LAWS OF FORM

$$\zeta(s) = \Sigma n^{-s} = \Pi(1-p^{-s})^{-1}$$

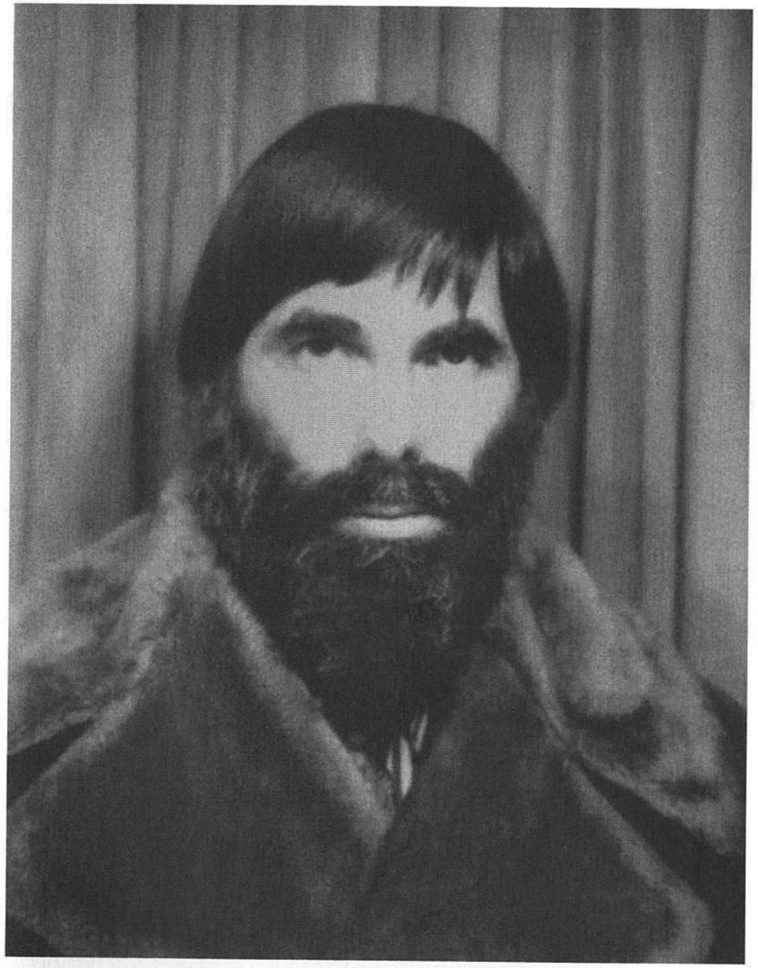
can be zero for nonreal s
only of the form $\frac{1}{2}+\mathrm{i}y$.

THE NEW EDITION OF THIS CLASSIC WITH THE FIRST-EVER PROOF OF RIEMANN'S HYPOTHESIS

GEORGE SPENCER-BROWN

BOHMEIER VERLAG

LAWS OF FORM



The author from a passport photograph in 1973

GEORGE SPENCER-BROWN

LAWS OF FORM

BOHMEIER VERLAG

Printing History

First published in London by George Allen and Unwin Ltd., 17th April 1969. Second impression 1971.

First American edition, Julian Press, September 1972. The Library of Science Bookclub edition October 1972. The Computer and Information Science Bookclub edition November 1972.

Bantam paperback edition November 1973.

Reprinted in August 1977 by Julian Press, a division of Crown Publishers, Inc..

Reprinted in 1979 by E.P.Dutton, a Division of Elsevier-Dutton Publishing Co., Inc., New York.

First German edition by Joh. Bohmeier Verlag 1997. Reprinted with additional text 1999.

This revised fifth English edition produced in 2011 by Bohmeier Verlag, Germany-04315 Leipzig, Konstantinstraße 6, Tel.: +49 (0) 341-681 2811, Fax: +49 (0) 341-681 1837, Web: http://bohmeier-verlag.de, Email: info@magick-pur.de.

Jointly produced with Spencer-Brown publishing, 156 West Common, Horningsham Warminster, BA 12 7 LT Tel/fax +44 1985 844 855.

ISBN 978-3-89094-580-4

Contents

To his best beloved

I think always of you, and I think 'I love you' every day.

Preface to the fifth English edition

As is now well known, *Laws of Form* took ten years from its inception to its publication, four years to write it and six years of political intrigue to get it published.

Typically of all unheralded best sellers from relatively obscure authors, it was turned down by six publishers, including Mark Longman who published my earlier work on probability. Even Sir Stanley Unwin refused to publish it until his best author, Bertrand Russell, told him he must.*

This crucial recommendation was not achieved without intrigue, and required me (not unwillingly) to sleep with one of Russell's granddaughters, who asked me in the morning,

'What exactly do you want from Bertie?'

'To endorse what he said about the book when he first read it in typescript,' I told her.

'He never will!' she exclaimed. 'You'll have to twist his arm, you'll have to blackmail him. How can I help?'

The next few years were spent in vigorous arm-twisting and incessant blackmail from us both. One of her threats was to invite me to Plas Penrhyn as her guest while Bertie and Edith were away in London. This sent Bertie into a paroxysm of terror of what the neighbours might think. He also had an irrational fear of spoiling his reputation as a mathematician, which was not good anyway, by recommending a book that had not yet been tried by the critics. He seemed totally unaware that any book he recommended, however ridiculous, would have no effect whatever on this.**

When we finally got him cornered, in my next visit to Plas Penrhyn, he carefully avoided mentioning the subject during the whole of my stay, and I considered it too dangerous to mention it myself. The next morning I was due to depart while Bertie and Edith were still in bed, and I thought I had failed miserably. But no! I missed my train because they had not ordered me a taxi to the station, which was their way of telling me that my visit was to be prolonged by another day.

* *Laws of Form* was the only work in the entire history of the planet to which Russell gave his unqualified approval. It is the one book he had always wanted to write, and actually tried to, but unfortunately it came out as *Principia*. This was, for him, a major tragedy, and was largely responsible for his spending his latter days in fruitless political protests. He once asked me, 'Do you think I wasted ten years of my life writing Principia?' What could I say? I couldn't say it was good, because he knew it was bad. In the end I said, rather lamely, 'No, Bertie. If you hadn't written the Principia, I couldn't have written the Laws.' More to the point, I wouldn't have because I wouldn't have needed to. Russell knew the Principia was useless because it was wrongly based on logic, and that I had written the Laws to correct it.

** There is no question that Russell regarded me as his successor as the next great English philosopher, and frequently told his grandchildren so. But to risk saying it in public at a time when nobody believed it was another matter.

The evening of this extra day came, and still nothing was mentioned. Ten o' clock bedtime arrived, and I thought I had failed again, when Bertie suddenly said,

'What exactly do you want of me?'

'To endorse what you said about the book three years ago,' I told him.

'You must remind me what it was,' he said.

I produced a verbatim report of his remarks, neatly typed out, and thrust it in his face.

'Are you sure this is all you want?' he said. 'Don't you want me to write a detailed introduction to the work, as I did for Wittgenstein?'

I told him that that would be very nice, but that this was all I needed just now.

He contemplated the page of typescript for a moment, and then a wicked gleam lit up his face, and he rubbed his hands.

'Supposing I don't?' he grinned.

'Then,' I heard myself saying, 'it might delay the publication for a year or so, but the book will still be published in the end, and you won't be associated with it.'

'Oh,' he said. 'I never thought of that. How would you like me to sign it?'

When the book finally came out, in 1969 April 17, its effect was sensational. The Whole Earth Catalog ordered 500 copies, which was half the edition, and other big dealers followed suit. The first printing was sold out before it reached the shops, and the publisher had to order a hurried reprint to meet the demand.

Nobody had seen anything like it. Here was an upstart author explaining the mysteries of mathematics that the so-called greats of the science in the last 8000 years (at least) had never noticed, and in language that a child of six could follow.

Having achieved my life's ambition of composing and publishing a nearly perfect work of literature by the age of 46, I was suddenly confronted by the problem of what to do with the rest of my life. I knew, and so did everybody else, that I could never top this achievement, so with what significant purpose could I carry on?

One thing I could and did do was learn some mathematics. In ten years I had learned enough to become a full professor in the University of Maryland, although I still thought I knew very little. Math is almost impossible to master without personal tuition, and I was lucky to strike up friendships with D H Lehmer and J C P Miller, both, as it happened, experts on Riemann's hypothesis, in which I had no interest whatever, nor in analytic number theory in general. It was only on being told by my former student James Flagg, who is the best-informed scholar of mathematics in the world, that I had in effect proved Riemann's hypothesis in Appendix 7, and again in Appendix 8, that persuaded me to think I had better learn something about it.

I am an intensely competitive person, which comes from being repeatedly told by my mother that I would never be any good. This forced me to spend my whole life attempting to prove her wrong. The tragedy of it is that however brilliantly I performed, it made no difference. Nothing I could do would change her mind. I beat her at chess when I was four, and all she did was refuse to play with me ever again, rather than admit that I was good.

If you solve a famous unsolved problem by mistake it doesn't count. You have to say 'I am going to solve this problem,' and then solve it. So I had to spend another ten years learning analytic number theory, which I hated because much of the literature is mathematically incompetent, e.g. see pp 218 and 219.

The result is so fascinating that it made the effort seem almost worth while, and the problem was so difficult that solving it gave me nearly as much pleasure as writing *Laws of Form*. The world of analysis is completely different from anywhere I had explored, the science of continuous variation rather than discontinuous jumping. And since Riemann's problem is solved by a marriage of the two, although the achievement of a solution cannot quite top what I did in *Laws of Form*, it runs it a close second, if not an equal first.

0100 hrs 23 06 2007 Saturday

Preface to the first American edition

Apart from the standard university logic problems, that the calculus published in this text renders so easy that we need not trouble ourselves further with them, perhaps the most significant thing, from the mathematical angle, that it enables us to do is to use complex values in the algebra of logic. They are the analogs, in ordinary algebra, to complex numbers $a + b\sqrt{-1}$. My brother and I had been using their Boolean counterparts in practical engineering for several years before realizing what they were. Of course, being what they are, they work perfectly well, but understandably we felt a bit guilty about using them, just as the first mathematicians to use 'square roots of negative numbers' had felt guilty, because they too could see no plausible way of giving them a respectable academic meaning. All the same, we were quite sure there was a perfectly good theory that would support them, if only we could think of it.

The position is simply this. In ordinary algebra, complex values are accepted as a matter of course, and the more advanced techniques would be impossible without them. In Boolean algebra (and thus, for example, in all our reasoning processes) we disallow them. Whitehead and Russell introduced a special rule, that they called the Theory of Types, expressly to do so. Mistakenly, as it now turns out. So, in this field, the more advanced techniques, although not impossible, simply don't yet exist. At the present moment we are constrained, in our reasoning processes, to do it the way it was done in Aristotle's day. The poet Blake might have had some insight into this, for in 1788 he wrote that 'reason, or the ratio of all we have already known, is not the same that it shall be when we know more.'

Recalling Russell's connexion with the Theory of Types, it was with some trepidation that I approached him in 1967 with the proof that it was unnecessary. To my relief he was delighted. The Theory was, he said, the most arbitrary thing he and Whitehead had ever had to do, not really a theory but a stopgap, and he was glad to have lived long enough to see the matter resolved.

Put as simply as I can make it, the resolution is as follows. All we have to show is that the self-referential paradoxes, discarded with the Theory of Types, are no worse than similar self-referential paradoxes, that are considered quite acceptable, in the ordinary theory of equations.

The most famous such paradox in logic is in the statement, 'This statement is false.'

Suppose we assume that a statement falls into one of three categories, true, false, or meaningless, and that a meaningful statement that is not true must be false, and one that is not false must be true. The statement under consideration does not appear to be meaningless (some philosophers have claimed that it is, but it is easy to refute this), so

it must be true or false. If it is true, it must be, as it says, false. But if it is false, since this is what it says, it must be true.

It has not hitherto been noticed that we have an equally vicious paradox in ordinary equation theory, because we have carefully guarded ourselves against expressing it this way. Let us now do so.

We will make assumptions analogous to those above. We assume that a number can be either positive, negative, or zero. We assume further that a nonzero number that is not positive must be negative, and one that is not negative must be positive. We now consider the equation

$$x^2 + 1 = 0.$$

Transposing, we have

$$x^2 = -1$$

and dividing both sides by x gives

$$x = \frac{-1}{x}.$$

We can see that this (like the analogous statement in logic) is self-referential: the root-value of x that we seek must be put back into the expression from which we seek it.

Mere inspection shows us that x must be a form of unity, or the equation would not balance numerically. We have assumed only two forms of unity, +1 and –1, so we may now try them each in turn. Set $x = +1$.
This gives

$$+1 = \frac{-1}{+1} = -1$$

which is clearly paradoxical. So set $x = -1$. This time we have

$$-1 = \frac{-1}{-1} = +1$$

and it is equally paradoxical.

Of course, as everybody knows, the paradox in this case is resolved by introducing a fourth class of number, called *imaginary*, so that we can say the roots of the equation above are $\pm\, i$, where i is a new kind of unity that consists of a square root of minus one.

What we do in Chapter 11 is extend the concept to Boolean algebras, which means that a valid argument may contain not just three classes of statement, but four: true, false, meaningless, and imaginary. The implications of this, in the fields of logic, philosophy, mathematics, and even physics, are profound.

What is fascinating about the imaginary Boolean values, once we admit them, is the light they apparently shed on our concepts of matter and time. It is, I guess, in the nature of us all to wonder why the universe appears just the way it does. Why, for example, does it not appear more symmetrical? Well, if you will be kind enough, and patient enough, to bear with me through the argument as it develops itself in this text, you will I think see, even though we begin it as symmetrically as we know how, that it becomes, of its own accord, less and less so as we proceed.

G Spencer-Brown
Cambridge, England
Maundy Thursday 1972

Preface

The exploration on which this work rests was begun towards the end of 1959. The subsequent record of it owes much, in its early stages, to the friendship and encouragement of Lord Russell, who was one of the few men at the beginning who could see a value in what I proposed to do. It owes equally, at a later stage, to the generous help of Dr J C P Miller, Fellow of University College and Lecturer in Mathematics in the University of Cambridge, who not only read the successive sets of printer's proofs, but also acted as an ever-available mentor and guide, and made many suggestions to improve the style and accuracy of both text and context.

In 1963 I accepted an invitation of Mr H G Frost, Staff Lecturer in Physical Sciences in the Department of Extra-mural Studies in the University of London, to give a course of lectures on the mathematics of logic. The course was later extended and repeated annually at the Institute of Computer Science in Gordon Square, and from it sprang some of the context in the notes and appendices to this essay. I was also enabled, through the help of successive classes of pupils, to extend and sharpen the text.

The publishers (including their readers and their technical artist) were particularly cooperative, as were the printers, and, before this, Mrs Peter Bragg undertook the exacting task of preparing a typescript. Finally I should mention that an original impetus to the work came from Mr I V Idelson, General Manager of Simon-MEL Distribution Engineering, the techniques here recorded being first developed not in respect of questions of logic, but in response to certain unsolved problems in engineering.

Richmond, August 1968

Acknowledgment

The author and publisher acknowledge the kind permission of Mr J Lust of the University of London School of Oriental and African Studies, to photograph part of a facsimile copy of the 12th Century Fujian print of the Dao De Jing in the old Palace Museum, Beijing.

Introduction

A principal intention of this essay is to separate what are known as algebras of logic from the subject of logic, and to re-align them with mathematics.

Such algebras, commonly called Boolean, appear mysterious because accounts of their properties at present reveal nothing of any mathematical interest about their arithmetics. Every algebra has an arithmetic, but Boole designed[1] his algebra to fit logic, which is a possible interpretation of it, and certainly not its arithmetic. Later authors have, in this respect, copied Boole, with the result that nobody hitherto appears to have made any sustained attempt to elucidate and to study the primary, non-numerical arithmetic of the algebra in everyday use that now bears Boole's name.

When I began, some seven years ago, to see that such a study was needed, I found myself upon what was, mathematically speaking, untrodden ground. I had to explore it inwards to discover the missing principles. They are of great depth and beauty, as we shall presently see.

In recording this account of them, I have aimed to write so that every special term shall be either defined or made clear by its context. I have assumed on the part of the reader no more than a knowledge of the English language, of counting, and of how numbers are commonly represented. I have allowed myself the liberty of writing somewhat more technically in this introduction and in the notes and appendices that follow the text, but even here, since the subject is of such general interest, I have endeavoured, where possible, to keep the account within the reach of a non-specialist.

Accounts of Boolean algebras have up to now been based on sets of postulates. We may take a postulate to be a statement that is accepted without evidence, because it belongs to a set of such statements from which it is possible to derive other statements that it happens to be convenient to assume. The chief characteristic that has always marked such statements has been an almost total lack of any spontaneous appearance of truth[2]. Nobody pretends, for example, that Sheffer's equations[3] are mathematically evident, for their evidence is not apparent apart from the usefulness of equations that follow from them. But in the primary arithmetic developed in this essay, the initial equations can be seen to represent two very simple laws of indication that, whatever our views on the nature of their self-evidence, at least recommend themselves to the findings of common sense. I am thus able to present (Appendix 1), apparently for the first time, proofs of

1 George Boole, *The mathematical analysis of logic*, Cambridge, 1847
2 Cf Alfred North Whitehead and Bertrand Russell, *Principia mathematica,* Vol. I, 2nd edition, Cambridge, 1927, p v.
3 Henry Maurice Sheffer, *Trans. Amer. Math. Soc.,* 14 (1913) 481-8.

each of Sheffer's postulates, and hence of all Boolean postulates, as theorems about an axiomatic system that is seen to rest on the fundamental ground of mathematics.

Working outwards from this fundamental source, the general form of mathematical communication, as we understand it today, tends to grow quite naturally under the hand that writes it. We have a definite system, we name its parts, and we adopt, in many cases, a single sign to represent each name. In doing this, forms of expression are called inevitably out of the need for them, and the proofs of theorems, that are at first seen to be little more than a relatively informal direction of attention to the complete range of possibilities, become more and more formal as we proceed from our original conception. At the half-way point the algebra, in all its representative completeness, is found to have grown imperceptibly out of the arithmetic, so that by the time we have started to work in it we are already fully acquainted with its formalities and possibilities without anywhere having set out with the intention of describing them as such.

One of the merits of this form of presentation is the gradual building up of mathematical notions and common forms of procedure without any apparent break from common sense. The discipline of mathematics is seen to be a way, powerful in comparison with others, of revealing our internal knowledge of the structure of the world, and only by the way associated with our common ability to reason and compute.

Even so, the orderly development of mathematical conventions and formulations stage by stage has not been without its problems on the reverse side. A person with mathematical training, who may automatically use a whole range of techniques without questioning their origin, can find himself in difficulties over an early part of the presentation, in which it has been necessary to develop an idea using only such techniques as have already been identified. In some of these cases we need to derive a concept for which the procedures already developed are only just adequate. The argument, which is maximally elegant at such a point, may thus be conceptually difficult to follow.

One such case, occurring in Chapter 2, is the derivation of the second of the two primitive equations of the calculus of indications. There seems to be such universal difficulty in following the argument at this point, that I have restated it less elegantly in the notes on this chapter at the end of the text. When this is done, the argument is seen to be so simple as to be almost trivial. But it must be remembered that, according to the rigorous procedure of the text, no principle may be used until it has been either called into being or justified in terms of other principles already adopted. In this particular instance, we make the argument easy by using ordinary substitution. But at the stage in the essay where it becomes necessary to formulate the second primitive equation, no principle of substitution has yet been called into being, since its use and justification, that we find later in the essay, depends in part upon the existence of the very equation we want to establish.

In Appendix 2, I give a brief account of some of the simplifications that can be made through using the primary algebra as an algebra of logic. For example, there are no primitive propositions. This is because we have a basic freedom, not granted to other algebras of logic, of access to the arithmetic whenever we please. Thus each of Whitehead and Russell's five primitive implications [2,pp 96-7] can be equated with a single constant. The constant, if it were a proposition, would be *the* primitive implication. But in fact, being arithmetical, it cannot represent a proposition.

A point of interest in this connexion is the development of the idea of a variable solely from that of the operative constant. This comes from the fact that the algebra represents our ability to consider the form of an arithmetical equation irrespective of the appearance, or otherwise, of this constant in certain specified places. And since, in the primary arithmetic, we are not presented, apparently, with two kinds of constant, such as 5, 6, etc and +, x, etc, but with expressions made up, apparently, of similar constants each with a single property, the conception of a variable comes from considering the irrelevant presence or absence of this property. This lends support to the view, suggested[4] by Wittgenstein, that variables in the calculus of propositions do not in fact represent the propositions in an expression, but only the truth-values of these propositions, since the propositions themselves cannot be equated with the mere presence or absence of a given property, while the possibility of their being true or not true can.

Another point of interest is the clear distinction, with the primary algebra and its arithmetic, that can be drawn between the proof of a theorem and the demonstration of a consequence. The concepts of theorem and consequence, and hence of proof and demonstration, are widely confused in current literature, where the words are used interchangeably. This has undoubtedly created spurious difficulties. As will be seen in the statement of the completeness of the primary algebra (theorem 17), what is to be proved becomes strikingly clear when the distinction is properly maintained. (A similar confusion is apparent, especially in the literature of so-called symbolic logic, of the concepts of axiom and postulate.)

It is possible to develop the primary algebra to such an extent that it can be used as a restricted (or even as a full) algebra of numbers. There are several ways of doing this, the most convenient of which is to limit condensation in the arithmetic, and thus to use a number of crosses in a given space to represent either the corresponding number or its image. When this is done we can see plainly some at least of the evidence for Gödel's

4 Ludwig Wittgenstein, *Tractatus logico-philosophicus,* London, 1922, propositions 5 sq.

and Church's theorems[5,6] of decision. But with the rehabilitation of the paradoxical equations undertaken in Chapter 11, the meaning and application of these theorems now stands in need of review. They certainly appear less destructive than was hitherto supposed.

I aimed in the text to carry the development only so far as to be able to consider reasonably fully all the forms that emerge at each stage. Although I indicate the expansion into complex forms in Chapter 11, I otherwise try to limit the development so as to render the account, as far as it goes, complete.

Most of the theorems are original and their proofs therefore new. But some of the later algebraic and mixed theorems, occurring in what is at this stage familiar ground, are already known and have, in other forms, been proved before. In all of these cases I have been able to find what seem to be clearer, simpler, or more direct proofs, and in most cases the theorems I prove are more general. A case in point is my theorem 16. It was only after contemplating this theorem for some two years that I found the beautiful key by which it is seen to be true for all possible algebras, Boolean or otherwise.

In arriving at proofs, I have often been struck by the apparent alignment of mathematics with psycho-analytic theory. In each discipline we attempt to find out, by a mixture of contemplation, changes in presentation, communion, and communication, what it is we already know. In mathematics, as in other forms of self-analysis, we do not have to go exploring the physical world to find what we are looking for. Any child of ten, who can multiply and divide, already knows, for example, that the sequence of prime numbers is endless. But if he is not shown Euclid's proof, it is unlikely that he will ever find out, before he dies, that he knows.

This analogy suggests that we have a direct awareness of mathematical form as an archetypal structure. I try in the final chapter to illustrate the nature of this awareness. In any case, questions of pure probability alone would lead us to suppose that some degree of direct awareness is present throughout mathematics.

We may take it that the number of statements that might or might not be provable is unlimited, and it is evident that, in any large enough finite sample, untrue statements, of those bearing any useful degree of significance, heavily outnumber true statements. Thus in principle, if there were no innate sense of rightness, a mathematician would attempt to prove more false statements than true ones. But in practice he seldom attempts to prove any statement unless he is already convinced of its truth. And since he has not yet proved it, his conviction must arise, in the first place, from considerations other than proof.

5 Kurt Gödel, *Monatshefte für Mathematik und Physik*, 38 (1931) 172-98.
6 Alonzo Church, *J. symbolic Logic*, 1 (1936) 40-1, 101-2.

Thus the codification of a proof procedure, or of any other directive process, although at first useful, can later stand as a threat to further progress. For example, we may consider the largely unconscious, but now codified, limitation of the reasoning (as distinct from the computative) parts of proof structures to the solution of Boolean equations of the first degree. As we see in Chapter 11, and in the notes thereto, the solution of equations of higher degree is not only possible, but has been undertaken by switching engineers on an ad hoc basis for some half a century or more. Such equations have hitherto been excluded from the subject matter of ordinary logic by the Whitehead-Russell theory of types [2, pp 37 sq, e.g. p 77].

I show in the text that we can construct an implicit function of itself so that it re-enters its own space at either an odd or an even depth. In the former case we find the possibility of a self-denying equation of the kind these authors describe. In such a case, the roots of the equation so set up are imaginary. But in the latter case we find a self-confirming equation that is satisfied, for some given configuration of the variables, by two real roots.

I am able, by this consideration, to rehabilitate[7] the formal structure hitherto discarded with the theory of types. As we now see, the structure can be identified in the more general theory of equations, behind which there already exists a weight of mathematical experience.

One prospect of such a rehabilitation, that could repay further attention, comes from the fact that, although Boolean equations of the first degree can be fully represented on a plane surface, those of the second degree cannot be so represented. D J Spencer-Brown and I found evidence, in unpublished work undertaken in 1962 through 1965, suggesting that both the four-colour theorem and Goldbach's theorem are undecidable with a proof structure confined to Boolean equations of the first degree, but decidable if we are prepared to avail ourselves of equations of higher degree.

One of the motives prompting the furtherance of the present work was the hope of bringing together the investigations of the inner structure of our knowledge of the universe, as expressed in the mathematical sciences, and the investigations of its outer structure, as expressed in the physical sciences. Here the work of Einstein, Schrödinger, and others seems to have led to the realization of an ultimate boundary of physical knowledge in the form of the media through which we perceive it. It becomes apparent that if certain facts about our common experience of perception, or what we might call the inside world, can be revealed by an extended study of what we call, in contrast, the outside world, then an equally extended study of this inside world will reveal, in turn, the

7 For a history of the earlier essays to rehabilitate, on a logical rather than a mathematical basis, something of what was discarded, see Abraham A Fraenkel and Yehoshua Bar-Hillel, *Foundations of set theory*, Amsterdam, 1958, pp 136-95

facts first met with in the world outside: for what we approach, in either case, from one side or the other, is the common boundary between them.

I do not pretend to have carried these revelations very far, or that others, better equipped, could not carry them further. I hope they will. My conscious intention in writing this essay was the elucidation of an indicative calculus, and its latent potential, becoming manifest only when the realization of this intention was already well advanced, took me by surprise.

I break off the account at the point where, as we enter the third dimension of representation with equations of degree higher than unity, the connexion with the basic ideas of the physical world begins to come more strongly into view. I had intended, before I began writing, to leave it here, since the latent forms that emerge at this, the fourth departure from the primary form (or the fifth departure, if we count from the void) are so many and so varied that I could not hope to present them all, even cursorily, in one book.

Medawar observes[8] that the standard form of presentation required of an ordinary scientific paper represents the very reverse of what the investigator was in fact doing. In reality, says Medawar, the hypothesis is first posited, and becomes the medium through which certain otherwise obscure facts, later to be collected in support of it, are first clearly seen. But the account in the paper is expected to give the impression that such facts first suggested the hypothesis, irrespective of whether this impression is truly representative.

In mathematics we see this process in reverse. The mathematician, more frequently than he is generally allowed to admit, proceeds by experiment, inventing and trying out hypotheses to see if they fit the facts of reasoning and computation with which he is presented. When he has found a hypothesis that fits, he is expected to publish an account of the work in the reverse order, so as to deduce the facts from the hypothesis.

I would not recommend that we should do otherwise, in either case. By all accounts, to tell the story backwards is convenient and saves time. But to pretend that the story was actually lived backwards can be extremely mystifying.

In view of this apparent reversal, Laing suggests[9] that what in empirical science are called *data*, being in a real sense *arbitrarily* chosen by the nature of the hypothesis already formed, could more honestly be called *capta*. By reverse analogy, the facts of mathematical science, appearing at first to be arbitrarily chosen, and thus *capta*, are not really arbitrary at all, but absolutely determined by the nature and coherence of our being. In this view we might consider the facts of mathematics to be the real *data* of

8 P B Medawar, Is the Scientific Paper a Fraud, *The Listener*, 12. September 1963, pp 377-8.
9 R D Laing, *The politics of experience and the bird of paradise*, London, 1967, pp 52 ff.

experience, for only these appear to be, in the final analysis, inescapable.

Although I have undertaken, to the best of my ability, to preserve, in the text itself, what is thus inescapable, and thereby timeless, and otherwise to discard what is temporal, I am under no illusion of having entirely succeeded on either count. That one can *not*, in such an undertaking, succeed perfectly, seems to me to reside in the manifest imperfection of the state of *particular existence*, in any form at all. (Cf Appendix 2.) The work of any human author must be to some extent idiosyncratic, even though he may know his personal ego to be but a fashionable garb to suit the mode of the present rather than the mean of past and future in which his work will come to rest.

A major aspect of the language of mathematics is the degree of its formality. Although it is true that we are concerned, in mathematics, to provide a shorthand for what is actually said, this is only half the story. What we aim to do, in addition, is to provide a more general form in which the ordinary language of experiences is seen to rest. As long as we confine ourselves to the subject at hand, without extending our consideration to what it has in common with other subjects, we are not availing ourselves of a truly mathematical mode of presentation.

What is encompassed, in mathematics, is a transcendence from a given state of vision to a new, and hitherto unapparent, vision beyond it. When the present existence has ceased to make sense, it can still come to sense again through the realization of its form.

Thus the subject matter of logic, however symbolically treated, is not, in as far as it confines itself to the ground of logic, a mathematical study. It becomes so only when we are able to perceive its ground as a part of a more general form, in a process without end. Its mathematical treatment is a treatment of the form in which our way of talking about our ordinary living experience can be seen to be cradled. It is the laws of this form, rather than those of logic, that I have attempted to record.

In making the attempt, I found it easier to acquire an access to the laws themselves than to determine a satisfactory way of communicating them. In general, the more universal the law, the more it seems to resist expression in any particular mode.

Some of the difficulties apparent in reading, as well as in writing, the earlier part of the text come from the fact that, from Chapter 5 backwards, we are extending the analysis through and beyond the point of simplicity where language ceases to act normally as a currency for communication. The point at which this break from normal usage occurs is in fact the point where algebras are ordinarily taken to begin. To extend them back beyond this point demands a considerable unlearning of the current descriptive superstructure that, until it is unlearned, can be mistaken for the reality.

The fact that, in a book, we have to use words and other signs in an attempt to express what the use of words and other signs has hitherto obscured, tends to make demands of an extraordinary nature on both writer and reader, and I am conscious, on my side, of

how imperfectly I succeed in rising to them. But at least, in the process of undertaking the task, I have become aware (as Boole himself became aware) that what I am saying has nothing to do with me, or anyone else, at the personal level. It, as it were, records itself and, whatever the faults in the record, *that which* is so recorded is not a matter of opinion. The only credit I feel entitled to accept in respect of it is for the instrumental labour of making a record that may be articulate and coherent enough to be understood in its temporal context.

