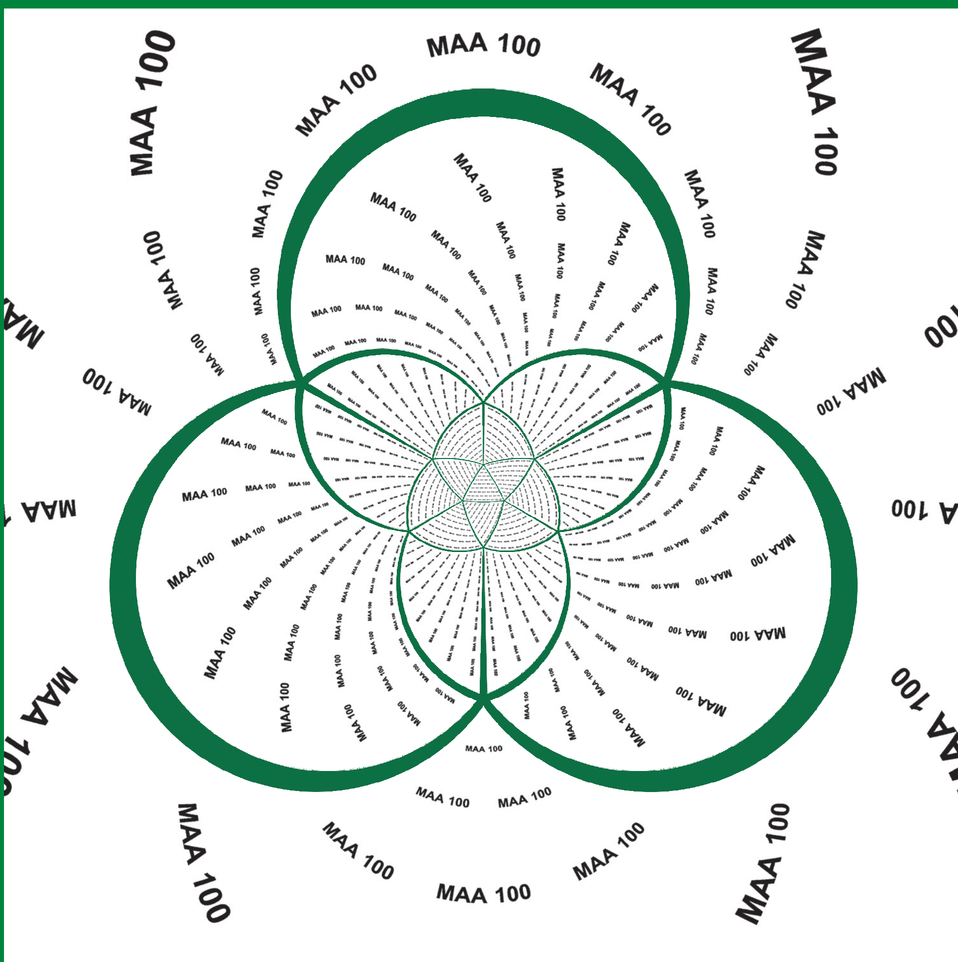
Vol. 88, No. 3, June 2015

MAA

MATHEMATICS MAGAZINE



- An interview with Lynn Steen, the Idea Man
- How lucky is too lucky? Winning too frequently in the lottery
- What does adding selectively have to do with Cantor sets?
- MAA Anniversary Crossword

EDITORIAL POLICY

Mathematics Magazine aims to provide lively and appealing mathematical exposition. The *Magazine* is not a research journal, so the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for the Magazine. Articles should include examples, applications, historical background, and illustrations, where appropriate. They should be attractive and accessible to undergraduates and would, ideally, be helpful in supplementing undergraduate courses or in stimulating student investigations. Manuscripts on history are especially welcome, as are those showing relationships among various branches of mathematics and between mathematics and other disciplines.

Submissions of articles are required via the *Mathematics Magazine*'s Editorial Manager System. The name(s) of the author(s) should not appear in the file. Initial submissions in pdf or LaTeX form can be sent to the editor at www.editorialmanager.com/mathmag/.

The Editorial Manager System will cue the author for all required information concerning the paper. Questions concerning submission of papers can be addressed to the editor at mathmag@maa.org. Authors who use LaTeX are urged to use the Magazine article template. However, a LaTeX file that uses a generic article class with no custom formatting is acceptable. The template and the Guidelines for Authors can be downloaded from www.maa.org/pubs/mathmag.

MATHEMATICS MAGAZINE (ISSN 0025-570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, NW, Washington, DC 20036 and Lancaster, PA, in the months of February, April, June, October, and December.

Microfilmed issues may be obtained from
University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.

Address advertising correspondence to

MAA Advertising
1529 Eighteenth St. NW
Washington, DC 20036
Phone: (202) 319-8461
E-mail: advertising@maa.org

Further advertising information can be found online at www.maa.org.

Change of address, missing issue inquiries, and other subscription correspondence can be sent to:

The MAA Customer Service Center
P.O. Box 91112
Washington, DC 20090-1112
(800) 331-1622
(301) 617-7800
maaservice@maa.org

Periodicals postage paid at Washington, D.C. and additional mailing offices.

Postmaster: Send address changes to Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, NW, Washington, DC 20036-1385.

Printed in the United States of America.

COVER IMAGE

Stereographic MAA © 2015 David A. Reimann (*Albion College*). Used by permission.

As with the MAA logo, this artwork features the icosahedron. Each face of the icosahedron was patterned with nine lines containing the text "MAA 100''. The icosahedron was first projected onto the circumscribing sphere and then stereographically projected onto the plane. The resulting transformation maps line segments into circular arcs.

Vol. 88, No. 3, June 2015

MATHEMATICS MAGAZINE

LETTER FROM THE EDITOR

At MathFest 2015 in Washington D.C., the MAA will celebrate its 100th anniversary. The lead article in this issue is an interview with Lynn Steen. As you may know, Lynn Steen was President of the MAA, was editor of THIS MAGAZINE, and was a proponent of quantitative literacy. But did you know that he did research with undergraduate students ... in the 1960s? You'll hear more of his ideas in the interview with Steen conducted by Deanna Haunsperger, Stephen Kennedy, Matthew Richey, and Paul Zorn that was transcribed and edited by Haunsperger and Kennedy. Steen passed away on June 21, 2015. This interview serves as a tribute to him and to his ideas.

Sandwiched between two proofs without words (one by Hasan Unal, and another by Victor Oxman and Moshe Stupal) is an article on the possible sums of series. Rather than consider partial sums in which the first n terms in a sequence are added, R. John Ferdinands considers adding any finite or infinite number of terms in a sequence. What such selective sums are possible? At times, the answer relates to Cantor sets.

Nachum Dershowitz and Christian Rinderknecht simplify some of the steps needed for calculating the average height of Catalan trees—a common data structure from computer science—by using elementary combinatorics that involve lattice-path counting. An online supplement to the article appears on the MAA website.

Skip Garibaldi, Richard Arratia, Lawrence Mower, and Philip B. Stark discuss the mathematics behind some recent investigative reporting which determined that some lottery winners were not as much lucky as they were dishonest. Their analysis uses the BKR inequality, named for van den Berg–Kesten–Reimer, that provides a bound on the probability of winning dependent bets even when which bets were made is not known.

You may recall Cavalieri's principle from calculus class, where it is usually introduced to show that two regions have the same volume. Zsolt Lengvarszky drops down a dimension and uses Cavalieri's principle to prove Pythagoras' theorem.

Daniel López-Aguayo was inspired by Problem 1942 that appeared in the April 2014 issue of THIS MAGAZINE. He asks a question whether or not there are an infinite number of certain binomial sums are nonintegral. Problem 1942 provided one such nonintegral sum, and López-Aguayo includes three more.

Mathematicians love games and enjoy using mathematics to analyze them. In their article, Darren Glass and Todd Neller determine optimal distributions of defensive armies in a linear version of the board game RISK.

The issue concludes with an anniversary crossword (Editors Past and President, Part 1) by Tracy Bennett, the Problems section, and the Reviews section. Once again, there are anecdotes from past editors of the MAGAZINE to continue the anniversary celebration. David Reimann provides MAA-inspired cover art—a stereographic projection of an icosahedron.

Hope to see you in Washington D.C.

Michael A. Jones, Editor

ARTICLES

The Idea Man: An Interview with Lynn Steen

DEANNA HAUNSPERGER
Carleton College
dhaunspe@carleton.edu

STEPHEN KENNEDY
Carleton College
skennedy@carleton.edu

At St. Olaf they have a word for it: you have been "Steened" if you got drafted by Lynn Steen to help him implement his latest great idea. Everyone in the math department there has stories of having been Steened. Steen, in a career spanning half a century at St. Olaf College, has been a never-ending fountain of innovative ideas. He led an undergraduate research summer program in the 1960s when conventional wisdom declared that impossible. In the 1980s he was a driving force in the Calculus Reform movement. More recently he helped create the nation-wide Quantitative Literacy effort. He was one of the architects of the phenomenal success of the St. Olaf program—Ted Vessey, longtime chair of that department, says the most interesting part of his job was every day hearing three new ideas from Lynn: one would be brilliant and should be implemented immediately, one would be brilliant but completely impractical, and the third would be just crazy. The challenge was deciding which was which.

Steen has been co-editor of THIS MAGAZINE, MAA President, Chair of the Committee on the Undergraduate Program in Mathematics, Fellow of the American Mathematical Society, Executive Director of the Mathematical Sciences Education Board, Chair of the Council of Scientific Society Presidents and occupied a leadership position in just about every organization concerned with US undergraduate mathematics education. He has put his inquisitive and creative mind to work attacking the problems of math education and has had a profound effect on teaching and learning in this country. We sat down with Lynn a few years ago at St. Olaf College with his colleagues Matthew Richey and Paul Zorn to ask him about his life and work.

> As this article went to press we learned that Lynn died on June 21, 2015. This is not only a great loss to our community, but also a deeply personal one for us as we have lost a valued friend and mentor. We hope this interview can stand as a fitting memorial to this deeply kind, wise man and his incredible contributions to mathematics. –Deanna Haunsperger and Stephen Kennedy

Q: *Can you tell us about your childhood, growing up, and any early interest in math?*

LS: Well, I grew up partly in the Midwest and partly in New York. I was born in Chicago, and my father died of tuberculosis when I was an infant. My mother was

Math. Mag. **88** (2015) 163–176. doi:10.4169/math.mag.88.3.163. © Mathematical Association of America
MSC: Primary 01S70

an opera singer, and she remarried a man whose name I now have (Steen). He was a musician who was in the Navy during World War II; when the war ended he got a job with Luther College in Iowa. We were there for a couple of years in the late 1940s and then moved to New York where he got a job with Wagner College on Staten Island.

My mother tells me that because both my parents were active with careers, they were constantly hiring college students to help babysit and do things around the house. Apparently, when I was in fifth or sixth grade, one of my jobs was drying dishes while one of these babysitters was washing dishes; she would keep me from being too bored or obnoxious by asking me math questions.

I went to Curtis High School on Staten Island. At Curtis, which was a pretty large New York City public school, there were about 2,000 students. There were three tracks of students there: a general track, a commercial track, and an academic track. The academic students took pretty much a standard curriculum with English, history, science, math, and foreign language. And then we took these Regents Exams and our way of knowing what we were good at was when we got a 94 or a 98 on the Regents Exams. Except for history, I was getting in the mid-90s for everything. I was more interested in math but I wasn't standing out; I wasn't the best math student in the school by any means.

Q: *Tell us about going to Luther.*

LS: Luther was a family connection. My stepfather had gone there, we lived there when I was young, and that was largely the reason that I went there. I was there from 1957 to 1961. It was a good liberal arts college but the East Coast had a slightly more advanced, more accelerated high school curriculum than the Midwest did at that point. I think it's fair to say that the kind of preparation I had in the New York City school system was better than the average preparation of many of the high schools in the Midwest.

I started with calculus, but that was relatively uncommon. I think there were three or four students who did that. I bet there were another 30 or 40 who could have started with calculus but it wasn't the typical pattern. At Luther I majored in both mathematics and physics and almost had a major in philosophy as well. That was part of the reason I wasn't sure what I was going to be doing. To be honest, I didn't really think about what I would do after college.

In mathematics we had a pretty standard course sequence: I took two years of calculus and advanced calculus, a year of probability and statistics, one semester of abstract algebra, and a course in number theory. I took a course in differential equations at Wagner College in summer school, and not much else. No complex analysis, no topology.

The thing that was interesting at Luther, which was also true at St. Olaf when I got here, is that both of these schools, and several other Midwestern liberal arts colleges, had faculty who had immigrated because of the turmoil in Europe. These European-educated faculty from Estonia, Latvia, northern Germany, and other places tended to be driving a different set of standards than the American-educated faculty. I don't think they had anybody with an American Ph.D. in the Luther math department most of the time I was there.

Q: *Were you prepared for graduate study?*

LS: The big thing that I realized shortly after I entered MIT was there was nearly a century-long gap between the curriculum I had as an undergraduate and what they were doing in graduate school. The most recent material that I had studied, for the most part, was from the late nineteenth century and the MIT faculty were teaching

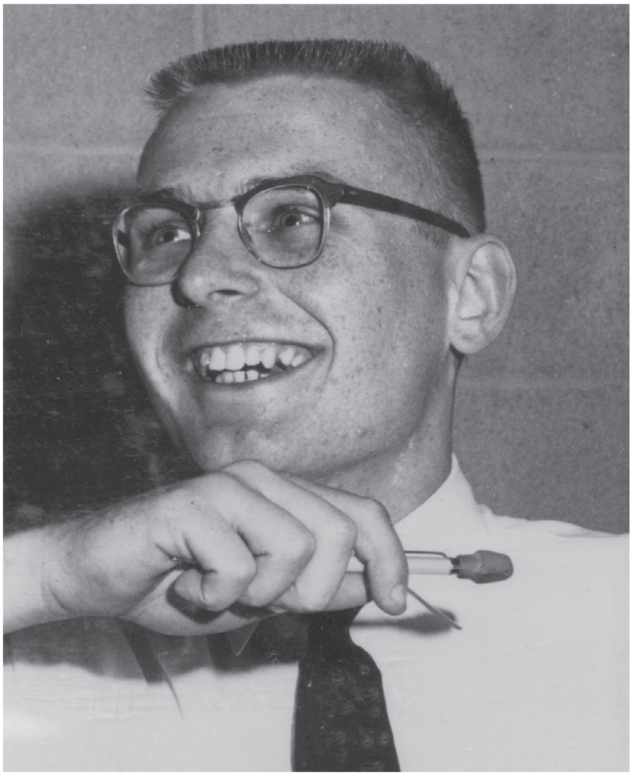


Figure 1 Lynn Steen at Luther College in 1961

material starting in the 1950s and beyond. The professors at MIT skipped over all that earlier material so I missed all the mathematics from the years in between. For a long time I didn't even recognize the subject; I had no clue that this was the same field that I'd thought I was in. But, somehow, it worked.

George Thomas was the adviser for first-year graduate students when I was there. When he looked at my undergraduate transcript he said, "Well, everybody who enrolls in the graduate program has got to take the graduate real analysis course and the graduate algebra course in their first year. But, since you haven't had it, you also have to take the undergraduate real analysis course. So you'll just have to take both at the same time." You do learn that way. It was a sink-or-swim business. I wasn't prepared mathematically, but I was prepared in the discipline of studying.

You know preparation—or lack thereof—is an interesting issue. I was rooming with a fellow first-year math graduate student whom I didn't know at all. He came from Tulane. I always had the impression he was a whole lot stronger than I was in his preparation and was doing better. But then, he never returned in the second year, and I did. It's a matter of persevering. I spent a lot of time in the library trying to learn things the instructors assumed. I was discouraged sometimes but I didn't have anything else to do. I just kept at it.

Q: *You had an NSF fellowship. How did that come about?*

LS: There was a dean at Luther telling seniors you have to apply for this and that. As a result, while I was getting money from NSF for my graduate fellowship, I also got a Danforth fellowship. That was a wonderful program that was going on for about twenty-five or thirty years from the Danforth Foundation in St. Louis. It was specifically focused on giving fellowships to students who were planning to become teachers at liberal arts colleges. When I was in Cambridge, faculty from Harvard, MIT, and Brandeis advised the fellowship students. Every six weeks or so we had an organized dinner for all the Danforth fellows in all different fields who were currently studying in the Cambridge-Boston area. So, I got to mix in with these people who would be talking about issues of higher education and teaching and learning. These events helped keep you interested in the career as a teacher. Also, there was a requirement that every summer, the last week in August, you had to spend a week at bootcamp on the eastern shore of Lake Michigan with all the other Danforth fellows. This was all focused entirely on preparing students to teach in liberal arts colleges.

Q: *So does that mean you knew you wanted to do that when you were an undergraduate?*

LS: Pretty much. To apply for the fellowship I had to write an essay convincing them that I really wanted to do this. Writing helps to sort out your thinking.

Q: *Back to graduate school: who was your thesis adviser and what did you do?*

LS: Kenneth Hoffman. And I was working in function algebras on something called Gleason parts (which are named after Andy Gleason). When I was at MIT I wandered around a lot trying to figure out what area I wanted to work in and who I wanted to study under. I think I wound up with Ken Hoffman largely because he was modeling the kind of teaching that I liked; a lot of my teachers weren't all that good. Several of the visiting foreign scholars I could hardly understand, but Ken was always very clear about what he was doing with his classes and he just had this kind of style.

When I was working on my thesis, I had these every-week or every-other-week appointments to check in to see what I was doing. Most of the time when I showed up Ken had an excuse that he had to go somewhere because he had to go see the Provost or he had to go see the Assistant Dean, so would I please walk with him. I don't know if you've been to MIT, but it's like the Pentagon; I mean it's got corridors all over the place—you could walk for three miles and never follow the same path twice around that place. And so I'd walk with him wherever he was going, and we'd take care of business while walking.

Finally it dawned on me: these appointments were just façades to get me away from the blackboard and his office. He wanted to make sure that I could explain a problem totally verbally and not have to rely on the crutch of drawing funny sketches on the board. It worked. It forced me to internalize what I was learning and to ask more and more questions. We've all had the experience of explaining a problem to somebody and solving it while you're explaining it; our walks were an even more extreme version.

Q: *How'd you get from MIT to St. Olaf?*

LS: Well, you know, that's interesting, especially looking at the way academia is today with all the difficulties people have in searching for jobs. The Danforth Foundation circulated a list of all of the Danforth Fellows to all the liberal arts colleges in the country. So, we didn't have to apply for jobs at all. I received fifty letters within the space of about three weeks from colleges who said "we've got a job." I actually wound up talking with Carleton and St. Olaf, partly because when I'd been at Luther, I'd heard about them and knew by their reputations that they were good places. Also I was a little bit nervous, having been in Cambridge where there's so much activity going on with Harvard and MIT, about going off to some place like a Luther which was one isolated little college in the middle of the cornfields. I figured Northfield, which I'd never been to, would be a good compromise: two colleges, near the Twin Cities. It turns out Carleton didn't have a position and St. Olaf did.

One of the other reasons I was attracted to this job was precisely because most of the people in the department were going to retire in three years so it was a chance to build a department rather than being one new person in a department where you couldn't get much done. In 1965 when Arthur Seebach and I were hired, we joined a department of four; in the space of four or five years, the older three people retired.

A few years earlier the Ford Foundation had done a study of small liberal arts colleges and concluded that if they didn't have a good endowment they'd have trouble surviving at the small size. So they gave grants to colleges that would be willing to grow. St. Olaf made a commitment to grow from 1500 to 3000. Every year they added a hundred students, the tuition from a hundred students, and the faculty that went with that. This went on from 1965 to 1980 and then St. Olaf more or less leveled off. We have been at 3000 ± 100, ever since. This strategy enabled the St. Olaf math department to quickly double in size.

It wasn't as if I knew what I wanted to help build when I got here. I think it was the reverse. Having been at Luther and having talked to other people at these Danforth conferences about liberal arts colleges, one of the difficulties that many of the smaller schools have is that if they only add one person every seven or eight years in a department that's already set in its ways, it's very difficult for that one person to do much of anything. So, the opportunity to come into a place where you knew there would be a big turnover was more attractive.

Q: *But you had some vision of what a private liberal arts college math program could, or should, move towards.*

LS: My own experience revealed that I had at least a fifty-year gap in preparation for MIT. I also knew that it was common for similar liberal arts colleges to regularly send students in chemistry or physics to really good graduate programs such as Harvard and Berkeley. But, they weren't doing it in mathematics. My sense was we ought to be able to make an undergraduate program that would actually work for the top graduate programs. How to do it was a different matter.

The first issue was getting involved in undergraduate research. Now, to be honest, that wasn't my original idea because as a graduate student I was in no way connected with the idea of undergraduate research in mathematics: at that time, undergraduate math majors were just taking courses, even at MIT. But, when I came to St. Olaf the dean who hired me, Al Finholt, was a chemist. At St. Olaf, chemistry was then the strongest program (other than music) and everybody in the college knew it. He was dean because he had the vision to make the rest of the college as strong as chemistry. And he made it clear to me that what he wanted new faculty to do was to develop undergraduate research programs.

This was quite a bit before it became a common thing in mathematics. The first week I was at St. Olaf we submitted an NSF-URP grant to support undergraduate research. That first try, hastily written, didn't succeed; but the next year Arthur Seebach and I concocted the idea of using point-set topology as the focus. We figured this approach could overcome most of the objections to doing research in mathematics: this field was close enough to the surface and it was full of little problems that weren't solved simply because people weren't interested in figuring out how to solve them. It was a good place for students to work. We got a grant from that idea that worked for about four or five years. The effect was seen in the students who were freshmen when I came here: we had three students in that class who had all been in the summer undergraduate research program who won National Science Foundation Fellowships.

Figure 2 Arthur Seebach at St. Olaf in 1986

One went to Columbia, one went to Harvard, and one went to Michigan. That served as an existence proof.

Undergraduate research is still realistic for only certain subfields of mathematics. What's going on now is that the faculty have become convinced, the NSF is getting convinced, and people are beginning to actively think about areas where they can work with students. Previously mathematicians would say, "Well, you can't do it in my field," which is true in many cases. It wasn't as if they were dissembling; in most areas, undergraduates could not do mathematical research. The reason we did point-set topology was that neither one of us had studied or knew much about it: Arthur did his thesis work in category theory, mine was in function algebras. We just picked something in the middle that neither one of us knew very well so that we could be actively and honestly researching along with the students.

Our book, *Counterexamples in Topology,* came out of that. That was sort of a take off on Gelbaum & Olmsted's *Counterexamples in Analysis* that came out about seven or eight years earlier and which we'd been using as a reference book in analysis courses. It seemed like a really handy kind of thing.

Q: *Was it an intentional decision of yours not to pursue a research track?*

LS: I really published only two short research papers. Basically, I was too busy early on to delve deeply into it because I wound up focusing my summers on undergraduate research. And then, on my first leave, I went to the Mittag-Leffler Institute in Sweden, thinking that would offer a chance to re-engage with my earlier work and get out of point-set topology. But I got distracted by something totally different when I discovered that the Mittag-Leffler Institute had this absolutely marvelous library of older mathematics books. Because Mittag-Leffler was a book collector, he bought up all the European math books. The Institute is on his estate and there's this round tower enclosing a multistoried library where you could walk up and take off the shelves original editions of just about anything you wanted. They weren't behind glass, they weren't locked up; they were just there for the visitors.

I wound up writing a couple of history papers in the *Monthly* together with a paper in *Scientific American;* one of them I finished when I was there, the others I finished after returning home. So that got me started writing expository pieces, and since at that time nobody else was doing it in mathematics, I just stayed with that for ten years.

I should mention one other thing about mathematical exposition. When I was at MIT, given the fact that I didn't know what I needed to know for background and at least some of the instructors weren't providing much of any background, I was left at the mercy of the library. I discovered quickly that the best sources of all of this intermediate mathematics—that is from 1900 to 1970—were Russians. Enough of it was translated that I was able to study really good material. Russian mathematicians had a tradition of writing good expository mathematics. These Russian translations really helped me a lot as a graduate student and made me value this kind of exposition.

Q: *How did your relationship with the MAA begin?*

LS: Well, it began with Telegraphic Reviews. In the mid-1960s Kenneth May became the Reviews Editor for the *Monthly* and inherited an enormous backlog of books. So he invented the Telegraphic Reviews as a way of managing this job—single-handedly reviewing several hundred books each year. Ken was then at Toronto, but had earlier been at Carleton. When Ken decided to retire from this arduous position, he suggested to Harley Flanders, then the editor of the *Monthly*, that he seek an editorial team in Northfield since Ken knew that between Carleton and St. Olaf, Northfield had a large number of mathematicians representing a variety of specialties.

So Harley approached Arthur Seebach and me about organizing this effort. The invitation came just when I was getting ready to go on leave. So Arthur set up the whole thing when I was in Stockholm and then I took it over when he went on leave the following year. The Telegraphic Reviews office moved to St. Olaf in 1970 and continued for twenty years. As one result, I wound up attending meetings of the Committee on Publications and gradually became more involved in that part of MAA's work.

Q: *How did you get started doing* MATHEMATICS MAGAZINE*?*

LS: Ed Beckenbach of UCLA was the chair of the Committee on Publications, so he knew Arthur and me from those meetings. At the time he asked us to be co-editors of MATH MAGAZINE, we had never even seen it. The MAA had only taken over the MAGAZINE in the early 1960s and it wasn't very visible. By around 1974 or so, it had sort of fallen into disarray in that the editor wasn't responding and they were having problems. So, it needed to be restarted in some way; Ed Beckenbach asked us what we could do to make it more interesting and lively.

We had the idea of trying to make it more of a public magazine, more expository, but on the other hand we realized it was on shaky ground because it had its own narrow constituency of people who read it and wrote for it and if you upended that too much you might lose the whole thing. We wanted to make sure we kept the same writers there, so we only made cosmetic changes. The biggest issue, which took the longest discussion, was deciding if it would be okay to put illustrations instead of the table of contents on the cover.

Figure 3 MATH MAGAZINE cover illustration from the issue containing "Logic from A to G" by Paul Halmos

Because of our experience with the Telegraphic Reviews, and our MAA-supported secretarial help at St. Olaf, we added a typewritten 'late-breaking news' section to the MAGAZINE which allowed us to put in letters to the editor that were actually timely and relevant to the article they were commenting on rather than six months delayed. I was, at that point, also doing some freelance work for *Science News,* and I got a tip from Ron Graham about the solution to the four-color problem and was able to get an article on that into MATH MAGAZINE several months before it appeared in other places because we typewrote it and dropped it into the next issue.

Q: *I heard a story that you are the inventor of Math Horizons. That you actually made a mockup of something called Math World.*

LS: After I did that historical work at Mittag-Leffler and published several expository papers and received Ford awards from MAA for two of them, I wrote a grant proposal to NSF to see if I could get funding to do more expository writing. They didn't have any programs for it and they quite naturally turned me down because, as they said, "We publish original research; we don't support exposition." A year later I got a call from the Conference Board of Mathematical Sciences (CBMS), an umbrella organization of the math societies, at that point chaired by Saunders Mac Lane. They thought mathematics wasn't getting enough visibility in Washington because no journalists were covering it. So they got a grant from NSF to explore why this was. They called to ask me to run this project for them.

I found out many years later that the reason they called me was that one of the people working on this project had been a reviewer of the proposal I'd submitted to NSF, and he remembered my interest in expository writing. I've mentioned this many times to students and Project Kaleidoscope faculty as evidence of the importance of applying for grants: if you have a good idea, even if it is a long shot, it might turn out later to be of some value.

CBMS set up this project that went on for about a year to eighteen months; I worked half-time on it. I talked to the Editor of *Science News* then went to the International Congress and wrote some freelance pieces for them. I also worked on the volume

Mathematics Today, for which Saunders Mac Lane and Felix Browder co-wrote the preface.

One part of that work was to make a proposal for a magazine in mathematics that would be much more public-oriented than the *Monthly* or anything else. That's what *Mathematical World* was. We made a mock-up by collecting articles that were in existence and presented it to the MAA and asked if they wanted to own it and support it and run the thing. It was turned down by a one-vote margin by the Board of Governors—partly because it was a big financial risk but also because at that same time Springer Verlag was starting this little yellow pamphlet that evolved into *The Mathematical Intelligencer*.

When the *Intelligencer* started it was just a public relations vehicle for Springer books. When it grew into a real magazine, it had good exposition, but advanced. It really was for the insider, like the *Monthly*. Recognizing that the *Intelligencer* was no substitute for *Mathematical World,* Don Albers launched *Math Horizons*, which has thrived and reached an important student audience.

Figure 4 A cartoon drawing of Lynn Steen and Arthur Seebach prepared by Don Albers's friend David Logothetti to celebrate their time as Editors of MATH MAGAZINE.

Q: *Moving on, how did you become President of the MAA?*

LS: In 1980, near the end of our term as editor of MATHEMATICS MAGAZINE, Dave Roselle, who was MAA Secretary at that point, came up to me at a meeting and asked if I had any interest in doing administrative work for MAA. I said "Well, I'm done with the MAGAZINE so I'd be willing to do something different." A little while later they nominated me for Vice-President, which mostly involved attending meetings of the Executive Committee. I didn't know whether I'd win or not, but I assume it was a close election because they didn't announce the result for a long time. (This was in the era of paper ballots.) What I didn't realize then was that a year or so later somebody said that they expected that I would be nominated for President because they had this tradition, which is no longer true, that the Vice-President became President.

Q: *Did you have an agenda as President?*

LS: In the early 1970s enrollment figures in mathematics started being affected by computer science and there was this huge decline—undergraduate, graduate, and Ph.D., all declined from 1970 down into the mid-1980s. The NSF set up this big study that was going on while I was President dealing with a whole range of undergraduate science issues. One of the key concerns was what to do about the decline in enrollment in mathematics which pretty much paralleled the corresponding increase in computer science. That's when the whole calculus reform movement started up.

My main concern was dealing with the issue of renewing the undergraduate curriculum, which was related in a way to my transition to graduate school. Even though the undergraduate programs in general were much better then than when I was an undergraduate, they weren't keeping up with the exploding uses of mathematics in other fields. So the frontier of mathematics, applied mathematics especially, was growing enormously and people were using mathematics for all sorts of stuff. But you still had a core curriculum that had been defined by CUPM in the 1960s going on in the 1980s when students were not too interested in it. My agenda was to make people aware of the problem. I decided as a personal endeavor to take every opportunity to tell that story with the data and the issues, assuming that if you get the information out there, a lot of people with good ideas are going to start working on it.

Several years before I became president of MAA, Ronald Reagan ran on a platform to eliminate the U.S. Department of Education. Instead, the commission charged with doing this issued a report called *A Nation At Risk* which argued for major improvement in K-12 education. As a consequence, the National Council of Teachers of Mathematics (NCTM) started their effort to develop standards for school mathematics—way before the rest of the country was talking about such things.

Several years after I finished my Ph.D., Ken Hoffman had become chairman of the math department at MIT at the same time as Frank Press was chairman of the geoscience department. When *A Nation At Risk* appeared, Frank was President of the National Academy of Sciences. Frank and Ken cooked up this idea that the way for scientists to deal with *A Nation at Risk* was to set up a board of science and mathematics education at the National Academy. They wanted to have a small group plan it, and they arranged for me to be in this group. So connecting MAA to this undertaking also became part of my agenda when I was president.

Q: *I read somewhere, something that you wrote, that one of the problems was that professional mathematicians were not involved enough in setting those standards and they kind of reacted badly to having those standards set by, what they thought of as, school teachers.*

LS: That's an interesting, complicated issue. What you just said, I think, I would agree with.

