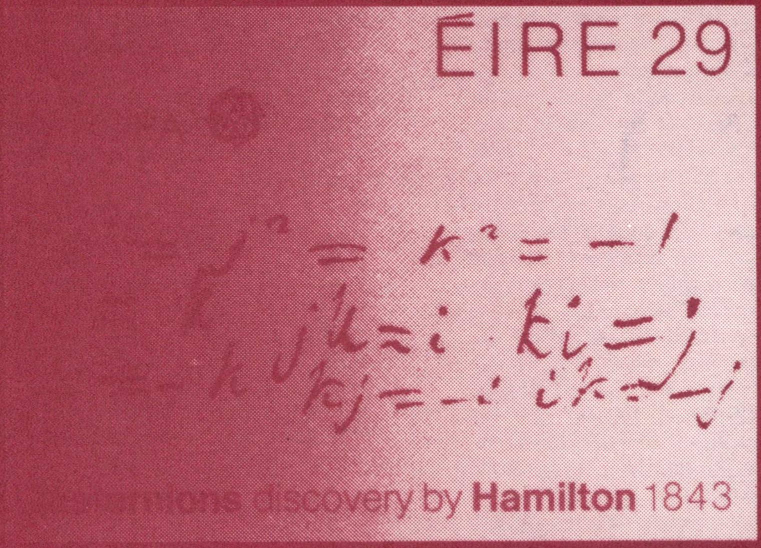
Vol. 62, No. 5 December 1989

# MATHEMATICS MAGAZINE




- Hamilton, Rodrigues, and Quaternions
- Sylvester's Four Point Problem
- Mathematics of the Casimir Effect

An Official Publication of The MATHEMATICAL ASSOCIATION OF AMERICA

## EDITORIAL POLICY

The aim of *Mathematics Magazine* is to provide lively and appealing mathematical exposition. This is not a research journal and, in general, the terse style appropriate for such a journal (lemma-theorem-proof-corollary) is not appropriate for an article 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. Articles on pedagogy alone, unaccompanied by interesting mathematics, are not suitable. Neither are articles consisting mainly of computer programs unless these are essential to the presentation of some good mathematics. Manuscripts on history are especially welcome, as are those showing relationships between various branches of mathematics and between mathematics and other disciplines.

The full statement of editorial policy appears in this *Magazine*, Vol. 54, pp. 44–45, and is available from the Editor. Manuscripts to be submitted should not be concurrently submitted to, accepted for publication by, nor published by another journal or publisher.

Send new manuscripts to: Martha Siegel, Editor-elect, Mathematics Magazine, Towson State University, Towson, MD 21204. Manuscripts should be typewritten and double spaced and prepared in a style consistent with the format of *Mathematics Magazine*. Authors should submit the original and one copy and keep one copy. Illustrations should be carefully prepared on separate sheets in black ink, the original without lettering and two copies with lettering added.

## AUTHORS

**Simon Altmann** was born and schooled in Buenos Aires but he did his Ph.D. at the Theoretical Physics Department of King's College, London. He was a postdoctoral instructor for five years at the Mathematical Institute in Oxford until he was appointed University Lecturer in the Theory of Metals in 1959. He is also a Fellow and Tutor in Mathematical Physics at Brasenose College. He has combined research on metals and semiconductors with applied group theory, on which he has written two books and a much xeroxed section of another. He stumbled into quaternions when trying to solve some intractable problems on the symmetrization of spinor wave functions required for solid-state calculations.

Vol. 62, No. 5 December 1989

# MATHEMATICS MAGAZINE

The *MATHEMATICS MAGAZINE* (ISSN 0025–570X) is published by the Mathematical Association of America at 1529 Eighteenth Street, N.W., Washington, D.C. 20036 and Montpelier, VT, bimonthly except July/August.
The annual subscription price for the MATHEMATICS MAGAZINE to an individual member of the Association is $13 included as part of the annual dues. (Annual dues for regular members, exclusive of annual subscription prices for MAA journals, are $29. Student, unemployed and emeritus members receive a 50% discount; new members receive a 30% dues discount for the first two years of membership.) The nonmember/library subscription price is $45 per year. Bulk subscriptions (5 or more copies) are available to colleges and universities for classroom distribution to undergraduate students at $5.40 per copy (minimum order 5 copies).
Subscription correspondence and notice of change of address should be sent to the Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036. Microfilmed issues may be obtained from University Microfilms International, Serials Bid Coordinator, 300 North Zeeb Road, Ann Arbor, MI 48106.
Advertising correspondence should be addressed to Ms. Elaine Pedreira, Advertising Manager, The Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20036.


Second class postage paid at Washington, D.C. and additional mailing offices.
Postmaster: Send address changes to Mathematics Magazine Membership/Subscriptions Department, Mathematical Association of America, 1529 Eighteenth Street, N.W., Washington, D.C. 20077-9564.
PRINTED IN THE UNITED STATES OF AMERICA

# ARTICLES

# Hamilton, Rodrigues, and the Quaternion Scandal

*What went wrong with one of the major mathematical discoveries of the nineteenth century**

SIMON L. ALTMANN
Brasenose College
Oxford OX1 4AJ, UK

Some of the best minds of the nineteenth century—and this was the century that saw the birth of modern mathematical physics—hailed the discovery of quaternions as just about the best thing since the invention of sliced bread. Thus James Clerk Maxwell, [**31**, p. 226], the discoverer of electromagnetic theory, wrote:

> The invention of the calculus of quaternions is a step towards the knowledge of quantities related to space which can only be compared, for its importance, with the invention of triple coördinates by Descartes. The ideas of this calculus, as distinguished from its operations and symbols, are fitted to be of the greatest use in all parts of science.

Not everybody, alas, was of the same mind, and some of the things said were pretty nasty:

> Quaternions came from Hamilton after his really good work had been done; and, though beautifully ingenious, have been an unmixed evil to those who have touched them in any way, including Clerk Maxwell. (Lord Kelvin, letter to Hayward, 1892; see [**38**, vol. II, p. 1070].)

Such robust language as Lord Kelvin's may now be largely forgotten, but the fact remains that the man in the street is strangely averse to using quaternions. Side by side with matrices and vectors, now the *lingua franca* of all physical scientists, quaternions appear to exude an air of nineteenth-century decay, as a rather unsuccessful species in the struggle-for-life of mathematical ideas. Mathematicians, admittedly, still keep a warm place in their hearts for the remarkable algebraic properties of quaternions, but such enthusiasm means little to the harder-headed physical scientist.

This article will attempt to highlight certain problems of interpretation as regards quaternions which may seriously have affected their progress, and which might explain their present parlous status. For claims were made for quaternions which quaternions could not possibly fulfil, and this made it difficult to grasp what quaternions are excellent at, which is handling rotations and double groups. It is

*This article follows closely material from Chapters 1 and 12 of *Rotations, Quaternions, and Double groups*, by Simon L. Altmann, Clarendon Press, Oxford, 1986.

essentially the relation between quaternions and rotations that will be explored in this paper and the reader interested in double groups will find this question fully discussed in my recent book [**1**].

## The Men Involved: Hamilton and Rodrigues

It is not possible to understand the quaternions' strange passage from glory to decay unless we look a little into the history of the subject, and the history of quaternions, more perhaps than that of any other nineteenth-century mathematical subject, is dominated by the extraordinary contrast of two personalities, the inventor of quaternions, Sir William Rowan Hamilton, Astronomer Royal of Ireland, and Olinde Rodrigues, one-time director of the Caisse Hypothécaire (a bank dedicated to lending money on mortgages) at the Rue Neuve-Saint-Augustin in Paris [**4**, p. 107].

Hamilton was a very great man indeed; his life is documented in minute detail in the three volumes of Graves [**15**]; and a whole issue in his honour was published in 1943, the centenary of quaternions, in the *Proceedings of the Royal Irish Academy*, vol. A50 and, in 1944 in *Scripta Mathematica*, vol. 10. There are also two excellent new biographies [**23**], [**33**] and numerous individual articles (see, e.g. [**26**]). Of Hamilton, we know the very minute of his birth, precisely midnight, between 3 and 4 August, 1805, in Dublin. Of Olinde Rodrigues, despite the excellent one-and-only published article on him by Jeremy Gray [**16**], we know next to nothing. He is given a mere one-page entry in the Michaud *Biographie Universelle* [**32**] as an 'economist and French reformer'. So little is he known, indeed, that Cartan [**6**, p. 57] invented a nonexistent collaborator of Rodrigues by the surname of Olinde, a mistake repeated by Temple [**37**, p. 68]. Booth [**4**] calls him *Rodrigue* throughout his book, and Wilson [**41**, p. 100] spells his name as *Rodriques*.

Nothing that Rodrigues did on the rotation group—and he did more than any man before him, or than any one would do for several decades afterwards— brought him undivided credit; and for much of his work he received no credit at all. This Invisible Man of the rotation group was probably born in Bordeaux on 16 October 1794, the son of a Jewish banker, and he was named Benjamin Olinde, although he never used his first name in later life. The family is often said to have been of Spanish origin, but the spelling of the family name rather suggests Portuguese descent (as indeed asserted by the *Enciclopedia Universal Illustrada Espasa-Calpe*). He studied mathematics at the École Normale, the École Polytechnique not being accessible to him owing to his Jewish extraction. He took his doctorate at the new University of Paris in 1816 with a thesis which contains the famous 'Rodrigues formula' for Legendre polynomials, for which he is mainly known [**14**].

The next 24 years or so until, out of the blue, he wrote the paper on rotations which we shall discuss later, are largely a blank as far as Rodrigues's mathematics is concerned. But he did lots of other things. The little that we know about Rodrigues relates to him mainly as a paranymph of Saint-Simon, the charismatic Utopian Socialist, whom he met in May, 1823, two months after Saint-Simon's attempted suicide. So, we read [**40**, p. 30] that the banker Rodrigues helped the poor victim in his illness and destitution, and supported him financially until his death in 1825. That Rodrigues must have been very well off we can surmise from Weill's reference to him as belonging to high banking circles, on a par with the wealthy Laffittes [**40**, p. 238]. After Saint-Simon died, with Rodrigues by his bedside, the latter shared the headship of the movement with Prosper Enfantin, an old friend and disciple of Saint-Simon. Thus he became *Père Olinde* for the acolytes. But the union did not last very long: in

1832 Rodrigues repudiated Enfantin's extreme views of sexual freedom and he proclaimed himself the apostle of Saint-Simonism. In August of that year he was charged with taking part in illegal meetings and outraging public morality, and was fined fifty francs [4]. Neither of the two early historians of Saint-Simonism, Booth and Weill, even mention that Rodrigues was a mathematician: the single reference to this is that in 1813 he was Enfantin's tutor in mathematics at the École Polytechnique. Indeed, all that we know about him in the year 1840 when he published his fundamental paper on the rotation group is that he was 'speculating at the Bourse' [4, p. 216].

Besides his extensive writings on social and political matters, Rodrigues published several pamphlets on the theory of banking and was influential in the development of the French railways. He died in Paris almost forgotten, however [32]. Even the date of his death is uncertain: 26 December 1850, according to the *Biographie Universelle* [32], or 17 December 1851, according to Larousse [27]. Sébastien Charléty [9, pp. 26, 294], although hardly touching upon Rodrigues in his authoritative history of Saint-Simonism, gives 1851 as the year of Rodrigues's death, a date which most modern references seem to favour.

Hamilton survived Rodrigues by fourteen years and had the pleasure, three months before his death in 1865, to see his name ranked as that of the greatest living scientist in the roll of the newly created Foreign Associates of the American National Academy of Sciences. And quite rightly so: his achievements had been immense by any standards. In comparison with Rodrigues, alas, he had been born with no more than a silver-plated spoon in his mouth: and the plating was tarnishing. When he was three the family had to park various children with relatives and William was sent to his uncle, the Rev. James Hamilton, who ran the diocesan school at Trim. That was an intellectually explosive association of child prodigy and eccentric pedant: at three William was scribbling in Hebrew and at seven he was said by an expert at Trinity College, Dublin, to have surpassed the standard in this language of many Fellowship candidates. At ten, he had mastered ten oriental languages, Chaldee, Syriac, and Sanscrit amongst them plus, of course, Latin and Greek and various European languages. This is, at least, the received wisdom on Hamilton and it may contain an element of legend: it is pretty clear, e.g., that his knowledge of German was not all that strong in later life and the veracity of the reports on these linguistic feats is disputed by O'Donnell [33]. Mathematics—if one does not count mental arithmetic, at which he was prodigious—came late but with a bang when, at seventeen, reading on his own Laplace's *Mécanique Céleste*, he found a mistake in it which he communicated to the President of the Irish Academy. His mathematical career was already set in 1823 when, still seventeen, he read a seminal paper on caustics before the Royal Irish Academy.

From then on Hamilton's career was meteoric: Astronomer Royal of Ireland at 22, when he still had to take two quarterly examinations as an undergraduate, knight at 30. Like Oersted, the Copenhagen pharmacist who had stirred the world in 1820 with his discovery of the electromagnetic interaction, Hamilton was a Kantian and a follower of the *Naturphilosophie* movement then popular in Central Europe. For Hamilton 'The design of physical science is... to learn the language and to interpret the oracles of the universe' (Lecture on Astronomy, 1831, see [15, vol. I, p. 501]). He discusses in 1835, prophetically (because of the later application of quaternions in relativity theory), "Algebra as the Science of Pure Time". He writes copiously both in prose and in stilted verse, engages in a life-long friendship with Wordsworth, and goes to Highgate in the spring of 1832 to meet Coleridge, whom he visits and with whom he corresponds regularly in the next few years, the poet praising him for his understanding 'that Science...needs a Baptism, a regeneration in Philosophy' or Theosophy [**15**, vol. I, p. 546].

## The Discovery of Quaternions

Hamilton had been interested in complex numbers since the early 1830's and he was the first to show, in 1833, that they form an algebra of couples. (See [**22**, vol. III].) I shall review briefly his ideas so as to lead the way to quaternions, but, here and hereafter, I shall use my own notation in order to avoid ambiguities. First, we define imaginary units, $\mathbf{1}$ and $\mathbf{i}$ with the well-known multiplication rules in TABLE 1. Then the elements of the algebra are the complex numbers $\mathbb{A} = a\mathbf{1} + A\mathbf{i}$, with $a$ and $A$ real.

TABLE 1. Multiplication table of the imaginary units.

| | $\mathbf{1}$ | $\mathbf{i}$ |
|---|---|---|
| $\mathbf{1}$ | $\mathbf{1}$ | $\mathbf{i}$ |
| $\mathbf{i}$ | $\mathbf{i}$ | $-\mathbf{1}$ |

Of course, to say that they form an algebra merely means that the formal rules of arithmetical operations are valid for the objects so defined. Thus, given $\mathbb{A}$ and a similarly defined $\mathbb{B}$, their product is

$$\mathbb{A}\mathbb{B} = ab - AB + \mathbf{i}(aB + bA). \tag{1}$$

We can now write the complex numbers $\mathbb{A}$ and $\mathbb{B}$ as *couples* (or *ordered pairs*)

$$\mathbb{A} = [\![a, A]\!], \qquad \mathbb{B} = [\![b, B]\!], \tag{2}$$

and their product is also a couple:

$$\mathbb{A}\mathbb{B} = [\![ab - AB, aB + bA]\!]. \tag{3}$$

Hamilton also recognized that the real number $a$ can be written as the complex couple

$$a = [\![a, 0]\!]. \tag{4}$$

For the next ten years Hamilton's mind was occupied, if not obsessed, with two problems. On the one hand, Hamilton tried to extend the concept of the complex number as a couple in order to define a triple, with one real and two imaginary units. This however, not even he could do. On the other hand, the concept of a vector was beginning to form in his mind. It must be remembered that in the 1830's not even the word *vector* existed, although people were playing about, in describing forces and such other quantities, with concepts that we would recognize today as at least vector like. It is pretty clear that during this gestation period, as a result of which Hamilton would eventually invent the notion of vector, he had built up in his mind a picture of the addition and of some form of multiplication of vectors, but there was an operation which baffled him in the extreme: coming down the stairs for breakfast, Hamilton often could hear his elder son asking: 'Father, have you now learned how to divide vectors?'. Out of this preoccupation Hamilton was to invent the most beautiful algebra of the century, but he was also to feed the fever that eventually led him to corrupt his own invention.

We must now come to Monday, 16 October 1843, one of the best documented days in the history of mathematics and which, by one of those ironies of fate, happened to be the 49th birthday of Olinde Rodrigues, whose work, however ignored, was to give a new meaning to Hamilton's creation. Hamilton's letter to his youngest son, of 5 August 1865 [**15**, vol. II, p. 434], is almost too well known, but bears brief repetition. The morning of that day Hamilton, accompanied by Lady Hamilton, was walking along the Royal Canal in Dublin towards the Royal Irish Academy, where Hamilton was to preside at a meeting. As he was walking past Broome Bridge (referred to as Brougham Bridge by Hamilton and called by this name ever since), Hamilton, in a flash of inspiration, realized that three, rather than two, imaginary units were needed, with the following properties:

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1, \qquad \mathbf{ij} = \mathbf{k}, \qquad \mathbf{ji} = -\mathbf{k}, \tag{5}$$

and cyclic permutation. As everyone knows, and de Valera was to do almost one century later on his prison wall, Hamilton carved these formulae on the stone of the bridge: poor Lady Hamilton had to wait. Armed now with four units, Hamilton called the number

$$\mathbb{A} = a1 + A_x\mathbf{i} + A_y\mathbf{j} + A_z\mathbf{k}, \tag{6}$$

where the coefficients here are all real, a *quaternion*. Thus were quaternions born and baptized: it was entered on the Council Books of the Academy for that day that Mr. W. R. Hamilton was given leave to read a paper on quaternions at the First General Meeting of the Session, 13 November 1843.

One of the various falsehoods which have to be dispelled about quaternions is the origin of their name, since entirely unsupported sources are often quoted, in particular Milton, *Paradise Lost*, vol. 181 [**28**, p. 70] and the *Vulgate*, Acts 12:4 [**37**, p. 46]. Of course, we know that Milton was a favourite poet of Hamilton at 24 [**15**, vol. I, p. 321], and to suggest that he was not aware of *Acts* and the apprehension of Peter by a quaternion of soldiers would be absurd. (These references appear in fact in Dr. Johnson's *Dictionary*, which was familiar to every schoolboy of the time.) But no one with the slightest acquaintance with Hamilton's thought would accept the obvious when the recondite will do. In his *Elements of Quaternions* [**21**, p. 114]) we find our first clue: 'As to the *mere word*, *quaternion*, it signifies primarily (as is well known), like its Latin original, "Quaternio" or the Greek noun τετραχτυς, a Set of Four'. The key word here is 'tetractys,' and there is evidence for this coming from Hamilton's closest, perhaps his only real pupil, P. G. Tait, who, writing in the *Encyclopaedia Britannica* (see article on Quaternions in the XIth edition) says: 'Sir W. R. Hamilton was probably influenced by the recollection of its Greek equivalent the Pythagorean Tetractys..., the mystic source of all things...'. That Tait very much believed in this is supported by the unattributed epigraph in Greek in the title page of his own treatise on quaternions [**36**]: these are verses 47 and 48 of *Carmen Aureum* (*Golden Song*), a Hellenistic Pythagorean poem much in vogue in the Augustan era, the full text of which appears in Diehl [**11**, p. 45]. Of course, the concept of the tetractys embodying, as we shall see, multiple layers of meaning in a single word, must have attracted Hamilton: for Pythagoras, having discovered that the intervals of Greek music are given by the ratios 1 : 2. 3 : 2. 4 : 3 made it appear that *kosmos*, that is, order and beauty, flow from the first four digits, 1, 2, 3, 4, the sum of which gives the perfect number 10, and is symbolized by the sacred symbol, the tetractys:

0
0 0
0 0 0
0 0 0 0

(A famous depiction of the tetractys can be seen in the *School of Athens*, the fresco by Raphael at the Vatican where, anachronistically, the sacred symbol is given in Latin numerals in the figure held in front of Pythagoras.) The Pythagoreans used to take an oath by the tetractys, as recorded by Sextus Empiricus (see [**24**, p. 233]): 'The Pythagoreans are accustomed sometimes to say "All things are like number" and sometimes to swear this most potent oath: 'Nay, by him that gave to us the *tetractys*, which contains the fount and root of ever-flowing nature.' That the tetractys exercised the imagination of Hamilton, there is no doubt: besides the cryptic footnote in the *Elements*, already quoted, we find Augustus DeMorgan (with whom Hamilton entertained a very copious correspondence) acknowledging on 27 December 1851 a sonnet from Hamilton (apparently lost) on the tetractys. It is tempting to speculate that Hamilton might have been introduced to the tetractys by Coleridge, who called it 'the adorable tetractys, or tetrad' (see [**2**, p. 252]) and who referred to it many times.

## In Praise of Hamilton: the Algebra of Quaternions

In comparison with the binary form (2) of a complex number, the quaternion (6) can also be written as a couple of a real number $a$ and a vector $\mathbf{A}$ of components $A_x$, $A_y$, $A_z$, (as already said, we use modern rather than historical notation):

$$\mathbb{A} = [\![a, \mathbf{A}]\!], \quad \mathbf{A} = (A_x, A_y, A_z). \tag{7}$$

Just as for the complex numbers, in order to multiply two such objects we need the multiplication table of the quaternion units, which follows at once from eqn (5) and it is given in Table 2. Consider now a second quaternion $\mathbb{B}$, write both $\mathbb{A}$ and $\mathbb{B}$ as in eqn (6) and, on using the table, their product follows at once in the same manner as that in (3):

$$\mathbb{A}\mathbb{B} = [\![ab - \mathbf{A}.\mathbf{B}, a\mathbf{B} + b\mathbf{A} + \mathbf{A} \times \mathbf{B}]\!]. \tag{8}$$

Although Hamilton did not give names or symbols for these operations, it is here that the scalar and vector products of two vectors appear for the first time in history. We can now go back to TABLE 2 and reflect a little about why Hamilton made the product $\mathbf{ij}$ noncommutative. Not only was this the first time that a noncommutative product appeared in mathematics, but this was a true stroke of genius. Remember that Hamilton wanted to divide vectors: he never really achieved this (neither was it worth trying) but the point is that, because of this, he was after a *division algebra*, i.e., one in which the quotient of an element of the algebra by any other nonnull element always exists. Now, a necessary condition for a division algebra is this: the product of two elements of the algebra must vanish if and only if one of the factors is the null element. TABLE 2 is designed so that this happens, as can easily be verified from the resulting multiplication rule, given by (8). A counterexample will be instructive. Suppose we take $\mathbf{ij}$ and $\mathbf{ji}$ as equal in TABLE 2, and, similarly, for the other products. Then, under the new multiplication rules, it is very easy to verify that the product of the two nonnull elements

$$\mathbb{A} = \mathsf{i} + \mathsf{j}, \qquad \mathbb{B} = -\mathsf{i} + \mathsf{j}, \tag{9}$$

vanishes.

TABLE 2. Multiplication table of the quaternion units.

| | 1 | i | j | k |
|---|---|---|---|---|
| 1 | 1 | i | j | k |
| i | i | −1 | k | −j |
| j | j | −k | −1 | i |
| k | k | j | −i | −1 |

Hamilton's everlasting monument (see [30]) is his construction of objects which, except for commutativity, obey the same algebra as that of the real and complex numbers and which therefore, like them, form a division algebra: and Hamilton was aware of this—although he could not foresee that his brain-child was going to receive at the hands of Frobenius in 1878 the supreme accolade of being proved to be the only possible algebra, in addition to the real and complex numbers, with this property.

## The Trouble Starts

Now back to 16 October 1843. That same evening Hamilton wrote a long and detailed draft of a letter to his friend John Graves, first published by A. J. McConnell [**29**] and included in Hamilton [**22**, vol. III]. Next day a final letter was written and sent, later published in the *Philosophical Magazine* [**17**]. The November report to the Irish Academy was published almost at the same time [**18**]. We can thus follow almost hour by hour Hamilton's first thoughts on quaternions. Although in the morning of the glorious day he had been led to the discovery through the algebra of the quaternions, by the evening (and in this he acknowledges the influence of Warren [**39**]), he had been able to recognize a relation between quaternions and what we now call rotations. And in this, sadly, we cannot but see the germ of the canker that eventually consumed the quaternion body. Three separate themes, ever present in Hamilton's mind, contributed to this infirmity.

As regards the first theme: as in (4), Hamilton identified a real number with a real quaternion:

$$a = [\![a, \mathbf{0}]\!]. \tag{10}$$

Nothing wrong here, but it invited Hamilton to go on and identify a *pure quaternion* (a quaternion with a null scalar part) with a *vector*, a word which he invented for this purpose in 1846 [**19**, p. 54]:

$$\mathbf{A} = [\![0, \mathbf{A}]\!]. \tag{11}$$

As the inventor of the vector he was entitled to call this object anything he wanted but the problem is that by this time people were already thinking about forces and such like objects very much as we think of vectors today and that the identification of Hamilton's 'vectors' with what they had in mind created a great deal of confusion. The apparently innocent convention (11) entails in fact two serious problems. That something here was a serious worry must have been evident for decades, as Klein [**25**, p. 186], one of the leading nineteenth-century geometers, implied himself. Yet, the first explicit statement to the effect that something is wrong here which I have been able to find is as recent as 1958 by Marcel Riesz [**34**, p. 21]: 'Hamilton and his school professed that the quaternions make the study of vectors in three-space unnecessary since every vector can be considered as the vectorial part...of a quaternion...this interpretation is grossly incorrect since the vectorial part of a quaternion behaves with respect to coordinate transformations like a bivector or "axial" vector and not like an ordinary or "polar" vector.' However damning this statement is, it is only half the story, since the pure quaternion (11) is not anything like a vector at all: we shall see that it is a *binary rotation*, that is a rotation by $\pi$. The left-hand side of (11) should be written as $\mathbb{A}$ and carefully distinguished from the vector $\mathbf{A}$. The fact that neither Hamilton, nor his successors to the present day, introduced any notational distinction between these two objects is the source of extraordinary confusion, as we shall soon witness ourselves.

Hamilton's second theme was closely connected to his first and has already been mentioned: he wanted to find a definition of the quotient of two vectors and however grateful we must be for this obsession, which has given us the last possible division algebra, we shall soon see that it led Hamilton to an interpretation of quaternions and of their operations which is not right.

Hamilton's capacious mind could not be at rest until he understood not just the formalities of his work but also what went on behind the scenes, and he had to understand the physical or geometrical meaning of equating the square of the imaginary or quaternion unit, $\mathbf{i}^2$, with $-1$. This was his third everlasting theme, for which he took a cue from Argand, who had observed in 1806 that the imaginary unit $\mathbf{i}$ rotates what we would now call a vector in the Argand plane by $\pi/2$, which made it possible to visualize the relation in question. From that point of view, in fact, $\mathbf{i}^2$ should be a rotation by $\pi$ which, duly enough, multiplies each vector of the plane by the factor $-1$. For this reason, Hamilton always identified the quaternion units with *quadrantal rotations*, as he called the rotations by $\pi/2$ (see [**20**, p. 64, art. 71]). Clifford [**10**, p. 351] associates himself with this interpretation which he presents with beautiful clarity. The sad truth is that, however appealing this argument is, to identify the quaternion units with rotations by $\pi/2$ is not only not right, but it is entirely unacceptable in the study of the rotation group: we shall see, in fact, that they are nothing else except binary rotations.