London, August 1967

A note on the mathematical approach

The theme of this book is that a universe comes into being when a space is severed or taken apart. The skin of a living organism cuts off an outside from an inside. So does the circumference of a circle in a plane. By tracing the way we represent such a severance, we can begin to reconstruct, with an accuracy and coverage that appear almost uncanny, the basic forms underlying linguistic, mathematical, physical, and biological science, and can begin to see how the familiar laws of our own experience follow inexorably from the original act of severance. The act is itself already remembered, even if unconsciously, as our first attempt to distinguish different things in a world where, in the first place, the boundaries can be drawn anywhere we please. At this stage the universe cannot be distinguished from how we act upon it, and the world may seem like shifting sand beneath our feet.

Although all forms, and thus all universes, are possible, and any particular form is mutable, it becomes evident that the laws relating such forms are the same in any universe. It is this sameness, the idea that we can find a reality independent of how the universe actually appears, that lends such fascination to the study of mathematics. That mathematics, in common with other art forms, can lead us beyond ordinary existence, and can show us something of the structure in which all creation hangs together, is no new idea. But mathematical texts generally begin the story somewhere in the middle, leaving the reader to pick up the thread as best he can. Here the story is traced from the beginning.

Unlike more superficial forms of expertise, mathematics is a way of saying less and less about more and more. A mathematical text is thus not an end in itself, but a key to a world beyond the compass of ordinary description.

An initial exploration of such a world is usually undertaken in the company of an experienced guide. To undertake it alone, although possible, is perhaps as difficult as to enter the world of music by attempting, without personal guidance, to read the score-sheets of a master composer, or to set out on a first solo flight in an aeroplane with no other preparation than a study of the pilots' manual.

Although the notes at the end of the text may to some extent make up for, they cannot effectively replace, such personal guidance. They are designed to be read in conjunction with the text, and it may in fact be helpful to read them first.

The reader who is already familiar with logic, in either its traditional or its "symbolic" form, may do well to begin with Appendix 2, referring through the Index of Forms to the text whenever necessary.

Mathematics is the science of what we *know* of what we have defined. It has no place for opinions or beliefs of any kind.

無名天地之始

1

The form

We take as given the idea of distinction and the idea of indication, and that we cannot make an indication without drawing a distinction. We take, therefore, the form of distinction for the form.

Definition

*Distinction is perfect continence.**

That is to say, a distinction is drawn by arranging a boundary with separate sides so that a point on one side cannot reach the other side without crossing the boundary. For example, in a plane space a circle draws a distinction.

Once a distinction is drawn, the spaces, states, or contents on each side of the boundary, being distinct, can be indicated.

There can be no distinction without motive, and there can be no motive unless contents are seen to differ in value.

If a content is of value, a name can be taken to indicate this value.

Thus the calling of the name can be identified with the value of the content.

* There have been so many jokes about this that I had better point out that I was employing "continence" in its original sense of "containment", not in its later usage as "sexual abstinence" – Author, reviewing this text 2000-06-27

Axiom 1. The law of calling

The value of a call made again is the value of the call.

That is to say, if a name is called and then is called again, the value indicated by the two calls taken together is the value indicated by one of them.

That is to say, for any name, to recall is to call.

Equally, if the content is of value, a motive or an intention or instruction to cross the boundary into the content can be taken to indicate this value.

Thus, also, the crossing of the boundary can be identified with the value of the content.

Axiom 2. The law of crossing

The value of a crossing made again is not the value of the crossing.

That is to say, if it is intended to cross a boundary and then it is intended to cross it again, the value indicated by the two intentions taken together is the value indicated by none of them.

That is to say, for any boundary, to recross is not to cross.

2

Forms taken out of the form

Construction

Draw a distinction.

Content

Call it the first distinction

Call the space in which it is drawn the space severed or cloven by the distinction.

Call the parts of the space shaped by the severance or cleft the sides of the distinction or, alternatively, the spaces, states, or contents distinguished by the distinction.

Intent

Let any mark, token, or sign be taken in any way with or with regard to the distinction as a signal.
Call the use of any signal its intent.

First canon. Convention of intention

Let the intent of a signal be limited to the use allowed to it.

Call this the convention of intention. In general, *what is not allowed is forbidden.*

Knowledge

Let a state distinguished by the distinction be marked with a mark

$$\urcorner$$

of distinction.

Let the state be known by the mark

Call the state the marked state.

Form

Call the space cloven by any distinction, together with the entire content of the space, the form of the distinction.

Call the form of the first distinction the form.

Name

Let there be a form distinct from the form.

Let the mark of distinction be copied out of the form into such another form.

Call any such copy of the mark a token of the mark.

Let any token of the mark be called as a name of the marked state.

Let the name indicate the state.

Arrangement

Call the form of a number of tokens considered with regard to one another (that is to say, considered in the same form) an arrangement.

Expression

Call any arrangement intended as an indicator an expression.

Value

Call a state indicated by an expression the value of the expression.

Equivalence

Call expressions of the same value equivalent.

Let a sign

$$=$$

of equivalence be written between equivalent expressions.

Now, by axiom 1,

$$\overline{}| \ \overline{}| = \overline{}| .$$

Call this the form of condensation.

Instruction

Call the state not marked with the mark the unmarked state.

Let each token of the mark be seen to cleave the space into which it is copied. That is to say, let each token be a distinction in its own form.

Call the concave side of a token its inside.

Let any token be intended as an instruction to cross the boundary of the first distinction.

Let the crossing be from the state indicated on the inside of the token.

Let the crossing be to the state indicated by the token.

Let a space with no token indicate the unmarked state.

Now, by axiom 2

$$\overline{\overline{}|}| = \quad .$$

Call this the form of cancellation.

Equation

Call an indication of equivalent expressions an equation.

Primitive equation

Call the form of condensation a primitive equation.
Call the form of cancellation a primitive equation.
Let there be no other primitive equation.

Simple expression

Note that the three forms of arrangement, $\overline{}| \ \overline{}|$, $\overline{\overline{}|}|$, $\overline{}|$ and the one absence of form, $\quad$, taken from the primitive equations are all, by convention, expressions.

Call any expression consisting of an empty token simple.
Call any expression consisting of an empty space simple.

Let there be no other simple expression

Operation

We now see that if a state can be indicated by using a token as a name it can be indicated by using the token as an instruction subject to convention. Any token may be taken, therefore, to be an instruction for the operation of an intention, and may itself be given a name

cross

to indicate what the intention is.

Relation

Having decided that the form of every token called cross is to be perfectly continent, we have allowed only one kind of relation between crosses: continence.

Let the intent of this relation be restricted so that a cross is said to contain what is on its inside and not to contain what is not on its inside.

Depth

In an arrangement a standing in a space s, call the number n of crosses that must be crossed to reach a space s_n from s the depth of s_n with regard to s.

Call a space reached by the greatest number of inwards crossings from s a deepest space in a.

Call the space reached by no crossing from s the shallowest space in a.

Thus

$$s_0 = s.$$

Let any cross standing in any space in a cross c be said to be contained in c.

Let any cross standing in the shallowest space in c be said to stand under, or to be covered by, c.

Unwritten cross

Suppose any s_0 to be surrounded by an unwritten cross.

Call the crosses standing under any cross c, written or unwritten, the crosses pervaded by the shallowest space in c.

Pervasive space

Let any given space s_n be said to pervade any arrangement in which s_n is the shallowest space.

Call the space s pervading an arrangement a, whether or not a is the only arrangement pervaded by s, the pervasive space of a.

3

The conception of calculation

Second canon. Contraction of reference

1 Construct a cross.

2 Mark it with c.

3 Let c be its name.

4 Let the name indicate the cross.

Let the four injunctions (two of constructive intent, two of conventional intent) above be contracted to the one injunction (of mixed intent) below.

1 Take any cross c.

In general, *let injunctions be contracted to any degree in which they can still be followed.*

Third canon. Convention of substitution

In any expression, let any arrangement be changed for an equivalent arrangement.

Step

Call any such change a step.

Let a sign

$$\rightharpoonup$$

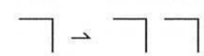
stand for the words

is changed to.

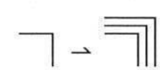
Let a barb in the sign indicate the direction of the change.

Direction

A step may now be considered not only with regard to its kind, as in



rather than



but also with regard to its direction, as in

$$\urcorner \rightharpoonup \urcorner\;\urcorner,$$

rather than

$$\urcorner \leftharpoondown \urcorner\;\urcorner.$$

Fourth canon. Hypothesis of simplification

Suppose the value of an arrangement to be the value of a simple expression to which, by taking steps, it can be changed.

Example. To find a value of the arrangement

$$\overline{\urcorner\;\urcorner}\!|\;\urcorner$$

take simplifying steps

$$\overline{\urcorner\;\urcorner}\!|\;\urcorner \rightharpoonup \overline{\urcorner}\!|\;\urcorner \qquad \text{condensation}$$

$$\overline{\urcorner}\!|\;\urcorner \rightharpoonup \urcorner \qquad \text{cancellation}$$

to change it for a simple expression. Now, by the hypothesis of simplification, its value is supposed to be the marked state.

Thus a value for any arrangement can be supposed if the arrangement can be simplified. But it is plain that some arrangements can be simplified in more than one way, and it is conceivable that others might not simplify at all. To show, therefore, that the hypothesis of simplification is useful determinant of value we shall need to show, at some stage, that any given arrangement will simplify and that all possible procedures for simplifying it will lead to an identical simple expression.

Fifth canon. Expansion of reference

The names hitherto used for the primitive equations suggest steps in the direction of simplicity, and so are not wholly suitable for steps that may in fact be taken in either direction. We therefore expand the form of reference.

condensation	$\urcorner\;\urcorner \rightharpoonup \urcorner$	number
confirmation	$\urcorner \rightharpoonup \urcorner\;\urcorner$	

cancellation	$\overline{\urcorner}\!| \rightharpoonup$	order
compensation	$\rightharpoonup \overline{\urcorner}\!|$	

In general, a contraction of reference accompanies an expansion of awareness, and an expansion of reference accompanies a contraction of awareness. If what was done through awareness is to be done by rule, forms of reference must grow (that is to say, divide) to accommodate rules.

Like contraction of reference, of which it is an image, expansion of reference happens, originally, of its own accord. It might at first seem to be a strange procedure, therefore, to call into being a rule permitting it. But we see, if we consider it, that we must call a rule for any process that happens of its own accord, in order to save the convention of intention.

Thus, in general, *let any form of reference be divisible without limit.*

Calculation

Call calculation a procedure by which, as a consequence of steps, a form is changed for another, and call a system of constructions and conventions that allows calculation a calculus.

Initial

The forms of step allowed in a calculus can be defined as all the forms that can be seen in a given set of equations. Call the equations so used to determine these forms the initial equations, or initials, of the calculus.

The calculus of indications

Call the calculus determined by taking the two primitive euations

$$\urcorner\;\urcorner = \urcorner \qquad \text{number}$$

$$\overline{\urcorner}\!| = \qquad \text{order}$$

as initials the calculus of indications.

Call the calculus limited to the forms generated from direct consequences of these initials the primary arithmetic.

4

The primary arithmetic

Initial 1. Number

$$\overline{}|\ \overline{}| = \overline{}| \qquad \begin{array}{c}\text{condense}\\ \rightleftharpoons\\ \text{confirm}\end{array}$$

Initial 2. Order

$$\overline{\overline{}|\,}| = \qquad \begin{array}{c}\text{cancel}\\ \rightleftharpoons\\ \text{compensate}\end{array}$$

We shall proceed to distinguish general statements, called theorems, that can be seen from formal considerations of these initials.

Theorem 1. Form

The form of any finite cardinal number of crosses can be taken as the form of an expression.

That is to say, any conceivable arrangement of any integral number of crosses can be constructed from a simple expression by the initial steps of the calculus.

We may prove this theorem by finding a procedure for simplification: since what can be reduced to a simple expression can, by retracing the steps, be constructed from it.

Proof

Take any such arrangement a in a space s.

Procedure. Find any deepest space in a. It can be found with a finite search since in any given a the number of crosses, and thereby the number of spaces, is finite.

Call the space s_d.

Now s_d is either contained in a cross or not contained in a cross.

If s_d is not contained in a cross, then s_d is s, and there is no cross in s, and so a is already simple.

If s_d is in a cross c_d, then c_d is empty, since if c_d were not empty s_d would not be deepest.

Now c_d either stands alone in s or does not stand alone in s.

If c_d stands alone in s, then a is already simple.

If c_d does not stand alone in s, then c_d must stand either

(*case 1*) in a space together with another empty cross (if the other cross were not emty s_d would not be deepest) or

(*case 2*) alone in the space under another cross.

Case 1. In this case c_d condenses with the other empty cross. Thereby, one cross is eliminated from a.

Case 2. In this case c_d cancels with the other cross. Thereby, two crosses are eliminated from a.

Now, since each repetition of the procedure used in *case 1* or *case 2* (that is to say, the procedure for an arrangement that is not simple) results in a new arrangement with one or two fewer crosses, there will come a time when, after a finite number of repetitions, a has been either reduced to one cross or eliminated completely.

Thus, in any case, a is simplified.

Therefore, the form of any finite cardinal number of crosses can be taken as the form of an expression.

Theorem 2. Content

If any space pervades an empty cross, the value indicated in the space is the marked state.

Proof

Consider an expression consisting of a part p in a space with an empty cross c_e. It is required to prove that in any case

$$pc_e = c_e.$$

Procedure. Simplify p.

If the procedure reduces p to an empty cross, then the empty cross condenses with c_e, and only c_e remains.

If the procedure eliminates p, then only c_e remains.

Thereby, the simplification of every form of pc_e is c_e.

But c_e indicates the marked state.

Therefore, if any space pervades an empty cross, the value indicated in the space is the marked state.

Theorem 3. Agreement

The simplification of an expression is unique.

That is to say, if an expression e simplifies to a simple expression e_s *then* e cannot simplify to a simple expression other than e_s.

In simplifying an expression, we may have a choice of steps. Thus the act of simplification cannot be a unique determinant of value unless we can find in it a form independent of this choice.

Now it is clear that, for some expressions, the hypothesis of simplification does provide a unique determinant of value, and we shall proceed to use this fact to show that it provides such a determinant for all expressions.

Let m stand for any number, greater than zero, of such expressions indicating the marked state.

Let n stand for any number of such expressions indicating the unmarked state.

By axiom 1

$$mm = m$$

and

$$nn = n$$

and by simplification or the use of theorem 2,

$$mn = m.$$

Call the value of m a dominant value, and call the value of n a recessive value.

These definitions and considerations may now be summarized in the following rule.

Sixth canon. Rule of dominance

If an expression e in a space s shows a dominant value in s, then the value of e is the marked state. Otherwise, the value of e is the unmarked state.

Also, by definition,

(i) $$m = \overline{\quad}|$$

and

(ii) $$n = \quad$$

so that

$$\overline{m\,}| = n \qquad \text{(i) cancellation (ii)}$$

and

$$\overline{n\,}| = m \qquad \text{(i),(ii).}$$

Proof of theorem 3

Let e stand in the space s_0.

Procedure. Count the number of crossings from s_0 to the deepest space in e. If the number is d, call the deepest space s_d.

By definition, the crosses covering s_d are empty, and they are the only contents of s_{d-1}.

Being empty, each cross in s_{d-1} can be seen to indicate only the marked state, and so the hypothesis of simplification uniquely determines its value.

1 Make a mark m on the outside of each cross in s_{d-1}.

We know by (i), that

$$m = \overline{\quad}|.$$

Thus no vaue in s_{d-1} is changed, since

$$\overline{\quad}| \rightharpoonup \overline{\quad}|\, m \qquad \text{procedure}$$
$$= \overline{\quad}|\,\overline{\quad}| \qquad \text{(i)}$$
$$= \overline{\quad}| \qquad \text{condensation.}$$

Therefore the value of e is unchanged.

2 Next consider the crosses in s_{d-2}.

Each cross in s_{d-2} either is empty or covers one or more crosses already marked with m.

If it is empty, mark it with m so that the considerations in *1* apply.

If it covers a mark m, mark it with n.

We know by (ii) that

$$n = \quad .$$

Thus no value in s_{d-2} is changed.

Therefore the value of e is unchanged.

3 Consider the crosses in s_{d-3}.

Any cross in s_{d-3} either is empty or covers one or more crosses already marked with m or n.

If it does not cover a mark m, mark it with m.

If it covers a mark m, mark it with n.

In either case, by the considerations in *1* and *2*, no value in s_{d-3} is changed, and so the value of e is unchanged.

The procedure in subsequent spaces to s_0 requires no additional consideration.

Thus, by the procedure, each cross in e is uniquely marked with m or n.

Therefore, by he rule of dominance, a unique value of e in s_0 is determined.

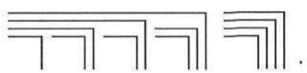
But the procedure leaves the value of e unchanged, and the rules of the procedure are taken from the rules of simplification.

Therefore, the value of e determined by the procedure is the same as the value of e determined by simplification.

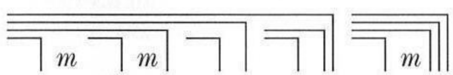
But e can be any expression.

Therefore, the simplification of an expression is unique.

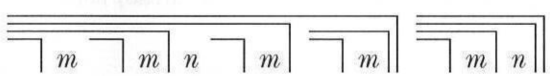
Illustration. Let e be



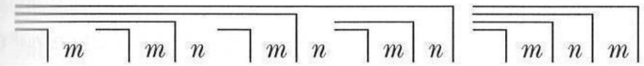
The deepest space in e is s_4, so mark crosses first in s_3



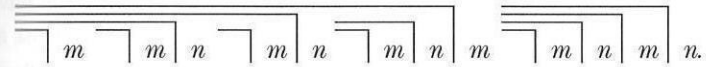
next in s_2



next in s_1

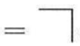


and finally in s_0



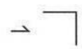
There is a dominant value in s_0.

Therefore

$e = m =$ 

Check by simplification.

 condensation

 cancellation (five times).

We have shown that the indicators of the two values in the calculus remain distinct when we take steps towards simplicity, thereby justifying the hypothesis of simplification. For completeness we must show that they remain similarly distinct when we take steps away from simplicity.

Theorem 4. Distinction

The value of any expression constructed by taking steps from a given simple expression is distinct from the value of any expression constructed by taking steps from a different simple expression

Proof

Consider any compound expression e_c constructed as a consequence of steps from a simple expression e_s.

Since each step in the construction of e_c can be retraced, there exists a simplification of e_c that leads to e_s.

But, by theorem 3, all simplifications of e_c agree. Hence all simplifications of e_c lead to e_s.

Thus, by the hypotheses of simplification, the use of which is justified in the proof of theorem 3, the only possible value of e_c is the value of e_s.

But e_s must be one of the simple expressions ⌉ or , that by definition have distinct values.

Therefore, the value of any expression constructed by taking steps from a given simple expression is distinct from the value of any expression constructed by taking steps from a different simple expression.

Consistency

We have now shown that the two values that the forms of the calculus are intended to indicate are not confused by any step allowed in the calculus and that, therefore, the calculus does in fact carry out its intention.

If, in a calculus intending several indications, they are anywhere confused, then they are everywhere confused, and if they are confused they are not distinguished, and if they are not distinguished they cannot be indicated, and the calculus thereby makes no indication.

A calculus that does not confuse a distinction it intends will be said to be consistent.

A classification of expressions

Expressions of the marked state may be called dominant. The letter m, unless otherwise employed, may be taken to indicate a dominant expression.

Expressions of the unmarked state may be called recessive. The letter n, unless otherwise employed, may be taken to indicate a recessive expression.

Theorem 5. Identity

Identical expressions express the same value.

In any case

$$x = x.$$

Proof

By theorems 3 and 4 we see that no step from an expression x can change the value expressed by x.

Therefore, any expression that can be reached by steps from x must have the same value as x.

But an expression identical with x can be reached by taking steps from x and then retracing them.

Thus any expression identical with x must express the same value as x.

Therefore

$$x = x$$

in any case.*

Theorem 6. Value

Expressions of the same value can be identified.

Proof

If x expresses the same value as y, then both x and y will simplify to the same simple expression, call it e_s.

Let $v = e_s$. Thus v will also simplify to e_s, and so v can be reached from either x or y by taking steps to e_s and then retracing the simplification of v.

Thus

$$x = v$$

and

$$y = v.$$

Therefore, by the convention of substitution, both x and y may be changed for an identical expression v in each case.

But x and y can be any equivalent expressions.

Therefore, expressions of the same value can be identified.

Theorem 7. Consequence

Expressions equivalent to an identical expression are equivalent to one another.

In any case, if

$$x = v$$

and

$$y = v,$$

then

$$x = y.$$

* Previous authors had to take $x = x$ as axiomatic. I am the first author to prove it from his axioms. – Author, 2000 06 27.

Proof

Let e_s be simple, and let $v = e_s$.

Now, since $x = v$ and $y = v$, e_s can be reached by steps from x and by steps from y.

Procedure. Take the steps from x to e_s and from e_s retrace the steps from y to e_s.

Thus y is reached by steps from x.

Therefore, if

$$x = v$$

and

$$y = v,$$

then

$$x = y$$

in any case.

Theorem 8. Invariance

If succesive spaces s_n, s_{n+1}, s_{n+2} are distinguished by two crosses, and s_{n+1} pervades an expression identical with the whole expression in s_{n+2}, then the value of the resultant expression in s_n is the unmarked state.

In any case,

$$\overline{\overline{p\,}|\;p\,}| = \qquad .$$

Proof

Let $p = \overline{\;\;}|$. In this case

$$\overline{\overline{p\,}|\;p\,}| = \overline{\overline{\overline{\;\;}|}|\;\overline{\;\;}|}| \qquad \text{substitution}$$

$$= \qquad \text{order (twice).}$$

Now let $p = \qquad$. In this case

$$\overline{\overline{p\,}|\;p\,}| = \overline{\overline{\;\;}|}| \qquad \text{substitution}$$

$$= \qquad \text{order.}$$

There is no other case of p, — theorem 1.

There is no other way of substituting any case of p, — theorems 5, 6.

Therefore,

$$\overline{\overline{p\,}|\;p\,}| =$$

in any case.

Theorem 9. Variance

If successive spaces s_n, s_{n+1}, s_{n+2} are arranged so that s_n, s_{n+1} are distinguished by one cross, and s_{n+1}, s_{n+2} are distinguished by two crosses (s_{n+2} being thus in two divisions), then the whole expression e in s_n is equivalent to an expression, similar in other respects to e, in which an identical expression has been taken out of each division of s_{n+2} and put into s_n.

In any case

$$\overline{\overline{pr\,}|\;\overline{qr\,}|}| = \overline{\overline{p\,}|\;\overline{q\,}|}|\;r.$$

Proof

Let $r = \overline{\;\;}|$.

Thus

$$\overline{\overline{pr\,}|\;\overline{qr\,}|}| = \overline{\overline{p\;\overline{\;\;}|}|\;\overline{q\;\overline{\;\;}|}|}| \qquad \text{substitution}$$

$$= \overline{\overline{\;\;}|\;\overline{\;\;}|}| \qquad \text{theorem 2 (twice)}$$

$$= \overline{\;\;}| \qquad \text{order (twice)}$$

and

$$\overline{\overline{p\,}|\;\overline{q\,}|}|\;r = \overline{\overline{p\,}|\;\overline{q\,}|}|\;\overline{\;\;}| \qquad \text{substitution}$$

$$= \overline{\;\;}| \qquad \text{theorem 2.}$$

Therefore, in this case,

$$\overline{\overline{pr\,}|\;\overline{qr\,}|}| = \overline{\overline{p\,}|\;\overline{q\,}|}|\;r \qquad \text{theorem 7.}$$

Now let $r = \qquad$.

Thus

$$\overline{\overline{pr\,}|\;\overline{qr\,}|}| = \overline{\overline{p\,}|\;\overline{q\,}|}| \qquad \text{substitution}$$

and

$$\overline{\overline{p\,}|\;\overline{q\,}|}|\;r = \overline{\overline{p\,}|\;\overline{q\,}|}| \qquad \text{substitution.}$$

Therefore, in this case,

$$\overline{\overline{pr\,}|\;\overline{qr\,}|}| = \overline{\overline{p\,}|\;\overline{q\,}|}|\;r \qquad \text{theorem 7.}$$

There is no other case of r, — theorem 1.

There is no other way of substituting any case of r, theorems 5, 6.

Therefore,

$$\overline{\overline{pr\,}|\ \overline{qr\,}|}\,| = \overline{\overline{p\,}|\ \overline{q\,}|}\,|\ r$$

in any case.

A classification of theorems

The first four theorems contain a statement of completeness and consistency of representation. Their proofs comprise a justification of the use of the primary arithmetic as a system of indicators of the states distinguished by the first distinction. We call them theorems of representation.

The next three theorems justify the use of certain procedural contractions without which subsequent justifications might become intolerably cumbersome. We call them theorems of procedure.

The last two theorems will serve as a gate of entry into a new calculus. We call them theorems of connexion.

The new calculus will itself give rise to further theorems that, when they describe aspects of the new calculus without direct reference to the old, will be called pure algebraic theorems, or theorems of the second order.

In addition we shall find theorems about the two calculi considered together. The bridge theorem and the theorem of completeness are examples, and we may call them mixed theorems.

5

A calculus taken out of the calculus

Let tokens of variable form

$$a,\ b,\ \ldots$$

indicate expressions in the primary arithmetic.

Let their values be unknown except in as far as, by theorem 5,

$$a = a,\ b = b,\ \ldots$$

Let tokens of constant form

$$\overline{}|$$

indicate instructions to cross the boundary of the first distinction according to the conventions already called.

Call any token of variable form after its form.

Call any token of constant form

cross.

Let indications used in the description of theorem 8 be taken out of context so that

$$\overline{\overline{p\,}|\ p\,}| = \quad .$$

Call this the form of position.

Let indications used in the description of theorem 9 be taken out of context so that

$$\overline{\overline{pr\,}|\ \overline{qr\,}|}\,| = \overline{\overline{p\,}|\ \overline{q\,}|}\,|\ r.$$

Call this the form of transposition.

Let the forms of position and transposition be taken as the initials of a calculus.

Let the calculus be seen as a calculus for the primary arithmetic.

Call it the primary algebra.

Algebraic calculation

For algebras, two rules are commonly accepted as implicit in the use of the sign $=$.

Rule 1. Substitution

If $e = f$, and if h is an expression constructed by substituting f for any appearance of e in g, then $g = h$.

Justification. This rule is a restatement of the arithmetical convention of substitution together with an inference from the theorems of representation.

Rule 2. Replacement

If $e = f$, and if every token of a given independent variable expression v in $e = f$ is replaced by an expression w, it not being necessary for v, w to be equivalent or for w to be independent or variable, and if as a result of this procedure e becomes j and f becomes k, then $j = k$.

Justification. This rule derives from the fact, proved with the theorems of connexion, that we can find equivalent expressions, not identical, that, considered arithmetically, are not wholly revealed. In an equation of such expressions each independent variable indicator stands for an expression that, being unknown except in as far as, by theorem 5, its value must be taken to be the same wherever its indicator appears, may be changed at will. Hence its indicator may also be changed at will, provided only that the change is made to every appearance of the indicator.*

Indexing

Numbered members of a class of findings will henceforth be indexed by a capital letter denoting the class followed by a figure denoting the number of the member. The classes will be indexed thus.

Consequence	C
Initial of the promary arihmetic	I
Initial of the primary algebra	J
Rule	R
Theorem	T

Certain equations, designated by E, will also be indexed, but the reference in each chapter will be confined to a separate set. Thus E1, etc, in Chapter 9 will not intentionally be the same equations as E1, etc, in Chapter 8.

* This is the first time the two universal rules of algebra were correctly stated. In other books they are not properly distinguished , or mentioned at all. Only one other author, Tarski, attempted to state them, and he got them wrong. – Author

6

The primary algebra

Initial 1. Position

J1 $\qquad \overline{\overline{p}\rceil\, p}\rceil = \qquad$ take out $\rightleftharpoons$ put in

Initial 2. Transposition

J2 $\qquad \overline{\overline{pr}\rceil\, \overline{qr}\rceil}\rceil = \overline{\overline{p}\rceil\, \overline{q}\rceil}\rceil\, r \qquad$ collect $\rightleftharpoons$ distribute

We shall proceed to distinguish particular equations, called consequences, that can be seen to condense certain sequences of these initials.

Consequence 1. Reflexion

C1 $\qquad \overline{\overline{a}\rceil}\rceil = a \qquad$ reflect $\rightleftharpoons$ reflect

Demonstration

We first find

$$\overline{\overline{a}\rceil}\rceil = \overline{\overline{\overline{a}\rceil}\rceil\, \overline{a}\rceil}\rceil\, \overline{\overline{a}\rceil}\rceil$$

by J1. We use R2 to convert $\overline{\overline{p}\rceil\, p}\rceil =$ to $\overline{\overline{\overline{a}\rceil}\rceil\, \overline{a}\rceil}\rceil =$ by changing every appearance of p to an appearance of $\overline{a}\rceil$. We next use R1 to change an appearance of to an appearance of $\overline{\overline{\overline{a}\rceil}\rceil\, \overline{a}\rceil}\rceil$ in the space with the original expression $\overline{\overline{a}\rceil}\rceil$, thus finding $\overline{\overline{a}\rceil}\rceil = \overline{\overline{\overline{a}\rceil}\rceil\, \overline{a}\rceil}\rceil\, \overline{\overline{a}\rceil}\rceil$.

We next find

$$\overline{\overline{\overline{a}\rceil}\rceil\, \overline{a}\rceil}\rceil\, \overline{\overline{a}\rceil}\rceil = \overline{\overline{\overline{\overline{a}\rceil}\rceil\, \overline{a}\rceil}\rceil\, \overline{\overline{\overline{a}\rceil}\rceil\, a}\rceil}\rceil$$

by J2. We make use of the license allowed in the definition (p 4) of $=$, to convert $\overline{\overline{pr}\rceil\, \overline{qr}\rceil}\rceil = \overline{\overline{p}\rceil\, \overline{q}\rceil}\rceil\, r$ to $\overline{\overline{p}\rceil\, \overline{q}\rceil}\rceil\, r = \overline{\overline{pr}\rceil\, \overline{qr}\rceil}\rceil$.

We next use the license allowed in the definition (p 5) of relation to change this to $\overline{\overline{p}\rceil\,\overline{q}\rceil}\rceil\, r = \overline{\overline{rp}\rceil\,\overline{rq}\rceil}\rceil$. We then use R2 to change every appearance of p in this equation to an appearance of $\overline{a}\rceil$, thus finding $\overline{\overline{\overline{a}\rceil}\rceil\,\overline{q}\rceil}\rceil\, r = \overline{\overline{r\,\overline{a}\rceil}\rceil\,\overline{rq}\rceil}\rceil$. We use R2 again to change every appearance of q in this equation to an appearance of a, and then again to change every appearance of r to an appearance of $\overline{\overline{a}\rceil}\rceil$, thus finding $\overline{\overline{\overline{a}\rceil}\rceil\,\overline{a}\rceil}\rceil\,\overline{\overline{a}\rceil}\rceil = \overline{\overline{\overline{\overline{a}\rceil}\rceil\,\overline{a}\rceil}\rceil\,\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil}\rceil$.