Q: *You said it... I'm paraphrasing.*

LS: I know; I'm just glad I still agree with it. The people who worked on the original NCTM *Standards* were primarily faculty in university mathematics education. Tom Romberg at Wisconsin was the leader of it and Wisconsin and Michigan and Georgia had the biggest math education programs in the United States. Most of the people who did all the work were university people but they were in math education departments. NCTM had various oversight committees that had representatives, one from MAA (which was me), one from AMS (I don't remember who that was).

At some point between the penultimate and final draft, NCTM President John Dossey sent letters to a whole range of research mathematicians asking them to review the document (or perhaps certain chapters). From what I gather after the fact, most of them ignored this request. Many had never even heard of NCTM. So even though NCTM made an effort, there hadn't been enough of an effort to satisfy the mathematicians. Whether there was enough input to make the document good enough for its purposes is another issue. There obviously were mistakes in it, but that would still have been true even if they'd had more mathematicians.

Before publication, AMS, MAA, and several other professional societies were asked to join NCTM in supporting the "vision of school mathematics" described in the document. Harvey Keynes from the University of Minnesota was chair of the AMS review committee for that decision. I was at the meeting of the AMS Council when this was brought up for endorsement. Harvey stood up and said a few nice words; members of the Council all smiled and said, "this sounds like a good thing" and approved it—which is what most boards do with most things like this most of the time. It was just normal committee behavior. There was no reason anybody had to believe that it would mean much of anything because at that point nobody in the country was even willing to use the phrase "national standards."

At that time, neither the NSF nor the Department of Education would fund work on nation-wide standards. The NCTM members paid for the whole thing. They had a special extra fee that they imposed on themselves to pay for this. Later the government decided that in fact the idea of trying to create national standards might be worthwhile, and then they funded the science standards, the history standards, and all these other ones. But in the beginning NCTM was all by itself.

Q: *And they got smacked for it.*

LS: Well they certainly did by the mathematicians. If you read the critiques that came out about the government-funded history standards and English standards—that same reaction was evident. I mean, can you imagine the canon issues with regard to English and the question of what is the truth about American history? When Lynn Cheney was Director for the National Endowment for the Humanities those standards were just as controversial, if not more so.

Q: *Is there value in this exercise?*

LS: Oh, there's some value, but not as much as many people invest in it. The standards aren't going to teach anybody anything: you need teachers to do that. And you need people to teach the teachers, and you need good problems for students to work on. I wrote an editorial for the *Notices* once suggesting that mathematicians would be putting their energy to better use if, instead of criticizing the standards, they would spend time inventing good problems because that's what students really need to wrestle with. Most math textbooks are full of mostly dumb problems and because they're so dumb, students don't get interested in the subject.

Q: *What mathematicians have stood out to you during your career as being exceptionally interesting or helpful, progressive, any luminaries in mathematics?*

LS: In the early part of my career the one who is clearly a distinctive luminary and who I sought out for the purpose of having the experience was Norbert Wiener. He was still teaching at MIT when I was a graduate student there, so I took a course from him, principally because I wanted the experience of taking a course from Norbert Wiener. The course he was teaching was on Fourier Series; at that point his eyesight was so bad he could hardly write, so this course was delivered entirely orally. He would talk Fourier Series at you. At least at the beginning I knew enough to transcribe his talking but after a while I resorted to reading some of his work. He was definitely interesting.

In terms of strong influence obviously Ken Hoffman had a lot of influence on me because I worked with him closely. When I was working with the *Monthly* there were a lot of good, strong editors that I learned from, especially Paul Halmos and Alex Rosenberg. I only dealt with them indirectly at meetings but you could get a sense of standards and context, what you should be looking for in things, and how they make judgments. More recently, when I was involved with the Mathematical Sciences Education Board (MSEB), I worked a lot with Hyman Bass who is a remarkable person; he's got an enormous range of both mathematics and education background. Recently he's been doing a lot with math education but he got into that because of the work we did with MSEB. He was one of the few members of the National Academy of Sciences that they could actually get to be chair of MSEB because most of the mathematicians there had little or no interest in this stuff at all.

Q: *Do you want to talk about quantitative literacy and how that started and where that's going?*

LS: For quite a number of years I was on the College Board Committee on Mathematics which advises on all of their tests: they have special committees for each test but they have one overarching committee for each subject that looks over everything. At one point the science committee came to the math committee and said "We're beginning to believe that the quantitative content of the College Board's science exams [that is the SAT, the AP tests, and the CLEP exams] isn't quite up to snuff with the way science is evolving." Could we, as the mathematics committee, give them some advice as to what was important to put on the science exams?

I suggested to Bob Orrill, who was at that time the executive director of the College Board's Office of Academic Affairs, that this kind of question would be better addressed by asking a lot of people who were actually using mathematics what they think is really important. So, I spent a year developing a big project on what came to be called quantitative literacy (QL). Bob got a grant from the Pew Charitable Trusts to start a literacy project with quantitative literacy being the lead topic and asked me to head up that part. Then I roped Bernie Madison into it; we organized conferences and edited volumes and all this other stuff.

Bob Orrill approaches all these problems by figuring out ways to get campuses to bite the bullet and take on an issue. So that's what our goal was with the quantitative literacy. It was kind of like earlier with my MAA presidency. Then it had been getting people aware of the problem of declining enrollment. Here our problem was to get people aware of the fact that there was a huge issue with civic literacy that dealt with numbers that was not being addressed by the regular curriculum. None of the standards in the K-12 system dealt with it; neither did the typical college math curriculum. It just wasn't anywhere. Lots of projects came sweeping out of this, including MAA's special interest group on QL, the National Numeracy Network, and the on-line journal *Numeracy*.

Figure 5 A recent photo of Lynn Steen

In preparing this interview for publication, we got back in touch with Lynn and asked him what he sees as the most important issue facing the MAA today.

LS: The landscape of collegiate mathematics—the focus of MAA's work—has changed enormously in recent years. In many respects, mathematics has prospered: it is relied on like never before to serve a whole range of careers and is regularly cited as one of the best choices of undergraduate majors. On the other hand, financial support for higher education has weakened; debates about school mathematics have become politically rancorous; and large numbers of voters distrust inferences based on data and scientific reasoning. Moreover, decades of effort have failed to reduce either the debilitating socioeconomic gaps in performance or the frustrating need for remediation among entering college students.

Each of these problems affects collegiate mathematics and the members of MAA, but none can be addressed effectively *within* MAA. The Association's main activities—meetings, journals, books—speak primarily to its own members and do not normally engage outsiders whose assumptions, agendas, and activities differ from those of most MAA members. Yet the professional problems MAA members face derive significantly from decisions made by individuals who know little or nothing about the reality faced by mathematics faculty in today's colleges and universities. As MAA begins its second century, it would probably be beneficial to undertake a sustained dialogue with other sectors of society with whom we have common interests.

MAA is not without resources for such an endeavor. Some of its members, many former members, and large numbers of former undergraduate mathematics majors are leaders in other fields—universities, technology, finance, energy, climate, even politics. As colleges and universities have discovered the benefits of connecting students with alumni to help enhance career preparation, perhaps MAA should seek, similarly, to connect with individuals in other fields whose education and early careers were advanced by members of MAA. I think we'd be pleasantly surprised by the many areas of common interest, and even where our agendas differ, working together to reduce these differences will itself be helpful to the Association and its goals.

Summary. An interview with Lynn Arthur Steen, this article questions this consummate idea man about his childhood, editing MATHEMATICS MAGAZINE with Arthur Seebach, being a member of the St. Olaf Math Department through its period of phenomenal growth, serving as MAA President, being on the front lines of math education—writing and being involved in the conversations about the NCTM Standards and Quantitative Literacy, and also commenting on what he sees as the most important issue facing the MAA today.

DEANNA HAUNSPERGER (MR Author ID: 321814) is Professor of Mathematics at Carleton College. She was colleagues with Lynn Steen many years ago, has been co-editor of Math Horizons, has co-directed, since 1995, the Carleton Summer Mathematics Program for women, and has been co-chair of the MAA Centennial Planning Committee, all with her co-author. She has served the MAA in a variety of ways, including as Second Vice President, and currently as Chair of the Council on Outreach Programs.

STEPHEN KENNEDY (MR Author ID: 321925) is Professor of Mathematics at Carleton College. He has always regretted that H comes before K; see his co-author's biography. He is currently Senior Acquisitions Editor for MAA Books.

MAA 100 MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

From the Files of Past MAGAZINE Editors

Paul Zorn 1996–2000

When Paul Zorn was editor of the MAGAZINE, most submissions were on paper (TEX and LATEX were beginning to be used); articles would be sent as big stacks of paper, in the order to be printed, to Harry Waldman. The compositor would type them in, and send bluelines back to Paul, which were actually off-white paper, slightly larger than the final product so that notes could be made in the margins, printed in light blue (as opposed to black) ink with a funny smell. The MAGAZINE still refers to the final page proof (now electronic) of the entire issue as the bluelines, despite the lack of blue ink and the lack of a smell. The terminology is from when photographic proof from negatives would show all colors in blue (or another color). Zorn would read the bluelines, as would Mary Kay Peterson, and once she caught that the entire magazine had had the footer running with the year 1098 on the bottom of every page.

Proof Without Words: Sums of Triangular Numbers

HASAN UNAL
Yildiz Technical University
Istanbul, Turkey
hunal@yildiz.edu.tr

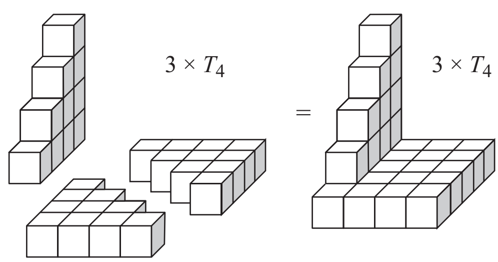
Theorem. Let $T_k = 1 + 2 + 3 + \cdots + k$, then

$$\sum_{k=1}^{n} T_k = \frac{n(n+1)(n+2)}{6}.$$

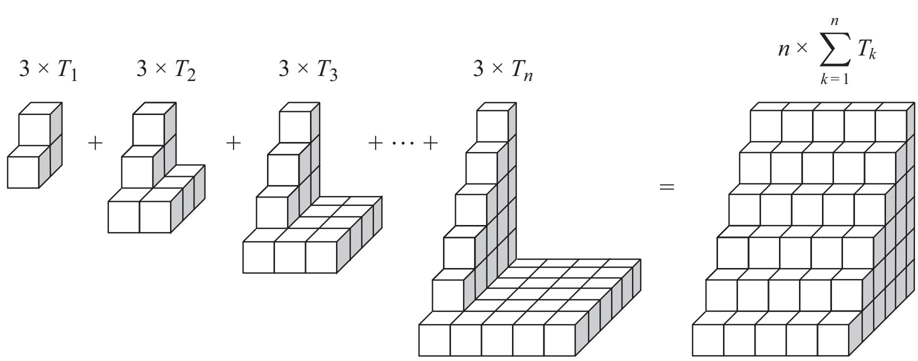
Proof. Notice that

$$3 \times \sum_{k=1}^{n} T_k = nT_{n+1} = \frac{n(n+1)(n+2)}{2},$$

because

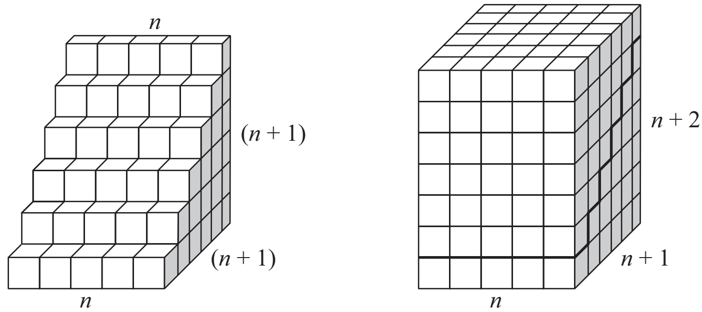


and



Math. Mag. **88** (2015) 177–178. doi:10.4169/math.mag.88.3.177. © Mathematical Association of America
MSC: Primary 05A19

Consequently,



$$\sum_{k=1}^{n} T_k = \frac{n(n+1)(n+2)}{6}.$$

DR. HASAN UNAL (MR Author ID: 1059881) received his masters and Ph.D. from Florida State University in 2005. He is currently working at Yildiz Technical University as an Associate Professor in Istanbul, Turkey. He is interested in visualization in teaching and learning mathematics.

MAA100
MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

From the Files of Past MAGAZINE Editors

Allen Schwenk 2006–2008

Former MAGAZINE editor Allen Schwenk said that editing the MAGAZINE was personally educational and stimulating, but a few exchanges with crackpots stick in his mind. One fellow was determined to disprove Euclid's Parallel Postulate. Consequently, he would send a diagram with 20 or 30 labeled angles, and then apply relations like opposite angles, interior-exterior angles, the sum of the angles in a triangle, complementary angles, and so on. After pages of algebra, he always ended with a contradiction. The first three or four times he submitted something like this, Schwenk painstakingly followed the argument until (inevitably) there would make a sign error. Schwenk tried to explain that his quest didn't make sense, since there is both Euclidean and non-Euclidean geometry, so one can accept or reject the parallel postulate and still have a consistent theory. He said he knew about that, but still believed that plane geometry did not satisfy the parallel postulate.

Selective Sums of an Infinite Series

R. JOHN FERDINANDS
Calvin College
Grand Rapids, MI 49546
ferd@calvin.edu

As every Calculus II student knows, the geometric series $\sum_{n=1}^{\infty} \frac{1}{2^n}$ converges with sum 1. But what happens if we sum only a subset of the terms of the series? For instance, the sum of the even terms is $\frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + \cdots = \frac{1/4}{1-1/4} = \frac{1}{3}$. Or we could sum finitely many terms; say $\frac{1}{2^3} + \frac{1}{2^7} = \frac{17}{128}$. The numbers $\frac{1}{3}$ and $\frac{17}{128}$ are examples of what Bonar and Khoury call the **selective sums** of the series [**1**, pp. 79–81]. Before you read on, you might want to guess what the set of all possible selective sums might be for the series $\sum_{n=1}^{\infty} \frac{1}{2^n}$, and also for $\sum_{n=1}^{\infty} \frac{2}{3^n}$.

It turns out to be helpful to formalize the concept of selective sums with the following definition.

Definition. *Let* $\sum_{n=1}^{\infty} a_n$ *be a series. A real number* x *is a selective sum of this series if* $x = \sum_{n=1}^{\infty} c_n a_n$*, where* $\{c_n\}$ *is a sequence consisting only of 0's and 1's.*

For the series $\sum_{n=1}^{\infty} \frac{1}{2^n}$, a selective sum must be a number between 0 and 1, because it is of the form $\sum_{n=1}^{\infty} \frac{c_n}{2^n}$, where $c_n = 0$ or 1 for each n. That is, each selective sum is of the form $0.c_1c_2c_3\ldots$ in binary notation. But every real number in [0, 1] can be written in this form. Therefore the set S of selective sums is [0, 1], so that S is as large as it could be for a series with positive terms and sum 1. We shall say that S is *maximal*.

The sum of the series $\sum_{n=1}^{\infty} \frac{2}{3^n}$ is 1 also. A selective sum is of the form $\sum_{n=1}^{\infty} \frac{2c_n}{3^n}$, where $2c_n = 0$ or 2 for each n. Hence the set S consists of all numbers in [0, 1] which can be expressed in base 3 notation without a 1. That is, S is the Cantor set. It is well known (see, for instance [**2**, p. 143]) that the Cantor set has measure zero.

These two geometric series illustrate the extreme cases: the case where S is maximal, and the case where it has measure zero. This suggests the question: for what types of series are the selective sums like those in the two extreme cases? And are there other possibilities?

Answers to these questions have appeared as a problem in Pólya and Szego [**5**], and in a paper by Menon [**3**]. The current paper obtains their results using a different approach.

For now we shall consider only convergent series whose terms are *positive* and *nonincreasing*. (In the last section we will return to more general series.) For convergent series with positive nonincreasing terms, we state a technical definition, which will simplify the arguments.

Math. Mag. **88** (2015) 179–185. doi:10.4169/math.mag.88.3.179.
MSC: Primary 40A05

Definition. *Let* $\sum_{n=1}^{\infty} a_n$ *be a convergent series such that the terms of the sequence* $\{a_n\}$ *are positive and nonincreasing. We say that a positive real number* x *is adequate with respect to the series if* $\sum_{i=n}^{\infty} a_i \geqslant x$*, where* n *is the smallest positive integer such that* $a_n \leqslant x$*. If for such* n*,* $\sum_{i=n}^{\infty} a_i < x$*, we say that* x *is inadequate with respect to the series. (Note that* n *must exist, since the series converges, and so* $\{a_n\}$ *has limit 0.)*

A different way to say this is that x is inadequate if, after we exclude all the terms of the series which are strictly greater than x, the sum of the remaining terms is strictly less than x. Since the terms are positive, we have the following necessary condition for a number to be a selective sum.

Necessity Lemma. *If* x *is inadequate, then* x *is not a selective sum. Equivalently, if* x *is a selective sum, then* x *is adequate.*

It is important to note that being adequate is not a sufficient condition for x to be a selective sum. For a counterexample, consider the series $\sum_{n=1}^{\infty} a_n$, where $a_1 = 6$, $a_2 = 4$, and $a_n = \frac{1}{2^{n-2}}$ for $n \geqslant 3$. The sum is $6 + 4 + \sum_{n=3}^{\infty} \frac{1}{2^{n-2}} = 10 + 1 = 11$. The number 8 is adequate, since $a_1 < 8$ and $\sum_{n=1}^{\infty} a_n > 8$. But it is easily seen that 8 is not a selective sum. Here is a sufficient condition for a number to be a selective sum.

Sufficiency Lemma. *Let* $\sum_{n=1}^{\infty} a_n$ *be a convergent series of positive and nonincreasing terms, and let* $x > 0$*. If every number in the interval* $(0, x]$ *is adequate, then* x *is a selective sum of the series.*

Proof. Suppose that the hypothesis is true for x. Assume that x is not equal to the sum of finitely many terms of the series. Let n_1 be the smallest integer such that $a_{n_1} < x$. Since x is adequate, $\sum_{i=n_1}^{\infty} a_i \geqslant x$. If equality holds we are done, so assume that $\sum_{i=n_1}^{\infty} a_i > x$. Keep adding terms, starting at a_{n_1}, until the first time that the sum $a_{n_1} + a_{n_1+1} + \cdots$ exceeds x. That is, find the largest positive integer $N_1 \geqslant n_1$ such that $\sum_{i=n_1}^{N_1} a_i < x < \sum_{i=n_1}^{N_1+1} a_i$.

Let n_2 be the smallest integer such that $a_{n_2} < x - \sum_{i=n_1}^{N_1} a_i$. (Note that $n_2 > N_1$.) Since every number in $(0, x]$ is adequate, and $x - \sum_{i=n_1}^{N_1} a_i < x$, it follows that $x - \sum_{i=n_1}^{N_1} a_i$ is adequate. Therefore $\sum_{i=n_2}^{\infty} a_i > x - \sum_{i=n_1}^{N_1} a_i$. (Again, if equality holds, we are done.) As in the previous step, keep adding terms until we reach a term a_{N_2} such that $\sum_{i=n_2}^{N_2} a_i < x - \sum_{i=n_1}^{N_1} a_i < \sum_{i=n_2}^{N_2+1} a_i$. Then $x - \left(\sum_{i=n_1}^{N_1} a_i + \sum_{i=n_2}^{N_2} a_i\right)$ is adequate, and $x - \left(\sum_{i=n_1}^{N_1} a_i + \sum_{i=n_2}^{N_2} a_i\right) < a_{N_2+1}$.

We repeat this process to get a subseries $\sum_{i=n_1}^{N_1} a_i + \sum_{i=n_2}^{N_2} a_i + \cdots + \sum_{i=n_k}^{N_k} a_i + \cdots +$ such that $x - \left(\sum_{i=n_1}^{N_1} a_i + \sum_{i=n_2}^{N_2} a_i + \cdots + \sum_{i=n_k}^{N_k} a_i\right)$ is adequate, and $x - \left(\sum_{i=n_1}^{N_1} a_i + \sum_{i=n_2}^{N_2} a_i + \cdots + \sum_{i=n_k}^{N_k} a_i\right) < a_{N_k+1}$. Since the partial sums are increasing and bounded above by x, the subseries converges. Since $a_{N_k+1} \to 0$ as $k \to \infty$, the sum of the subseries is x, and therefore x is a partial sum. This completes the proof of the lemma. ■

The key to determining the set S of selective sums is to observe how each term of the series compares to the sum of the succeeding terms. Let $R_n = \sum_{i=n+1}^{\infty} a_i$, where n is a nonnegative integer. If $a_n \leqslant R_n$ for all n, then S is maximal. If $a_n > R_n$ for all n, then S has properties which resemble the Cantor set. These cases are treated in Theorems 1 and 2, respectively. Theorem 1 is found in [**5**], and Theorem 2 in [**3**].

Theorem 1. *Let* $\sum_{n=1}^{\infty} a_n$ *be a convergent series of positive and nonincreasing terms, and let s be the sum of the series. The set S of selective sums is equal to* $[0, s]$ *if and only if* $a_n \leqslant R_n$ *for all* $n \geqslant 1$.

Proof. Suppose that $a_n \leqslant R_n$ for all $n \geqslant 1$. For x in $(0, s]$, let n be the smallest integer such that $a_n \leqslant x$. If $n = 1$, then $\sum_{i=1}^{\infty} a_i = s \geqslant x$, so x is adequate. If $n > 1$, then $a_n < x < a_{n-1}$, and $\sum_{i=n}^{\infty} a_i \geqslant a_{n-1} > x$, which implies that x is adequate. Therefore every number in $(0, s]$ is adequate. It follows from the sufficiency lemma that every number in $(0, s]$ is a selective sum. Since 0 is obviously a selective sum, $S = [0, s]$. ■

Now suppose that there is some $n \geqslant 1$ such that $a_n > R_n$. Then if x is a number for which $R_n = \sum_{i=n+1}^{\infty} a_i < x < a_n$, then x is inadequate, and hence is not a selective sum. Therefore, if $S = [0, s]$, then $a_n \leqslant R_n$ for all $n \geqslant 1$.

Theorem 2. *Let* $\sum_{n=1}^{\infty} a_n$ *be a convergent series of positive and nonincreasing terms, and s be the sum of the series. Suppose that* $a_n > R_n$ *for all* $n \geqslant 1$. *Then* $S = \bigcap_{n=1}^{\infty} F_n$ *where* F_n *is the union of the* 2^n *disjoint closed intervals of the form* $\left[\sum_{i=1}^{n} c_i a_i, \sum_{i=1}^{n} c_i a_i + R_n\right]$ *with each* c_i *equal to 0 or 1.*

Proof. Before proving the result, we show that the F_n's can be obtained in a way analogous to the construction of the Cantor set. Begin by removing the open interval (R_1, a_1) from $[0, s]$, leaving closed intervals of length R_1 on either side. (We are using the fact that $R_1 < a_1$.) Then $F_1 = [0, R_1] \cup [a_1, s]$. Next, remove an open interval from each interval of F_1, leaving intervals of length R_2 on either side; this gives $F_2 = [0, R_2] \cup [a_2, a_2 + R_2] \cup [a_1, a_1 + R_2] \cup [a_1 + a_2, s]$. We repeat the process to get a sequence $F_1 \supset F_2 \supset F_3 \supset \cdots$. It may be shown by induction that F_n has the form stated in the theorem. We leave the details to the reader.

Now we prove that $S = \bigcap_{n=1}^{\infty} F_n$. Let $x \in S$; by definition, $x = \sum_{i=1}^{\infty} c_i a_i$ for some sequence $\{c_i\}$ of 0's and 1's. For each n, $\sum_{i=1}^{n} c_i a_i \leqslant \sum_{i=1}^{\infty} c_i a_i = x = \sum_{i=1}^{n} c_i a_i + \sum_{i=n+1}^{\infty} c_i a_i \leqslant \sum_{i=1}^{n} c_i a_i + R_n$. Hence $x \in \left(\sum_{i=1}^{n} c_i a_i, \sum_{i=1}^{n} c_i a_i + R_n\right) \subseteq F_n$, and thus $S \subseteq \bigcap_{n=1}^{\infty} F_n$.

Now suppose that $x \in F_n$ for every $n \geqslant 1$. Define a sequence $\{c_n\}$ as follows. For each $n \geqslant 1$, x belongs to an interval in F_n, and is not in the open interval that is removed from it. If x is on the left of the deleted open interval, let $c_n = 0$; if it is on the right, let $c_n = 1$. It is straightforward to prove by induction that $\sum_{i=1}^{n} c_i a_i \leqslant x \leqslant \sum_{i=1}^{n} c_i a_i + R_n$ for every $n \geqslant 1$. Since $\sum_{n=1}^{\infty} a_n$ converges, $\lim_{n\to\infty} R_n = 0$. It follows that $x = \sum_{i=1}^{\infty} c_i a_i$, and hence $x \in S$. ■

To summarize the theorems, the set S of selective sums is maximal if and only if $a_n \leqslant R_n$ for all n, while if $a_n > R_n$ for all n, then S resembles the Cantor set. However S need not have measure zero, as the following corollary shows. (This fact is also found in [**3**].)

Corollary. *Under the hypotheses of Theorem 2, the measure of S is*

$$m(S) = \lim_{n\to\infty} (2^n R_n).$$

Proof. This follows easily from the fact that F_n consists of 2^n intervals, each of length R_n, and also from a standard result from measure theory [**2**, Proposition 14.8, p. 404], by which we see that $m(S) = m\left(\bigcap_{n=1}^{\infty} F_n\right) = \lim_{n\to\infty} m(F_n) = \lim_{n\to\infty} (2^n R_n)$. ■

Some examples

1. If $0 < r < 1$, the geometric series $\sum_{n=1}^{\infty} r^n$ converges with sum $\frac{r}{1-r}$. Observe that $r^n \leqslant \sum_{i=n+1}^{\infty} r^i$ if and only if $r^n \leqslant \frac{r^{n+1}}{1-r}$, if and only if $\frac{1}{2} \leqslant r$. By Theorem 1, the set of selective sums is $S = \left[0, \frac{r}{1-r}\right]$ if and only if $\frac{1}{2} \leqslant r \leqslant 1$. If $0 < r < \frac{1}{2}$, then $r^n > \sum_{i=n+1}^{\infty} r^i$ for all $n \geqslant 1$. Hence S has the structure described in Theorem 2, and its measure is $m(S) = \lim_{n\to\infty}\left(2^n \frac{r^{n+1}}{1-r}\right) = \frac{r}{1-r} \lim_{n\to\infty} (2r)^n = 0$, since $0 < 2r < 1$.

2. Here is an example of a series $\sum_{n=1}^{\infty} a_n$ for which S has the structure described in Theorem 2, but its measure is not zero. Let $a_1 = 2$, and $a_n = \frac{n^2+1}{2^n n(n-1)}$ for $n \geqslant 2$. Then $a_n = \frac{n}{(n-1)2^{n-1}} - \frac{n+1}{n2^n}$ for $n \geqslant 2$, yielding a telescoping series. It is easy to verify that $\sum_{i=n+1}^{\infty} a_i = \frac{n+1}{n2^n}$ for all $n \geqslant 1$, and that $a_n > \sum_{i=n+1}^{\infty} a_i$ for all $n \geqslant 1$. Hence the hypotheses of Theorem 2 are satisfied. The measure of S is given by $m(S) = \lim_{n\to\infty}\left(2^n \frac{n+1}{2^n n}\right) = 1$. More generally, it is shown in [**3**] how to construct a series satisfying the hypotheses of Theorem 2 for which S has any given measure α for $0 \leqslant \alpha \leqslant s$, where s is the sum of the series.

3. The p-series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges if and only if $p > 1$. By considering left sums for the graph of $f(x) = \frac{1}{x^p}$, we see that $\sum_{i=n+1}^{\infty} \frac{1}{i^p} > \int_{n+1}^{\infty} \frac{1}{x^p} dx = \frac{1}{(p-1)(n+1)^{p-1}} > \frac{1}{n^p}$ for all n greater than some N that depends on p. For all such N by Theorem 1, the set of selective sums of $\sum_{n=N}^{\infty} \frac{1}{n^p}$ is $[0, s]$, where $s = \sum_{n=N}^{\infty} \frac{1}{n^p}$.

The p-series gives examples of a third possibility for a_n and R_n: $a_n > R_n$ for only finitely many n. In this case, the set S of selective sums of $\sum_{n=1}^{\infty} a_n$ is the union of finitely many closed intervals. To see this, note that there is some $N > 1$ such that $a_n \leqslant R_n$ for all $n \geqslant N$. As we saw above, Theorem 1 implies that the set of selective sums of $\sum_{n=N}^{\infty} a_n$ is the interval $I = \left[0, \sum_{n=N}^{\infty} a_n\right]$. The set S of selective sums of $\sum_{n=1}^{\infty} a_n$ will consist of I and the intervals of the form $\sum_{i=1}^{N-1} c_i a_i + I$, where each c_i is equal to 0 or 1. This result is found in [**3**] also.

Selective sums for other series

Bonar and Khoury [**1**, pp. 80–81] have shown that every positive real number is a selective sum of the harmonic series $\sum_{n=1}^{\infty} \frac{1}{n}$. This suggests that series more general than those which have been considered in this paper (that is, divergent series, or series with terms that are not positive or nonincreasing) may also have interesting selective sums. For example, here is a result about absolutely convergent series (whose terms may be positive or negative).

Theorem 3. *Let $\sum_{n=1}^{\infty} a_n$ be an absolutely convergent series. Then the set S of selective sums is compact.*

Proof. Let $\{0, 1\}^{\omega}$ be the set of real sequences whose terms consist only of 0's and 1's. Define $f : \{0, 1\}^{\omega} \to S$ by $f(\{c_n\}) = \sum_{n=1}^{\infty} c_n a_n$. Note that f is well defined, since, by the comparison test, $\sum_{n=1}^{\infty} c_n a_n$ converges absolutely. Furthermore, f is continuous. (See Exercise 6 at the end of this paper.) By Tychonoff's theorem, $\{0, 1\}^{\omega}$ is compact. (See [**4**, p. 232].) Therefore $f(\{0, 1\}^{\omega}) = S$ is compact. ■

Questions for future work

It would be interesting to investigate the nature of the selective sums for other types of series. Here are some possible questions:

(a) Can the result for geometric series found in Example 1 be generalized to series $\sum_{n=1}^{\infty} a_n$ for which $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = r$, where $0 < r < 1$?

(b) What can be said about the selective sums of an alternating series?

(c) What can be said about the selective sums of a series of complex numbers?

(d) If $a_n > R_n$ for infinitely many n, and also $a_n \leqslant R_n$ for infinitely many n, what must be true about the set of selective sums? Exercise 4 in the Exercises at the end of this paper gives an example of such a series.

(e) A result in [**3**] implies that if $a_n > R_n$ for all $n \geqslant 1$, then the representation of a selective sum in the form $\sum_{n=1}^{\infty} c_n a_n$ is unique. Under what conditions could the number of representations of a selective sum be countably or uncountably infinite?

We end with some exercises for the reader.

Exercises

1. If $\{a_n\}$ is a sequence of positive terms with limit zero, prove that there exists a rearrangement of the terms such that they are nonincreasing. (Note that this result shows that if we are considering a convergent series with positive terms, there is no loss of generality in assuming that the terms are nonincreasing.)

2. Prove that for the series $\sum_{n=1}^{\infty} \frac{1}{n!} = e - 1$, the set of selective sums has measure zero. (Hint: Show that $\sum_{i=n+1}^{\infty} \frac{1}{i!} < \frac{1}{n!}$.)