## Quaternions and Rotations: the First Steps

Already during the first day of his creation Hamilton knew what he had to do in order to define rotations. Since rotations must leave the lengths of vectors invariant, and since for Hamilton a vector was a particular case of a quaternion, the first thing we had to do is to define the *norm* or length $|\mathbb{A}|$ of a quaternion. He defined for this purpose the *conjugate quaternion*

$$[\![a,\mathbf{A}]\!]^* = [\![a,-\mathbf{A}]\!]. \tag{12}$$

The norm is now defined as follows:

$$\mathbb{A}\mathbb{A}^* = [\![a,\mathbf{A}]\!][\![a,-\mathbf{A}]\!] = [\![a^2+A^2,0]\!] = a^2+A^2 = |\mathbb{A}|^2. \tag{13}$$

A quaternion of unit norm is called *a normalized quaternion* and, although Hamilton also considered more general quaternions, these will be the only quaternions which we shall need for the purposes of this article. It was easy for Hamilton to prove that *the product of two normalized quaternions is a normalized quaternion*. (See, e.g. [**1**, p. 208].) We are now ready to accompany Hamilton in performing an extraordinary piece of legerdemain.

## An Optical Illusion: the Rectangular Rotation

A rotation is an operation which transforms a unit, that is normalized, position vector (a vector with its tail at the origin), **r**, into another unit position vector **r**′. If we go along with Hamilton and identify the vector **r** with the pure quaternion $\mathbb{R}$ equal to $[\![0, \mathbf{r}]\!]$, the latter is clearly normalized. In order to achieve a rotation and keep the normalization requirement of $\mathbb{R}$, all that we need is to act on $[\![0, \mathbf{r}]\!]$ with a normalized quaternion. (See the italicized statement in the previous paragraph.) Let us, therefore, choose for this purpose the quaternion

$$\mathbb{A} = [\![\cos\alpha, \sin\alpha\mathbf{n}]\!], \qquad |\mathbf{n}| = 1, \tag{14}$$

which is clearly normalized. (Here **n** is a unit position vector.) This is not the end of the story, however, because we must require that the product $\mathbb{A}\mathbb{R}$ be not only normalized, but also a pure quaternion $\mathbb{R}'$ of the form $[\![0, \mathbf{r}']\!]$ which Hamilton would identify with the rotated position vector **r**′. This is what Hamilton envisioned on the same evening of Creation Day, and he also realized that for this idea to work it was necessary that the vector **n**, which he called the *axis of the quaternion*, be normal to the vector **r**. (This is why this is called the *rectangular transformation*.) To verify that this works is child's play on using the quaternion multiplication rule (8). Given that $\mathbf{r}, \mathbf{n} = 0$, then

$$\begin{aligned}\mathbb{A}\mathbb{R} &= [\![\cos\alpha, \sin\alpha\mathbf{n}]\!][\![0, \mathbf{r}]\!] \\ &= [\![0, \cos\alpha\mathbf{r} + \sin\alpha(\mathbf{n}\times\mathbf{r})]\!] = [\![0, \mathbf{r}']\!] = \mathbb{R}'.\end{aligned} \tag{15}$$

If we briefly avert our gaze while Hamilton rewrites this equation as

$$\mathbb{A}\mathbf{r} = \mathbf{r}', \tag{16}$$

then the job is done, that is, the quaternion $\mathbb{A}$ transforms the unit position vector **r** into another unit position vector **r**′ and, therefore, has rotated **r** into **r**′. What is more: it is clear from FIGURE 1 that the angle of rotation is $\alpha$. Thus Hamilton identified the

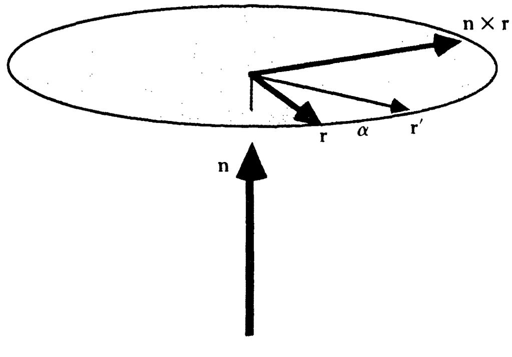



FIGURE 1
**The rectangular transformation. The vector r′ is defined in eqn (15).**

quaternion (14) with a rotation around the axis **n** of the quaternion by the angle $\alpha$ of rotation of the quaternion.

All this is so wonderfully convincing that it is difficult to believe that there is anything wrong here. Moreover, Hamilton immediately obtained confirmation that one of his themes was coming out all right, for it was clear to him that the quaternion units in this picture are quadrantal rotations, as it is immediate from (14): on comparing this equation with, say, the quaternion unit **k**, given by $[\![0, \mathbf{k}]\!]$, one can see at once that the rotation angle to be associated with **k** must be $\pi/2$. His other theme was also coming out well here, since from (16) the quaternion $\mathbb{A}$ can be considered as the quotient of the vector $\mathbf{r}'$ by the vector $\mathbf{r}$. This picture of quaternions was thus so near Hamilton's heart that in his *Lectures on Quaternions* [**20**, p. 122], and ever after, the primary definition of a quaternion which he used was 'The quotient of two vectors, or the operator which changes one vector into another,' as later adopted by the *Oxford English Dictionary*, and this definition became the core of the quaternion dogma, thus causing endless damage. We shall see, in fact, that a quaternion can never operate on a vector, as (16) implies, and that this equation must always be understood as the quaternion product in (15).

## The Comical Transformation (this heading contains a misprint)

The conical transformation was the means by which nature began to make its protest against Hamilton. Even accepting that a pure quaternion *is* a vector, we must ensure, in order to have a rotation, that the transform of a normalized pure quaternion is another normalized pure quaternion. Thus, a general rotation cannot be written as $\mathbb{A}\mathbb{R}$ because this product is not a pure quaternion unless, as we have seen, the axis of $\mathbb{A}$ is normal to **r**. Hamilton and his colleagues, therefore, searched for a quaternion transformation of a pure quaternion $\mathbb{R}$ under a normalized quaternion $\mathbb{A}$ which would always produce a normalized pure quaternion $\mathbb{R}'$. The result of this search was the following transformation:

$$\mathbb{A}\mathbb{R}\mathbb{A}^* = \mathbb{R}'. \tag{17}$$

It is, in fact, quite easy by means of (8) to verify that the left-hand side of this equation is normalized and pure. With a little bit of geometry (see [**1**, p. 214]), and assuming that $\mathbb{A}$ is given by (14), it can be proved that **r**, **n**, and $\mathbf{r}'$ are related as shown in FIGURE 2, i.e., that the vector **r** is rotated around **n** by the angle $2\alpha$.

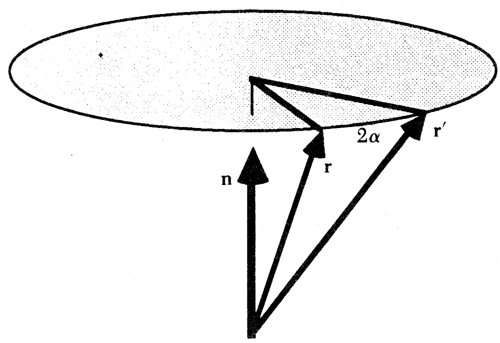



FIGURE 2
**The conical transformation.**

There are two problems here: first, the form of (17) has nothing at all to do with that of (15), so that it is no longer possible to say that the quaternion operates on a vector transforming it into another vector; even less that it is the quotient of two vectors. The second problem is this: which is the angle of rotation to be associated with a quaternion (14), $\alpha$, as in the rectangular transformation, or $2\alpha$, as it turns out in the case of a general rotation? It is, perhaps, significant that Hamilton obtained (17) (writing the conjugate quaternion as the reciprocal, as it is valid for a normalized quaternion) but did not publish it for some time. Cayley [7] was the first to go into print, although in his collected papers [8, vol. I, p. 586, note 20] he concedes priority to Hamilton. Cayley notices that the components of $\mathbf{r}'$ are 'precisely those given for such transformations by M. Olinde Rodrigues.... It would be an interesting question to account, *a priori* for the appearance of these coefficients here'. Let us see what Hamilton has to say about this: 'The SYMBOL OF OPERATION $q(\ )q^{-1}$, where $q$ may be called (as before) the operator quaternion, while the symbol (suppose $r$) of the operand quaternion is conceived to occupy the place marked by the parentheses...'can be regarded as' a conical transformation of the operand round the axis of the operator, through *double the angle thereof.*' [**20**, p. 271, my italics]. It is clear that Hamilton, rather than accepting the result of the more general transformation (17) to recognize that the angle of rotation of the quaternion (14) is $2\alpha$, gives greater weight to the transformation (16) and keeps talking of the angle $\alpha$ in (14) as the angle of rotation. Naturally, whereas (16) had the shape that he expected, the form of (17), as Cayley stated, could not be explained. It is, perhaps, because of this that, although FIGURE 2 is nothing else than the most general rotation of a vector, Hamilton refers to it with the *ad hoc* name of *conical rotation*, as if it were a particular case of the transformation of a vector. It had, instead, been Rodrigues who had recognized, three years before Hamilton's invention of quaternions, that the angle $\alpha$ in (14) is not the rotation angle but only half of it. But his paper, which had puzzled Cayley, was almost certainly never read by Hamilton and it was never again quoted by any of the major quaternionists. As for Cayley's question, it was probably never answered until 1986. (See [**1**, p. 214].)

## The Rodrigues Programme

Hamilton constructed quaternions as an algebra, whence the elements of the algebra were given a dual role as operators (rotations) and operands (vectors). This was very lucidly explained by Clifford [**10**], but it must be clearly appreciated that, as we have already asserted, the status of vectors in this scheme is highly dubious, of which more later. Be that as it may, in Hamilton's approach rotations become subservient to the algebra, which opens the door to a variety of misinterpretations.

Historically, however, a treatment of rotations and quaternions had been going on for some years before 1843, quite independently of Hamilton and taking a diametrically opposed view to his. This treatment was entirely geometrical, and because it tried to do a simple job in a simple way it was clear and precise and it was entirely successful; but it was largely ignored by everyone.

Let us consider rotations of a unit sphere with fixed centre about various axes. The first problem which arises is whether, if we apply one rotation after another, the net result is a rotation of the sphere around some unique axis by some unique angle. Euler [**12**] proved algebraically that this is so, but he did not provide either a geometric or a constructive solution (i.e., a solution in which the axis and angle of the resultant rotation are determined geometrically or algebraically). It was the paper by Rodrigues in 1840 [35] which solved all these aspects of the problem. In its § 8 he describes most clearly, without a figure, a geometrical construction which, given the angles and axes of two successive rotations, determines the orientation of the resultant axis of rotation and the geometrical value of the angle of rotation. This construction is usually called in the literature the *Euler construction*, although Euler had nothing to do with it. Not only was this construction ignored by the quaternionists but it is not even mentioned in modern books on the rotation group. Although Hamilton himself rediscovered, geometrically [**20**, p. 328], the results of the Rodrigues construction, this is not a theorem to which either he or his commentators paid much attention (see [**1**, pp. 19–20]).

Let us represent a rotation around the axis $\mathbf{p}$ by the angle $\phi$ with the symbol $R(\phi\mathbf{p})$. Then, if we use the Rodrigues construction for the following product of rotations,

$$R(\alpha\mathbf{l})R(\beta\mathbf{m}) = R(\gamma\mathbf{n}), \tag{18}$$

it turns out that the axes $\mathbf{l}$, $\mathbf{m}$, and $\mathbf{n}$ form a spherical triangle with the angles shown in FIGURE 3. (Remember that in the left-hand side of (18) the rotation around $\mathbf{m}$ is applied first and it is followed by that around $\mathbf{l}$: this is the usual convention for reading operators.)

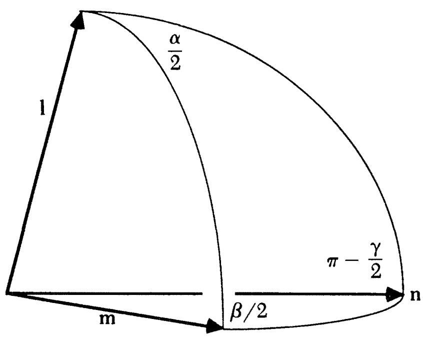



FIGURE 3
The product of the rotations $R(\alpha\mathbf{l})$ and $R(\beta\mathbf{m})$ is the rotation $R(\gamma\mathbf{n})$.

What is very remarkable about this very simple triangle is that the angles of the rotations appear in it as *half-angles*, and this is the first time that half-angles occur in the study of rotations. Their importance is absolutely crucial, as we shall see, and yet they were ignored by Euler and were never considered by Hamilton or his followers: it took more than forty years before their significance was appreciated. One would shirk nowadays at the solution of the spherical triangle in FIGURE 3 (see [**1**, p. 157]) but mathematical training in France on surveying and such was very good and Rodrigues was able to obtain quite easily expressions for the angle and axis of the resultant rotation, that is, the one on the right-hand side of (18), in terms of those of the factors which appear on its left. The following are Rodrigues's formulae exactly as he gave them, except that I have introduced vector notation, which was nonexistent in his time:

$$\cos\frac{\gamma}{2} = \cos\frac{\alpha}{2}\cos\frac{\beta}{2} - \sin\frac{\alpha}{2}\sin\frac{\beta}{2}\mathbf{l}.\mathbf{m}, \tag{19}$$

$$\sin\frac{\gamma}{2}\mathbf{n} = \sin\frac{\alpha}{2}\cos\frac{\beta}{2}\mathbf{l} + \cos\frac{\alpha}{2}\sin\frac{\beta}{2}\mathbf{m} + \sin\frac{\alpha}{2}\sin\frac{\beta}{2}\mathbf{l}\times\mathbf{m}. \tag{20}$$

These formulae immediately suggest that a rotation $R(\alpha\mathbf{l})$ can be represented by a couple of a scalar and a vector (although this notation was not used by Rodrigues)

$$R(\alpha\mathbf{l}) = [\![\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2}\mathbf{l}]\!], \tag{21}$$

so that the product (18) is written as follows

$$[\![\cos\frac{\alpha}{2}, \sin\frac{\alpha}{2}\mathbf{l}]\!][\![\cos\frac{\beta}{2}, \sin\frac{\beta}{2}\mathbf{m}]\!] = [\![\cos\frac{\gamma}{2}, \sin\frac{\gamma}{2}\mathbf{n}]\!], \tag{22}$$

with the parameters on the right-hand side of this equation being given by (19) and (20). It can immediately be seen that the multiplication rule for the couples so defined is identical with the multiplication rule (8) of Hamilton's quaternions! Rodrigues's couples are, therefore, quaternions, but the difference in parametrization between (21) and (14) is profound. We see at once that the conical transformation, which gives the angle of rotation as twice the angle which appears in the quaternion was right, and that Hamilton committed a serious error of judgement in basing his parametrization on the special case of the rectangular transformation.

Simple as the distinction is, the consequences are dramatic, and never more so than when we consider pure quaternions. From (21), it is clear that for a quaternion to be pure the angle of rotation must be $\pi$, that is, a pure quaternion is nothing other than a binary rotation:

$$[\![0, \mathbf{r}]\!] = R(\pi\mathbf{r}). \tag{23}$$

Thus, it is entirely wrong ever to identify a pure quaternion with a vector, as Hamilton had done in (11). This simple fact will exorcise all the demons so far lurking into our story. Before we do this, we must mention that the (four) rotation parameters in (21) are called the Euler-Rodrigues parameters in the literature. The reasons for this are entirely disreputable (see [1, p. 20]), since Euler never came near them: in particular, he never used half-angles which, as demonstrated by Rodrigues, are an essential feature of the parametrization of rotations.

## The Resolution of the Paradoxes

Although quaternions are always rotations and never vectors, they allow us to mark points in space very much as a position vector does. Consider the unit sphere centered and fixed at the origin. A rotation of it determines a single point of the sphere which is called the *pole of the rotation*. This is the point of the sphere which is left invariant by the rotation and such that from outside it the rotation is seen as positive (counterclockwise). If we want to mark a point in space by means of rotation poles, it is sensible to use always binary rotations for this purpose, since, as it follows from (23) these are the nearest things to vectors that we can get within the quaternion algebra. (It should be stressed that this is purely a matter of convenience: the pole of any arbitrary rotation is just as good to denote a point of the unit sphere and thus to masquerade as a vector.) Let us now look again at (15) with the quaternion $\mathbb{A}$ in it given by

$$\mathbb{A}\mathbb{R} = \mathbb{R}', \qquad \mathbb{A} = [\![\cos\alpha, \sin\alpha\mathbf{n}]\!]. \tag{24}$$

If we compare the left-hand side of this line with (18) and (21) it says this: the product of a rotation by $2\alpha$ around the axis $\mathbf{n}$, times a binary rotation around the axis r, is a binary rotation around the axis r′:

$$R(2\alpha\mathbf{n})R(\pi\mathbf{r}) = R(\pi\mathbf{r}'). \tag{25}$$

The angle of rotation, which we already know, and the orientation of the axis r′ can, of course, be obtained from the quaternion multiplication rules, as given, e.g., in (19) and (20), but it will be instructive to obtain an independent geometrical verification, since this will show up the paradox involved in Hamilton's interpretation of the rectangular transformation. We do this in FIGURE 4. In order to multiply rotations we transform the unit sphere, whose intersection with the plane of the drawing is shown in FIGURE 4. The rotation axis n is perpendicular to and above the plane of the drawing. The rotation axis r is in the plane of the drawing and thus, as it must be in the rectangular transformation, is normal to n. A point above the plane of the drawing is represented with a cross and those below with a circle. In order to multiply the two rotations on the left of (25) we start with point 1 above the plane of the drawing. The first rotation to act on it (remember to read the left-hand side of (25) from right to left) is a rotation by $\pi$ around r which takes it into the point 2 below the plane of the drawing. The rotation around n by $2\alpha$ takes 2 to 3. Thus, the two combined operations take the point 1 above the drawing to the point 3 below the drawing, which is the effect of a binary rotation around the axis r′. Notice that the angle between the axes r and r′ is $\alpha$ and that this angle *is not the angle of rotation*. We can now see how Hamilton's optical illusion was performed. If in (24) we identify the quaternions $\mathbb{R}$ and $\mathbb{R}'$ with their corresponding vectors r and r′, FIGURE 4 now reads as the rotation of r into r′ by $\alpha$. Incidentally, the correct reading of FIGURE 4 as stating that a rotation axis by $2\alpha$ and a perpendicular binary axis determine another perpendicular binary axis at an angle $\alpha$ to the first one is so fundamental in crystallography that the whole of this science would collapse like a pack of cards if it were not true.

What about the conical transformation? I cannot go into all the details of the theory but a sketch will suffice. We must accept the following result (see [1, p. 215]). If we take a pole of a binary rotation r and we rotate this pole about an axis n by an angle $\alpha$, the new pole thus obtained, r′, is the pole of another binary rotation given in the following form:

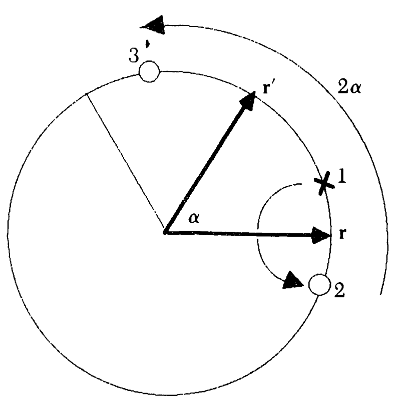



FIGURE 4
Product of a rotation by $2\alpha$ around the axis n normal to the plane of the drawing, with a binary rotation around the axis r.

$$R(\alpha\mathbf{n})R(\pi\mathbf{r})R(-\alpha\mathbf{n})=R(\pi\mathbf{r}'). \tag{26}$$

Because the corresponding quaternions must multiply in the same manner, we get immediately (17), since the inverse and conjugate quaternions are identical in our case. We must remember, though, to use the Rodrigues parametrization of the quaternion, as in (21), and not Hamilton's (14):

$$\mathbb{A}\mathbb{R}\mathbb{A}^*=\mathbb{R}', \qquad \mathbb{A}=[\![\cos\frac{\alpha}{2},\sin\frac{\alpha}{2}\mathbf{n}]\!] \tag{27}$$

In this case it is heuristically possible to substitute $\mathbf{r}$ and $\mathbf{r}'$ for the quaternions $\mathbb{R}$ and $\mathbb{R}'$, with the poles of the binary rotations masquerading successfully as position vectors. This substitution, however, must never be done anywhere else. In particular, one must never attempt to operate with a quaternion on a vector, as is shown by the disastrous results of the crude interpretation of the rectangular transformation.

We must now discuss again the significance of the quaternion units. Because they are pure quaternions they must now be identified with binary rotations (rotations by $\pi$). This, for Hamilton, must have been absurd: the relation $\mathbf{i}^2$ equal to $-1$ must still be satisfied. But the product of two rotations by $\pi$ about the same axis is a rotation by $2\pi$. This is clearly the identity operation, i.e., one which does not change any vector, whereas we are now saying that it is equal to $-1$, i.e., that it changes the sign of all vectors in space. I believe that this is the reason Hamilton was forced to accept his parametrization, since this agreed with his picture of quaternion units as quadrantal rotations. Rodrigues, practical man as bankers must be, knew better than to worry about this strange result of his geometry—he did not carry, like Hamilton, all the world's problems on his shoulders. Nature and history, alas, were playing games with Hamilton. How was he to know that Cartan was going to discover in 1913 [**5**] objects (spinors) which are indeed multiplied by $-1$ under a rotation by $2\pi$, exactly as Rodrigues's parametrization requires? Moreover, when the topology of the rotation group became understood in the 1920s through the work of Hermann Weyl, it became natural to accept that the square of a binary rotation multiplies the identity by $-1$ and thus behaves like the quaternion units. Though this should have shown the enormous importance of quaternions in the rotation group, they were by that time somewhat discredited, so that other much less effective parametrizations of the rotation group were in universal use.

It must be stressed that the Rodrigues approach to rotations, by emphasizing their multiplication rules and by regarding them entirely as operators, fully reveals the group properties of the set of all orthogonal rotations, the full orthogonal group $SO(3)$, as it is now called. The set of all normalized quaternions (in the Rodrigues parametrization) is a group homomorphic to $SO(3)$ and it is its *covering group*. Although I cannot go into the mathematical significance of this statement, its practical importance in quantum mechanics, e.g., can be easily understood: it permits the study of the transformation properties of the wave functions of the electron spin. It is for this purpose that quaternions are superb, because their use in dealing with rotations makes the work not only simpler but also more precise than with any other method.

## The Decline

Hamilton was still under forty when he discovered quaternions, but he had more than twenty years of very productive research past him and was already showing the signs of having passed his prime. Financial and even sentimental worries are often mentioned, as well as overwork and an increasing consumption of alcohol [**33**]. I am inclined to believe that a major factor was that Hamilton himself was in some way aware of the internal contradictions of his doctrine and that he could not rest until he could peel off all layers of reality one by one to reach to the core. This was always beyond his power, since he was not prepared to renounce the, for him, essential picture of the quaternion units as quadrantal rotations. Be it as it may, his writing became more and more obscure: *even his supporters found his books unreadable. And* he himself became more isolated and eccentric.

E. T. Bell [3, Ch. 19] labelled the last twenty years of Hamilton 'The Irish Tragedy. Lanczos [26] compares them with Einstein's fruitless search for a unified field theory in his own last two decades. The truth is probably somewhere between these two views. For Hamilton suffered the weight of his own greatness: it was not enough for him to have an algebra, it was not enough to have a geometry, he had to 'interpret the oracles of the Universe' and the oracles trailed in front of him false clues that no one was to unravel for another three-score years after his death. If only he had known about spinors! The result, however, was that Hamilton, and to a much greater degree his followers, became dogmatic and intolerant (see [25, p. 182]) and that a great deal of sterile discussion ensued.

The last years of Hamilton, despite his immense fame, were not without worries: Continentals were spreading rumours that the great Gauss had actually discovered quaternions but had never bothered to publish. (They were right, as shown by Gauss's notes from 1819, published in 1900; see [13, vol. VIII, pp. 357–362].) In letters to De Morgan of January 1852 [15, vol. II, p. 490, vol. III, p. 330), Hamilton attacks these allegations. Curiously enough, of Rodrigues, who in 1840 not only had invented quaternions bar their name, but also published his formulae, there was never a word. Who would pay attention to a Socialist banker in matters mathematical?

After Hamilton's death his work began to give fruits but not in the direction which he had expected. His ideas of vectors and of their scalar and vector products were much too important so that people began to try and graft a new skin onto them in order to make these concepts usable. Grassmann in Germany and Heaviside in Britain moved some way in this direction, but one must admit that they were not much more transparent than Hamilton himself. It was left to Willard Gibbs of Yale to produce not only the first coherent picture of vectors and of their operations but also a good and successful working notation. This hardened the response of Hamilton's followers, who adopted a truly Byzantine posture, intent on stopping the flood of rebellion from across the Atlantic. Thus P. G. Tait [36, p. vi]:

> Even Prof. Willard Gibbs must be ranked as one of the retarders of quaternion progress, in virtue of his pamphlet on *Vector Analysis*, a sort of hermaphrodite monster, compounded of the notations of Hamilton and of Grassmann.

The kiss of life for quaternions, alas, much too late, came with the foundation in 1895 of an International Association for Promoting the Study of Quaternions and Allied Systems of Mathematics: an acknowledgement that quaternions were a corpse in need of resuscitation. Alexander Macfarlane, who taught at Texas, became the leading force of the Association, which actually published a Bulletin from 1900 to 1923. The influence of this group extended as far as Japan, where Kimura in 1907 became one of the major influences of the Association. Nothing that they did, however, succeeded in preventing the rise of vectors and the consequent decline of quaternions.

A number of applications of quaternions went on appearing from time to time (see

[1, p. 18].) Ironically, however, by the time in the late twenties when quantum mechanics made the study of the rotation group crucial, thus giving the quaternions their real *raison d'être*, they had been submerged for much too long in the murky waters of their battle against vectors to be able to come to the surface again. They are much too useful in this context, though, for their time not to return.

There is a moral to this story: Rodrigues's applied mathematics yields a more accurate picture of the quaternions than that afforded by the pure mathematics of their inventor: it is probably a myth that pure mathematics is either born or can stand entirely on its own, although the aesthetic appeal of pure mathematics makes us often think otherwise.