We then find

$$\overline{\overline{\overline{\overline{a}\rceil}\rceil\,\overline{a}\rceil}\rceil\,\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil}\rceil = \overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil$$

by J1. We found $\overline{\overline{\overline{a}\rceil}\rceil\,\overline{a}\rceil}\rceil = \quad$ for the first equation, and we now only need to use R1 to change the appearance of $\overline{\overline{\overline{a}\rceil}\rceil\,\overline{a}\rceil}\rceil$ in the space with $\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil$ to an appearance of $\quad$ in $\overline{\overline{\overline{\overline{a}\rceil}\rceil\,\overline{a}\rceil}\rceil\,\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil}\rceil$ to find $\overline{\overline{\overline{\overline{a}\rceil}\rceil\,\overline{a}\rceil}\rceil\,\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil}\rceil = \overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil$.

We then find

$$\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil = \overline{\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil\,\overline{\overline{a}\rceil\,a}\rceil}\rceil$$

by J1. We use R2 to convert $\overline{\overline{p}\rceil\,p}\rceil = \quad$ to $\overline{\overline{a}\rceil\,a}\rceil = \quad$ by changing all p to a, and then use R1 to change $\quad$ to $\overline{\overline{a}\rceil\,a}\rceil$ in the space with $\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil$, thus finding $\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil = \overline{\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil\,\overline{\overline{a}\rceil\,a}\rceil}\rceil$.

We then find

$$\overline{\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil\,\overline{\overline{a}\rceil\,a}\rceil}\rceil = \overline{\overline{\overline{\overline{a}\rceil}\rceil}\rceil\,\overline{\overline{a}\rceil}\rceil}\rceil\,a$$

by J2, using R2 to change all p to $\overline{\overline{a}\rceil}\rceil$, and then all q to $\overline{a}\rceil$, and then all r to a.

And lastly, we find

$$\overline{\overline{\overline{\overline{a}\rceil}\rceil}\rceil\,\overline{\overline{a}\rceil}\rceil}\rceil\,a = a$$

by J1. We find $\overline{\overline{\overline{\overline{a}\rceil}\rceil}\rceil\,\overline{\overline{a}\rceil}\rceil}\rceil = \quad$ by using R2 to change all p to $\overline{\overline{a}\rceil}\rceil$, and then use R1 to change $\overline{\overline{\overline{\overline{a}\rceil}\rceil}\rceil\,\overline{\overline{a}\rceil}\rceil}\rceil$ to $\quad$ in the space with a, thus finding $\overline{\overline{\overline{\overline{a}\rceil}\rceil}\rceil\,\overline{\overline{a}\rceil}\rceil}\rceil\,a = a.$

This completes a detailed account of each of six steps. We may now use T7 five times to find

$$\overline{\overline{a}\rceil}\rceil = a$$

and this completes the demonstration.

We repeat this demonstration, and give subsequent demonstrations, with only the key indices to the procedure.

$\overline{\overline{a}\rceil}\rceil$		
	$= \overline{\overline{\overline{a}\rceil}\rceil\,\overline{a}\rceil}\rceil\,\overline{\overline{a}\rceil}\rceil$	J1
	$= \overline{\overline{\overline{\overline{a}\rceil}\rceil\,\overline{a}\rceil}\rceil\,\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil}\rceil$	J2
	$= \overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil$	J1
	$= \overline{\overline{\overline{\overline{a}\rceil}\rceil\,a}\rceil\,\overline{\overline{a}\rceil\,a}\rceil}\rceil$	J1
	$= \overline{\overline{\overline{\overline{a}\rceil}\rceil}\rceil\,\overline{\overline{a}\rceil}\rceil}\rceil\,a$	J2
	$= a$	J1.

Consequence 2. Generation

C2	$\overline{ab}\rceil\,b = \overline{a}\rceil\,b$	degenerate $\rightleftharpoons$ regenerate

Demonstration

$$\overline{ab\,}|\; b$$

$$= \overline{\overline{\overline{a\,}|}|\; b\,}|\; b \qquad \text{C1}$$

$$= \overline{\overline{\overline{a\,}|}|\; \overline{\overline{b\,}|}|}|\; b \qquad \text{C1}$$

$$= \overline{\overline{\overline{a\,}|\; b\,}|\; \overline{\overline{b\,}|\; b\,}|}| \qquad \text{J2}$$

$$= \overline{\overline{\overline{a\,}|\; b\,}|}| \qquad \text{J1}$$

$$= \overline{a\,}|\; b \qquad \text{C1.}$$

Consequence 3. Integration

C3 $$\overline{\;\;}|\; a = \overline{\;\;}|$$ reduce ⇌ augment

Demonstration

$$\overline{\;\;}|\; a$$

$$= \overline{a\,}|\; a \qquad \text{C2}$$

$$= \overline{\overline{\overline{a\,}|\; a\,}|}| \qquad \text{C2}$$

$$= \overline{\;\;}| \qquad \text{J1.}$$

Consequence 4. Occultation

C4 $$\overline{\overline{a\,}|\; b\,}|\; a = a$$ conceal ⇌ reveal

Demonstration

$$\overline{\overline{a\,}|\; b\,}|\; a$$

$$= \overline{\overline{a\,}|\; ba\,}|\; a \qquad \text{C2}$$

$$= \overline{\overline{ab\,}|\; ba\,}|\; a \qquad \text{C2}$$

$$= a \qquad \text{J1.}$$

Consequence 5. Iteration

C5 $$aa = a$$ iterate ⇌ reiterate

Demonstration

$$aa$$

$$= \overline{\overline{a\,}|}|\; a \qquad \text{C1}$$

$$= a \qquad \text{C4.}$$

Consequence 6. Extension

C6 $$\overline{\overline{a\,}|\; \overline{b\,}|}|\; \overline{\overline{a\,}|\; b\,}| = a$$ contract ⇌ expand

Demonstration

$$\overline{\overline{a\,}|\; \overline{b\,}|}|\; \overline{\overline{a\,}|\; b\,}|$$

$$= \overline{\overline{\overline{\overline{a\,}|\; \overline{b\,}|}|\; \overline{\overline{a\,}|\; b\,}|}|}| \qquad \text{C1}$$

$$= \overline{\overline{\overline{\overline{b\,}|}|\; \overline{b\,}|}|\; \overline{a\,}|}| \qquad \text{J2}$$

$$= \overline{\overline{a\,}|}| \qquad \text{J1}$$

$$= a \qquad \text{C1.}$$

Consequence 7. Echelon

C7 $$\overline{\overline{\overline{a}|\,b}|\,c}| = \overline{ac}|\ \overline{\overline{b}|\,c}|$$ break ⇌ make

Demonstration

$$\overline{\overline{\overline{a}|\,b}|\,c}|$$

$$= \overline{\overline{\overline{a}|\ \overline{\overline{b}|}|}|\,c}| \qquad \text{C1}$$

$$= \overline{\overline{\overline{ac}|\ \overline{\overline{b}|\,c}|}|}| \qquad \text{J2}$$

$$= \overline{ac}|\ \overline{\overline{b}|\,c}| \qquad \text{C1.}$$

Consequence 8. Modified transposition

C8 $$\overline{\overline{a}|\ \overline{br}|\ \overline{cr}|}| = \overline{\overline{a}|\ \overline{b}|\ \overline{c}|}|\ \overline{\overline{a}|\ \overline{r}|}|$$ collect ⇌ distribute

Demonstration

$$\overline{\overline{a}|\ \overline{br}|\ \overline{cr}|}|$$

$$= \overline{\overline{a}|\ \overline{\overline{\overline{br}|\ \overline{cr}|}|}|}| \qquad \text{C1}$$

$$= \overline{\overline{a}|\ \overline{\overline{\overline{b}|\ \overline{c}|}|\,r}|}| \qquad \text{J2}$$

$$= \overline{\overline{a}|\ \overline{b}|\ \overline{c}|}|\ \overline{\overline{a}|\ \overline{r}|}| \qquad \text{C7.}$$

Consequence 9. Crosstransposition

C9 $$\overline{\overline{\overline{b}|\ \overline{r}|}|\ \overline{\overline{a}|\ \overline{r}|}|\ \overline{\overline{x}|\,r}|\ \overline{\overline{y}|\,r}|}| = \overline{\overline{r}|\,ab}|\ \overline{rxy}|$$ crosstranspose (collect) ⇌ crosstranspose (distribute)

Demonstration

$$\overline{\overline{\overline{b}|\ \overline{r}|}|\ \overline{\overline{a}|\ \overline{r}|}|\ \overline{\overline{x}|\,r}|\ \overline{\overline{y}|\,r}|}|$$

$$= \overline{\overline{\overline{b}|\ \overline{r}|}|\ \overline{\overline{a}|\ \overline{r}|}|\ \overline{\overline{xy}|\,r}|}| \qquad \text{C1, J2, C1}$$

$$= \overline{ba\ \overline{\overline{xy}|\,r}|}|\ \overline{r\ \overline{\overline{xy}|\,r}|}| \qquad \text{C8, C1 (thrice)}$$

$$= \overline{ba\ \overline{\overline{xy}|\,r}|}|\ \overline{rxy}| \qquad \text{C2, C1}$$

$$= \overline{ba\ \overline{\overline{xy}|\,r}|\ \overline{rxy}|}|\ \overline{rxy}| \qquad \text{C2}$$

$$= \overline{\overline{r}|\,ab}|\ \overline{rxy}| \qquad \text{C6.}$$

The classification of consequences

In classifying these consequences, there is no need to confine them rigidly to the forms above. The name of a consequence may indicate a part of the consequence as in

$$\overline{a}|\,a = \overline{\quad}| \qquad \text{integration.}$$

In another case it may include reflexions as in

$$\overline{\overline{a}|\,r}|\ \overline{\overline{b}|\,r}| = \overline{\overline{ab}|\,r}| \qquad \text{transposition or echelon}$$

and

$$\overline{\overline{b}|\,a}|\ \overline{ba}| = \overline{a}| \qquad \text{extension.}$$

In yet another case it may indicate a crosstransposed form such as

$$\overline{\overline{ab}|\ \overline{a\ \overline{b}|}|}| = a \qquad \text{extension.}$$

Nor, as we already see in one case, are the classes of consequence properly distinct. What we are doing is indicating larger and larger numbers of steps in a single indication. This is the dual form of the contraction of a reference, notably the expansion of its content. We shed the labour of calculation by taking a number of steps as one step.

Thus if we consider the equivalence of steps, we find

$$\rightharpoonup\rightharpoonup\; = \;\rightharpoonup.$$

Also, since to retrace a step can be considered as not to take it, we find

$$\rightharpoonup\leftharpoondown\; = \quad .$$

But now if we allow steps in the indication of steps, we find that the resulting calculus is inconsistent.

Thus

$$\rightharpoonup\rightharpoonup\leftharpoondown\; = \;\rightharpoonup\leftharpoondown\; = \quad ,$$

or

$$\rightharpoonup\rightharpoonup\leftharpoondown\; = \;\rightharpoonup,$$

according to which step we take first.

Therefore

$$\quad = \;\rightharpoonup,$$

which suggests that, in any calculation, we regard any number of steps, including zero, as a step.

This agrees with our idea of the nature of a step that, as we have already determined, is not intended to cross a boundary.

A further classification of expressions

The algebraic consideration of the calculus of indications leads to a further distinction between expressions.

Expressions of the marked state may be called integral. The letter m, unless otherwise employed, may be taken to indicate an integral expression.

Expressions of the unmarked state may be called disintegral. The letter n, unless otherwise employed, may be taken to indicate a disintegral expression.

Expressions of a state consequent on the states of their unknown indicators may be called consequential. The letter v, unless otherwise employed, may be taken to indicate a consequential expression.

7

Theorems of the second order

Theorem 10

The scope of J2 can be extended to any number of divisions of the space s_{n+2}.

In any case,

$$\overline{\overline{a\,\rvert}\;\overline{b\,\rvert}\;\ldots\,\rvert}\;r = \overline{\overline{ar\,\rvert}\;\overline{br\,\rvert}\;\ldots\,\rvert}.$$

Proof

We consider the cases in which s_{n+2} is divided into 0, 1, 2, and more than 2 divisions respectively. In case 0

$$\overline{\ \ \rvert}\;r = \overline{\ \ \rvert} \qquad \text{C3.}$$

In case 1

$$\overline{\overline{a\,\rvert}\,\rvert}\;r = ar \qquad \text{C1}$$

$$= \overline{\overline{ar\,\rvert}\,\rvert} \qquad \text{C1.}$$

In case 2

$$\overline{\overline{b\,\rvert}\;\overline{a\,\rvert}\,\rvert}\;r = \overline{\overline{br\,\rvert}\;\overline{ar\,\rvert}\,\rvert} \qquad \text{J2.}$$

In case more than 2

$$\overline{\ldots\;\overline{c\,\rvert}\;\overline{b\,\rvert}\;\overline{a\,\rvert}\,\rvert}\;r$$

$$= \overline{\overline{\overline{\overline{\overline{\ldots\;\overline{c\,\rvert}\,\rvert}\,\rvert}\;\overline{b\,\rvert}\,\rvert}\,\rvert}\;\overline{a\,\rvert}\,\rvert}\;r \qquad \text{C1 (as often as necessary)}$$

$$= \overline{\overline{\overline{\overline{\overline{\ldots\;\overline{cr\,\rvert}\,\rvert}\,\rvert}\;\overline{br\,\rvert}\,\rvert}\,\rvert}\;\overline{ar\,\rvert}\,\rvert} \qquad \text{J2 (as often as necessary)}$$

$$= \overline{\ldots\;\overline{cr\,\rvert}\;\overline{br\,\rvert}\;\overline{ar\,\rvert}\,\rvert} \qquad \text{C1 (as often as before).}$$

This completes the proof.

Theorem 11

The scope of C8 can be extended as in T10.

$$\overline{\overline{a}|\;\overline{br}|\;\overline{cr}|\;\ldots}| \;=\; \overline{\overline{a}|\;\overline{b}|\;\overline{c}|\;\ldots}|\;\overline{\overline{a}|\;\overline{r}|}|$$

Theorem 12

The scope of C9 can be extended as in T10.

$$\overline{\ldots\;\overline{\overline{b}|\;\overline{r}|}|\;\overline{\overline{a}|\;\overline{r}|}|\;\overline{\overline{x}|\;r}|\;\overline{\overline{y}|\;r}|\;\ldots}|$$

$$=\;\overline{\overline{r}|\;ab\ldots}|\;\overline{rxy\ldots}|$$

Proofs of T11 and T12 follow from demonstrations as in C8 and C9, using T10 instead of J2.

Theorem 13

The generative process in C2 can be extended to any space not shallower than that in which the generated variable first appears.

Proof

We consider cases in which a variable is generated in spaces 0, 1, and more than 1 space deeper than the space of the variable of origin. In case 0

$$\overline{\overline{\overline{\ldots c}|\;b}|\;a}|\;g$$

$$=\;\overline{\overline{\overline{\ldots c}|\;b}|\;a}|\;gg \qquad \text{C5.}$$

In case 1

$$\overline{\overline{\overline{\ldots c}|\;b}|\;a}|\;g$$

$$=\;\overline{\overline{\overline{\ldots c}|\;b}|\;ag}|\;g \qquad \text{C2.}$$

In case more than 1

$$\overline{\overline{\overline{\ldots c}|\;b}|\;a}|\;g$$

$$=\;\overline{\overline{\overline{\ldots c}|\;b}|\;ag}|\;g \qquad \text{C2}$$

$$=\;\overline{\overline{\overline{\ldots c}|\;bg}|\;ag}|\;g \qquad \text{C2}$$

$$=\;\overline{\overline{\overline{\ldots c}|\;bg}|\;a}|\;g \qquad \text{C2.}$$

and so on. Clearly no additional consideration is needed for further generation of g, and it is plain that any space not shallower than that in which g stands can be reached.

It is convenient to consider J2, C2, C8, and C9 as extended by their respective theorems, and to let the name of the initial or consequence denote also the theorem extending it.

Theorem 14. Canon with respect to the constant

From any given expression, an equivalent expression not more than two crosses deep can be derived.

Proof

Suppose that a given expression e has i deepest spaces of the depth d, and that $d > 2$.

We carry out a depth-reducing procedure with C7. Inspection of possibilities shows that not more than $2^i - 1$ steps are needed to find $e = e_1$ so that e_1 has (say) j deepest spaces of depth $d - 1$. (The maximum number of steps is needed in case the part of s_{d-2} in e is the only part containing s_d, and each division of s_d is contained in a separate division of s_{d-1}.) If $(d - 1) > 2$ we continue the procedure with at most $2^j - 1$ additional steps to find $e_1 = e_2$ so that e_2 is only $d - 2$ crosses deep. We see that the procedure can be continued until we find $e = e_{d-2}$ so that e_{d-2} is only $d - (d - 2) = 2$ crosses deep, and this completes the proof.

Theorem 15. Canon with respect to a variable

From any given expression, an equivalent expression can be derived so as to contain not more than two appearances of any given variable.

Proof

The proof is trivial for a variable not contained in the original expression e, and so we may confine our consideration to the case of a variable v contained in e.

Now by C1 and T14

$$e = \ldots\;\overline{\overline{vb}|\;q}|\;\overline{\overline{va}|\;p}|\;f\;\overline{vx}|\;\overline{vy}|\;\ldots,$$

in which $a, b, \ldots, p, q, \ldots, x, y, \ldots$ and f stand for arrangements appropriate to the expression e,

$$= \ldots\ \overline{\overline{v}|\ q}|\ \overline{\overline{b}|\ p}|\ \overline{\overline{v}|\ p}|\ \overline{\overline{a}|\ p}|\ f\ \overline{vx}|\ \overline{vy}|\ldots \qquad \text{C1, J2, C1 (each as often as necessary)}$$

$$= \ldots\ \overline{\overline{v}|\ q}|\ \overline{\overline{v}|\ p}|\ g\ \overline{vx}|\ \overline{vy}|\ldots \qquad \text{calling } g = f\ \overline{\overline{a}|\ p}|\ \overline{\overline{b}|\ p}|\ldots$$

$$= \overline{\overline{\overline{p}|\ \overline{q}|\ \ldots}|\ \overline{v}|}|\ \overline{\overline{\overline{x}|\ \overline{y}|\ \ldots}|\ v}|\ g \qquad \text{C1, J2 (twice each)}$$

and this completes the proof.

8

Re-uniting the two orders

Content, image, and reflexion

Of any expression e, call e the content, call $\overline{e}|$ the image, and call $\overline{\overline{e}|}|$ the reflexion.

Since $\overline{\overline{e}|}| = e$, the act of reflexion is a return from an image to its content or from a content to its image.

Suppose e is a cross. The content of e is the content of the space in which it stands, not the content of the cross that marks the space.

In general, a content is where we have marked it, and a mark is not inside the boundary shaping its form, but inside the boundary surrounding it and shaping another form. Thus in describing a form, we find a succession,

$$\text{content}$$

$$\overline{\text{content}}|\ \text{image}$$

$$\overline{\overline{\text{content}}|\ \text{image}}|\ \text{content}$$

$$\overline{\overline{\overline{\text{content}}|\ \text{image}}|\ \text{content}}|\ \text{image}\ \ldots$$

Indicative space

If s_0 is the pervasive space of e, the value of e is its value to s_0. If e is the whole expression in s_0, s_0 takes the value of e and we can call s_0 the indicative space of e.

In evaluating e we imagine ourselves in s_0 with e and thus surrounded by the unwritten cross that is the boundary to s_{-1}.

Seventh canon. Principle of relevance

*If a property is common to every indication, it need not be indicated.**

An unwritten cross** is common to every expression in the calculus of indications and

* In fact it *cannot* be indicated, since every attempted indication of it would be devoid of meaning. In short, we cannot say anything of everything. Cf. J M Robertson, *Letters on reasoning*, London 1902, 'There can be no rational ascription of a single mode to the totality of things.'

** Nonmathematicians have attributed some deep philosophical meaning to the unwritten cross, but it is merely a device to collect the whole expression into one bracket. – Author.

so need not be written. Similarly, a recessive value is common to every expression in the calculus of indications and also, by this principle, has no necessary indicator there.

In the form of any calculus, we find the consequences in its content and the theorems in its image.

Thus

$$\overline{\overline{\overline{\;\;}|}|\;\overline{\;\;}|}|\;\overline{\;\;}|\;\overline{\;\;}| = \overline{\;\;}|$$

is a consequence in, and therefore in the content of, the primary arithmetic.

Demonstration

$$\begin{aligned} &\overline{\overline{\overline{\;\;}|}|\;\overline{\;\;}|}|\;\overline{\;\;}|\;\overline{\;\;}| && \\ &= \overline{\overline{\;\;}|}|\;\overline{\;\;}|\;\overline{\;\;}| && \text{I2} \\ &= \overline{\;\;}|\;\overline{\;\;}| && \text{I2} \\ &= \overline{\;\;}| && \text{I1.} \end{aligned}$$

A consequence is acceptable because we decided the rules. All we need to show is that it follows through them.

But demonstrations of any but the simplest consequences in the content of the primary arithmetic are repetitive and tedious, and we can contract the procedure by using theorems, that are about, or in the image of, the primary arithmetic. For example, instead of demonstrating the consequence above, we can use T2.

T2 is a statement that all expressions of a certain kind, which it describes without enumeration, and of which the expression above can be recognized as an example, indicate the marked state. Its proof may be regarded as a simultaneous demonstration of all the simplifications of expressions of the kind it describes.

But the theorem itself is not a consequence. Its proof does not proceed according to the rules of the arithmetic, but follows, instead, through ideas and rules of reasoning and counting that, at this stage, we have done nothing to justify.

Thus if any person will not accept a proof, we can do no better than try another. A theorem is acceptable because what it states is evident, but we do not as a rule consider it worth recording if its evidence does not need, in some way, to be made evident. This rule is excepted in the case of an axiom, which may appear evident without further guidance. Both axioms and theorems are more or less simple statements about the ground on which we have chosen to reside.

Since the initial steps in the algebra were taken to represent theorems about the arithmetic, it depends on our point of view whether we regard an equation with variables as expressing a consequence in the algebra or a theorem about the arithmetic. Any demonstrable consequence is alternatively provable as a theorem, and this fact may be of use where the sequence of steps is difficult to find. Thus, instead of demonstrating in algebra the equation

$$\overline{\overline{\overline{a}|\;\overline{b}|}|\;\overline{a\;\overline{c}|}|}| = \overline{\overline{a}|\;b}|\;\overline{ac}|,$$

we can prove it by arithmetic.

Call

E1 $$\overline{\overline{\overline{a}|\;\overline{b}|}|\;\overline{a\;\overline{c}|}|}| = x$$

and

E2 $$\overline{\overline{a}|\;b}|\;\overline{ac}| = y.$$

Set $a = \overline{\;\;}|$. Thus

$$\begin{aligned} x &= \overline{\overline{\overline{\overline{\;\;}|}|\;\overline{b}|}|\;\overline{\overline{\;\;}|\;\overline{c}|}|}| && \text{substitution in E1} \\ &= \overline{\overline{\overline{b}|}|\;\overline{\overline{\;\;}|\;\overline{c}|}|}| && \text{I2} \\ &= \overline{\overline{\overline{b}|}|\;\overline{\overline{\;\;}|\;\overline{b}|\;\overline{c}|}|}| && \text{T2, T7} \\ &= \overline{\overline{\;\;}|\;\overline{\overline{\;\;}|\;\overline{c}|}|}|\;\overline{b}| && \text{T9} \\ &= \overline{\overline{\;\;}|\;\overline{\overline{\;\;}|}|}|\;\overline{b}| && \text{T2} \\ &= \overline{b}| && \text{I2 (twice)} \end{aligned}$$

and

$$\begin{aligned} y &= \overline{\overline{\overline{\;\;}|}|\;b}|\;\overline{\overline{\;\;}|\;c}| && \text{substitution in E2} \\ &= \overline{\overline{\overline{\;\;}|}|\;b}|\;\overline{\overline{\;\;}|}| && \text{T2} \\ &= \overline{b}| && \text{I2 (twice)} \end{aligned}$$

and so x = y in this case T7.

Now set $a =$. Thus

$$x = \overline{\overline{\overline{\;\;}|\,\overline{b}|}|\,\overline{\overline{c}|}|}| \qquad \text{substitution in E1}$$

$$= \overline{\overline{\overline{\;\;}|\,\overline{b}|\,\overline{c}|}|\,\overline{\overline{c}|}|}| \qquad \text{T2, T7}$$

$$= \overline{\overline{\overline{\;\;}|\,\overline{b}|}|\,\overline{\;\;}|}|\,\overline{c}| \qquad \text{T9}$$

$$= \overline{\overline{\overline{\;\;}|}|\,\overline{\;\;}|}|\,\overline{c}| \qquad \text{T2}$$

$$= \overline{c}| \qquad \text{I2 (twice)}$$

and

$$y = \overline{\overline{\;\;}|\,b}|\,\overline{c}| \qquad \text{substitution in E2}$$

$$= \overline{\overline{\;\;}|}|\,\overline{c}| \qquad \text{T2}$$

$$= \overline{c}| \qquad \text{I2}$$

and so $x = y$ in this case, T7.

There is no other case, T1.

Therefore $x = y$.

By their origin, the consequences in the algebra are arithmetically valid, so we may use them as we please to shorten the proof.

Abridged proof

Set $a = \overline{\;\;}|$. Thus

$$x = \overline{\overline{\overline{\overline{\;\;}|}|\,\overline{b}|}|\,\overline{\overline{\;\;}|\,\overline{c}|}|}| \qquad \text{substitution in E1}$$

$$= \overline{b}| \qquad \text{C3, C1 (thrice)}$$

and

$$y = \overline{\overline{\overline{\;\;}|}|\,b}|\,\overline{\overline{\;\;}|\,c}| \qquad \text{substitution in E2}$$

$$= \overline{b}| \qquad \text{C3, C1 (twice)}$$

and so $x = y$ in this case, T7.

Set $a =$. Thus

$$x = \overline{\overline{\overline{\;\;}|\,\overline{b}|}|\,\overline{\overline{c}|}|}| \qquad \text{substitution in E1}$$

$$= \overline{c}| \qquad \text{C3, C1 (twice)}$$

and

$$y = \overline{\overline{\;\;}|\,b}|\,\overline{c}| \qquad \text{substitution in E2}$$

$$= \overline{c}| \qquad \text{C3, C1}$$

and so $x = y$ in this case, T7.

There is no other case, T1.

Therefore $x = y$.

In these proofs we evidently supposed the irrelevance of variables other than the one we fixed arithmetically. It may not at first be obvious that we can ignore the possible values of the other variables, but the supposition is in fact justified in all instances (and, indeed, in all algebras), as the following proof will show.

Theorem 16. The bridge

If expressions are equivalent in every case of one variable, they are equivalent.

Let a variable v in a space s_q oscillate between the limits of its value m, n.

If the value of another indicator in s_q is n, the oscillation of v will be transmitted through s_q and seen as a variation in the value of the boundary of s_q to s_{q-1}.

Under this condition call s_q transparent.

If the value of any other indicator in s_q is m, nothing will be transmitted through s_q.

Under this condition call s_q opaque.

The transmission from v is the alternation between transparency and opacity in s_q and in any more distant space in which this alternation can be detected. It may at any point be absorbed in transmissions from other variables in the space through which it passes. On condition that this absorption is total, call the band of space in which it occurs opaque. Under any other condition, call it transparent.

From these definitions and considerations we can see the following principle.

Eighth canon. Principle of transmission

With regard to an oscillation in the value of a variable, the space outside the variable is either transparent or opaque.

Proof of theorem 16

Let s, t be the indicative spaces of e, f respectively.

Let either of e, f contain a variable v, and let v oscillate between the limits of its value m, n.

Consider the condition under which both e and f are opaque to transmission from v. If e and f are equivalent after a change in the value of v, they were equivalent before.

Thus $e = f$ under this condition.

Consider either e or f transparent.

Suppose the oscillation of v is transmitted to one indicative space and not to the other. By selecting an appropriate value of v, we could make e not equivalent to f, and this is contrary to our hypothesis that they are equivalent. Thus if either of e or f is transparent, both are transparent.

Thus any change in the value of v is transmitted to s and t.

Therefore, if e and f are equivalent after a change in v, they were equivalent before.

Thus $e = f$ under this condition.

But, by the principle of transmission, there is no other condition.

Therefore $e = f$ under any condition, and hence in any case.

This completes the proof.*

* Again, no other author was able to prove this before, so it was taken as axiomatic -Author.

9

Completeness

We have seen that any demonstrable consequence in the algebra must indicate a provable theorem about the arithmetic. In this way consequences in the algebra may be said to represent properties of the arithmetic. In particular, they represent the properties of the arithmetic that can be expressed in forms of equation.

We can question whether the algebra is a complete or only a partial account of these properties. That is to say, we can ask whether or not every form of equation that can be proved as a theorem about the arithmetic can be demonstrated as a consequence in the algebra.

Theorem 17. Completeness

The primary algebra is complete.

That is to say, if $\alpha = \beta$ can be proved as a theorem about the primary arithmetic, then it can be demonstrated as a consequence for all equivalent α, β in the primary algebra.

We prove this theorem by induction. We first show that if all cases of $\alpha = \beta$ are algebraically demonstrable with less than a certain positive number n of distinct variables, then so is any case of $\alpha = \beta$ with n distinct variables. We then show that the condition of complete demonstrability in cases of less than n variables does in fact hold for some positive value of n.

Proof

Suppose that the demonstrability $\alpha = \beta$ is established for all α, β, containing an aggregate of less than n distinct variables.

Let a given equivalent α, β contain between them n distinct variables.

Procedure. Reduce the given α, β to their canonical forms, say α', β', with respect to a variable v.

We see in the proofs of T14 and T15 that this reduction is algebraic, so that $\alpha = \alpha'$ and $\beta = \beta'$ are both demonstrable, and that no distinct variable is added during the course of it.