3. Let $\sum_{n=1}^{\infty} a_n$ be a convergent series with positive nonincreasing terms. If $a_n > R_n$ for all but finitely many n, describe the set of selective sums.

4. Define a series $\sum_{n=1}^{\infty} a_n$ by $a_{2n-1} = \frac{2}{3}\left(\frac{1}{6}\right)^{n-1}$ for $n \geqslant 1$, and $a_{2n} = \left(\frac{1}{6}\right)^{n}$ for $n \geqslant 1$.

   (a) Prove that for all $n \geqslant 1$, $a_{2n-1} > \sum_{i=2n}^{\infty} a_i$ and $a_{2n} = \sum_{i=2n+1}^{\infty} a_i$.

   (b) Find the set of selective sums of the series, and show that it has measure zero.

5. Let $\sum_{n=1}^{\infty} a_n$ be a series which satisfies the hypotheses of Theorem 2. Show that the set S of selective sums is perfect, and totally disconnected.

   (The results of Exercise 5 are due to [**3**]. See [**2**, p. 147] for the definition of a perfect set. When proving that S is perfect, consider separately the cases where an element of S can, and cannot, be written as the sum of finitely many terms of the series.)

6. Prove that the function $f : \{0, 1\}^{\omega} \to S$ by $f(\{c_n\}) = \sum_{n=1}^{\infty} c_n a_n$ defined in the proof of Theorem 3 is continuous. (Hint: Define a sequence of functions $\{f_n\}$ by $f_n(\{c_n\}) = c_n a_n$, and use the Weierstrass M-Test [**2**, p. 247].)

Acknowledgment I thank the four anonymous referees for their many helpful and detailed comments. I am especially grateful to the referee who pointed out that Theorems 1 and 2 are found in [**3**] and [**5**]. Finally, thanks to the Editor for his help and encouragement.

REFERENCES

1. D D. Bonar M J. Khoury, *Real Infinite Series*. MAA Textbooks, Washington, DC, 2006.
2. S G. Krantz, *Real Analysis and Foundations*. Second edition. Chapman and Hall/CRC Press, London, 2005.

3. P. K Menon, On a class of perfect sets, *Bull Amer Math Soc* **54** no. 8 (1948) 706–711.
4. J R. Munkres, *Topology: A First Course*. Prentice-Hall. Englewood Cliffs, NJ, 1975.
5. Polya and Szego, *Problems and Theorems in Analysis II*. Springer-Verlag, New York, 1990, Part 1, Chapter 3, no. 131.

Summary. The sum of a subset of the terms of an infinite series is called a selective sum of the series. We describe the set of all selective sums for some series, and we show that for a series which satisfies certain conditions, the set of selective sums can be obtained in a way analogous to the construction of the Cantor set.

R. JOHN FERDINANDS (MR Author ID: 616048) received his Ph.D. in mathematics from Purdue University under the direction of Reinhardt Schultz in 1988. He is a professor of mathematics at Calvin College.

1 D	2 O	3 F	4 F	■	5 B	6 U	7 S	8 T	■	■	9 Z	10 O	11 R	12 N
13 A	V	E	R	■	14 I	N	C	H	■	15 S	A	U	T	E
16 T	A	L	E	■	17 G	I	R	O	■	18 E	G	R	E	T
19 A	L	L	E	20 N	D	O	E	R	21 F	E	R	■	■	■
■	■	■	22 B	O	O	N	E	■	23 U	S	E	24 F	25 U	26 L
27 A	28 L	29 B	E	R	S	■	30 C	31 O	G	■	32 B	A	R	E
33 S	E	E	E	M	■	34 S	H	R	U	35 G	■	36 C	A	T
37 H	E	D	R	I	38 C	K	■	39 S	E	E	40 B	A	C	H
41 O	R	D	■	42 E	R	A	43 T	O	■	44 N	A	D	I	A
45 R	E	E	46 D	■	47 U	T	A	■	48 E	T	B	E	L	L
49 E	D	D	I	50 E	D	■	51 M	52 A	N	L	Y	■	■	■
■	■	■	53 A	L	E	54 X	A	N	D	E	R	55 S	56 O	57 N
58 S	59 M	60 E	L	L	■	61 E	L	K	S	■	62 U	L	N	A
63 P	A	R	E	E	■	64 N	E	H	I	■	65 T	A	U	S
66 F	O	R	D	■	■	67 A	S	S	T	■	68 H	Y	P	H

Proof Without Words: An Elegant Property of the Equilateral Triangle

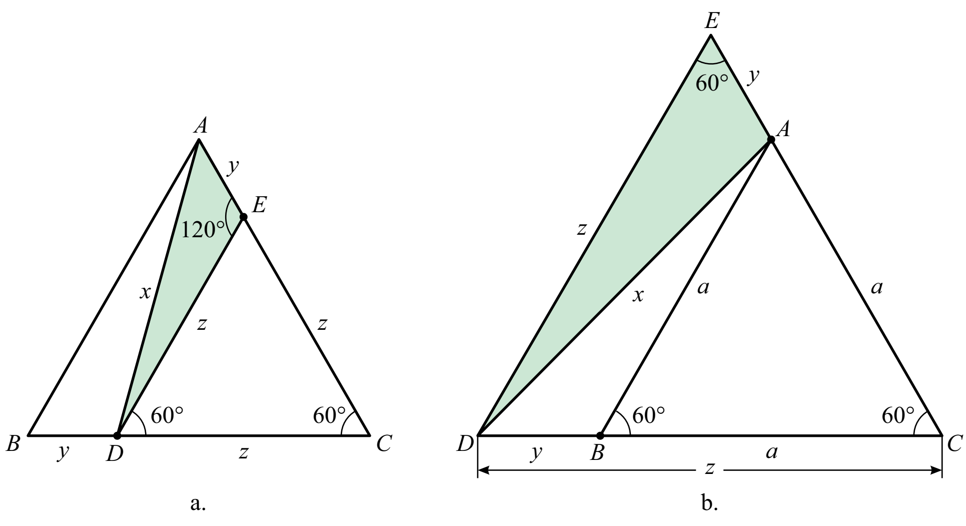
VICTOR OXMAN
Western Galilee College,
Acre, Israel
victor.oxman@gmail.com

MOSHE STUPEL
Shaanan College,
Haifa, Israel
stupel@bezeqint.net

For an equilateral triangle ΔABC and a point D on the ray CB, let x, y, z denote the distances from the point D to the vertices A, B, and C, respectively. The following relationships hold:

a. If $x \geq z$, then $x^2 = y^2 + z^2 + y \cdot z$.
b. If $x < z$, then $x^2 = y^2 + z^2 - y \cdot z$.

Proof:



Summary. Figures are used to show that for an equilateral triangle ABC and some point D on the ray CB, the distances from the point D to the vertices of the triangle satisfy different polynomial equations depending on whether the point D is between B and C or not.

VICTOR OXMAN (MR Author ID: 756975) teaches at Western Galilee College. His research interests are in geometry and mathematical education.

MOSHE STUPEL (MR Author ID: 1028564) teaches at Shaanan College. His research interests are also in geometry and mathematical education.

Math. Mag. **88** (2015) 186. doi:10.4169/math.mag.88.3.186.
MSC: Primary 51M04.

The Average Height of Catalan Trees by Counting Lattice Paths

NACHUM DERSHOWITZ
School of Computer Science
Tel Aviv University
Tel Aviv, Israel
nachumd@post.tau.ac.il

CHRISTIAN RINDERKNECHT
Montpellier, France
rinderknecht@free.fr

In memoriam Philippe Flajolet, friend and colleague

Structured documents, like books, articles, and web pages, are composed of chapters, sections, paragraphs, figures, appendices, indices, etc. The occurrences of these components are mutually constrained; for instance, it is understood that a section is part of a chapter and that appendices are located at the end of a document. This hierarchical layout is meant to facilitate reading, and it supports the search for specific items of information. When considering computer systems, these data must be uniformly encoded by means of a formal language.

Consider, for instance, an email message. It contains at least the sender's address, a subject or title, the recipient's address, and a body of text. These elements correspond to *nodes* arranged in a structure called a *Catalan tree*, a.k.a. an ordered tree or rooted plane tree. For example, the email

```
From: Me
Subject: Homework
To: You

A deadline is a due date for a
homework.
```

can be modeled by the tree in FIGURE 1, where the topmost node ("email") is called the *root* and the framed pieces of text are *leaves*. Note that, for historical reasons, computer scientists grow their trees upside down, with the root at the top. The inner (nonleaf) nodes hold "metadata," or "markup," that is, information about the nature of the data contained in the subtree.

Catalan trees are a pervasive data structure in computer science, in that they are a natural representation for hierarchical data. For example, in XML (eXtensible Markup Language), textual information is stored in leaves, and, consequently, its retrieval requires the traversal of the tree from the root to a leaf. The *height* of a tree is the number of nodes on a maximal path from root to leaf; for example, travel down the path with nodes depicted as ○ in the tree of height 5 in FIGURE 2.

In general, the maximum cost of a search is proportional to the height of the tree, and the determination of the average height becomes relevant when performing a series of random searches [**16**]. The mathematical study of this average quantity often relies on advanced analytical tools, and the purpose of the present article is to propose a partial simplification of these approaches by using elementary combinatorics.

Math. Mag. **88** (2015) 187–195. doi:10.4169/math.mag.88.3.187.
MSC: Primary 05A19, Secondary 05A16.

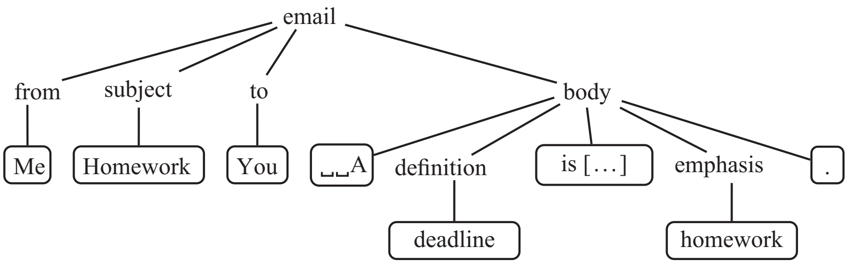


Figure 1 An email viewed as a Catalan tree.

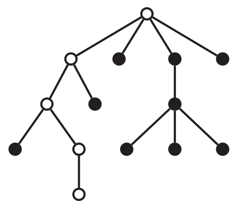


Figure 2 Catalan tree of height 5 and size 13.

The analytical derivation

We measure the *size* of a tree by the number of its edges; for example, the tree in FIGURE 2 is of size 13. Let h_n be the average height of Catalan trees of size n and H_n^h the number of Catalan trees of size n and height h. We then have $h_n = S_n/C_n$, where $S_n := \sum_{h\geqslant 1} h\, H_n^h$, and $C_n := \binom{2n}{n}/(n+1)$ is the number of Catalan trees of size n. The height of a tree with n edges can range from 2 (all leaves directly below the root) to $n+1$ (one straight path from root to a lone leaf).

To get a handle on the sum S_n, we may define $H_n^{<h}$ as being the number of trees with n edges and height less than h. Then $H_n^h = H_n^{<h+1} - H_n^{<h}$. Of course, we have $H_n^{<h} = H_n^{<n+2} = C_n$, if $h > n+1$. Formulas can be further simplified by letting $H_n^{\geqslant h}$ be the number of trees with n edges and height greater than or equal to h. Now we have:

$$S_n = \sum_{h\geqslant 1} h\left(H_n^{<h+1} - H_n^{<h}\right) = \sum_{h\geqslant 1} h\left(H_n^{\geqslant h} - H_n^{\geqslant h+1}\right) = \sum_{h\geqslant 1} H_n^{\geqslant h}. \tag{1}$$

Knuth, de Bruijn, and Rice [**12**] published a landmark paper in 1972, where they obtained the asymptotic approximation of the average height h_n. They started by modeling the problem with a generating function [**17**] that satisfies a recurrence equation whose solution expresses the generating function in terms of continued fractions of Fibonacci polynomials. Integration over complex numbers is then utilized to obtain the formula

$$H_n^{\geqslant h} = \sum_{k\geqslant 1}\left[\binom{2n}{n+1-kh} - 2\binom{2n}{n-kh} + \binom{2n}{n-1-kh}\right]. \tag{2}$$

The authors conclude by employing real and complex analysis to obtain asymptotic expansions of $H_n^{\geqslant h}$, S_n, and h_n. As we will see, the main term is $h_n \sim \sqrt{\pi n}$, where $f(n) \sim g(n)$ means $\lim_{n\to\infty} f(n)/g(n) = 1$, wherever f and g are defined.

The purpose of the present note is to show how to circumvent heavy analytic techniques in the derivation of equation (2). Instead, we propose an elementary combinatorial proof based on the enumeration of the Dyck paths of a certain height, which are in bijection with Catalan trees of a related height. We find this bijective proof to be more intuitive, in particular to computer scientists, for whom the result matters for the analysis of algorithms. Technically, our approach is in tune with Mohanty [**13**], as well as Dershowitz and Zaks [**3**].

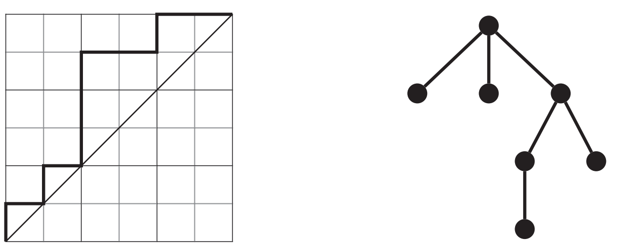
Counting Catalan trees

Before we determine $H_n^{\geqslant h}$, let us solve a related and easier question: deriving the number C_n of Catalan trees with n edges, called the *Catalan number*.

In 1984, Kemp [**10**, p. 64] (see also [**6**]) derived equation (2) by analytical means too, but, instead of working directly with Catalan trees, he used certain lattice paths in an integer grid. *Monotonic lattice paths* [**9, 13**] are made up of two kinds of steps, oriented upward and oriented rightward, starting at $(0, 0)$ with an upward step. *Dyck paths* of length $2n$ are monotonic paths ending at (n, n) that never venture below the diagonal; an example for $n = 6$ is shown in FIGURE 3(a).

A bijection with Dyck paths Crucially, there is a bijection between Dyck paths of length $2n$ and Catalan trees with n edges [**11**].

This bijection is shown on an example in FIGURE 3. To construct the lattice path in FIGURE 3(a) from the tree in FIGURE 3(b), we imagine that the tree is a roadmap and our avatar plans a tour starting at the root as follows: We take the rightmost unvisited road (from the avatar's viewpoint), else we backtrack. In the end, we have taken each road twice: there and back again. More technically, in FIGURE 4, we follow the dotted arrows: Each downward arrow in the tree corresponds to a step up $\uparrow$ (called a *rise*) in the lattice, and an upward arrow in the tree to a step right $\rightarrow$ (called a *fall*). In the tree, the series $\downarrow \uparrow \downarrow \uparrow \downarrow \downarrow \downarrow \uparrow \uparrow \downarrow \uparrow \uparrow$ becomes $\uparrow \rightarrow \uparrow \rightarrow \uparrow \uparrow \uparrow \rightarrow \rightarrow \uparrow \rightarrow \rightarrow$ in the lattice. If we follow the latter from the start at the bottom left corner $(0, 0)$, we obtain the path in FIGURE 3(a). This kind of traversal is called *preorder*, or "document" order, because it is the way we would read the document represented by the tree, from cover



(a) Dyck path of length 12. (b) Catalan tree with 6 edges.

Figure 3 Bijection between Dyck paths and Catalan trees.

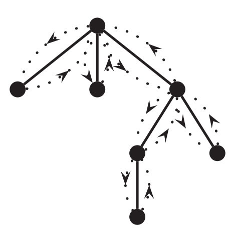


Figure 4 Preorder traversal

to cover. Note as well that there are always $n + 1$ nodes if, and only if, there are n edges in the tree because there is precisely one edge per node going up, save for the topmost node (*root*).

The inclusion-exclusion principle The previous bijection allows us to count the Catalan trees with n edges by counting instead the Dyck paths of length $2n$.

It is easy to count all the monotonic paths of length $2n$ because there are as many as choices of n rises among $2n$ steps, that is, $\binom{2n}{n}$. To count only the Dyck paths, we need to subtract the number of paths that start with a rise but cross below the diagonal at some point.

This approach is a simple instance of the method known as the *inclusion-exclusion principle*, whereby the direct and difficult enumeration of a set is replaced by an easier enumeration of a strict superset and the subtraction of the cardinality of a strict subset so that the resulting sets are equal.

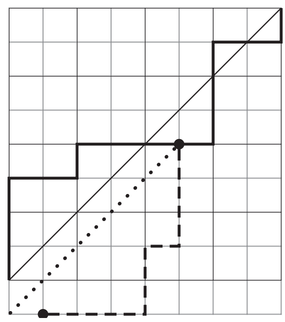


Figure 5 Reflection of a prefix with respect to $y = x - 1$.

An example of a path that is not a Dyck path is shown in FIGURE 5, drawn in bold. The first point reached below the diagonal is used to plot a dotted line parallel to the diagonal back to the y-axis. All the steps on the path from that point back to $(0, 0)$ are then changed into their counterpart: A rise is replaced by a fall and vice versa. The resulting segment is drawn as connected dashed lines. This operation is called a *reflection* [**14**]. The crux of the matter is that we can reflect each monotonic path crossing the diagonal into a distinct path from $(1, -1)$ to (n, n). These reflected paths can, in turn, be reflected back into their original counterpart when they reach the dotted line. In other words, the mapping is bijective. (Another intuitive and visual approach to the same result has been published by Callan [**1**].) Consequently, there are as many monotonic paths from $(0, 0)$ to (n, n) that cross the diagonal as there are monotonic paths from $(1, -1)$ to (n, n). The latter are readily enumerated: $\binom{2n}{n-1}$. In conclusion, the number of Dyck paths of length $2n$ is

$$C_n = \binom{2n}{n} - \binom{2n}{n-1} = \binom{2n}{n} - \frac{(2n)!}{(n-1)!(n+1)!} \tag{3}$$
$$= \binom{2n}{n} - \frac{n}{n+1}\,\frac{(2n)!}{n!n!} = \binom{2n}{n} - \frac{n}{n+1}\binom{2n}{n} = \frac{1}{n+1}\binom{2n}{n}.$$

Using Stirling's formula for the asymptotic equivalence, we draw the conclusion

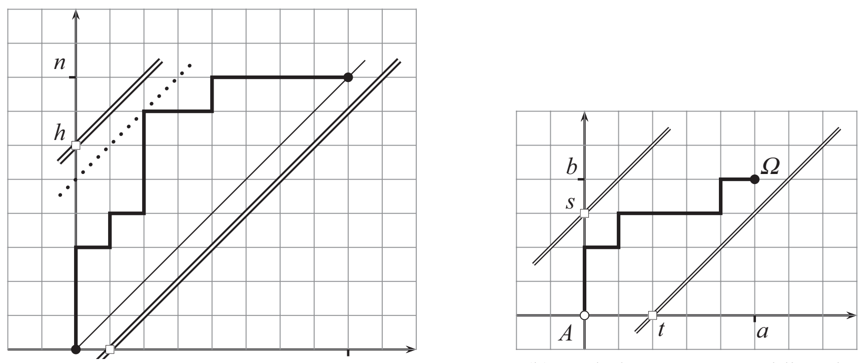
$$C_n = \frac{1}{n+1}\binom{2n}{n} \sim \frac{4^n}{n\sqrt{\pi n}}, \quad \text{as } n \to \infty. \tag{4}$$

A combinatorial proof

In 1996, Sedgewick and Flajolet [**7, 15**] derived the enumerations of Catalan trees by height, also using analytic combinatorics, but they employed real analysis to obtain the asymptotic approximation of $H_n^{\geqslant h}$. They write [**15**, p. 260]:

> *This analysis is the hardest nut that we are cracking in this book. It combines techniques for solving linear recurrences and continued fractions, generating function expansions, especially by the Lagrange inversion theorem, and binomial approximations and Euler-Maclaurin summations.*

To avoid the aforementioned advanced techniques used to derive equation (2), we use again a bijection between Dyck paths and Catalan trees, but, this time, the key point is that Catalan trees of size n and height h are in bijection with Dyck paths of length $2n$ and height $h-1$. This simple observation allows us to reason about the height of the Dyck paths and transfer our findings back to Catalan trees.



(a) Dyck path of length $2n$ and height $h-1$.

(b) Path from A to Ω avoiding the boundaries $y=x+s$ and $y=x-t$.

Figure 6 Paths avoiding diagonal boundaries.

With the determination of $H_n^{<h}$ in mind, let us consider a Dyck path of length $2n$ and height $h-1$, as in FIGURE 6(a). The double lines are boundaries that may not be attained by the path. This is in fact a special case of a general monotonic path between two diagonal boundaries, as shown in FIGURE 6(b), where s denotes the vertical distance from A, and t the horizontal distance from A. It is well known that the number of monotonic paths from $A(0,0)$ to $\Omega(a,b)$ avoiding the boundaries $y=x+s$ and $y=x-t$ is

$$|\mathcal{L}(a,b;t,s)| = \sum_{k\in\mathbb{Z}}\left[\binom{a+b}{b+k(s+t)} - \binom{a+b}{b+k(s+t)+t}\right]. \tag{5}$$

The proof by Mohanty [**13**, p. 6] of this formula is based on the reflection principle and the principle of inclusion and exclusion, which we used earlier. We quote his proof here verbatim because it is rarely found in print nowadays.

Proof (Mohanty [**13**]) For brevity, call the boundaries $x=y+t$ and $x=y-s$, $\mathcal{L}^+$ and $\mathcal{L}^-$, respectively. Denote by A_1 the set of paths that reach $\mathcal{L}^+$, by A_2 the set of paths that reach $\mathcal{L}^+$, $\mathcal{L}^-$ in that order, and in general by A_i the set of paths reaching $\mathcal{L}^+$, $\mathcal{L}^-$, $\mathcal{L}^+$, … (i times) in the specified order. Similarly, let B_i be the set of paths reaching $\mathcal{L}^-$, $\mathcal{L}^+$, $\mathcal{L}^-$, ... (i times) in the specified order. An application of the usual inclusion-exclusion method yields

$$|\mathcal{L}(a, b; t, s)| = \binom{a+b}{b} + \sum_{i \geqslant 1} (-1)^i \left(|A_i| + |B_i|\right), \tag{6}$$

where $|A_i|$ and $|B_i|$ are evaluated by using the reflection principle repeatedly. For example, consider A_3. Since every path in A_3 must reach $\mathcal{L}^+$, A_3 when reflected about $\mathcal{L}^+$ becomes the set of paths from $(t, -t)$ to (a, b) each of which reaches $\mathcal{L}^+$ after reaching $\mathcal{L}^-$. Another reflection about $\mathcal{L}^-$ would make A_3 equivalent to the set of paths from $(-s-t, s+t)$ to (a, b) that reach $\mathcal{L}^+$, which in turn can be written as $R(a+s+t, b-s-t; 2s+3t)$. [Note: $R(a, b; t)$ is the set of paths from $(0, 0)$ to (a, b) reflected about $\mathcal{L}^+$.] Thus, since $|R(a, b; t)| = \binom{a+b}{a-t}$, we have

$$|A_3| = \binom{a+b}{a-s-2t},$$

and, more generally,

$$\left|A_{2j}\right| = \binom{a+b}{a+j(s+t)} \quad \text{and} \quad \left|A_{2j+1}\right| = \binom{a+b}{a-j(s+t)-t}.$$

The expressions for $|B_{2j}|$, $|B_{2j+1}|$, $j = 0, 1, 2, \ldots$, with $|A_0|$, $|B_0|$ being $\binom{a+b}{b}$, are obtained by interchanging a with b and s with t. Substitution of these values in (6) yields (5) after some simplifications. □

Resuming our argument, if we match the subfigures in FIGURE 6, we find $s = h$, $t = 1$, $a = b = n$, hence $a + b = 2n$ and $b + k(s + t) = n + k(h + 1)$, which we plug into formula (5) and change h into $h - 1$:

$$H_n^{<h} = \sum_{k \in \mathbb{Z}} \left[\binom{2n}{n+kh} - \binom{2n}{n+1+kh}\right].$$

After splitting the sum into $k < 0$, $k = 0$, and $k > 0$, then changing the sign of k in the first case, using $\binom{p}{q} = \binom{p}{p-q}$ in the second, and lastly gathering the remaining sums ranging over $k \geqslant 1$, we reach

$$H_n^{<h} = -\sum_{k \geqslant 1} \left[\binom{2n}{n+1-kh} - 2\binom{2n}{n-kh} + \binom{2n}{n-1-kh}\right] + \binom{2n}{n} - \binom{2n}{n-1}.$$

Recognizing C_n from equation (3), we simplify as follows:

$$C_n - H_n^{<h} = \sum_{k \geqslant 1} \left[\binom{2n}{n+1-kh} - 2\binom{2n}{n-kh} + \binom{2n}{n-1-kh}\right].$$

Finally, recalling that $H_n^{\geqslant h} = C_n - H_n^{<h}$, we arrive at

$$H_n^{\geqslant h} = \sum_{k \geqslant 1} \left[\binom{2n}{n+1-kh} - 2\binom{2n}{n-kh} + \binom{2n}{n-1-kh}\right],$$

which is none other than our target, equation (2).

In this way, we have achieved our goal merely by enumerating lattice paths and hopefully have, in the process, made this classic result less daunting.

Asymptotics

We could stop here, but we would like to give a hint as to how the asymptotic approximation is carried out. The approximation will give us a practical handle on the expected height of Catalan trees, which in turn tells us what to expect by way of performance of algorithms, like search, that traverse down paths in arbitrary trees.

Equation (1) entails $S_n = \sum_{h\geqslant 1} H_n^{\geqslant h}$; therefore,

$$S_n = \sum_{k'\geqslant 1} d(k')\left[\binom{2n}{n+1-k'} - 2\binom{2n}{n-k'} + \binom{2n}{n-1-k'}\right],$$

where $d(k')$ is the number of positive divisors of k', but complex analysis is needed [**5**, **12**]. Another way is to express the binomials in terms of $\binom{2n}{n-kh}$:

$$\begin{aligned}\binom{2n}{n-m+1} &= \frac{(2n)!}{(n-m+1)!\,(n+m-1)!}\\ &= \frac{(2n)!\,(n+m)}{(n-m)!\,(n-m+1)(n+m)!} = \frac{n+m}{n-m+1}\binom{2n}{n-m},\\ \binom{2n}{n-m-1} &= \frac{(2n)!}{(n-m-1)!\,(n+m+1)!}\\ &= \frac{(2n)!\,(n-m)}{(n-m)!\,(n+m)!\,(n+m+1)} = \frac{n-m}{n+m+1}\binom{2n}{n-m}.\end{aligned}$$

Therefore,

$$\binom{2n}{n-m+1} - 2\binom{2n}{n-m} + \binom{2n}{n-m-1} = 2\,\frac{2m^2-(n+1)}{(n+1)^2-m^2}\binom{2n}{n-m}.$$

Let $F_n(m) = (2m^2-n)/(n^2-m^2)$. We have

$$S_n = 2\sum_{h\geqslant 1}\sum_{k\geqslant 1} F_{n+1}(kh)\binom{2n}{n-kh}.$$

From equation (4) and $h_n = S_n/C_n$, we deduce $h_n = (n+1)S_n/\binom{2n}{n}$, hence we must approximate $(n+1)F_{n+1}(m)$ and $\binom{2n}{n-m}/\binom{2n}{n}$. On the one hand, we have

$$F_{n+1}(m) \sim \frac{2m^2-n}{n^2} \sim \frac{2m^2-n}{n(n+1)},$$

so $(n+1)F_{n+1}(kh) \sim 2k^2h^2/n - 1$. On the other hand, Sedgewick and Flajolet [**15**, 4.6, 4.8] show

$$\binom{2n}{n-m}\bigg/\binom{2n}{n} \sim e^{-m^2/n}.$$

Assuming that the tails (the implicit error terms) of the two previous approximations decrease exponentially, we have

$$h_n \sim \sum_{h\geqslant 1}\sum_{k\geqslant 1}(4k^2h^2/n-2)e^{-k^2h^2/n} = \sum_{h\geqslant 1} H(h/\sqrt{n}),$$

where $H(x) = \sum_{k\geqslant 1}(4k^2x^2-2)e^{-k^2x^2}$. Finally, Sedgewick and Flajolet [**15**, §5.9], on the one hand, and Graham, Knuth, and Patashnik [**8**, §9.6], on the other hand, use real analysis to conclude

$$h_n \sim \sum_{h\geqslant 1} H(h/\sqrt{n}) \sim \sqrt{n}\int_0^\infty H(x)dx \sim \sqrt{\pi n}.$$

The end of this derivation is difficult because the error terms in the bivariate asymptotic approximations must be carefully checked, so it is unlikely to be simplified further.

Remarkably, the main term $\sqrt{\pi n}$ in the asymptotic value for the average height can also be obtained by simple lattice-path arguments [**4**], as shown in the online supplement to this article [**2**].

REFERENCES

1. D. Callan, Pair them up! A visual approach to the Chung–Feller theorem, *College Math. J* **26** no. 3 (1995) 196–198.
2. N. Dershowitz, C. Rinderknecht, A Purely Combinatorial Derivation of the Average Height of Catalan Trees, 2015. http://www.maa.org/mathmag/supplements.
3. N. Dershowitz, S. Zaks, Applied tree enumerations, in *Proceedings of the Sixth Colloquium on Trees in Algebra and Programming,* Vol 112, Lecture Notes in Computer Science, Springer, Berlin, 1981.
4. N. Dershowitz, S. Zaks, The cycle lemma and some applications, *European J. Combin.* **11** no. 1 (1990) 35–40.
5. P. Flajolet, X. Gourdon, P. Dumas, MellinTransforms and asymptotics: Harmonic sums, *Theoret. Comput. Sci.*, **144** (1995) 3–58.
6. P. Flajolet, M. Nebel, H. Prodinger, The Scientific Works of Rainer Kemp (1949–2004), *Theoret. Comput. Sci.* **355** no. 3 (2006) 371–381.
7. P. Flajolet, R. Sedgewick, *Analytic Combinatorics.* Cambridge Univ. Press, Cambridge, MA 2009.
8. R. L. Graham, D. E. Knuth, O. Patashnik, *Concrete Mathematics.* Second edition. Addison-Wesley, Reading, MA, 1994.
9. K. Humphreys, A history and a survey of lattice path enumeration, *J. Statist. Plann. Inference* **140** no. 8 (2010) 2237–2254.
10. R. Kemp, *Fundamentals of the Average Case Analysis of Particular Algorithms.* Wiley-Teubner Series in Computer Science, John Wiley & Sons, B. G. Teubner, Stuttgart, 1984.
11. D. A. Klarner, Correspondence between plane trees and binary sequences, *J. Comb. Theory* **9** (1970) 401–411.
12. D. E. Knuth, N. G. de Bruijn, S. O. Rice. *The Average Height of Planted Plane Trees,* Academic Press, New York, 1972. 15–22. Republished in Selected Papers on the Analysis of Algorithms, CSLI Lecture Notes 102, Stanford University, CA, 2000. 215–223.
13. S. G. Mohanty, *Lattice Path Counting and Applications.* Vol. 37, Probability and Mathematical Statistics, Academic Press, New York, 1979.
14. M. Renault, Lost (and found) in translation: André's actual method and its application to the generalized ballot problem, *Amer. Math. Monthly* **155** no. 4 (2008) 358–363.
15. R. Sedgewick, P. Flajolet, *An Introduction to the Analysis of Algorithms.* Second edition. Addison-Wesley, Boston, MA, 2013.
16. J. S. Vitter, P. Flajolet, Average-case analysis of algorithms and data structures in *Algorithms and Complexity,* Vol. A, Handbook of Theoretical Computer Science, Elsevier Science, 1990.
17. H. S. Wilf, *Generating functionology.* Academic Press, 1990.