## Epilogue

After this article was communicated, a book was announced which contains new information about Rodrigues. This is the *Dictionnaire du Judaïsme Bordelais aux XVIII$^e$ et XIX$^e$ Siècles*, by Jean Cavignac (Archives Départementales de La Gironde, Bordeaux, 1987). This book contains a family tree of Rodrigues, which shows that his great-grandfather, Isaac Rodrigues-Henriques, was born in Spain, around 1689–91 and died in Bordeaux in 1767. He was indeed a banker but, contrary to previous belief, Olinde's father was an accountant. Surprisingly, the so-far universally accepted date of birth of Rodrigues is not right (so that the extraordinary coincidence with the day of the discovery of quaternions becomes a second-order effect). The correct date is 6 October 1795, and this is now unimpeachable, since his birth certificate is fully transcribed in a paper on Rodrigues by Paul Courteault (*Un Bordelais Saint-Simonien*) which I, like most people so far, had missed, since it was published in an obscure journal (*Revue Philomatique de Bordeaux*, Octobre–Décembre 1925, pp. 151–166). In accordance to this certificate, the date of birth was 14 *vendémiare* in the year IV of the Republic, at 1 p.m. Courteault (and Cavignac), instead, both agree with Michaud's date of death, 26.12.1850. Courteault also gives evidence that, although Olinde tried to enter the École Normale, he did not succeed in so doing, being prevented by his religion, so that how he learnt his advanced mathematics remains an unsolved mystery. Even worse, it appears that Rodrigues did not even attend the local secondary school (*Lycée*) at Bordeaux, so that we do not yet know anything at all about his formative years. It is of some value, however, that some of the traditional wisdom about this period, as repeated, I am afraid, in my paper, is now known to be worthless.

**Acknowledgements.** I am grateful to Drs. I. Grattan-Guinness and J. Gray for useful comments on various historical matters and to Dr. D. R. Altmann for very useful comments and stylistic advice. I am very grateful to M. Gérard Nahon, of the Centre National de la Recherche Scientifique for putting me in touch with M. Cavignac, who very kindly provided me with a copy of Courteault's paper and with advanced proofs of his book. I am also indebted to my friend, Dr. Richard Cooper, for providing me with an independent transcription of Rodrigues's birthdate from the Republican calendar.

## REFERENCES

1. S. L. Altmann, *Rotations, Quaternions, and Double Groups*, Clarendon Press, Oxford, 1986.
2. O. Barfield, *What Coleridge Thought*, Oxford University Press, London, 1972.
3. E. T. Bell, *Men of Mathematics*, Gollancz, London, 1937.
4. A. J. Booth, *Saint-Simon and Saint-Simonism, a Chapter in the History of Socialism in France*, Longmans, Green, Reader & Dyer, London, 1871.

5. E. Cartan, Les groupes projectifs qui ne laissent invariante aucune multiplicité plane, *Bull. Soc. Math. de France* 41 (1913), 53–96.
6. ______, *Leçons sur la Théorie des Spineurs*, vol. I, Hermann, Paris, 1938.
7. A. Cayley, On certain results relating to quaternions, *Phil. Mag.* 26 (1845), 141–145.
8. ______, *The Collected Mathematical Papers*, 13 volumes plus Index, University Press, Cambridge, 1888–1898.
9. S. Charléty, *Histoire du Saint-Simonisme (1825–1864)*, 2nd ed., Hermann, Paris, 1936 (1st ed., 1896).
10. W. K. Clifford, Applications of Grassmann's extensive algebra, *Amer. J. of Math. Pure and Applied* 1 (1878), 350–58.
11. E. Diehl, *Anthologia Lyrica Græca* (3rd ed.), Fasc. 2, Teubner, Leipzig, 1940.
12. L. Euler, Formulae generales pro translatione quacunque corporum rigidorum, *Novi. Comm. Acad. Sci. Imp. Petrop.* 20 (1775), 189–207.
13. C. F. Gauss, *Werke*, 12 vols, Königliche Gesellschaft der Wissenschaften, Göttingen, 1863–1929.
14. I. Grattan-Guinness, Personal communication, 1983.
15. R. P. Graves, *Life of Sir William Rowan Hamilton*, 3 vols, Hodges, Figgis & Co., Dublin, 1882–1889.
16. J. Gray, Olinde Rodrigues' paper of 1840 on transformation groups, *Arch. Hist. Exact Sci.* 21 (1980), 375–385.
17. W. R. Hamilton, On quaternions; or a new system of imaginaries in algebra, *Phil. Mag.*, 3rd ser. 25 (1844), 489–495.
18. ______, On a new species of imaginary quantities connected with the theory of quaternions, *Proc. Roy. Irish Acad.* 2 (1844), 424–434.
19. ______, On symbolical geometry, *Cambridge and Dublin Mathematical Journal* 1 (1846), 45–57.
20. ______, *Lectures on Quaternions*, Hodges & Smith, Dublin, 1853.
21. ______, *Elements of Quaternions*, (2nd edn, ed C. J. Jolly), 2 vols, Longmans, Green & Co., London, 1899–1901.
22. ______, *The Mathematical Papers of Sir William Rowan Hamilton*, vol. I, Geometrical Optics (eds. A. W. Conway and J. L. Synge); vol. II, Dynamics (eds. A. W. Conway and A. J. McConnell; vol. III, Algebra (eds, H. Halberstam and R. E. Ingram), Cambridge University Press, Cambridge, 1931, 1940, 1967.
23. T. L. Hankins, *Sir William Rowan Hamilton*, Johns Hopkins University Press, Baltimore, 1980.
24. G. S. Kirk, J. E. Raven, and M. Schofield, *The Presocratic Philosophers*, (2nd ed.), Cambridge University Press, Cambridge, 1963.
25. F. Klein, *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, 2 vols, Springer, Berlin, 1926, 1927.
26. C. Lanczos, William Rowan Hamilton-An appreciation, *American Scientist* 55 (1967), 129–143.
27. P. Larousse, *Grand Dictionnaire Universel du XIX$^e$ Siècle*, Larousse, Paris, vol. I, 1866; vol. XIII, no date.
28. A. L. Mackay, *The Harvest of a Quiet Eye. A Selection of Scientific Quotations*, (ed. M. Ebison), The Institute of Physics, Bristol, 1977.
29. A. J. McConnell, The Dublin Mathematical School in the first half of the nineteenth century, *Proc. Roy. Irish Acad.* A50 (1943), 75–92.
30. C. C. MacDuffee, Algebra's debt to Hamilton, *Scripta Math.* 10 (1944), 25–36.
31. J. C. Maxwell, Remarks on the mathematical classification of physical quantities, *Proc. London Math. Soc.* 3 (1869), 224–232.
32. Michaud, *Biographie Universelle Ancienne et Moderne*. Nouvelle édition. Desplaces, Paris. Vol. I, 1843, vol. XXXVI (no date), pp 288–9.
33. S. O'Donnell, *William Rowan Hamilton. Portrait of a Prodigy*. Boole Press, Dublin, 1983.
34. M. Riesz, Clifford Numbers and Spinors. Lecture Series No 38, The Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, Maryland, 1958.
35. O. Rodrigues, Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et la variation des coordonnées provenant de ses déplacements considérés indépendamment des causes qui peuvent les produire, *J. de Mathématiques Pures et Appliquées* 5 (1840), 380–440.
36. P. G. Tait, *An Elementary Treatise on Quaternions*, (3rd ed.), Cambridge University Press, Cambridge, 1890.
37. G. Temple, *Cartesian Tensors. An Introduction*, Methuen, London, 1960.
38. S. P. Thompson, *The Life of William Thomson, Baron Kelvin of Largs*, Macmillan, London, 1910.
39. J. Warren, On the geometrical representation of the powers of quantities, whose indices involve the square roots of negative quantities, *Phil. Trans. Roy. Soc.* 119 (1829), 339–359.
40. G. Weill, *Saint Simon et son Œuvre. Un Précurseur du Socialisme*, Perrin et Cie, Paris, 1894.
41. E. Wilson, *To the Finland Station. A Study in the Writing and Acting of History*, Secker & Warburg, London, 1941.

# NOTES

## The Historical Development of J. J. Sylvester's Four Point Problem

RICHARD E. PFIEFER
San Jose State University
San Jose, CA 95192

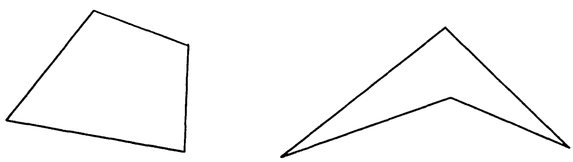
*Historically, it would seem that the first question given on local probability, since Buffon, was the remarkable four-point problem of Prof. Sylvester.*

M. W. Crofton (1885) [6]

**1. The Early History** In the *Educational Times* of 1864 [15], question 1491, J. J. Sylvester proposed what became known as his four point problem:

> Show that the chance of four points forming the apices of a reentrant quadrilateral is $1/4$ if they be taken at random in an indefinite plane, but $1/4 + e^2 + x^2$, where $e$ is a finite constant and $x$ a variable quantity, if they be limited by an area of any magnitude and of any form.

The limiting area mentioned above is understood to be a convex region [18].



convex quadrilateral | reentrant quadrilateral

FIGURE 1.

Since there are two questions stated, we first report the results on the probability, $P$, that four points taken at random in the plane form a reentrant quadrilateral. The readers of the *Educational Times* set to work on this problem and here is a list of some of the solvers and their published answers [9].

| SOLVER | PROBABILITY, $P$ |
|---|---|
| Cayley and Sylvester | $1/4$ |
| G. C. DeMorgan | $1/2$ |
| J. M. Wilson | $1/3$ |
| C. M. Ingleby | $P < 1/2$ |
| (Name Unknown) | $3/8$ |
| W. S. B. Woolhouse | $35/12\pi^2$ |

At the time, no one could detect whether any of the probabilities computed above was *the* solution, and J. J. Sylvester concluded "This problem does not admit of a determinate solution" [16].

Cayley and Sylvester "solved" the problem by assuming that $A$, $B$, and $C$ are the three points which form the largest triangle, then the fourth point $D$ gives a reentrant quadrilateral only if it falls within the original triangle, $\triangle ABC$, out of the four equal-area triangles as shown in FIGURE 2. Cayley and Sylvester knew that this argument was insufficient in that it was possible, by an equally good argument, to obtain an inconsistent result [18].

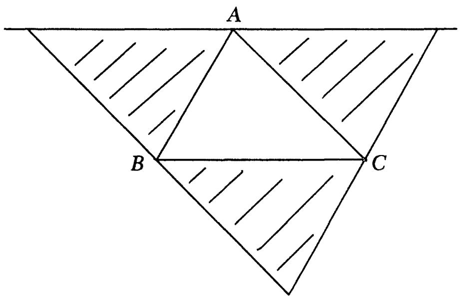



FIGURE 2.

In his solution to Question 1491, W. S. B. Woolhouse decided that he could assume that the four points were contained inside a circle, $K$, of radius $r > 0$, then compute the probability for four points inside $K$, and finally take the limit as $r \to \infty$. In other words, he wanted to treat the plane as a circle of infinite radius. (See FIGURE 3.) He noted that:

$$\begin{aligned} P(K) &= \text{Prob (reentrant quadrilateral)} \\ &= 4 \text{ Prob (One point is inside the triangle formed by the other 3 points.)} \\ &= 4 \text{ Mean (Expected) Triangle Area of 3 points/Area of } K \\ &= 4\, M(K)/A(K), \text{ where} \end{aligned}$$

$$M(K) = 1/(A(K))^3 \iint_{P\in K} \iint_{Q\in K} \iint_{R\in K} \frac{1}{2} \left\| \begin{matrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{matrix} \right\| dy_3\, dy_2\, dy_1\, dx_3\, dx_2\, dx_1. \quad (1)$$

This formula defines $M(K)$ for any closed, bounded convex plane region $K$. When $K$ is a circle of radius $r$, $A(K) = \pi r^2$.

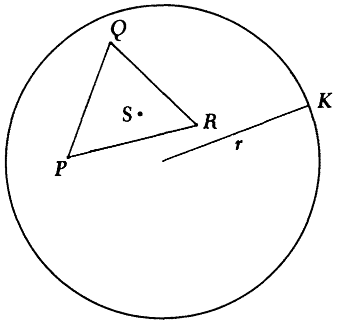



FIGURE 3.

The computation of $M(K)$, for $K$ a circle of radius $r$, will be left to the interested reader with the hints: use Crofton's Formula and polar coordinates. (See Solomon [14] for Crofton's Formula and some related computations.) After a few pages, you will obtain:

$$M(K) = \left(1/(\pi r^2)^3\right)(35/48\pi^2)(\pi r^2)^4, \quad \text{and } P(K) = 35/12\pi^2.$$

Noting that $P(K)$ is independent of the radius $r$, Woolhouse concluded that the solution to Question 1491 was $35/12\pi^2$.

The culprit responsible for these inconsistent results is, of course, the phrase "at random in the plane." From this sample of two of the solutions, we see that each of the solvers above had used his own intuitive interpretation of the phrase and arrived at different answers. In subsequent issues of the *Educational Times*, there was a great deal of spirited discussion of what "at random in the plane" should mean. Crofton [5], in 1868, wrote about these differences of opinion and the discordant results in the four-point problem:

> this arises, not from any inherent ambiguity in the subject matter, but from the weakness of the instrument employed; our undisciplined conceptions of a novel subject requiring to be repeatedly and patiently reviewed, tested, and corrected by the light of experience and comparison, before they are purged from all latent error.

The discussion, of course, would not be completely resolved until probability theory in terms of appropriate measures was developed in the following century. For a discussion of the necessity of such a measure in geometric probability see the monograph by Kendall and Moran [11, pp. 9–13].

**2. The Variational Four-Point Problem** Now we concentrate on the computation of the probability, $P(K)$, of the four points forming a re-entrant quadrilateral when they are taken at random inside a closed, bounded, convex plane region $K$. As previously shown, $P(K) = 4M(K)/A(K)$ where $M(K)$, defined by formula (1), is the mean (or expected) area of the triangle formed by three points taken at random in $K$ and $A(K)$ is the area of $K$.

As noted above, Woolhouse obtained

$$M(K) = (35/48\pi^2)(\pi r^2), \quad \text{and} \quad P(K) = 35/12\pi^2$$

when $K = D$ is a circular disk.

Question No. 1229 from the *Educational Times* of 1865 [17] proposed by S. Watson and solved by J. J. Sylvester, was to show that $M(K) = (1/12)A(K)$ when $K$ is a triangle. It follows that $P(K) = 1/3$ when $K = \Delta$ is a triangle. By 1867, Woolhouse [19] had computed $M(K)$ and $P(K)$ for $K$ a square, and $K$ a regular hexagon. We summarize these values in TABLE 4.

It should be noted that, if $T$ is a non-singular affine transformation of the plane, then $P(T(K)) = P(K)$. This explains why squares and parallelograms, or circles and

TABLE 4

| $K$ | Triangle | Square or Parallelogram | Regular Hexagon | Circle or Ellipse |
|---|---|---|---|---|
| $M(K)$ | $A(K)/12$ | $11A(K)/144$ | $289A(K)/3888$ | $35A(K)/48\pi^2$ |
| $P(K)$ | $1/3$ | $11/36$ | $289/972$ | $35/12\pi^2$ |

ellipses have the same probability. Making the convention that $A(K)=1$ will simplify TABLE 4, so we will adopt that convention and only allow area-preserving affine transformations $T$.

Notice (from TABLE 4) that $1/3 > 11/36 > 289/972 > 35/12\pi^2$. J. J. Sylvester asked for the shape of the regions $K$ that gave the maximum and minimum probabilities $P(K)$ (see [6]), and the conjecture was:

(i) $P(D) \leqslant P(K)$, when $D$ is bounded by a circle or an ellipse;

(ii) $P(\Delta) \geqslant P(K)$, when $\Delta$ is bounded by a triangle.

The computational problem of finding $P(K)$ and this new variational problem (or conjecture) are now collectively known as Sylvester's Four Point Problem [2, 3, 11, 12, 13, 14].

The first (not quite rigorous) proof of (i) was given by M. W. Crofton [6] in 1885. But it was not until 1917 that a complete proof of both (i) and (ii) was given by W. Blaschke [2]. Blaschke gave another proof in 1923 [3]. The two proofs given by Blaschke use the same geometric ideas so we will outline Blaschke's solution emphasizing these ideas.

**3. The Geometry of Blaschke's Solution of Sylvester's Four Point Problem** W. Blaschke actually proved the equivalent conjecture for the mean value $M(K)$, namely:

(i′) $M(D) \leqslant M(K)$, where $D$ is bounded by a circle or an ellipse;

(ii′) $M(\Delta) \geqslant M(K)$, where $\Delta$ is bounded by a triangle. And he showed that equality holds if and only if $K$ is an ellipse in (i′) or $K$ is a triangle in (ii′).

To simplify the expression for $M(K)$, we continue the convention that $A(K)=1$.

Blaschke's solution depends on the geometry of the multiple integral

$$M(K)=\int_a^b\int_a^b\int_a^b\int_{\alpha_1}^{\beta_1}\int_{\alpha_2}^{\beta_2}\int_{\alpha_3}^{\beta_3}\frac{1}{2}\left\|\begin{matrix}1 & x_1 & y_1\\ 1 & x_2 & y_2\\ 1 & x_3 & y_3\end{matrix}\right\| dy_3\,dy_2\,dy_1\,dx_3\,dx_2\,dx_1,$$

and, more specifically, the "inside" triple integral,

$$I(x_1,x_2,x_3)=\int_{\alpha_1}^{\beta_1}\int_{\alpha_2}^{\beta_2}\int_{\alpha_3}^{\beta_3}|Ay_1+By_2+Cy_3|\,dy_3\,dy_2\,dy_1.$$

(See FIGURE 5.)

Here $A=(x_3-x_2)$, $B=(x_1-x_3)$, and $C=(x_2-x_1)$. We may assume $x_1, x_2, x_3$ are distinct.

Since we cannot sketch the surface given by the integrand $f(y_1,y_2,y_3)=|Ay_1+By_2+Cy_3|$, let's look at the integral

$$I=\int_{\alpha_1}^{\beta_1}\int_{\alpha_2}^{\beta_2}|Ay_1+By_2|\,dy_2\,dy_1,$$

as shown in FIGURE 6.

Since the integral, $I$, represents the volume under the surface $z=|Ay_1+By_2|$ and above the rectangle $R=[\alpha_1,\beta_1]\times[\alpha_2,\beta_2]$, we can see from FIGURE 6 that the value of $I$ is a strictly increasing function of the distance, $d$, of the center of $R$, $(m_1,m_2)=(\alpha_1+\beta_1,\alpha_2+\beta_2)/2$, from the line $Ay_1+By_2=0$ in the $(y_1,y_2)$ plane. The corresponding fact holds for our integral $I(x_1,x_2,x_3)$.

In addition, the convexity of $K$ forces the determinants

$$l=\begin{vmatrix}1 & x_1 & \alpha_1\\ 1 & x_2 & \alpha_2\\ 1 & x_3 & \alpha_3\end{vmatrix} \quad \text{and} \quad u=\begin{vmatrix}1 & x_1 & \beta_1\\ 1 & x_2 & \beta_2\\ 1 & x_3 & \beta_3\end{vmatrix}$$

to have opposite signs. (Remember, $x_1, x_2, x_3$ may be in any order.) This is illustrated in FIGURE 5.

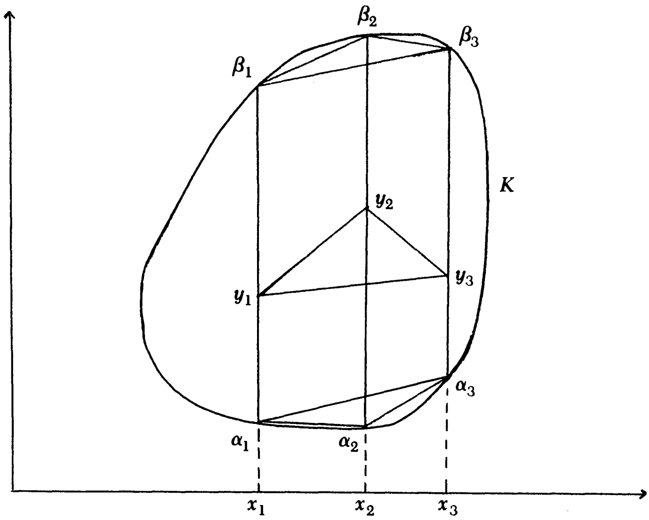



FIGURE 5.

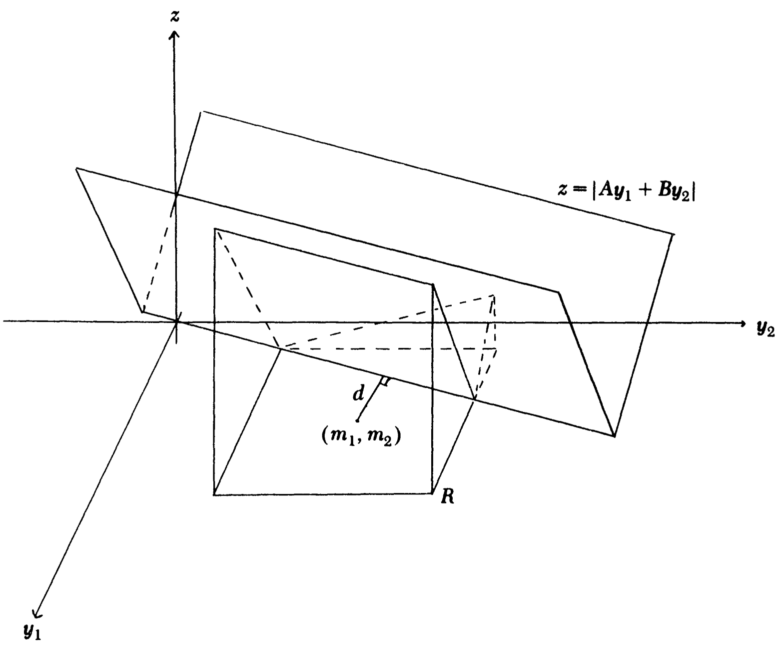



FIGURE 6.

Using the formula for distance, $d$, from the center $(m_1, m_2, m_3) = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \alpha_3 + \beta_3)/2$ to the plane $Ay_1 + By_2 + Cy_3 = 0$, we have

$$d = \frac{|Am_1 + Bm_2 + Cm_3|}{\sqrt{A^2+B^2+C^2}} = \frac{\left\|\begin{matrix} 1 & x_1 & \alpha_1+\beta_1 \\ 1 & x_2 & \alpha_2+\beta_2 \\ 1 & x_3 & \alpha_3+\beta_3 \end{matrix}\right\|}{2\sqrt{A^2+B^2+C^2}} = \frac{\left\|\begin{matrix} 1 & x_1 & \alpha_1 \\ 1 & x_2 & \alpha_2 \\ 1 & x_3 & \alpha_3 \end{matrix}\right\| + \left\|\begin{matrix} 1 & x_1 & \beta_1 \\ 1 & x_2 & \beta_2 \\ 1 & x_3 & \beta_3 \end{matrix}\right\|}{2\sqrt{A^2+B^2+C^2}}$$
$$= |l+u| \Big/ 2\sqrt{A^2+B^2+C^2}.$$

Note that $d$ is a constant multiple of the area of the "midpoint triangle" with vertices $(x_1, m_1), (x_2, m_2), (x_3, m_3)$.

Now, if we allow the three intervals in FIGURE 5 to vary vertically, subject to the constraint that $l$ and $u$ have opposite sign (or one or both are zero), we see that $d$ attains its minimum value of zero when $l = -u$ and $d$ attains its maximum when $l = 0$ or $u = 0$. In other words, $I(x_1, x_2, x_3)$ attains the smallest value, call it $I^*(x_1, x_2, x_3)$, if the three intervals in FIGURE 5 have their midpoints $(x_1, m_1), (x_2, m_2), (x_3, m_3)$ on a line $N$. $I(x_1, x_2, x_3)$ attains its maximum value, call it $\bar{I}(x_1, x_2, x_3)$, when the lower endpoints $(x_1, \alpha_1), (x_2, \alpha_2), (x_3, \alpha_3)$ lie on a line $\bar{N}$ (i.e. $l = 0$). See FIGURE 7, where we have used the $x$-axis for both lines $N$ and $\bar{N}$. We may always use the $x$-axis for the line $N$ (or $\bar{N}$) because of the invariance of $I(x_1, x_2, x_3)$ under transformations of the form $T(x, y) = (x, y - (mx + b))$.

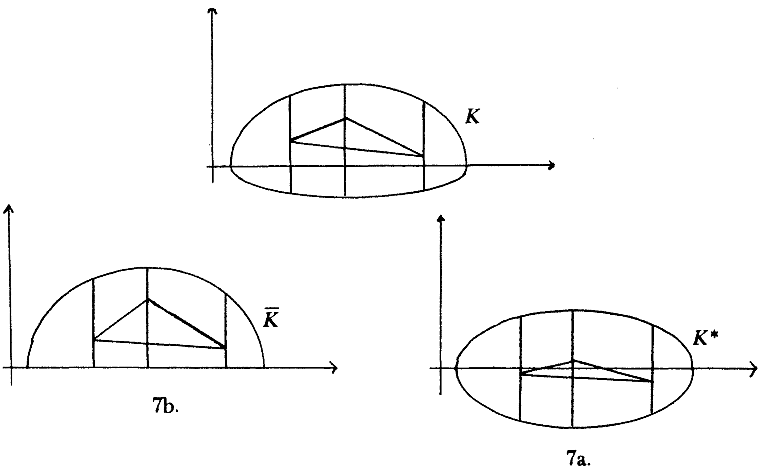



FIGURE 7.

Since $x_1, x_2, x_3$ was an arbitrary triple, centering all of the vertical line segments of $K$ in a line will form a new set $K^*$, (FIGURE 7a) with the property $M(K^*) \leqslant M(K)$. Setting all of the vertical line segments on top of the $x$-axis will form a new set $\bar{K}$, (FIGURE 7b) with the property $M(\bar{K}) \geqslant M(K)$. These two operations, of forming $K^*$ and $\bar{K}$ from a given set $K$, are called the Steiner Symmetrization of $K$ and the Schüttelung of $K$, respectively, in the line $N$, and are well-known in the geometry of convex sets [1, 4, 8]. It is easy to see that $K$, $K^*$, and $\bar{K}$ all have the same area and it is not difficult to show that $K^*$ and $\bar{K}$ are convex whenever $K$ is convex. It is also well known that there exists a sequence of Steiner Symmetrizations of $K$ (respectively Schüttelung operations of $K$) in a sequence of lines $N_1, N_2, \ldots,$ (resp. $\bar{N}_1, \bar{N}_2, \ldots,$) which converges to a circle [4, 8] (respectively, triangle [1]). See FIGURE 8.