By the proof of T15 we may suppose the canonical form of α to be $\overline{\overline{v|}\ A_1}|\ \overline{v\ A_2}|\ A_3$, and that of β to be $\overline{\overline{v|}\ B_1}|\ \overline{v\ B_2}|\ B_3$. Hence

E1 $$\alpha = \overline{\overline{v|}\ A_1}|\ \overline{v\ A_2}|\ A_3$$

and

E2 $$\beta = \overline{\overline{v}|\;B_1}|\;\overline{v\,B_2}|\;B_3$$

are both demonstrable. Thus

$$\overline{\overline{v}|\;A_1}|\;\overline{v\,A_2}|\;A_3 = \overline{\overline{v}|\;B_1}|\;\overline{v\,B_2}|\;B_3$$

is true, although we do not yet know if it is demonstrable. But by substituting constant values for v we find

E3 $$\overline{A_1}|\;A_3 = \overline{B_1}|\;B_3$$

E4 $$\overline{A_2}|\;A_3 = \overline{B_2}|\;B_3.$$

Now each of E3, E4, having at most $n-1$ distinct variables, is demonstrable by hypothesis. Hence E1 through E4 are all demonstrable, and we can demonstrate

$$\alpha = \overline{\overline{v}|\;A_1}|\;\overline{v\,A_2}|\;A_3 \qquad \text{E1}$$

$$= \overline{\overline{\overline{v}|\;\overline{A_1}|\;v\;\overline{A_2}|}|}|\;A_3 \qquad \text{C9}$$

$$= \overline{\overline{\overline{v}|\;\overline{\overline{A_1}|\;A_3}|\;\overline{v\;\overline{A_2}|\;A_3}|}|}| \qquad \text{J2}$$

$$= \overline{\overline{\overline{v}|\;\overline{\overline{B_1}|\;B_3}|\;\overline{v\;\overline{B_2}|\;B_3}|}|}| \qquad \text{E3, E4}$$

$$= \beta \qquad \text{J2, C9, E2.}$$

Thus $\alpha = \beta$ is demonstrable with n variables on condition that it is demonstrable with fewer than n variables.

It remains to show that there exists a positive value of n for which $\alpha = \beta$ is demonstrable for all equivalent α, β with fewer than n variables.

It is sufficient to prove the condition for $n = 1$. Thus we need to show that if $\alpha = \beta$ contains no variable, it is demonstrable in the algebra.

If α, β contain no variable, they may be considered as expressions in the primary arithmetic.

We see in the proofs of T1 through T4 that all arithmetical equations are demonstrable in the arithmetic. It remains to show that they are demonstrable in the algebra.

In C3 let $a = \overline{\;\;}|$ to give

$$\overline{\;\;}|\;\overline{\;\;}| = \overline{\;\;}|$$

and this is I1.

In C1 let $a = \quad$ to give

$$\overline{\overline{\;\;}|}| =$$

and this is I2.

Thus the initials of the arithmetic are demonstrable in the algebra, and so if $\alpha = \beta$ contains no variable it is demonstrable in the algebra.

This completes the proof.

10

Independence

We call the equations in a set independent if no one equation can be demonstrated from the others.

Theorem 18. Independence

The initials of the primary algebra are independent.

That is to say, given J1 as the only initial, we cannot find J2 as a consequence, and given J2 as the only initial, we cannot find J1 as a consequence.

Proof

Suppose J1 determine the only transformation allowed in the algebra. It follows from the convention of intention that no expression other than of the form $\overline{\overline{p}|\,p}|$ can be put into or taken out of any space.

But in J2, r is taken out of one space and put into another, and r is not necessarily of the form $\overline{\overline{p}|\,p}|$.

Therefore, J2 cannot be demonstrated as a consequence of J1.

Next suppose J2 determine the only transformation allowed in the algebra.

Inspection of J2 reveals no way of eliminating any distinct variable.

But J1 eliminates a distinct variable.

Therefore, J1 cannot be demonstrated as a consequence of J2, and this completes the proof.

11

Equations of the second degree

Hitherto we have obeyed a rule (theorem 1) that requires that any given expression, in either the arithmetic or the algebra, shall be finite. Otherwise, by the canons so far called, we should have no means of finding its value.

It follows that any given expression can be reached from any other given equivalent expression in a finite number of steps. We shall find it convenient to extract this principle as a rule to characterize the process of demonstration.

Ninth canon. Rule of demonstration

A demonstration rests in a finite number of steps.

One way to see that this rule is obeyed is to count steps. We need not confine its application to any given level of consideration. In an algebraic expression each variable represents an unknown (or immaterial) number of crosses, and so it is not possible in this case to count arithmetical steps. But we can still count algebraic steps.

We may note that, according to the observation in Chapter 6 on the nature of a step, it does not matter if several counts disagree, as long as at least one count is finite.

Consider the expression $\overline{\overline{a}|\,b}|$. We propose now to generate a step-sequence of the following form.

$$\overline{\overline{a}|\,b}|$$

$$= \overline{\overline{a}|\,b}|\;\overline{\overline{a}|\,b}| \qquad \text{C5}$$

$$= \overline{\overline{a}|\;\overline{\overline{b}|}|}|\;\overline{\overline{a}|\,b}| \qquad \text{C1}$$

$$= \overline{\overline{\overline{\overline{a}|\,b}|\,a}|\;\overline{\overline{\overline{a}|\,b}|\;\overline{b}|}|}| \qquad \text{J2}$$

$$= \overline{\overline{\overline{\overline{a}|\,b}|\,a}|\;\overline{\overline{b}|}|}| \qquad \text{C4}$$

$$= \overline{\overline{\overline{\overline{a}|\,b}|\,a}|\,b}| \qquad \text{C1}$$

$$= \overline{\overline{\overline{\overline{a|}\;b|}\;\overline{\overline{\overline{a|}\;b|}\;a|}\;b|}} \qquad \text{C5}$$

$$= \overline{\overline{\overline{\overline{a|}\;\overline{\overline{b|}|}|}\;\overline{\overline{\overline{a|}\;b|}\;a|}\;b|}} \qquad \text{C1}$$

$$= \overline{\overline{\overline{\overline{\overline{a|}\;b|}\;a|}\;\overline{\overline{\overline{a|}\;b|}\;\overline{\overline{b|}|}|}\;a|}\;b|} \qquad \text{J2}$$

$$= \overline{\overline{\overline{\overline{\overline{a|}\;b|}\;a|}\;\overline{\overline{b|}|}|\;a|}\;b|} \qquad \text{C4}$$

$$= \overline{\overline{\overline{\overline{\overline{\overline{a|}\;b|}\;a|}\;b|}\;a|}\;b|} \qquad \text{C1}$$

etc. There is no limit to the possibility of continuing the sequence, and thus no limit to the size of the echelon of alternating a's and b's with which $\overline{\overline{a|}\;b|}$ can be equated.

Let us imagine, if we can, that the order to begin the step-sequence is never countermanded, so that the process continues timelessly. In space this will give us an echelon without limit, of the form

$$\overline{\overline{\overline{\overline{\ldots a|}\;b|}\;a|}\;b|}\,.$$

Now, since this form, being endless, cannot be reached in a finite number of steps from $\overline{\overline{a|}\;b|}$, we do not expect it to express, necessarily, the same value as $\overline{\overline{a|}\;b|}$. But we can, by means of an exhaustive examination of possibilities, ascertain what values it might take in the various cases of a, b, and compare them with those of the values it might take in the various cases of a, b, and compare them with those of the finite expression.

Re-entry

The key is to see that the crossed part of the expression at every even depth is identical with the whole expression, that can thus be regarded as re-entering its own inner space at any even depth. Thus

$$f = \overline{\overline{\overline{\overline{\ldots a|}\;b|}\;a|}\;b|}$$

E1
$$= \overline{\overline{fa|}\;b|}\,.$$

We can now find, by the rule of dominance, the values that f may take in each possible case of a, b

$$\overline{\overline{fa|}\;b|} = f \qquad \text{E1}$$

$$\overline{\overline{fm|}\;m|} = n$$

$$\overline{\overline{fm|}\;n|} = m$$

$$\overline{\overline{fn|}\;m|} = n$$

$$\overline{\overline{fn|}\;n|} = m \text{ or } n.$$

For the last case suppose $f = m$. Then $\overline{\overline{mn|}\;n|} = m$ and E1 is satisfied. Now suppose $f = n$,. Then $\overline{\overline{nn|}\;n|} = n$ and so E1 is again satisfied. Thus the equation, in this case, has two solutions.

It is evident, then, that, by an unlimited number of steps from a given expression e, we can reach an expression e' that is not equivalent to e.

We see, in such a case, that the theorems of representation no longer hold, since the arithmetical value of e' is not, in every possible case of a, b, uniquely determined.

Indeterminacy

We have thus introduced into e' a degree of indeterminacy in respect of its value that is not (as it was in the case of indeterminacy introduced merely by cause of using independent variables) necessarily resolved by fixing the value of each independent variable. But this does not preclude our equating such an expression with another, provided that the degree of indeterminacy shown by each expression is the same.

Degree

We define the degree of an expression as the number of distinct stable states it can be shown to have. A paradoxical (= 'oscillatory') state is defined to be 'stable' for this purpose.

It is evident that J1 and J2 hold for all equations, whatever their degree. It is thus possible to use the ordinary procedure of demonstration (outlined in Chapter 6) to verify an equation of degree > 1. But we are denied the procedure (outlined in Chapter 8) of referring to the arithmetic to confirm a demonstration of any such equation, since the excursion to infinity undertaken to produce it has denied us our former access to a complete knowledge of where we are in the form. Hence it was necessary to extract, before departing, the rule of demonstration, for this now becomes, with the rule of dominance, a guiding principle by which we can still find our way.

Imaginary state

Our loss of connexion with the arithmetic is illustrated by the following example. Let

E2 $$f_2 = \overline{\overline{f_2}\,|}\,|\,,$$

E3 $$f_3 = \overline{f_3}\,|\,.$$

Plainly, each of E2, E3 can be represented, in arithmetic, by equating either f with the same infinite expression, thus

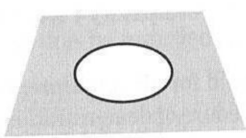
$$f_2, f_3 = \overline{\overline{\overline{\ \ldots\ }\,|}\,|}\,|\,.$$

But equally plainly, whereas E2 is open to the arithmetical solutions $\overline{\quad}|$ or   , each of which satisfies it without contradiction, E3 is satisfied by neither of these solutions, and cannot, thereby, express the same value as E2. And since $\overline{\quad}|$ and   represent the only states of the form hitherto envisaged, if we wish to pretend that E3 has a solution, we must allow it to have a solution representing an imaginary state, not hitherto envisaged, of the form.

Time

Since we do not wish, if we can avoid it, to leave the form, the state we envisage is not in space but in time. (It being possible to enter a state of time without leaving the state of space in which one is already lodged.)

One way of imagining this is to suppose that the transmission of a change of value through the space in which it is represented takes time to cover distance. Consider a cross



in a plane. An indication of the marked state is shown by the shading.

Now suppose the distinction drawn by the cross to be destroyed by a tunnel under the surface in which in appears. In Figure 1 we see the results of such destruction at intervals $t_1, t_2, \ldots$

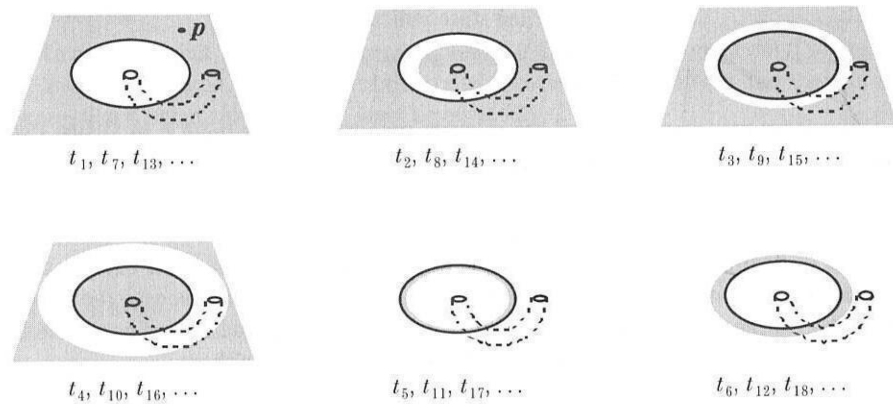


Figure 1

Frequency

If we consider the speed at which the representation of value travels through the space of the expression to be constant, then the frequency of its oscillation is determined by the length of the tunnel. Alternatively, if we consider this length to be constant, then the frequency of the oscillation is determined by the speed of its transmission through space.

Velocity

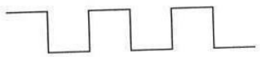
We see that once we give the transmission of an indication of value a speed, we must also give it a direction, so that it becomes a velocity. For if we did not, there would be nothing to stop the propagation proceeding as represented to t_4 (say) and then continuing towards the representation shown in t_3 instead of that shown in t_5.

Function

We shall call an expression containing a variable v alternatively a function of v. We thus see expressions of value or functions of variables, according to from which point of view we regard them.

Oscillator function

In considering the indications of value at the point p in Figure 1, we have, in time, a succession of square waves of a given frequency.



. . . marked state indicated

. . . unmarked state indicated

Suppose we now arrange for all the relevant properties of the point p in Figure 1 to appear in two successive spaces of expression, thus.

$$\overline{p}|\;p$$

We could do this by arranging similarly undermined distinctions in each space, supposing the speed of transmission to be constant throughout. In this case the superimposition of the two square waves in the outer space, one of them inverted by the cross, would add up to a continuous representation of the marked state there.

Real and imaginary value

The value represented at (or by) the point (or variable) p, being indeterminate in space, may be called imaginary in relation with the form. Nevertheless, as we see above, it is real in relation with time and can, in relation with itself, become determinate in space, and thus real in the form.

We have considered thus far a graphical representation of E3. We will now consider E1 and its limiting case E2 on similar lines.

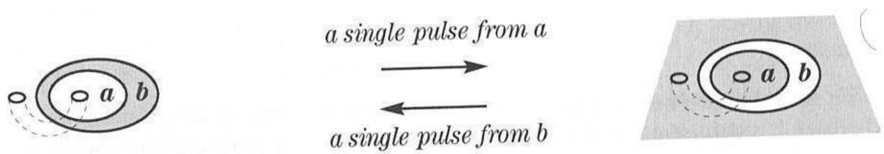


Figure 2

Memory function

The present value of the function f in E1 may depend on its past value, and thus on past values of a and b. In effect, when a, b both indicate the unmarked state, it remembers which of them last indicated the marked state. If a, then $f = m$. If b, then $f = n$

Subversion

A way to make the set-up illustrated in Figure 2 behave exactly like the f in E1, is to arrange that effective transmission through the tunnel shall be only from outside to inside. We shall call such a partial destruction of the distinctive properties of constants a subversion.

We may note that, if we wish to avail ourselves of the memory property of f, where f is an evenly subverted function, certain transformations, allowable in the case of an expression without this property, must be avoided.

We may, for example, allow

$$\overline{\overline{\overline{\overline{a}|\;\overline{b}|}|\,f}|\;c}| \rightharpoonup \overline{\overline{fa}|\;\overline{fb}|\;c}| \qquad \text{J2, C1}$$

but must avoid

$$\overline{\overline{\overline{\overline{a}|\;\overline{b}|}|\,f}|\;c}| \rightharpoonup \overline{\overline{\overline{\overline{a}|\;\overline{b}|}|\,fc}|\;c}| \qquad \text{C2}$$

since the latter transformation is from an expression by which an indication of the marked state by c can be reliably remembered, to an expression in which the memory is apparently lost.

Time in finite expressions

The introduction of time into our deliberations did not come as an arbitrary choice, but as a necessary measure to further the inquiry.

The degree of necessity of a measure adopted is the extent of its application. The measure of time, as we have introduced it here, can be seen to cover, without inconsistency, all the representative forms hitherto considered.

This can be illustrated by reconsidering E1. Here we can test the use of the concept of time by finding whether it leads to the same answer (i.e. whether it leads to the same memory of dominant states of a, b) in the expanded version of f as it does in the contracted version in Figure 2. For the purpose of illustration, we shall consider a finite expression first.

It is seen from Figure 3 that such a finite expression is stable in one condition, and has a finite memory of the other, of duration proportional to the degree of its extension. It is plain that an endless extension of the echelon allows and endless memory of either condition, so that the concept of time is a key by which the contracted and expanded forms of f in E1 are made patent to one another.

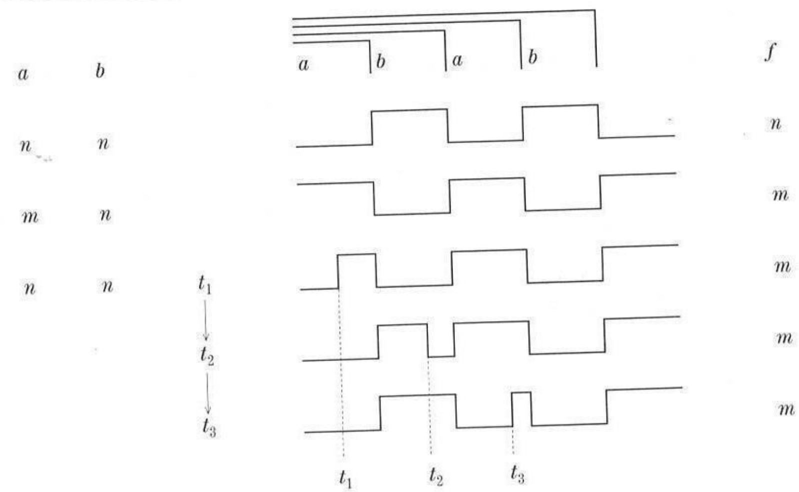


Figure 3

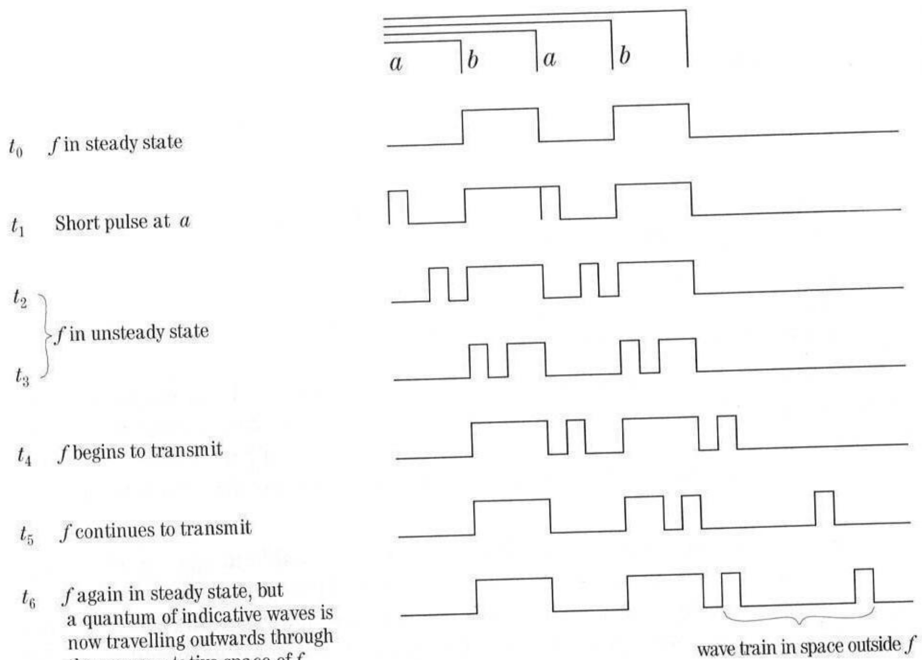


Figure 4

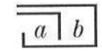
A condition of special interest emerges if the dominant pulse from a is of sufficiently short duration. In this condition the expression emits a wave train of finite length and duration, as illustrated in Figure 4.

The duration of the wave train, the frequency of its components, etc, depend on the nature and extent of the expression from which it is emitted. From an infinitely extended expression comes a potentially endless emission, and here again, the two ways (contracted or expanded) of expressing E1 in relation with time give the same answer. Without the key of time, only the contracted expression makes sense.

Crosses and markers

Consider the case where the expression in E1 represents a part of a larger expression. It now becomes necessary not only to indicate where a re-insertion takes place, but also to designate the part of the expression re-inserted. Since the whole is no longer the part re-inserted, it will be necessary in each case either to name the part re-inserted or to indicate it by direct connexion.

The latter is less cumbersome. Thus we can rewrite the expression in E1



so that it can be placed, without ambiguity, within a larger expression.

In a simple subverted expression of this kind neither of the non-literal parts are, strictly speaking, crosses, since they represent, in a sense, the same boundary. It is convenient, nevertheless, to refer to them separately, and for this purpose we call each separate non-literal part of any expression a marker. Thus a cross is a marker, but a marker need not be a cross.

Modulator* function

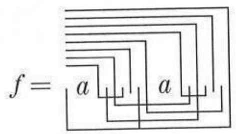
We have seen that functions of the second degree can either oscillate or remember. We can also write an equation of degree 2 that will not only remember, but count.

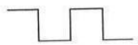
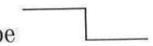
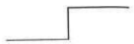
A way of picturing counting is to consider it as the contrary of remembering. A memory function remembers the same response to the same signal: a counting function counts it different each time.

Another way to picture counting is as a modulation of a wave structure. This is the way we shall picture it here.

The simplest modulation is to a wave structure of half the frequency of the original. To achieve this with a function using only real values, we can use eight markers, thus.

* In engineering applications I now call these expressions *reductors* – Author.

E4 

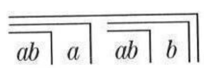
If the wave structure of a is  then that of f wil be  or , depending on how the expression is originally set before a starts to oscillate.

We are now in difficulties through attempting to write in two dimensions what is clearly represented in three. We ought to be writing in three dimensions. We can at least devise a better system of drawing three-dimensional representations in two.

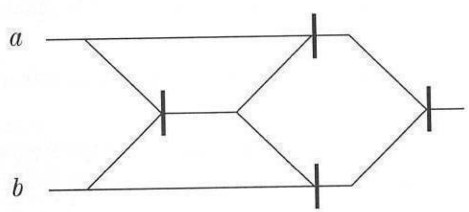
Let a marker be represented by a vertical stroke, thus.

Let what is under the marker be seen to be so by lines of connexion, called leads, thus.

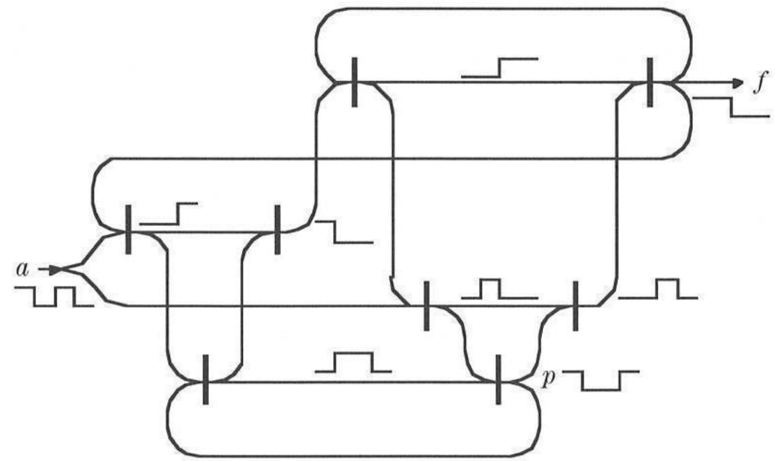
Let the value indicated by the marker be led from the marker by a lead, which may, in the expression, divide to be entered under other markers. Now, for example, the expression



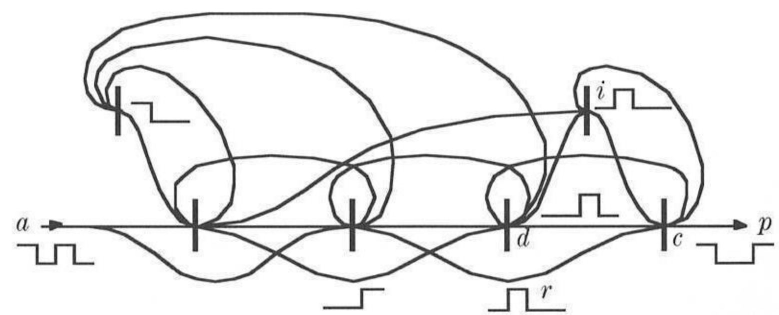
can be represented thus.



Transfigured in this way, E4 appears in a form in which it is easier to follow how the wave structure of a is taken apart and recombined to give that of f.



We see that the wave structure at p constitutes a similar modulation with the phase displaced. By using imaginary components of some wave structures, it is possible to obtain the wave structure at p with only six markers. This is illustrated in the following equation.



Here, although the real wave structure at i is identical with that at r, the imaginary* component at i ensures that the memory in markers c and d is properly set. Similar con-

* An imaginary value is one that switches itself off when it has set the memory in the required direction.

siderations apply to other memories in the expression. (Four of the leads in this expression are redundant. I did not notice this when I first wrote the chapter. It will be a nice exercise for the reader to discover which they are. – Author 2008 01 17)

Coda

At this point, before we have gone so far as to forget it, we may return to consider what it is we are deliberating.

We are, and have been all along, deliberating the form of a single construction (commanded on p 3), notably the first distinction. The whole account of our deliberations is an account of how it can appear, in the light of various states of mind that we put upon ourselves.

By the canon of expanding reference (p 10), we see that the account can be continued endlessly.

This book is not endless, so we have to break it off somewhere. We now do so here with the words

and so on.

Before departing, we return for a last look at the agreement with which the account was opened.

12

Re-entry into the form

The conception of the form lies in the desire to distinguish.

Granted this desire, we cannot escape the form, although we can see it any way we please.

The calculus of indications is a way of regarding the form.

We can see the calculus by the form and the form in the calculus unaided and unhindered by the intervention of laws, initials, theorems, or consequences.

The experiments below illustrate one of the indefinite number of possible ways of doing this.

We may note that in these experiments the sign

$$=$$

may stand for the words

is confused with.

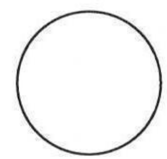
We may also note that the sides of each distinction experimentally drawn have two kinds of reference.

The first, or explicit, reference is to the value of a side, according to how it is marked.

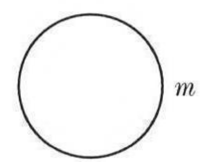
The second, or implicit, reference is to an outside observer. That is to say, the outside is the side from which a distinction is supposed to be seen.

First experiment

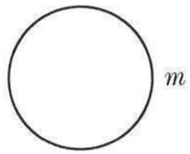
In a plane space, draw a circle.



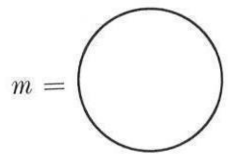
Let a mark m indicate the outside of the circumference.



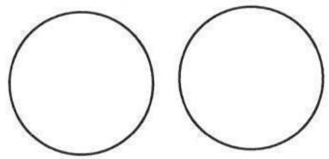
Let no mark indicate the inside of the circumference.



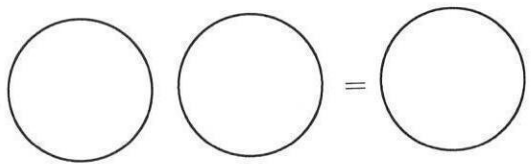
Let the mark m be a circle.



Re-enter the mark into the form of the circle.

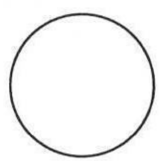


Now the circle and the mark cannot (in respect of their relevant properties) be distinguished, and so

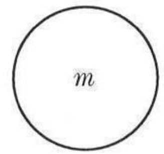


Second experiment

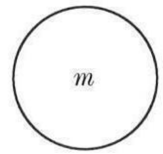
In a plane space, draw a circle.



Let a mark m indicate the inside of the circumference.



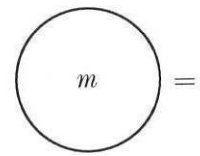
Let no mark indicate the outside of the circumference.



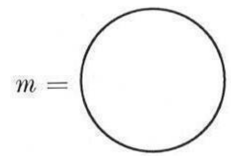
Let the value of a mark be its value to the space in which it stands. That is to say, let the value of a mark be to the space outside the mark.

Now the space outside the circumference is unmarked.

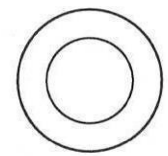
Therefore, by valuation,



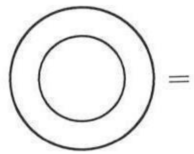
Let the mark m be a circle.



Re-enter the mark into the form of the circle.

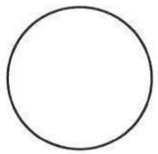


Now, by valuation,

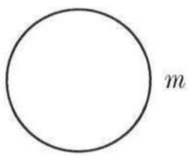


Third experiment

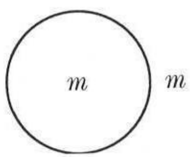
In a plane space, draw a circle.



Let a mark m indicate the outside of the circumference.

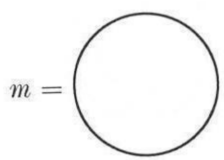


Let a similar mark m indicate the inside of the circumference.

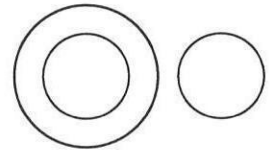


Now, since a mark m indicates both sides of the circumference, they cannot, in respect of value, be distinguished.

Again let the mark m be a circle.



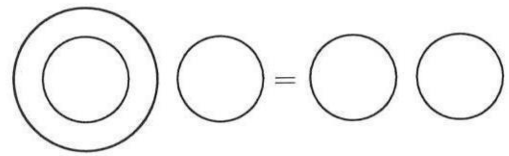
Re-enter the mark into the form of the circle.



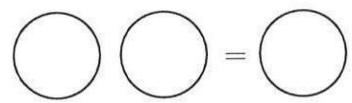
Now, because of identical markings, the original circle cannot distinguish different values.

Therefore, it is not, in this respect, a distinction.

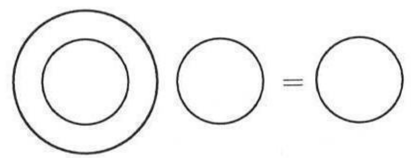
Therefore it may be deleted without loss or gain to the space in which it stands.



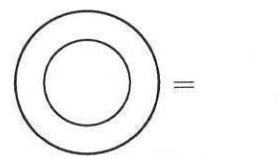
But we found in the first experiment that



Therefore,



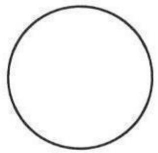
and this is not inconsistent with the finding of the second experiment that



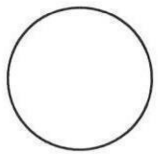
since we have done here in two steps what was done there in one.

Fourth experiment

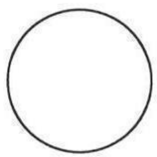
In a plane space, draw a circle.



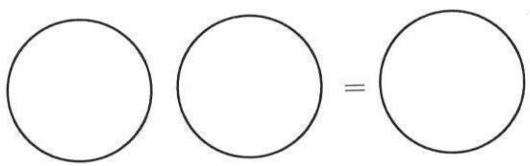
Let the outside of the circumference be unmarked.



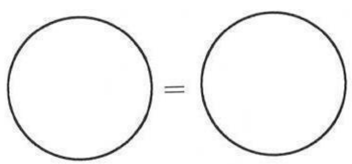
Let the inside of the circumference be unmarked.



But we saw in the first experiment that



and that therefore, by reversing the purifying procedure there,

 *m.*

The value of a circumference to the space outside must be, therefore, the value of the mark, since the mark now distinguishes this space.