Summary. The average height of Catalan trees of a given size is a structural parameter important in the analysis of algorithms, as it measures the expected maximum cost of a search in a tree. This parameter has been studied

first with generating functions and complex variable theory, yielding an asymptotic approximation. Later on, real analysis was used instead of complex analysis. We have further reduced the conceptual difficulty by replacing generating functions with the enumeration of monotonic lattice paths, whose graphical representations make the derivation much more intuitive.

NACHUM DERSHOWITZ (MR Author ID: 56990) is a professor of computational logic at Tel Aviv University. His graduate degrees in applied mathematics are from the Weizmann Institute of Science. He is coauthor of *Calendrical Calculations* (Cambridge Univ. Press, Fourth edition, forthcoming) and *Calendrical Tabulations* (Cambridge Univ. Press, 2002) and was elected to Academia Europaea in 2013.

CHRISTIAN RINDERKNECHT (MR Author ID: 899221) is an expert in compilers. He was an assistant professor in France, Korea, and Hungary. He graduated from Université Pierre et Marie Curie in Paris, and did his doctoral studies at INRIA. He is the author of *Design and Analysis of Purely Functional Programs* (College Publications, Third edition, London).

MAA100
MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

From the Files of Past MAGAZINE Editors

Frank Farris 2001–2005; 2008–2009

Former MAGAZINE editor Frank Farris had a way of working with papers (and with the people who wrote them) that was both firm and kind.

When you are an editor, you have to make a lot of hard decisions, and it's only human that you worry about what people will think about those decisions. Frank used to brace himself for that inevitable encounter when someone would come up to him at the math meetings and accost him for the decisions he had made.

One meeting he went to, it happened. A mathematician he knew fairly well stopped him and said, "You just rejected the paper of one of my colleagues." There was a long pause (. . . wait for it, wait for it . . .) and then the mathematician added, "He told me it was the kindest and most helpful letter he had ever received."

Some People Have All the Luck

RICHARD ARRATIA
University of Southern California
Los Angeles, CA 90089-2532
rarratia@usc.edu

SKIP GARIBALDI
University of California, Los Angeles
Los Angeles, California 90095-7121
skip@member.ams.org

LAWRENCE MOWER
Palm Beach Post
West Palm Beach, FL 33405
lawrence_mower@pbpost.com

PHILIP B. STARK
University of California
Berkeley, CA 94720-3860
stark@stat.berkeley.edu

It is unusual to win a lottery prize worth \$600 or more. No one we know has. But ten people have each won more than 80 such prizes in the Florida Lottery. This seems fishy. Someone might get lucky and win the Mega Millions jackpot (a 1-in-259 million chance) having bought just one ticket. But it's implausible that a gambler would win many unlikely prizes without having bet very many times.

How many? We pose an optimization problem whose answer gives a lower bound on any sensible estimate of an alleged gambler's spending: over all possible combinations of Florida Lottery bets, what is the minimum amount spent so that, if *every* Florida resident spent that much, the chance that *any* of them would win so many times is still less than one in a million? If that amount is implausibly large compared to that gambler's means, we have statistical evidence that she is up to something.

Solving this optimization problem in practice hinges on two math facts:

- an inequality that lets us bound the probability of winning dependent bets in some situations in which we do not know precisely which bets were made, and
- log-concavity of the regularized Beta function, which lets us show that any local minimizer attains the global minimal value.

We conclude that 2 of the 10 suspicious gamblers could just be lucky. The other 8 are chiseling or spending implausibly large sums on lottery tickets. These results were used by one of us (LM) to focus on-the-ground investigations and to support an exposé of lax security in the Florida lottery [**19**]. We describe what those investigations found, and the policy consequences in Florida and other states.

How long can a gambler gamble?

Is there a non-negligible probability that a pathological gambler of moderate means could win many \$600+ prizes? If not, we are done: our suspicion of these 10 gamblers is justified.

Math. Mag. **88** (2015) 196–211. doi:10.4169/math.mag.88.3.196. © Mathematical Association of America
MSC: Primary 91A60, Secondary 97K80.

So, suppose a gambler starts with a bankroll of S_0 and buys a single kind of lottery ticket over and over again. If he spends his initial bankroll and all his winnings, how much would he expect to spend in total and how many prizes would he expect to collect before going broke?

Let the random variable X denote the value of a ticket, payoff minus cost. We assume that

$$\mathbb{E}(X) < 0, \tag{1}$$

because that is the situation in the games where our suspicious winners claimed prizes. (It does infrequently happen that lottery tickets can have positive expectation, see [**14**] or [**1**].) Assumption (1) and the Law of Large Numbers say that a gambler with a finite bankroll eventually will run out of money, with probability 1. The question is: *how fast?*

Write $c > 0$ for the cost of the ticket, so that

$$\mathbb{P}(X \geq -c) = 1 \quad \text{and} \quad \mathbb{P}(X = -c) \neq 0. \tag{2}$$

To illustrate our assumptions and notation, let's look at a concrete example of a Florida game, Play 4. It is based on the *numbers* or *policy* game formerly offered by organized crime, described in [**16**] and [**22**]. Variations on it are offered in most states that have a lottery.

Example 1. (Florida's Play 4 game) Our ten gamblers claimed many prizes in Florida's Play 4 game, although in 2012 it only accounted for about 6% of the Florida Lottery's \$4.45 billion in sales. Here are the rules, simplified in ways that don't change the probabilities.

The Lottery draws a 4-digit random number twice a day. A gambler can bet on the next drawing by paying $c = \$1$ for a ticket, picking a 4-digit number, and choosing "straight" or "box."

If the gambler bets "straight," she wins \$5000 if her number matches the next 4-digit number exactly (which has probability $p = 10^{-4}$). She wins nothing otherwise. The expected value of a straight ticket is $\mathbb{E}(X) = \$5000 \times 10^{-4} - \$1 = -\$0.50$.

If a gambler bets "box," she wins if her number is a permutation of the digits in the next 4-digit number the Lottery draws. She wins nothing otherwise. The probability of winning this bet depends on the number of distinguishable permutations of the digits the gambler selects.

For instance, if the gambler bets on 1112, there are 4 possible permutations, 1112, 1121, 1211, and 2111. This bet is a "4-way box." It wins \$1198 with probability $1/2500 = 4 \times 10^{-4}$, since 4 of the 10,000 equally likely outcomes are permutations of those four digits. If the gambler bets on 1122, there are 6 possible permutations of the digits; this bet is called a "6-way box." It wins \$800 with probability 6×10^{-4}. (The 6-way box is relatively unpopular, accounting for less than 1% of Play 4 tickets.) Buying such a ticket has expected value $\mathbb{E}(X) \approx -\$0.52$. Similarly, there are 12-way and 24-way boxes.

In the abstract setting, the gambler's bankroll after t bets is

$$S_t := S_0 + X_1 + X_2 + \cdots + X_t,$$

where $X_1, \ldots, X_t$ are i.i.d. (independent, identically distributed) random variables with the same distribution as X, and X_i is the net payoff of the ith ticket. The gambler can no longer afford to keep buying tickets after the Tth one, where T is the smallest $t \geq 0$ for which $S_t < c$.

Proposition 1. *In the notation of the preceding paragraph,*

$$\frac{S_0 - c}{|\mathbb{E}(X)|} < \mathbb{E}(T) \le \frac{S_0}{|\mathbb{E}(X)|},$$

with equality on the right if S_0 and all possible values of X are integer multiples of c.

In most situations, S_0 is much larger than c, and the two bounds are almost identical. In expectation, the gambler spends a total of $c\mathbb{E}(T)$ on tickets, including all of his winnings, which amount to $c\mathbb{E}(T) - S_0$.

Proof. By the definition of T and (2),

$$0 \le \mathbb{E}(S_T) < c, \tag{3}$$

with equality on the left in case S_0 and X are integer multiples of c. Now the crux is to relate $\mathbb{E}(T)$ to $\mathbb{E}(S_T)$. If T were constant (instead of random), then $T = \mathbb{E}T$ and we could simply write

$$\mathbb{E}(S_T) = \mathbb{E}\left(S_0 + \sum_{i=1}^{\mathbb{E}T} X_i\right) = S_0 + \mathbb{E}(T)\,\mathbb{E}(X). \tag{4}$$

Combining (4) with (3) would give the claim. The key is that equation (4) holds even though T is random—this is Wald's equation (see, e.g., [**9**, §5.4]). The essential property is that T is a *stopping time*, i.e., for every $k > 0$, whether or not one places a kth bet is determined just from the outcomes of the first $k - 1$ bets.

You might recognize that in this discussion that we are considering a version of the gambler's ruin problem but with an unfair bet and where the house has infinite money; for bounds on gambler's ruin without these hypotheses, see, e.g., [**10**].

A ticket with just one prize The proposition lets us address the question from the beginning of this section. Suppose a ticket pays j with probability p and nothing otherwise; the expected value of the ticket, $\mathbb{E}(X) = pj - c$, is negative; and j is an integer multiple of c. If a gambler starts with a bankroll of S_0 and spends it all on tickets, successively using the winnings to buy more tickets, then by Proposition 1 the gambler should expect to buy $\mathbb{E}(T) = S_0/(c - pj)$ tickets, which means winning

$$\frac{c\mathbb{E}(T) - S_0}{j} = \frac{pS_0}{c - pj}.$$

prizes.

Example 2. How many prizes might a compulsive gambler of "ordinary" means claim? Surely some gamblers have lost houses, so let us say he starts with a bankroll worth $S_0 = \$175{,}000$, an amount between the median list price and the median sale price of a house in Florida [**24**]. If he always buys Play 4 6-way box tickets and recycles his winnings to buy more tickets, the previous paragraph shows that he can expect to win about

$$pS_0/(c - pj) = 6 \times 17.5/0.52 \approx 202 \text{ times.}$$

This is big enough to put him among the top handful of winners in the history of the Florida lottery.

Hence, the number of wins alone does not give evidence that a gambler cheated. We must take into account the particulars of the winning bets.

A toy version of the problem

From here on, a "win" means a win large enough to be recorded; for Florida, the threshold is \$600. Suppose for the moment that a gambler only buys one kind of lottery ticket, and that each ticket is for a different drawing, so that wins are independent. Suppose each ticket has probability p of winning.

A gambler who buys n tickets spends cn and, on average, wins np times. This is intuitively obvious, and follows formally by modeling a lottery bet as a Bernoulli trial with probability p of success: in n trials we expect np successes.

We don't know n, and the gambler is unlikely to tell us. But based on the calculation in the preceding paragraph, we might guess that a gambler who won W times bought roughly W/p tickets. Indeed, an unbiased estimate for n is $\hat{n} := W/p$, corresponding to the gambler spending $c\hat{n}$ on tickets. Since p is very small, like 10^{-4}, the number $\hat{n}$ is big—and so is the estimated amount spent, $c\hat{n}$. (Note that this estimate includes any winnings "reinvested" in more lottery tickets.)

A gambler confronted with $\hat{n}$ might quite reasonably object that she is just very lucky, and that the true number of tickets she bought, n, is much smaller. Under the assumptions in this section, her tickets are i.i.d. Bernoulli trials, and the number of wins W has a binomial distribution with parameters n and p, which lets us check the plausibility of her claim by considering

$$D(n; w, p) := \boxed{\begin{array}{l}\text{probability of at least } w\\ \text{wins with } n \text{ tickets}\end{array}} = \sum_{k=w}^{n} \binom{n}{k} p^k (1-p)^{n-k}. \tag{5}$$

Modeling a lottery bet as a Bernoulli trial is precisely correct in the case of games like Play 4. But for scratcher games, there is a very large pool from which the gambler is sampling without replacement by buying tickets; as the pool is much larger than the values of n that we will consider, the difference between drawing tickets with and without replacement is negligible.

Example 3. (Louis Johnson) Of the 10 people who had won more than 80 prizes each in the Florida Lottery, the second most-frequent prize claimant was Louis Johnson. He claimed $W = 57$ \$5,000 prizes from straight Play 4 tickets (as well as many prizes in many other games that we ignore in this example). We estimate that he bought $\hat{n} = W/p = 570{,}000$ tickets at a cost of \$570,000.

What if he claimed to only have bought $n = 175{,}000$ tickets? The probability of winning at least 57 times with 175,000 tickets is

$$D(175000; 57, 10^{-4}) \approx 6.3 \times 10^{-14}.$$

For comparison, by one estimate there are about 400 billion stars in our galaxy [**15**]. Suppose there were a list of all those stars, and two people independently pick a star at random from that list. The chance they would pick the same star is minuscule, yet it is still 40 times greater than the probability we just calculated. It is utterly implausible that a gambler wins 57 times by buying 175,000 or fewer tickets.

What this has to do with Joe DiMaggio

The computation in Example 3 does not directly answer whether Louis Johnson is lucky or up to something shady. The most glaring problem is that we have calculated the probability that a *particular* innocent gambler who buys \$175,000 of Play 4 tickets would win so many times. The news media have publicized some lottery coincidences as astronomically unlikely, yet these coincidences have turned out to be relatively unsurprising given the enormous number of people playing the lottery; see, for example, [**7**, esp. p. 859] or [**23**] and the references therein.

Among other things, we need to check whether so many people are playing Play 4 so frequently that it's reasonably likely at least one of them would win at least 57 times. If so, Louis Johnson might be that person, just like with Mega Millions: no particular ticket has a big chance of winning, but if there are enough gamblers, there is a big chance *someone* wins.

We take an approach similar to how baseball probability enthusiasts attempt to answer the question, *Precisely how amazing was Joe DiMaggio?* Joe DiMaggio is famous for having the longest hitting streak in baseball: he hit in 56 consecutive games in 1941. (The modern player with the second longest hitting streak is Pete Rose, who hit in 44 consecutive games in 1978.) One way to frame the question is to consider the probability that a randomly selected player gets a hit in a game, and then estimate the probability that there is at least one hitting streak at least 56 games long in the entire history of baseball. If a streak of 56 or more games is likely, then the answer to the question is "not so amazing"; DiMaggio just happened to be the person who had the unsurprisingly long streak. If it is very unlikely that there would be such a long streak, then the answer is: DiMaggio was truly amazing. (The conclusions in DiMaggio's case have been equivocal, see the discussion in [**13**, pp. 30–38].)

Let's apply this reasoning to Louis Johnson's 57 Play 4 wins (Example 3). Suppose that N gamblers bought Play 4 tickets during the relevant time period, each of whom spent at most \$175,000. Then an upper bound on the probability that at least one such gambler would win at least 57 times is the chance of at least one success in N Bernoulli trials, each of which has probability no larger than $p \approx 6.3 \times 10^{-14}$ of success. (Louis Johnson represents a success.) The trials might not be independent, because different gamblers might bet on the same numbers for the same game, but the chance that at least one of the N gamblers wins at least 57 times is at most Np by the Bonferroni bound (for any set of events $A_1, \dots, A_N$, $\mathbb{P}(\cup_{i=1}^N A_i) \leq \sum_{i=1}^N \mathbb{P}(A_i)$).

What is N? Suppose it's the current population of Florida, approximately 19 million. Then the chance at least one person would win at least 57 times is no larger than $19 \times 10^6 \times 6.3 \times 10^{-14} = 0.0000012$, just over one in a million.

This estimate is crude because the estimated number of gamblers is very rough and of course the estimate is not at all sharp (it gives a lot away in the direction of making the gambler look less suspicious) because most people spend nowhere near \$175,000 on the lottery. We are giving even more away because Louis Johnson won many other bets (his total winnings are, of course, dwarfed by the expected cash outlay). Considering all these factors, one might reasonably conclude that either Louis Johnson has a source of hidden of money—perhaps he is a wealthy heir with a gambling problem—or he is up to something.

Example 4. (Louis Johnson 2) In Example 3 we picked the \$175,000 spending level almost out of thin air, based on Florida house prices as in Example 2. Instead of starting with a limit on spending and deducing the probability of a number of wins, let's start with a probability, $\varepsilon = 5 \times 10^{-14}$, and infer the minimum spending required to have at least that probability of so many wins.

If Johnson buys n tickets, then he wins at least 57 times with probability $D(n; 57, 10^{-4})$. We compute n_0, the smallest n such that

$$D(n; 57, 10^{-4}) \geq \varepsilon,$$

which gives $n_0 = 174{,}000$.

Using the Bonferroni bound again, we find that the probability, if *everyone* in Florida spent \$174,000 on straight Play 4 tickets, the chance that *any* of them would win 57 times or more is less than one in a million.

Multiple kinds of tickets

Real lottery gamblers tend to wager on a variety of games with different odds of winning and different payoffs. Suppose they place b different kinds of bets. (It might feel more natural to say "games," but a gambler could place several dependent bets on a single Play 4 drawing: straight, several boxes, etc.)

Number the bets $1, 2, \ldots, b$. Bet i costs c_i dollars and has probability p_i of winning. The gambler won more than the threshold on bet i w_i times. We don't know n_i, the number of times the gambler wagered on bet i. If we did know the vector $\vec{n} = (n_1, n_2, \ldots, n_b)$, then we might be able to calculate the probability:

$$P(\vec{n}; \vec{w}, \vec{p}) := \begin{pmatrix}\text{probability of winning at least } w_i \text{ times on bet } i \text{ with} \\ n_i \text{ tickets, for all } i\end{pmatrix}. \qquad (6)$$

As in Example 4, we can find a lower bound on the amount spent to attain w_i wins on bet i, $i = 1, \ldots, b$, by solving

$$\vec{c} \cdot \vec{n}^* = \min_{\vec{n}} \vec{c} \cdot \vec{n} \quad \text{such that} \quad n_i \geq w_i \quad \text{and} \quad P(\vec{n}; \vec{w}, \vec{p}) \geq \varepsilon. \qquad (7)$$

For a typical gambler that we study, this lower bound $\vec{c} \cdot \vec{n}^*$ will be in the millions of dollars. Thinking back to the "Joe DiMaggio" justification for why (7) is a lower bound, it is clear that not every resident of Florida would spend so much on lottery tickets, and our gut feeling is that a more refined justification would produce a larger lower bound for the amount spent.

But how can we find $P(\vec{n}; \vec{w}, \vec{p})$? If the different bets were on independent events (say, each bet is a different kind of scratcher ticket), then

$$P(\vec{n}; \vec{w}, \vec{p}) = \prod_{i=1}^{b} \begin{pmatrix}\text{probability of winning at least } w_i \\ \text{times on bet } i \text{ with } n_i \text{ tickets}\end{pmatrix} = \prod_{i=1}^{b} D(n_i; w_i, p_i). \qquad (8)$$

But gamblers can make dependent bets, in which case (8) does not hold. Fortunately, it is possible to derive an upper bound for the typical case, as we now show.

No dependent wins is almost as good as independent bets

For most of the 10 gamblers, we did not observe wins on dependent bets, such as a win on a straight ticket and a win on a 4-way box ticket for the same Play 4 drawing. We seek to prove Proposition 2 (below), which gives an upper bound on the probability that we observe at least so many wins but no wins on dependent bets.

Abstractly, we envision a finite number d of independent drawings, such as a sequence of Play 4 drawings. For each drawing j, $j = 1, \ldots, d$, the gambler may bet any amount on any of b different bets (such as 1234 straight, 1344 6-way box, etc.), whose outcomes—for drawing j—may be dependent, but whose outcomes on different draws are independent. We write p_i for the probability that a bet on i wins in any particular drawing; p_i is the same for all drawings j.

For $i = 1, \ldots, b$ and $j = 1, \ldots, d$, let $n_{ij} \in \{0, 1\}$ be the indicator that the gambler wagered on bet i in drawing j, so that ith row sum, $n_i := \sum_j n_{ij}$, is the total number of bets on i. We call the entire system of bets B, represented by the b-by-d zero-one matrix $B = [n_{ij}]$.

Proposition 2. *Suppose that, for each i, a gambler wagers on bet i in n_i different drawings, as specified by B, above. Given the bets B, consider the events*

$$W_i := (\textit{gambler wins bet } i \textit{ at least } w_i \textit{ times with bets } B),$$

and the event

$$I := (\textit{in each drawing } j\textit{, the gambler wins at most one bet}).$$

Then

$$\mathbb{P}(I \cap W_1 \cap \cdots \cap W_b) \leq \prod_{i=1}^{b} \mathbb{P}(W_i). \tag{9}$$

In our case, $\mathbb{P}(W_i) = D(n_i; w_i, p_i)$, so we restate (9) as

$$\mathbb{P}(I \cap W_1 \cap \cdots \cap W_b) \leq \prod_{i=1}^{b} D(n_i; w_i, p_i). \tag{10}$$

Proposition 2 is intuitively plausible: even though the bets are not independent, the drawings are, and event I guarantees that any single drawing helps at most one of the events $\{W_i\}$ to occur. We prove Proposition 2 as a corollary of an extension of a celebrated result, the BKR inequality, named for van den Berg–Kesten–Reimer, conjectured in [**4**], and proved in [**21**] and [**3**] (or see [**5**]). The remainder of this section provides the details. The original BKR inequality is stated as Theorem 1. We separate the purely set-theoretic aspects of the discussion, from the probabilistic aspects.

The BKR operation $\Box$ Let S be an arbitrary set, and write S^d for the Cartesian product of d copies of S. Since our application is probability, we call an element $\omega = (\omega_1, \ldots, \omega_d) \in S^d$ an *outcome*, and we call any $A \subseteq S^d$ an *event*.

For a subset $J \subseteq \{1, \ldots, d\}$ and an outcome $\omega \in S^d$, the *J-cylinder of ω*, denoted (J, ω), is the collection of $\omega' \in S^d$ such that $\omega'_j = \omega_j$ for all $j \in J$. For events $A_1, A_2, \ldots, A_b$, let $A_1 \Box A_2 \Box \cdots \Box A_b \subseteq S^d$ be the set of ω for which there exist pairwise disjoint $J_1, J_2, \ldots, J_b \subseteq \{1, \ldots, d\}$ such that $(J_i, \omega) \subseteq A_i$ for all i. The case $b = 2$, where one combines just two events, is the context for the original BKR inequality as in [**4**, p. 564]; the operation with $b > 2$ is new and is the main study of this section.

Here is another definition of $\Box$ that might be more transparent. Given an event $A \subseteq S^d$ and a subset $J \subseteq \{1, \ldots, d\}$, define the event

$$[A]_J := \{\omega \in A \mid (J, \omega) \subseteq A\} = \bigcup_{\{\omega \mid (J,\omega) \subseteq A\}} (J, \omega).$$

Informally, $[A]_J$ consists of the outcomes in A, such that by looking only at the coordinates indexed by J, one can tell that A must have occurred. Evidently, for $A, B \subseteq S^d$,

$$A \subseteq B \text{ implies } [A]_J \subseteq [B]_J \quad \text{and} \quad J \subseteq K \text{ implies } [A]_J \subseteq [A]_K. \tag{11}$$

The definition of $\Box$ becomes

$$\mathop{\Box}_{1 \leq i \leq b} A_i := \bigcup_{\text{pairwise disjoint } J_1, \ldots, J_b \subseteq \{1, \ldots, d\}} [A_1]_{J_1} \cap [A_2]_{J_2} \cap \cdots \cap [A_b]_{J_b}. \tag{12}$$

We read the above definition as "$\square_{1\le i\le b} A_i$ is the event that all b events occur, with b disjoint sets of reasons to simultaneously certify the b events." Informally, the outcome ω, observed only on the coordinate indices in J_i, supplies the "reason" that we can certify that event A_i occurs.

Our notation $\square_{1\le i\le b} A_i \equiv A_1 \,\square\, A_2 \,\square \cdots \square\, A_b$ is intentionally analogous to the notations for set intersection, $\bigcap_{1\le i\le b} A_i \equiv A_1 \cap A_2 \cap \cdots \cap A_b$, and set union, $\bigcup_{1\le i\le b} A_i \equiv A_1 \cup A_2 \cup \cdots \cup A_b$. The multi-input operator $\square$ is, like set intersection $\bigcap$ and set union $\bigcup$, fully commutative, i.e., unchanged by any re-ordering of the inputs. Unlike intersection and union, $\square$ is not associative, as we now show.

Example 5. Take $S = \{0, 1\}$, $d = 3$, and

$$A = (0, *, *) \cup (1, 0, *), \quad B = (0, *, *) \cup (1, 1, *), \quad C = (*, 0, 1),$$

where we write for example $(1, 0, *) = \{(1, 0, 0), (1, 0, 1)\} = (\{1, 2\}, (1, 0, s))$ for $s = 0, 1$ and $(0, *, *) = \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1)\}$. Note that $|A| = |B| = 6$. Then $A \,\square\, B = (0, *, *)$, $(A \,\square\, B) \,\square\, C = \{(0, 0, 1)\}$—using $J_1 = \{1\}$ and $J_2 = \{2, 3\}$ in (12)—but $B \,\square\, C = \{(0, 0, 1)\}$ and $A \,\square\, (B \,\square\, C) = \emptyset$. Also, $A \,\square\, B \,\square\, C = \emptyset$.

The connection between lottery drawings and $\square$ Before continuing to discuss the BKR operation $\square$ in the abstract, we consider what it means for lottery drawings. We take $S = \{0, 1, \ldots, 2^b - 1\}$ to encode the results of a single draw: an element $s \in S$ answers, for each of the b bets, whether that bet wins or not. The sample space for our probability model is S^d; the jth coordinate ω_j reports the results of the b bets on the jth draw.

It is easy to see that, in the notation of Proposition 2,

$$(I \cap W_1 \cap \cdots \cap W_b) \subseteq \square_{1}^{b} W_i. \tag{13}$$

Indeed, given an outcome $\omega \in I \cap W_1 \cap \cdots \cap W_b$, we can take, for $i = 1$ to b, $J_i := \{j \mid \text{on draw } j, \text{bet } i \text{ wins and } n_{ij} = 1\}$. Since $\omega \in I$, the sets $J_1, \ldots, J_b$ are mutually disjoint; and since $\omega \in W_i$, $|J_i| \ge n_i$. Hence, $(J_i, \omega) \subseteq W_i$, and thus $\omega \in [W_i]_{J_i}$, for $i = 1$ to b.

Example 6. The left hand side of (13) can be a strict subset of the right hand side. For example, with $b = 2$ bets and $d = 2$ draws, suppose that $w_1 = w_2 = 1$ and the gambler lays both bets on both draws. The outcome where both bets win on both draws is not in the left side of (13) but is in $W_1 \,\square\, W_2$.

To write this example out fully, we think of the binary encoding, $S = \{0, 1, 2, 3\}$ corresponding to $\{00, 01, 10, 11\}$, so that, for example, $0 \in S$ represents a draw where both bets lose, $1 \in S$ represents the outcome 01 where the first bet loses and the second bet wins, $2 \in S$ represents the outcome 10 where the first bet wins and the second bet loses, and $3 \in S$ represents the outcome 11 where both bets win.

The event I is the set of $\omega = (\omega_1, \omega_2)$ for which no coordinate ω_j is equal to 3. The event W_1 is the set of ω such that at least one of the coordinates is equal to 2 or 3, and the event W_2 is the set of ω such that at least one of the coordinates is equal to 1 or 3. Certainly,

$$I \cap W_1 \cap W_2 = \{(1, 2), (2, 1)\},$$

yet

$$W_1 \,\square\, W_2 = \{(1, 2), (2, 1), (1, 3), (2, 3), (3, 1), (3, 2), (3, 3)\}.$$

Set theoretic considerations related to the BKR inequality It is obvious that, for events $B_1, \dots, B_r \subseteq S^d$ and $J \subseteq \{1, \cdots, d\}$,

$$\left[\bigcap_{1\le i\le r} B_i\right]_J = \bigcap_{1\le i\le r} [B_i]_J \quad \text{and} \quad \left[\bigcup_{1\le i\le r} B_i\right]_J \supseteq \bigcup_{1\le i\le r} [B_i]_J. \tag{14}$$

For unions, the containment may be strict, as in Example 5, where $A \cup B = S^d$ hence $[A \cup B]_{\{1\}} = S^d$, whereas $[A]_{\{1\}} = [B]_{\{1\}} = (0, *, *)$.

Lemma 1. (Composition of cylinder operators) *For $A \subseteq S^d$ and $J, K \subseteq \{1, \cdots, d\}$,*

$$[[A]_J]_K = [A]_{J\cap K}.$$

Proof. Suppose first that $\omega \in [[A]_J]_K$. That is, $(K, \omega) \subseteq [A]_J$: if $\omega'' \in S^d$ agrees with ω on K, then $(J, \omega'') \subseteq A$. We must show that ω is in $[A]_{J\cap K}$; i.e., if ω'' is in $(J \cap K, \omega)$, then ω'' is in A.

Given $\omega'' \in (J \cap K, \omega)$, pick ω' to agree with ω on K and ω'' on $S^d \setminus K$. Then ω' agrees with ω'' on $(S^d \setminus K) \cup (J \cap K)$, so on J, i.e., $\omega'' \in A$, proving $\subseteq$.

We omit the proof of the containment $\supseteq$, which is easier.

Proposition 3. *For $A_1, A_2, \dots, A_b \subseteq S^d$, we have*

$$\mathop{\square}_1^b A_i \subseteq (\cdots((A_1 \square A_2) \square A_3) \cdots \square A_{b-1}) \square A_b.$$

Proof. By induction, using (11), it suffices to prove that

$$\left(\mathop{\square}_1^b A_i\right) \subseteq \left(\mathop{\square}_1^{b-1} A_i\right) \square A_b.$$

With unions over $K \subseteq \{1, \dots, d\}$ and pairwise disjoint $J_1, J_2, \dots$,

$$\left(\mathop{\square}_1^{b-1} A_i\right) \square A_b = \bigcup_K \left(\left[\mathop{\square}_{i=1}^{b-1} A_i\right]_K \cap [A_b]_{K^c}\right) \tag{15}$$

$$= \bigcup_K \left(\left[\bigcup_{J_1,\dots,J_{b-1}} \bigcap_{i=1}^{b-1} [A_i]_{J_i}\right]_K \cap [A_b]_{K^c}\right) \tag{16}$$

$$\supseteq \bigcup_K \left(\left(\bigcup_{J_1,\dots,J_{b-1}} \bigcap_{i=1}^{b-1} [[A_i]_{J_i}]_K\right) \cap [A_b]_{K^c}\right) \tag{17}$$

$$= \bigcup_K \left(\left(\bigcup_{J_1,\dots,J_{b-1}} \bigcap_{i=1}^{b-1} [A_i]_{J_i\cap K}\right) \cap [A_b]_{K^c}\right) \tag{18}$$

$$= \bigcup_{K_1,\dots,K_b} \bigcap_{i=1}^{b} [A_i]_{K_i} = \mathop{\square}_1^b A_i. \tag{19}$$

The justifications are as follows. Line (15) follows by using a $b = 2$ version of the definition (12), where K^c denotes the complement of K. Line (16) follows by using the definition (12) with b replaced by $b - 1$. The set inclusion in line (17) results from applying both parts of (14). Line (18) follows by applying Lemma 1 on the composition of cylinder operators. Line (19) is just relabeling the indices: the previous line is a union, indexed by pairwise disjoint $J_1, \ldots, J_{b-1}$, and a set K; for $i = 1$ to $b - 1$, $K_i = J_i \cap K$, and for index b, we take $K_b = K^c$—the set of possible indices $\alpha = (J_1 \cap K, \ldots, J_{b-1} \cap K, K^c)$ is identical to the set of $\alpha = (K_1, \ldots, K_b)$, with $i \neq j$ implies $K_i \cap K_j = \emptyset$.

Probability considerations related to the BKR inequality References to the BKR inequality were given just after Equation (10).