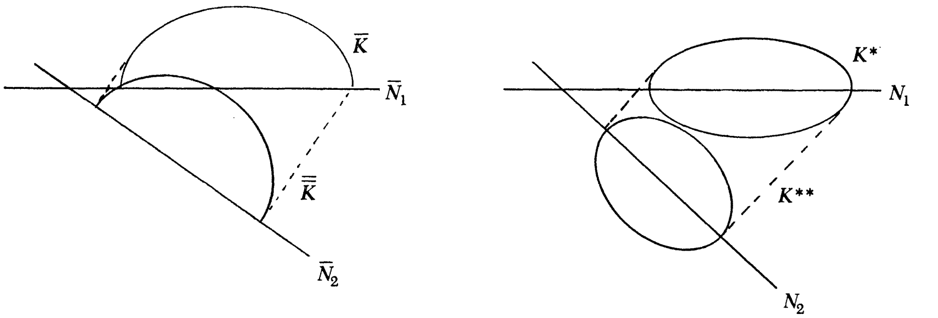



FIGURE 8.

If we denote these sequences by $K_0 = K$, $K_1 = K^*$, $K_2 = (K^*)^*$, etc., and $K^0 = K$, $K^1 = \overline{K}$, $K^2 = (\overline{\overline{K}})$, etc., and observe that $M(K)$ is a continuous function of $K$, we see that $\{M(K_i)\}$ is a decreasing sequence with limit $M(D)$, where $D$ is bounded by a circle, and $\{M(K^i)\}$ is an increasing sequence with limit $M(\triangle)$, where $\triangle$ is bounded by a triangle. This completes the outline of the proof of (i′) and (ii′). The equality conditions are established by observing that the first Symmetrization (resp., Schüttelung) of $K$ can be made to strictly decrease (resp., increase) $M(K)$ unless $K$ is an ellipse (resp., triangle).

**4. The Generalization of Sylvester's Four Point Problem to Three Dimensions; an Unsolved Problem** If we let $K$ be a three dimensional, compact, convex set of volume 1, and define $M(K)$ to be the mean (or expected) value of the volume of the tetrahedron formed by 4 points taken at random (uniform distribution) from $K$, the natural generalization of Sylvester's Four Point Problem would be the conjecture:

(a) $M(K) \geqslant M(D)$, where $D$ is bounded by a sphere or an ellipsoid; and equality holds if and only if $K$ is a solid ellipsoid;

(b) $M(K) \leqslant M(\triangle)$ where $\triangle$ is a solid tetrahedron, and equality holds if and only if $K$ is a tetrahedron.

Blaschke stated in 1917 [2] that both (a) and (b) were true, and that the proofs, as outlined in section 3, would carry over to higher dimensions. For (a), he was correct. This was verified by Groemer [7] in 1973. However, conjecture (b) is still unsolved! Why doesn't the proof of (ii′) carry over to the three-dimensional problem? In two dimensions, the convexity of $K$ forced the "lower" and "upper" determinants $l$ and $u$ to have opposite sign. For Blaschke's proof to work in three dimensions, the convexity of $K$ must force the corresponding determinants

$$l = \begin{vmatrix} 1 & x_1 & y_1 & \alpha_1 \\ 1 & x_2 & y_2 & \alpha_2 \\ 1 & x_3 & y_3 & \alpha_3 \\ 1 & x_4 & y_4 & \alpha_4 \end{vmatrix}, \quad \text{and} \quad u = \begin{vmatrix} 1 & x_1 & y_1 & \beta_1 \\ 1 & x_2 & y_2 & \beta_2 \\ 1 & x_3 & y_3 & \beta_3 \\ 1 & x_4 & y_4 & \beta_4 \end{vmatrix}$$

to have opposite sign. But this is not the case! Referring to FIGURE 9, $K$ is a tetrahedron and

$$l = \begin{vmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 1 \end{vmatrix} = 4, \quad \text{and} \quad u = \begin{vmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & -1 & 2 \end{vmatrix} = 4.$$

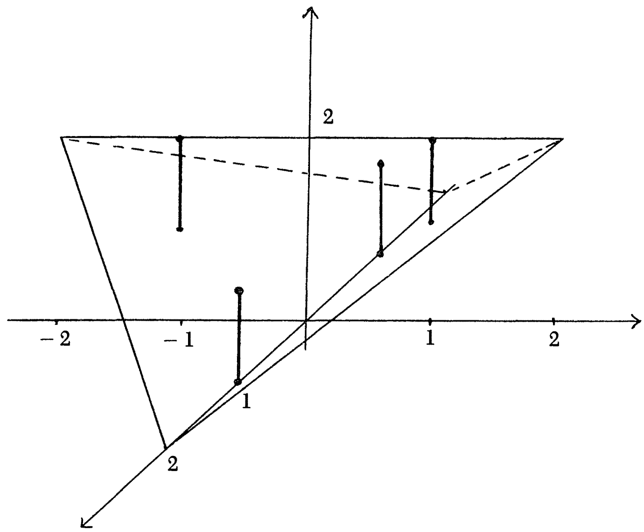



FIGURE 9.

If we shifted the four intervals shown in FIGURE 9, using the Schüttelung process, the corresponding function $d$ would *decrease* to 0. (The verification of this is left to the reader.) For Blaschke's solution of (ii′) to carry over to (b), $d$ should increase! This indicates that the Schüttelung procedure may not be the correct procedure. (Perhaps (b) is not the correct generalization of (ii′).) If you would care to work on this problem and/or a restatement of (b), please do.

By the way, the thought may have occurred to the reader that perhaps a maximum does not exist. That is, maybe there is no set $\Delta$ of volume one such that $M(K) \leqslant M(\Delta)$ for all $K$. That would not be correct. Using the result of John [**10**], that every three-dimensional, compact, convex set of $K$ of volume 1 is contained inside an ellipsoid of volume $3^3 = 27$, it is (relatively) easy to show that such a maximum set exists. This, too, will be left to the reader.

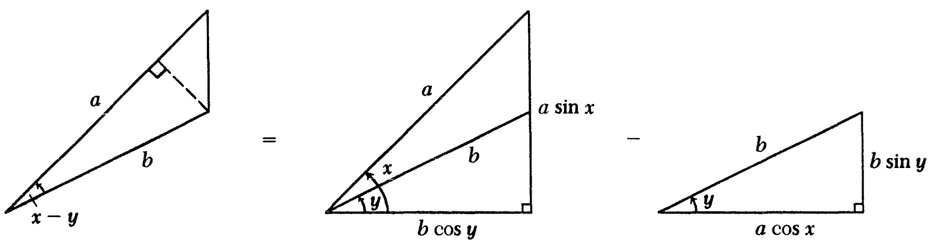
**5. Concluding Remarks** We have followed a single linear sequence of events in the development of Sylvester's Four Point Problem in order to arrive at the current status of only one of its offspring. There have been many other relatives along the way. There is a survey of results in the book by Santaló [**13**]. Solomon's book [**14**] gives related results and concentrates on some of the computations using Crofton's Theorem. For more historical remarks, the article by Klee [**12**] is recommended.

REFERENCES

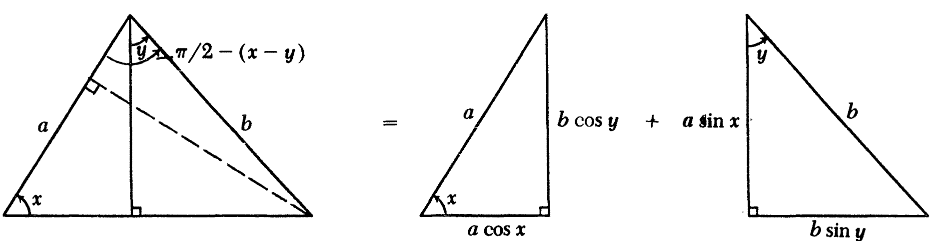

1. T. Biehl, Über Affine Geometrie XXXVIII, Über die Schüttlung von Eikörpern, *Abh. math. Semin. Hamburg Univ.* Bd 2 (1923), 69–70.
2. W. Blaschke, Über affine Geometrie XI: Lösung des "Vierpunktproblems" von Sylvester aus der Theorie der geometrischen Wahrscheinlichkeiten, *Leipziger Berichte* 69 (1917), 436–453.
3. ______, *Vorlesungen über Differentialgeometrie II: Affine Differentialgeometrie*, Springer, Berlin, 1923.
4. T. Bonnesen and W. Fenchel, *Theorie Der Konvexen Körper*, Chelsea, New York, 1948.
5. M. W. Crofton, On the Theory of Local Probability, *Royal Society of London*, *Philosophical Transactions* 158 (1868), 181–200.
6. ______, Probability, *Encyclopedia Brittanica*, 9th ed., 19 (1885), 768–788.

7. H. Groemer, On some mean values associated with a randomly selected simplex in a convex set, *Pacific J. Math.* 45 (1973), 525–533.
8. H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1957.
9. C. M. Ingleby, *Mathematical Questions and their Solutions from the Educational Times*, W. J. C. Miller, Editor, 5 (1865), 108.
10. F. John, *Extremum Problems with Inequalities as Subsidiary Conditions, Studies and Essays Presented to R. Courant*. Interscience Publishers, Inc., New York, 1948.
11. M. G. Kendall and P. A. P. Moran, Geometrical Probability, *Griffin's Statistical Monographs and Courses*, No. 10, M. G. Kendall, Editor, Hafner Publishing Company, New York, 1963.
12. V. Klee, What is the expected volume of a simplex whose vertices are chosen at random from a given convex body?, *Amer. Math. Monthly* 76 (1969), 286–288.
13. L. A. Santaló, *Integral Geometry and Geometric Probability*, Addison-Wesley Publ. Co., Reading, MA, 1976.
14. H. Solomon, *Geometric Probability*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1978.
15. J. J. Sylvester, *The Educational Times*, London, April (1864).
16. ______, *Report of the British Association* 35 (1865), 8–9.
17. S. Watson, Question No. 1229, *Mathematical Questions and their Solutions from the Educational Times*, W. J. C. Miller, editor, 4 (1865), 101.
18. W. S. B. Woolhouse, Some Additional Observations on the Four-Point Problem, *Mathematical Questions and their Solutions from the Educational Times*, W. J. C. Miller, editor, 7 (1867), 81.
19. ______, Question no. 2471, *Mathematical Questions and their Solutions from the Educational Times*, W. J. C. Miller, editor, 8 (1867), 100–105.

# Proof without Words: Area and Difference Formulas




$\sin(x - y) = \sin x \cos y - \cos x \sin y$




$\cos(x - y) = \cos x \cos y + \sin x \sin y$

—SIDNEY H. KUNG
JACKSONVILLE UNIVERSITY

# Parametric Integration Techniques

AUREL J. ZAJTA
University of Houston-Downtown
Houston, TX 77002

SUDHIR K. GOEL
Valdosta State College
Valdosta, GA 31698

In this note, we present an integration technique that evaluates integrals through the manipulation of a parameter. By "parametric integrals" we mean single integrals of bivariate functions with respect to one of the variables (the variable of integration), while the other variable is referred to as a *parameter*. By "manipulation" we mean operations under the integral sign in terms of the parameter, such as: differentiation, integration, or some other kind of limiting processes (like summation over an infinite index set, etc.). We illustrate the method with the help of some selected examples. These examples will show that the parametric integration technique only requires the mathematical maturity of Calculus III level and provides a very straightforward method to evaluate difficult integrals for which one commonly uses the method of *contour integration*. However, this technique does not appear to be widely used. We begin with a simple example.

*Example* 1. Consider the well-known integral

$$\int_0^\infty \frac{dx}{p^2+x^2} = \frac{\pi}{2p}, \qquad (p>0).$$

Considering $p$ as a parameter and differentiating with respect to it, we obtain

$$\int_0^\infty \frac{-2p\,dx}{\left(p^2+x^2\right)^2} = -\frac{\pi}{2p^2}, \qquad (p>0)$$

or

$$\int_0^\infty \frac{dx}{\left(p^2+x^2\right)^2} = \frac{\pi}{4p^3}, \qquad (p>0)$$

as a new integral formula. The standard method to derive this is integration by parts, but that would take a bit longer.

Continuing this process of differentiating with respect to $p$ and simplifying afterwards, one obtains the formulas for the values of the integrals of $1/(p^2+x^2)^3$, $1/(p^2+x^2)^4$, etc., or $1/(p^2+x^2)^n$ by induction.

A natural question arises: How should one introduce a parameter within the integrand? Usually, this is not a problem. In many integrals, especially in the formulas of integral transforms, parameters are already present. However, there exist cases when at the outset an integral may contain no parameter. In such cases, a parameter is generally introduced by changing the variable of integration using a substitution that contains a parameter. For example, consider the well-known integral

$$\int_0^\infty dx/(1+x^2) = \pi/2$$

(a special case of the integral in Example 1 above with $p=1$). In this example, we can introduce a parameter $p$ by substituting $x/p$ for $x$.

The application of limiting processes under the integral sign requires the interchangeability of two limiting processes, one being the integration itself and the other one a limiting process (differentiation or integration, etc.) in terms of the parameter. More precisely, we must assume the following:

(1) Differentiation with respect to a parameter under the integral sign is permitted.
(2) Integration with respect to a parameter under the integral sign is permitted.
(3) Limit and an integral sign can be interchanged.

Theorems establishing conditions when the order of these processes may be interchanged have long been known; some can be traced back to Leibniz. Studying these (see [3], pp. 286–292), we shall find that the conditions are quite general (that is, general enough to cover the evaluation of a broad class of integrals). The *uniformity* of the limiting processes is a usual requirement, but this is evident in most cases when the integrals themselves exist. However, care must be exercised to avoid absurdities (see Example 6 below).

We present the rest of the examples under the above assumptions.

*Example* 2. Show that $\Gamma(n+1)=\int_0^\infty x^n e^{-x}\,dx=n!$ for a nonnegative integer $n$.

We begin with the well-known integral

$$\int_0^\infty e^{-ax}\,dx=\frac{1}{a},\qquad (a>0).$$

Repeated differentiation with respect to the parameter '$a$' gives

$$\begin{aligned}\int_0^\infty xe^{-ax}\,dx&=\frac{1}{a^2},\\ \int_0^\infty x^2e^{-ax}\,dx&=\frac{2}{a^3},\\ &\vdots\\ \int_0^\infty x^ne^{-ax}\,dx&=\frac{n!}{a^{n+1}}.\end{aligned}$$

Choosing $a=1$, we obtain the desired formula.

*Example* 3. Show that

$$\int_0^1 x^n(\log x)^k\,dx=\frac{(-1)^k k!}{(n+1)^{k+1}},\qquad (n\neq -1).$$

We begin with the well-known formula

$$\int_0^1 x^n\,dx=\frac{1}{n+1},\qquad (n\neq -1).$$

Considering $n$ as a parameter instead of a constant and differentiating both sides with respect to it, we obtain

$$\int_0^1 x^n\log x\,dx=-\frac{1}{(n+1)^2}=\frac{(-1)^1\cdot 1!}{(n+1)^{1+1}}.$$

Thus, the result holds for $k=1$. To obtain the formula for the general case, we continue differentiating and apply induction.

*Example* 4. Show that

$$\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}.$$

This integral is generally evaluated by using contour integration and thus requires the theory of complex functions. However, it can be computed easily by applying parametric integration. To show this, we need to consider the integral of the product of $\sin(x)$ with another function involving a parameter that will:

(i) guarantee the existence (convergence) of the integral over the interval $(0,\infty)$ and

(ii) render a denominator of $x$ from integrating under the integral sign with respect to the parameter $p$.

An integral satisfying these requirements is readily available in the form of the Laplace transform of $\sin x$, that is

$$\int_0^\infty e^{-px}\sin x\,dx = \frac{1}{1+p^2}, \qquad (p>0). \qquad (*)$$

Integrating both sides with respect to the parameter from 0 to $p$, we obtain

$$\int_0^\infty \frac{1-e^{-px}}{x}\sin x\,dx = \arctan p.$$

The last formula is almost what we wanted, with the exponential term yet to be removed. To this end, we apply the limit as $p \to \infty$.

*Example* 5. We can use Example 4 to show that

$$\int_0^\infty \left(\frac{\sin t}{t}\right)^2 dt = \frac{\pi}{2}.$$

Substituting $t = \frac{x}{p}$ $(p>0)$ in Example 4, we obtain

$$\int_0^\infty \left(\frac{\sin pt}{t}\right) dt = \frac{\pi}{2}, \qquad (p>0). \qquad (**)$$

Integrating both sides with respect to the parameter from 0 to $p$, we obtain

$$\int_0^\infty \left(\frac{1-\cos pt}{t^2}\right) dt = \frac{\pi p}{2}. \qquad (***)$$

Finally, substituting $p=2$ and dividing both sides by 2, we obtain the result,

$$\int_0^\infty \left(\frac{\sin t}{t}\right)^2 dt = \frac{\pi}{2}.$$

*Example* 6. In this example, we illustrate why care must be exercised to avoid absurdities. Consider the integral given by $(**)$ in Example 5 above, that is

$$\int_0^\infty \frac{\sin pt}{t}\,dt = \frac{\pi}{2}, \qquad (p>0).$$

Differentiating both sides with respect to the parameter $p$, we obtain

$$\int_0^{\infty} \cos pt\, dt = 0,$$

which is absurd as the integral on the left-hand side does not converge.

The true convenience of parametric integration becomes most evident when one compares the use of parametric integration with the derivations resulting from contour integration. Among others, the list of integrals we have checked includes the following:

$$\int_0^{\infty} \left( \frac{x^{-\alpha}}{1+x} \right) dx = \frac{\pi}{\sin \pi\alpha}, \qquad (0 < \alpha < 1) \tag{1}$$

$$F_1 = \int_{-\infty}^{\infty} \sin x^2\, dx = \sqrt{\frac{\pi}{2}}, \tag{2}$$

$$F_2 = \int_{-\infty}^{\infty} \cos x^2\, dx = \sqrt{\frac{\pi}{2}}, \tag{3}$$

$$\int_0^{\infty} \frac{\cos px}{1+x^2}\, dx = \frac{\pi}{2} \cdot e^{-p}, \qquad (p \geqslant 0) \tag{4}$$

$$F(p) = \int_0^{\infty} \frac{\sin px}{x(1+x^2)}\, dx = \frac{\pi}{2} \cdot (1 - e^{-p}), \qquad (p \geqslant 0). \tag{5}$$

Formula (1) is extremely useful. It embodies a host of other integral formulas by assigning specific numerical values to $\alpha$ and changing the integration variable. To derive it, we consider the geometric series expansions of the function $f(x) = 1/(1+x)$, that is

$$1/(1+x) = 1 - x + x^2 - x^3 + \cdots,$$

and

$$1/(1+x) = \frac{1}{x} - \frac{1}{x^2} + \frac{1}{x^3} - \frac{1}{x^4} + \cdots,$$

according to whether $|x|$ is less than 1 or greater than 1. Hence, we split the interval of integration in two parts, namely $(0,1)$ and $(1,\infty)$, and evaluate the integrals of the infinite series termwise. In this example, the role of parameter is taken by the discrete variable summation index. The result is the partial fraction series of $\pi/\sin \pi\alpha$.

The integrals $F_1$ and $F_2$ are known as Fresnel integrals. We sketch the derivation of $F_1$; the derivation of $F_2$ is analogous. Substituting $t = x^2$ in $F_1$, we obtain

$$F_1 = \int_0^{\infty} \frac{\sin t}{\sqrt{t}}\, dt.$$

Replacing $p$ by $p^2$ in formula $(*)$ above, we obtain

$$\int_0^{\infty} e^{-p^2 t} \sin t\, dt = \frac{1}{1+p^4}.$$

Integrating both sides with respect to $p$ from 0 to infinity gives

$$\int_0^{\infty} \sin t \left( \int_0^{\infty} e^{-p^2 t}\, dp \right) dt = \int_0^{\infty} \frac{1}{1+p^4}\, dp.$$

Since

$$\int_0^\infty e^{-p^2 t}\,dp = \frac{1}{2}\sqrt{\frac{\pi}{t}}$$

(a well-known formula from probability theory) and

$$\int_0^\infty \frac{dp}{1+p^4} = \frac{\pi}{2\sqrt{2}},$$

(obtained from (1) above by substituting $p^4$ for $x$ and $3/4$ for $\alpha$), the value of $F_1$ is evident.

Formulas (4) and (5) are interrelated; one can be obtained from the other by integration or differentiation, respectively, with respect to the parameter $p$. Differentiating the integral (5) with respect to $p$ twice, we can easily verify that it satisfies the differential equation:

$$F''(p) - F(p) = -\frac{\pi}{2},$$

with initial values $F(0) = 0$ and $F'(0) = \pi/2$. The unique solution of this differential equation provides the value of the integral (5).

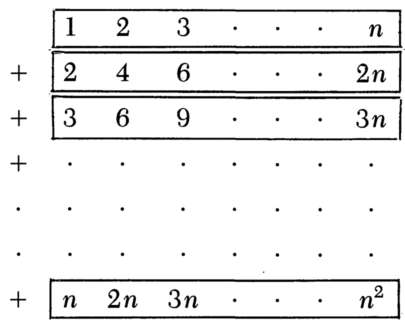
These and other examples clearly demonstrate the power and efficiency of the parametric technique as well as its superiority over alternative methods (e.g., contour integration). As a method, it is not unknown in the literature (see [1]–[7]) and every now and then there are some instances of its applications as cited in ([8]–[10]). However, it appears that the topic is generally not included in the undergraduate mathematics curriculum.

## REFERENCES

1. Tom M. Apostol, *Mathematical Analysis*, second edition, Addison-Wesley Publishing Company, Reading, Massachusetts, 1975.
2. J. C. Burkhill and H. Burkhill, *A Second Course in Mathematical Analysis*, Cambridge University Press, 1970.
3. Serge Lange, *Undergraduate Analysis*, second edition, Springer-Verlag, New York, 1983.
4. Shanti Narayan, *Integral Calculus*, ninth edition, S. Chand and Co., India, 1975.
5. M. H. Protter and C. B. Morrey, *A First Course in Real Analysis*, Springer-Verlag, New York, 1977.
6. Walter Rudin, *Principles of Mathematical Analysis*, third edition, McGraw-Hill, New York, 1977.
7. Kennan T. Smith, *Primer of Modern Analysis*, Springer-Verlag, New York, 1983.
8. K. F. Anderson, Solution of problem 6417, proposed by R. P. Boas, *Amer. Math. Monthly* 91 (1984), 447.
9. A. C. King, Integration under the integral sign, *Math. Gaz.* (June 1984), 131–134.
10. Kenneth S. Williams, Letter to the Editor, *Amer. Math. Monthly* 91 (1984), 455.

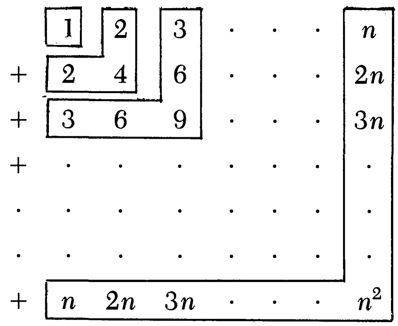
# Proof without Words

## Suggestion 1: Combinatorial Proof




$$= \sum_{i=1}^{n} i + 2\sum_{i=1}^{n} i + \cdots + n\sum_{i=1}^{n} i$$

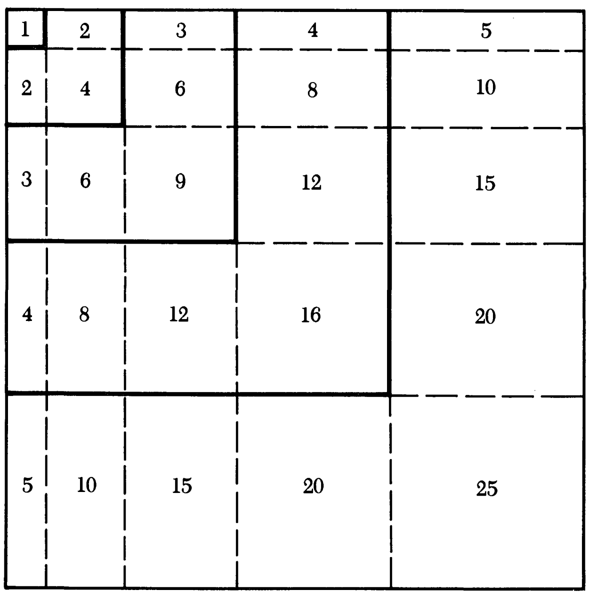
$$= \Big(\sum_{i=1}^{n} i\Big)^2$$




$$= 1^2 + 2(2)^2 + \cdots + n(n)^2$$

$$= \sum_{i=1}^{n} i^3$$

## Suggestion 2: Geometric Proof




$$\left(\sum_{i=1}^{n} i\right)^2 = 1\cdot(1)^2 + 2\cdot(2)^2 + \cdots + n\cdot(n)^2 = \sum_{i=1}^{n} i^3$$

—Farhood Pouryoussefi
*Tehran, Iran*

# The Mathematics of the Casimir Effect

JONATHAN P. DOWLING
Department of Physics
University of Colorado
Boulder, CO 80309

**Introduction** In 1948 H. G. B. Casimir, a mathematical physicist, made a rather startling prediction. He calculated from the theory of quantum electrodynamics that if two massless, perfectly-reflecting plane mirrors were positioned parallel to each other a short distance apart at zero temperature, in a perfect vacuum, then there would arise between them a force of attraction inversely proportional to the fourth power of their separation. The force would be due solely to the interaction of the plates with the vacuum of space [**1**].

That the vacuum could exert a force on anything came as a bit of a surprise to most physicists. (To mathematicians it must sound like just another one of those incoherent ramblings usually attributed to physicists.) In fact, many physicists at first simply dismissed what has become known as the **Casimir effect** as patent nonsense; but that was before, when, in the mid-1950's, the effect was verified in a series of precision experiments. (Two gold mirrors are actually suspended a few microns apart and the force measured directly.) The numbers agreed with those predicted by Casimir—right on the button. (A discussion of this and other odd properties of the vacuum is given in [**2**].)

But what does this all have to do with mathematics? In Casimir's original theoretical calculation there comes a point when one must compare an improper integral $\int f$ to an infinite series $\Sigma f$, where the functional form of $f$ is the same in both cases. In particular one needs to know $\Sigma f - \int f$ in order to arrive at a final numerical result. It turns out that there are two very nice formulas from classical analysis which will do the trick; one is due to Euler and Maclaurin, and the other attributed to the two mathematicians Abel and Plana.

In the course, then, of discussing some interesting physics and some beautiful mathematics I hope to convey the idea that it is possible for a few people (such as a humble mathematical physicist like myself) to do and to enjoy doing both.

**QED and the Casimir effect** Quantum electrodynamics, or more playfully, QED, is a theory which describes how point, charged particles interact with light [**3**]. *Some* of the theoretical predictions of QED agree with experiment to within one part in $10^{12}$ —making it the most accurate physical theory ever invented. Although "thus it has been demonstrated" that "QED" gives some very good results, the theory is also in the habit of predicting things which are quite nonsensical, or so it would seem. To arrive at the very accurate predictions of QED, physicists blatantly add and subtract badly divergent integrals as if they were finite quantities, a prescription which makes even the strongest of mathematicians wince. Worse, QED predicts that the energy density of empty space, at zero temperature, is infinite. Yes, infinite.