An observer, since he distinguishes the space he occupies, is also a mark.

In the experiments above, imagine the circles to be forms and their circumferences to be the distinctions shaping the spaces of these forms.

In this conception a distinction drawn in any space is a mark distinguishing the space. Equally and conversely, any mark in a space draws a distinction.

We see now that the first distinction, the mark, and the observer are not only interchangeable, but, in the form, identical.

Notes

Chapter 1

Although it says somewhat more, all that the reader needs to take with him from Chapter 1 are the definition of distinction as a form of closure, and the two axioms that rest with this definition.

Cahpter 2

It may be helpful at this stage to realize that the primary form of mathematical communication is not description, but injunction. In this respect it is comparable with practical art forms like cookery, in which the taste of a cake, although literally indescribable, can be conveyed to a reader in the form of a set of injunctions called a recipe. Music is a similar art form, the composer does not even attempt to describe the set of sounds he has in mind, much less the set of feelings occasioned through them, but writes down a set of commands that, if they are obeyed by the reader, can result in a reproduction, to the reader, of the composer's original experience.

Where Wittgenstein says [4, proposition 7]:

> whereof one cannot speak
> thereof one must be silent

he seems to be considering descriptive speech only. He notes elsewhere that the mathematician, descriptively speaking, says nothing. The same may be said of the composer, who, if he were to attempt a *description* (i.e. a limitation) of the set of ecstasies apparent *through* (i.e. unlimited by) his *composition*, would fail miserably and necessarily. But neither the composer nor the mathematician must, for this reason, be silent.

In his introduction to the *Tractatus*, Russell expresses what thus seems to be a justifiable doubt in respect of the rightness of Wittgenstein's last proposition when he says [p 22],

> what causes hesitation is the fact that, after all, Mr Wittgenstein manages to say a good deal about what cannot be said, thus suggesting to the sceptical reader that possibly there may be some loophole through a hierarchy of languages, or by some other exit.

The exit, as we have seen it here, is evident in the injunctive faculty of language.

Even natural science appears to be more dependent upon injunction than we are

usually prepared to admit. The professional initiation of the man of science consists not so much in reading the proper textbooks, as in obeying injunctions such as 'look down that microscope'. But it is not out of order for men of science, having looked down the microscope, now to describe to each other, and to discuss amongst themselves, what they have seen, and to write papers and textbooks describing it. Similarly, it is not out of order for mathematicians, each having obeyed a given set of injunctions, to describe to each other, and to discuss amongst themselves, what they have seen, and to write papers and textbooks describing it. But in each case, the description is dependent upon, and secondary to, the set of injunctions having been obeyed first.

When we attempt to realize a piece of music composed by another person, we do so by *illustrating*, to ourselves, with a musical instrument of some kind, the composer's commands. Similarly, if we are to realize a piece of mathematics, we must find a way of illustrating, to ourselves, the commands of the mathematician. The normal way to do this is with some kind of scorer and a flat scorable surface, for example a finger and a tide-flattened stretch of sand, or a pencil and a piece of paper. Taking such an aid to illustration, we may now begin to carry out the commands in Chapter 2.

First we may illustrate a form, such as a circle or near-circle. A flat piece of paper, being itself illustrative of a plane surface, is a useful mathematical instrument for this purpose, since we happen to know that a circle in such a space does in fact draw a distinction. (If, for example, we had chosen to write upon the surface of a torus, the circle might not have drawn a distinction.)

When we come to the injunction

> let there be a form distinct from the form

we can illustrate it by taking a fresh piece of paper (or another stretch of sand). Now, in this separate form, we may illustrate the command

> let the mark of distinction be copied
> out of the form into such another form.

It is not necessary for the reader to confine his illustrations to the commands in the text. He may wander at will, inventing his own illustrations, either consistent or inconsistent with the textual commands. Only thus, by his own explorations, will he come to see distinctly the bounds or laws of the world from which the mathematician is speaking. Similarly, if the reader does not follow the argument at any point, it is never necessary for him to remain stuck at that point until he sees how to proceed. We cannot fully understand the beginning of anything until we see the end. What the mathematician aims to do is to give a complete picture, the order *of what* he presents being essential, the order *in which* he presents it being to some degree arbitrary. The reader may quite legitimately change the arbitrary order as he pleases.

We may distinguish, in the essential order, *commands*, that call something into being, conjure up some order of being, call to order, and that are usually carried in permissive forms such as

let there be so-and-so

or occasionally in more specifically active forms like

drop a perpendicular;

names, given to be used as reference points or tokens; in relation with the operation of *instructions*, which are designed to take effect within whatever universe has already been commanded or called to order. The institution or ceremony of naming is usually carried in the form

call so-and-so such-and-such,

and the call may be transmitted in both directions, as with the sign =, so that by calling so-and-so such-and-such we may also call such-and-such so-and-so. Naming may thus be considered to be without direction, or, alternatively, pan-directional. By contrast, instruction is directional, in that it demands a crossing from a state or condition, with its own name, to a different state or condition, with another name, such that the name of the former may not be called as a name of the latter.

The more important structures of command are sometimes called canons. They are the ways in which the guiding injunctions appear to group themselves in constellations, and are thus by no means independent of each other. A canon bears the distinction of being outside (i.e. describing) the system under construction, but a command to construct (e.g. 'draw a distinction'), even though it may be of central importance, is not a canon. A canon is an order, or set of orders, to permit or allow, but not to construct or create.

The instructions that are to take effect, within the creation and its permission, must be distinguished as those in the actual text of calculation, designated by the constants or operators of the calculus, and those in the context, that may themselves be instructions to name something with a particular name so that it can be referred to again without redescription.

Later on (Chapter 4) we shall come to consider what we call the proofs or justifications of certain statements. What we shall be showing, here, is that such statements are implicit in, or follow from, or are permitted by, the canons or standing orders hitherto convened or called to presence. Thus, in the structure of a proof, we shall find injunctions of the form

consider such-and-such,
suppose so-and-so,

that are not commands, but *invitations* or *directions* to a way in which the implication can be clearly and wholly followed.

In conceiving the calculus of indications, we begin at a point of such degeneracy as to find that the ideas of description, indication, name, and instruction can amount to the same thing. It is of some importance for the reader to realize this for himself, or he will find it difficult to understand (although he may follow) the argument (p 5) leading to the second primitive equation.

In the command

let the crossing be to the
state indicated by the token

we at once make the token doubly meaningful, first as an instruction to cross, secondly as an indicator (and thus a name) of where the crossing has taken us. It was an open question, before obeying this command, whether the token would carry an indication at all. But the command determines without ambiguity the state to which the crossing is made and thus, without ambiguity, the indication that the token will henceforth carry.

This double carry of name-with-instruction and instruction-with-name is usually referred to (in the language of mathematics) as a structure in which ideas or meanings *degenerate*. We may also refer to it (in the language of psychology) as a place where the ideas *condense* in one symbol. It is this condensation that gives the symbol its power. For in mathematics, as in other disciplines, the power of a system resides in its elegance (literally, its capacity to pick out or elect), which is achieved by condensing as much as is needed into as little as is needed, and so making that little as free from irrelevance (or from elaboration) as is allowed by the necessity of writing it out and reading it in with ease and without error.

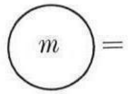
We may now helpfully distinguish between an elegance in the calculus, that can make it easy to use, and an elegance in the descriptive context, that can make it hard to follow. We are accustomed, in ordinary life, to having indications of what to do confirmed in several different ways, and when presented with an injunction, however clear and unambiguous, which, stripped to its bare minimum, indicates what to do once and in one way only, we might refuse it. (We may consider how far, in ordinary life, we must observe the spirit rather than the letter of an injunction, and must develop the habitual capacity to interpret any injunction we receive by screening it against other indications of what we ought to do. In mathematics we have to unlearn this habit in favour of accepting an injunction literally and at once. This is why an author of mathematics must take such great pains to make his injunctions mutually permissive. Otherwise these pains, that rightly rest with the author, will fall with sickening import upon the reader, who, by virtue of his relationship with respect to the author, may be in no position to bear them.)

The second of the two primitive equations of the primary arithmetic can be derived less elegantly, but in a way that is possibly easier to follow, by allowing substitution prematurely.

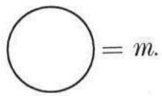
Suppose we indicate the marked state by a token m, and, as before, let the absence of a token indicate the unmarked state.

Let a bracket round any indicator indicate, in the space outside the bracket, the state other than that indicated inside the bracket.

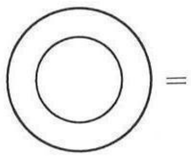
Thus



and



Substituting, we find

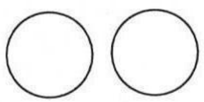


and this is the second prmitive equation.

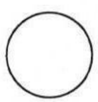
The condition that one of the primary states shall be nameless is mandatory for this elimination.

The first primitive equation can also be derived a different way.

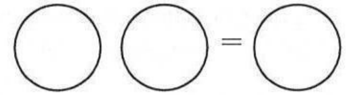
Imagine a blind animal able only to distinguish inside from outside. A space with what appears to us as a number of distinct insides and one outside, such as



will appear to it, upon exploration, to be indistinguishable from



The ideas described in the text at this point do not go beyond what this animal can find out for itself, and so in its world, such as it is,



We may note that even if this animal can count its crossings, it still will not be able to distinguish two divisions from one, although it will now have an alternative way of distinguishing inside from outside that no longer depends on knowing which is which.

Reconsidering the first command,

draw a distinction,

we note that it may equally well be expressed in such ways as

let there be a distinction,
find a distinction,
see a distinction,
describe a distinction,
define a distinction,

or

let a distinction be drawn,

for we have here reached a place so primitive that active and passive, as well as a number of other more peripheral opposites, have long since condensed together, and almost any form of words will suggest more categories than there really are.

Chapter 3

The hypothesis of simplification is the first *overt* convention that is put to use before it has been justified. But it has a precursor in the injunction 'let a state indicated by an expression be the value of the expression' in the last chapter, which allows value to an expression only in case not less and not more than one state is indicated by the expression. The use of both the injunction and the convention are eventually justified in the theorems of representation. Other cases of delayed justification will be found later, a notable example being theorem 16.

We may ask why we do not justify such a convention at once when it is given. The answer, in most cases, is that justification (although valid) would be meaningless until we had first become acquainted with the *use* of the principle that requires justifying. In other words, before we can reasonably justify a deep-lying principle, we first need to be familiar with how it works.

We might suppose this practice of deferred justification to be operative elsewhere. It is a notable fact that in mathematics very few *useful* theorems remain unproved. By 'useful' I do not necessarily mean with practical application outside mathematics. A theorem can be useful mathematically, for example to justify another theorem.

One of the most 'useless' theorems in mathematics is Goldbach's conjecture. We do not frequently find ourselves saying 'if only we knew that every even number greater than 2 could be represented as a sum of two prime numbers, we should be able to show that . . .' D J Spencer Brown, in a private communication, suggested that their apparent uselessness is not exactly a reason why such theorems cannot be proved, but is a reason for supposing that if a valid proof were given today, nobody would recognize it as such, since nobody is yet *familiar* with the *ground* on which such a proof would rest. I shall have more to say about this in the notes to Chapters 8 and 11.

Chapter 4

In all mathematics it becomes apparent, at some stage, that we have for some time been following a rule without being consciously aware of the fact. This might be described as the use of a *covert* convention. A recognizable aspect of the advancement of mathematics consists in the advancement of the consciousness of what we are doing, whereby the covert becomes overt.

The nearer we are to the beginning of what we set out to achieve, the more likely we are to find, there, procedures that have been adopted without comment. Their use can be considered as the presence of an arrangement in the absence of an agreement. For example, in the statement and proof of theorem 1 it is arranged (although not agreed) that we shall write on a plane surface. If we write on the surface of a torus the theorem is not true. (Or to make it true, we must be more explicit.)

The fact that men have for centuries used a plane surface for writing means that, at this point in the text, both author and reader are ready to be conned into the assumption of a plane writing surface without question. But, like any other assumption, it is not unquestionable, and the fact that we can question it here means that we can question it elsewhere. In fact we have found a common but hitherto unspoken assumption underlying what is written in mathematics, notably a plane surface (more generally, a surface of genus 0, although we shall see later (pp 102 sq) that this further generalization forces us to recognize another hitherto silent assumption). Moreover, it is now evident that if a different surface is used, what is written on it, although identical in marking, may be not identical in meaning.

In general there is an order of precedence amongst theorems, so that theorems that can be proved more easily with the help of other theorems are placed so as to be proved after such other theorems. This order is not rigid. For example, having proved theorem 3, we use what we found in the proof to prove theorem 4. But theorems 3 and 4 are

symmetrical, their order depending only on whether we wish to proceed from simplicity to complexity or from complexity to simplicity. The reader might try, if he wishes, to prove theorem 4 first without the aid of theorem 3, after which he will be able to prove theorem 3 analogously to the way theorem 4 is proved in the text.

It will be observed that the literal representation of theorem 8 is less strong than the theorem itself. The theorem is consistent with

$$\overline{\overline{p\,|}\;pq\,|} = \qquad ,$$

whereas we prove the weaker version

$$\overline{\overline{\mathrm{p}\,|}\;\mathrm{p}\,|} = \qquad .$$

The stronger version is plainly true, but we shall find that we are able to demonstrate it as a consequence in the algebra. We therefore prove, and use as the first algebraic initial, the weaker version.

In theorem 9 we see the difference between our use of the verb *divide* and our use of the verb *cleave*. Any division of a space results in *otherwise indistinguishable divisions of a state,* that are all at the same level, whereas a severance or cleavage shapes *distinguishable states*, that are at different levels.

An idea of the relative strenghts of severance and division may be gathered from the fact that the rule of number is sufficient to unify a divided space, but not to void a cloven space.

Chapter 5

In eliciting rules for algebraic manipulation the text explicitly refers to the existence of systems of calculation other than the system described. This reference is both deliberate and inessential. It marks the level at which these systems are usually fitted out with their false, or truncated, or postulated, origins.

It is deliberate to inform the reader that, in the system of calculation we are building, we are not departing from the basic methods of other systems. Thus what we arrive at, in the end, will serve to elucidate them, as well as to fit them with their true origin. But, at the same time, it is important for the reader to see that the reference to other systems is inessential to the development of the argument in the text. For here it stands or falls on its own merit, dependent in no way for its validity upon agreement or disagreement with other systems. Thus rules 1 and 2, as can be seen from their justifications, say nothing that has not, in the text, already been said. They merely summarize the commands and instructions that will be relevant to the new kind of calculation we are about to undertake.

The replacement referred to in rule 2 is usually confined to independent variable expressions of simple (i.e. literal) form, and is in fact so confined in the text. But the greater licence granted by the rule is not devoid of significant application, if required.

Chapter 6

By the revelation and incorporation of its own origin, the primary algebra provides immediate access to the nature of the relationship between operators and operands. An operand in the algebra is merely a conjectured presence or absence of an operator.

This partial identity of operand and operator, which is not confined to boolean algebras, can in fact be seen if we extend more familiar descriptions, although in these descriptions it is not so obvious. For example, we can find it by taking the boolean operators $\vee$ (usually interpreted as the logical 'or', but here used purely mathematically) and $\wedge$ (usually interpreted as the logical 'and', but here again used purely mathematically), freeing their scope (as, by the principle of relevance, we may), freeing the order of the variables within their scope (as, by the same principle, we also may), and extrapolating mathematically to the case of no variable,

...	$(a\,b\,c)$ $\vee\wedge$	$(a\,b)$ $\vee\wedge$	(a) $\vee\wedge$	$()$ $\vee\wedge$
permute	111 11	11 11	1 11	01
permute	110 10	10 10	0 00	
permute	100 10	00 00		
permute	000 00			

which shows quite plainly that we have no need of the arithmetical forms 0, 1 (or z, u, or F, T, etc), since we can equate them with $()\vee$ and $()\wedge$ respectively. We can now write a boolean variable of the form a, b, etc wherever we conjecture the presence of one of these two fundamental particles, but are not sure (or don't care) which. The functional tables for $\vee$ and $\wedge$ of two variables thus become

(a b)	$\vee$	$\wedge$
($()\vee$ $()\vee$)	$()\vee$	$()\vee$
($()\vee$ $()\wedge$)	$()\wedge$	$()\vee$
($()\wedge$ $()\wedge$)	$()\wedge$	$()\wedge$,

the permutation being assumed.

J1, J2 are not the only two initials that may be taken to determine the primary algebra. We see[10] from Huntington's fourth postulate-set that we could have used C5, C6.*

The demonstration of J1, J2 from C5, C6 is both difficult and tedious. This is evidently

10 Edward V Huntington, *Trans. Amer. Math. Soc.*, 35 (1933) 280-5

* It is now known that Huntington's postulates are not independent, so we can derive the primary algebra from C6 alone – Author

because we find two basic algebraic principles, in one of which a variable is transplanted in the expression, and in the other of which it is eliminated from it. Provided we keep these two principles apart, subsequent demonstrations are not difficult. If, as in Huntington's two equations, they are inter-mingled, then their subsequent unravelling can be difficult.

Our expression here of Huntington's equations in the form of C5, C6 is not in the form in which he originally expressed them. He was hampered by the crippling assumptions of order relevance and binary scope, with which we have not at any stage weakened the primary algebra. For this reason he found it necessary to give two more equations to complete the set. C5 and C6, considered as initials, are of interest chiefly because they employ only two distinct variables, whereas J1 and J2 employ three.

I had at first supposed the demonstration of C1 to be impossible from J1 and J2 as they stand. In 1965 a pupil, Mr John Dawes, produced a rather long proof to the contrary, so the following year I set the problem to my class as an exercise, and was rewarded with a most elegant demonstration by Mr D A Utting. I use Mr Utting's demonstration, slightly modified, in the text.

Although, superficially, it may look less efficient, it is, eventually, more natural and convenient to use names rather than numbers to identify the more important consequences, as indeed it is with theorems, since they do not in general form an ordered set.

In naming such consequences I have aimed to find what seems appropriate as a description of the named process, as it appears in the algebra, without doing violence to its arithmetical origin. In some places both the forms and the names are recognizably similar to those of other authors who have determined boolean algebras. In most such cases hitherto, the commonly used name describes only one of the directions in which the step can be taken. What is called boolean expansion is an example. In such a case, where the name is appropriate to the step as taken in one direction only, I have introduced an antonym for the other direction, and given a generic name to cover both. In other recognizable cases I have found what seems to me to be a more appropriate name, such as occultation for what Whitehead called[11] absorption. The occulting part of the expression is not so much absorbed in the remainder as eclipsed by it. This can be seen quite plainly in the arithmetic, or alternatively if the expression is illustrated with a Venn diagram. To the best of my knowledge, Peirce was the only previous author to recognize, as such, what I call position. He called[12] it erasure, thus again drawing attention to only one direction of application.

I do not suppose all the names will always stick. Familiarity tends to produce a kind of in-slang, often more appropriate, in its place, than what is deemed to be academically

11 Alfred North Whitehead, *A treatise on universal Algebra*, Vol. I, Cambridge, 1898, p 36.

12 Charles Sanders Peirce, *Collected papers*, Vol. IV, Cambridge, Massachusetts, 1933, pp 13-18.

proper or seemly. For example, the engineering application of consequence 2 has produced the more homely 'breed' for 'regenerate', and 'revert' for 'degenerate', and it is of interest to note that the transformations of this consequence are immediate images of what Proclus called[13] *πρόοδος* and *ἐπιστροφή*, tranalated by Dodds into *procession* and *reversion*.

The fact that descriptive names such as 'transposition' and 'integration' are differently applied elsewhere in mathematics (and, indeed, elsewhere in this book) does not appear to be a reason for avoiding their use in the senses defined in this chapter. The deeper the level of investigation, the harder it becomes to find words strong enough to cover what is found there, and in all cases my use of language to describe primitive processes draws on a greater power of signification than is needed for its more superficial and specialized uses.

One of the most beautiful facts emerging from mathematical studies is this very potent relationship between the mathematical process and ordinary language. There seems to be no mathematical idea of any importance or profundity that is not mirrored, with an almost uncanny accuracy, in the common use of words, and this appears especially true when we consider words in their original, and sometimes long forgotten, senses.

The fact that a word may have different, but related, meanings at different, but related, levels of consideration does not normally render communication impossible. On the contrary, it is evident that communication of any but the most trivial ideas would be impossible without it.

Since at this point in the text the fundamental forms of mathematical communication are now practically complete, it may be a revealing exercise to retranslate into longhand some of the shorthand forms developed by application of the canon of contracting reference. For this purpose we take the statement and demonstration of consequence 9 (p 28 and 29). In words and figures it could run thus.

The ninth consequence, called crosstransposition, or C9 for short, may be stated as follows.

b cross *r* cross cross all *a*
cross *r* cross cross 2 *x* cross
r cross 2 *y* cross *r* cross 2
cross all

expresses the same value as

r cross *ab* cross all *rxy* cross 3.

When the step allowed by this equation is taken from the former to the latter expression, it is called to crosstranspose or collect, and when taken in reverse it is called

13 *ΠΡΟΚΛΟΥ ΔΙΑΔΟΧζΟΥ ΣΤΟΙΧΕΙΩΣΙΣ ΘΕΟΛΟΓΙΚΗ* with a translation by E R Dodds, 2nd edition, Oxford, 1963.

to crosstranspose or distribute.

The equation can be demonstrated thus.

b cross *r* cross cross all *a*
cross *r* cross cross 2 *x* cross
r cross 2 *y* cross *r* cross 2
cross all

may be changed to

b cross *r* cross cross all *a*
cross *r* cross cross 2 *xy*
cross 2 *r* cross 2 cross all

by using C1, J2, and then C1 again. This in turn may be changed to

baxy cross 2 *r* cross 2 cross
all *rxy* cross 2 *r* cross 2
cross 2

by C8 and then by applying C1 three times, etc.

We may observe that, in expressions, the mathematical language has become entirely visual, there is no proper spoken form, so that in reverbalizing it we must *encode* it in a form suitable for ordinary speech. Thus, although the mathematical form of an expression is clear, the reverbalized form is obscure.

The main difficulty in translating from the written to the spoken form comes from the fact that in mathematical writing we are free to mark the two dimensions of the plane, whereas in speech we can mark only the one dimension of time.

Much that is unnecessary and obstructive in mathematics today appears to be vestigial of this limitation of the spoken word. For example, in ordinary speech, to avoid direct reference to a plurality of dimensions, we have to fix the scope of constants such as 'and' and 'or', and this we can most conveniently do at the level of the first plural number. But to carry the fixation over into the written form is to fail to realize the freedom offered by an added dimension. This in turn can lead us to suppose that the binary scope of operators assumed for the convenience of representing them in one dimension is something of relevance to the actual form of their operation, which, in the case of these simple operators, it is not.*

* For this reason it is more-correct to describe the algebra in this text as *brownian* rather than 'boolean', since boolean and all previous algebras were constrained by the crippling limitations of order-relevance and binary scope. – Author, 2000 06 27

Chapter 7

In the description of theorem 14 'the constant' refers to the operative constant. There are two constants in the calculus, a mark or operator, and a blank or void. Reference to 'the constant' without qualification will usually be taken to denote the operator rather than the void.

Chapter 8

We have already distinguished, in the text, between demonstration and proof. In making this distinction, that appears quite natural, we see at once that a proof can never be justified in the same way as a demonstration. Whereas in a demonstration we can see that the instructions already recorded are properly obeyed, we cannot avail ourselves of this procedure in the case of a proof.

In a proof we are dealing in terms that are outside of the calculus, and thus not amenable to its instructions. In any attempt to render such proofs themselves subject to instruction, we succeed only at the cost of making another calculus, inside of which the original calculus is cradled, and outside of which we shall again see forms that are amenable to proof but not demonstration.

The validity of a proof thus rests not in our common motivation by a set of instructions, but in our common experience of a state of affairs. This experience usually includes the ability to reason that has been formalized in logic, but is not confined to it. Nearly all proofs, whether about a system containing numbers or not, use the common ability to compute, i.e. to count in either direction, and ideas stemming from our experience of this ability.

It seems open to question why we regard the proof of a theorem as amounting to the same degree of certainty as the demonstration of a consequence. It is not a question that, at first sight, admits of an easy answer. If an answer is possible, it would seem to lie in the concept of experience. We gain experience of living representative processes, in particular of argument and of counting forwards and backwards in units, and through this experience become quite certain, in our own minds, of the validity of using it to substantiate a proof. But since the procedures of the proof are not, themselves, yet codified in a calculus (although they may eventually become so), our certainty at this stage must be deemed to be intuitive. We can achieve a demonstration simply by following instructions, although we may be unfamiliar with the system in which the instructions are obeyed.* But in proving a theorem, if we have not already codified the structure of

* I don't know why I am making such a meal of this. A theorem is a statement in ordinary language, therefore cannot be an element in the calculus. A consequence is the result of a calculation, therefore must be an element in the calculus.
– Author 2007 09 17 0700

the proof *in* the form of a calculus, we must at least be familiar with, or experienced in, whatever it is we take to be the *ground* of the proof, otherwise we shall not *see* it *as* a proof.

Another way of regarding the relationship between demonstration and proof, that adds support to the proposition that the degree of certainty of a proof is equal to that of a demonstration, is to consider it as the boundary dividing the state of proof from the state of demonstration. A demonstration, we remember, occurs inside the calculus, a proof outside. The boundary between them is thus a shared boundary, and is what is approached, in one or the other direction, according to whether we are demonstrating a consequence or proving a theorem. Thus consequences and theorems can be seen to bear to each other a fitting relationship.

But the boundary marking their relationship, although shared, is (like the existential boundary (see pp 124 sq)) seen from one side only, since if we know the ground on which a demonstration rests (i.e. provided we understand the formal, as distinct from the pragmatic, reasons for the initial equations we employ and so do not have to postulate them), the demonstration can be seen as a proof by implication, although a proof is never seen as a demonstration. We observe, in fact, that demonstration bears the same relationship to proof as initial equation bears to axiom, but we should also note that the relationship is evident for arithmetic only, and is lost when we make the departure into algebra. This appears to be why algebras are commonly presented without axioms, in any proper sense of the word.

The fact that a proof is a way of making apparently obvious what was already latently so is of some mathematical interest. Although there are any number of distinct proofs of a given theorem, they can all, even so, be hard to find. In other words, we can set about trying to prove a theorem in a large number of wrong ways before coming across a right way.

Even the analogy of seeking something cannot, in this context, be quite right. For what we find, eventually, is something we have known, and may well have been consciously aware of, all along. Thus we are not, in this sense, seeking something that has ever been hidden. The idea of performing a search can be unhelpful, or even positively obstructive, since searches are in general organized to find something that has been previously hidden, and is thus not open to view.

In discovering a proof, we must do something more subtle than search. We must come to see the *relevance*, in respect of whatever statement it is we wish to justify, of some fact in full view, and of which, therefore, we are already constantly aware. Whereas we may know how to undertake a search for something we can *not* see, the subtlety of the technique of trying to 'find' something that we already can see may more easily escape our efforts.

This might be a helpful moment to introduce a distinction between following a course of argument and understanding it. I take understanding to be the experience of what is understood in a wider context. In this sense, we do not fully understand a theorem until we are able to contain it in a more general theorem. We can nevertheless follow its proof, in the sense of coming to see its evidence, without understanding it in the wider sense in which it may rest.

Following and understanding, like demonstrating and proving, are sometimes wrongly taken as synonymous. Very often a person is regarded as not understanding an argument, a process, a doctrine, when all that is certain is that he has not followed it. But his failure to follow may be quite deliberate, and may arise from the fact that he has understood what was presented to him, and does not follow it because he sees a shorter, or otherwise more acceptable, path, although he might not, yet, know how to communicate it.

Following may thus be associated particularly with doctrine, and doctrine demands an adherence to a particular way of saying or doing something. Understanding has to do with the fact that what ever is said or done can always be said or done a different way, and yet all ways remain the same.

Chapter 9

We observe that the idea of completeness cannot apply to a calculus as a whole, but only to a representation of one determination of it by another. What is questioned, in fact, is the completeness of an alternative form of expression.

The paragon of such an alternative is the algebraic representation of an arithmetic, although we do in fact find a more central case of it in the arithmetical representation of a form. In the latter case, as we see from the theorems of representation, the idea of completeness condenses with that of consistency. In the less central case, the two ideas come apart. Thus the most primitive example of completeness, in its pure form, is to be found in algebraic representation.

A fact to which Gödel drew attention [5] is that an algebra that includes representations of addition and multiplication *cannot* fully account for an arithmetic of the natural numbers in which these operations are taken as elementary. Thus, in number theory, although certain relationships can be proved, no algebra can be constructed in which all such relationships are demonstrable.

The advent of Gödel's theorem has never seemed to me to be a reason for despair, as some investigators have taken it to be, but rather an occasion for celebration, since it confirms what men of mathematics have found from experience, notably that ordinary arithmetic is a richer ground for investigation than ordinary algebra.*

* Gödel was himself confused about what he had shown. He gave the impression that some true statements cannot be proved. This is not so. All true statements can be proved, but in certain algebras some of them cannot be demonstrated.

Chapter 10

It is usual to prove the independence of initial equations indirectly.[14] It is not commonly observed, although it becomes evident when consider it, that with a set of only two initials, a direct proof of their independence is always available, and I give such a proof in the text.

An independence proof may be properly considered as an incompleteness proof of the calculus with the missing initial.

Chapter 11

The question of whether or not functions of themselves are allowable has been discussed at wearisome length by many authorities [cf 8] since *Principia mathematica* was published. The Whitehead-Russell argument for disallowing them is well known. It is the subject of a number of comments by Wittgenstein [4, propositions 5.241 sq]. (I use the Pears-McGuiness translation for what follows.)

An operation, says Wittgenstein, is not the mark of a form, but of a relation between forms. Wittgenstein here sees what I call the mark of distinction between states, which he calls forms, and also sees its connexion with the idea of operation. He then remarks [5.251] that

> A function cannot be its own argument, whereas an
> operation can take one of its own results as its base.

This applies only, in the strict sense, to single-valued functions. If we allow inverse and implicit functions, then the assertion above is untrue. A function of a variable, in the wider meaning with which it is defined in this chapter, is the result of a possible set of operations on the variable. Thus if an operation can take its own result as a base, the function determined by this operation can be its own argument

I shall proceed, in the light of this relaxation, to examine in some detail the analogy between boolean equations and those of an ordinary numerical algebra.

Boole maintained[15] that the equation with which he defined what he called the law of duality, notably

$$x^2 = x,$$

is of the second degree. So it is, as stated, but by it he determines that, in his notation, all equations of degree >1 shall be reduced to the first degree. In other words, it is an equation of the second degree only at the descriptive level, not in the algebra itself.

The spuriousness of its alleged degree, considered in the algebra itself, is revealed by

14 following Edward V Huntington, *Trans. Amer. Math. Soc.*, 5 (1904) 288-309.
15 George Boole, *An investigation of the laws of thought*, Cambridge, 1854, p 50.