Theorem 1. (The original BKR inequality) *Let S be a finite set, and let $\mathbb{P}$ be a probability measure on S^d for which the d coordinates are mutually independent. (The coordinates might have different distributions.) For any events $A, B \subseteq S^d$, with the event $A \,\square\, B$ as defined by* (12)*,*

$$\mathbb{P}(A \,\square\, B) \leq \mathbb{P}(A)\mathbb{P}(B).$$

Corollary 1. *Under the hypotheses of Theorem 1, for $b = 2, 3, \ldots$ and $A_1, \ldots, A_b \subseteq S^d$,*

$$\mathbb{P}(A_1 \,\square\, A_2 \,\square \cdots \square\, A_b) \leq \prod_{i=1}^{b} \mathbb{P}(A_i). \tag{20}$$

Proof. For $b = 2$, (20) is the original BKR inequality. For $b \geq 3$, we apply Proposition 3 to see that

$$\mathbb{P}(A_1 \,\square \cdots \square\, A_b) \leq \mathbb{P}((\cdots((A_1 \,\square\, A_2) \,\square\, A_3) \cdots \square\, A_{b-1}) \,\square\, A_b).$$

Applying the $b = 2$ case and induction proves the claim.

We can now prove Proposition 2, which from our new perspective is a simple corollary of the extended BKR inequality, Corollary 1.

Proof of Proposition 2 In view of the containment (13), we have

$$\mathbb{P}(I \cap W_1 \cap \cdots \cap W_b) \leq \mathbb{P}\left(\mathop{\square}_{1}^{b} W_i\right),$$

and by Corollary 1

$$\mathbb{P}\left(\mathop{\square}_{1}^{b} W_i\right) \leq \prod \mathbb{P}(W_i).$$

The optimization problem we actually solve

In order to exploit the material in the previous section, we *replace* definition (6) of P with

$$P(\vec{n}; \vec{w}, \vec{p}) := \begin{pmatrix}\text{probability of winning at least } w_i \text{ times on bet } i \text{ with} \\ n_i \text{ tickets, for all } i, \text{ and no wins on dependent bets}\end{pmatrix};$$

from Proposition 2, we know that then

$$P(\vec{n}; \vec{w}, \vec{p}) \leq \prod_{i=1}^{b} D(n_i; w_i, p_i). \tag{21}$$

We will find a lower bound $\vec{c}\cdot\vec{n}^*$ on the amount spent by a gambler who did not win dependent bets by solving not (7), but rather

$$\vec{c}\cdot\vec{n}^* = \min_{\vec{n}} \vec{c}\cdot\vec{n} \quad \text{such that} \quad n_i \geq w_i \quad \text{and} \quad \prod_{i=1}^{b} D(n_i; w_i, p_i) \geq \varepsilon. \tag{22}$$

We also relax the requirement that the numbers of bets, the n_i's, be integers and we extend the domain of D to include positive real values of n_i as in [**2**, p. 945, 26.5.24]

$$D(n; w, p) = I_p(w, n-w+1), \quad \text{where} \quad I_x(a,b) := \frac{\int_0^x t^{a-1}(1-t)^{b-1}\,\mathrm{d}t}{\int_0^1 t^{a-1}(1-t)^{b-1}\,\mathrm{d}t} \tag{23}$$

is the *regularized Beta function*. The function I_x, or at least its numerator and denominator, are available in many scientific computing packages, including Python's SciPy library. Extending the domain of the optimization problem to nonintegral n_i can only decrease the lower bound $\vec{c}\cdot\vec{n}^*$, and it brings two benefits, which we now describe.

In our examples, $\prod D(w_i; w_i, p_i)$ is much less than ε, and consequently $n_i^* > w_i$ for some i. As $D(n; w, p)$ is monotonically increasing in n, we have an *equality* $\prod D(n_i^*; w_i, p_i) = \varepsilon$. This is the first benefit, and it implies by (21) an inequality $P(\vec{n}^*; \vec{w}, \vec{p}) \leq \varepsilon$. Therefore, as in the discussion of Joe DiMaggio, if all N people in the gambling population spent at least $\vec{c}\cdot\vec{n}^*$ on tickets, the probability that one or more of the gamblers would win at least w_i times on bet i for all i is at most $N\varepsilon$. To say it differently: *the solution $\vec{c}\cdot\vec{n}^*$ to* (22) *is an underestimate of the minimum plausible spending required to win so many times.*

The second benefit of extending the domain of the optimization problem is to make the problem convex instead of combinatorial. The convexity allows us to show that any local minimum (as found by the computer) attains the global minimal value.

Proposition 4. *A local minimizer $\vec{n}^*$ for the optimization problem* (22) *(relaxed to include noninteger values of n_i) attains the global minimal value.*

Proof. We shall show that the set of values of $\vec{n}$ over which we optimize, the *feasible set*,

$$\left\{\vec{n}\in\mathbb{R}^b \mid n_i \geq w_i \text{ for all } i\right\} \cap \left\{\vec{n}\in\mathbb{R}^b \mid \prod\nolimits_i D(n_i; w_i, p_i) \geq \varepsilon\right\}, \tag{24}$$

is convex. As the *objective function* $\vec{c}\cdot\vec{n}$ is linear in $\vec{n}$, the claim follows.

The first set in (24) defines a polytope, which is clearly convex. Because the intersection of two convex sets is convex, it remains to show that the second set is also convex.

The logarithm is a monotonic function, so taking the log of both sides of an inequality preserves the inequality, and we may write the second set in (24) as

$$\left\{\vec{n}\in\mathbb{R}^b \mid \sum\nolimits_i \log D(n_i; w_i, p_i) \geq \log\varepsilon\right\}. \tag{25}$$

For $0 \leq x \leq 1$ and α, β positive, the function

$$\beta \mapsto \log I_x(\alpha, \beta)$$

is concave by [**12**, Cor. 4.6(iii)]. Hence $\log D(n_i; w_i, p_i)$ is concave for $n_i \geq w_i$. A sum of concave functions is concave, so the set (25) is convex, proving the claim.

Example 7. (Louis Johnson 3) If we solve (22) for Louis Johnson's wins—including not only his Pick 4 wins but also many of his prizes from scratcher games—we find a minimum amount spent of at least \$2 million for $\varepsilon = 5 \times 10^{-14}$.

Monotonicity Some of the gamblers we studied for the investigative report claimed prizes in more than 50 different lottery games. In such cases it is convenient to solve (22) for only a subset of the games to ease computation by reducing the number of variables. Since removing restrictions results in minimizing the same function over a set that strictly includes the original set, the resulting "relaxed" optimization problem still gives a lower bound for the gambler's minimum amount spent.

The man from Hollywood

Louis Johnson's astounding 252 prizes is beaten by a man from Hollywood, Florida, whom we refer to as "H." During the same time period, H claimed 570 prizes, more than twice as many as Johnson did. Yet Mower's news report [**19**] stimulated a law enforcement action against Johnson but not against H. Why?

All but one of H's prizes are in Play 4, which is really different from scratcher games: if you buy \$100 worth of scratcher tickets for a single \$1 game, this amounts to 100 (almost) independent Bernoulli trials, each of which is like playing a single \$1 scratcher ticket. In Play 4, you can bet any multiple of \$1 on a number to win a given drawing; if you win (which happens with probability $p = 10^{-4}$), then you win 5000 times your bet. If you bet \$100 on a single Play 4 draw, your odds of winning remain 10^{-4}, but your possible jackpot becomes \$500,000, and if you win, the Florida Lottery records this in the list of claimed prizes as if it were 100 separate wins. Clearly, these are wins on dependent bets.

So, to infer how much H had to spend on the lottery for his wins to be unsurprising, first we have to estimate how much he bet on each drawing. Unfortunately, we cannot deduce this from the list of claimed prizes, because it includes the date the prize was claimed but not the specific drawing the ticket was for. (Louis Johnson's Play 4 prizes were all claimed on distinct dates, so it is reasonable to assume they were bets on different draws.) The Palm Beach Post paid the Florida Lottery to retrieve a sample of H's winning tickets from their archives. We think H's winning plays were as in TABLE 1.

TABLE 1: H's Play 4 wins during 2011–2013

Date	Number played	Amount wagered
12/6/2011	6251	\$52
??	????	\$1
11/11/2012	4077	\$101
12/31/2012	1195	\$2
2/4/2013	1951	\$212
3/4/2013	1951	\$200

To find a lower bound on the amount H spent by solving the optimization problem (22), we imagine that he played several different Play 4 games, distinguished by their bet size. For simplicity, let us pretend that a player can bet \$1, \$50, \$100, or \$200, and suppose we observed H winning these bets 2, 1, 1, and 2 times, respectively. Using these as the parameters in (22) and the same probability cutoff $\varepsilon = 5 \times 10^{-14}$ gives a minimum amount spent of just \$96,354.

But we can find a number tied more closely to H's circumstances. In 2011–2013 he claimed \$2.84 million in prizes. These are subject to income tax. If his tax rate was about 35%, he would have taken home about \$1.85 million. If he spent that entire sum on Play 4 tickets, what is the probability that he would have won so much? We can find this by solving the following optimization problem with $p = 10^{-4}$, $\vec{w} = (2, 1, 1, 2)$, and $\vec{c} = (1, 50, 100, 200)$:

$$\max_{\vec{n}} \prod_{i=1}^{4} D(n_i; w_i, p) \quad \text{s.t.} \quad w_i \le n_i \quad \text{and} \quad \vec{c} \cdot \vec{n} \le 1.85 \times 10^6.$$

The solution is about 0.0016, or one-in-625: it is plausible that H was just lucky. That's because he made large, dependent bets, while we know from the examples above that betting a similar sum on smaller, independent bets is less likely to succeed.

This illustrates a principle of casino gambling from [**8**, p. 170] or [**18**, #37]: *bold play is better than cautious play*. If you are willing to risk \$100 betting red-black on a game of roulette, and you only care about doubling your money at the end of the evening, you are better off wagering \$100 on one spin and then stopping, rather than placing 100 \$1 bets.

The real world

How did this paper come to be? One of us, Lawrence Mower, is an investigative reporter in Palm Beach, Florida. His job is to find interesting news stories and spend 4–6 months investigating them. He wondered whether something might be going on with the Florida Lottery, so he obtained the list of prizes and contacted the other three of us to help analyze the data. Below we describe some of the nonmathematical aspects.

What some people get up to Various schemes can result in someone claiming many prizes.

Clerks at lottery retailers have been known to scratch the wax on a ticket lightly with a pin, revealing just enough of the barcode underneath to be able to scan it, as described in [**17**, paragraph 75]. If they scan it and it's not a winner, they'll sell it to a customer, who may not notice the very faint scratches on the card. Lottery operators in many states replaced the linear barcode with a 2-dimensional barcode to make this scam more difficult, but it still goes on: a California clerk was arrested for it on 9/25/14.

Sometimes gamblers will ask a clerk to check whether a ticket is a winner. If it is, the clerk might say it's a loser, or might say the ticket is worth less than it really is, then claim the prize at the lottery office—and become the recorded winner. Of course, most clerks are honest, but this scheme is popular; see, for example, [**17**, paragraphs 47, 48, 80, 146].

Another angle, *ticket aggregation*, goes as follows. A gambler who wins a prize of \$600 or more may be reluctant to claim the prize at the lottery office. The office might be far away; the gambler might be an illegal alien; or the gambler might owe child support or back taxes, which the lottery is required to subtract from the winnings. In such cases, the gambler might sell the winning ticket to a third party, an *aggregator*, who claims the prize and is recorded to be the winner. The aggregator pays the gambler less than face value, to cover income tax (paid by the aggregator) and to provide the aggregator a profit. The market rate in Florida is \$500–\$600 for a \$1000 ticket.

Some criminals have acted as aggregators to launder money. They pay the gambler in cash, but the lottery pays them with a check, "clean" money because it is already in the banking system. Notorious Boston mobster Whitey Bulger [**6**] and Spanish politician Carlos Fabra [**11**] are alleged to have used this dodge.

When questioned by Mower, some of our suspects confessed to aggregating tickets, which is a crime in Florida (Florida statute 24.101, paragraph 2).

Outcomes in Florida Before Mower's story appeared, he interviewed Florida Lottery Secretary Cynthia O'Connell about these gamblers. She answered that they could be lucky: "That's what the lottery is all about. You can buy one ticket and you become a millionaire" [**19**]. Our calculations show that for most of these 10 gamblers, this is an implausible claim. O'Connell and the Florida Lottery have since announced reforms to curb the activities highlighted here [**20**]. They stopped lottery operations at more than 30 stores across the state and seized the lottery terminals at those stores.

More news stories and outcomes in other states Further stories about "too frequent" winners have now appeared in California (KCBS Los Angeles 10/30/14, KPIX CBS San Francisco 10/31/14), Georgia (Atlanta Fox 5 News 9/12/14, Atlanta Journal-Constitution 9/18/14), Indiana (ABC 6 Indianapolis, 2/19/15), Iowa (The Gazette, 1/23/15), Kentucky (WLKY, 11/20/14), Massachusetts (Boston Globe, 7/20/14), Michigan (Lansing State Journal, 11/18/14), New Jersey (Asbury Park Press, 12/5/14 & 2/18/15; USA Today, 2/19/15), Ohio (Dayton Daily News 9/12/14), and Pennsylvania (CBS Philly 5/19/15). In Massachusetts, ticket aggregation is not illegal per se. In California, the lottery makes no effort to track frequent winners.

In Georgia, ticket aggregation is illegal but the law had not been enforced. The practice was so widespread that elementary calculations (much simpler than those presented in this article) cast suspicion on 125 people. This gap in enforcement, in principle easy to detect, came to light as a consequence of the much more challenging investigation in Florida described here. This led to a change in policy announced by the Georgia Lottery Director, Debbie Alford, on 9/18/14: "We believe that most of these cases involved retailers agreeing to cash winning tickets on behalf of their customers—a violation of law, rules, and regulations."

Acknowledgment We are grateful to Don Ylvisaker, Dmitry B. Karp, and an anonymous referee for helpful comments and insight. The second author's research was partially supported by NSF grant DMS-1201542.

REFERENCES

1. A. Abrams, S. Garibaldi, Finding good bets in the lottery, and why you shouldn't take them, *Amer. Math. Monthly* **117** (2010) 3–26.
2. R. E. Barlow, F. Proschan, *Statistical Theory of Reliability and Life Testing: Probability Models.* International Series in Decision Processes. Holt, Rinehart and Winston, 1975.
3. J. van den Berg, U. Fiebig, On a combinatorial conjecture concerning disjoint occurrences of events, *Ann. Probab.* **15** no. 1 (1987) 354–374.
4. J. van den Berg, H. Kesten, Inequalities with applications to percolation and reliability, *J. Appl. Prob.* **22** (1985) 556–569.
5. J. T. Chayes, A. L. Puha, T. Sweet, Independent and dependent percolation, in *Probability Theory and Applications*. Vol. 6. IAS/Park City Math. Ser. Amer. Math. Soc., Providence, RI, 1999.
6. K. Cullen, US orders lottery to hold Bulger's winnings, *Boston Globe*, July 18, 1995, http://www.boston.com/news/packages/whitey/globe_stories/1995/0718_us_orders_lottery_to_hold_bulger_s_winnings.htm.
7. P. Diaconis, F. Mosteller, Methods for studying coincidences, *J. Amer. Stat. Assoc.* **84** (1989) 853–861.
8. L. E. Dubins, L. J. Savage, *How to Gamble If You Must: Inequalities for Stochastic Processes.* McGraw-Hill, New York, 1965.
9. R. Durrett, *Essentials of Stochastic Processes.* Springer Texts in Statistics. Second edition. Springer, New York, 2013.
10. S. N. Ethier, D. Khoshnevisan, Bounds on gambler's ruin probabilities in terms of moments, *Methodol. Comput. Appl. Probab.* **4** (2002) 55–68.

11. M. Fabra, The law finally catches up with former Castellón cacique Fabra, *El País in English* October 2, 2013, http://elpais.com/elpais/2013/10/02/inenglish/1380734613_504452.html.
12. H. Finner, M. Roters, Log-concavity and inequalities for chi-square, F and Beta distributions with applications in multiple comparisons, *Statist. Sinica* **7** (1997) 771–787.
13. C. M. Grinstead, W. P. Peterson, J. L. Snell, *Probability Tales*. Vol. 57. Student Mathematical Library. American Mathematical Society, Washington, DC, 2011.
14. K. R. Grote, V. A. Matheson, In search of a fair bet in the lottery, *Eastern Econom. J.* **32** (2006) 673–684.
15. P. W. Hodge, *Galaxies.* Harvard Univ. Press, Cambridge, MA, 1986.
16. L. J. Kaplan, J. M. Maher, The economics of the numbers game, *Amer. J. Econom. Sociol.* **29** no. 4 (1970) 391–408.
17. A. Marin, A game of trust: Investigation into the Ontario Lottery and Gaming Corporation's protection of the public from fraud and theft. Ombudsman Report, March 2007, http://www.ombudsman.on.ca/Investigations/SORT-Investigations/Completed/Lotteries---em-A-Game-of-Trust--em-.aspx?lang=en-CA.
18. F. Mosteller, *Fifty Challenging Problems in Probability with Solutions.* Reprint of the 1965 Addision-Wesley edition. Dover, New York, 1987.
19. L. Mower, Gaming the lottery. *The Palm Beach Post* March 2014, http://www.mypalmbeachpost.com/gaming-the-lottery/.
20. C. O'Connell, Point of view: Lottery working to improve security, fight fraud. *The Palm Beach Post* April 2014, Opinion Page, http://www.palmbeachpost.com/news/news/opinion/point-of-view-lottery-working-to-improve-security-/nfQmq/.
21. D. Reimer, Proof of the van den Berg–Kesten Conjecture, *Combin. Probab. Comput.* **9** no. 1 (2000) 27–32.
22. T. Sellin, Organized crime: A business enterprise, *Ann. Amer. Acad. Political Social Sci.* **347** (1963) 12–19.
23. L. A. Stefanski, The North Carolina lottery coincidence, *Amer. Stat.* **62** no. 2 (2008) 130–134, http://www.doi.org/10.1198/000313008X306844.
24. Zillow.com, "Florida home prices and home values." January 2014, www.zillow.com/local-info/FL-home-value/r

Summary. We looked at the Florida Lottery records of winners of prizes worth $600 or more. Some individuals claimed large numbers of prizes. Were they lucky, or up to something? We distinguished the "plausibly lucky" from the "implausibly lucky" by solving optimization problems that took into account the particular games each gambler won. Plausibility was determined by finding the minimum expenditure so that if every Florida resident spent that much, the chance that any of them would win as often as the gambler did would still be less than one in a million. Dealing with dependent bets relied on the BKR inequality; solving the optimization problem numerically relied on the log-concavity of the regularized Beta function. Subsequent investigation by law enforcement confirmed that the gamblers we identified as "implausibly lucky" were indeed behaving illegally.

RICHARD ARRATIA (MR Author ID: 27395) is a professor of mathematics at the University of Southern California. He received his BS from MIT and Ph.D. from UW Madison. His highest career honor, so far, was captaining the Stuyvesant High School math team, back in 1968, and that year of course Stuyvesant won the city, NYC. His research focus is probability, especially the boundary between dependence and independence, in regard to number theory and combinatorics. He coauthored the book "Logarithmic Combinatorial Structures, a Probabilistic Approach" with Andrew Barbour and Simon Tavaré. His hobbies are playing table tennis, swimming without a wetsuit in the open ocean, playing wind instruments, including tuba, and tending fruit trees, all of which he does like a fish.

SKIP GARIBALDI (MR Author ID: 622970) is associate director of UCLA's Institute for Pure and Applied Mathematics and a professor in Emory University's Department of Mathematics & Computer Science. His research is primarily on algebraic groups—including co-authoring the book "Cohomological invariants in Galois cohomology" with Alexander Merkurjev and Jean-Pierre Serre—but evidently he also works on more applied subjects. His previous work on the lottery was the subject of a chapter in the popular book "Brain Trust" and spurred interviews on 20/20, CNN, and Fox & Friends. He is featured in a museum exhibit about mathematics that has been touring US science museums since 2012, and he was recently appointed to the Air Force Scientific Advisory Board.

LAWRENCE MOWER (MR Author ID: 1104768) is an investigative reporter with The Palm Beach Post. He joined The Post in 2013, after working for the Las Vegas Review-Journal, where his yearlong investigation into Las Vegas police shootings sparked a Department of Justice investigation and led to reforms in policy and oversight. The five-part series was awarded by the National Headliner Awards, Investigative Reporters and Editors, and the ACLU of Nevada, and in 2012 he was named Nevada's Outstanding Journalist by the Nevada Press Association. He is a 2006 graduate of the University of Nevada, Las Vegas.

PHILIP B. STARK (MR Author ID: 266672) is professor and chair of the Department of Statistics at UC Berkeley. He received his AB from Princeton University and his Ph.D. from UCSD. He was a Presidential Young Investigator and received the Chancellor's Award for Research in the Public Interest. His current research includes causal inference, cosmology, ecology, election integrity, food equity and sustainability, online education, risk assessment, seismology, teaching assessment, and uncertainty quantification. He has consulted for many firms and government agencies, including the U.S. Departments of Justice, Agriculture, Commerce, Housing and Urban Development, and Veterans Affairs.

MAA 100
MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

From the Files of Past MAGAZINE Editors

Allen Schwenk 2006–2008

Perhaps the most memorable interaction former MAGAZINE editor Allen Schwenk had with a mathematical crank was with a retired physics professor, who expressed interest in number theory, and mentioned problems and number theorists that convinced Schwenk that while the author may be a mathematical amateur, surely he was well read. After being encouraged to submit his manuscript, Schwenk received 40 separate messages from him the next morning, each one an intended submission. The most intriguing title: "A Short Proof of Fermat's Last Theorem." The proof is a classic. Here it is. We know that $x^2 + y^2 = z^2$ has many solutions in positive integers, and we may assume, without loss of generality, that $1 \le x < y < z$. Now for any $n \le 3$, multiply by y^{n-2} to get $x^2 y^{n-2} + y^n = z^2 y^{n-2}$.

Since $x^{n-2} < y^{n-2} < z^{n-2}$, we deduce that $x^n + y^n = x^2 x^{n-2} + y^n < x^2 y^{n-2} + y^n = z^2 y^{n-2} < z^2 z^{n-2} = z^n$. A sample of two or three other files proved to be of similar quality.

ACROSS

1. Remove, as a hat
5. Head-and-shoulders sculpture
9. Past MAA president and MAGAZINE editor Paul M.
13. Affirm as fact
14. Foot part
15. Quickly pan-fry
16. Narrative
17. Early whirlybird, for short
18. Leggy, white wader
19. Past MAA President Carl B. with an eponymous award honoring articles in this MAGAZINE
22. Daniel of the old frontier
23. Handy
27. Past CMJ editor Donald J. who co-edited *Mathematical People* with 53-Across
30. Gear's tooth
32. Unclothed
33. Column Like You ___ (Plants vs. Zombies mini-game)
34. Shoulder gesture
36. Schrödinger's ___
37. First MAA President Earle R.
39. J. Arthur who co-edited this MAGAZINE with Steen
41. California's historic Fort ___
42. Sappho's poetic muse
44. Olympic gymnast Comăneci
45. Oboist's accessory
47. Theater legend Hagen
48. Past MAA president also known as John Taine
49. Swirled liquidly
51. Embodying machismo
53. Past MAA President and MAGAZINE editor Gerald L. who also presided over the Fibonacci Association
58. Common sense?
61. Antlered buglers
62. Humerus neighbor
63. "Gay ___" (*Victor/Victoria* song)
64. Classic grape soda favored by Radar of *M*A*S*H*
65. Two π's?
66. Past MAA President and MONTHLY editor Lester R.
67. Office helper: Abbr.
68. "-": Abbr.

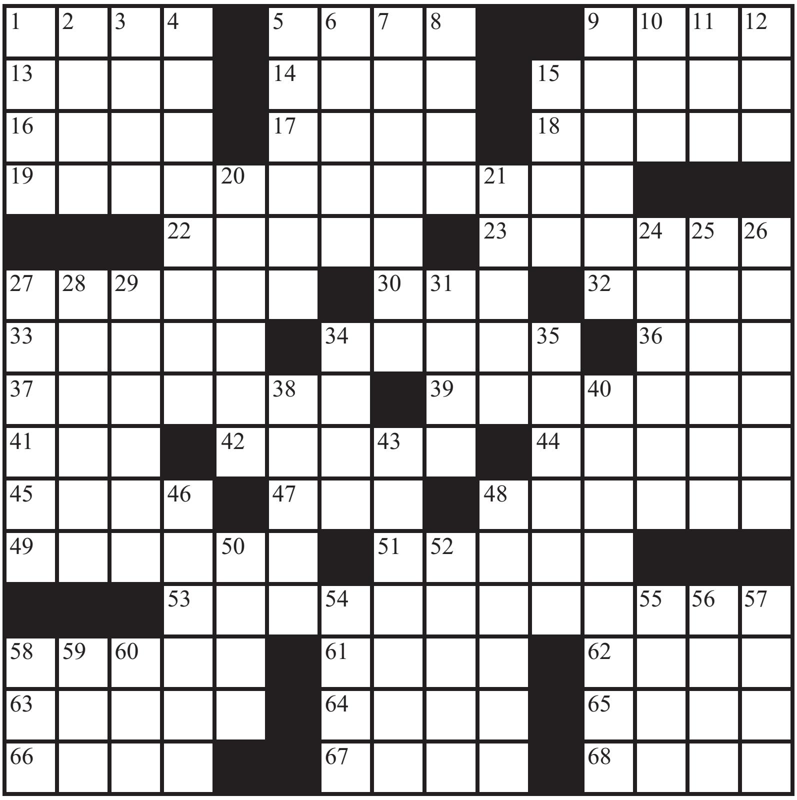
DOWN

1. What statisticians crunch
2. Flattened circle
3. Tumbled down
4. Offer a party animal can't resist
5. Bouffants and beehives, say
6. $\cup$
7. ___ to a halt
8. Norse thunder god
9. Croatia's capital
10. ___ *Town* (Wilder play)
11. GPS display
12. Neural ___
15. Understands
20. Cliff's nickname for Peterson on *Cheers*
21. Bach composition that may embody the golden ratio
24. False front
25. Adenine, guanine, cytosine and ___
26. Deadly
27. No longer at sea
28. Ogled offensively
29. ___ down for the night (tucked in)
31. ". . . ___ I've been told"
34. Old three-player card game
35. Like a little lamb, say
38. Unsophisticated
40. Candy bar ostensibly not named for legendary slugger
43. Mexican fare steamed in corn husks
46. Used a rotary phone
48. What $\Box$ does to a proof
50. Vogue competitor
52. "Breath of life" cross symbols of Egypt
54. Warrior princess of 1990s TV
55. Do in, like a dragon killer
56. "Movin' ___ " (*The Jeffersons* theme song)
57. Game theorist depicted in *A Beautiful Mind*
58. UV-blocking letters
59. "Women hold up half the sky" proclaimer
60. Misspell someone's name, say

Math. Mag. **88** (2015) 212–213. doi:10.4169/math.mag.88.3.212.

Editors Past and President, Part I

TRACY BENNETT
Mathematical Reviews
Ann Arbor, MI
tgp@ams.org



Clues start at left, on page 212. The Solution is on page 185.

Extra copies of the puzzle can be found at the Magazine's website, www.maa.org/mathmag/supplements.

MAA 100
MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

From the Files of Past MAGAZINE Editors

Allen Schwenk 2006–2008

Former MAGAZINE editor Allen Schwenk remembers when an author sent numerous proofs of "the true value of π." He sent several versions terminating with a 228-page treatise. In case you are wondering, the true value seems to be $\pi = \frac{14-\sqrt{2}}{4}$.

Pythagoras via Cavalieri

ZSOLT LENGVÁRSZKY
Louisiana State University in Shreveport
Shreveport, LA
Zsolt.Lengvarszky@lsus.edu

Some proofs of the Pythagorean theorem utilize more advanced, seemingly unrelated results. One example (see C. Alsina and R. B. Nelsen [**1**]) is a proof based on Heron's theorem which states that the area of a triangle is given by the formula $A = \sqrt{s(s-a)(s-b)(s-c)}$, where a, b, c are the side lengths of the triangle and $s = (a+b+c)/2$. Heron's theorem is commonly proved with the help of the Pythagorean theorem, however this is not necessary. Hence, it is legitimate, and not circular, to use Heron to prove the Pythagorean theorem. Recent proofs of the Pythagorean theorem have applied infinite geometric series (see [**3**]) and differentials (see [**2, 5**]). These proofs are also noncircular because the results employed can be proved independently of the Pythagorean theorem.

In this note we present a proof that is based on the well-known principle named after the Italian mathematician Bonaventura Cavalieri (1598?–1642). Cavalieri's principle is discussed in most calculus texts, and most often in a three-dimensional setting in order to find volumes of solids. However, it applies to areas in the plane as well. Cavalieri's principle states that if parallel planes (lines) give equal cross-sectional areas (lengths), then the volumes (areas) are equal (see G. Zill et al. [**6**], for example). A more general version of the principle is obtained by replacing equality in the statement by a given ratio to yield proportional volumes (areas).

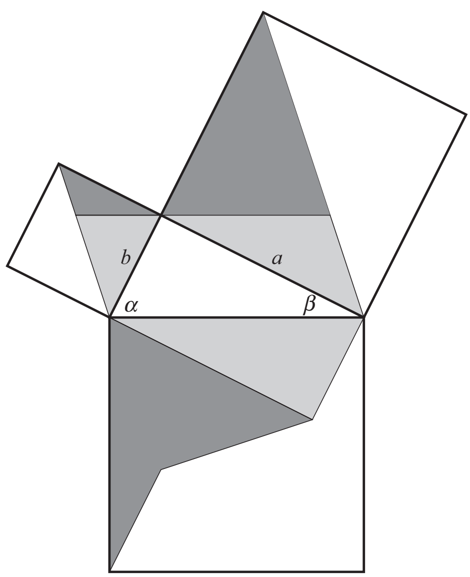


Figure 1 Proof by Cavalieri's principle.

FIGURE 1 shows the proof of the Pythagorean theorem using Cavalieri's principle. Notice that half of each of the three squares erected on the sides of the right triangle

Math. Mag. **88** (2015) 214–216. doi:10.4169/math.mag.88.3.214. © Mathematical Association of America

are shaded in the diagram. This is achieved by drawing diagonals on the squares on the legs of the triangle, and drawing a "zigzag" including two translations of the shorter leg b of $\triangle ABC$ for the square on the hypotenuse of the triangle. Furthermore, we divide the regions corresponding to $a^2/2 + b^2/2$ and $c^2/2$ into two parts, one with lighter shading and one with darker shading, as shown in FIGURE 1.

At this point we are ready to apply Cavalieri's principle: Draw a horizontal line across each of the two regions with lighter shading, and note that the (total) cross-sectional lengths are the same if applied at the same levels. Indeed, for both regions, the cross-sectional lengths change from c to 0 as linear functions of height, which in turn has the same range of values (0 to $a \sin \beta$, the altitude of the $\triangle ABC$) in both cases. Next consider cross sections of the regions with darker shading; horizontal ones for $a^2/2 + b^2/2$, and vertical cross sections for $c^2/2$. In this case the heights (distances of the cross sections from the horizontal and vertical sides, respectively) vary from 0 to $a \sin \alpha$, and the (total) cross-sectional lengths decrease from c to 0 as piecewise linear functions of height for both regions. The change in rate occurs at the same time (when the height is $b \sin \beta$), and both rates are the same for both regions due to the fact that the two angles at the vertical side of the dark quadrilateral are congruent, respectively, to the two angles of the dark triangles at their common vertex. Again, we find that the cross-sectional lengths are the same, hence, the similarly shaded regions are equal in area, and $a^2/2 + b^2/2 = c^2/2$ follows.