QED is like the "girl with the curl" ($\nabla \times \mathbf{A}$?) from the old nursery rhyme; for when she is good she is very, *very* good— but when she is bad she is *horrid*.

So for a time, in order to keep the universe from vaporizing in a blast of radiant vacuum energy, physicists pragmatically subtracted this infinite quantity from their formulas and proceeded as though it did not exist. Then along came Casimir, who resurrected the vacuum energy and haunted the pragmatists with a proof of its physical reality.

Consider the following *Gedankenexperiment*. A rectangular parallelopiped of dimensions $a \times b \times c$ is constructed from perfectly reflecting mirrors and nestled against the coordinate axes in the first octant. (See FIGURE 1a.) Call the energy in the box $E_{\mathrm{I}}$. (Remember $E_{\mathrm{I}} = \infty$.) Now place a perfectly-reflecting, mirrored partition in the box, parallel to the $xy$-plane, a small distance $R$ from the origin. (See FIGURE 1b.) We assume that $R \ll a$, $b$, or $c$. Now the volume of the box is divided into two volumes; call the vacuum energy in the smaller volume $E_{\mathrm{II}}$ and that in the larger $E_{\mathrm{III}}$. One can then operationally define an effective potential energy between the partition and the $xy$-plane as $\Delta E := (E_{\mathrm{II}} + E_{\mathrm{III}}) - E_{\mathrm{I}}$. If the $E$'s were finite, this would of course be zero. But as things stand, $\Delta E$ has the indeterminate form $\infty - \infty$. (Indeterminate, not undefined; like the indeterminate form $0/0$, there is still hope.) Since $\Delta E$ is a function of the separation $R$, an effective force of attraction is gotten using $F = -\,\partial/\partial R(\Delta E)$.

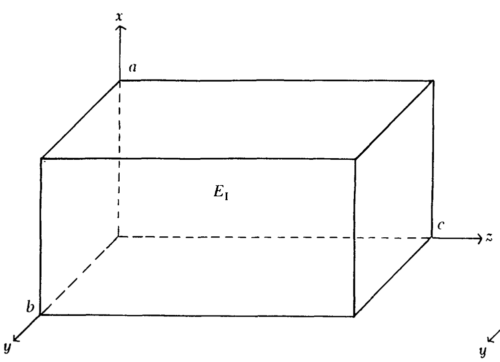



FIGURE 1a

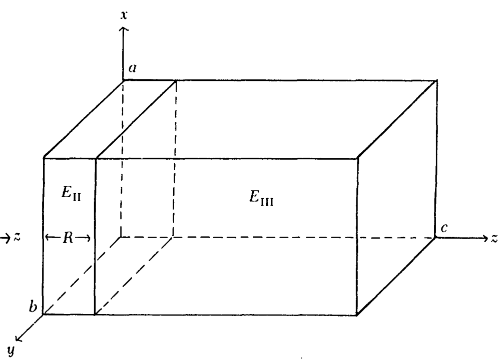



FIGURE 1b

It turns out that $E_{\mathrm{I}}$ and $E_{\mathrm{III}}$ are proportional to divergent improper integrals and $E_{\mathrm{II}}$ to a divergent series. They diverge only cubically, and so it is possible to induce convergence by multiplying the integrands (or summand) by a factor $e^{-\lambda x}(e^{-\lambda n})$ with the understanding that $\lambda \to 0^+$ at the end of the whole calculation. When this is done, one finds it necessary to evaluate the now finite expression

$$\Delta_0^\infty(f) := \sum_{n=0}^{\infty} f(n) - \int_0^\infty f(x)\,dx, \tag{1}$$

where for this problem $f(x) = x^3 e^{-\lambda x}$. More generally, we can replace $f(x)$ by $f(x;\lambda) := f(x)g(x;\lambda)$, where $g(x;\lambda)$ is some cutoff function which satisfies

$$\lim_{x\to\infty} g(x;\lambda) = 0 \qquad \lim_{\lambda\to 0^+} g(x;\lambda) = 1,$$

where the limits converge separately and uniformly. (For more details on the allowable class of functions $f$ see [**5**].)

So we are done with the physics; the mathematicians may stop snickering now, and we can proceed with the investigation of the functional $\Delta(f)$.

**The Euler–Maclaurin summation formula** The first formula we shall consider for $\Delta(f)$ is the **Euler–Maclaurin summation formula** (EMSF). A pedestrian derivation of this formula is given in the complex analysis text of Carrier, Krook, and Pearson [4]. I am not sure what the opposite of *pedestrian* is—but whatever it is, Hardy is it. He gives a comprehensive treatment in his wonderful book, *Divergent Series* [5]. The formula was found independently by Euler in 1732 and Maclaurin in 1742. (Euler's productivity is measured by the truckload, and often his results took a while to diffuse throughout Europe.) But this is not a history lesson, I'm afraid, and so no dusty morsels gathered from yellowed manuscripts in subbasement $e$ of the Göttingen archives will be forthcoming.

Speaking of ancient history, though, I'd like to relate here how I first started thinking about the difference between the sum of a function and its integral. The question arose in my first course on elementary calculus ($n$ years ago, where $n$ is large), when I learned the **integral test for convergent series:** *If $f(x)$ decreases on the interval $(n_0, \infty)$, where $n_0$ is a positive integer, then the series $\sum_{n=n_0}^{\infty} f(n)$ and the integral $\int_{n_0}^{\infty} f(x)\,dx$ converge or diverge together*. Fair enough. But at the time I had a wild idea: "Do they give the same answer?" A quick counterexample dashed my foolish hopes:

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \sum_{n=1}^{\infty}\left(\frac{1}{n} - \frac{1}{n+1}\right) = 1$$

$$\int_1^{\infty} \frac{dx}{x(x+1)} = \int_1^{\infty}\left(\frac{1}{x} - \frac{1}{x+1}\right) dx = \ln(2),$$

where the series is telescopic. (Nice, though, that the same partial-fraction technique allows you to do both.) I soon realized that there was no hope of getting the same answer. The sum is always just an approximation to the integral and the area under the rectangles is never the same as the area under the curve. (See FIGURE 2.) But still, I figured, they had to be related. Not being Euler... I went on to other things.

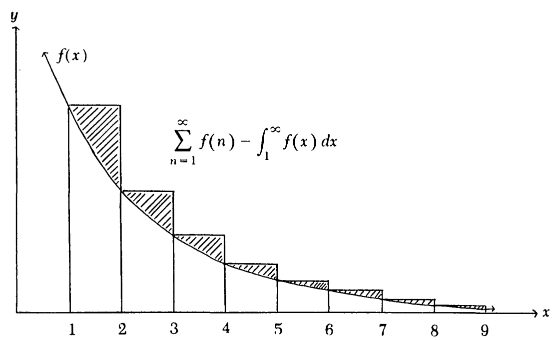



FIGURE 2

So now let us put ourselves in Euler's shoes; the saying goes something like: "As most men breathe, Euler calculated." Euler enjoyed manipulations with series. His instincts may have been to transform the integral term in $\Delta(f)$ into one; so that both terms would be on the same footing. This is easily done. Let $N$ be a positive integer, then

$$\begin{aligned}\Delta_0^N(f) &:= \sum_{n=0}^{N} f(n) - \int_0^N f(x)\,dx \\ &= \frac{1}{2}[f(N)+f(0)] + \sum_{n=0}^{N-1} \frac{1}{2}[f(n+1)+f(n)] - \sum_{n=0}^{N-1}\int_n^{n+1} f(x)\,dx \\ &= \frac{1}{2}[f(N)+f(0)] + \sum_{n=0}^{N-1}\left\{\frac{1}{2}[f(n+1)+f(n)] - \int_n^{n+1} f(x)\,dx\right\}. \quad (2)\end{aligned}$$

So now what? Now a typical Eulerian trick. The expression in { } above can be rewritten as

$$\{\ \} = \frac{1}{2}[f(n+1)+f(n)] - xf(x)\Big|_n^{n+1} + \int_n^{n+1} xf'(x)\,dx$$
$$= -\left(n+\frac{1}{2}\right)[f(n+1)-f(n)] + \int_n^{n+1} xf'(x)\,dx$$
$$= -\left(n+\frac{1}{2}\right)\left[\int_n^{n+1} f'(x)\,dx\right] + \int_n^{n+1} xf'(x)\,dx$$
$$= \int_n^{n+1}\left[x-n-\frac{1}{2}\right]f'(x)\,dx.$$

If we notice that for all $x \in (n, n+1)$ that $[x] = n$, where $[x]$ is the usual "greatest integer less than $x$," then we can write equation (2) as

$$\Delta_0^N(f) = \frac{1}{2}[f(n)+f(0)] + \sum_{n=0}^{N-1}\int_n^{n+1}\left(x-[x]-\frac{1}{2}\right)f'(x)\,dx. \tag{3}$$

The factor $(x - [x] - 1/2)$ multiplying $f'(x)$ is a *sawtooth function* which we denote by $S_1(x)$. We have tacitly assumed $f$ to be once differentiable in order to pull off the integration by parts we did. In order to continue we shall assume that $f$ is $K$ times differentiable on $(0, N)$. If we define $\{S_k(x)\}_{k=1}^K$ to be any sequence of functions on $(0, N)$ satisfying $S_k'(x) = S_{k-1}(x)$, we can integrate the expression (3) for $\Delta_0^N(f)$ by parts $K$ times to get

$$\Delta_0^N(f) = \frac{1}{2}[f(N)-f(0)] + \sum_{k=2}^{K-1}(-1)^k S_k(x) f^{(k)}(x)\Big|_0^N$$
$$+(-1)^{K+1}\int_0^N S_K(x) f^{(K)}(x)\,dx, \tag{4}$$

where $f^{(k)}$ is as usual the $k$th derivative of $f$. It would be nice to pin down the functions $S_k(x)$ at this point. Euler probably reached for (or invented) the **Euler polynomials** $E_k(x)$, which have the convenient property that $E_k'(x)/k! = E_{k-1}(x)/(k-1)!$ Instead it is more common these days to invoke the **Bernoulli polynomials** $B_k(x)$ which have exactly the same property. (See [4] and [5].)

The Bernoulli polynomials can be defined implicitly via a generating function

$$\frac{te^{xt}}{e^t-1} := \sum_{k=0}^{\infty} B_k(x)\frac{t^k}{k!} \qquad (t < 2\pi)$$

and as noted above they have the property that $B_k'(x)/k! = B_{k-1}(x)/(n-1)!$ for $n = 1, 2, \ldots,$ which can be proved with minimal pain from the definition. The **Bernoulli numbers** are defined as $B_k := B_k(0)$ and the first few are: $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, and $B_4 = -1/30$. In fact all odd-order $B$ numbers beyond $B_1$ are identically zero, a fact we shall be needing. We shall be needing also the relationship $B_k = (-1)^k B_k(1)$.

Now if we define the functions $S_k(x) := B_k(x)/k!$ for $k = 2, 3, \ldots$ on the interval $(0, 1)$, and $S_k(x)$ periodic thereafter with period 1—we can then apply all the above stated properties to expression (4) to arrive at the long-awaited

EULER–MACLAURIN SUMMATION FORMULA. *Let $f(x)$ be $K$ times differentiable on the interval $(0, N)$, with $N$ a positive integer. Then*

$$\Delta_0^N(f) = \frac{1}{2}[f(N) + f(0)] + \sum_{m=1}^{M(K)} \frac{B_{2m}}{(2m)!}\left[f^{(2m-1)}(N) - f^{(2m-1)}(0)\right] + R_K(z), \quad (6)$$

*where*

$$M(K) := \begin{cases} \dfrac{K}{2} & (K \text{ even}) \\ \dfrac{K-1}{2} & (K \text{ odd}), \end{cases}$$

*and we have used the mean value theorem to write the remainder as $R_K(z) = (-1)^{K+1}NS_K(z)f^{(K)}(z)$ for some $z \in (0, N)$.*

Recall that Casimir had $f(x) = x^3e^{-\lambda x}$. For this function we clearly may take both $N$ and $K$ to $\infty$. Notice that $f^{(k)}(\infty) = 0$ for all $k = 0, 1, \ldots$ and that $f(0) = 0$, $f'(0) = 0$, $f'''(0) = 6$, and $f^{(k)}(0)$ is proportional to some positive power of $\lambda$ for $k = 4, 5, \ldots$.

If we finally take the limit of $\lambda \to 0^+$ as well, the series then truncates and we are left with the result $\Delta(f) = -1/120$. This number becomes incorporated into a constant of proportionality for the parallel-plate force law $F \propto 1/R^4$, where $R$ was the plate separation. In particular

$$F(R) = \left[-\frac{1}{120}\right]\frac{hcA\pi}{4}\frac{1}{R^4}, \quad (7)$$

where $h$ is Planck's constant, $c$ is the speed of light, and $A$ the area of the plates. The negative sign indicates that the force is one of attraction. The value of $1/120$ and the fact that the force is attractive agrees well with experiments in which two gold, mirrored plates are suspended a small distance apart in a good vacuum and the resultant force measured.

**The Abel–Plana summation formula** Recently some physicists have recalculated the Casimir effect under the additional (hypothetical) assumption that photons have a nonzero rest mass. This modification alters the calculation somewhat and at the point where one tries to apply the EMSF... it does not work. (The series does not truncate nicely as before.) A better formula was needed. So out of the treasure chest of complex analysis was hauled the **Abel–Plana summation formula** (APSF). The only two places I have ever seen this formula are in Hardy's book [5] and in the writings of the "massive photon" people—who also got it from Hardy. We may still venture a guess of the sort: what would Abel have done? With the EMSF already known, and with his substantial experience in the field of integral equations, Abel might have leaned toward an integral formula to express the functional $\Delta(f)$.

As a start in this direction, we now detour to the most beautiful formula in elementary complex analysis

CAUCHY'S INTEGRAL FORMULA. *Let $f(z)$ be analytic on some simply-connected region $R$ and let $\Gamma$ be some closed contour contained in $R$. If $\zeta$ is any point enclosed by $\Gamma$ then*

$$f(\zeta) = \frac{1}{2\pi i}\oint_\Gamma \frac{f(z)}{z - \zeta}\,dz. \quad (8)$$

From this formula all sorts of wonderful things flow. It's beautiful because it tells you in just what way an analytic function is "nice." Analyticity is not determined pointwise, but rather by a coherent unison of all the points throughout the domain. For how else could the function's value at any point $\zeta$ be determined by the function's values along *any* enclosing contour $\Gamma$? But enough meta-mathematics. (I think that Will Rogers once said: "I never meta-mathematician I didn't like.")

So now we recall that the function $\pi\cot(\pi z)$ has an infinite partial fraction expansion, also due to Euler, given by

$$\pi\cot(\pi z)=\sum_{n=-\infty}^{\infty}\frac{1}{z-n}.$$

Hence if we choose contours $\Gamma$ and $\{\Gamma_n\}_{n=0}^{\infty}$ as shown in FIGURE 3, we can apply the Cauchy formula an infinite number of times to get

$$\sum_{n=0}^{\infty}f(n)=\frac{1}{2\pi i}\sum_{n=0}^{\infty}\oint_{\Gamma_n}\frac{f(z)}{z-n}\,dz$$

$$=\frac{1}{2i}\oint_{\Gamma}\cot(\pi z)f(z)\,dz \tag{9}$$

and so now $\sum f$ is a complex-contour integral. To make it a real integral we rotate the upper arm of $\Gamma$ by $\pi/2$ and the lower by $-\pi/2$. The substitution: $z=x\pm i\varepsilon\to\pm iy\pm i\varepsilon$ will do this nicely, since $e^{\pm i\pi/2}=\pm i$. Taking $\varepsilon\to 0^+$ we arrive at

$$\sum_{n=0}^{\infty}f(n)=\frac{1}{2}\operatorname*{res}_{z=0}\,[f(z)\cot(\pi z)]-\frac{1}{2}\int_0^{\infty}[f(iy)-f(-iy)]\cot(i\pi y)\,dy$$

$$=\frac{1}{2}f(0)+\frac{i}{2}\int_0^{\infty}[f(iy)-f(-iy)]\coth(\pi y)\,dy, \tag{10}$$

where the first term arises from a residue at the origin, and we have used the identity between the hyperbolic and ordinary cotangent given by: $\coth(z)=i\cot(iz)$. We have implicitly assumed that $f(z)$ is analytic on the 1/2 plane $\operatorname{Re}\{z\}>-\delta$, for some $\delta>0$. A further condition we shall need is that

$$\lim_{y\to\infty}e^{-2\pi|y|}|f(x+iy)|=0$$

uniformly in any finite interval of $x$. (See [5].)

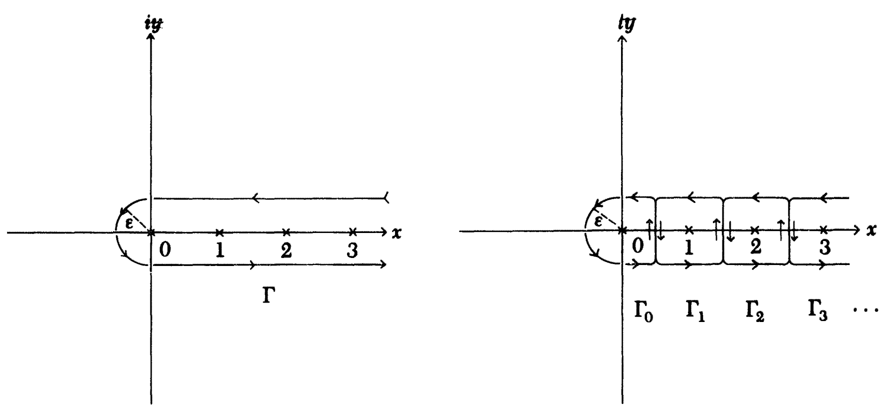



FIGURE 3

Under these same conditions we can transform the $\int f$ portion of $\Delta(f)$ to resemble the form given above for $\Sigma f$. We use the following trick. (A *trick* is something you use only once—a *technique* is something you use more than once.)

$$\begin{aligned}\int_0^\infty f(x)\,dx &= \frac{1}{2}\int_0^\infty f(x)\,dx + \frac{1}{2}\int_0^\infty f(x)\,dx \\ &= \frac{1}{2}\int_0^\infty f(iy)\,d[iy] + \frac{1}{2}\int_0^\infty f(-iy)\,d[-iy] \\ &= \frac{i}{2}\int_0^\infty [f(iy) - f(-iy)]\,dy, \end{aligned} \tag{11}$$

which amounts to formally duplicating the integral and rotating the first one upwards into the complex plane by an angle of $\pi/2$, and the second downwards by the same amount. (As it stands, this manipulation is valid for a smaller class of functions than for which the final result will actually hold. For example: if $f(x) = e^{-x}$ on the left-hand side of expression (11) then the integral on the right-hand side becomes $-\int_0^\infty \sin(x)\,dx$ which does not converge in the normal sense. One can remedy this short-coming by introducing cut-off functions $g(x;\lambda)$ as described before—or avoid the problem altogether by using a more careful and rigorous method of deforming contours, as in [5].) With the identity $\coth(x) - 1 = 1/[e^{2x} - 1]$ we carry out the subtraction in $\Delta(f)$ and we arrive, finally, at

THE ABEL–PLANA SUMMATION FORMULA. *If $f(z)$ is analytic on the 1/2 plane $\mathrm{Re}\{z\} > -\delta$, for some $\delta > 0$, and if both $\Sigma_{n=0}^{\infty} f(n)$ and $\int_0^\infty f(x)\,dx$ converge, then*

$$\Delta_0^\infty(f) = \frac{1}{2}f(0) + i\int_0^\infty \frac{f(iy) - f(-iy)}{e^{2\pi y} - 1}\,dy. \tag{12}$$

This is a rather elegant result. It is also a little bit strange; the difference between a *real* sum and integral of a function of a *real* variable is now expressed as a single integral with the function having a pure imaginary argument. As a check, let us apply this formula to Casimir's function $f(x) = x^3 e^{-\lambda x}$:

$$\begin{aligned}\Delta_0^\infty(x^3 e^{-\lambda x}) &= -2\int_0^\infty \frac{y^3 \cos(\lambda y)}{e^{2\pi y} - 1}\,dy & (\lambda \to 0^+) \\ &= -\frac{2}{(2\pi)^4}\int_0^\infty \frac{t^3}{e^t - 1}\,dt & (t = 2\pi y) \\ &= -\frac{2}{(2\pi)^4}3!\zeta(4) \\ &= -\frac{1}{120},\end{aligned}$$

the same number as before. We have used an integral representation for the **Riemann zeta function**

$$\zeta(n) = \frac{1}{(n-1)!}\int_0^\infty \frac{t^{n-1}}{e^t - 1}\,dt \qquad (n = 1, 2, \ldots) \tag{13}$$

and the fact that $\zeta(4) = \pi^4/90$. (See [4] or [5].) It is curious that the result of $-1/120$ comes from $B_4$ in the EMSF and from $\zeta(4)$ here. Are the two related? Indeed they are; notice that the generating function for the $B_n(x)$ has a form very similar to the integral representation for $\zeta(n)$. It in fact can be shown that

$$\zeta(2m) = \frac{(2\pi)^{2m}}{2(2m)!}|B_{2m}| \qquad (m = 1,2,\dots). \tag{14}$$

(See, for example, [5].) Could this somehow be used to relate the two different summation formulas? The answer is yes.

**A relationship between the two formulas** Recall that in the APSF we assumed that $f(z)$ is analytic in some open 1/2 plane $\mathrm{Re}\{z\} > -\delta$, which includes the origin. Thus $f(z)$ has (appropriately enough) a **Maclaurin expansion** in this region given by: $f(z) = \sum_{n=0}^{\infty} f^{(n)}(0)z^n/n!$ If we insert this expansion into the AP prescription, and further restrict ourselves to functions $f(z)$ which allow the integral in the APSF to converge uniformly so that we may interchange the $\Sigma$ and $\int$, we can integrate the series term-by-term

$$\begin{aligned}\Delta_0^{\infty}(f) &= \frac{1}{2}f(0) + i\sum_{n=0}^{\infty}\left\{\frac{f^{(n)}(0)}{n!}\left[\int_0^{\infty}\frac{(iy)^n - (-iy)^n}{e^{2\pi y} - 1}dy\right]\right\}\\ &= \frac{1}{2}f(0) + i\sum_{k=1}^{\infty}\left\{\frac{(-1)^k f^{(2k-1)}(0)}{(2k-1)!(2\pi)^{2k}}\int_0^{\infty}\frac{t^{2k-1}}{e^t - 1}dt\right\}\\ &= \frac{1}{2}f(0) - i\sum_{k=1}^{\infty}\frac{B_{2k}}{(2k)!}f^{(2k-1)}(0),\end{aligned}$$

which is precisely the Euler-Maclaurin formula! (In the special case where $N$ and $K \to \infty$, i.e., $f(x)$ is integrable in the AP expression and has continuous derivatives of all orders—which of course it does since it is in a region where it is analytic.)

**Conclusions** So we have come full circle. We have seen how a question posed in theoretical physics has led us to consider some really nice, old formulas from classical analysis. These we saw could be used to compute results to take us back to the physics again—this time in the laboratory. It is to be hoped that in the process we have gained insight into why, for some people at least, doing mathematics and physics together can be more stimulating than doing either one separately—not to mention it's downright *fun*.

## REFERENCES

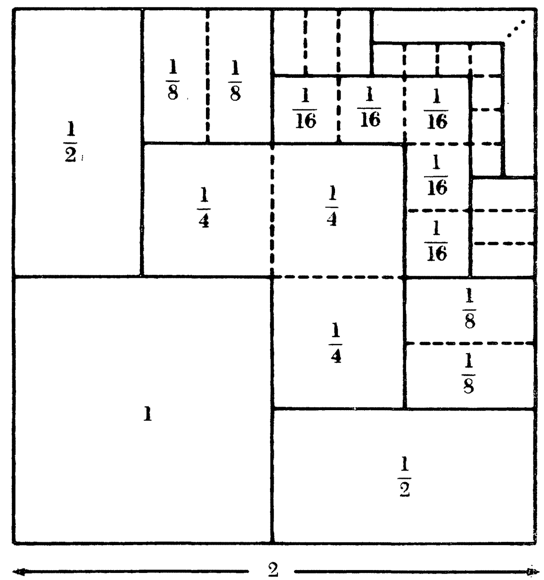
1. H. G. B. Casimir, On the attraction of two perfectly conducting plates, *Proc. Kon. Akad. Ned. Wet.* 51 (1948), 793–795.
2. T. H. Boyer, The classical vacuum, *Scientific American* 253 (1985), 56–63.
3. E. A. Power, *Introductory Quantum Electrodynamics*, Longmans, London, 1964, Appendix 1. (A more accessible account of the Casimir effect, as well as a good beginning test for QED. Warning! Some knowledge of physics may be required.)
4. G. P. Carrier, M. Krook and C. E. Pearson, *Functions of a Complex Variable*, McGraw-Hill, New York, 1966.
5. G. H. Hardy, *Divergent Series*, Oxford Press, London, 1963, Chapter XIII. (The last book that Hardy wrote; a masterpiece of classical analysis.)