Boole's assertion in a footnote [p 50] that an equation of the third degree has no interpretation in his algebra. It has, as we shall presently see, but Boole appears at this point to have been overcome by his notation, which uses numerical forms for an algebra that is essentially non-numerical.

Boole's equation

$$x^2 = x$$

is an analogue, in the primary algebra, of

$$aa = a.$$

This, as we see, is an equation of the first degree, being expressible without subversion. The real form of the analogy with a numerical algebra may be illustrated as follows.

Suppose

$$px^2 + qx + r = 0$$

where p, q, r may stand for rational numbers. We can re-write this equation in the form

F1 $$x^2 + ax + b = 0$$

by calling $q/p = a$ and $r/p = b$ and it may then be further transposed into

$$x = -a + \frac{-b}{x}$$

F2

$$= -a + \cfrac{-b}{-a + \cfrac{-b}{-a + \cfrac{-b}{\cdots}}}$$

In a boolean algebra we are properly denied the mode of F1, but permitted the mode of F2, which is either continuous or, if we wish to see it so, subversive. Thus an equation of any degree is both constructible and meaningful in a boolean algebra, although not necessarily in the primary form of it. To reach a higher degree, all we need to do is to add another stable state. The two modulator equations at the end of the chapter are both of degree 2. They were first developed in 1961, in collaboration with Mr D J Spencer-Brown, for special-purpose computer circuits.

The circuits represented by these equations, the latter being presently in use by British Railways, comprise, as far as we know, a first application of each of two inventions, notably the first construction of a device that counts entirely by 'logic' (i.e. with switches only, and with no artificial time delays such as electrical condensers) and, in addition, the first use, in a switching circuit, of imaginary boolean values in the course of the construction of a real answer. (An imaginary value is one that the desired output switches off.)

The fact that imaginary values *can* be used to reason towards a real and certain answer, coupled with the fact that they *are not* so used in mathematical reasoning today, and also coupled with the fact that certain equations plainly *cannot* be solved without the use of imaginary values, means that *there* **must** *be mathematical statements* (whose truth or untruth is in fact perfectly decidable) *that cannot be decided by the methods of reasoning to which we have hitherto restricted ourselves.**

Generally speaking, if we confine our reasoning to an interpretation of boolean equations of the first degree only, we should expect to find theorems that will always defy decision, and the fact that we do seem to find such theorems in common arithmetic may serve, here, as a practical confirmation of this obvious prediction. To confirm it theoretically, we need only to prove (1) that such theorems *cannot* be decided by reasoning of the first degree, and (2) that they *can* be decided by reasoning of a higher degree. (2) would of course be proved by providing such a proof of one of these theorems.

I may say that I believe that at least one such theorem will shortly be decided by the methods outlined in the text. In other words, I believe that I have reduced their decision to a technical problem that is well within the capacity of an ordinary mathematician who is prepared, and who has the patronage or other means, to undertake the labour.

Any evenly subverted equation of the second degree might be called, alternatively, evenly informed. We can see it over a sub-version (turning under) of the surface upon which it is written, or alternatively, as an in-formation (formation within) of what it expresses.

Such an expression is thus informed in the sense of having its own form within it, and at the same time informed in the sense of remembering what has happened to it in the past.

We need not suppose that this is exactly how memory happens in an animal, but there are certainly memories, so-called, constructed this way in electronic computers, and engineers have constructed such in-formed memories with magnetic relays for the greater part of the present century.

We may perhaps look upon such memory, in this simplified in-formation, as a precursor of the more complicated and varied forms of memory and information in man and the higher animals. We can also regard other manifestions of the classical forms of physical or biological science in the same spirit.

Thus we do not imagine the wave train emitted by an excited finite echelon to be exactly like the wave train emitted from an excited physical particle. For one thing the wave form from an echelon is square, and for another it is emitted without energy. (We should need, I guess, to make at least one more departure from the form before arriving at a

* The reader may refer to Appendix 8, where I use an argument of degree 2 to prove that there are primes between squares.

conception of energy on these lines.) What we see in the forms of expression at this stage, although recognizable, might be considered as simplified precursors of what we take, in physical science, to be the real thing. Even so, their accuracy and coverage is striking. For example, if, instead of considering the wave train emitted by the expression in Figure 4, we consider the expression itself, in its quiescent state, we see that it is composed of standing waves. If, therefore, we shoot such an expression through its own representative space, it will, upon passing a given point, be observable at that point as a simple oscillation with a frequency proportional to the velocity of its passage. We have thus already arrived, even at this stage, at a remarkable and striking precursor of the wave properties of material particles.

We may look upon such manifestations as the formal seeds, the existential forerunners, of what must, in a less central state, under less certain conditions, come about. There is a tendency, especially today, to regard existence as the source of reality, and thus as a central concept. But as soon as it is formally examined (cf Appendix 2), existence* is seen to be highly peripheral and, as such, especially corrupt (in the formal sense) and vulnerable. The concept of truth is more central, although still recognizably peripheral. If the weakness of present-day science is that it centres round existence, the weakness of present-day logic is that it centres round truth.

Throughout the essay, we find no need of the concept of truth, apart from two avoidable appearances (true = open to proof) in the descriptive context. At no point, to say the least, is it a necessary inhabitant of the calculating forms. These forms are thus not only precursors of existence, they are also precursors of truth.

It is, I would say, the intellectual block that most of us come up against at the points where, to experience the world clearly, we must abandon existence to truth, truth to indication, indication to form, and form to void, that has so held up the development of logic and its mathematics.

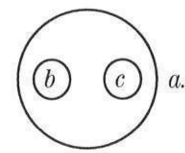
What status, then, does logic bear in relation with mathematics? We may anticipate, for a moment, Appendix 2, from which we see that the arguments we used to justify the calculating forms (e.g. in the proofs of theorems) *can themselves be justified by putting them in the form of the calculus.* The process of justification can be thus seen to feed upon itself, and this may comprise the strongest reason against believing that the codification of a proof procedure lends evidential support to the proofs in it. All it does is provide them with coherence. A theorem is no more proved by logic and computation than a sonnet is written by grammar and rhetoric, or than a sonata is composed by harmony and counterpoint, or a picture painted by balance and perspective. Logic and computation, grammar and rhetoric, harmony and counterpoint, balance and perspective,

* *ex* = out, *stare* = stand.

can be seen in the work *after* it is created, but these forms are, in the final analysis, parasitic on, they have no existence apart form, the creativity of the work itself. Thus the relation of logic to mathematics is seen to be that of an applied science to its pure ground, and all applied science is seen as drawing sustenance from a process of creation with which it can combine to give structure, but that it cannot appropriate.

Chapter 12

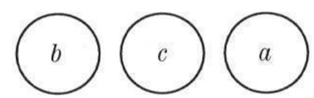
Let us imagine that, instead of writing on a plane surface, we are writing on the surface of the Earth. Ignoring rabbit holes, etc, we may take it to be a surface of genus 0. Suppose we write



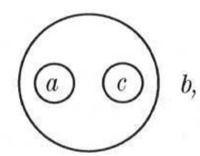
To make it readable from another planet, we write it large. Suppose we draw the outer bracket round the Equator, and make the brackets containing *b* and *c* follow the coastlines of Australia and the South Island of New Zealand respectively.

Above is how the expression will appear from somewhere in the Northern Hemisphere, say London. But let us travel.

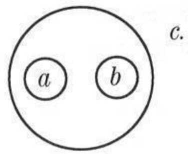
Arriving at Capetown we see



Sailing on to Melbourne, we see



and proceeding from there to Christchurch, we see



These four expressions are distinct and not equivalent. Thus it is evidently not enough merely to write down an expression, even on a surface of genus 0, and expect it to be understood. We must also indicate where the observer is supposed to be standing in relation to the expression. Writing on a plane, the ambiguity is not apparent because we tend to see the expression from outside of the outermost bracket. When it is written on the surface of a sphere, there may be no means of telling which of the brackets is supposed to be outermost. In such a case, to make an expression meaningful, we must add to it an indicator to present a place from which the observer is invited to regard it.

We observe in the third experiment an alternative way (although here less powerful) of using the principle of relevance. By the normal use of the principle we could obliterate the additional markings (since every state is identically marked) and arrive at the single circle in one step, whereas in the experiment we take the weaker course of obliterating the line of distinction between the markings, and then need one more step to reach the single circle.

Note that both of these ways of simplification are *different* from the methods of cancellation and condensation adopted for the calculus, although arising from, and thus not inconsistent with, them. From the experiment we begin to see in fact how all the constellar principles by which we navigate our journeys out from and in to the form spring from the ultimate reducibility of numbers and voidability of relations. It is only by arresting or fixing the use of these principles at some stage that we manage to maintain a universe in any form at all, and our understanding of such a universe comes not from discovering its present appearance, but in remembering what we originally did to bring it about.

In this way the calculus itself can be realized as a direct recollection. As we left the central state of the form, proceeding outwards and imagewise towards the peripheral condition of existence, we saw how the laws of calling and crossing, that set the stage of our journey through representative space, became fixed stars in the familiar play of time. Our projected hopes and fears of their ultimate atonement, that we called theorems, became their supporting cast. In the end, as we reenter the form, they are all justified and expended. They were needed only as long as they were doubted. When they cannot be doubted, they can be discarded.

Returning, briefly, to the idea of existential precursors, we see that if we accept their form as endogenous to the less primitive structure identified, in present-day science, with reality, we cannot escape the inference that what is commonly now regarded as real consists, in its very presence, merely of tokens or expressions. And since tokens or expressions are considered to be *of* some (other) substratum, so the universe itself, as we know it, may be considered to be an expression of a reality other than itself.

Let us then consider, for a moment, the world as described by the physicist. It consists of a number of fundamental particles that, if shot through their own space, appear as waves, and are thus (as in Chapter 11) of the same laminated structure as pearls or onions, and other wave forms called electromagnetic that it is convenient, by Occam's razor, to consider as travelling through space with a standard velocity. All these appear bound by certain natural laws that indicate the form of their relationship.

Now the physicist himself, who describes all this, is, in his own account, himself constructed of it. He is, in short, made of a conglomeration of the very particulars he describes, no more, no less, bound together by and obeying such general laws as he himself has managed to find and to record.

Thus we cannot escape the fact that the world we know is constructed in order (and thus in such a way as to be able) to see itself.

This is indeed amazing.

Not so much in view of what it sees, although this may appear fantastic enough, but in respect of the fact that it *can* see *at all*.

But *in order* to do so, evidently it must fist cut itself up into at least one state that sees, and at least one other state that is seen. In this severed and mutilated condition, whatever it sees is *only partially* itself. We may take it that the world undoubtedly is itself (i.e. is indistinct from itself), but, in any attempt to see itself as an object, it must, equally undoubtedly, act* so as to make itself distinct from, and therefore false to, itself. In this condition it will always partially elude itself.

It seems hard to find an acceptable answer to the question of how or why the world conceives a desire, and discovers and ability, to see itself, and appears to suffer the process. That it does so is sometimes called the original mystery. Perhaps, in view of *the form* in which *we* presently *take* ourselves *to exist*, the mystery *arises from* our insistence on *framing* a question where there is, in reality, *nothing* to question. However it may appear, if such desire, ability, and sufferance be granted, the state or condition that arises as an outcome is, according to the laws here formulated, absolutely unavoidable. In this respect, at least, there is no mystery. We, as universal representatives, *can* record universal law far enough to say

* Cf ἀγωνιστής = actor, antagonist.

and so on, and so on you will eventually construct the universe, in every detail and potentiality, as you know it now; but then, again, what you will construct will not be all, for by the time you will have reached what now is, the universe will have expanded into a new order to contain what will then be.

In this sense, in respect of its own information, the universe must expand to escape the telescopes through which we, who are it, are trying to capture it, which is us.

Thus the world, when ever it appears as a physical universe*, must always seem to us, its representatives, to be playing a kind of hide-and-seek with itself. What is revealed will be concealed, but what is concealed will again be revealed. And since we ourselves represent it, this occultation will be apparent in our life in general, and in our mathematics in particular. What I try to show, in the final chapter, is the fact that we really knew all along that the two axioms by which we set our course were mutually permissive and agreeable. At a certain stage in the argument, we somehow cleverly obscured this knowledge from ourselves, in order that we might then navigate ourselves through a journey of rediscovery, consisting in a series of justifications and proofs with the purpose of again rendering, to ourselves, irrefutable evidence of what we already knew.

Coming across it thus again, in the light of what we had to do to render it acceptable, we see that our journey was, in its preconception, unnecessary, although its formal course, once we had set out upon it, was inevitable.

* *unus* = one, *vertere* = turn. Any given (or captivated) universe is what is seen as the result of a making of one turn, and thus *is the appearance* of any first distinction.

Appendix 1

Proofs of Sheffer's postulates

Sheffer's postulates [3, p 482] for boolean algebras are chosen for proof because they comprise, amongst those that are widely known, the least such set. They do not constitute, under the constraints he adopted, the least possible such set.

Sheffer's description, quoted below, is in fact complete (although not proved to be so at the time), so that proofs of the postulates in it will serve to prove all postulates in every description of boolean algebra. None, as far as I know, has been proved before.

He assumes

I. A class K,

II. A binary K-rule of combination $|$,

III. The following properties of K and $|$:

1. There are at least two distinct K-elements.
2. Whenever a and b are K-elements, $a \mid b$ is a K-element.

Def. $$a' = a \mid a$$

3. Whenever a and the indicated combinations of a are K-elements,
$$(a')' = a.$$
4. Whenever a, b, and the indicated combinations of a and b are K-elements,
$$a \mid (b \mid b') = a'.$$
5. Whenever a, b, c, and the indicated combinations of a, b and c are K-elements
$$(a \mid (b \mid c))' = (b' \mid a) \mid (c' \mid a).$$

We aim to prove each of the propositions numbered 1 through 5.

Proofs

1. Let the class K be the set of indicators of the states distinguished by the first distinction. There are two such states, from which the first proposition follows.

2. Let $a \mid b$ be written for $\overline{ab\,}\rceil$. The second proposition evidently follows.

Let a' be written for $\overline{a\,}\rceil$. Sheffer's definition
$$a' = a \mid a$$
follows since
$$\overline{a\,}\rceil = \overline{aa\,}\rceil \qquad \text{C5.}$$

Now, if each literal variable is a K-element,

3. $$\overline{\overline{a}\rceil}\rceil = a$$ C1

may be written

$$(a')' = a,$$

4. $$\overline{a\;\overline{b\;\overline{b}\rceil}\rceil}\rceil = \overline{a}\rceil$$ J1

may be written

$$a\,|\,(b\,|\,b') = a'$$

and

5. $$\overline{\overline{a\;\overline{b\;c}\rceil}\rceil}\rceil = \overline{\overline{\overline{b}\rceil\;a}\rceil\;\overline{\overline{c}\rceil\;a}\rceil}\rceil$$ C1 (thrice), J2

may be written

$$(a\,|\,(b\,|\,c))' = (b'\,|\,a)\,|\,(c'\,|\,a).$$

This accounts for the third, fourth, and fifth propositions, and completes the proofs.

Note 1. By the principle of relevance, the stroke in Sheffer's notation may be omitted. Proof of this, which is left with the reader, is perhaps somewhat harder than the immediate apprehension of its truth.

Note 2. . Sheffer explicitly assumes the restriction of his operator to a binary scope, and also, implicitly, assumes the relevance of the order in which the variables under operation appear. Each of these assumptions is in fact less central to mathematics than is commonly supposed, and neither is necessary at this stage. Sheffer was therefore forced to design his initial equations so ingeniously as to contradict them both. The latter he can contradict explicitly, without the disorder becoming too apparent, by allowing $a\,|\,b = b\,|\,a$ as a consequence, but he cannot explicitly contradict the former without obviously denying a rule already recorded, and this would appear foolish, although it is, in fact, now the best way out of the deep trouble that such an ill-considered rule brings in its train. By allowing it to stand, Sheffer's description is rendered practically useless as a calculus.

To understand why Sheffer did not see this, let us take the unusual course of considering his position in the light of the social forces at work around him.

Discoveries of any great moment in mathematics and other disciplines, once they are discovered, are seen to be extremely simple and obvious, and make everybody, including their discoverer, appear foolish for not having discovered them before. It is all too often

forgotten that the ancient symbol for the prenascence of the world is a fool, and that foolishness, being a divine state, is not a condition to be either proud or ashamed of.

Unfortunately we find systems of education today that have departed so far from the plain truth, that they now teach us to be proud of what we know and ashamed of ignorance. This is doubly corrupt. It is corrupt not only because pride is the first of the so-called mortal sins, but also because to teach pride in knowledge is to put up an effective barrier against any advance upon what is already known, since it makes us ashamed to look beyond the bonds imposed by our ignorance.

To any person prepared to enter with respect into the realm of his great and universal ignorance, the secrets of being will eventually unfold, and they will do so in a measure according to his freedom from natural and indoctrinated shame in his respect of their revelation.

In the face of the strong, and indeed violent, social pressures against it, few people have been prepared to take this simple and satisfying course towards sanity. And in a society where a prominent psychiatrist can advertise that, given the chance, he would have treated Newton to electric shock therapy, who can blame any person for being afraid to do so?

To arrive at the simplest truth, as Newton knew and practised, requires *years* of *contemplation.* Not activity. Not reasoning. Not calculating. Not busy behaviour of any kind. Not reading. Not talking. Not making an effort. Not thinking. Simply *bearing in mind* what it is one needs to know. And yet those with the courage to tread this path to real discovery are not only offered practically no guidance on how to do so, they are actively discouraged and have to set about it in secret, pretending meanwhile to be diligently engaged in the frantic diversions and to conform with the deadening personal opinions that are being continually thrust upon them.

In these circumstances, the discoveries that any person is able to undertake represent the places where, in the face of induced psychosis, he has, by his own faltering and unaided efforts, returned to sanity. Painfully, and even dangerously, maybe. But nonetheless returned, however furtively.

We may note in this connexion that Peirce [13], who discovered, some thirty years ahead of Sheffer, that the logic of propositions could be done with one constant, did not publish this discovery, although its importance must have been evident to him; that Stamm, who himself discovered and published[16] this fact two years before Sheffer, omits, in his paper, to make a simple and obvious substitution that would have put his claim beyond doubt; and that Sheffer [3], who ignores Stamm's paper, is wrongly credited with the major discovery recorded in it.

16 E. Stamm, *Monatshefte für Mathematik und Physik,* 22 (1911) 137-40

Appendix 2

The calculus interpreted for logic

The calculus of indications consists of a set of ways of indicating one or the other of the two states distinguished by the first distinction, so we shall be able to find an application of it to the indicative forms of any clear distinction of this kind. It must, for example, apply to cases where doors can be open or shut, or where switches can be on or off, or where lines can be clear or blocked. It will also apply to a language structure in which sentences can be true or false.

Considering the question of its application in the light of the direction from which we have come, it is not immediately obvious that the calculus will have a *useful* or *revealing* application to any of these cases, even though we can see it will apply. The calculus has been built up, in the essay, in a series of forms and departures, and although what we have found there may seem curious, why we took the trouble to look for it may seem equally so.

The fact is, in undertaking the development of the calculus in this direction, the author is making the journey a second time, whereas it may be the first journey for his reader. The author's previous journey was in the opposite direction, from the forms of interpretation we are now about to discuss, towards the form of indication from which they arise. So he is aware, although his reader may not yet be, of how and where it will end, and of the clarifications and simplifications he had to undertake in order to find the way to the place from which he is now returning. He knows, also, that these clarifications will become strengthened on the return journey, although he may still have to convey his vision of their clarity and impression of their strength to the reader.

In interpreting a calculus, what we do is match the values or states or elements allowed in the calculus to a similar set of values or states or elements in what is to become its interpretation. An interpretation is properly matched if each element in it is associated with an identifiable element in the calculus, and the elements in each case have similar distinctions between them. Even so, although there must be this degree of similarity between a calculus and an interpretation of it, in any case of a calculus of more than one value, the calculus and the interpretation are distinct. The fact of their distinction is made plain by the plurality of ways in which a given interpretation can be applied.

With a calculus representing n distinct values there are evidently $n!$ different ways of matching them with n distinct values represented in the interpretation, and thus $n!$ different forms that such an interpretation can take. In interpreting the calculus of indications for sentential logic, we shall match one each of the states of the primary

distinction with one each of the states distinguished by what is true and what is not true, which will offer us $2! = 2$ possible interpretative choices.

The fact that a calculus and an interpretation of it are distinct entities is of crucial importance. By failure to make use of it, we cut ourselves off from forms of simplification that are otherwise readily available. One such form, recognizably frequent in mathematics, consists in the underlying use, when required, of a construction that is devoid of interpretation within the particular application, but that can nevertheless be used to shorten the way to an answer there. A notable example, from outside the field of logic, is the use, as an operator, of $i = \sqrt{-1}$ in electromagnetic theory.

We see, in logic, that 'not true' means the same as 'false', and that 'not false' also means 'true'. So we have a choice of whether to associate the unmarked state with truth and the marked state with untruth, or to associate the marked state with truth and the unmarked state with untruth. Although it is quite immaterial, from the point of view of calculation, which we do, the latter arrangement is in fact easier from the point of view of interpretation.

Accordingly, we identify the marked state, and thereby an empty cross, with true, and the unmarked state, and therein a blank space, with false.

We can now let variables $a, b, \ldots$ stand for the possible truth values of the various simple sentences in a compound sentence, and for this purpose we may allot a distinct variable to each distinct simple sentence.

Next we now must find forms, in the primary algebra, that will properly represent the constants, in the sentential calculus, by which these values are related.

It is clear that we can interpret $\sim a$ or *not a*, through $\overline{a\,|}$ It is also clear that a truth table for $a \vee b$ or *a and/or b*, has exactly the same form as that displayed by the rule of dominance, so that $a \vee b$ can be represented simply by ab. All other forms can now be built up from these. Thus

in words	*in the sentential calculus*	*in the primary algebra*
not a	$\sim a$	$\overline{a\,\vert}$
a or b	$a \vee b$	ab
a and b	$a \wedge b$	$\overline{\overline{a\,\vert}\;\overline{b\,\vert}\,\vert}$
a implies b	$a \supset b$	$\overline{a\,\vert}\; b.$

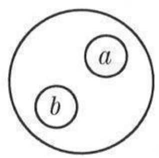
It is the simplicity, in this interpretative choice, of the representation of implication that renders it easier than the alternative, in which $a \supset b$ must be written $\overline{\overline{b}|\; a}|$.

In examining the interpretation as thus set out, we at once see two sources of power that are both unavailable to the standard sentential calculus. They are, notably, the condensation of a number of representative forms into one form, and the ability to proceed, where required, beyond logic through the primary arithmetic.

Regarding the first of these sources, we may take, for the purpose of illustration, the forms for logical conjunction. In the sentential calculus they are

$$a \wedge b$$
$$b \wedge a$$
$$\sim(\sim a \vee \sim b)$$
$$\sim(\sim b \vee \sim a)$$
$$\sim(a \supset \sim b)$$
$$\sim(b \supset \sim a).$$

Each of these six distinct expressions is written, in the primary algebra, in only one way,



This is a proper simplification, since the object of making such sentences correspond with these signs is not representation, but calculation. Thus, by the mere principle of avoiding an unnecessary prolixity in the representative form, we make the process of calculation considerably less troublesome.

But the power granted to us through this simplicity, although great, is itself small compared with the power available through the connexion of the primary algebra with its arithmetic. For this faculty enables us to dispense with a whole set of lengthy and tedious calculations, and also with their no less troublesome alternatives, such as the exhaustive (and mathematically weak) procedures of truth tabulation, and the graphical (and thus mathematically unsophisticated) methods of Venn diagrams and their modern equivalents.

This is made possible by the fact that the three classes of algebraic expression, integral, disintegral, and consequential, that correspond, in the interpretation, with true (tautologous), false (contradictory), and contingent, are readily distinguishable by manipulation.

Example. Classify the following compound sentences in respect of their truth, untruth, or contingency.

1. $(q \supset r) \supset ((p \vee q) \supset (r \vee p))$.
2. $((r \supset p) \supset (\sim(p \vee q))) \wedge p$.
3. $((p \supset q) \wedge (r \supset s) \wedge (q \vee s)) \supset (p \vee r)$.

1.	$= \overline{\overline{q}|\; r}|\;\; \overline{pq}|\;\; rp$	transcription
	$= \overline{\overline{q}|\; r}|\;\; \overline{q}|\;\; rp$	gen (C2)
	$= \overline{\;\;}|\;\; \overline{q}|\;\; rp$	gen
	$= \overline{\;\;}|$	int (C3).

True.

2.	$= \overline{\overline{\overline{\overline{r}|\; p}|\;\; \overline{pq}|}|\;\; \overline{p}|}|$	transcription
	$= \overline{\overline{\overline{\overline{r}|\; p}|\;\; p}|\;\; \overline{\overline{\overline{p}|\;\; pq}|}|}|$	tra (J2)
	$= \overline{\overline{\overline{\overline{\;\;}|\; r}|\; p}|\;\; \overline{\overline{\overline{\;\;}|\;\; pq}|}|}|$	gen (C2) (twice)
	$= \overline{\overline{\overline{\;\;}|}|\;\; \overline{\overline{\overline{\;\;}|}|}|}|$	int (C3) (thrice)
	$=$	pos (J1) (thrice).

False.

3.	$= \overline{\overline{\overline{p}|\; q}|\;\; \overline{\overline{r}|\; s}|\;\; \overline{qs}|}|\;\; pr$	transcription
	$= \overline{\overline{p}|\; q}|\;\; \overline{\overline{r}|\; s}|\;\; \overline{qs}|\;\; pr$	ref (C1)
	$= \overline{qs}|\;\; pr$	occ (C4) (twice).

Contingent.

These calculations, conducted in the primary algebra, are so simple as to be trivial. That is to say, the moment each of the sentences is written down in the calculus of indications, the answer, to any person familiar with this form, becomes obvious to mere inspection. I have here done the calculations slowly, in very small steps, on the assumption that the reader is not yet familiar with the form.

The consequences of this arithmetical availability are sweeping. All forms of primitive implication become redundant, since both they and their derivations are easily constructed from, or tested by reduction to, a single cross. For example, everything in pp 98-126 of *Principia mathematica* can be rewritten without formal loss in the one sign

$$\overline{\quad\;}|$$

provided, at this stage, the formalities of calculation and interpretation are implicitly understood, as indeed they are in *Principia*. Allowing some 1500 signs to the page, this represents a reduction of the mathematical noise-level by a factor of more than 40000.

With such a huge gain in the formal clarity of expressions, the invalidity of a false argument is similarly open to immediate confirmation. We illustrate such an argument below, offered[17] by Maurant as a dilemma.

> If we are to have a sound economy, we must not inflate the currency. But if we are to have an expanding economy, we must inflate the currency. Either we inflate the currency or we do not inflate the currency. Therefore, we shall have neither a sound economy nor an expanding economy.

Let

s	stand for	*we have a sound economy*
c	stand for	*we inflate the currency*
e	stand for	*we have an expanding economy.*

Transcribing it into the primary algebra, we find

$$\overline{\overline{\overline{\overline{s|}\;\overline{c|}\,|}\;\overline{\overline{e|}\;c\,|}\;\overline{c\;\overline{c|}\,|}\,|}\,|}\;\overline{se|}$$

$$= \overline{\overline{s|}\;\overline{c|}\,|}\;\overline{\overline{e|}\;c\,|}\;\overline{se|} \qquad \text{pos (J2), ref (C1).}$$

This expression is consequential, but if the fact is not yet apparent, we may use the

17 John A Maurant, *Formal logic,* New York, 1963, p 169.

converse of the theorem 16 with an arbitrary variable and constant, say $c = \overline{\quad}|$, giving

$$\overline{\overline{s|}\;\overline{\overline{\quad}|\,|}\,|}\;\overline{\overline{e|}\;\overline{\quad}|\,|}\;\overline{se|}$$

$$= \overline{e|}\;s \qquad \text{int (C3), ref (C1) (thrice), gen (C2).}$$

This plainly cannot be reduced, so nor can the original. Thus there is no dilemma. Other characteristics of the argument are also illuminated, especially the utter irrelevance of the premiss ‘either c or not c’.

If we stand back for a moment to regard the structure of an implicational logic, such as Whitehead and Russell’s, we see that it is fully contained in that of an equivalence logic. The difference is in the kind of step used. In the one case expressions are detached at the point of implication, in the other they are detached at the point of equivalence.

If an expression is detached at the point of implication, it of course need not be equivalent to the expression from which it is derived. But if it is a tautology it can be implied only by another tautology, so that, in such cases, the sign of implication can always be replaced by a sign of equivalence. Thus an implicational logic in fact degenerates into an equivalence logic in respect of the class of true statements, with which such logics are most intimately concerned.

The completeness theorem for the primary algebra in the text is what, interpreted in logic, is called a strong completeness theorem, since it includes Post’s original[18] weaker theorem. The weaker version merely asserts that all true statements are implied by the true statements initially given as primitive. Since, in the case of true statements, implication is equivalent to equivalence, we see that such a theorem must be included in a theorem that states the completeness of all forms of equivalence, irrespective of whether the statements interpreted from them are true, false, or contingent.

We may turn, now, to consider how the calculus of indications can be applied to the traditional logic of classes. Before doing so, it is of interest to state another hitherto silent (or relatively silent) assumption to the effect that, in the absence of instructions to the contrary, we assume the premisses of an argument to be related by logical conjunction. For example, in transcribing the alleged dilemma above, we first cross the transcription of each individual premiss and then cross the result to give the conjunction, and finally cross all this again for the implication. We have, in fact, habitually come to regard ‘and’ as the proper interstitial constant. But we could, for example, rephrase both sentential and class logic on the assumption that ‘or’, instead of ‘and’, is

18 Emil L Post, *Amer. J. Math,* 43 (1921) 163-85

the constant relating premisses. The reader might like to attempt a proof of this. It is a revealing exercise, especially with respect to the logic of classes, and it is not difficult.

All universal forms of the traditional logic of classes can be accommodated within the logic of sentences, so we will consider these forms first. To accommodate them, we use the pattern in the following key.

for *all a are b* use $(x \in a) \supset (x \in \mathrm{b})$

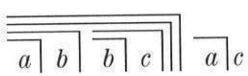
for *no a is b* use $(x \in a) \supset (x \in \text{not-}b)$

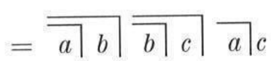
and other forms accordingly. To avoid the use of distinct letters for sentences and classes, we can allow, in the calculating forms, any simple literal variable v to stand for the sentence '$x \in v$' i.e. 'x is a member of the class v'. This will not lead to unintentional confusion, since the sign v, as used to denote the class, does not enter the calculation, which is undertaken with v representing only the truth value of the corresponding sentence.

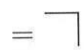
Taking the form of a syllogism in Barbara, and putting the minor premiss first, as Whitehead and Russell do, we find

if all a are b
and all b are c
then all a are c

which we can represent by



F1  ref

 gen (thrice), int.