The novelty of our proof lies in the application of Cavalieri's principle. Note that the diagram in FIGURE 1 easily transforms into a dissection proof: just slide over the isosceles right triangle on the left hand side as shown in FIGURE 2; the regions with lighter/darker shade are now translations/rotations of each other. This dissection proof of the Pythagorean theorem is a variation of the well-known proof attributed to J. E. Böttcher in Loomis' collection of 367 proofs of the Pythagorean theorem (see [**4**]).

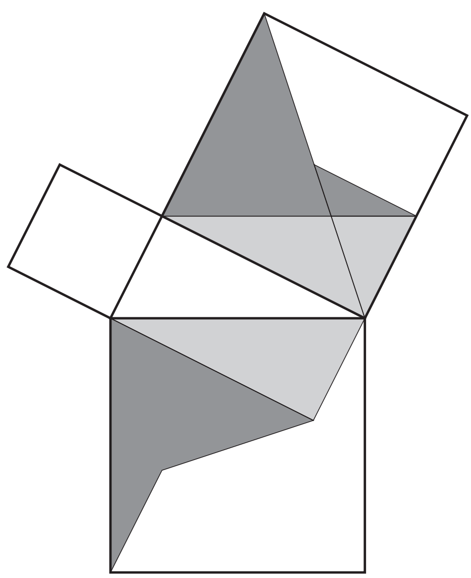


Figure 2 Dissection proof.

REFERENCES

1. C. Alsina, R. B. Nelsen, *Charming Proofs, A Journey Into Elegant Mathematics*, MAA Dolciani Mathematical Expositions, Exercise 5.14. Mathematical Association of America, Washington, DC, 2010.
2. M. Hardy, Pythagoras made difficult, *Math. Intelligencer* **10** no. 3 (1988) 31.
3. Z. Lengvárszky, Proving the Pythagorean theorem via infinite dissections, *Amer. Math. Monthly* **120** no. 8 (2013) 751–753.
4. E. S. Loomis, *The Pythagorean Proposition*. National Council of Teacher's of Mathematics, Washington, DC, 1968.
5. M. Staring, The Pythagorean proposition: A proof by means of calculus, *Math. Mag.* **69** no. 1 (1996) 45–46.
6. D. G. Zill, S. Wright, W. S. Wright, *Calculus: Early Transcendentals.* Third edition. Jones and Bartlett Learning, Burlington, MA, 2009, p. xxvii.

Summary. We present a proof of the Pythagorean theorem via Cavalieri's principle.

ZSOLT LENGVÁRSZKY (MR Author ID: 112490) studied at the University of Szeged, Hungary, and at the University of South Carolina. He has been at Louisiana State University in Shreveport since 2008. His hobbies include reading and listening to music.

MAA 100
MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

From the Files of Past MAGAZINE Editors

J. Arthur Seebach and Lynn Arthur Steen 1976–1980

Former MAGAZINE co-editor Lynn Steen introduced a typewritten letter to the editor column, allowing comments to get into the same issue of the MAGAZINE often with the article they were commenting on. Mary Kay Peterson produced typed, camera-ready copy just before the MAGAZINE went to press.

When Steen heard from Ron Graham about the four-color theorem, Steen was able to get an article about the computer-assisted proof into the MAGAZINE, long before other journals were able to comment on it. See L. Steen, Solution of the four color problem in News and Letters, *Math. Mag.* **52** no. 3 (1976) 219–222.

Optimal Defensive Strategies in One-Dimensional *RISK*

DARREN GLASS
Gettysburg College
300 N. Washington St.
Gettysburg, PA 17325
dglass@gettysburg.edu

TODD NELLER
Dept of Computer Science
Gettysburg College
300 N. Washington St.
Gettysburg, PA 17325
tneller@gettysburg.edu

RISK is a popular board game invented by Albert Lamorisse and released in 1957. The board depicts a stylized political world map divided into territories, each occupied by one or more of a player's army units, which we will refer to simply as armies. The bulk of play consists of a turn-based series of attacks between player armies occupying adjacent territories in an effort to occupy the entire world. In the late game, it is common for one player to attempt to eliminate another player along a chain of territories.

In this note, we consider the problem of how a defensive player should distribute his armies to maximize the probability of survival. In particular, we will consider a one-dimensional version of the game, which takes place on a chain of $m + 1$ consecutive territories, as depicted in FIGURE 1. We now describe the rules of our version of the game. Experienced *RISK* players will note that this is a significantly simplified version of the game, but we believe that we have captured most of the spirit of the original game. At the end of the paper, we will discuss some ways in which the differences in the actual game of *RISK* might affect our proposed strategies.

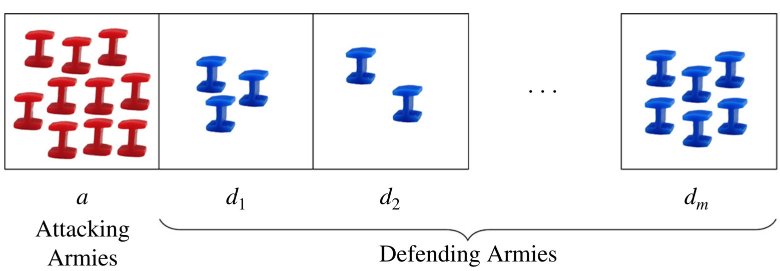


Figure 1 One-dimensional *RISK* board

At the beginning of the game we assume that the attacker has a positive number of armies (labeled a in FIGURE 1) on a territory at one end and the defender has a number of armies (labeled $d_1, \ldots, d_m$) distributed among the other m territories so that there is at least one army per territory. The goal of the attacker is to take over all m of the defender's territories while the goal of the defender is to prevent this from happening.

Math. Mag. **88** (2015) 217–230. doi:10.4169/math.mag.88.3.217. © Mathematical Association of America
MSC: Primary 00A08, 11B37, 60J10

The attacker begins by rolling a number of dice that is the lesser of 3 and the number of attacking armies minus 1. The defender then simultaneously rolls a number of dice that is the lesser of 2 and the number of defending armies. (We note that in the actual game of *RISK* each player may choose to roll a smaller number of dice than allowed. We assume that they always choose to roll the maximum number allowed.)

The dice rolled by attacker and defender are then sorted and compared. The maximum attacking die roll is compared to the maximum defending die roll. If a second attacker–defender die pair was rolled, the second-to-maximum die rolls are compared as well. For each attacker–defender die pair, the attacker wins if and only if the attacking die roll is greater than the defending die roll, i.e., the maximum roll wins with ties going to the defender. For each loss, the attacker or defender removes a single army from the relevant territory. Thus, the result of a roll of three attacking dice versus two attacking dice can result $(a_{\text{loss}}, d_{\text{loss}}) \in \{(0, 2), (1, 1), (2, 0)\}$, where a_{loss} and d_{loss} are the number of armies lost for the attacker and defender, respectively.

As long as the attacker has two or more armies remaining and the defender has any armies remaining, we assume that the attacker will repeat the attack roll process. If the attacker drops below two armies then we declare the defending army the winner. If the defending armies are successfully eliminated, the territory is captured, and the attacker must occupy the territory. We assume that the attacker leaves precisely one army behind on her territory and moves the remaining armies onto the defender's territory. Assuming she has moved at least two armies, she continues to attack the next territory in the defender's chain. The process repeats until the attacking player either drops below two armies on her leading territory or she captures all of the territories in the defending player's chain.

This version of the game captures the essence of the situation in the actual game of *RISK* in which players have large numbers of army reinforcements (i.e., from occupation of entire continents) or are playing *RISK* variations where one can obtain a large number of armies by trading in card sets. In those situations, the maximum number of armies is often moved into a territory after one successful capture. The attacker then immediately seeks to capture a territory adjacent to the captured territory. This process often has the goal to occupy a continent or possibly eliminate an opponent from the game. It is thus of great interest to understand how the defender army distribution affects the probability of defender survival.

In the spirit of two earlier articles in THIS MAGAZINE, we implement a Markov chain model of the game of *RISK*. In [**4**], Tan develops such a model in order to answer the question of when it is worthwhile for a player to attack an adjacent territory and what the expected damage to the attacker in such a battle will be. Several years later, Osborne found that Tan's model made overly strong assumptions of independence, and in [**2**], he corrects those assumptions and addresses additional questions of strategy under his corrected Markov chain model. However, he continues to only look at the strategy in situations where one territory attacks another, while we look at the more general question of how to distribute armies among multiple territories in order to optimize one's chances of survival. Looking at the numerical results of this model, as seen below, we have formulated the following conjecture.

Conjecture. *In order to best survive an attack on a chain of m territories with d armies where $d \geq 2m$ and the number of attacking armies is sufficiently large, the best defense is to place two armies on each of the first $m - 1$ territories and the remaining armies on the last territory.*

Pieces of this conjecture have been found in the folklore of the game of *RISK*, and in particular it is asserted without justification in [**1**] that one wants to defend with an even number of armies whenever possible. We note that this conjecture is not true if the number of attacking armies is small: In particular, we will show that a different strategy is optimal if the attacker only has $m + 1$ armies, which is the smallest number that can be used in a campaign against m territories with any hope of success. While we are unable to prove the full conjecture, we will prove the following theorem.

Theorem 1. *In order to best survive an attack on a chain of m territories with n armies, the best defense will be to place either one or two armies on each of the first $m - 1$ territories and the remaining armies on the last territory.*

We will also consider a variant on our question in which it is assumed that the attacker has an overwhelmingly large number of armies and is therefore always rolling three dice. Under this assumption, we will consider how the defender should place her armies so as to cause the most damage to the attacker, even though he will eventually lose all of his armies. While this question is not quite the same situation as we would like to consider in the game of *RISK*, it shares many similarities and the fact that the best strategy is the same as in the conjecture gives us some evidence that the conjecture is correct. A final section discusses other considerations in the actual gameplay of *RISK* that may change the way a player would apply our results in practice.

Our numerical experiments and theoretical computations rely on the following probabilities, computed by Osborne in [**2**]. Suppose that for a given battle the defending player has one army on her territory and the attacking player rolls three dice. It is an elementary (if tedious) probability calculation to see that the probability that the value of the defender's roll is at least as high as the highest of the attacker's roll is $\frac{441}{1296}$, and we denote this probability by $\widetilde{p}_0$. Similarly, the probability that the defender loses the one army in this battle is $\frac{855}{1296}$, which we denote by $\widetilde{p}_1$. Osborne further computes the probabilities of all possible outcomes of a given set of die rolls, which we give in FIGURE 2. In our notation, the letter is dependent on the number of dice rolled by the attacker, with p meaning three dice, q meaning two dice, and r meaning a single die. The presence or lack of a tilde depends on whether the defender is rolling one or two dice, respectively, and the subscript is the number of armies lost by the defender.

Att. Dice	Def. Dice	Def. Loss	Att. Loss	Prob.	Notation
3	2	2	0	$\frac{2890}{7776}$	p_2
3	2	1	1	$\frac{2611}{7776}$	p_1
3	2	0	2	$\frac{2275}{7776}$	p_0
3	1	1	0	$\frac{855}{1296}$	$\widetilde{p}_1$
3	1	0	1	$\frac{441}{1296}$	$\widetilde{p}_0$
2	2	2	0	$\frac{295}{1296}$	q_2
2	2	1	1	$\frac{420}{1296}$	q_1
2	2	0	2	$\frac{581}{1296}$	q_0
2	1	1	0	$\frac{125}{216}$	$\widetilde{q}_1$
2	1	0	1	$\frac{91}{216}$	$\widetilde{q}_0$
1	2	1	0	$\frac{55}{216}$	r_1
1	2	0	1	$\frac{161}{216}$	r_0
1	1	1	0	$\frac{15}{36}$	$\widetilde{r}_1$
1	1	0	1	$\frac{21}{36}$	$\widetilde{r}_0$

Figure 2 Probabilities for different outcomes based on the number of dice rolled

Empirical observations

In this section, we use a Markov chain model of the game of *RISK* in order to gather some data that we then use to formulate our conjecture. Recall that our version of *RISK* assumes that the attacker armies begin amassed on a single territory adjacent to a chain of m territories occupied by the defending player. We begin by simulating the battle between the attacker and the defender in the first territory in the defender's chain. If the attacker wins this battle, we assume that she maximally occupies the territory, leaving behind a single occupying army. She proceeds to attack the defender's second territory with the remaining armies, continuing as she wins each additional territory.

We have performed brute-force computations to gain a sense of what distribution of the defender's armies will lead to the highest probability that he survives the full attack. For a given number of total armies, we consider all nontrivial divisions between the number of attacker armies a and defender armies d. In this context, "nontrivial" means that the defender has more than one possible distribution to consider, and the optimal probability for survival is neither 0 nor 1. For each distribution of defending armies, we computed the probability that the defending player had armies remaining at the end of the battle.

To get a flavor of these experiments, consider the case where an attacker with 30 armies seeks to eliminate a defender with 30 armies distributed along an adjacent chain of five defender territories. If we distribute the defending armies uniformly along the first four chain territories and place the remainder on the last chain territory, we can see that all even-numbered uniform army distributions yield greater survival probabilities than each odd-numbered uniform army distribution, as computed from our Markov chain model and described in FIGURE 3.

Defender Distribution					Survival Probability
1	1	1	1	26	0.4986
2	2	2	2	22	0.5367
3	3	3	3	18	0.5165
4	4	4	4	14	0.5271
5	5	5	5	10	0.5216
6	6	6	6	6	0.5242

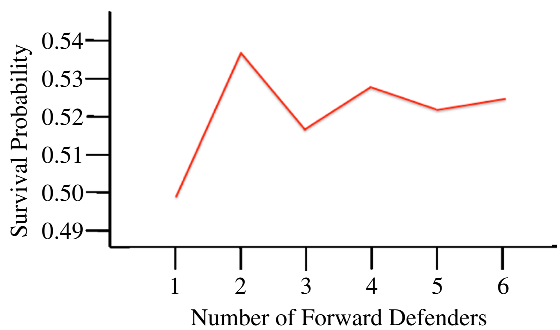


Figure 3 Scenarios where defender distributes armies uniformly

Given a number of defending armies, we consider each length m of defender territory chains that lead to legal, nontrivial distributions. Each territory must be occupied, so this amounts to the condition $2 \leq m \leq d - 1$. Given a, d, and m, we then consider all possible distributions of defender armies, compute the survival probability of each distribution, and test hypotheses on the optimal distributions. We note that when d is close to m, then the defender does not have much flexibility in distributing the armies, and in this situation we will say that the defender is "highly constrained." Similarly, when a is close to m, we will say that the attacker is "highly constrained."

The results of this experiment are contained in FIGURE 4 with axes a, d, and m. Each glyph represents a class of optimal distribution for a given a, d, and m. In considering all possible scenarios up to 46 total armies, we observe several patterns that were the basis for our conjecture and theorems.

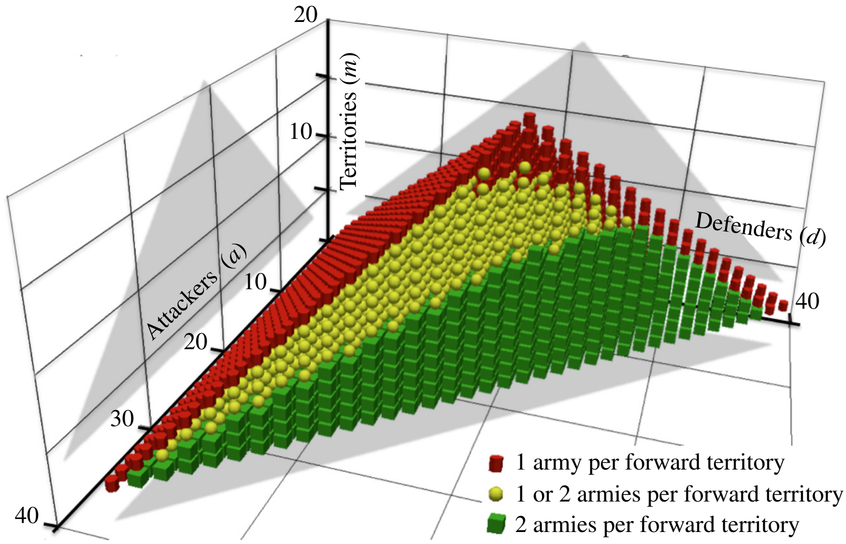


Figure 4 Optimal distributions

- In all cases, the optimal strategy consists of placing either one or two armies on all but the final territory. We prove that this is the case as Theorem 3.
- When $a \leq m + 2$, the defender's optimal army distribution has a minimal forward defense of one army per territory with all remaining armies on the last territory. We call this the *one-army strategy* and Theorem 2 shows that it is optimal in the case where $a = m + 1$.
- When the attacker has a sufficiently large force, the defender's optimal army distribution has a forward defense of two armies per territory with all remaining armies on the last territory, as long as the defender has enough armies to use this distribution (i.e., $d \geq 2m$). We refer to this as the *two-army strategy* and while we have not been able to find sharp bounds for what we mean by "sufficiently large," the number appears to be bounded below by the total number of territories plus a linear function of the number of attacker armies.

To see situations in the "in-between" cases where the defender uses a mix of the one- and two-army defense, we let $e = a - m$ be the number of attacker armies in excess of the number of defender territories. For each $4 \leq e \leq 10$, FIGURE 5 shows the smallest number of total armies for which an optimal distribution calls for fewer than two armies per forward space when it is possible to defend with two armies per forward space. In each case, note that the number of attacking armies is slightly less than half the total number of armies.

We observe that one-army distributions occur only in situations in which the defender is highly constrained. Usually, either the attacker must have just enough armies to occupy all territories given perfect attacking rolls, or the defender has one or two armies to distribute beyond the minimum one-army per space for occupation. Exceptions occur when the number of attacking armies and defending armies are nearly equal and both players are highly constrained. All cases of optimal one-army distributions are highly constrained for at least one of the players.

We next observe that in most other cases where the defender can defend with two armies per forward territory, it is optimal to do so. Most mixed distribution cases are out of necessity; the defender hasn't enough armies to defend with two per territory.

Defender Territories (m)	Attacker Armies (a)	Defender Armies (d)	$e = a - m$	Optimal Distribution
4	8	9	4	2, 1, 2, 4
5	10	11	5	2, 1, 2, 2, 4
6	12	13	6	2, 2, 1, 2, 2, 4
8	15	17	7	2, 2, 2, 2, 1, 2, 2, 4
9	17	19	8	1, 2, 2, 2, 2, 2, 2, 2, 4
10	19	21	9	2, 1, 2, 2, 2, 2, 2, 2, 2, 4
11	21	23	10	2, 2, 1, 2, 2, 2, 2, 2, 2, 2, 4

Figure 5 "Mixed strategy" distributions that are optimal

We also note that the division between mixed and two-army distributions lies almost exactly along the plane defined by $d = 2m$, with exceptions occurring only in cases where the attacker is highly constrained. Most importantly, we note that if neither player is highly constrained, and the defender can defend with a two-army distribution, it is usually optimal to do so.

These observations led us to make the conjecture in the introduction, and related results are discussed in the final sections of this article. FIGURES 6, 7, and 8 depict several cross sections of the three-dimensional array in FIGURE 4 that we found helpful in understanding the situations in which various strategies were optimal.

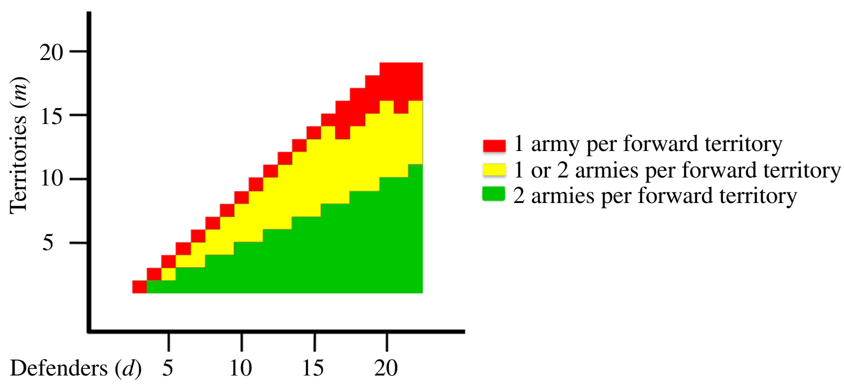


Figure 6 Cross section where $a = 20$

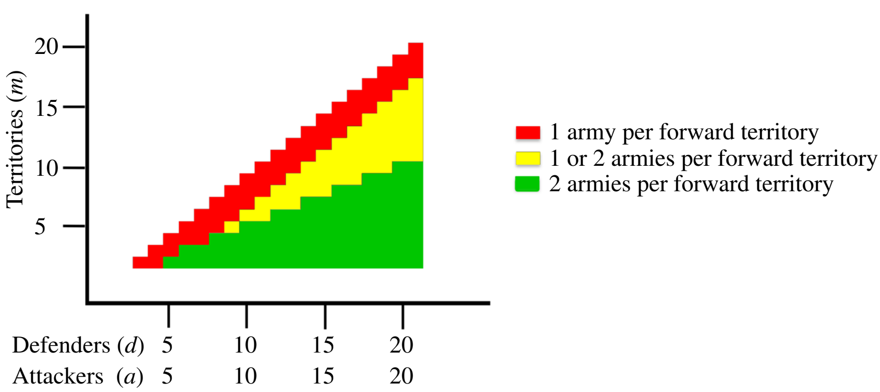


Figure 7 Cross section where $a = d$

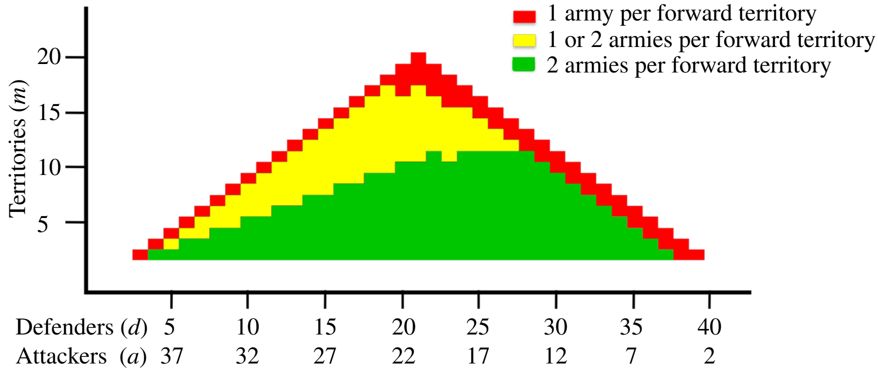


Figure 8 Cross section where $a + d = 42$

Holding off a limited number of attacking armies

In this section, we consider strategies based on a fixed number a of attacking armies and a chain of length m compared to a.

Defending $a - 1$ territories We begin by considering the case where $m = a - 1$ and therefore the attacker is highly constrained. In particular, in this case, we are able to prove the following theorem giving an explicit optimal strategy.

Theorem 2. *In order to best survive an attack on a chain of m territories where the number of attacking armies is $m + 1$, the best defense is to place one army on each of the first $m - 1$ territories and the remaining armies on the last territory.*

We note that if the attacker is successful, then she will need all $m + 1$ of her armies in order to occupy the $m + 1$ total spaces involved. In particular, the defending player will be successful in defending his chain if and only if he manages to defeat a single attacking army during the campaign.

If $m \geq 3$, then in the battle over the first territory, the attacking army will have at least four armies and therefore will be able to roll the full three dice. In particular, the probability that the attacker will defeat k armies without losing a single one of her own armies is given by $p_2^{\lfloor \frac{k}{2} \rfloor} \widetilde{p_1}^{\overline{k}}$ where $\overline{k} = 0$ if k is even and $\overline{k} = 1$ if k is odd. This same formula will hold in each of the first $m - 2$ territories. For the battle over the $(m - 1)$st territory, the formula is similar, only now the attacker is only allowed to roll two dice, so it becomes $q_2^{\lfloor \frac{k}{2} \rfloor} \widetilde{q_1}^{\overline{k}}$. For the final territory, the attacker is only allowed to roll one die, and therefore the probability she is successful in a battle against k armies is $r_1^{k-1}\widetilde{r_1}$.

Putting this all together, we see that if the defending player distributes his d armies so that there are d_i armies on the i^{th} territory for $i = 1, \ldots, m - 1$ and the remaining $d_m = d - \sum d_i$ armies on the final territory, then the probability that the defender will lose all of his territories is given by the following expression:

$$F(d_1, d_2, \ldots, d_{m-1}) = p_2^{\lfloor \frac{d_1}{2} \rfloor} \widetilde{p_1}^{\overline{d_1}} \cdots p_2^{\lfloor \frac{d_{m-2}}{2} \rfloor} \widetilde{p_1}^{\overline{d_{m-2}}} q_2^{\lfloor \frac{d_{m-1}}{2} \rfloor} \widetilde{q_1}^{\overline{d_{m-1}}} r_1^{d-\sum d_i - 1} \widetilde{r_1}.$$

In particular, we note that increasing d_1 by 1 leads to the following identity:

$$F(d_1 + 1, d_2, \ldots, d_{m-1}) = \begin{cases} \frac{p_2}{\widetilde{p_1} r_1} F(d_1, d_2, \ldots, d_{m-1}) \approx 2.21 F(d_1, d_2, \ldots, d_{m-1}) & \text{if } d_1 \text{ odd} \\ \frac{\widetilde{p_1}}{r_1} F(d_1, d_2, \ldots, d_{m-1}) \approx 2.59 F(d_1, d_2, \ldots, d_{n-1}) & \text{if } d_1 \text{ even.} \end{cases}$$

Increasing d_1 by one will always lead to decreasing the defender's probability of success, and therefore the defender should choose d_1 to be as small as possible. Specifically, he should place a single army on the first territory. The exact same computation holds for territories 2 through $m-2$. To consider the $(m-1)$st territory, we see from our formula that

$$F(d_1, d_2, \ldots, d_{m-1}+1) = \begin{cases} \frac{q_2}{\widetilde{q}_1 r_1} F(d_1, d_2, \ldots, d_{m-1}) \approx 1.54 F(d_1, d_2, \ldots, d_{m-1}) & \text{if } m \text{ is odd} \\ \frac{\widetilde{q}_1}{r_1} F(d_1, d_2, \ldots, d_{m-1}) \approx 2.27 F(d_1, d_2, \ldots, d_{m-1}) & \text{if } m \text{ is even,} \end{cases}$$

and again we see that the defender should choose d_{m-1} to be as small as possible.

Defending $a-2$ territories In situations in which the attacker has more armies and can withstand losses, the formula to compute the probability that the attacker wins is not quite as simple as in the previous theorem and in particular breaks down into different cases depending on how many armies the attacker loses and when she loses these armies. In particular, when attacking a territory that defended by k armies, then one can see that there are $\lceil \frac{k}{2} \rceil$ ways in which the attacker can lose one army. Therefore, if the defending army spreads its d armies with d_1 armies on the first territory, d_2 on the second, etc., then there are a total of $\lceil \frac{d_1}{2} \rceil + \cdots + \lceil \frac{d_m}{2} \rceil$ ways in which the campaign can play out with the attacker losing a single army and defeating all m territories. We note that this sum depends only on the parities of the d_i. In particular, if $d_1 \geq 3$, then we note that there are the same number of cases to consider if we instead distribute the armies with $d_1 - 2$ armies on the first territory, $d_m + 2$ armies on the final territory, and d_i armies on the ith territory for all other i. Moreover, these cases pair up in a natural way depending on when the attacker loses the single army—in one case, she loses it on the first roll of the dice, in another the second roll, etc. Depending on when this loss occurs with respect to the breaks between territories, the probability of a given case occurring might be different, as we see in the following example.

Example 1. We consider two scenarios in which the defending army has eight armies to defend four territories against six attacking armies. In Scenario A, the defender splits the armies 3/2/2/1 and in Scenario B he splits the armies 1/2/2/3. In each scenario, there are five battles in which the attacker might lose a single army and therefore the defender wins overall. The following table gives the probability of each of these cases for each scenario, as well as the case where the attacker does not lose any armies, along with the relative attacker advantage given by Scenario A.

Battle lost	*Scenario A*	*Scenario B*	*Attacker Advantage for A*
None	$p_2 \widetilde{p}_1 p_2 p_2 \widetilde{q}_1$	$\widetilde{p}_1 p_2 p_2 q_2 \widetilde{q}_1$	$\frac{p_2}{q_2}$
First	$p_1 p_2 p_2 q_2 \widetilde{r}_1$	$\widetilde{p}_0 \widetilde{p}_1 p_2 q_2 r_1^2 \widetilde{r}_1$	$\frac{p_1 p_2}{\widetilde{p}_0 \widetilde{p}_1 r_1^2}$
Second	$p_2 \widetilde{p}_0 \widetilde{p}_1 p_2 q_2 \widetilde{r}_1$	$\widetilde{p}_1 p_1 \widetilde{p}_1 q_2 r^2 \widetilde{r}_1$	$\frac{p_2^2 \widetilde{p}_0}{p_1 \widetilde{p}_1 r_1^2}$
Third	$p_2 \widetilde{p}_1 p_1 \widetilde{p}_1 q_2 \widetilde{r}_1$	$\widetilde{p}_1 p_2 p_1 \widetilde{q}_1 r_1^2 \widetilde{r}_1$	$\frac{\widetilde{p}_1 q_2}{\widetilde{q}_1 r_1^2}$
Fourth	$p_2 \widetilde{p}_1 p_2 p_1 \widetilde{q}_1 r_1$	$\widetilde{p}_1 p_2 p_2 q_1 r_1 \widetilde{r}_1$	$\frac{p_1 \widetilde{q}_1}{q_1 \widetilde{r}_1}$
Fifth	$p_2 \widetilde{p}_1 p_2 p_2 \widetilde{q}_0 \widetilde{r}_1$	$\widetilde{p}_1 p_2 p_2 q_2 \widetilde{q}_0 \widetilde{r}_1$	$\frac{p_2}{q_2}$

In all six cases, one can compute that the attacker's advantage is greater than one in Scenario A. In particular, because the probability of the attacker winning the campaign is the sum of these six cases, each of which prefers the attacker in Scenario A, it is clear that the defending player should prefer Scenario B and therefore divide the armies as 1/2/2/3.

More generally, one can consider what the various possibilities are for the change in the attacker's probability of winning when two armies are shifted from the last territory to the first territory if one fixes the battle in which the attacker suffers her only casualty. In the following table, we consider all of the possibilities of where this loss can occur as well as the parity of the number of defending armies in this territory as well as the previous territory and the impact that will be felt if the defender moves these two armies.