## Proof without Words: Differentiated geometric series

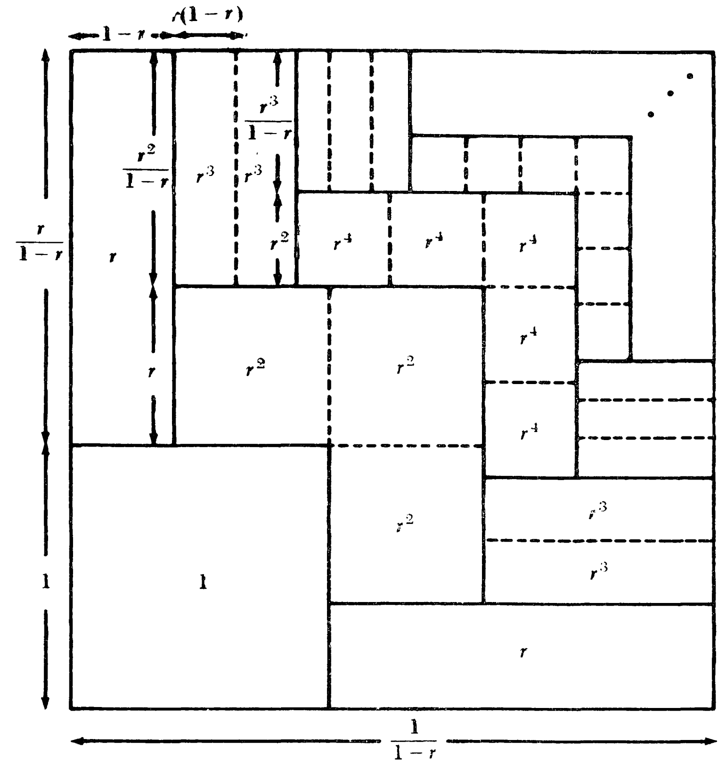
$$1+2\left(\frac{1}{2}\right)+3\left(\frac{1}{4}\right)+4\left(\frac{1}{8}\right)+\cdots=4$$

$$1=4-3(1)$$
$$1+2\left(\frac{1}{2}\right)=4-4\left(\frac{1}{2}\right)$$
$$1+2\left(\frac{1}{2}\right)+3\left(\frac{1}{4}\right)=4-5\left(\frac{1}{4}\right)$$
$$1+2\left(\frac{1}{2}\right)+3\left(\frac{1}{4}\right)+4\left(\frac{1}{8}\right)=4-6\left(\frac{1}{8}\right)$$
$$\vdots$$
$$1+2\left(\frac{1}{2}\right)+3\left(\frac{1}{4}\right)+\cdots+n\left(\frac{1}{2}\right)^{n-1}=4-(n+2)\left(\frac{1}{2}\right)^{n-1}$$

$$1+2\left(\frac{1}{2}\right)+3\left(\frac{1}{4}\right)+\cdots=4$$



$$1+2r+3r^2+4r^3+\cdots=\left(\frac{1}{1-r}\right)^2,\qquad 0\le r<1$$




$$1=\left(\frac{1}{1-r}\right)^2-r(2-r)\left(\frac{1}{1-r}\right)^2$$

$$1+2r=\left(\frac{1}{1-r}\right)^2-r^2(3-2r)\left(\frac{1}{1-r}\right)^2$$

$$1+2r+3r^2=\left(\frac{1}{1-r}\right)^2-r^3(4-3r)\left(\frac{1}{1-r}\right)^2$$

$$\vdots$$

$$1+2r+3r^2+\cdots+nr^{n-1}=\left(\frac{1}{1-r}\right)^2-r^n(n+1-nr)\left(\frac{1}{1-r}\right)^2$$

$$1+2r+3r^2+\cdots=\left(\frac{1}{1-r}\right)^2$$

—Roger B. Nelsen
*Lewis and Clark College*

# Difference Equations and a Principle of Double Induction

RODOLFO SUAREZ
Universidad Autonoma Metropolitana-Iztapalapa
Apartado Postal 55-534, CP. 09340, Mexico

The theory of partial difference equations, that is, difference equations in two or more variables, has received increasing attention during recent years. This is due to the large number of applications in modern engineering fields such as digital filtering, digital picture processing, seismic data processing, X-ray image enhancement, the enhancement and analysis of aerial photographs for detection of forest fires or crop damage, the analysis of satellite weather photos, and multipass processes (machining operations where the material or workpiece involved is processed by a sequence of passes of the processing tool). Many papers have been published in this area, especially in the two-dimensional linear case (2-D systems). An updated bibliography of these topics may be found in [2]–[5].

Another reason for turning our attention to the forgotten field of partial difference equations (DE) is the use of the finite-difference method for the numerical solution of partial differential equations (PDE). This method transforms many continuous problems into a DE (see [3]–[11]). This fact greatly extends the wide range of possible applications of DE. However, in the study of DE associated with PDE, the emphasis has been put on the approximation of solutions. The analysis of DE properties has been omitted. There have been only a few papers concerning the application of DE to the study of continuous systems described by PDE (see [8]).

To illustrate, we present an example of a DE coming from a PDE and one example of a DE modelling a discrete problem.

*Example 1.* Consider the first order hyperbolic equation

$$\frac{\partial T(x,t)}{\partial x} = -\frac{\partial T(x,t)}{\partial t} - T(x,t) \tag{1}$$

with initial and boundary conditions

$$T(x,0) = f_1(x), \quad T(0,t) = f_2(t), \tag{2}$$

which describe some thermal processes, for example, in chemical reactors, heat exchangers and pipe furnaces [8]. $T(x,t)$ is the temperature at $x$(space) $\in [0, x_f]$ and $t$(time) $\in [0, \infty]$.

A DE associated to (1) is obtained if we replace the continuous range of the arguments, $(x,t)$, by a computational grid, $(m\,\Delta x, n\,\Delta t)$, and, instead of the function $T(x,t)$, we consider the discrete function $T(m,n) = T(m\,\Delta x, n\,\Delta t)$. Replacing the partial derivatives $\partial T/\partial x$ and $\partial T/\partial t$ appearing in (1) by the backward and forward difference quotients

$$\frac{\partial T}{\partial x}(x,t) \simeq \frac{T(m,n) - T(m-1,n)}{\Delta x}, \qquad \frac{\partial T}{\partial t}(x,t) \simeq \frac{T(m,n+1) - T(m,n)}{\Delta t},$$

we obtain the partial difference equation

$$T(m,n+1) = a_1 T(m,n) + a_2 T(m-1,n), \tag{3}$$

where $a_1 = 1 - \lambda - \Delta t$, $a_2 = \lambda$ and $\lambda = \Delta t/\Delta x$, with initial and boundary conditions

$$T(m,0) = f_1(m\,\Delta x), \qquad T(0,n) = f_2(n\,\Delta t). \tag{4}$$

*Example* 2. Consider the problem of Parcours [4]: A point is moving a distance of one unit on the $x$ axis for each unit of time, the direction of motion being random at each step. Starting from $x$, it may advance one step to $x+1$, or it may go back one step to $x-1$. The probabilities of the first and second events are $p, q$, respectively, where $p+q=1$. The problem is to find the probability that, if started at $x$, the point in $n$ steps will be at $x_1$. Let $f(x,n)$ be the required probability.

To determine $f(x,n)$ let us remark that if at the first move the point has advanced (gone back) one step, then it has to cover a distance of $x_1 - x - 1(x_1 - x + 1)$ in $n-1$ steps. The probability of both events is $pf(x+1, n-1)$ and $qf(x-1, n-1)$. This shows that the probability satisfies the following difference equation:

$$f(x,n) = pf(x+1,n-1) + qf(x-1,n-1) \tag{5}$$

with the initial condition

$$f(x,0) = \begin{cases} 1 & \text{if } x = x_1 \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

If we have the additional restriction that the point should reach $x_1$ without touching in its movement the point $x=0$, then we must impose the boundary condition

$$f(0,n) = 0. \tag{7}$$

Having constructed the DE modelling these problems, we ask if, by means of (3) and (5), it is possible to calculate, step by step, all the values of the temperature $T(m,n)$ and the probability $f(x,n)$, starting with the initial and boundary conditions (4) and (6)–(7), respectively. In order to answer this question, we shall prove an existence and uniqueness theorem for DE.

In the proof of the existence and uniqueness of solutions for DE, the manner in which initial or boundary conditions are defined is essential. For this reason we consider a particular type of DE which we call temporal difference equations (TDE), as in [10]. These are partial difference equations that can be written explicitly with causal dependence on one of their variables (see definition 1). We call them TDE because in general one of the variables is time. Time is a noninherently anticipatory (causal) variable.

In this note we consider the set of natural numbers as $\mathbb{N} = \{0,1,2,3,\dots\}$.

*Definition* 1. Let $X$ be a nonempty set. Let $F$ be a function from $X^k$ to $X$ and let $f$ be a function from $\mathbb{N}\times\mathbb{N}$ to $X$. The equation

$$f(m,n) = F(f(m+i_1, n+j_1), \dots, f(m+i_k, n+j_k)) \tag{8}$$

is called a temporal difference equation in two independent variables if $i_l$, $j_l$ are integer numbers that satisfy:

1) $j_l \leqslant 0$ for all $l$, such that $1 \leqslant l \leqslant k$.
2) If for any $s$, we have that if $j_s = 0$ then $i_s < 0$.

Condition 1 establishes causal dependence of (8) with respect to the second variable $n$. Condition 2 will ensure, in particular, that $f(m,n)$ does not appear on the right side of the equation, that, in other words, equation (8) is in explicit form.

As we have observed, the difficulty in solving the TDE is related to the existence of initial and boundary conditions which allow us to calculate the function $f(m, n)$ with the aid of (8). In fact we are looking for a minimal set $I \subset \mathbb{N} \times \mathbb{N}$ such that once we know the values of $f$ on $I$, then the function can be extended to $\mathbb{N} \times \mathbb{N}$ by (8).

This question may be interpreted geometrically in terms of computational molecules. For the equation (3) we take a computational molecule consisting of three points labelled $\{(m, n+1),(m, n),(m-1, n)\}$ that can be represented by the diagram in FIGURE 1. For problem (5), we take $\{(x, n),(x-1, n-1),(x+1, n-1)\}$ represented by the diagram in FIGURE 2.

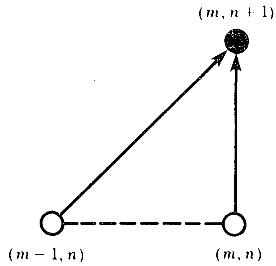



FIGURE 1

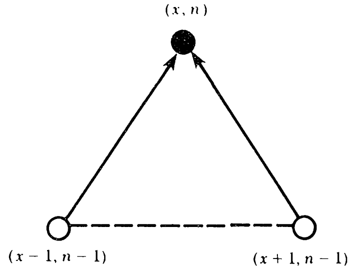



FIGURE 2

To illustrate this idea further we have three more examples. To count the number of basic Lie elements of a given degree (see [7]) we use the difference equation

$$f(r, m)=f(r, m-1)+f(r-1, m). \tag{9}$$

Corresponding to equation (9) we have the computational molecule $\{(r, m), (r, m-1),(r-1, m)\}$ (see FIGURE 3). The Laplace equation can be approximated by the discrete scheme [1]

$$\frac{u(m+1, n)-2u(m, n)+u(m-1, n)}{(\Delta x)^2} + \frac{u(m, n+1)-2u(m, n)+u(m, n-1)}{(\Delta y)^2}=0. \tag{10}$$

The computational molecule is $\{(m, n+1),(m+1, n),(m, n),(m-1, n),(m, n-1)\}$ (see FIGURE 4). For the equation

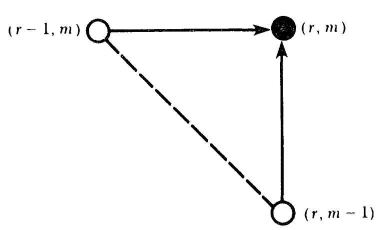



FIGURE 3

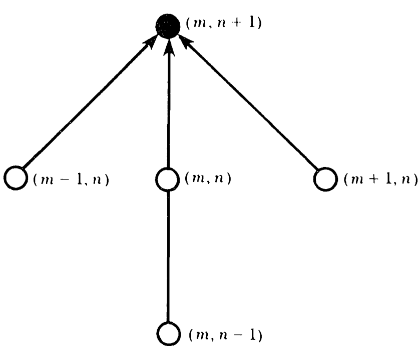



FIGURE 4

$$f(m,n)=F(f(m+1,n),f(m-1,n),f(m,n-1)) \tag{11}$$

the associated computational molecule is $\{(m,n),(m+1,n),(m-1,n),(m,n-1)\}$ (see FIGURE 5). Equation (11) does not satisfy definition 1.

Hence, the geometrical interpretation of the main question is to find an initial set from which, using the respective computational molecule, we can fill all the grid $\mathbb{N}\times\mathbb{N}$ with black balls. A picture of the construction is shown in FIGURE 6, for equation (9). We can observe that there is not an initial set for (11).

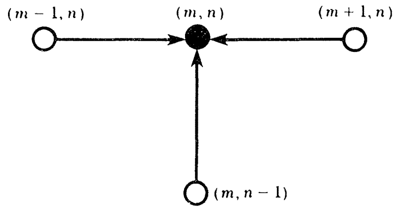



FIGURE 5

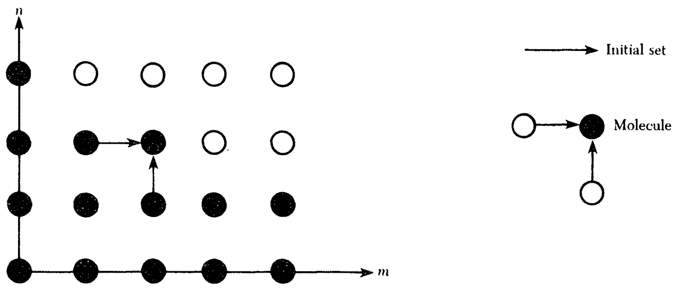



FIGURE 6

The question will be answered in the next theorem. A set of ordered pairs satisfying conditions 1 and 2 of definition 1 will be called a generative set.

*Definition* 2. Let $\{(i_l, j_l)\colon 1\leqslant l\leqslant k\}$ be a generative set. For each pair of natural numbers $(m,n)$ we define $M(m,n)$, the molecule of $(m,n)$, as

$$M(m,n)=\{(m+i_l,n+j_l)\}_{l=1}^{k} \tag{12}$$

and the initial set $I$ as

$$I=\{(i,j)\colon 0\leqslant i\leqslant i_0-1,\ j\in\mathbb{N}\}\cup\{(i,j)\colon 0\leqslant j\leqslant j_0-1,\ i\in\mathbb{N}\} \tag{13}$$

where

$$i_0=\max_{i\leqslant l\leqslant k}\{-i_l,0\},\quad j_0=\max_{i\leqslant l\leqslant k}\{-j_l,0\}.$$

The generative set, the molecule, and the initial set can be easily obtained in each of the introductory examples. Specifically, for equation (3), we have the generative set $\{(0,-1),(-1,-1)\}$; the molecule $M(m,n)=\{(m,n-1),(m-1,n-1)\}$; and the initial set $I=\{(0,j)\colon j\in\mathbb{N}\}\cup\{(i,0)\colon i\in\mathbb{N}\}$. Further, $\{(1,-1),(-1,-1)\}$ is the generative set of (5); its molecule is $M(m,n)=\{(m+1,n-1),(m-1,n-1)\}$; and its initial set is $I=\{(0,j)\colon j\in\mathbb{N}\}\cup\{(i,0)\colon i\in\mathbb{N}\}$. We have for equation (9) the sets

$\{(0,-1),(-1,0)\}$; $M(m,n)=\{(m,n-1),(m-1,n)\}$; and $I=\{(0,j)\colon j\in\mathbb{N}\}\cup\{(i,0)\colon i\in\mathbb{N}\}$. Finally, if we put $u(m,n+1)$ in equation (10) as a function of the grid values of $u$, we obtain the corresponding generative set $\{(1,-1),(0,-1),(-1,-1),(0,-2)\}$; the molecule $M(m,n)=\{(m+1,n-1),(m,n-1),(m-1,n-1),(m,n-2)\}$; and the initial set $I=\{(0,j)\colon j\in\mathbb{N}\}\cup\{(i,j)\colon 0\leqslant j\leqslant 1, i\in\mathbb{N}\}$.

The proof of the existence and uniqueness theorem for TDE will be achieved using an induction principle in two variables (principle of double induction). This is quite natural since mathematical induction is an indispensable tool for establishing results in recursively defined structures, as is the case of equation (8) with its initial set $I$. The principle of double induction to be established is important by itself since a large number of propositions containing two or more integer variables usually arise in combinatorial problems (see [9]).

THEOREM 1. (*Principle of Double Induction*). *Let* $\{(i_l, j_l)\colon 1\leqslant l\leqslant k\}$ *be a generative set and* $I$ *its initial set. Let* $A$ *be a subset of* $\mathbb{N}\times\mathbb{N}$, *that satisfies*

1. $I\subset A$
2. $M(m,n)\subset A$ *implies* $(m,n)\in A$;

*then* $A=\mathbb{N}\times\mathbb{N}$.

*Proof.* We define in $\mathbb{N}\times\mathbb{N}$ the following lexicographical order

$$(m_1,n_1)<(m_2,n_2)\quad\text{if}\quad\begin{cases}n_1<n_2\\ \text{or}\\ n_1=n_2\quad\text{and}\quad m_1<m_2.\end{cases}\tag{14}$$

If we suppose that $\mathbb{N}\times\mathbb{N}\backslash A$ is not empty, then there is a $(m_0,n_0)\in\mathbb{N}\times\mathbb{N}\backslash A$ such that

$$(m_0,n_0)\leqslant(m,n)\quad\text{for all}\quad(m,n)\in\mathbb{N}\times\mathbb{N}\backslash A.\tag{15}$$

Now, because of the lexicographical order definition (14) and by virtue of conditions 1 and 2 from Definition 1 we have

$$(m,n)<(m_0,n_0)\quad\text{for all}\quad(m,n)\in M(m_0,n_0),\tag{16}$$

from which it follows that $M(m_0,n_0)\subset A$. Finally property 2 of $A$ implies that $(m_0,n_0)\in A$ which is a contradiction, so that the theorem is proved.

In order to present this principle in a more familiar form we state:

THEOREM 2. *Let us suppose that for every* $(m,n)\in\mathbb{N}\times\mathbb{N}$, $P(m,n)$ *is a proposition. If we wish to prove that each of the propositions* $P(m,n)$ *is true, it is sufficient to exhibit a generative set, with molecules* $M(m,n)$ *and initial set* $I$, *for which*:

1. $P(m,n)$ *is true for all* $(m,n)\in I$;
2. *if* $P(i,j)$ *is true for all* $(i,j)\in M(m,n)$, *then* $P(m,n)$ *is true.*

*Proof.* If we define the set

$$A=\{(m,n)\in\mathbb{N}\times\mathbb{N}\colon P(m,n)\text{ is true}\}$$

the proof is a consequence of Theorem 1.

We can see Theorem 2 as a consequence of a more general principle. It is the principle of structural induction on well-founded sets, which is a principle of induction applicable to partially ordered sets. (See [**6**].)

Just as the principle of double induction has been proved, we should be able to prove the existence and uniqueness theorem for temporal difference equations. The proof is similar to that found in [12] for a difference equation in one variable.

THEOREM 3. *Let* $\{(i_l, j_l)\colon 1 \leqslant l \leqslant k\}$ *be a generative set and* $I$ *its initial set. Let* $X$ *be a nonempty set and* $F$ *a function from* $X^k$ *to* $X$. *Then, for every function* $f_0\colon I \to X$, *there exists a unique function* $f\colon \mathbb{N} \times \mathbb{N} \to X$ *such that*

$$f(m,n) = F(f(m+i_1, n+j_1), \dots, f(m+i_k, n+j_k))$$

*and which satisfies the initial condition*

$$f(m,n) = f_0(m,n) \quad \textit{for all} \quad (m,n) \in I. \tag{17}$$

*Proof.* The proof scheme is the following:

Let $C$ be the family of all subsets $H$ of $\mathbb{N} \times \mathbb{N} \times X$ which have the properties:

1. For all $(m, n) \in I$, $(m, n, f_0(m, n)) \in H$,
2. If $\{(m + i_l, n + j_l, x_l)\colon 1 \leqslant l \leqslant k\} \subset H$, then $(m, n, F(x_1, \dots x_k)) \in H$,

and let $f = \bigcap\{H\colon H \in C\}$. Then obviously $f$ satisfies the properties 1 and 2 and it is only necessary to prove that $f$ is a function. For that we define the set $S = \{(m, n) \in \mathbb{N} \times \mathbb{N}\colon$ there is just one $x \in X$ such that $(m, n, x) \in f\}$ and using Theorem 1, it follows that $S = \mathbb{N} \times \mathbb{N}$.

To establish uniqueness, suppose there exists another function $g\colon \mathbb{N} \times \mathbb{N} \to X$ with the same properties as $f$ and prove, using the same generative set, that the set $\{(m, n) \in \mathbb{N} \times \mathbb{N}\colon g(m, n) = f(m, n)\}$ is equal to $\mathbb{N} \times \mathbb{N}$.

To conclude, we note that all definitions and theorems admit a natural generalization to the case of TDE of three or more independent variables. However, in the case of other types of partial difference equations we find, even in two dimensions, an increase in the complexity of the problems encountered.

**Acknowledgement.** I would like to express my thanks to J. Goddard, A Sestier, and the anonymous referees for their remarks which improved the presentation of this paper.

## REFERENCES


1. Richard L. Burden, J. Douglas Faires, and Albert C. Reynolds, *Numerical Analysis*, PWS publishers, Boston, 1981, p. 509.
2. John Chiasson and E. Bruce Lee, Control of multipass processes using 2-D system theory, *Proceedings of the 24th Conference on Decision and Control*, Ft. Lauderdale, FL, December, 1985, pp. 1734–1737.
3. Donald Greenspan, *Discrete Models*, Addison-Wesley, Reading, MA, 1973.
4. Charles Jordan, *Calculus of Finite Differences*, Chelsea, New York, 1960, p. 630.
5. Tadeusz Kaczorek, *Two-Dimensional Linear Systems. Lecture Notes in Control and Information Sciences*, Springer-Verlag, Heidelberg, 1985.
6. Leon S. Levy, *Discrete Structures of Computer Science*, John Wiley, New York, 1976, p. 205.
7. Wilhelm Magnus, Abraham Karras, and Donald Solitar, *Combinatorial Group Theory*, Dover, New York, 1976.
8. Wieslaw Marszalek, Two-dimensional state-space discrete models for hyperbolic partial differential equations, *Appl. Math. Modelling* 8 (1984), 11–14.
9. Georgina Pulido and Rodolfo Suárez, Principle of double induction, Reporte de Investigación 136, UAM-Azcapotzalco, 1984.
10. Robert P. Roesser, A discrete state-space model for linear image processing, *IEEE Transactions on Automatic Control*, AC-20, No. 1 (1975), 1–10.
11. Gordon D. Smith, *Numerical Solution of Partial Differential Equations: Finite Difference Methods*, 2nd. ed., Clarendon Press, Oxford, 1978.
12. A. J. White, *Real Analysis: An Introduction*, Addison-Wesley, London, 1968, p. 25.

# The Group of Units in $Z_m$

DAVID J. DEVRIES
Georgia College
Milledgeville, GA 31061

A very elementary proof of part of the primitive root theorem can be given using a basic result from group theory.

THEOREM. *If $m$ is not of the form 2, 4, $p^k$, or $2p^k$ where $p$ is an odd prime, then $Z_m$ has an element of multiplicative order 2 in addition to $m-1$.*

*Proof.* Case I: Suppose $m = 4h$ for some $h > 1$. Let $a = 2h + 1$. Clearly $a$ has order 2 and $1 < a < m - 1$.

Case II: Suppose $m = p^e q^f h$, where $p$ and $q$ are odd primes and $(pq,h) = 1$. Let $j$ be the least positive solution to $(p^e h)x \equiv -2 \mod q^f$. Let $a = p^e hj + 1$. Easily $1 < a < m - 1$. Also $(a+1)(a-1) \equiv 0 \mod m$. Thus $a$ has order 2. □

It follows that for such $m$ the group of units of $Z_m$ cannot be a cyclic group since such groups have at most one element of order 2. It is interesting to show that the rest of the primitive root theorem cannot be proved using this method. In this connection, it is a nice exercise to exhibit a noncyclic group with a unique element of order 2.

---

# Classicists and Constructivists: A Dilemma

F. ALEXANDER NORMAN
University of North Carolina
Charlotte, NC 28223

An existential dilemma (in two parts, naturally):

Conditions for existence:

a) If A is constructivist then for A a dialectic existence proof is not sufficient.
b) If B is classicist then for B a constructive proof is sufficient, but not necessary.

The proof is trivial.

An anecdotal account of a classic constructible dilemma.

I was sitting in the grass outside Cabell Hall reading from Lakatos when I spied Cohn and Klaus ambling down the lawn, as was their wont each weekday during the lunch hour. Cohn, recently converted to a constructivist philosophy, had involved Klaus in an animated discussion.

"I can no longer—I dare say I will NEVER—accept existence merely on the presumption that nonexistence is contradictory." Cohn exclaimed. "To paraphrase Bishop, if one proves that a thing exists then one should show how to find it."

To which Klaus responded, "That's ridiculous! Brouwer notwithstanding, I will NEVER require that existence is dependent upon absolute constructibility. What a notion!"

The two had gone round and round like this—Brouwerian counterexamples to the use of the principle of the excluded middle, the claim that Hilbertian formalization leads to trivialization of knowledge and loss of reference, doubts about the validity of the constructivist assumption that the integers represent universally held, indisputable knowledge, etc.—when one of the two (I could not tell which) spotted an abandoned chess game a dozen or so meters away and directed the attention of the other to it.

Cohn commented, "As the king still appears to be standing, the game most likely has been interrupted. Perhaps you would like to play it out?"

"Indeed, I would!" replied Klaus.

As they neared the board the details of the position [2] became more visible; with hardly a glance one could tell White held a tremendous advantage. Cohn, not having examined the board closely, offered, "My dear Klaus, I wager I can exhibit a mate in two for White!" "You're on!"

After a few moments of analysis Cohn announced, "You are done in, my friend. I can prove there is a mate in two."

"Indeed, I imagine so," grinned Klaus mischievously. "Be so kind as to show me."

"Well, I move Ke6, and if you don't castle then g8 provides the mate."

"So it would. In which case, I should castle, don't you think?" replied Klaus.

"Of course, but if you can castle that means that neither your King nor Rook has moved, so you must have moved the Black pawn from e7 to e5. If so, then as a first move I would capture the pawn *en passant* and, if you then castled, b7 would provide the mate; if you did not castle, g8 mates."

"Very clever. But of course, you, of all people, see the 'obvious' fallacy."

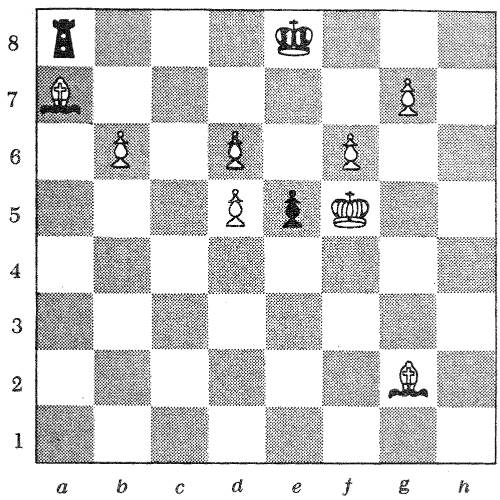
"What?—Either you can castle or you can't. If you can then White can mate. If you can't then White can mate. What's the big deal?" queried Cohn, not a little put out at this offhanded dismissal of his brilliant—well, ingenious—analysis.

"Such liberal use of non-constructivist principles—the excluded middle, dialectic existence . . . ."

"Now wait a minute, Klaus! You are the one who said you'd NEVER make constructivity a prerequisite for existence."

"And it was you, Cohn, who swore NEVER to accept an argument which . . . ."

It was getting late and I had a class to teach. Tucking *Proofs and Refutations* [1] under my arm I headed toward the nearest building. I could not help smiling. As the door closed quietly behind, I could hear the muted, yet lively, voices of Cohn and Klaus still engaged in their absurd debate.



Admittedly this dialogue is an imperfect presentation of the philosophical prejudices of constructivists and classicists. The implication here is that the working mathematician does not let philosophical considerations deprive him or her of results. However, a more compelling theme emerges; one that has been eloquently articulated by Lakatos and others. As one colleague wrote: "...the current chess position is not the whole story of the game, similarly no formal axiomatic presentation of a branch of mathematics is the whole story of the objects with which it deals."

REFERENCES

1. Imre Lakatos, *Proofs and Refutations: The Logic of Mathematical Discovery*, Cambridge University Press, Cambridge, 1986.
2. This position, in which it can be *proved* that White can play and mate in two but the mate can not be exhibited, is the creation of Raymond Smullyan, and can be found in his delightful *The Chess Mysteries of Sherlock Holmes*, Alfred A. Knopf, New York, 1982, p. 103.