The sentential form is thus seen to be a tautology and the argument thereby valid. In the case of an invalid argument, the algebraic expression will not reduce to a cross, so we have a reliable system for testing the validity of any universal argument in syllogism form. We shall later study a method that will determine the conclusion from the premisses alone. In the present form, as we see, although its validity can be tested, the conclusion, given the premisses alone, can be found only by trial.

Equivalence problems are similarly open to solution in this way.

Example[19] A club has the following rules

(a) The Financial Committee must be chosen from among the General Committee,

(b) No-one shall be a member of the General and Library Committees unless he is also on the Financial Committee,

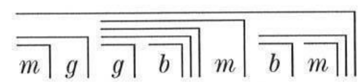
(c) No member of the Library Committee shall be on the Financial Committee.

Simplify these rules.

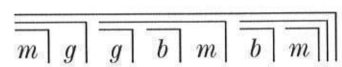
Procedure.

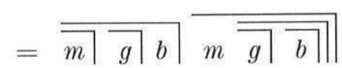
for *x is a member of the Financial Committee* write m
for *x is a member of the General Committee* write g
for *x is a member of the Library Committee* write b.

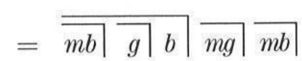
The interstitial constant of a set of rules is usually understood to be conjunction, so we may now transcribe them into the primary algebra as follows.

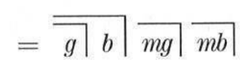


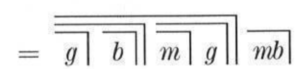
Our aim is to reduce this, if possible, to a simpler conjunctive form that is equivalent to, and may thus be used to replace, the original set of rules.

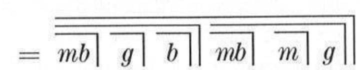
F2  ref

 cro (C9)

 gen, tra, ref

F3  gen

 cro

 tra

19 from B V Bowden, *Faster than thought*, London, 1953, p 36.

$$\text{F4} \qquad = \overline{\overline{\overline{\overline{g|}\; b|}|}\;\overline{\overline{m|}\; g|}|} \qquad \text{occ (twice).}$$

Retranscribing gives the answer

(1) The Financial Committee must be chose from among the General Committee,

(2) Nŏ member of the General Committee shall be on the Library Committee.

We may check this answer by theorem 16. Let $m = \overline{\;\;|}$. Now

$$\text{F2} = \overline{\overline{g|}\; b|}$$

$$\text{F4} = \overline{\overline{\overline{\overline{g|}\; b|}|}\;\overline{g|}|} = \overline{\overline{g|}\; b|}.$$

Let $m =$. Now

$$\text{F2} = \overline{\overline{\overline{\overline{g|}\;\overline{b|}|}|}|}$$

$$\text{F4} = \overline{\overline{\overline{\overline{g|}\;\overline{b|}|}|}|},$$

so the answer is correct, provided only that we have properly interpreted the problem.

We see that we can, from this answer, obtain an implication (not an equivalence) to the effect that no member of the library committee shall be on the financial committee, since by crossing F4 (for the implication) and reflecting we get

$$\overline{\overline{g|}\;\overline{b|}|}\;\overline{\overline{m|}\; g|},$$

and now adding our tentative conclusion gives

$$\overline{\overline{g|}\;\overline{b|}|}\;\overline{\overline{m|}\; g|}\;\overline{b|}\;\overline{m|}$$

$$= \overline{\;\;|}.$$

The mathematical structure illustrated in this sort of inference suggests the following proposition.

Interpretative theorem 1

If the primary algebra is interpreted so that integral expressions are true, and if each of a number of class-inclusion premisses is sententially transcribed in it, and if variables representing the same sentence at odd and even levels are cancelled, what remains, when retranscribed, is the logical conclusion.

The proof is not difficult and may be left with the reader. The theorem itself, as a short cut to inference, is of considerable power. We may take Lewis Carroll's last sorites to illustrate it.

The problem is to draw the conclusion from the following set of premisses.

(1) The only animals in this house are cats;

(2) Every animal is suitable for a pet, that loves to gaze at the moon;

(3) When I detest an animal, I avoid it;

(4) No animals are carnivorous, unless they prowl at night;

(5) No cat fails to kill mice;

(6) No animals ever take to me, except what are in this house;

(7) Kangaroos are not suitable for pets;

(8) None but carnivora kill mice;

(9) I detest animals that do not take to me;

(10) Animals, that prowl at night, always love to gaze at the moon.

The method employed hitherto to solve such a problem was to work it out by stages, but this can be quite time consuming. Using the theorem above, we simply adopt a distinct variable for each distinct (but not complementary) set, transcribe, cancel, and arrive at the answer instantaneously. Let us, then, proceed to adopt

h for *house, in this*

c for *cat*

p for *pet, suitable for*

d for *detested by me*

a for *avoided by me*

m for *moon, love to gaze at*

v for *carnivorous*

n for *night, prowl at*

k for *kill mice*

t for *take to me*

r for *kangaroo.*

We see from the principle of relevance that we do not need to adopt a variable for the set of animals. We now proceed to the transcription and cancellation

$$\overline{\cancel{h}\,|}\;\cancel{c}\;\overline{\cancel{m}\,|}\;\cancel{p}\;\overline{\cancel{d}\,|}\;a\;\overline{\cancel{v}\,|}\;\cancel{n}\;\overline{\cancel{c}\,|}\;\cancel{k}\;\overline{\cancel{t}\,|}\;\cancel{h}\;\overline{r\,|}\;\overline{\cancel{p}\,|}\;\overline{\cancel{k}\,|}\;\cancel{v}\;\cancel{t}\;\cancel{d}\;\overline{\cancel{n}\,|}\;\cancel{m}$$

which reveals $\overline{r\,|}\;a$. Therefore, all kangaroos are avoided by me.

So far we have considered how the calculus of indications, in the form of the primary algebra, may be used to clarify and simplify problems in sentential logic, and also those of universal, or non-existential, import in class logic or set theory. We shall turn now to consider its extension, in class logic, to problems of existential, or particular, import.

We resolved the question of how to represent a universal statement such as

all a are b

by translating it into an equivalent statement in the sentential calculus. The question we must now seek to answer is, can an existential statement, such as

some a are b

be similarly translated?

We first note that, to contradict the general assertion that all a are b, it is sufficient to find some a that is not b. We may note by the way that the statement.

no a is b

does not contradict

all a are b,

since, in case a is non-existent, both assertions are true.

Transcribing according to the principles already adopted, we take

some a are not b,

to say

not all a are b,

and so represent it by

$$\overline{\overline{a\,|}\;b\,|}\,.$$

Similarly we present

some a are b

by

$$\overline{\overline{a\,|}\;\overline{b\,|}\,|}\,.$$

To see how this works out, we transcribe another syllogism, this time of existential import. Thus

all a are b

some a are c

$\therefore$ some b are c

becomes

$$\overline{\overline{\overline{\overline{a\,|}\;b\,|}\;\overline{\overline{\overline{\overline{a\,|}\;\overline{c\,|}\,|}\,|}\,|}\,|}\,|}\;\overline{\overline{b\,|}\;\overline{c\,|}\,|}$$

$$=\overline{\overline{a\,|}\;b\,|}\;\overline{a\,|}\;\overline{c\,|}\;\overline{\overline{b\,|}\;\overline{c\,|}\,|}$$

$$=\overline{\;\,|}\,.$$

Since we can see otherwise that this syllogism is valid, it appears to be properly represented. But using the same rules we can represent

some a are b

some b are c

$\therefore$ some a are c

which we know to be invalid, by

$$\overline{\overline{\overline{\overline{\overline{a\,|}\;\overline{b\,|}\,|}\,|}\;\overline{\overline{\overline{b\,|}\;\overline{c\,|}\,|}\,|}\,|}\,|}\;\overline{\overline{a\,|}\;\overline{c\,|}\,|}$$

$$=\overline{a\,|}\;\overline{b\,|}\;\overline{b\,|}\;\overline{c\,|}\;\overline{\overline{a\,|}\;\overline{c\,|}\,|}$$

$$=\overline{\;\,|}\,,$$

in which it appears to be valid. How do we resolve this seeming contradiction?

Let us be clear of one thing. The question is answered (implicitly, since it is not usually asked) in the textbooks as it was originally answered by Aristotle, by giving a more or less complicated set of rules that disallow this inference. But a set of rules to say *that* one must not do something is not an explanation of *why* one must not, and nor does the fact that, if we allow the inference, it may mislead us to an improper conclusion, meet with the high degree of understanding required of all explanatory accounts in this book. We have found an area in which an apparently impeccable interpretative procedure has suddenly let us down, and all rules that say we must therefore avoid this area, however well they may work in practice, have an unsatisfactorily *ad hoc* flavour.

It is no use, either, appealing to graphical forms such as Venn diagrams, since these, in common with other graphs, offer a picturesque realization that is peripheral, not central, to the question. To answer it, we must find an altogether subtler approach.

We begin with the observation that statements about the universe of discourse, e.g.

if there is an a in it

then there is also a b in it,

assert no claim to the existence of anything in it, although they may be taken, at a different level, to claim the existence of the universe that has these conditional properties. But to deny such a statement, we claim that

there *exists* at least one a in it that is not a b.

Now the distinction between existing and not existing is not applied like the distinction between true and not true. If a statement s is true, then its complementary statement $\sim s$ is false. But if a thing t exists, then its complementary thing not-t is not necessarily non-existent. In the universe of England, the complement of London is the country. Both, at the time of writing, apparently exist. Thus no existence follows from another existence, so that from a statement, or a list of statements, asserting existence only, no proper conclusion can be drawn.

So far we are still at the periphery. That is to say, we are still examining the form of the interpretation, without finding exactly how and where it breaks faith with the mathematics.

In relating the mathematics and the interpretation, we found forms such as

$$\overline{a|}\; b,$$

that say nothing, in their interpretation, about existence, neither asserting nor denying it. But such forms, when crossed,

$$\overline{\overline{a|}\; b\,|},$$

now do say something about existence, at least in the interpretation we have allowed them.

The expression $\overline{a|}\; b$ is universal because it limits the shape of the universe so that there is no space in it for an a that is not a b. At least, that is how we take it. But we could (although we don't) take it to mean, simply, that in this universe there just happens not to exist an a that is a b, although there is the space, if we need it, to hold one. In other words, we *could* (although we don't) interpret it existentially.

Similarly we *could* interpret the expression $\overline{\overline{a|}\; b\,|}$ universally. The statement that some a are not b, although *sufficient* to contradict the statement that all a are b, is not *necessary.* An alternative way to contradict it would be simply to deny that the universe is of such a form as to demand of any a that it shall be also a b, without requiring the existence of an a to prove it.

In this alternative we have a means of confining all interpretations to a non-existential import. Let us see how it works out in the case of the invalid syllogism. We should now write

some a are b

some b are c

$\therefore$ some a are c

in the form

it is not the case that no a is b

it is not the case that no b is c

$\therefore$ it is not the case that no a is c,

making explicit the requirement than no statement is to be taken existentially.

Even so, at first sight, we are not entirely out of trouble. For although, from such a description, the universe appears compelled to reserve space for a's that might be b's and for b's that might be c's, it does not appear *compelled* (as, by the implication, it should be) to reserve any space for a's that might be c's.

But a universe without such space would contain at most six different departments, since it would be missing a department for a's that are also b's that are also c's, and for a's that are also not-b's that are also c's. Now there is a well-known theorem, a proof of which was published [14, p 309] by Huntington in 1904, according to which *the number of elements in every finite logical field must be* 2^m (*m an integer* >0). Thus an algebra suitable for such a logical field cannot, without further constraint, represent a form in which the number of elements is not a natural integral power of 2. Such a constraint, when required, is normally imposed through the premisses. That is to say, if any of the possible 2^m spaces is required to be absent from the universe, it must be positively (i.e. referentially) excluded. None of the possible eight spaces is excluded by the premisses of the syllogism above, and so all eight must be presumed to exist, or the mathematical form cannot be properly interpreted. And if they exist, the conclusion follows.

Another, and perhaps easier, way to see *in what sense* the traditionally invalid syllogism above is valid, is to return to our original method of interpretation. Using standard sentential constants it becomes

$$\sim((x \in a) \supset (x \in \text{not-}b) \wedge \sim(\sim(x \in \text{not-}b) \supset (x \in \text{not-}c)) \supset \sim(x \in a) \supset (x \in \text{not-}c)$$

and is of course, in this form, true.

Let there be no mistake, we do not assert, by this, that the syllogism taken, isolated, within the ordinary meaning of 'some a are b, some b . . . etc' is anything other than invalid. It is just that, in trying to place it in a deeper mathematical foundation, we come across (or up against) the inconsequential relation, apparent in ordinary speech, between a form and its content, occasioned by the partly accidental fact that the existence of a particular content can serve to negate a general form.

It remains for us to extricate ourselves, as elegantly as we can, from the unintentional confusion that follows in the train of such a state of affairs: or alternatively, if we have so extricated ourselves, to devise the most peaceful set of rules by which the possibility of such confusion can be laid to rest. The rules that, by tradition, are enlisted to serve this purpose are too numerous for what is a basically simple ambiguity, and they can surely be reduced.

Such a reduction, as we have seen, will be mathematically powerful if it can be taken to a point of degeneration. In this case the ideal degeneration would be at a place where the two kinds of denial, universal or existential, of a universal proposition amount to the same thing. At such a point we could use the calculus freely, without fear of its letting us down.

We have observed that as long as inferences or equations in class logic are universally interpreted, the primary algebra can be freely used to determine them. In other words, the sentential form into which we placed universal statements about classes or sets can be seen to accommodate them exactly, without formal loss or gain. It is the denials of such statements, when we wish to interpret them existentially, that present the difficulty, which arises evidently from a formal gain, since we find a need to constrain the calculus in this respect, rather than to relax it.

Let us return, for a moment, to examine our procedure for solving Bowden's problem about the club rules. In the algebraic path to its solution we find an expression

$$\overline{\overline{g|}\;b|}\;\;\overline{mg|}\;\;\overline{mb|},$$

marked F3. Taken existentially, it would mean

either	some g are not b
or	some things are neither m nor g
or	some things are neither m nor b.

But in fact the whole argument depends on not taking F3, or any other intermediate expression, this way. Algebraically, of course, it doesn't matter, we have no choice, and arrive at the answer willy-nilly. It is only on retracing the path by which we got there, and stopping on the way to look at the pitfalls, that we see the alarming prospect of the interpretative dangers that it effectively by-passed.

The first rule that suggests itself, therefore, is never to make an existential interpretation unless the argument demands it. No such demand is evident in Bowden's problem, and so, in solving it, we can effectively avoid existence, and thereby avoid the pitfalls it

brings in its train. The question that then frames itself is how far we can take this avoidance, or, considered in reverse, in what circumstances, and at what place, during the course of solving a problem, do we ever need to make an existential interpretation?

The answer is none. Existential interpretations, where they are necessary at all, can be confined to entering and leaving the problem, and need never occur in the course of solving it.

To see how this comes about, we may return to the syllogism in Barbara, taken in the form

F1 $$\overline{\overline{a|}\;b|}\;\;\overline{\overline{b|}\;c|}\;\;\overline{a|}\;c\,.$$

Since the order of each of the three complexes in F1 is irrelevant to the meaning of the whole expression, we may transpose it to find

F1' $$\overline{\overline{a|}\;b|}\;\;\overline{a|}\;c\;\;\overline{\overline{b|}\;c|}$$

$$= \overline{\overline{\overline{\overline{a|}\;b|}\;\;\overline{\overline{\overline{\overline{a|}\;c|}|}|}|}|}\;\;\overline{\overline{b|}\;c|}$$

which can be retranscribed

all a are b

some a are not c

$\therefore$ some b are not c

Transposing it yet again, we find

F1'' $$\overline{a|}\;c\;\;\overline{\overline{b|}\;c|}\;\;\overline{\overline{a|}\;b|}$$

which will give

some a are not c

all b are c

$\therefore$ some a are not b

So we see that the representative form of a syllogism in Barbara *is also* the representative form (remembering that we have in each case put what is called the minor premiss first) of syllogisms in Bocardo and Baroco. The three syllogisms above, being effectively reducible to the same mathematical expression, must therefore represent, at this level, an *identical* form of argument.

This is both interesting and fascinating. It is interesting because, from it, we shall be able to obtain a much simplified rule-structure for existential arguments, and fascinating because of the light it sheds on what we are doing when we argue from existence. We may note in passing, as Prior reminds[20] us, that glimpses of the path to this identity are apparent in the work of Aristotle, who refers to a form lately more fully described[21] by Ladd-Franklin, in which what she calls an antilogism condenses three syllogisms. Here we elucidate a further stage, in which the three-in-one nature of the syllogism is evident from its transcription alone, without recourse to an image or anti-logism.

From the conversion (or converse) of what we have just recounted, we observe the following proposition.

Interpretative theorem 2

An existential inference is valid only in as far as its algebraic structure can be seen as a universal inference.

For example, each of the existential arguments transcribed from F1' and F1'' is valid because of the validity of the universal argument transcribed from F1.

This single rule takes care of all the separate rules for syllogisms, their parts, and their extensions. It even includes the provision that there shall be not more than one particular premiss, for with more than one, no representation as a universal argument is possible.

We have here found the degeneration we were seeking, at the place where the existential condenses with the universal. This degeneration, like the one undertaken earlier for the sentential calculus, is a release from the bond of the particular, and through it we see the whole syllogistic structure in the one prototype

$$\overline{\overline{a|}\; b|}\ \ \overline{\overline{b|}\; c|}\ \ \overline{a|}\; c\,.$$

In this prototype, not only can we transpose each complex, we can also independently cross each literal variable, finding, by a combination of these means, a set of 24 distinguishable valid arguments. Formally there is no difference between them. If we distinguish any, we should distinguish all. In fact not all twenty-four are distinguished in logic, which arrives somewhat arbitrarily at the number fifteen.

Thus, by realizing a condensation, we no longer need to remember, for syllogisms or related arguments, the wearisome rules of their construction and validity. All these are now subsumed in, and can be reconstructed from, the simple basic form and interpretation to which we have here reduced them.

20 A N Prior, *Formal logic*, 2nd edition, Oxford, 1964, p 113.
21 C F Ladd-Franklin, *Mind*, 37 (1928) 532-4

We may return, for a moment, to reconsider the sorites, which is the general form under which the syllogism is the primary member. In the light of the degeneration undertaken above, we see that the method we developed for revealing a conclusion by cancellation applies equaly whether the argument is universal or existential. For a universal sorites we have

$$\overline{a|}\; b,\quad \overline{b|}\; c,\ \ldots,\ \overline{p|}\; q,$$

$$\therefore \qquad \overline{a|}\; q.$$

To convert it into an existential one we simply negate the conclusion and transpose it with one of the premisses which, itself negated, becomes the new conclusion. So from

$$\overline{a|}\; b,\quad \overline{b|}\; c,\quad \overline{c|}\; d,$$

$$\therefore \qquad \overline{a|}\; d$$

we find, for example, the set of premisses

$$\overline{a|}\; b,\quad \overline{b|}\; c,\quad \overline{\overline{a|}\; d|},$$

from which, as before, the conclusion can be revealed by cancellation,

$$\overline{\cancel{a}|}\; \cancel{b}\ \ \overline{\cancel{b}|}\; c\ \ \overline{\overline{\cancel{a}|}\; d|}.$$

All we need to remember is that it will now be existential, and so should in this case be written

$$\overline{\overline{c|}\; d|}.$$

We leave the account here, where the interested reader will be able to continue it at his or her pleasure. The problems solved so far, and the questions answered, are simple ones, although the calculus is, in practice, successfully applied[22] to the solution of problems of great complexity. So much, at the primitive level, is commonly overlooked, and what is seen is normally recounted in a fashion so fragmentary as to be hardly coherent. The very act of dwelling for a while with even a simple form can evidently tax the whole of one's powers, so that to leave the simple forms before one is properly familiar with them can result in many unrewarding, or largely unrewarding, mathematical excursions.

In fact this whole book is concerned with what can be found, if we seek it, at a level of extreme simplicity, and is in the way of being beyond elementary, but beyond on the side of simplicity, not complexity. This does not necessarily make it easy, but will, I feel, allow it to be rewarding.

22 Cf. G Spencer-Brown, British Patent Specifications 1006018 and 1006019 (1965).

Index of references

Note. In context, a page reference is confined to what is of particular interest to the discussion. In this index it is expanded to include the whole work.

[1] George Boole, *The mathematical analysis of logic*, Cambridge, 1847. xiv

[2] Alfred North Whitehead and Bertrand Russell, *Principia mathematica*, Vol. I, 2nd edition, Cambridge, 1927 xiv

[3] Henry Maurice Sheffer, *Trans. Amer. Soc.*, 14 (1913) 481-8. xiv

[4] Ludwig Wittgenstein, *Tractatus logico-philosophicus*, London, 1922. xvi

[5] Kurt Gödel, *Monatshefte für Mathematik und Physik*, 38 (1931), 172-98. xvii

[6] Alonzo Church, *J. Symbolic Logic*, 1 (1936), 40-1, 101-2. xvii

[7] Abraham A Fraenkel and Yehoshua Bar-Hillel, *Foundations of set theory*, Amsterdam, 1958, pp 136-95. xviii

[8] P B Medawar, Is the Scientific Paper a Fraud, *The Listener*, 12 September 1963, pp 377-8. xix

[9] R D Laing, *The politics of experience and the bird of paradise*, London, 1967. xix

[10] Edward V Huntington, *Trans. Amer. Math. Soc.*, 35 (1933), 274-304. 72

[11] Alfred North Whitehead, *A treatise on universal algebra*, Vol. I, Cambridge, 1898. 73

[12] Charles Sanders Peirce, *Collected papers*, Vol. IV, Cambridge, Massachusetts, 1933. 73

[13] *ΠΡΟΚΛΟΥ ΔΙΑΔΟΧΟΥ ΣΤΟΙΧΕΙΩΣΙΣ ΘΕΟΛΟΓΙΚΗ* with a translation by E R Dodds, 2nd edition, Oxford, 1963. 74

[14] Edward V Huntington, *Trans. Amer. Math. Soc.*, 5 (1904), 288-309. 79

[15] George Boole, *An investigation of the laws of thought*, Cambridge, 1854. 79

[16] E Stamm, *Monatshefte für Mathematik und Physik*, 22 (1911), 137-49. 89

[17] John A Maurant, *Formal logic*, New York, 1963. 94

[18] Emil L Post, *Amer. J. Math.*, 43 (1921), 163-85. 95

[19] B V Bowden, *Faster than thought*, London, 1953. 97

[20] A N Prior, *Formal logic*, 2nd edition, Oxford, 1964. 105

[21] C F Ladd-Franklin, *Mind*, 37 (1928), 532-4. 105

[22] G Spencer-Brown, British Patent Specifications 1006018 and 1006019 (1965). 107

Index of forms

Note. A theorem marked with an asterisk has a true converse.

Theorems

representative

*T1	The form of any finite cardinal number of crosses can be taken as the form of an expression.	10
T2	If any space pervades an empty cross, the value indicated in the space is the marked state.	11
T3	The simplification of an expression is unique.	12
T4	The value of any expression constructed by taking steps from a given simple expression is distict from the value of any expression constructed by taking steps from a different simple expression.	15

procedural

T5	Identical expressions express the same value.	16
*T6	Expressions of the same value can be identified.	17
*T7	Expressions equivalent to an identical expression are equivalent to one another.	17

connective

T8	If successive spaces s_n, s_{n+1}, s_{n+2} are distinguished by two crosses, and s_{n+1} pervades an expression identical with the whole expression in s_{n+2}, then the value of the resultant expression s_n is the unmarked state.	18
T9	If successive spaces s_n, s_{n+1}, s_{n+2} are arranged so that s_n, s_{n+1} are distinguished by one cross, and s_{n+1}, s_{n+2} are distinguished by two crosses, then the whole expression e in s_n is equivalent to an expression, similar in other respects to e, in which an identical expression has been taken out of each division of s_{n+2} and put into s_n.	19

algebraic

T10	The scope of J2 can be extended to any number of divisions of the space s_{n+2}.	31
T11	The scope of C8 can be extended as in T10.	32
T12	The scope of C9 can be extended as in T10.	32
T13	The generative process in C2 can be extended to any space not shallower than that in which the generated variable first appears.	32
T14	From any given expression, an equivalent expression not more than two crosses deep can be derived.	33
T15	From any given expression, an equivalent expression can be derived so as to contain not more than two appearances of any given variable.	33

mixed

*T16	If expressions are equivalent in every case of one variable, they are equivalent.	39
T17	The primary algebra is complete.	41

algebraic

T18	The initials of the primary algebra are independent.	44

Rules of Substitution and Replacement

R1	If $e = f$, and if h is an expression constructed by substituting f for any appearence of e in g, then $g = h$.	21
R2	If $e = f$, and if every token of a given variable expression v in $e = f$ is replaced by an expression w, it not being necessary for v, w to be equivalent or for w to be independent or variable, and if as a result of this procedure e becomes j and f becomes k, then $j = k$.	21

Consequences

C1	$\overline{\overline{a}\|}\| = a$	ref	23
C2	$\overline{ab}\|\, b = \overline{a}\|\, b$	gen	25
C3	$\overline{\ \ }\|\, a = \overline{\ \ }\|$	int	26
C4	$\overline{\overline{a}\|\, b}\|\, a = a$	occ	26
C5	$aa = a$	ite	27
C6	$\overline{\overline{a}\|\ \overline{b}\|}\|\ \overline{\overline{a}\|\, b}\| = a$	ext	27
C7	$\overline{\overline{\overline{a}\|\, b}\|\, c}\| = \overline{ac}\|\ \overline{\overline{b}\|\, c}\|$	ech	28
C8	$\overline{\overline{a}\|\ \overline{br}\|\ \overline{cr}\|}\| = \overline{\overline{a}\|\ \overline{b}\|\ \overline{c}\|}\|\ \overline{\overline{a}\|\ \overline{r}\|}\|$	mod	28
C9	$\overline{\overline{\overline{b}\|\ \overline{r}\|}\|\ \overline{\overline{a}\|\ \overline{r}\|}\|\ \overline{\overline{x}\|\, r}\|\ \overline{\overline{y}\|\, r}\|}\| = \overline{\overline{r}\|\, ab}\|\ \overline{rxy}\|$	cro	28

Appendix 3

Bertrand Russell and the *Laws of Form*

Bertrand Russell loved my poems, and his favourite was my first, the one about the talking cows. He was terrified of my mathematics, but I was not to know this until he informed the world of it in his autobiography.

In April 1965 he told me he wished to acquire a more complete understanding of my new work in the algebras of logic, and invited me to stay through the inside of a week to give him tutorials on it morning and afternoon.

I was myself so terrified at this prospect, that his own fears about it were not apparent to me. He hid them under a mask of testiness, sharpness, and irritability. It never occured to me that all such conduct is symptomatic of terror.

The irony of my position was not lost on me – that this great author and scholar, my senior by more than half a century, whom I had idolized at school, and whose works I knew by heart, had unexpectedly engaged me, like a college tutor, to teach him algebra. The thought of it overwhelmed me, and I did not see how I could possibly succeed.

He was ninety-three years old. If he hadn't written *Principia Mathematica*, in which he did everything a different way, and in my view the wrong way, there might have been a vestige of hope. As it was, I could see none.

The Cambrian Coast Express – now, I think, sadly discontinued – from Paddington to Penrhyndeudraeth, covered some of the most astonishing stretches of track in existence. It took all day, beginning among the grimy suburbs of London, continuing through miles of conventional countryside, and then, just when the traveller is thinking he can no longer remain conscious, it suddenly bursts out to thread its way through the most magical coastline imaginable. One moment in a tunnel, the next out on a narrow ledge of rock with the waves pounding dangerously below, then more tunnels and escarpments. And then, when the traveller is beginning to expect more of the same, it unexpectedly takes off into space, out across an extraordinary single-track viaduct, so narrow as to be invisible from the coach windows at either side, and on which the train appears to glide like a low-flying aircraft over miles of winding river estuaries and deserted mud flats.

At the final destination, it is as if one has arrived on a new planet. Everything is similar, yet somehow strangely different. A group of children, lonely for want of a leader, gather at the station exit, in the hope of catching a glimpse of any extraterrestrial being that might emerge from the train. I was always the sole passenger to alight, and whenever I did, two or three of the girls in the party would propose to me. Aged about ten.

The train was scheduled to arrive in time for dinner, but was always exactly an hour late. And the village taxi driver, impatiently brushing the offers of marriage aside, was always in a hurry to load my luggage and drive me at break-neck speed to Plas Penrhyn, where Bertie and Edith would always have started dinner without me.

Invariably, Bertie would greet me at the gate.

'We couldn't wait,' he would say, thrusting a large triple whisky into my hand, even before I got inside, so that I could catch up.

The morning tutorials were to be at eleven thirty. Bertie explained they could not be earlier, because before that he would be making love to Edith. I knew it was true, because they left the bedroom door open. The afternoon tutorials would be at two.

Each morning, on the dot, a manservant came up to my room.

'His Lordship will see you now, Sir.'

He would escort me down to the famous drawing room, where Bertie did all his television interviews, open the door for me, motion me inside, and shut it behind me.

Although Bertie was physically small and slight, his charismal presence was enormous.

'Well?' he barked. The tutorial had begun.

Needless to say, it went very badly. Bertie would not leave his home ground, the way he had done it in *Principia*. And until he did so, I could not show him the way I did it in the *Laws*. As the week went by, he grew more and more irritable, and I more despairing. Come Friday, I had resigned myself to complete failure. The morning tutorial went particularly badly, and for the afternoon one, the last, I was prepared merely to go through the motions.

When it came, he was even more testy than usual.

'And how do you deal with propositional functions?' he snapped.

'Bertie I don't,' I replied.

'Oh?' he said. 'And why not?'

'Because I'm not a good enough mathematician,' I said.

'Oh,' he said. His manner changed, and he grew confidential.

'You know, I have never admitted this to anyone before, but I, too, am a *very* bad mathematician. Come, let me see what you have written.'

I handed him the typescript, and he began to read it from the beginning, page by page, in complete silence. Three quarters through, he looked up at me. He said,

'I think I can take the rest as read. I see, now, what you have done. In the history of mathematics, it is very rare indeed. You have made a new calculus, of great power and simplicity, and I congratulate you!'

Very few of my friends had understood it, and even those who did, found it so difficult

that they never thought to congratulate me. Bertie was my one link with a bygone age, when great men were generous in their regard for one another, and were mindful of the appropriate social formalities in respect of a significant advancement by one of their number. He knew that congratulations were in order, and, for the first time, I too saw that they were.

When wrong is done we sometimes laugh, but when right is done we cry. The tears were streaming down my face, and Bertie was weeping too. He opened a fresh bottle of Scotch, and filled two tumblers to the brim, one for me and one for him.

I knew Bertie could always drink me under the table. He consumed at least a bottle of whisky a day, as well as countless cigarettes, which Edith had specially imported from the USA, and which he would not allow me to refuse. It generally took me several days to recover physically from his hospitality.