Territory of Loss	Battle of Loss Within Terr.	Parity of Defenders in Terr. of Loss	Parity of Defenders in Prior Terr.	Advantage for Attacker	Numerical Advantage
No loss				$\frac{p_2}{q_2}$	1.63
Final	Not First	Any	Any	$\frac{p_2}{q_2}$	1.63
Final	First	Any	Even	$\frac{p_1\widetilde{q_1}}{q_1 r_1}$	2.35
Final	First	Any	Odd	$\frac{p_2\widetilde{p_0}\widetilde{q_1}}{\widetilde{p_1}q_1 r_1}$	1.34
Penultimate	Last	Any	Any	$\frac{p_1 q_2}{\widetilde{p_0}\widetilde{p_1}r_1^2}$	5.33
Penultimate	Middle	Any	Any	$\frac{q_2}{r_1^2}$	3.5
Penultimate	First	Even	Even	$\frac{\widetilde{p_1}q_2}{\widetilde{q_1}r_1^2}$	3.93
Penultimate	First	Even	Odd	$\frac{p_2\widetilde{p_0}q_2}{p_1\widetilde{q_1}r_1^2}$	2.28
Penultimate	First	Odd	Even	$\frac{\widetilde{p_1}\widetilde{q_1}}{r_1^2}$	5.78
Penultimate	First	Odd	Odd	$\frac{p_2\widetilde{p_0}\widetilde{q_1}}{p_1 r_1^2}$	3.35
Other	Last	Any	Any	$\frac{p_1 p_2}{\widetilde{p_0}\widetilde{p_1}r_1^2}$	8.69
Other	Middle	Any	Any	$\frac{p_2}{r_1^2}$	5.72
Other	First	Even	Even	$\frac{p_2}{r_1^2}$	5.72
Other	First	Odd	Even	$\frac{\widetilde{p_1}^2}{r_1^2}$	6.49
Other	First	Even	Odd	$\frac{p_2^2\widetilde{p_0}}{p_1\widetilde{p_1}r_1^2}$	3.31
Other	First	Odd	Odd	$\frac{p_2\widetilde{p_0}\widetilde{p_1}}{p_1 r_1^2}$	3.76

We can see that, in all cases, the attacker is better off if the defender shifts armies to the front of the line, and in some cases she will be significantly better off, suggesting that the defender should choose to move pairs of armies to the final territory in the chain whenever possible.

Defending fewer territories In situations where the attacker has more than $m + 2$ armies and therefore can lose more than a single army during the campaign, an enumeration of cases becomes significantly more complicated. However, it is clear that those cases will be a combination of the possibilities enumerated above for situations where the attacker only loses one army at any given time and situations where the attacker loses a pair of armies at the same time, which will affect the probabilities by giving advantages to the attacker of $\frac{p_2}{q_2}$, $\frac{p_2}{r_1^2}$, or $\frac{q_2}{r_1^2}$, all of which are bigger than one. Therefore, we conclude that the defender should always move armies in pairs to the final territory if possible. More specifically, we get the following theorem.

Theorem 3. *In order to best survive an attack on a chain of m territories with n armies, the best defense will be to place either one or two armies on each of the first $m-1$ territories and the remaining armies on the last territory.*

We have seen that in the case where the attacker has only $m+1$ armies that it will be best for the defender to place a single army on each of the first $m-1$ territories. In the following sections, we will see that the opposite conclusion holds if the attacker has an overwhelmingly large number of armies.

Holding off a large number of attacking armies

This section considers the situation in which the attacker has an overwhelmingly large number of attacking armies. In particular, we assume that the defender has a negligible chance of defeating the attacker and instead ask what distribution of the defending armies will do the most damage to the attacker during the campaign. This question is clearly different from our initial question, but we will see that the answer should be similar to the optimal distribution in that case.

Defending a single territory The first case we wish to consider is when the defending player is trying to defend a single territory with d armies on it. Let $E(d)$ be the expected change that a player with a large number of armies will have when attacking a territory with d armies. We note that $E(d)$ will be negative.

Lemma 4. $E(0) = -1$.

In particular, to take over an empty territory, the attacker needs to move a single army onto that territory, depleting her ranks by one.

Lemma 5. $E(1) = \frac{-1}{\widetilde{p}_1}$.

Proof. If the defending player has a single army, then on the first round there will be two possibilities: With probability $\widetilde{p}_1$, the attacker will win and therefore only "lose" the army she needs to use to take over the territory, and with probability p_0 the attacker will lose an army and have to face the defender in the same situation once again. Thus, $E(1) = -\widetilde{p}_1 + \widetilde{p}_0(-1 + E(1))$. The lemma follows from a simple calculation, noting that $\widetilde{p}_1 + \widetilde{p}_0 = 1$. ∎

When facing more than one army, there are three possible outcomes of a given attack. Considering these three cases, one can see that for $d \geq 2$ we have

$$E(d) = p_2 E(d-2) + p_1(E(d-1) - 1) + p_0(E(d) - 2),$$

which simplifies to give the following recursive formula for $E(d)$:

$$E(d) = \frac{p_2}{1-p_0} E(d-2) + \frac{p_1}{1-p_0} E(d-1) - \frac{p_1 + 2p_0}{1-p_0}.$$

There are several ways to approach nonhomogeneous recurrence relations such as this one; we proceed by noting that one can also see that

$$E(d+1) = \frac{p_2}{1-p_0}E(d-1) + \frac{p_1}{1-p_0}E(d) - \frac{p_1+2p_0}{1-p_0}.$$

Subtracting these two formulae from each other and solving for $E(d+1)$, one obtains the following homogeneous recursion formula for $E(d+1)$:

$$E(d+1) = \left(1 + \frac{p_1}{1-p_0}\right)E(d) + \left(\frac{p_2-p_1}{1-p_0}\right)E(d-1) - \frac{p_2}{1-p_0}E(d-2).$$

The characteristic equation of this recurrence relation is $x^3 - \left(1+\frac{p_1}{1-p_0}\right)x^2 - \left(\frac{p_2-p_1}{1-p_0}\right)x + \frac{p_2}{1-p_0}$, which factors as $(x-1)^2(x+\frac{p_2}{1-p_0})$. It follows from standard results in recurrence relations (see, for example, [**3**, Ch. 6]) that the generic solution to the recurrence relation takes the form

$$E(d) = c_1 + c_2 d + c_3\left(\frac{-p_2}{1-p_0}\right)^d.$$

Using Osborne's values for the p_i, it is easy to compute values for $E(2)$. Using the values of $E(0)$, $E(1)$, and $E(2)$, one can solve for the c_i and obtain the following result.

Theorem 6. *In a given battle in which the defending player has d armies and the attacking player starts with a large enough number of armies that she rolls three dice throughout the battle, the expected number of armies the attacker will lose before taking over the territory is* $E(d) = c_1 + c_2 d + c_3\alpha^d$ *where we define the constants* $c_1 = -\frac{578702951}{743204855}$, $c_2 = -\frac{2387}{2797}$, $c_3 = -\frac{164501904}{743204855}$, *and* $\alpha = -\frac{2890}{5501}$.

We note that our formula differs somewhat from the formula obtained in [**1**], in which they state the results as $E(n) = c_1 + c_2 n + c_3\alpha^n$ with $c_1 = 0.22134, c_2 = -0.85341, c_3 = -0.22134$ (and the same value of α). The difference arises because we are including the army that the attacker must "leave behind" when she moves onto the new territory.

Defending a chain of two territories We next consider the case where the defending player has two territories that he wishes to defend. Moreover, he can split the d armies between these two territories, although according to the rules of *RISK* he cannot vacate either territory. In particular, he must choose k with $1 \le k \le d-1$ and place k armies on the first territory and $d-k$ armies on the second territory. If he wishes to do this to maximize damage to his opponent, then he is trying to choose k to minimize the function $F(k) = E(k) + E(d-k)$, where E is the function defined in the previous section. In particular, one can use Theorem 6 to compute that $F(k) = 2c_1 + c_2 d + c_3(\alpha^k + \alpha^{d-k})$ where α and the c_i are the constants given in the statement of Theorem 6. In particular, we note that the first two terms in this formula are constants with respect to k and moreover that c_3 is negative, so it will suffice for the defending player to maximize the function $\hat{F}(k) = \alpha^k + \alpha^{d-k}$.

We note that the function F is symmetric in the sense that the attacker's expected losses will be the same if the defender places k armies on the first territory and $n-k$ armies on the second territory or the other way around. This is slightly different from actual gameplay in *RISK*, as our computations show that it is actually advantageous to place the smaller number of armies on the first territory and the larger number on the second territory. To understand why this is so, consider that, in our Markov chain model, the attacker that captures the first territory must leave behind a single army when occupying the captured territory. Thus, the attacker is effectively weakened for the second territory attack, no matter how well the first territory attack proceeds. If defender armies are unevenly distributed between the two spaces, the defender would be more likely to survive if the bulk of the defender armies were met by a weakened attacker. Although in cases with large numbers of armies, this weakening of armies left behind can be subtle, it is nonetheless a measurable advantage. At the end of this article, we will discuss further reasons why our model does not capture actual gameplay with complete accuracy.

Lemma 7. *If $\beta > 0$ and $m > 0$, then the function $\hat{G}(x) = \beta^x + \beta^{m-x}$ is minimized at $x = \frac{m}{2}$ and maximized at both $x = 0$ or $x = m$.*

Proof. The proof of this lemma is a straightforward calculus exercise, as one notes that $\hat{G}'(x) = \ln(\beta)(\beta^x - \beta^{m-x})$. If $0 < \beta < 1$, then $\ln(\beta) < 0$ and $\beta^x - \beta^{m-x}$ will be positive exactly when $x < m - x$. If $\beta > 1$, then both of these signs will be reversed. In either case, $\hat{G}$ will be decreasing for values of $x < \frac{m}{2}$ and increasing for values of x that are greater than $\frac{m}{2}$. ■

Our function is more interesting, as the value of α that we are working with is negative and therefore the function $\hat{F}$ is only defined at integer values of k. To maximize this function, we wish to consider two cases based on the parity of d.

We first consider the case where d is even. Note that for all even values of k the function $\hat{F}(k) = \alpha^k + \alpha^{d-k} = (\alpha^2)^{k/2} + (\alpha^2)^{d/2-k/2}$. Therefore, if k is an even integer, then the function $\hat{F}(k)$ agrees with the function $\hat{G}(k/2)$ as defined in the previous lemma where $\beta = \alpha^2$ and $m = d/2$. This function is maximized by choosing the smallest or largest values of k possible, which in this case must be $k = 2$ or $k = d - 2$ given the restrictions on k.

On the other hand, if k is odd, then we note that $\hat{F}(k) = \alpha(\alpha^{k-1} + \alpha^{d-k-1}) = \alpha((\alpha^2)^{(k-1)/2} + (\alpha^2)^{(d-k-1)/2})$. Because α is negative, we note that this number will always be negative. In particular, it will always be less than $\hat{F}(k)$ for any even choice of k. In particular, we can conclude that if d is even, then $\hat{F}(k)$ is maximized when $k = 2$ (or $k = d - 2$ due to symmetry).

If d is odd, then we begin by noting that for even values of k that the function $\hat{F}(k)$ is equal to $(\alpha^2)^{k/2} + \alpha \cdot (\alpha^2)^{(d-k-1)/2}$. Because $0 < \alpha^2 < 1$, one can compute that this is a continuous function whose derivative is always negative, and therefore the function is strictly decreasing. One can similarly show that $\hat{F}(k)$ is strictly increasing for odd values of k. The fact that $\hat{F}(1) < 0$ and $\hat{F}(2) > 0$ then implies that, for integers between 1 and $d - 1$, this function is maximized at $k = 2$ (or $k = d - 2$ due to symmetry).

We have therefore proved the following theorem.

Theorem 8. *In a situation in which the defending player is trying to defend two consecutive territories with d armies and the attacking army has significantly more than d armies, then the defending player will cause the most damage if he places two armies on one territory and $d - 2$ armies on the other territory.*

It is interesting to note that while the best strategy is to place two armies on one territory and $d - 2$ on the other, the *worst* strategy is actually to divide your armies with one army on one territory and $d - 1$ armies on the other territory and that in general one wishes to place an even number of armies on each territory.

Defending longer chains In this section, we wish to show that the optimal defense for defending a chain of m territories, assuming that one has $d \geq 2m$ armies, will be to place two armies on all but one of the territories and $d - 2(m - 1)$ on the remaining territory. Having shown this in the case where $m = 2$, we now wish to proceed by induction.

Given an arrangement $(d_1, \dots, d_m)$ of armies on the m territories, we define the expected loss of armies by the attacker by the function

$$F_m(d_1, \dots, d_m) = E(d_1) + \cdots + E(d_m)$$
$$= mc_1 + dc_2 + c_3(\alpha^{d_1} + \cdots + \alpha^{d_m})$$

where α, c_1, c_2, c_3 are as defined in Theorem 6. In particular, this function will be minimized exactly when the function $\hat{F}_m(d_1, \dots, d_m) = \alpha^{d_1} + \cdots + \alpha^{d_m}$ is maximized and the defender's goal is to choose the constants $d_1, \dots, d_m$ with $1 \leq d_i \leq d$ and $\sum d_i = d$ so that $\hat{F}_m(d_1, \dots, d_m)$ is maximized.

We note that for any fixed choice of d_m, we have that $\hat{F}_m(d_1, \dots, d_m) = \alpha^{d_m} + \hat{F}_{m-1}(d_1, \dots, d_{m-1})$. Therefore, this function will be maximized when $\hat{F}_{m-1}(d_1, \dots, d_{m-1})$ is maximized. However, this describes the situation in which a defender is trying to defend $m - 1$ territories with $d - d_m$ armies, and by the inductive hypothesis we know that this will be maximized when all but one of the entries is equal to 2.

Another way of seeing this is by contradiction. Assume that $\hat{F}_m(a_1, \dots, a_m)$ is a maximum value of $\hat{F}_m$ over all m-tuples with $\sum a_i = n$ and further assume that more than one of the a_i is not equal to 2: Without loss of generality, let us assume that $a_1 \neq 2$ and $a_2 \neq 2$. Then $\hat{F}_{m-1}(a_1, \dots, a_{m-1})$ will be a maximum for $\hat{F}_{m-1}$ given the restriction that $\sum a_i = n - a_m$. However, this contradicts the inductive hypothesis, which states that $\hat{F}_{m-1}$ will be maximized when all but one of the entries is equal to two.

We note that the proof of the above theorem is actually quite general and in fact is quite robust in the values of the p_i. In particular, we note that one can conclude that the same strategy is the best if the attacking player always rolls exactly two dice. Given that the actual play in the game of *RISK* is a linear combination of these two scenarios, it seems natural that the best possible strategy to cause the most damage to your opponent in the actual game will be to place two armies on each but the final territory due to a convexity type of argument.

Concluding thoughts

Our model suggests, and proves under various hypotheses, that the best strategy a defending player has in order to fend off an attacker is to place two armies on each territory except the last territory, where he will place the remaining armies. However, there are several aspects of the game of *RISK* that our model does not accommodate for, and we conclude this paper by briefly discussing some of these considerations.

In practice, one will usually reverse the distribution of defending armies, putting the main force at the front. The reason for this is that while this leads to a slight decrease in survival probability, it gives the defending player an increased expectation of retained territories leading to an increased expectation in earned army reinforcements. In the game of *RISK*, a player receives army reinforcements at the beginning of each turn according to the total number of territories the player occupies modulo 3 (with a three army reinforcement minimum) plus bonuses for complete continents occupied.

Thus, while survival is important, a small tradeoff of the probability of immediate survival is advisable to increase the expectation of territory occupation and thus army reinforcements and thus the probability of longer-term survival. Such long-term strategic considerations are beyond the scope of this tactical paper and would indeed be interesting future work. However, our work has given substantial evidence that the two-army-per-territory defense is the most efficient in certain situations and provided further insight to extreme cases.

Acknowledgment We would like to thank Clifton Presser for his preparation of the graphics used in this paper.

REFERENCES

1. O. Lyne, *RISK* FAQ—Version 5.61, `http://www.kent.ac.uk/IMS/personal/odl/riskfaq.htm`.
2. J. A. Osborne, Markov chains for the *RISK* board game revisited. *Math. Mag.* **76** no. 2 (2003) 129–135.
3. F. S. Roberts, B. Tesman, *Applied Combinatorics*. Second ed. CRC Press, Boca Raton, FL, 2009.
4. B. Tan, Markov chains and the *RISK* board game. *Math. Mag.* **70** no. 5 (1997) 349–357.

Summary. We consider a one-dimensional version of the board game *RISK* and discuss the problem of how a defending player might choose to distribute his armies along a chain of territories in order to maximize the probability of survival. In particular, we analyze a Markov chain model of this situation and run computer simulations in order to make conjectures as to the optimal strategies. The latter sections of the paper analyze this strategy rigorously and use results on recurrence relations and probability theory in order to prove a related result.

DARREN GLASS (MR Author ID: 699957) received his B.A. in mathematics and economics at Rice University and his Ph.D. at the University of Pennsylvania, where he worked on questions in arithmetic geometry and Galois theory. Since joining the math department at Gettysburg College in 2005, his research interests have expanded to include number theory, graph theory, sabermetrics, cryptography, and whatever interesting questions colleagues wander into his office to ask. This paper is evidence of the latter.

TODD W. NELLER (MR Author ID: 750401) received his computer science Ph.D. at Stanford University with distinction in teaching in 2000, where he was awarded a Stanford University Lieberman Fellowship, and the George E. Forsythe Memorial Award for excellence in teaching. His dissertation concerned simulation-based search for hybrid system control and analysis. A game enthusiast, Neller has since enjoyed a variety of artificial intelligence game and puzzle challenges, including the computation of optimal play for dice games (e.g., Pass the Pigs, Dudo) and the computer-aided design of logic mazes.

MAA 100
MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

Among Notable Authors who have Published in the MAGAZINE

Theodore Kaczynski, the Unabomber, published T. Kaczynski, Note on a problem of Alan Sutcliffe, *Math. Mag.* **41** no. 2 (1968) 84–86.

Non-Integrality of Binomial Sums and Fermat's Little Theorem

DANIEL LÓPEZ-AGUAYO
Centro de Ciencias Matemáticas UNAM
Morelia, Mexico
dlopez@matmor.unam.mx

In the April 2014 issue of THIS MAGAZINE, Marcel Chiriță proposed Problem 1942 which asked to prove that

$$\sum_{k=0}^{n} \frac{k}{k+r}\binom{n}{k} \tag{1}$$

is not an integer for $n \geq 2$ and $r = 1$. Two solutions to this problem appear in this issue (on page 238), and both apply Fermat's little theorem. A natural question is to ask if there are infinitely many integers r for which Equation (1) is not an integer. In this note we show that Equation (1) is nonintegral for $r = 2$, 3, and 4. Our results also hinge on Fermat's little theorem, as it may be used to prove the following useful lemma.

Lemma. *For every integer $r \geq 2$ we have that $2^r \not\equiv 1 \mod r$.*

All proofs below follow the same methodology. Rearrange the sum and apply the binomial theorem repeatedly to yield an expression that is shown to be nonintegral via the lemma.

Proposition 1. *Let $n \geq 2$. Then $\sum_{k=0}^{n} \frac{k}{k+2}\binom{n}{k}$ is not an integer.*

Proof. The summation may be rearranged so that

$$\begin{aligned}\sum_{k=0}^{n} \frac{k}{k+2}\binom{n}{k} &= \sum_{k=0}^{n}\left(1-\frac{2}{k+2}\right)\binom{n}{k} = 2^n - 2\sum_{k=0}^{n}\frac{1}{k+2}\binom{n}{k} \\ &= 2^n - \frac{2}{(n+1)(n+2)}\sum_{k=0}^{n}(k+1)\binom{n+2}{k+2} \\ &= 2^n - \frac{2}{(n+1)(n+2)}\left[\sum_{k=0}^{n}k\binom{n+2}{k+2} + \sum_{k=0}^{n}\binom{n+2}{k+2}\right].\end{aligned}$$

To simplify the above expression, we first evaluate $\sum_{k=0}^{n} k\binom{n+2}{k+2}$. Differentiating $\sum_{k=0}^{n+2}\binom{n}{k}x^k = (1+x)^{n+2}$ and letting $x = 1$ yields

$$\sum_{k=0}^{n+2} k\binom{n}{k} = (n+2)2^{n+1}. \tag{2}$$

Math. Mag. **88** (2015) 231–234. doi:10.4169/math.mag.88.3.231.
MSC: Primary 11B65

By a change of variables,

$$\sum_{k=0}^{n} k\binom{n+2}{k+2} = \sum_{w=2}^{n+2}(w-2)\binom{n+2}{w} = \sum_{w=2}^{n+2} w\binom{n+2}{w} - \sum_{w=2}^{n+2} 2\binom{n+2}{w}. \quad (3)$$

Using Equation (2) and the binomial theorem, Equation (3) simplifies to

$$2^{n+1}(n+2) - (n+2) - 2\left[2^{n+2} - 1 - (n+2)\right] = 2^{n+1}n - 2^{n+2} + n + 4. \quad (4)$$

Because $\sum_{k=0}^{n}\binom{n+2}{k+2} = 2^{n+2} - 1 - (n+2)$, then

$$\sum_{k=0}^{n} \frac{k}{k+2}\binom{n}{k} = 2^n - \frac{2(2^{n+1}n+1)}{(n+1)(n+2)}.$$

It suffices to show that $\frac{2(2^{n+1}n+1)}{(n+1)(n+2)}$ is not an integer. Suppose otherwise. If $n+1$ is odd, then $2^{n+1}n + 1 \equiv 0 \mod (n+1)$. This implies that $2^{n+1}n + 1 \equiv 1 - 2^{n+1} \mod (n+1)$, which contradicts the lemma. If $n+1$ is even, then $n+2$ is odd and $2^{n+1}n + 1 \equiv 0 \mod (n+2)$. This implies that $2^{n+1}n + 1 \equiv 1 - 2^{n+2} \mod (n+2)$, which contradicts lemma. Therefore, the requisite sum is nonintegral. ■

Proposition 2. *Let $n \geq 2$. Then $\sum_{k=0}^{n} \frac{k}{k+3}\binom{n}{k}$ is not an integer.*

Proof. Similar applications of the binomial theorem as in the proof of Proposition 1 gives

$$\sum_{k=0}^{n} \frac{k}{k+3}\binom{n}{k} = 2^n - \frac{6(2^n n^2 + 2^n n + 2^{n+1} - 1)}{(n+1)(n+2)(n+3)}.$$

The calculations may be found on the online supplement to this article [**1**].

We need to show that $\frac{6(2^n n^2+2^n n+2^{n+1}-1)}{(n+1)(n+2)(n+3)}$ is not an integer. Suppose first that $n+1$ is even, then $n+3$ is even, too, which implies that $(n+1)(n+3)$ is evenly divisible by 4. But $6(2^n n^2 + 2^n n + 2^{n+1} - 1)$ is not evenly divisible by 4, because

$$\begin{aligned} 6(2^n n^2 + 2^n n + 2^{n+1} - 1) &\equiv 2(2^n n^2 + 2^n n + 2^{n+1} - 1) \mod 4 \\ &\equiv 2^{n+1}[n(n+1) + 2] - 2 \mod 4 \\ &\equiv -2 \mod 4. \end{aligned}$$

Hence, in this case, $\frac{6(2^n n^2+2^n n+2^{n+1}-1)}{(n+1)(n+2)(n+3)}$ is not an integer.

Suppose that $n+1$ is odd, but is not a multiple of 3. It must be the case that $n+1$ and 6 are relatively prime. Notice that

$$\begin{aligned} 2^n n^2 + 2^n n + 2^{n+1} - 1 &= 2^n n(n+1) + 2^{n+1} - 1 \\ &\equiv 2^{n+1} - 1 \mod (n+1). \end{aligned}$$

For $\frac{6(2^n n^2+2^n n+2^{n+1}-1)}{(n+1)(n+2)(n+3)}$ to be an integer, then $n+1$ must evenly divide $2^n n^2 + 2^n n + 2^{n+1} - 1$ or, equivalently, $2^{n+1} \equiv 1 \mod (n+1)$, which contradicts the lemma.

For $n+1$ odd and a multiple of 3, then $n+3$ is odd and not divisible by 3. This time $n+3$ and 6 are relatively prime. In a similar approach to the above case,

$$\begin{aligned} 2^n n^2 + 2^n n + 2^{n+1} - 1 &\equiv 2^n (n^2 + n + 2) - 1 \mod (n+3) \\ &\equiv 2^n [(n-2)(n+3) + 8] - 1 \mod (n+3) \\ &\equiv 2^{n+3} - 1 \mod (n+3). \end{aligned}$$

For $\frac{6(2^n n^2 + 2^n n + 2^{n+1} - 1)}{(n+1)(n+2)(n+3)}$ to be an integer, then $n+3$ must evenly divide $2^n n^2 + 2^n n + 2^{n+1} - 1$ or, equivalently, $2^{n+3} \equiv 1 \mod (n+3)$, which contradicts the lemma.

Therefore, $\frac{6(2^n n^2 + 2^n n + 2^{n+1} - 1)}{(n+1)(n+2)(n+3)}$ is not an integer, which completes the proof. ■

Proposition 3. *Let $n \geq 2$. Then $\sum_{k=0}^{n} \frac{k}{k+4}\binom{n}{k}$ is not an integer.*

Proof. Using a similar method as before, it follows that

$$\sum_{k=0}^{n} \binom{n}{k} \frac{k}{k+4} = 2^n - \frac{8(2^n n^3 + 3 \cdot 2^n n^2 + 2^{n+3} n + 3)}{(n+1)(n+2)(n+3)(n+4)}. \tag{5}$$

Suppose first that $n+3$ is odd but not divisible by 3. Because $n+3$ and 8 are relatively prime, if $\frac{8(2^n n^3 + 3 \cdot 2^n n^2 + 2^{n+3} n + 3)}{(n+1)(n+2)(n+3)(n+4)}$ is an integer, then $n+3$ must divide $2^n n^3 + 3 \cdot 2^n n^2 + 2^{n+3} n + 3$. Notice that $n^3 + 3n^2 + 8n = (n+3)(n^2+8) - 24$ and therefore $n^3 + 3n^2 + 8n \equiv -24 \mod (n+3)$. Consequently,

$$\begin{aligned} 2^n n^3 + 3 \cdot 2^n n^2 + 2^{n+3} n + 3 &\equiv 2^n(-24) + 3 \mod (n+3) \\ &\equiv 3 - 3 \cdot 2^{n+3} \mod (n+3). \end{aligned}$$

Thus $n+3$ divides $3(2^{n+3} - 1)$ and because $n+3$ and 3 are relatively prime, then $2^{n+3} - 1 \equiv 0 \mod (n+3)$, which contradicts the lemma.

Suppose now that $n+3$ is odd but divisible by 3. Then $n+1$ is odd but not divisible by 3. As before, because $n+1$ and 8 are relatively prime, then for the expression in Equation (5) to be an integer, then $n+1$ must divide $2^n n^3 + 3 \cdot 2^n n^2 + 2^{n+3} n + 3$. Now note that $n^3 + 3n^2 + 8n = (n+1)(n^2 + 2n + 6) - 6$ and thus $n^3 + 3n^2 + 8n \equiv -6 \mod (n+1)$. Therefore,

$$\begin{aligned} 2^n n^3 + 3 \cdot 2^n n^2 + 2^{n+3} n + 3 &\equiv 2^n(-6) + 3 \mod (n+1) \\ &\equiv 3 - 3 \cdot 2^{n+1} \mod (n+1). \end{aligned}$$

Therefore, $n+1$ divides $3(2^{n+1} - 1)$. Because $n+1$ and 3 are relatively prime, then $2^{n+1} - 1 \equiv 0 \mod (n+1)$, which once again contradicts the lemma.

Suppose now that $n+3$ is even and that $n+2$ is not divisible by 3. For the expression in Equation (5) to be an integer, then $n+2$ must divide $2^n n^3 + 3 \cdot 2^n n^2 + 2^{n+3} n + 3$. Since $n^3 + 3n^2 + 8n = (n+2)(n^2 + n + 6) - 12$, then $n^3 + 3n^2 + 8n \equiv -12 \mod (n+2)$. It follows that

$$\begin{aligned} 2^n n^3 + 3 \cdot 2^n n^2 + 2^{n+3} n + 3 &\equiv 2^n(-12) + 3 \mod (n+2) \\ &\equiv 3 - 3 \cdot 2^{n+2} \mod (n+2). \end{aligned}$$

Therefore, $n+2$ divides $3(2^{n+2} - 1)$. Because $n+2$ and 3 are relatively prime, then $2^{n+2} - 1 \equiv 0 \mod (n+2)$—another contradiction to the lemma.

Finally, assume that $n+3$ is even but $n+2$ is divisible by 3. In this case $n+4$ is odd and not divisible by 3. For the expression in Equation (5) to be an integer,

$n+4$ must divide $2^n n^3 + 3\cdot 2^n n^2 + 2^{n+3}n + 3$. Note that $n^3 + 3n^2 + 8n = (n^2 - n + 12)(n+4) - 48$ and thus $n^3 + n^2 + 8n \equiv -48 \mod (n+4)$. Therefore,

$$\begin{aligned} 2^n n^3 + 3\cdot 2^n n^2 + 2^{n+3}n + 3 &\equiv 2^n(-48) + 3 \mod (n+4) \\ &\equiv -2^n 2^4 \cdot 3 + 3 \mod (n+4) \\ &\equiv 3 - 3\cdot 2^{n+4} \mod (n+4). \end{aligned}$$

It follows that $n+4$ divides $3(2^{n+4}-1)$. Since $n+4$ and 3 are relatively prime, then $n+4$ divides $2^{n+4}-1$; this is the final contradiction of the lemma and it completes the proof. ■

Acknowledgment The author wishes to thank Moubariz Garaev for his helpful advice on this note.

REFERENCES

1. D. López-Aguayo, Online supplement to "Nonintegrality of Binomial Sums and Fermat's Little Theorem," www.maa.org/mathmag/supplements.

Summary. Problem 1942 from the April 2014 issue of the Magazine asks whether a particular binomial sum is nonintegral. We pose an open question whether there are infinitely many integers r for which the requisite sum is nonintegral when $k+1$ is replaced by $k+r$. We prove that the sum is nonintegral for $r = 2$, 3, and 4 by an application of Fermat's little theorem.

DANIEL LÓPEZ-AGUAYO (MR Author ID 1105399) is a Ph.D. candidate at the Center of Mathematical Sciences UNAM working under the supervision of Raymundo Bautista. His field of research is the representation theory of finite dimensional algebras. He also has a keen interest in homological algebra, module theory, and the history of mathematics.

PROBLEMS

BERNARDO M. ÁBREGO, *Editor*
California State University, Northridge

Assistant Editors: SILVIA FERNÁNDEZ-MERCHANT, California State University, Northridge; JOSÉ A. GÓMEZ, Facultad de Ciencias, UNAM, México; EUGEN J. IONASCU, Columbus State University; ROGELIO VALDEZ, Facultad de Ciencias, UAEM, México; WILLIAM WATKINS, California State University, Northridge

PROPOSALS

To be considered for publication, solutions should be received by November 1, 2015.

1971. *Proposed by George Apostolopoulos, Messolonghi, Greece.*

Find all pairs of integers (x, y) such that

$$x^8 + (y^2 + y - 1)(4 - 3x^4) = 2.$$

1972. *Proposed by Marcel Chirita, Bucharest, Romania.*

Let $n \geq 2$ be an integer. Determine all the continuous functions $f : [1, \infty) \to \mathbb{R}$ such that

$$\int_x^{x^n} f(t)dt = \int_1^x (t^{n-1} + t^{n-2} + \cdots + t) f(t)dt$$

for every $x \in [1, \infty)$.

1973. *Proposed by Arkady Alt, San Jose, CA.*

Let $\Delta(x, y, z) = 2(xy + yz + xz) - (x^2 + y^2 + z^2)$. Prove that for any positive real numbers a, b, and c, the following inequality holds:

$$\left(\Delta(a^2, b^2, c^2)\right)^2 \geq \Delta(a, b, c) \cdot \Delta(a^3, b^3, c^3).$$

Math. Mag. **88** (2015) 235–242. doi:10.4169/math.mag.88.3.235.

We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals must, in general, be accompanied by solutions and by any 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.

Solutions should be written in a style appropriate for this MAGAZINE.

Solutions and new proposals should be mailed to Bernardo M. Ábrego, Problems Editor, Department of Mathematics, California State University, Northridge, 18111 Nordhoff St, Northridge, CA 91330-8313, or mailed electronically (ideally as a LaTeX or pdf file) to mathmagproblems@csun.edu. All communications, written or electronic, should include **on each page** the reader's name, full address, and an e-mail address and/or FAX number.