# Solutions of $x^n + y^n = z^{n+1}$

LARRY HOEHN
Austin Peay State University
Clarksville, TN 37044

In students' first introduction to Fermat's Last Theorem (i.e, there are no nontrivial integral solutions to $x^n + y^n = z^n$ for $n > 2$), they frequently encounter several related diophantine equations. Some of these, such as $x^4 + y^4 = z^2$ (in most texts), $x^4 + 4y^4 = z^2$ [3], and $x^4 - y^4 = z^2$ [1], have no nontrivial solutions; whereas others, such as $x^2 + y^2 = z^2$ (Pythagorean triplets), $x^4 + 3y^4 = z^4$ [3], and $x^2 + y^2 = z^3$ [1], [3], have an infinite number of integral solutions.

There is, however, another equation, $x^n + y^n = z^{n+1}$, that is almost identical to Fermat's and which is very easy to deal with. To show that this equation has an infinite number of solutions, for any positive integers $a$ and $b$ define $z_0 = a^n + b^n$, $x_0 = az_0$, and $y_0 = bz_0$. By elementary algebra $x_0, y_0, z_0$ is a solution of $x^n + y^n = z^{n+1}$.

It may be noted also that $x = a(a^{n+1} + b^{n+1})$ and $y = b(a^{n+1} + b^{n+1})$ provide one solution in integers to $ax^n + by^n = z^{n+1}$.

One of the referees has pointed out that work on the equation $x^n + y^n = z^{n+1}$ goes back to at least 1908. In 1914, L. Aubry [2] gave the following more general result: If $\gcd(m, n) = 1$, then $x^m + y^m = z^n$ has the solution $x = a(a^m + b^m)^u$, $y = b(a^m + b^m)^u$, $z = (a^m + b^m)^v$, where $nv - mu = 1$.

REFERENCES

1. David M. Burton, *Elementary Number Theory*, Allyn and Bacon, Inc., Boston, 1980.
2. Leonard Eugene Dickson, *History of the Theory of Numbers*, vol. II, Chelsea Publishing Company, New York, 1952.
3. Kenneth H. Rosen, *Elementary Number Theory and Its Applications*, Addison-Wesley Publishing Company, Reading, Massachusetts, 1984.

# PROBLEMS

LOREN C. LARSON, *editor*
St. Olaf College

GEORGE GILBERT, *associate editor*
St. Olaf College

## Proposals

*To be considered for publication, solutions should be received by May 1, 1990.*

**1332.** *Proposed by Bernardo Recamán, Instituto Alberto Merani, Columbia, and Michael Zielinski, St. John's University, Collegeville, Minnesota.*

The permutation $(8,5,2,7,4,1,6,3)$ has the property that its partial sums, taken left to right modulo 8, also form a permutation, namely, $(8,5,7,6,2,3,1,4)$. For each $n$, let $P_n$ denote the set of all permutations of $\{1,2,\dots,n\}$ with this property. That is, $\mathbf{x}=(x_1,x_2,\dots,x_n)\in P_n$ if and only if $\mathbf{x}$ is a permutation of $\{1,2,\dots,n\}$ and $s(\mathbf{x})\equiv(s_1,s_2,\dots,s_n)$ is also a permutation where $s_i=x_1+x_2+\cdots+x_i \pmod n$, $1\leqslant s_i\leqslant n$.

a. Prove that $P_n\neq\varnothing$ if and only if $n$ is even.

b. Prove that if $\mathbf{x}\in P_n$ then $s(\mathbf{x})\notin P_n$.

**1333.** *Proposed by K. R. S. Sastry, Addis Ababa, Ethiopia.*

Prove that the quadrilateral formed by the adjacent quadrisectors of the angles of a rhombus is a square.

**1334.** *Proposed by William P. Wardlaw, U. S. Naval Academy, Annapolis, Maryland.*

Let

$$\mathbf{B}=\begin{pmatrix}5&5&2\\5&5&2\\1&1&6\end{pmatrix}.$$

Show that $\mathbf{B}$ is not the classical adjoint of any matrix with real entries.

---

ASSISTANT EDITORS: CLIFTON CORZATT and THEODORE VESSEY, *St. Olaf College. We invite readers to submit problems believed to be new and appealing to students and teachers of advanced undergraduate mathematics. Proposals should be accompanied by solutions, if at all possible, and by any other information that will assist the editors and referees. A problem submitted as a Quickie should have an unexpected, succinct solution. An asterisk (*) next to a problem number indicates that neither the proposer nor the editors supplied a solution.*

*Solutions should be written in a style appropriate for Mathematics Magazine. Each solution should begin on a separate sheet containing the solver's name and full address.*

*Solutions and new proposals should be mailed in duplicate to Loren C. Larson, Department of Mathematics, St. Olaf College, Northfield, MN 55057.*

**1335.** *Proposed by David Callan, University of Bridgeport, Bridgeport, Connecticut.*

Suppose a sequence of integers is defined by $u_0, u_1 \in \mathbf{Z}$, and $u_n = u_{n-1} - (n-1)u_{n-2}$ for $n \geqslant 2$. Show that $n-2$ divides $u_n$ for $n \geqslant 3$.

**1336.** *Proposed by N. J. Lord, Tonbridge School, Kent, England.*

It is known that the number of nonzero terms in the expansion of the $n \times n$ determinant having zeros in the main diagonal and ones elsewhere is equal to $d_n$, the number of derangements of $n$ objects. (Recall that a derangement is a permutation which 'moves' every object.)

a. Show that the numerical value of the determinant is $(-1)^{n-1}(n-1)$.

b. For which $n$ is it possible to allocate $\pm$ signs to the ones in such a way that $d_n$ is numerically equal to the value of the resulting determinant?

**1337.** *Proposed by G. Behforooz, Utica College of Syracuse University, Utica, New York.*

Prove that the least common multiple of $\{1, 2, 3, \dots, n\}$ is greater than or equal to $2^{n-1}$.

# Quickies

*Answers to the Quickies are on page 349.*

**Q754.** *Proposed by Harry D. Ruderman, Bronx, New York.*

Let $ABCD$ be a skew quadrilateral, let $P, Q, R, S$ be points on sides $AB, BC, CD, DA$, respectively; let $a = AP$, $b = BQ$, $c = CR$, $d = DS$, $p = PB$, $q = QC$, $r = RD$, $s = SA$. Show that $P, Q, R, S$ are coplanar if $abcd = pqrs$.

**Q755.** *Proposed by Chester Palmer, Auburn University, Montgomery, Alabama.*

Let $S$ be a set, $*$ a binary operation on $S$. Write a nontrivial equation (i.e., one not true of all $S$ and $*$) which must be true if $*$ is either commutative or associative on $S$.

**Q756.** *Proposed by Martin Feuerman, New Jersey Medical School, Newark, New Jersey, and Allen R. Miller, Naval Research Lab, Washington, D.C.*

Find

$$\lim_{x \to 0^+} \frac{x}{\sqrt{1 - e^{-x^2}}}.$$

# Solutions

## A Steiner-type inequality — December 1988

**1307.** *Proposed by Zun Shan and Edward T. H. Wang, Wilfrid Laurier University, Canada.*

Let $\triangle ABC$ be a triangle with altitudes $h_a$, $h_b$, and $h_c$, and let $P$ be a point inside (or on the boundary of) the triangle. Show that

$$PA+PB+PC \geqslant \frac{2}{3}(h_a+h_b+h_c)$$

with equality if and only if $\triangle ABC$ is equilateral and $P$ is its center.

*Solution by John Oman, University of Wisconsin-Oshkosh, Oshkosh, Wisconsin.*

$h_a \leqslant PA + p_a$ where $p_a$ is the perpendicular distance from $P$ to $a$, since $h_a$ is the shortest distance from $A$ to $a$. Also, equality holds if and only if $P$ is on the altitude. Then

$$h_a+h_b+h_c \leqslant PA+PB+PC+p_a+p_b+p_c \leqslant \frac{3}{2}(PA+PB+PC)$$

from the Erdös-Mordell Theorem. Equality holds for the Erdös-Mordell inequality if and only if $\triangle ABC$ is equilateral and $P$ is its center. (See N. D. Kazarinoff, *Geometric Inequalities*, Random House, 1961, for a discussion of the Erdös-Mordell Theorem.)

*Also solved by Nicolas K. Artémiadis (Greece), Anna Boettcher and Václav Konečný, Mostafa Ghandehari, Francis M. Henderson, Levent Kitis, John Oman (second solution), Michael Vowe (Switzerland), and the proposers.*

## Compositions and bijections — December 1988

**1308.** *Proposed by Nicholas A. Martin, Indiana University, Bloomington, Indiana.*

Let $\mathbf{N}$ be the set of natural numbers $\{1,2,3,\ldots\}$, let $g\colon \mathbf{N}\to\mathbf{N}$ be a bijection, and $a\in\mathbf{N}$ be an odd number.

a. Prove that there is no function $f$ such that $f(f(n)) = g(n)+a$ for all $n\in\mathbf{N}$.

b*. What if $a$ is even?

*Solution by David Callan, University of Bridgeport, Bridgeport, Connecticut.*

Let $h(n)=g(n)+a$. A function $f\colon \mathbf{N}\to\mathbf{N}$ as required, if it exists, must clearly be one-to-one. Any bijection from $\mathbf{N}$ to a subset of $\mathbf{N}$ can be represented, essentially uniquely, as a product of disjoint cycles. For example, if $g$ is the identity on $\mathbf{N}$ and $a=2$, then $h=(1,3,5,7,\ldots)(2,4,6,8,\ldots)$. More generally, for arbitrary $a\in\mathbf{N}$ and $g\in S(\mathbf{N})$, $h=(1,\ldots)(2,\ldots)\ldots(a,\ldots)h_1$, where $h_1$ consists of (possibly vanishing) cycles, each either finite or doubly-infinite (infinite in both directions). The first $a$ cycles listed are semi-infinite (infinite in one direction), reflecting the fact that the range of $h$ is $\mathbf{N}\setminus\{1,2,\ldots,a\}$. The required $f$ exists if and only if $h$ is a "square."

The following formulas characterize the square of the various kinds of cycles:

$$(a_1, b_1, a_2, b_2, \dots)^2 = (a_1, a_2, \dots)(b_1, b_2, \dots)$$
$$(\dots a_1, b_1, a_2, b_2, \dots)^2 = (\dots a_1, a_2, \dots)(\dots b_1, b_2, \dots)$$
$$(a_1, b_1, \dots, a_m, b_m)^2 = (a_1, \dots, a_m)(b_1, \dots, b_m)$$
$$(a_1, b_1, \dots, a_m, b_m, a_{m+1})^2 = (a_1, \dots, a_{m+1}, b_1, \dots, b_m).$$

It follows that the required $f$ exists if and only if the classes of cycles of $h$ that are semi-infinite, doubly-infinite, or finite of even length, each have an even (or infinite) cardinality. In particular, $a$ must be even. Also, the $h$ above has the required square root ($f = (1,2,3,4,\dots)$ is one such), and for $g = (1,2,3,4,5,6)$ and $a = 2$, we have $h = (1,4,7,9,11,13,\dots)(2,5,8,10,12,14,\dots)(3,6)$, whence no such $f$ exists.

*Also solved by N. K. Artemiadis (Greece), S. F. Barger, Duane M. Broline, Roger B. Eggleton (Brunei Darussalam), Russell Jay Hendel, L. R. King, Walter W. Kirchherr, Leroy F. Meyers, Andreas Müller (Switzerland), Mark Leeney, Marijo LeVan, Michael M. Parmenter, Arlo W. Schurle, Shailesh Shirali (India), Keith Wayland and Sham Oltikar (Puerto Rico), A. Zulauf, and the proposer.*

Meyers and Zulauf pointed out that the proposition is incorrect if $f$ is allowed to take negative integer values. For if $f_1$ is defined by $f_1(n) = -g(n)$ if $n > 0$, $f_1(n) = -n + a$ if $n \leqslant 0$, then $f_1(f_1(n)) = g(n) + a$ for all $n \in \mathbf{N}$.

Among the many generalizations provided was the following, obtained independently by S. F. Barger and L. R. King:

Let $a, k \in \mathbf{N}$ and let $g\colon \mathbf{N} \to \mathbf{N}$ be a bijection. Let $f^{(k)}$ denote the $k$th composite of $f\colon \mathbf{N} \to \mathbf{N}$. If $f^{(k)}(n) = g(n) + a$ for all $n \in \mathbf{N}$, then $a$ is divisible by $k$.

**Conics and period-six decimals** **December 1988**

**1309.** *Proposed by Edward Kitchen, Santa Monica, California.*

Let $0.a_1a_2a_3a_4a_5a_6\dots$, be a period six decimal expansion which determines six distinct lattice points $(a_i, a_{i+1})$, $i = 1,2,3,4,5,6$, with subscripts taken modulo 6. Prove that these six points lie pairwise symmetrically on a central conic (ellipse or hyperbola) if and only if $a_i + a_{i+3} = b$, $i = 1,2,3$, for some positive integer $b$. (See Problem 1248 and its solution, this MAGAZINE (October 1987), p. 245.)

*Solution by L. Kuipers, Sierre, Switzerland.*

Let the six distinct lattice points be $A = (a_1, a_2)$, $B = (a_2, a_3)$, $C = (a_3, b - a_1)$, $D = (b - a_1, b - a_2)$, $E = (b - a_2, b - a_3)$, $F = (b - a_3, a_1)$. The common center point of $AD$, $BE$, and $CF$ is $P = (b/2, b/2)$. Now there is one central conic $\Gamma$ with $P$ as center and containing $A$, $B$, and $C$. Then automatically the points $D$, $E$, and $F$ are on $\Gamma$.

Conversely, if the six points $(a_i, a_{i+1})$, $i = 1,2,3,4,5,6$, are on a conic with center $(c_1, c_2)$, in a symmetrical way, so that $(a_1, a_2)$ is symmetric with $(a_4, a_5)$, $(a_2, a_3)$ is symmetric with $(a_5, a_6)$, and $(a_3, a_4)$ is symmetric with $(a_6, a_1)$, say, then we have $a_1 + a_4 = a_2 + a_5 = a_3 + a_6 =$ a constant.

*Also solved by Michael Vowe (Switzerland) and the proposer.*

**A series of lcm powers** **December 1988**

**1310.** *Proposed by C. B. Khare, Cambridge, England.*

Show that $\sum_{n=1}^{\infty} 1/(\mathrm{lcm}\{1,2,3,\dots,n\})^{\delta}$ converges for all $\delta > 0$. ($\mathrm{lcm}\{1,2,3,\dots,n\}$ is the least common multiple of the set $\{1,2,3,\dots,n\}$.)

I. *Solution by J. Foster, Weber State College, Ogden, Utah.*

For any prime $p \leqslant n$, let $p^i$ be the highest power of $p$ that divides $\{1,2,3,\dots,n\}$.

Then $p^{i+1} > n$, or $p^i > n/p$. Hence

$$\operatorname{lcm}\{1,2,3,\dots,n\} > \prod_{p\leqslant n} n/p \geqslant \prod_{p\leqslant\sqrt{n}} n/p \geqslant \prod_{p\leqslant\sqrt{n}} n/\sqrt{n} = \sqrt{n}^{\,\pi(\sqrt{n})},$$

where $\pi(n)$ denotes the number of primes $\leqslant n$. Since there are infinitely many primes, $\delta\pi(\sqrt{n}) \geqslant 4$ for sufficiently large $n$ and the given series converges by comparison to $\Sigma_{n=1}^{\infty} 1/n^2$.

II. *Solution by G. A. Edgar, Ohio State University, Columbus.*

From Hardy and Wright, *An Introduction to the Theory of Numbers*, Oxford Science Publications, fifth edition, 1979, p. 342, we find that $\operatorname{lcm}\{1,2,3,\dots,n\} \geqslant 2^{n/4}$. (Also, see Problem #1337 in this issue.) So, for any $\delta > 0$,

$$\sum_{n=1}^{\infty} \frac{1}{(\operatorname{lcm}\{1,2,3,\dots,n\})^{\delta}} \leqslant \sum_{n=1}^{\infty} \left(\frac{1}{2^{\delta/4}}\right)^{n}.$$

The latter is a convergent geometric series, so the left-hand series converges by the comparison test.

III. *Solution by Robert Doucette, Iowa City, Iowa.*

Let $u_n = \operatorname{lcm}\{1,2,\dots,n\}$, and let $p_i$ denote the $i$th prime (so that $p_1 = 2$). Choose $m$ an integer $\geqslant 2$ such that $1/(m-1) < \delta$. Suppose that $n > \Pi_{i=1}^{m} 2^i$. It is clear that $u_n = \Pi_{i=1}^{r} p_i^{\alpha_i}$, where $r$ is chosen so that $p_r \leqslant n < p_{r+1}$ and $\alpha_i$ is an integer such that $p_i^{\alpha_i} \leqslant n < p_i^{\alpha_i+1}$, $i = 1,2,\dots,r$. By Bertrand's postulate, $2p_i > p_{i+1}$, from which we get that $p_{m+1} < 2^{m+1}$, so that $m+1 \leqslant r$ and $2 \leqslant r-m+1$. We also have $2^i p_{r-i+1} > p_{r+1}$, $i = 1,2,\dots,m$. Therefore,

$$u_n \geqslant \prod_{i=1}^{m} 2^i p_{r-i+1} > p_{r+1}^{m} > n^m.$$

Hence,

$$\left(\frac{1}{u_n}\right)^{\delta} < \left(\frac{1}{u_n}\right)^{1/(m-1)} < \left(\frac{1}{n}\right)^{m/(m-1)}$$

Since $\Sigma_{n=1}^{\infty}(1/n)^{m/m-1}$ converges, we may conclude that $\Sigma_{n=1}^{\infty}(1/u_n)^{\delta}$ converges as well.

IV. *Solution by H.-J. Seiffert, Berlin, West Germany.*

If $k = \Pi_{i=1}^{\infty} p_i^{a_{k,i}}$ is the prime factorization of $k \in \mathbf{N}$, then

$$\operatorname{lcm}\{1,2,\dots,n\} = \prod_{i=1}^{\infty} p_i^{\max\{a_{1,i},a_{2,i},\dots,a_{n,i}\}}. \tag{1}$$

Thus the sequence $(\operatorname{lcm}\{1,2,\dots,n\})_n$ is nondecreasing and Cauchy's condensation test shows that it suffices to prove that

$$\sum_{n=1}^{\infty} \frac{2^n}{(\operatorname{lcm}\{1,2,\dots,n\})^{\delta}} \tag{2}$$

converges for all $\delta > 0$. By Bertrand's Postulate, there exists a prime $p_n$ with $2^n < p_n < 2^{n+1}$. Equation (1) shows that $p_n$ divides $\operatorname{lcm}\{1,2,\dots,2^{n+1}\}$ but does not divide $\operatorname{lcm}\{1,2,\dots,2^n\}$. Again by using (1), one gets

$$\frac{\operatorname{lcm}\{1,2,\dots,2^n\}}{\operatorname{lcm}\{1,2,\dots,2^{n+1}\}} \leqslant \frac{1}{p_n} < \frac{1}{2^n}.$$

Now the convergence of (2) follows by applying the ratio test, since

$$\frac{2^{n+1}}{\left(\operatorname{lcm}\{1,2,\dots,2^{n+1}\}\right)^{\delta}} \frac{\left(\operatorname{lcm}\{1,2,\dots,2^{n}\}\right)^{\delta}}{2^n} < \frac{1}{2^{n\delta-1}}.$$

*Also solved by Hamza Yousef A. Ahmad (student; Kuwait), Seung-Jin Bang (Korea), G. Behforooz, W. E. Briggs, David Callan, Chico Problem Group, Jesse Deutsch, Roger B. Eggleton (Brunei Darussalam), H. K. Krishnapriyan, Kee-Wai Lau (Hong Kong), Harvey Schmidt, Jr., C. Wildhagen (The Netherlands), A. Zulauf (New Zealand), and the proposer.*

**A product of sums inequality** **December 1988**

**1311.** *Proposed by Mihály Bencze, Braşov, Romania.*

Let $0 < m \leqslant x_1, x_2, \dots, x_{2n+1} \leqslant M$. Prove that

$$(M-m)^2 + 4Mm\left(\sum_{k=1}^{2n+1} x_k\right)\left(\sum_{k=1}^{2n+1}\frac{1}{x_k}\right) \leqslant (2n+1)^2(M+m)^2.$$

*Solution by The Chico Problem Group, California State University, Chico.*

Consider the function

$$G = \left(\sum_{k=1}^{2n+1} x_k\right)\left(\sum_{k=1}^{2n+1}\frac{1}{x_k}\right).$$

A straightforward calculation shows that

$$\frac{\partial^2 G}{\partial x_j^2} = \sum_{\substack{k=1\\k\neq j}}^{2n+1} \frac{2x_k}{x_j^2},$$

so that

$$\frac{\partial^2 G}{\partial x_j^2} > 0, \qquad 1 \leqslant j \leqslant 2n+1.$$

If $x_k$ is fixed for all $k \neq j$, then, as a function of $x_j$, $G$ cannot have an interior maximum; the maximum must occur for $x_j = m$ or $M$. Since $j$ is arbitrary, we conclude that the maximum of $G$ occurs for each $x_k = m$ or $M$, $1 \leqslant k \leqslant 2n+1$.

Now we may assume that $x_k = m$ for $1 \leqslant k \leqslant p$, and $x_k = M$ for $p+1 \leqslant k \leqslant 2n+1$. Let $q = 2n+1-p$. We must show that

$$(M-m)^2 + 4Mm(pm+qM)\left(\frac{p}{m}+\frac{q}{M}\right) \leqslant (p+q)^2(M+m)^2.$$

Algebraic simplification yields the equivalent inequality

$$(M-m)^2 \leqslant (M-m)^2(p-q)^2. \tag{1}$$

Since $p+q=2n+1$ is odd, $p-q$ is odd, and the proof is complete.

Note that if the number of $x_i$'s is even, the inequality (1), and thus the original inequality, fail if $p=q$.

*Also solved by J. C. Binz (Switzerland), David Callan, Russell Jay Hendel, William Moser (Canada), Richard E. Pfiefer, and the proposer.*

# Answers

*Solutions to the Quickies on p. 344.*

**A754.** Place masses $bcd$, $qrs$, $brs$, and $bcs$ at $A$, $B$, $C$, and $D$, respectively. Then the center of mass for the entire system must be on both $PR$ and $QS$, so $PR$ must intersect $QS$.

An easy indirect proof can establish the converse. A similar result holds when $P$, $Q$, $R$, and $S$ are on extensions of the sides.

**A755.** Perhaps the simplest such equation is $x*(x*x)=(x*x)*x$.

**A756.** Applying L'Hôpital's rule directly to this limit yields the same indeterminate form. Although one can find the limit without using L'Hôpital's rule by expanding $e^{-x^2}$ as an infinite series, an easier way is to consider the limit

$$\lim_{x\to 0}\frac{x^2}{1-e^{-x^2}}=1,$$

which follows directly by L'Hôpital's rule. This yields

$$\lim_{x\to 0^+}\frac{x}{\sqrt{1-e^{-x^2}}}=1$$

as well as

$$\lim_{x\to 0^-}\frac{x}{\sqrt{1-e^{-x^2}}}=-1.$$

# REVIEWS

PAUL J. CAMPBELL, *editor*
Beloit College

*Assistant Editor: Eric S. Rosenthal, West Orange, NJ. Articles and books are selected for this section to call attention to interesting mathematical exposition that occurs outside the mainstream of the mathematics literature. Readers are invited to suggest items for review to the editors.*

Mathematics: As American as apple pi, *Time* (19 June 1989) 37.

Will copywriters ever get tired of tiresome puns? Be prepared for more of the same, as the record for calculation of digits of pi seesaws back and forth between US and Japanese teams. In June the Chudnovsky brothers of the US held the lead, at 480 million digits. Stay tuned.

Cipra, Barry A., How the Grinch stole mathematics, *Science* (11 August 1989) 595.

For some time "insiders" have known that hypergeometric series provide the key to mechanical and relatively effortless proofs of some kinds of combinatorial identities (e.g., virtually all that involve binomial coefficients). Cipra (whom *Science* probably prohobited from using the "technical" word "hypergeometric") notes that the secret is not only out, but you can use a special command in the computer algebra system Macsyma to do the remaining dirty work. Far from "taking the fun out of mathematics," the new approach has led to the discovery of several new identities.

Cipra, Barry A., Say it again in plain algebra, *Science* 245 (15 September 1989) 1190-1191.

Susan Landau (University of Massachusetts) has found a general algorithm to reduce any arithmetic expression with nested radicals to its least-nested equal. The problem is ancient, and the tools involved (the appropriate splitting field and its Galois group) have been well-known for a century; but it is today's computer algebra systems that gave impetus to solving the problem.

Steen, Lynn Arthur (ed.), *Reshaping College Mathematics: A Project of the Committee on the Undergraduate Program in Mathematics*, MAA, 1989; x + 125 pp (P).

Integrated and updated collection of excerpts from a decade of reports. "We don't need to look far for sound goals and objectives for college mathematics. Most of what we need can be found in this volume. What remains to be done—as much now as ever—is to find effective means for turning ideals into practice."

Connolly, Paul, and Teresa Vilardi (eds.), *Writing to Learn Mathematics and Science*, Teachers College Press, 1989; xviii + 307 pp.

Despite the title, almost all of the 23 essays in this book concern how students' understanding of *mathematics* can be enhanced through writing; three of the contributions were presented at the January 1988 MAA meeting in Atlanta.

Elmer-Dewitt, Philip, Time for some fuzzy thinking, *Time* (25 September 1989) 79.

Fuzzy sets allow objects to belong to them with any membership value between 0 and 1, and fuzzy mathematics proceeds from there to devise further fuzzy objects and their calculi. Those of us who may have dismissed fuzzy sets, fuzzy logic, fuzzy systems, etc., may need to take a second look. The Japanese are hot on the fuzzy trail, having patented fuzzy auto transmission and antiskid braking systems, introduced a fuzzy stock-market investment program (perhaps not so surprising after all!), with a fuzzy auto-focus camera to come. This is not to mention their fuzzy automobile traffic controller and the forthcoming fuzzy shower system (which prevents scaldings). Watch out for fuzzy expert systems! But is it really true that "the only barrier remaining to wider use of fuzzy logic is the philosophical resistance of the West"??

Cunningham-Green, Ray, Geometry, shoemaking, and the milk tray problem, *New Scientist* (12 August 1989) 50-53.

Popular exposition of the mathematics of fitting irregular shapes together economically, with industrial applications.

Stewart, Ian, Mock theta conjectures, *Nature* 339 (1 June 1989) 341.

Ramanujan defined by infinite series some functions that he claimed share many of the properties of Jacobi's theta functions ("the simplest elements out of which elliptic functions can be constructed"—Morris Kline). But are Ramanujan's "mock theta" functions just Jacobi's functions in disguise? Dean Hickerson has proved that they are indeed different.

