions of $\mathbb{Q}(z)$), but not for all number fields (algebraic extensions of $\mathbb{Q}$) [I hope by when you read this that the *Times* has corrected the misspelling of Abel's middle name.]

Maor, Eli, *Music by the Numbers: From Pythagoras to Schoenberg*, Princeton University Press, 2018; xvii + 159 pp, $24.95. ISBN 978-0-691-17690-1.

"I attempted to survey the musical-mathematical affinity from a historical perspective. ..." Maor takes us from the Pythagorean string theory to Schoenberg's 12-tone music, explaining musical concepts and connections to culture and mathematics. Along the way, Maor notes that the Slinky toy is useful for illustrating waves, but the mechanical metronome (two centuries old) and tuning fork (three centuries) have yielded to electronics. No mathematical background beyond high school algebra and trigonometry is required; ability to read musical notation is a plus but not essential.

Hossenfelder, Sabine, *Lost in Math: How Beauty Leads Physics Astray*, Basic Books, 2018; vii + 291 pp, $30. ISBN 978-0-465-09425-7.

Author Hossenfelder argues that there has not been a fundamental breakthrough in the foundations of physics for decades, and the reason is that physicists prefer beautiful and elegant theories (hmm. ... isn't that what mathematicians prefer?). "What failed physicists wasn't their math; it was their choice of math. They believed that Mother Nature was elegant, simple, and kind about providing clues. ... Whatever laws of nature govern our universe, they're not what physicists thought they were. ... The reliance of theoretical physicists on criteria of beauty and the resulting lack of progress represent a failure of science to self-correct."

Math. Mag. **91** (2018) 239–240. doi:10.1080/0025570X.2018.1460144

Su, Francis Edward, Mathematics for human flourishing, *American Mathematical Monthly* 124 (6) (June-July 2017) 483–493. https://www.dropbox.com/s/5022pc2niqpxy97/Su_Mathematics%20for%20Human%20Flourishing.pdf?dl=0. Video at https://www.youtube.com/watch?v=xEtDvc1SWm8.

Hartnett, Kevin, To live your best life, do mathematics, https://www.quantamagazine.org/math-and-the-best-life-an-interview-with-francis-su-20170202/.

Bleicher, Ariel, Why math is the best way to make sense of the world, https://www.quantamagazine.org/why-math-is-the-best-way-to-make-sense-of-the-world-20170911/.

Campbell, Paul J., Mathematics: Unreasonably ineffective? *The UMAP Journal of Undergraduate Mathematics and Its Applications* 39 (1) (2018) 1–4.

Much of the public interest in and support for mathematics derives from its utility in applications in fields with high-paying jobs. Apart from the potential for financial gain, what are the other benefits of studying and learning mathematics? "It trains the mind, it teaches how to think" is the standard answer, a conviction that has not been adequately documented (I apologize for citing in this connection a work of mine, but this is a topic that has been on my mind). Rebecca Goldin (George Mason Univ.) in her interview by Ariel Bleicher suggests that "quantitative literacy" allows for more public engagement and "getting past biases and beliefs and prejudices." But numeracy is not the heart of mathematics, nor statistics its soul. Francis Su (Harvey Mudd College), in his retiring address as president of the MAA, takes a grander view of what mathematics can do for a person. He asks, Why do mathematics? And his answer is: "*Mathematics helps people flourish*. … the practice of mathematics cultivates virtues [better thinking, perseverance, joy, hopefulness] that help people flourish. … It builds skills that allow people to do things they might otherwise not have been able to do or experience." He sees human flourishing in the Aristotelian terms of "activity in accord with virtue." "Why should anyone care about mathematics if it doesn't connect deeply to some human desire: to play, seek truth, pursue beauty, truth, fight for justice," or realize love. To pursue those desires is to flourish. The trouble, Su notes with examples, is that "we aren't helping all our students flourish"—and we should and must do better. This is an inspiring essay for the ages.

Mérő, László, *The Logic of Miracles: Making Sense of Rare, Really Rare, and Impossibly Rare Events*, Yale University Press, 2018; xii + 276 pp, $27.50. ISBN 978-0-300-23848-8.

This book offers a new take on rare events, like the "black swans" popularized by N.N. Taleb. Unlike Taleb's black swans, Mérő's miracles don't necessarily have to have huge impact. Mérő renames Taleb's Mediocristan "Mildovia" and his Extremistan—the world of unusual events—"Wildovia." Mérő suggests that we organize our lives as if we live in Mildovia (a world of Gaussian-distributed phenomena) rather than in Wildovia (where tail-heavy Cauchy distributions rule). Mérő "applies" Gödel's incompleteness theorem to the realm of totalitarian ideas and to esthetics; mentions fixed-point theorems in connection with economic equilibrium; and brings in chaos, scale-invariance, and fractals. He identifies factors that can bring on Wildovian phenomena: the Matthew effect, an increase in complexity, accumulation, and extreme competition. "[W]e must be. … familiar with the laws of both worlds so that we may calmly live our familiar Mildovian lives while being prepared to react to a Wildovian event at any moment." Mérő seems conflicted between the two worlds, and his prescriptions for preparedness appear ineffectual.

Darling, David, and Agnijo Banerjeeby, *Weird Math: A Teenage Genius and His Teacher Reveal the Strange Connections Between Math and Everyday Life*, Basic Books, 2018; xiii + 290 pp, $16.99. ISBN 978-1-54164479-3.

What book would you give a student that would offer a view of mathematics beyond the arithmetic, algebra, and geometry of the standard curriculum? Maybe this one. Written by a science writer and a 15-year-old student, this equationless survey of mathematics is a very readable account of ideas, challenges, and achievements.