1974. *Proposed by Boon Wee Ong, Behrend College, Erie, PA.*

Let $q \neq 1$ be a positive real number. Define for $n \geq 1$,

$$\nu_n = \frac{q^{n/2} - q^{-n/2}}{q^{1/2} - q^{-1/2}} \quad \text{and} \quad \mu_n = q^{n/2} + q^{-n/2}.$$

(a) Prove that $2\nu_{n+1} = \mu_1 \nu_n + \mu_n$ for all $n \geq 1$.

(b) Prove that $2\nu_{n-1} = \mu_1 \nu_n - \mu_n$ for all $n \geq 2$.

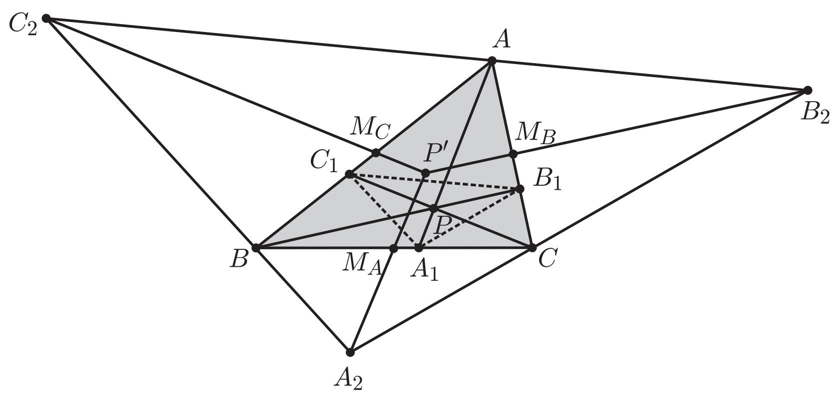
(c) Write the sum

$$\sum_{n=2}^{N} \frac{1}{\nu_n \nu_{n+1}}$$

in closed form.

1975. *Proposed by Sohail Farhangi, Virginia Polytechnic Institute and State University, Blacksburg, VA.*

Let ABC be a triangle in the plane. For every point P in the interior of $\triangle ABC$, construct point P' as follows. Let lines AP, BP, and CP intersect sides BC, CA, and AB at points A_1, B_1, and C_1, respectively. Let L_A be the line passing through A that is parallel to B_1C_1, and define lines L_B and L_C in a similar manner. Let A_2 be the intersection of L_B and L_C, and define points B_2 and C_2 in a similar manner. Let M_A, M_B, and M_C be the midpoints of sides BC, CA, and AB, respectively. Finally, let P' be the concurrent point of the lines A_2M_A, B_2M_B, and C_2M_C (which are guaranteed to concur by the cevian nest theorem.) Prove that for any two points R and S inside $\triangle ABC$ the lines RS and $R'S'$ are parallel.



Quickies

Answers to the Quickies are on page 242.

Q1051. *Proposed by Howard Cary Morris, Cordova, TN.*

Let $p(x) = \sum a_k x^{n_k}$ be a polynomial where the a_k are real, and the n_k are in ascending (or descending) order. According to Descartes sign rule, if m is the number of sign changes among consecutive a_k, then $p(x)$ has at most m positive zeros. Let $D(x) = \sum a_k e^{b_k x}$ be a Dirichlet polynomial where the a_k are real and the b_k are in ascending (or descending) order. Prove that if m is the number of sign changes among consecutive a_k, then $D(x)$ has at most m real zeros.

Q1052. *Proposed by Digby Smith, Mount Royal University, Calgary, AB, Canada.*

(a) Evaluate

$$\int \frac{e^{20x} - e^{15x}}{e^{20x} + e^{15x}} dx.$$

(b) Suppose A and B are constants with $A \neq B$. Evaluate

$$\int \frac{e^{Ax} - e^{Bx}}{e^{Ax} + e^{Bx}} dx.$$

Solutions

***Freshman* quotient rule and more** **April 2014**

1941. *Proposed by Spiros P. Andriopoulos, Third High School of Amaliada, Eleia, Greece.*

Find all pairs of nonzero real valued functions $(u(x), v(x))$ that satisfy both of the following equations:

$$\left(\frac{u}{v}\right)' = \frac{u'}{v'} \quad \text{and} \quad u'v' = uv.$$

Solution by Joseph DiMuro, Biola University, La Mirada, CA.

From the first and second equations, we have

$$\frac{u'v - uv'}{v^2} = \frac{u'}{v'} \quad \text{and} \quad u = \frac{u'v'}{v}.$$

From the first equation,we know that $v(x)$ is never 0. Plugging u in the second equation into the first equation gives

$$\frac{u'v^2 - u'(v')^2}{v^3} = \frac{u'}{v'}.$$

Assuming there are no regions where $u' = 0$ (and if there are, we must also have $u = 0$ to satisfy the two original equations), we may multiply both sides by (v'/u'), producing

$$\frac{v'}{v} - \left(\frac{v'}{v}\right)^3 = 1$$

Thus, v'/v must be a constant c, such that $c - c^3 = 1$. (And there is only one such constant, roughly -1.325.) Thus, we must have $v = Ae^{cx}$ for some nonzero constant A. Going back to the second given equation, in regions where u is never 0, we have $u'/u = v/v' = 1/c$. So we must have $u = Be^{x/c}$ for some nonzero constant B. It is impossible for u to jump between being a constant 0 and being a function of the form $Be^{x/c}$ since then u would not be everywhere differentiable. So if u is 0 anywhere, then u is 0 everywhere, and then v can be any differentiable function where v and v' are never 0. So those are the only possible solutions: Assuming u is not just a constant 0, we must have $u = Be^{x/c}$ and $v = Ae^{cx}$ for some nonzero constants A and B, or $u = 0$ and v is an arbitrary nonzero differentiable function with nonzero derivative. It is then easy to check that any such pair of functions satisfies the two given equations.

Editor's Note. Many of our solvers calculated the exact value of the constant c (or an equivalent related constant), it turns out to be equal to

$$c = -\frac{(9-\sqrt{69})^{1/3} + (9+\sqrt{69})^{1/3}}{18^{1/3}}.$$

Also solved by Arkady Alt, Hatef I. Arshagi, Seungyoon Baek (Korea) and Yugeun Shim(Korea), Michel Bataille (France), Stan Byrd and Roger Nichols, Robert Calcaterra, L. Cíosoǵ, Neil Curwen (United Kingdom), Gino T. Fala, Michael Goldenberg and Mark Kaplan, GWstat Problem Solving Group, Eugen A. Herman, Omran Kouba (Syria), Elias Lampakis (Greece), Charles Z. Martin, Missouri State University Problem Solving Group, Edward Schmeichel, Achilleas Sinefakopoulos (Greece), Nora S. Thornber, Traian Viteam (India), MIchael Vowe (Switzerland), and the proposer. There were three incorrect submissions.

A sum of binomials is never an integer **April 2014**

1942. *Proposed by Marcel Chiriţă, Bucharest, Romania.*

Let $n \geq 2$ be an integer. Prove that

$$\sum_{k=0}^{n} \frac{k}{k+1}\binom{n}{k}$$

is not an integer.

I. *Solution by Missouri State University Problem Solving Group, Missouri State University, Springfield, MO.*

The condition $n \geq 2$ can be relaxed to $n \geq 1$. Since $\sum_{k=0}^{n}\binom{n}{k} = 2^n$ and

$$\sum_{k=0}^{n} \frac{k}{k+1}\binom{n}{k} = \sum_{k=0}^{n}\binom{n}{k} - \sum_{k=0}^{n}\frac{1}{k+1}\binom{n}{k},$$

it suffices to show that $\sum_{k=0}^{n}\binom{n}{k}/(k+1)$ is not an integer. Now

$$\sum_{k=0}^{n}\frac{1}{k+1}\binom{n}{k} = \int_0^1 \sum_{k=0}^{n}\binom{n}{k}x^k\,dx = \int_0^1 (x+1)^n dx = \frac{2^{n+1}-1}{n+1},$$

so it suffices to show that $2^m \not\equiv 1 \pmod{m}$ for $m \geq 2$. Suppose to the contrary that $2^m \equiv 1 \pmod{m}$. This forces m to be odd. Let p be the smallest prime factor of m; note that p is odd. It follows from our assumption that $2^m \equiv 1 \pmod{p}$. We also know that $2^{p-1} \equiv 1 \pmod{p}$ by Fermat's little theorem. Hence, $2^{\gcd(m,p-1)} \equiv 1 \pmod{m}$. But since p is the smallest prime factor of m, it follows that $\gcd(m, p-1) = 1$, and we conclude that $2 \equiv 1 \pmod{p}$, which is a contradiction.

II. *Solution by Peter Hauber, University of Applied Sciences, Stuttgart, Germany.*

First we observe that the sum in question equals

$$\sum_{k=0}^{n}\frac{k}{k+1}\binom{n}{k} = \sum_{k=0}^{n}\left(1-\frac{1}{k+1}\right)\binom{n}{k} = \sum_{k=0}^{n}\binom{n}{k} - \frac{1}{n+1}\sum_{k=0}^{n}\frac{n+1}{k+1}\binom{n}{k}$$
$$= 2^n - \frac{1}{n+1}\sum_{k=0}^{n}\binom{n+1}{k+1} = 2^n - \frac{2^{n+1}-1}{n+1}.$$

Therefore, the claim will be proved if we show that $n+1$ never divides $2^{n+1}-1$ for $n \geq 1$. This can be proved as in the second part of the first solution.

Editor's Note. Many of our readers remarked that proving that n does not divide $2^n - 1$ (for $n > 1$) appeared as problem A5 in the 1972 William Lowell Putnam Exam.

Also solved by Daniel López-Aguayo (México), Adnan Ali (India), Arkady Alt, Michael Andreoli, Michael R. Bacon and Charles K. Cook, Josiah M. Banks, Michel Bataille (France), David Bressoud and Stan Wagon, Robert Calcaterra, John Christopher, L. Cíosoǵ, CMC 328, Neil Curwen, Joseph DiMuro, John Fitch, Daniel Fritze (Germany), Eugene A. Herman, Joel Iiams, Omran Kouba (Syria), Harris Kwong, Andrew Kwon, Elias Lampakis (Greece), Kee-Wai Lau (China), Kathleen E. Lewis (Republic of the Gambia), Reiner Martin (Germany), Kandasamy Muthuvel, Northwestern University Math Problem Solving Group, Moubinool Omarjee (France), Edward Omey (Belgium), Henry Ricardo, Greg Ronsse, Achilleas Sinefakopoulos (Greece), Dave Trautman, Jim Vandergriff, Traian Viteam (India), Timothy Woodcock, Titu Zvonaru and Neculai Stanciu (Romania), and the proposer. There were four incomplete or incorrect solutions.

An inequality of nested square roots **April 2014**

1943. *Proposed by Mihály Bencze, Brason, Romania.*

Let n be a positive integer, and $a_1, a_2, \ldots, a_n$ be positive real numbers. Set $A = \frac{1}{n}\sum_{k=1}^{n} a_k$ and $a_{n+k} = a_k$ for $1 \leq k \leq n-1$. Prove that

$$\sum_{k=0}^{n-1}\sqrt{a_{k+1}+\sqrt{a_{k+2}+\sqrt{a_{k+3}+\cdots+\sqrt{a_{k+n}}}}} \leq n\sqrt{A+\sqrt{A+\sqrt{A+\cdots+\sqrt{A}}}}.$$

Solution by Harris Kwong, SUNY Fredonia, Fredonia, NY.

The condition $a_{n+k} = a_k$ means the subscripts run cyclicly modulo n. For any integer $n \geq 1$, define

$$a_{i,n} = a_i + \sqrt{a_{i+1}+\sqrt{a_{i+2}+\cdots+\sqrt{a_{i+n-1}}}}$$

so that

$$a_{i,1} = a_i, \text{ and } a_{i,n} = a_i + \sqrt{a_{i+1,n-1}} \text{ for } n \geq 2,$$

and we can write the sum on the left of the stated inequality as $\sum_{i=1}^{n}\sqrt{a_{i,n}}$, where the subscript i runs cyclicly modulo n. Cauchy–Schwarz inequality asserts that

$$\frac{1}{n}\sum_{i=1}^{n}\sqrt{a_{i,n}} \leq \sqrt{\frac{1}{n}\sum_{i=1}^{n}a_{i,n}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(a_i+\sqrt{a_{i+1,n-1}}\right)}.$$

Because of the cyclic nature of the subscript i,

$$\sum_{i=1}^{n}\left(a_i+\sqrt{a_{i+1,n-1}}\right) = \sum_{i=1}^{n}a_i + \sum_{i=1}^{n}\sqrt{a_{i,n-1}}.$$

Hence,

$$\frac{1}{n}\sum_{i=1}^{n}\sqrt{a_{i,n}} \leq \sqrt{A+\frac{1}{n}\sum_{i=1}^{n}\sqrt{a_{i,n-1}}}.$$

A repetition of the argument leads to

$$\frac{1}{n}\sum_{i=1}^{n}\sqrt{a_{i,n}} \le \sqrt{A+\sqrt{A+\sqrt{A+\cdots+\sqrt{A}}}},$$

which yields the desired inequality.

Also solved by Arkady Alt, Michel Bataille (France), Robert Calcaterra, L. Cíosoǵ, Neil Curwen (United Kingdom), Eugene A. Herman, Omran Kouba (Syria), Elias Lampakis (Greece), Reiner Martin (Germany), Rituraj Nandan, Paolo Perfetti (Italy), Ángel Plaza (Spain), Edward Schmeichel, Yugeun Shim (Korea), and the proposer.

On the sum of the square-lengths of a triangle **April 2014**

1944. *Proposed by Mowaffaq Hajja and Mostafa Hayajneh, Yarmouk University, Irbid, Jordan.*

For any triangle XYZ with side lengths x, y, and z, let $\nu(\triangle XYZ) = x^2 + y^2 + z^2$.

(a) Let $ABCD$ be a convex quadrilateral with diagonals intersecting at P. Prove that if

$$\nu(\triangle PAB) = \nu(\triangle PBC) = \nu(\triangle PCD) = \nu(\triangle PDA), \tag{1}$$

then $ABCD$ is a rhombus.

(b) Classify the quadrilaterals $ABCD$ for which there exists a point P (not necessarily the intersection of the diagonals) such that (1) holds.

Solution by Seungyoon Baek, Institute of Science Education for the Gifted and Talented, Yonsei University, Seoul, Korea.

Suppose that there exists a point P such that (1) holds. Since $\nu(\triangle PAB) = \nu(\triangle PDA)$, it follows that

$$PA^2 + AB^2 + PB^2 = PA^2 + AD^2 + PD^2;$$

that is, $PB^2 - PD^2 = AD^2 - AB^2$. Similarly, we have $PB^2 - PD^2 = CD^2 - BC^2$. Hence, $AD^2 - AB^2 = CD^2 - BC^2$. It is well known that this equation characterizes the quadrilaterals $ABCD$ with perpendicular diagonals. Thus, AC and BD are perpendicular.

(a) Suppose that the point P is the intersection of the diagonals. As we saw before, $PB^2 - PD^2 = AD^2 - AB^2$. By the Pythagorean theorem, we further have

$$PB^2 - PD^2 = AD^2 - AB^2 = (AP^2 + PD^2) - (AP^2 + PB^2) = PD^2 - PB^2,$$

and thus $PB = PD$. Similarly, we have $PA = PC$. Hence, we conclude that $ABCD$ is a rhombus.

(b) Suppose that there exists a point P such that (1) holds. As we saw before, we know that AC and BD are perpendicular. Actually, this is a sufficient condition. To show this, suppose that AC and BD are perpendicular. Let O be the intersection of AC and BD. We may assume that $AO \le OC$ and $DO \le OB$. Let O_1 and O_2 be points on $\overline{OC}$ and $\overline{OB}$, respectively, such that $AO = O_1C$ and $DO = O_2B$. Let P be a point such that OO_2PO_1 is a rectangle. Now we will show that this point P satisfies (1). By the Pythagorean theorem,

$$\begin{aligned} DP^2 - BP^2 &= DO_2^2 - BO_2^2 = BO^2 - DO^2 \\ &= (BO^2 + OC^2) - (DO^2 + OC^2) = BC^2 - CD^2. \end{aligned}$$

Hence,

$$\nu(\triangle PCD) = DP^2 + CD^2 + PC^2 = BP^2 + BC^2 + PC^2 = \nu(\triangle PBC).$$

Similarly, we have $\nu(\triangle PDA) = \nu(\triangle PCD)$ and $\nu(\triangle PAB) = \nu(\triangle PDA)$, and this completes the proof.

Also solved by Michel Bataille (France), Robert Calcaterra, Dmitry Fleischman (Part (a)), Ahmad Habil (Syria), Omran Kouba (Syria), Elias Lampakis (Greece), Yugeun Shim (Korea), Michael Vowe (Switzerland) (Part(a)), and the proposer. There was one incorrect submission.

A mean value theorem for integrals **April 2014**

1945. *Proposed by Michael W. Botsko, Saint Vincent College, Latrobe, PA.*

Let f be continuous on [0, 1] and let $\int_0^1 f(x)dx = 0$. In addition let $\phi(x)$ be differentiable on [0, 1] with $\phi(0) = 0$ and $\phi'(x) > 0$ on (0, 1). Prove that there exists $x_0 \in (0, 1)$ such that

$$\int_0^{x_0} \phi(x) f(x) dx = 0.$$

Solution by Kee-Wai Lau, Hong Kong, China.

For $x \in [0, 1]$, let $F(x) = \int_0^x f(t)dt$ and $G(x) = \int_0^x \phi(t) f(t)dt$. We have $F(0) = F(1) = 0$. Suppose, on the contrary, that $G(x)$ does not vanish on (0, 1). Replacing f by $-f$ if necessary, we may assume that $G(x) > 0$ on (0, 1). By continuity, we have $G(1) \geq 0$. Integrating by parts, we obtain

$$G(x) = F(x)\phi(x) - \int_0^x F(t)\phi'(t)dt \tag{1}$$

so that

$$\int_0^1 F(t)\phi'(t)dt = -G(1) \leq 0.$$

Since $\phi'(x) > 0$ on (0, 1) and by continuity of F, there exists $x_1 \in (0, 1)$ such that $F(x_1) \leq 0$, and $F(x_1) \leq F(x)$ for any $x \in [0, 1]$. By (1), we have

$$\begin{aligned} 0 < G(x_1) &= F(x_1)\phi(x_1) - \int_0^{x_1} F(t)\phi'(t)dt \\ &\leq F(x_1)\phi(x_1) - \int_0^{x_1} F(x_1)\phi'(t)dt \\ &= F(x_1)\phi(x_1) - F(x_1)(\phi(x_1) - \phi(0)) = 0, \end{aligned}$$

which is false. Thus, there exists $x_0 \in (0, 1)$ such that $G(x_0) = 0$ as required.

Editor's Note. Robert Calcaterra solved the problem using the weaker hypothesis of requiring ϕ to be nonnegative and strictly increasing on [0, 1], instead of ϕ differentiable with positive derivative on (0, 1).

Also solved by Robert Calcaterra, Daniel Fritze (Germany), Kunwoo Kim (Korea), Omran Kouba (Syria), Elias Lampakis (Greece), Moubinool Omarjee (France), Paolo Perfetti (Italy), Ángel Plaza (Spain) and José M. Pacheco (Spain), Achilleas Sinefakopoulos (Greece), Haohao Wang, and the proposer. There was one incorrect submission.

Answers

Solutions to the Quickies from pages 236–237.

A1051. Without loss of generality, we may assume that a_1 is positive and that the b_k are in ascending order. For $m = 0$, the result is true as all the terms in the sum are strictly positive. Suppose that the result is true for Dirichlet polynomials with less than m changes of sign among consecutive a_k. Suppose $m \geq 1$ and let a_j be the first coefficient that is negative. Consider the Dirichlet polynomial

$$f(x) := e^{-b_j x} D(x) = \sum a_k e^{(b_k - b_j)x}.$$

Because $e^{-b_j x}$ is strictly positive, it follows that f and D have the same number of real zeros. Note that

$$f'(x) = \sum a_k (b_k - b_j) e^{(b_k - b_j)x};$$

thus, the first $j - 1$ coefficients are negative, the jth coefficient is zero, and the rest of the signs remain unchanged. Hence, the derivative has $m - 1$ changes of sign, and then by induction it has at most $m - 1$ real zeros. If $f(x)$ were to have more than m distinct real zeros, then by Rolle's theorem, f' would have at least m zeros, which is a contradiction.

A1052. Let $M = (A - B)/2$ and $N = (A + B)/2$. Then

$$\int \frac{e^{Ax} - e^{Bx}}{e^{Ax} + e^{Bx}} dx = \int \frac{e^{Nx}(e^{Mx} - e^{-Mx})}{e^{Nx}(e^{Mx} + e^{-Mx})} dx = \int \frac{e^{Mx} - e^{-Mx}}{e^{Mx} + e^{-Mx}} dx$$
$$= \int \tanh(Mx) = \frac{1}{M} \ln(\cosh(Mx)) + C.$$

In part (a), we have $A = 20$ and $B = 15$. Thus,

$$\int \frac{e^{20x} - e^{15x}}{e^{20x} + e^{15x}} dx = \frac{2}{5} \ln\left(\cosh\left(\frac{5}{2}x\right)\right) + C.$$

MAA 100
MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

From the Files of Past MAGAZINE Editors

Allen Schwenk 2006–2008

To counter balance the stories about submissions of incorrect mathematics, half a dozen times former MAGAZINE editor Allen Schwenk received submissions from high school students who had independently discovered some mathematical result. He took great pains drafting his response to these students. Not wanting to crush their initiative with a brief rejection, he first would praise their discovery and declare it good mathematical research. Then he would point out two or three college textbooks that had their result as a theorem or an exercise. In one case, Schwenk was able to send a photocopy of his midterm exam from a graduate course that he took; the question was on the exam.

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.

The Imitation Game. Directed by Morten Tyldum. Produced by Black Bear Pictures and Bristol Automotive. U.S. distributor: The Weinstein Company. Academy Award for Best Writing, Adapted Screenplay. Film, 2014, 114 min. DVD $29.98, Blu-ray $34.99, rental and pay-per-view also available.

Hodges, Andrew, *Alan Turing: The Enigma*, Princeton University Press, 1983; reprint with new preface, 2014. Print or electronic version, xxxii + 736 pp, $16.95(P). ISBN 978-0-691-16472-4.

Anderson, L.V., How accurate is *The Imitation Game*?, http://www.slate.com/blogs/browbeat/2014/12/03/the_imitation_game_fact_vs_fiction_how_true_the_new_movie_is_to_alan_turing.html.

Christie, Thony, The Renaissance Mathematicus: Why *The Imitation Game* is a disaster for historians, https://thonyc.wordpress.com/2015/03/10/why-the-imitation-game-is-a-disaster-for-historians/.

We live amid contention about civil rights for gays and lesbians, and in Britain there has been growing interest in the wartime code-breaking at Bletchley Park, with the site itself being restored. So it is natural to make a film about Alan Turing, who did key work on breaking Enigma codes at Bletchley but later suffered under British law against homosexual acts. The film, great entertainment, is advertised as "based on a true story," the history in Hodges' biography, but it is by no means a historical re-enactment, being only very loosely based on history. Since the film is what students and colleagues will learn about Turing, it is important to be able to inform them where gross "liberties" have been taken (including falsehoods). You can do so without reading all of Hodges' work since Anderson compares the film to the book; Christie points to other critical reviews. Neither the film nor the reviews mention Turing's major contribution to computer science—not decoding Enigma, not building a computer, but conceptualizing computability.

Debakcsy, Dale, In defense of useless math: Why the new Common Core isn't as liberating as it seems, *Skeptical Inquirer* 39 (1) (January/February 2015) 36–37.

The author, a high school mathematics teacher, points out positive aspects of the Common Core mathematics curriculum then notes that "merely *interesting* math has been discreetly thrown off the pier to make room for profitable mathematics." The new answer to why study math? "Because it will make you rich." The effect is to turn "the most exciting branch of human knowledge" into "a series of trade skills." "I don't see any of it leading to a larger curiosity ... getting kids excited enough ... to learn about it independently."

Contenta, Sandro, The Canadian who reinvented mathematics, Toronto Star (28 March 2015) http://torontostar.newspaperdirect.com/epaper/viewer.aspx, http://projects.thestar.com/math-the-canadian-who-reinvented-mathematics/.

Biographical sketch of Robert Langlands (now age 78), whose conjectures regarding Galois groups and automorphic forms set out a monumental research program still continuing today.

Math. Mag. **88** (2015) 243–244. doi:10.4169/math.mag.88.3.243. © Mathematical Association of America

The Great Math Mystery. Written by Dan McCabe. Produced by WGBH Boston. U.S. distributor: Public Broadcasting System. Film, 2015, 53:10 min. DVD $24.99 from `www.shoppbs.org`; Internet pay-per-view $1.99 (standard definition), $2.99 (high definition). `http://www.pbs.org/wgbh/nova/physics/great-math-mystery.html`.

Wigner, E. P., The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics* 13 (1960) 1–14.

Abbott, Derek, The reasonable ineffectiveness of mathematics, *Proceedings of the IEEE* 101 (10) (October 2013) 2147–2153. `http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6600840`.

Mathematics: discovered or invented? This film presents the issue as a “mystery” concerning “civilization’s greatest achievement.” Topics introduced include the occurrence of the Fibonacci numbers in botany (without mention of the “golden ratio”!), the appearance of pi in unexpected connections (including Buffon’s needle experiment), the numeracy of animals, Galileo’s cannonball experiment as a prelude to landing a Mars rover, a re-creation of Marconi’s invention of radio waves, and the discovery of the Higgs boson (whose existence had been predicted from mathematics 50 years earlier). Wigner’s classic essay on the “unreasonable effectiveness of mathematics ” is highlighted, as is Abbott’s “reasonable ineffectiveness” (mathematics solves only problems based on “simplified models or reality”). Setting on its head Kronecker’s famous quote about the integers being God-given, Mario Livio is interviewed in the film, suggesting that humans invented the integers and then discovered relationships among them. So the film ends with the happy conclusion—which will not satisfy many philosophers of mathematics—that mathematics is both discovered and invented.

Boardman, Michael E., and Roger B. Nelsen, *College Calculus: A One-Term Course for Students with Previous Calculus Experience*, MAA, 2015; ix + 371 pp, $60 (members: $48). ISBN 978-1-93951-206-2.

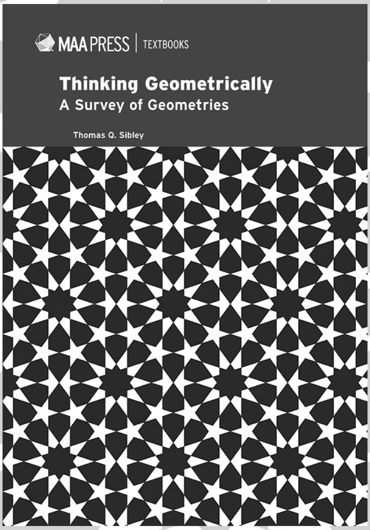
Bressoud, David, Calculus at crisis I: The pressures, `http://launchings.blogspot.com/`.

More students take calculus in high school than in college. Some take Advanced Placement (AP) calculus courses, either AB (first-semester) or BC (second-semester) to earn college credit and perhaps reduce college cost; others take ”college-placement calculus” to enhance their appeal to colleges, without any lasting interest in mathematics. Some colleges find that even students awarded college credit based on their AP score are poorly prepared for a succeeding calculus course, either because their understanding is weak, they are used to more intensive instruction, or the syllabus of the AP course does not match the supposedly corresponding college course. Authors Boardman and Nelsen, who have long been involved with AP calculus, note the overlaps between the AB syllabus and college Calculus II but also point out what is in the latter that AB students will not have seen. These latter topics are featured in their book, which is designed for students who received a 4 or a 5 on the AP Calculus AB exam (or a 4 or 5 calculus AB subscore on the BC exam). Unresolved is what should happen for students with a score of 3, who at many institutions receive credit (hence cannot retake Calculus I for credit) but may be ill-prepared for either Calculus II or the topics featured here. Meanwhile, David Bressoud, former president of the MAA, has started a new series of blog posts about new and continuing challenges in teaching calculus.

Kaplan, Daniel, *Start R in Calculus*, Project Mosaic, 2013; 81 pp, $16.58(P). ISBN 978-0-9839658-3.

What? Use a statistics language in calculus? R is a standard command-line language used by statisticians for calculation and graphics, but like other graphics and computer algebra systems, R can do more than statistics: It can plot functions, take derivatives and integrals in symbolic form, and solve ordinary differential equations. This booklet features RStudio, a GUI for R; both R and RStudio are available free for all platforms. There are short chapters with examples followed by exercises and endpapers that summarize the very compact set of commands used.

Save on textbooks
in the MAA eBooks Store



Thinking Geometrically
A Survey of Geometries

Thomas Q. Sibley

MAA Textbooks

Thinking Geometrically: A Survey of Geometries is a well written and comprehensive survey of college geometry that would serve a wide variety of courses for both mathematics majors and mathematics education majors.

Students with less mathematical preparation than upper-division mathematics majors can successfully study the topics needed for the preparation of high school teachers. There is a multitude of exercises and projects in those chapters developing all aspects of geometric thinking for these students as well as for more advanced students. These chapters include Euclidean Geometry, Axiomatic Systems and Models, Analytic Geometry, Transformational Geometry, and Symmetry. Topics in the other chapters, including Non-Euclidean Geometry, Projective Geometry, Finite Geometry, Differential Geometry, and Discrete Geometry, provide a broader view of geometry. The different chapters are as independent as possible, while the text still manages to highlight the many connections between topics.

The text is self-contained, including appendices with the material in Euclid's first book and a high school axiomatic system as well as Hilbert's axioms. Appendices give brief summaries of the parts of linear algebra and multivariable calculus needed for certain chapters. While some chapters use the language of groups, no prior experience with abstract algebra is presumed. The text will support an approach emphasizing dynamical geometry software without being tied to any particular software.

eISBN: 978-1-61444-619-4
2015, 584 pp

Price: $30.00

To order, visit maa.org/ebooks/TGSG

MAA 100
MATHEMATICAL ASSOCIATION OF AMERICA
CELEBRATING A CENTURY OF ADVANCING MATHEMATICS

MATHEMATICAL ASSOCIATION OF AMERICA
1529 Eighteenth St., NW · Washington, DC 20036

CONTENTS

ARTICLES

163 The Idea Man: An Interview with Lynn Steen *by Deanna Haunsperger and Stephen Kennedy*

177 Proof Without Words: Sums of Triangular Numbers *by Hasan Unal*

179 Selective Sums of an Infinite Series *by R. John Ferdinands*

186 Proof Without Words: An Elegant Property of the Equilateral Triangle *by Victor Oxman and Moshe Stupel*

187 The Average Height of Catalan Trees by Counting Lattice Paths *by Nachum Dershowitz and Christian Rinderknecht*

196 Some People Have All the Luck *by Richard Arratia, Skip Garibaldi, Lawrence Mower, and Philip B. Stark*

212 Anniversary Crossword: Editors Past and President, Part I *by Tracy Bennett*

214 Pythagoras via Cavalieri *by Zsolt Lengvárszky*

217 Optimal Defensive Strategies in One-Dimensional *RISK by Darren Glass and Todd Neller*

231 Non-Integrality of Binomial Sums and Fermat's Little Theorem *by Daniel López-Aguayo*

PROBLEMS AND SOLUTIONS

235 Proposals, 1971–1975

236 Quickies, 1051–1052

237 Solutions, 1941–1945

242 Answers, 1051–1052

REVIEWS

243 *The Imitation Game: not history; mathematics: unreasonable effectiveness vs. reasonable ineffectiveness; a statistics language for calculus?*