Bracewell, Ronald N., The Fourier transform, *Scientific American* (June 1989) 86-95.

Where the Fourier transform involves both real and imaginary components, the Hartley transform (named after its discoverer in the 1920s) provides the same information by using only a real component. Each has a corresponding "fast" version, the fast Fourier being due to Cooley and Tukey in 1965, and the fast Hartley to Bracewell himself in 1984. The latter transform, dubbed by its inventor the Bracewell transform, offers greater efficiency in calculation than the fast Fourier; it was granted one of the first software patents by the US Patent Office (a fact curiously unmentioned in the article).

Cipra, Barry A., Math team vaults over prime record, *Science* (25 August 1989) 815; Computing a prime champion, *Science News* (16 September 1989) 191.

The largest known prime is now $391{,}581 \times 2^{216{,}193} - 1$, which has 65,807 decimal digits. The successful search, by a team from Amdahl Corporation, was made possible by their developing a new algorithm for high-speed convolution (a key ingredient in multidigit multiplication), which itself may have applications in seismic research, weather prediction, aeronautical simulation, and the search for pulsars.

Apostol, Tom M., *The Theorem of Pythagoras*, MAA, 1988; 15-min videotape, $25, plus *Program Guide and Workbook*, 30 pp, $4 (P).

First in a series of modules that use computer animation to help teach basic concepts in secondary mathematics. The video motivates the Pythagorean theorem, describes its history, and shows several visual proofs. The graphics are excellent (the video received several gold medals at film festivals). The tape was funded by ACM/SIGGRAPH and produced by Mathematics!, a project at the California Institute of Technology.

# NEWS AND LETTERS

## EIGHTEENTH U.S.A. MATHEMATICAL OLYMPIAD (Solutions)

1. For each positive integer $n$, let

$$S_n = 1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \tfrac{1}{n},$$

$$T_n = S_1 + S_2 + S_3 + \cdots + S_n,$$

$$U_n = \frac{T_1}{2} + \frac{T_2}{3} + \frac{T_3}{4} + \cdots + \frac{T_n}{n+1}.$$

Find, with proof, integers $0 < a,b,c,d < 1000000$ such that $T_{1988} = aS_{1989} - b$ and $U_{1988} = cS_{1989} - d$.

*Sol.* More generally, we look for simple formulas expressing $T_n$ and $U_n$ in terms of $S_{n+1}$. First, we have

$$\begin{aligned}
T_n &= 1 \\
&\quad + \left(1 + \tfrac{1}{2}\right) \\
&\quad + \left(1 + \tfrac{1}{2} + \tfrac{1}{3}\right) \\
&\quad + \cdots \\
&\quad + \left(1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \tfrac{1}{n}\right) \\
&= \tfrac{n}{1} + \tfrac{n-1}{2} + \tfrac{n-2}{3} + \cdots + \tfrac{1}{n} \\
&= \left(\tfrac{n+1}{1} - \tfrac{1}{1}\right) + \left(\tfrac{n+1}{2} - \tfrac{2}{2}\right) + \left(\tfrac{n+1}{3} - \tfrac{3}{3}\right) + \\
&\qquad \cdots + \left(\tfrac{n+1}{n} - \tfrac{n}{n}\right) \\
&= (n+1)\left(1 + \tfrac{1}{2} + \tfrac{1}{3} + \cdots + \tfrac{1}{n}\right) \\
&\qquad - \left(\tfrac{1}{1} + \tfrac{2}{2} + \tfrac{3}{3} + \cdots + \tfrac{n}{n}\right) \\
&= (n+1)\left(S_{n+1} - \tfrac{1}{n+1}\right) - n \\
&= (n+1)S_{n+1} - (n+1).
\end{aligned}$$

Now, since $\dfrac{T_n}{n+1} = S_{n+1} - 1$, we have

$$\begin{aligned}
U_n &= \tfrac{T_1}{2} + \tfrac{T_2}{3} + \tfrac{T_3}{4} + \cdots + \tfrac{T_n}{n+1} \\
&= (S_2 - 1) + (S_3 - 1) + (S_4 - 1) + \cdots + \\
&\qquad (S_{n+1} - 1) \\
&= (S_2 + S_3 + S_4 + \cdots + S_{n+1}) - n \\
&= (-S_1 + T_n + S_{n+1}) - n \\
&= (-1 + [(n+1)S_{n+1} - (n+1)] + S_{n+1}) - n \\
&= (n+2)S_{n+1} - (2n+2).
\end{aligned}$$

Thus $T_{1988} = 1989S_{1989} - 1989$ and $U_{1988} = 1990S_{1989} - 3978$, and so we can take $(a, b, c, d) = (1989, 1989, 1990, 3978)$.

Note. The problem can also be solved by guessing the formulas for $T_n$ and $U_n$ as a result of experimentation with small $n$, and then proving them by mathematical induction.

2. The 20 members of a local tennis club have scheduled exactly 14 two-person games among themselves, with each member playing in at least one game. Prove that within this schedule there must be a set of 6 games with 12 distinct players.

*Sol.* Let $k$ be the largest integer for which the schedule includes a set of $k$ games with $2k$ distinct players.

No two of the other 20-2$k$ club members can play each other, for otherwise we could for a set of $k$+1 scheduled games with 2$k$+2 distinct players, contrary to the maximality of $k$. Since each of these other 20-2$k$ club members plays in at least one game on the schedule, but no two in the same game, there must be at least 20-2$k$ distinct games on the schedule in addition to the $k$ that we selected for our maximal set. Together these comprise (20-2$k$)+$k$ games within a 14-game schedule, therefore $14 \geq (20-2k)+k$ and $k \geq 6$.

Note. The numbers 14, 6 and 20 can be replaced by $m$, $n$ and $m+n$, respectively.

3. Let $P(z) = z^n + c_1 z^{n-1} + c_2 z^{n-2} + \cdots + c_n$ be a polynomial in the complex variable $z$, with real coefficients $c_k$. Suppose that $|P(i)| < 1$. Prove that there exist real numbers $a$ and $b$ such that $P(a+bi) = 0$ and $(a^2 + b^2 + 1)^2 < 4b^2 + 1$.

*Sol.* Let the roots of the polynomial be $r_1$, $r_2$, ..., $r_n$, so that

$$P(z) = (z - r_1)(z - r_2)\cdots(z - r_n).$$

The given inequality now becomes

$$|i - r_1||i - r_2|\cdots|i - r_n| < 1.$$

Since $|i - r| = \sqrt{1 + r^2} \geq 1$ for all real $r$, the product of the $|i - r_j|$ for <u>non-real</u> $r_j$ must be less than 1. Furthermore, these non-real roots $r_j$ occur in conjugate pairs, since the polynomial coefficients $c_k$ are real numbers. Therefore we must have $|i - r_j||i - \bar{r}_j| < 1$ for some root $r_j = a + bi$. Thus $P(a + bi) = 0$, and

$$\begin{aligned}
1 &> |i - (a + bi)||i - (a - bi)| \\
&= \sqrt{a^2 + (1 - b)^2}\,\sqrt{a^2 + (1 + b)^2} \\
&= \sqrt{(a^2+b^2+1) - 2b}\,\sqrt{(a^2+b^2+1) + 2b} \\
&= \sqrt{(a^2+b^2+1)^2 - 4b^2},
\end{aligned}$$

so

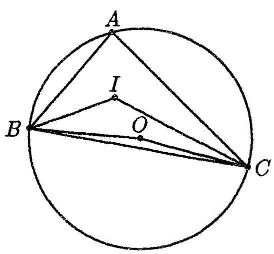
$$4b^2 + 1 > (a^2 + b^2 + 1)^2.$$

4. *Let* $ABC$ *be an acute-angled triangle whose side lengths satisfy the inequalities* $AB < AC < BC$. *If point* $I$ *is the center of the inscribed circle of triangle* $ABC$ *and point* $O$ *is the center of the circumscribed circle, prove that line* $IO$ *intersects line segments* $AB$ *and* $BC$.

*Sol.* We will use $A,B,C$ to denote the angles of the triangle, each less than 90°. Since a circular arc measures twice the inscribed angle it subtends, points $A,B,C$ divide the circumcircle into three arcs, each less than 180°. It follows that point $O$ lies in the interior of $\triangle ABC$. Of course, so does point $I$. Now if we could show that

$$(*) \qquad \angle OBC < \angle IBC \text{ and } \angle OCB < \angle ICB,$$

then it would follow that point $O$ lies in the interior of $\triangle$IBC, and hence that line $IO$ intersects segment $BC$.




The two inequalities (*) are established in a relatively straightforward manner.

(1) Angle $A$ subtends an arc with central angle $BOC$, so $\angle BOC = 2A$. Since $\triangle BOC$ is isosceles,
$\angle OBC = \angle OCB = \frac{1}{2}(180° - \angle BOC)$
$= 90° - A$.

(2) The incenter of $\triangle ABC$ lies on each of the angle bisectors, so
$\angle IBC = B/2$ and $\angle ICB = C/2$.

(3) We recall that the measures of the sides of a triangle occur in the same order as the measures of their opposite angles. Therefore
$C < B < A$.

Using observations (1), (2) and (3), we obtain

$$90° - A = \frac{A+B+C}{2} - A = \frac{C+(B-A)}{2} < \frac{C}{2} < \frac{B}{2}$$

and

$\angle OBC = \angle OCB = 90° - A < \angle ICB < \angle IBC$, thereby establishing (*).

Likewise, if we could show that $\angle$IAB < $\angle$OAB and $\angle$IBA < $\angle$OBA, then point $I$ would be located in the interior of $\triangle$OAB and the line $IO$ would intersect segment $AB$. These two inequalities follow from

$$90° - C = \frac{A+B+C}{2} - C = \frac{A+(B-C)}{2} > \frac{A}{2} > \frac{B}{2},$$

completing the proof.

5. Let $u$ and $v$ be real numbers such that

$$(u + u^2 + u^3 + \cdots + u^8) + 10u^9$$
$$= (v + v^2 + v^3 + \cdots + v^{10}) + 10v^{11} = 8.$$

Determine, with proof, which of the two numbers, $u$ or $v$, is larger.

*Sol.* Define $F_n(x) = (x + x^2 + x^3 + \cdots + x^{n-1}) + 10x^n = \frac{x + 10x^n(0.9 - x)}{1 - x}$. Since we are comparing $u$ and $v$, it might be helpful to consider the given equation with $u = v$. We find that

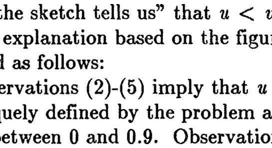
$$F_9(x) - F_{11}(x) = 9x^9 - x^{10} - 10x^{11}$$
$$= 10x^9(0.9 - x)(1 + x),$$

so there are in fact three places where $F_9 = F_{11}$:

$F_9(u) = F_{11}(v) = 0$ when $u = v = 0$;
$F_9(u) = F_{11}(v) = 9$ when $u = v = 0.9$;
$F_9(u) = F_{11}(v) = -10$ when $u = v = -1$.

This leads us to a meaningful sketch of $F_9(x)$ and $F_{11}(x)$, based on the following five observations:

(1) $F_9(x) - F_{11}(x) = 10x^9(0.9-x)(1+x) > 0$ when $0 < x < 0.9$;
(2) $F_9(0) = F_{11}(0) = 0$;
(3) $F_9(0.9) = F_{11}(0.9) = 9$;
(4) $F_9(x)$ and $F_{11}(x)$ are increasing continuous functions of $x$ when $x \geq 0$;
(5) $F_9(x)$ and $F_{11}(x)$ are both negative when $x < 0$.




Now "the sketch tells us" that $u < v$, and a formal explanation based on the figure could proceed as follows:

Observations (2)-(5) imply that $u$ and $v$ are uniquely defined by the problem and that they lie between 0 and 0.9. Observation (1) now implies that $F_{11}(u) < F_9(u) = F_{11}(v)$. Finally, observation (4) implies that $u < v$.

Note. To three decimal places, $u = 0.883$ and $v = 0.884$.

ACKNOWLEDGEMENTS

*In addition to our associate editors, the following have assisted the* Magazine *by refereeing papers during the past year. We appreciate the time and care they have given.*

Acu, Dumitru, *Sibiu, Romania*
Alavi, Yousef, *Western Michigan University*
Albers, Donald J., *Menlo College*
Alder, Henry L., *University of California, Davis*
Altmann, Simon L., *Brasenose College, Oxford*
Andrews, George E., *Pennsylvania State University, University Park*
Archibald, Thomas, *Acadia University*
Askey, Richard, *University of Wisconsin, Madison*
Aspray, William, *IEEE, New York*
Barbeau, Edward J., *University of Toronto*
Barksdale, James B., Jr., *Western Kentucky University*
Barnette, Dave, *University of California, Davis*
Barría, José, *Santa Clara University*
Baxley, John V., *Wake Forest University*
Beals, Richard W., *Yale University*
Beaver, Olga R., *Williams College*
Beezer, Robert A., *University of Puget Sound*
Beineke, Lowell, *Indiana-Purdue University at Ft. Wayne*
Boas, Ralph P., *Northwestern University*
Braden, Bart, *Northern Kentucky University*
Bremser, Priscilla, *Middlebury College*
Broline, Duane M., *Eastern Illinois University*
Brualdi, Richard A., *University of Wisconsin, Madison*
Brunson, Barry W., *Western Kentucky University*
Cabaniss, Sharon, *San Jose State University*
Castellanos, Dario, *Universidad de Carabobo, Valencia, Venezuela*
Chakerian, G.D., *University of California, Davis*
Chung, Fan, *Bell Communications Research, Morristown, NJ*
Clark, Dean, *University of Rhode Island*
Conrey, J. Brian, *Oklahoma State University*
Coxeter, H.S.M., *University of Toronto*
Cundy, H. Martyn, *Kendal, Cumbria, England*
deBouvère, Karel L., *Santa Clara University*
DeTemple, Duane W., *Washington State University*
Devaney, Robert L., *Boston University*
Diaconis, Persi, *Stanford University*
Dobbs, David E., *University of Tennessee*
Drobot, Vladimir, *Santa Clara University*
Drucker, Thomas, *Dickinson College*
Edgar, Gerald A., *Ohio State University*
Edgar, Hugh M.W., *San Jose State University*
Eisenberg, Bennett, *Lehigh University*
Epp, Susanna, *De Paul University*
Eves, Howard, *Oviedo, FL*
Farris, Frank, *Santa Clara University*
Farris, Steven L., *Ball State University*
Feroe, John, *Vassar College*
Firey, William J., *Oregon State University*
Fisher, Evan, *Lafayette College*
Fisher, J. Chris, *University of Regina*
Foster, Lorraine L., *California State University, Northridge*
Fraser, Craig, *Victoria College*
Fredricksen, Harold M., *Naval Postgraduate School*
Gallian, Joseph A., *University of Minnesota, Duluth*
Galovich, Steve, *Carleton College*
Gardner, Martin, *Hendersonville, NC*
Gibbs, Richard A., *Fort Lewis College*
Gilbert, William J., *University of Waterloo*
Goldstein, Jerome A., *Tulane University*
Goldston, Dan, *San Jose State University*
Golomb, Solomon, *University of Southern California*
Green, Euline "Bo", *Abilene Christian University*
Gretsky, Neil E., *University of California, Riverside*
Groeneveld, Richard A., *Iowa State University*
Grosshans, Frank D., *West Chester University*
Grünbaum, Branko, *University of Washington*
Guy, Richard K., *University of Calgary*
Hall, James E., *Westminster College*
Halmos, Paul R., *Santa Clara University*
Harary, Frank, *New Mexico State University*
Harborth, Heiko, *Technical University Carolo Wilhelmina, West Germany*
Hayashi, Elmer K., *Wake Forest University*
Henderson, James, *Colorado College*
Henrion, Claudia, *Middlebury College*
Hickling, Fred, *Santa Clara University*
Hillman, A. P., *Albuquerque, NM*
Hilton, Peter, *SUNY—Binghamton*
Honsberger, Ross, *University of Waterloo*
Howard, F.T., *Wake Forest University*
Jacobson, Bernard, *Franklin & Marshall College*
Jamison, Robert E., *Clemson University*
Janke, Steven, *Colorado College*
Jordan, Jim, *Washington State University*
Kalman, Dan, *Rancho Palos Verdes, California*
Katz, Victor J., *Silver Spring, Maryland*
Kiltinen, John O., *Northern Michigan University*
Kimble, Gerald, *University of Nevada*
Klamkin, Murray S., *University of Alberta*
Klosinski, Leonard, *Santa Clara University*
Konhauser, Joseph D.E., *Macalester College*
Kuzmanovich, James, *Wake Forest University*
Lagarias, Jeffrey C., *AT&T Bell Labs*
Lange, Lester H., *Moss Landing Laboratories*
Lautzenheiser, Roger G., *Rose-Hulman Institute of Technology*
Levasseur, Kenneth M., *University of Lowell*
Levine, Alan, *Franklin & Marshall College*
LoBello, Anthony, *Allegheny College*
Logothetti, Dave, *Santa Clara University*
Loud, Warren S., *University of Minnesota, Minneapolis*
Merrill, Kathy, *Colorado College*

Mitchem, John, *San Jose State University*
Monson, Barry, *University of New Brunswick*
Moore, David S., *Purdue University*
Morrison, John, *Towson State University*
Morrison, Kent E., *California Polytechnic State University*
Mugler, Dale, *University of Akron*
McCartney, Philip, *Northern Kentucky University*
McCleary, John, *Vassar College*
McCormick, Phillip, *Santa Clara University*
McCurley, Kevin, *IBM Research Division, San Jose, CA*
McDermot, Richard, *Allegheny College*
McDonald, Joanne M., *University of Regina*
McHugh, Joseph F., *LaSalle University*
Nelsen, Roger B., *Lewis & Clark College*
Niven, Ivan, *University of Oregon*
Notz, William I., *Ohio State University*
Osofsky, Barbara, *Rutgers University*
Owens, Robert, *Lewis & Clark College*
Pederson, Jean, *Santa Clara University*
Pfiefer, Richard, *San Jose State University*
Propp, James G., *Berkeley, CA*
Rabinowitz, Stanley, *Westford, MA*
Raimi, Ralph A., *University of Rochester*
Read, R. C., *University of Waterloo*
Reay, John, *Western Washington University*
Rebman, Kenneth R., *California State University, Hayward*
Redheffer, Raymond, *University of California at Los Angeles*
Redmond, Don, *Southern Illinois University*
Reznick, Bruce, *University of Illinois, Urbana-Champaign*
Rigby, John F., *University of Wales, College of Cardiff*
Riley, John, *Bloomsburg University*
Robertson, Jack M., *Washington State University*
Rosenstein, George M., *Franklin & Marshall College*
Rosier, Ronald, *Georgetown University*
Ross, Kenneth, *University of Oregon*
Ross, Peter, *Santa Clara University*
Roth, Richard, *University of Colorado*
Roy, Ranjan, *Beloit College*
Sagan, Hans, *North Carolina State University*
Sallee, G. T., *University of California, Davis*
Salwach, Chester J., *Lafayette College*
Sanchez, David A., *Lehigh University*
Sandefur, James T., *Georgetown University*
Saunders, Sam C., *Washington State University*
Schumer, Peter, *Middlebury College*
Senechal, Marjorie, *Smith College*
Shuchat, Alan H., *Wellesley College*
Sloughter, Dan, *Furman University*
Smolarski, Dennis, *Santa Clara University*
Spencer, Joel H., *Courant Institute of Mathematical Sciences*
Spiegel, Murray, *East Hartford, CT*
Stone, Alexander P., *University of New Mexico*
Sved, Marta, *University of Adelaide*
Thurston, Hugh, *University of British Columbia*
Tingley, Daryl, *University of New Brunswick*
Townsend, Douglas W., *Indiana-Purdue University at Ft. Wayne*
Tucker, Alan, *SUNY at Stony Brook*
Vanden Eynden, Charles L., *Illinois State University*
Van de Wetering, R.L., *San Diego State University*
Wakin, Shirley, *University of New Haven*
Warner, Hoyt D., *Western New England College*
Watkins, John J., *Colorado College*
Webb, William, *Washington State University*
Weiss, Asia Ivic, *York University*
Wessel, Alan, *Santa Clara University*
Wilker, J. B., *University of Toronto, Scarborough*
Wilson, Stephen, *Northern Arizona University*
Wise, M. Norton, *University of California, Los Angeles*
Yorke, James A., *University of Maryland, College Park*
Zerger, Monte, *Adams State College*
Zoeller, Bettina, *Western Kentucky University*
Zorn, Paul, *St. Olaf College*

# INDEX TO VOLUME 62

## AUTHORS

## TITLES



## PROBLEMS

P, S, Q, *and* C *refer to Proposals, Solutions, Quickies, and Comments, respectively. Page numbers are given in parentheses. Thus,* P1312(58) *refers to proposal number* 1312 *which appears on page* 58.

## NEWS & LETTERS

U.S. Postal Service

**STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION**

*Required by 39 U.S.C. 3685)*

| 1A. Title of Publication | 1B. PUBLICATION NO. | 2. Date of Filing |
|---|---|---|
| MATHEMATICS MAGAZINE | 0 0 2 5 5 7 0 X | Sept. 29, 1989 |

| 3. Frequency of Issue | 3A. No. of Issues Published Annually | 3B. Annual Subscription Price |
|---|---|---|
| bi-monthly except July and August | Five | Library $45<br>Indiv. Member $13 |

4. Complete Mailing Address of Known Office of Publication *(Street, City, County, State and ZIP+4 Code) (Not printers)*
1529 Eighteenth St., N.W., Washington, D.C. 20036

5. Complete Mailing Address of the Headquarters of General Business Offices of the Publisher *(Not printer)*
1529 Eighteenth St.,N.W., Washington, D.C. 20036

6. Full Names and Complete Mailing Address of Publisher, Editor, and Managing Editor *(This item MUST NOT be blank)*

Publisher *(Name and Complete Mailing Address)*
The Mathematical Association of America, 1529 Eighteenth St.,N.W., Washington, D.C. 20036

Editor *(Name and Complete Mailing Address)*
Gerald L. Alexanderson, Santa Clara Univ., Dept. of Mathematics, Santa Clara, CA 95053

Managing Editor *(Name and Complete Mailing Address)*
Harry Waldman, MAA, 1529 Eighteenth St.,N.W., Washington, D.C. 20036

7. Owner *(If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given. If the publication is published by a nonprofit organization, its name and address must be stated.) (Item must be completed.)*

| Full Name | Complete Mailing Address |
|---|---|
| The Mathematical Association of America,Inc. | 1529 Eighteenth St.,N.W. |
| | Washington, D.C. 20036 |

8. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages or Other Securities *(If there are none, so state)*

| Full Name | Complete Mailing Address |
|---|---|
| None | |

9. For Completion by Nonprofit Organizations Authorized To Mail at Special Rates *(DMM Section 423.12 only)*
The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes *(Check one)*

(1) ☒ Has Not Changed During Preceding 12 Months
(2) ☐ Has Changed During Preceding 12 Months
*(If changed, publisher must submit explanation of change with this statement.)*

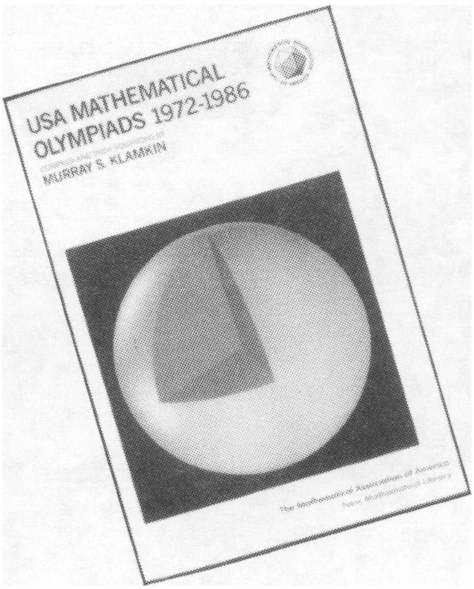
| 10. Extent and Nature of Circulation *(See instructions on reverse side)* | Average No. Copies Each Issue During Preceding 12 Months | Actual No. Copies of Single Issue Published Nearest to Filing Date |
|---|---|---|
| A. Total No. Copies *(Net Press Run)* | 17,153 | 17,633 |
| B. Paid and/or Requested Circulation<br>1. Sales through dealers and carriers, street vendors and counter sales | 0 | 0 |
| 2. Mail Subscription *(Paid and/or requested)* | 14,734 | 14,358 |
| C. Total Paid and/or Requested Circulation *(Sum or 10B1 and 10B2)* | 14,734 | 14,358 |
| D. Free Distribution by Mail, Carrier or Other Means Samples, Complimentary, and Other Free Copies | 444 | 522 |
| E. Total Distribution *(Sum of C and D)* | 15,178 | 14,880 |
| F. Copies Not Distributed<br>1. Office use, left over, unaccounted, spoiled after printing | 1,975 | 2,753 |
| 2. Return from News Agents | 0 | 0 |
| G. TOTAL *(Sum of E, F1 and 2—should equal net press run shown in A)* | 17,153 | 17,633 |

11. **I certify that the statements made by me above are correct and complete**

Signature and Title of Editor, Publisher, Business Manager, or Owner
Kay Lamont
Membership Manager

PS Form **3526**, Feb. 1989 — *(See instructions on reverse)*

# USA MATHEMATICAL OLYMPIADS




Every year 100 of the most mathematically talented high school students in the country compete in the USA Mathematical Olympiad (USAMO). The USAMO is the third stage of a three-tiered mathematical competition for high school students in the United States and Canada that begins with the AHSME taken by over 400,000 students, continues with the American Invitational Mathematics Exam involving 2,000 students, and culminates in the 100-contestant USAMO.

**USA MATHEMATICAL OLYMPIADS 1972-1986, PROBLEMS AND SOLUTIONS**

Compiled by Murray S. Klamkin

People delight in working on problems "because they are there, " for the sheer pleasure of meeting a challenge. This is a book full of such delights. In it, Murray S. Klamkin brings together 75 original USA Mathematical Olympiad (USAMO) problems for years 1972-1986, with many improvements, extensions, finger exercises, open problems, references and solutions, often showing alternative approaches. The problems are coded by subject and solutions are arranged by subject as an aid to those interested in a particular field. Contains a glossary of frequently used terms and theorems, and a comprehensive bibliography with items numbered and referred to in brackets in the text. The problems are intriguing and the solutions elegant and informative. Students and teachers will enjoy working these challenging problems. Indeed all those those who are mathematically inclined will find many delights and pleasant challenges in this book.

180 pp., 1988, ISBN-0-88385-634-4

List: $13.50 MAA Member: $12.50

Catalog Number NML-33

*Order from:*

The Mathematical Association of America
1529 Eighteenth Street, N. W.
Washington, D. C. 20036
(202) 387-5200

# CONTENTS




**THE MATHEMATICAL ASSOCIATION OF AMERICA**
**1529 Eighteenth Street, NW**
**Washington, D.C. 20036**