Knowing his capacity for alcohol, and my own, I decided, as I began on the full tumbler he had handed to me, to stay glued in my chair until my constitution had recovered sufficiently to be able to walk again.

Not so Bertie. He finished his glass in a few gulps, then got up and marched about the room making declamatory speeches in Latin. But the alcohol had unsteadied him, and to recover his balance he made a grab at the huge revolving bookcase that towered over his chair.

I have never seen such a bookcase. It stood about eight feet tall, at least a yard wide, and circular, like the tower of Pisa. It must have contained a thousand books. As he grabbed it, there came an ominous crack of splintering wood from the pedestal at the base, and the whole edifice tottered crazily towards him, showering down its contents. And as each section of shelves, relieved of the burden of its books, became lighter, the heavier sections, still full of books, revolved towards Bertie and showered their contents over him again. I watched, fascinated, as the process continued, with mathematical precision, until the entire bookcase was empty, and Bertie had disappeared under a pyramid of books about five feet high.

The noise and the shock, while it lasted, was like an earthquake, and it brought Edith hurrying in from another part of the house to see what had happened.

She saw me, still stuck in my chair with a half-finished tumbler of whisky, apparently alone in the room. Mindful of the earthquake and the thunderous noise, it appeared that perhaps I had just cast some sort of magic spell that had changed her husband into a pile of books.

For a moment all was still. Then a hand erupted from the apex of the pyramid, and waved to us. Then a voice in the pyramid announced, 'Here I am!'

Together we rapidly dug Bertie out, stood him up, and dusted him down. He turned to Edith and said,

'It's perfectly all right. There's no need to look so worried. We were just celebrating what clever fellows we are!'

Snettisham Beach, 5. September 1992

Introduction to Appendices 4 and 5

The following two Appendices are more closely related than at first meets the eye. For just as the properties of all numbers can be observed from studying their minimal forms, zero and unity, so the properties of all uncolourable networks can be seen in their smallest members, the dumbbell and the petersen.

The brief essay comprising appendix 4 was conceived in 1960 and written the following year, shortly after I had discovered the nature of formal mathematics. The reader will notice that it sometimes still uses the undiscriminating language of logic and philosophy, confusing consequences with theorems and demonstrations with proofs. It is one of many ways of constructing a brownian algebra of numbers, not necessarily the best, but the first I tried. It reduces the dozen or so postulates in current use to only two, but without any explanation of why these two unusal-looking postulates reproduce, in part, the behaviour of what we commonly regard as numbers. Exactly half of the elements in the system are not numbers in any language that has so far been conceived, but are their images or complements, the use of which greatly simplifies the necessary structure of definitions and rules of computation.

In the essay I point out that the whole interpretation can be reversed so that $\overline{\overline{}|}|$, $\overline{\overline{}|\,\overline{}|}|$, . . . are the numbers and $\overline{}|$, $\overline{}|\,\overline{}|$, . . . are their images. This leads to a conventional system more like common school algebra, in which

$$ab =_{\mathrm{df}} a \times b$$

$$\overline{\overline{a|}\,\overline{b|}}\,| =_{\mathrm{df}} a + b$$

and

$$\overline{\overline{a|}\,b}\,| =_{\mathrm{df}} a^b.$$

This interpretation may be preferred in practice, and the operative laws of the system itself are of course exactly the same. The difference is merely in the interpretation, which prefers the multiplicative ock to the additive ock.

Completing the calculus with a subtractive element, which also accounts for division, I did in a much longer document, now lost. This makes the full number field, with irrational, transcendental, and complex numbers. We introduce a second mark, called a *score* and written $\underline{}|$, so that $\underline{x}|\,x$ found together make an ock. This does not render the arithmetic easier to do than the way we do it conventionally, but increases our understanding of what is going on when calculations are conceived, since it and the less-complete calculus in Appendix 4 both use, for the first time, the complements of numbers as elements in the calculations. These not-numbers are of course not the negatives of numbers. Negative numbers are already familiar, and may be thought of simply as elements carrying an opposite charge, like a universe of anti-matter. All of this opens our eyes to what hugely complicated things numbers actually are.

My basic contribution to mathematics has been to discover the fundamental particles from which numbers and other, simpler, elements of mathematical systems can be made.

The essay in Appendix 5, written some eighteen years later, is much more sophisticated, and uses a pre-numerical formal arithmetic to solve a hitherto unsolved problem in 'conventional' mathematics. In it I show that the four-colour problem is a factorization problem of formal elements in the plane. Such problems, even in number theory, tend to be difficult, and the difficulties created in the elucidation of formal factorizations are exactly analogous to those that have taxed mathematicians, down the ages, in the discovery of theorems concerning numerical factorizations.

Just as we can make tests for prime numbers that identify them in almost all cases, but occasionally allow a composite number to slip through, so we can make tests for uncolourable networks that identify them in allmost all cases, but occasionally allow a colourable network to slip through.

The theorems that are true in general, such as the planar parity theorem, and that lead to a proof of the four-colour map theorem, are not really very complicated: it was picking my way through the pseudo-theorems, the ones that are true in nearly all cases, but are not true in general, that make so many traps for the unwary, and that took up most of my time exploring their ramifications. Kempe's 'fallacious' map-colouring algorithm is a case in point, and repays a more-detailed study, since there is in fact no map that a careful choice of the manner of its application cannot be made to colour.

Some of the pseudo-theorems are themselves of great beauty, and worth studying for their own sakes, and also for their undoubted applications in electronic apparatus that can reproduce what we call 'artificial intelligence'. What strikes the explorer is the enormous richness of the field of exploration, where even the theorems that are wholly true exist in a great variety of forms, as if we are classifying evolutionary species, and it becomes difficult, if not impossible, to decide what is the simplest and most representative description of what happens to be so. For the purposes of communication, a choice has to be made, but such a choice necessarily limits the presentation and ignores the unlimited delights of the exploration itself, that defies communication and can only be individually experienced by each new explorer, who comes therein to know much more than can ever be stated as mere matters of fact.

*London, 1996 April 23^d^*

Appendix 4

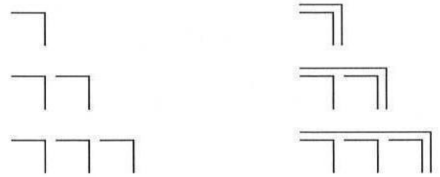
An algebra for the natural numbers

Prolegomena

In standard algebras of common arithmetic we cannot define the operators +, × either in terms of each other or in terms of the other elements 0, 1; $a, b, c, \ldots$. But an analysis of the nature of operation leads us to suppose that either of these essays may succeed, and, moreover, that in constructing an answer to one we shall have found an answer to the other. The form of the calculus described below has been dictated by the nature of the problem set here.

Definitions

The Arithmetic

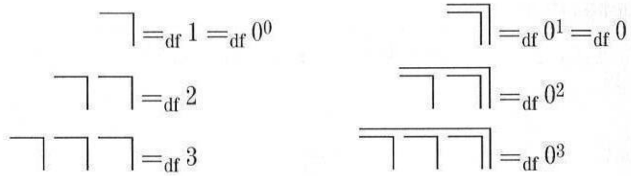


and so on are elements of calculation. Elements ⌉, ⌉⌉, . . . are called *1-elements* and elements ⌉‖, ⌉⌉‖, . . . are called *0-elements*.

The only relationship between elements is continence: a given element either is or is not contained in another given element.

Thus ⌉ abbreviates □, ⌉⌉‖ abbreviates ⧠□□⧠, $\overline{a ⌉ b ⌉ c}‖$ abbreviates ⧠a b c⧠.

Its Interpretation



and so on.

The element defined as 0^0 is called the universal element, and the element defined as 0^1 is called the null element.

A definition of an element may be regarded as an interpretation, and an interpretation of an element may be taken as its value. The values 0^2, 0^3, . . . are unequal but numerically indeteminate. They may be taken as logarithms to the base zero.

The Algebra

Letters a, b, . . . may stand for elements of calculation, and the values of such elements may be uncertain except that (1) $a =_{df} a$, $b =_{df} b$, . . . , and (2) if a is a 1-element then $\overline{a\,|}$ is a 0-element; if a is a 0-element then $\overline{a\,|}$ is a 1-element.

If and only if a, b, . . . are either (1) none of them 0-elements other than 0^1, or (2) none of them 1-elements, they are said to be homogeneous.

Its Interpretation

$$ab =_{df} a + b$$

$$\overline{\overline{a\,|}\;\overline{b\,|}\,|} =_{df} a \times b$$

The Calculus

Equivalence Postulates

1. Universe

$$\overline{\overline{\ \ |}\,|} = \quad ,$$

which means that the null element, as an element, need never be symbolized in the calculating forms.

2. Transfer

If and only if a, b, . . . are homogeneous and either t is not a 1-element or a, b, . . . are not 1-elements, then

$$\overline{\overline{a\,|}\;\overline{b\,|}\ldots|}\;t = \overline{\overline{at\,|}\;\overline{bt\,|}\ldots|}$$

Theorem

reflexion

$$\overline{\overline{a\,|}\,|} = a.$$

Proof

$$\overline{\overline{a\,|}\,|} = \overline{\overline{\ \ |}\,|}\;a \qquad \textit{transfer}$$

$$= a. \qquad \textit{universe}$$

Duality of Interpretations

By the principle of duality, the calculus will still apply if each element be given an opposite interpretation. [In this case, e.g $\overline{\overline{\ \ |}\;\overline{\ \ |}\,|}$ will represent 2 and $\overline{\ \ |}\;\overline{\ \ |}$ will represent 0^2. So the latter will look like a 0-element, but to make the system work we still have to call the former the 0-element. In fact 'one' and 'zero' refer to the calculus itself, not to its numeric interpretation, and designate expressions in its arithmetic that are of odd and even depths respectively. – Author, London 0200 02 02 1997.]

Standard Postulates

The following seven equivalence postulates of common algebra can now be proved.

1. $a + b = b + a$
2. $(a + b) + c = a + (b + c)$
3. $a \times b = b \times a$
4. $(a \times b) \times c = a \times (b \times c)$
5. $a \times (b + c) = (a \times b) + (a \times c)$
6. $a + 0 = a$
7. $a \times 1 = a$

The proofs are plain and need not be shown here.

The Representation of Powers

We can show that a^n may be represented by $\overline{n\,|}\,a$.

Demonstration.

$$a^n =_{df} (a)_1 \times (a)_2 \times \ldots \times (a)_n$$

$$=_{df} \overline{\overline{a\,|}_1\;\overline{a\,|}_2 \ldots \overline{a\,|}_n\,|}$$

$$=_{df} \overline{\overline{\ \ |}_1\;\overline{\ \ |}_2 \ldots \overline{\ \ |}_n\,|}\;a \qquad \textit{transfer}$$

But $\overline{\;\;}|_1\,\overline{\;\;}|_2 \ldots \overline{\;\;}|_n =_{\mathrm{df}} n$

Therefore

$$\overline{\overline{\;\;}|_1\,\overline{\;\;}|_2 \ldots \overline{\;\;}|_n}|\; a = \overline{n}|\, a$$

The Use of Brackets

The use of brackets is seldom necessary because of the continental form of the universal element itself. When there is a need to distinguish between interpretations, an expression may be broken by dots. Thus

$$\overline{n}|\; ab =_{\mathrm{df}} (a+b)^n$$

$$\overline{n}|\; a.b =_{\mathrm{df}} a^n + b$$

$$\overline{\;\;}|\; a =_{\mathrm{df}} 1 + a$$

$$\overline{\;\;}|.a =_{\mathrm{df}} a^0$$

The Null Power

Theorem

An element raised to the power of the null element takes the value of the universal element.

$$\overline{\;\;}|\,.\,a = \overline{\;\;}|$$

or, in common algebra,

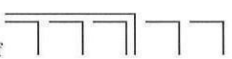
$$a^0 = 1.$$

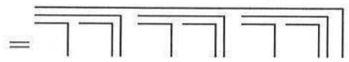
Proof

The null power theorem can be proved by analysis, but for this we must see clearly what happens. First we take an arithmetical example.

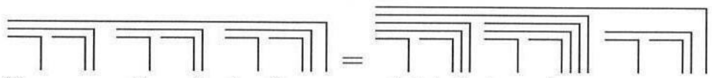
Theorem: $2^3 = 8$

Proof. To begin the calculation we take out the part of the expression representing 2 and distribute it into the inner elements of the part representing the power to which 2 is to be raised, i.e. 3.

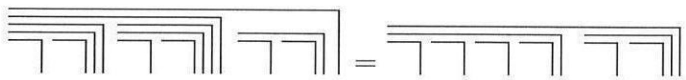
$2^3 =_{\mathrm{df}}$ 

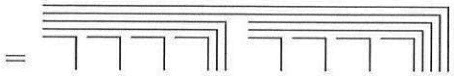
 *transfer*

Next we distribute one of the inner elements.

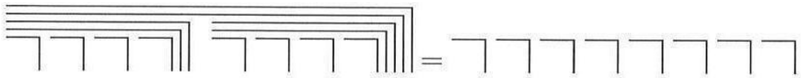
 *transfer*

Now we use the reflexion theorem and distribute again.

 *reflex*

 *transfer*

And finally we use the reflexion theorem to give the answer.

 *reflex*

$$=_{\mathrm{df}} 8$$

In this calculation the initial distribution is the step that determines the nature of the null-power theorem: for when the order of the power is zero, there are no inner elements in the expression for the power; yet the fact that the expression is still interpreted as a power means that the instruction for powers must be obeyed. Thus in

$$\overline{\;\;}|\,.\,a$$

the element a *must* be distributed into all the inner elements in the power-element; and since there are no inner elements in the power-element, the element a is by this means transferred into nonexistant space and ceases to be apparent there. But by the rules of the transfer the element a, having been transferred, can no longer appear in the space by the power-element, and so disappears completely. The power-element (which may now be interpreted as unity) is therefore all that remains.

Other Theorems for Powers

Standard theorems for powers can be proved without difficulty.

1. $(a^b)^c = a^{b\times c}$
2. $a^b \times a^c = a^{b+c}$

3. $$a^c \times b^c = (a \times b)^c$$

Proofs

1. $$(a^b)^c =_{\text{df}} \overline{c}\rvert\;\overline{b}\rvert\;a$$

$$= \overline{\overline{\overline{c}\rvert\;\overline{b}\rvert}\rvert}\rvert\;a \qquad \textit{reflex}$$

$$=_{\text{df}} a^{b \times c}$$

2. $$a^b \times a^c =_{\text{df}} \overline{\overline{\overline{\overline{b}\rvert\,a}\rvert\;\overline{\overline{c}\rvert\,a}\rvert}\rvert}\rvert$$

$$= \overline{\overline{\overline{b}\rvert}\rvert\;\overline{\overline{c}\rvert}\rvert}\rvert\;a \qquad \textit{trans}$$

$$= \overline{bc}\rvert\;a \qquad \textit{reflex}$$

$$=_{\text{df}} a^{b+c}$$

If $c = 0$ the transfer rule may still be applied. The reason will be apparent later. Alternatively

$$a^b \times a^0 =_{\text{df}} \overline{\overline{\overline{\overline{b}\rvert\,a}\rvert\;\overline{\overline{\;\;}\rvert\,.\,a}\rvert}\rvert}\rvert$$

$$= \overline{\overline{\overline{b}\rvert}\rvert\;\overline{\overline{\;\;}\rvert}\rvert}\rvert \qquad \textit{null power}$$

$$= \overline{\overline{\overline{b}\rvert\,a}\rvert}\rvert \qquad \textit{univ}$$

$$= \overline{b}\rvert\;a \qquad \textit{reflex}$$

$$=_{\text{df}} a^b$$

$$= a^{b+0}$$

3. $$a^c \times b^c =_{\text{df}} \overline{\overline{\overline{\overline{c}\rvert\,a}\rvert\;\overline{\overline{c}\rvert\,b}\rvert}\rvert}\rvert$$

$$= \overline{\overline{\overline{a}\rvert\;\overline{b}\rvert}\rvert}\rvert\;\overline{c}\rvert \qquad \textit{trans}$$

$$=_{\text{df}} (a \times b)^c$$

If $c > 1$ and $a = 0$ or $b = 0$, the expression represents a power >1 of 0 and is therefore numerically indeterminate. If $c = 0$ its cover may still be transferred since, in its context as a power-instructor, it is not a 1-element. Alternatively the case may be

regarded as special and the proof completed, as before, with the null power theorem.

We may note that if

$$\overline{\;\;}\rvert\,\overline{\;\;}\rvert$$

is always interpreted in the scale of powers instead of in the scale of integers, then

$$\overline{\;\;}\rvert\,\overline{\;\;}\rvert = \overline{\;\;}\rvert$$

and we have constructed the arithmetic for the primary algebra.

London, 5 May 1961

CAST AND FORMATION PROPERTIES OF MAPS

by

George Spencer-Brown

Cast and formation properties of maps

Appendix 5

Two proofs of the four-colour map theorem

The question of whether every plane map can be marked with not more than four colours so that no two regions with a border in common are coloured alike, has been the object of more false claims than any other unsolved problem of mathematics. Kempe's claim to have proved it in 1879 was accepted as valid, even by Cayley, until Heawood, in 1890, published a map that Kempe's algorithm failed to colour. The theorem has certainly been the object of more defective *professional* claims than any other: Kempe's attempt on it is now probably the most famous defective proof in history. Although Kempe was strictly an amateur, being (like Pierre de Fermat) by profession a lawyer, I include him with the professionals because he was certainly a better mathematician than most of them.

It is interesting to note that the first known reference to the theorem, in a letter dated 23d October 1852 from Augustus De Morgan to Sir William Hamilton, occurs at a time when Gauss was still alive. And in the present century Birkhoff, whose contributions to the literature rank in importance second only to those of Kempe and Heawood, considered that every mathematician at some stage attempts the problem, even if secretly, and at some stage believes he has solved it.

It has always been easy to reach a stage where the theorem is so 'obviously' true that one is tempted to lower one's standards of what constitutes a proof. Goldbach's conjecture is equally obvious if one examines the empirical evidence, but this does not mean it is proved. And of those who have had the courage to attempt the four-colour problem publicly, it has generally been the completion, the 'obvious' final step, at first considered to be of no great difficulty, that has defeated them in the end.

Once I had constructed the primary arithmetic in *Laws of Form*, I became aware that I had a technique that, suitably applied, would solve the map-colouring problem. I did not immediately apply it to the problem, because I felt that to make what would almost certainly be a difficult proof, in a completely novel system of mathematics, would occasion the hostility and disbelief of the more-superficial members of the mathematical profession, who would be unwilling (or unable) to learn the new discipline in which the proof was made manifest. This subsequently turned out to be the case.

In 1961 my brother came to visit me and I showed him the problem. He went away contemptuously, saying,

'Soon prove *that*!'

After a week he returned with the news that it was turning out to be more difficult than he had at first anticipated. But he had found an astonishing algebraic colouring algorithm. It applied my primary algebra to the map with certain relaxations to the rules that I cannot now recollect. I generalized the theory of it and applied it to what I call a pack-space, in which not more than n objects can be mutually related, and proved they could always be segregated into not more than n categories of objects that were mutually unrelated. I remember the proof was quite simple, showing that if there had to be more than n unrelated categories, it implied that more than n objects could all be related. I showed it to Bertrand Russell when I next visited him in 1962, and he could see nothing wrong with it. But the mathematics was so unusual, he said, that he thought we would have difficulty getting it accepted.

In 1975 my father died, and my brother the following year.

It was in these circumstances of despair and bereavement that I decided things could not get worse, so I might as well prove the four-colour theorem. My brother's algorithm was now lost, so I set about it the only way I knew how, using two elementary marks to set up the special case.

(The reader should turn to p177 for the much simpler proof I published in 1998.)

Part 0. Introduction

A word of introduction to the text that follows is in order.

The spelling is American English. It was prepared shortly after my appointment as Visiting Professor of Mathematics to the University of Maryland in the autumn 1979, and hand-written in my house in Cambridge England before I took up the Office. It is practically a first draft, there was no time to rewrite it more succinctly before I had to leave for Washington DC. The first to see it were Professors Frank Adams and J C P Miller. A xerographed copy of the MS was deposited by Professor Adams in the Library of the Royal Society in 1980 March 17.*

At Christmas last year I was invited by Mr Thomas Wolf, who had just completed his translation of *Laws of Form* into German, to visit him and his talented daughter Katharina in Vienna. It was agreed that in return I would show them one of my proofs of the map theorem.

I looked at the old MS and decided that the direct proofs there were all too difficult for anyone new to my modular methods, but that the indirect proofs implied by the contradiction between theorems 17 and 25 were not. I also considered that first-time readers would not wish to go beyond proving the 4-color theorem, so I sharpened the statement and proof of theorem 17, made theorem 25 more explicit and renumbered it 23, and ended the account with the two indirect proofs that result from the contradiction of these two theorems.

The remaining sixty pages of more-advanced mathematics, with some reluctance, I cut out entirely, though I did retain four of the coloring algorithms. In other respects the document, in parts I and II, is a generally-exact copy of the Royal Society MS, except for the omission of some redundant paragraphs and a change of the confusing terms 'cingle' and 'cingulus' to 'band' and 'circuit'.

I expanded the Notes in part III, and added three new names to the References in part IV, while deleting those whose work no longer seems to me to be relevant. In this context I should mention that Petersen attributed what has come to be known as 'his' graph to Sylvester, but it is too late now to change that name, and it will have to go down as one of the many false attributions, like Pell's equation and Sheffer's stroke, with which the history of mathematics is littered.

The reattribution of Petersen's graph to Sylvester, an English mathematician of great distinction, seems to confirm the 4-color problem as a particularly English proposition. Of course we must not forget Guthrie, the Scotsman who first drew attention to it, but even he, we are told, only thought of it while attempting to color a map of England. After him

* Spencer-Brown, G. MS. 734, Cast and formation properties of maps, 171p typ-ms photocopied.

the line of major English contributors, who either worked on the problem or assiduously broadcast it to the public notice, has been unbroken – De Morgan – Cayley – Kempe – Sylvester – Heawood. Perhaps it is fitting that their legacy should finally have fallen to me, finally in the sense that I have solved the problem, but of course not finally in respect of the rich new field of mathematics that my solution has lain open to exploration.

George Spencer-Brown
London, February 1996

Part I. The cast

I shall analyze the essential components of two kinds of definition of a map M.

M is completely defined *either* by indicating the location of each 1-cell or border b_i of M in the 2-space s in which b_i is imbedded, *or* by a proper coloration, called a *formation*, of each 2-cell or region of s defined by the borders. A coloration is proper when distinct regions on either side of a border are not colored alike.

I define the cast C of M to be the border-structure of M imbedded in a 3-space v. C is not a graph, since the nodes where ≥ 3 borders meet are not elements of C. Thus if a border of C with an origin in a 3-node is removed, the other borders at the node are subsequently counted as one.

C identifies with M in that, for each M in s, C is unique, but M may have a number of different formations $F_1, F_2, \ldots, F_m$ any two of which may be nominally and/or mathematically different while all define M and not any other map in s.

A cast C is *imbedded* in a space s if all of C is in s and no pair of borders intersect. Evidently all casts can be imbedded in the 3-space v.

I shall consider properties of casts and their possible formations in contexts of both s and v, and I shall present a number of new theorems. Each theorem representing what is, to the best of my information, a major departure from what was hitherto known, or at least from what was commonly realized in the context in which the theorem, if known or partially known, was usually presented, I shall mark with an asterisk *.

An outstanding property of a cast C is its connectivity.

Notice that regionality is a property of M but not of C, so that although M may have borders, C, strictly speaking, has none. What were borders in M, we call *links* in C. If C is imbedded in an appropriate s, it defines a map and all its links indicate borders, so there is no ambiguity in referring to a link of C as a border.

For a connected cast C, the connectivity k of C is defined as follows.

C is k-connected if it cannot be broken into two separate casts, or eliminated completely, by the removal of fewer than k links.

A nodeless map with just one border thus has a cast of connectivity $k = 1$, and *we further attribute the connectivity $k = 1$ to the casts of all nodeless maps with >1 borders*, even though the links of such casts are disconnected.

We distinguish cast-connectivity k and surface-connectivity c. *A surface s is c-connected if a maximum of c - 1 crosscuts can be made in s without separating it.*

A *crosscut* is defined as a continuous cut that begins and ends either by crossing itself or by crossing one or more other crosscuts. In the case of a bounded surface, a crosscut can begin and/or end at a boundary. A surface with no bound is said to be *closed.*

The formal plane is the closed surface with connectivity $c = 1$.

The question of the number of color-markings required for the formation of any kind of map in a surface s has been solved, or in principle can be solved, for all closed s with $c > 1$. The case of $c = 1$ has remained unsolved since 1852, when it was first conjectured that 4 color-markings were sufficient in this case. Thus

Theorem 0*. The four-color map theorem

Not more than four distinct colorings are sufficient to color any map, on the sphere or in the plane, so that no two regions with a common border are colored alike.

I have starred theorem 0 because, although the conjecture is well known, and solutions have from time to time been claimed, none of these claims has hitherto been substantiated.

The central purpose of this communication is to present a complete proof of theorem 0, and thus to settle this long-standing conjecture in the affirmative.

Other new theorems, not necessarily lemmata to theorem 0, but relevant to the title under which this communication is presented, are also offered, but without proofs.

[In the original document there followed theorems 1, 2, 3 that are not lemmata to my proof of the map theorem. In keeping with my present policy of not burdening the reader with inessentials before completing the map theorem proof, I have omitted these three theorems from the present document, but for referential reasons I have retained the numbering of the remaining theorems through to 22. – Author, Wien, April 2d 1997.]

We can draw any map with a series of draft-strokes that follow a set of crosscuts in the map-surface to their points of intersection.

If more than two crosscuts traverse any point, we shall find more than three borders of the map so drawn meeting at one node.

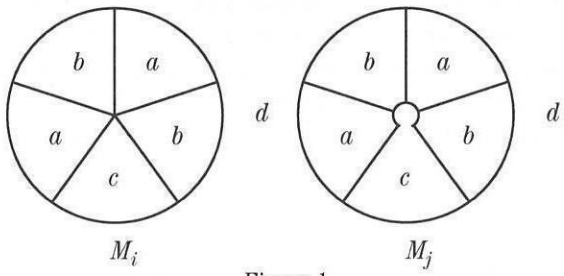


Figure 1

Consider the central node of M_i in figure 1, where > 3 borders meet. In such a case we can expand one of the regions into the point, as in M_j, so that not more than 3 borders meet at each node. We call this procedure, repeated at every >3-node until all nodes are 3-nodes, standardizing the map, and a map M in which all nodes are 3-nodes is called a *trivalent* or *standard* map. Its cast C is called a 3-cast. Clearly we may assume, without loss of generality, that M is connected.

Suppose we have colored $M = M_j$ as in figure 1 with four colors a, b, c, d. It is evident if M_j is properly colored, and we revert it, maintaining the same colors, to M_i, then M_i will be properly colored, since M_i can have no border-contact between regions that is not present in M_j. We have thus proved

Theorem 4. (Heawood)

If theorem 0 *is true for standard maps, then it is true for all maps.*

In map theory we are not interested in the lengths or shapes of borders or the areas of regions. These are the provinces of geography, geometry, and other branches of *metrical mathematics*.

Formal mathematics is concerned with the province in which line-segments have no length, and regions have no shape or size. Only the relationships between elements, which are not metrical in any usual sense, are preserved.

Nearly all the mathematics developed before 1900 was metrical, and formal problems were usually avoided or, in cases where they could not be ignored, inappropriate attempts were made to solve them with metrical techniques.

Thus the terms we have to employ in formal mathematics, and certainly in topology, often still carry metrical overtones. We speak of a surface with a hole as having a boundary 'curve', although in formal mathematics there is no such thing as a curve. We can speak of a boundary, and there we must end: the concept of shape is not necessary to the mathematics, and has not been introduced.

Topology is interested in metrical spaces in so far as it is interested in what kinds of metrical space can be topologically transformed into what kinds of other metrical space. In formal mathematics, apart from illustrations and applications, these questions do not concern us.

Pure formal mathematics is the science of space that has no metrical property.

Having discarded the metrical concept of size with respect to a map M, we are free to employ the term 'size' to denote any convenient *formal* property of M that seems appropriate. For example we can consider the size of M to be the number of its regions.

Clearly both the metrical plane and the spherical surface are topologically equivalent to the formal plane, and so, in proving the map theorem, we may confine our attentions to the formal plane.

We proceed to develop a set of related functions (f) of the standard map M in the formal plane, each function denoted by the form fM, or simply by f where the M is understood. Call

tM the numher of regions
uM the number of borders
vM the number of nodes

in M.

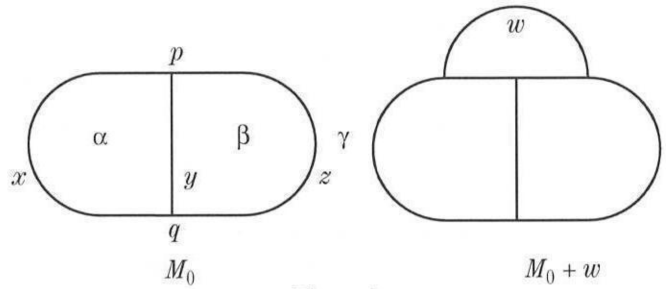


Figure 2

Figure 2 illustrates the smallest map, say M_0, in which each of these functions is positive. M_0 has 3 regions α, β, γ, 3 borders x, y, z, and 2 nodes p, q.

Since $c = 1$, the addition of a further stroke w (say) to M_0 must define another region and, with each further stroke, t is evidently increased by 1, u by 3, and v by 2. Thus by induction on M_0 we find

(1) $t = 1/2v+2 = 1/3u+2$
$u = 3/2v = 3(t - 2)$
$v = 2/3u = 2(t - 2).$

Indicate the formal size h of a region Q in M by hQ, where h is the number of borders, and thereby the number of nodes, in the perimeter of Q.

Consider now the size h of M expressed as the sum of the sizes of each of its regions. Suppose $M = Q_1 + Q_2 + \ldots + Q_t$. Then

(2) $hM = \Sigma hQ_i.$

Theorem 5. (Kempe)

Every plane map has a region bounded by not more than 5 borders.

Signally, $hQ \leq 5$ for some region Q in M.

Proof

Since each border in M is shared by two regions, and each node by three, we may use (1) to obtain

(3) $hM = 2u = 3v = 6t - 12$

Call $A = hM/t$ the average size of the regions in M. Then by (3)

(4) $A = 6 - (12/t)$

Thus $A < 6$ for all M and so $hQ \leq 5$ for some region Q in M.

The above proof, ostensibly for standard maps, is evidently sufficient to settle the question for all maps.

Consider a nonstandard map M' and its standard version M. Choose any region Q' in M' and its corresponding region Q in M. Clearly

$$hQ' \leq hQ$$

and so theorem 5 is true generally of maps in the plane.

Theorem 0 is an obvious corollary of the more general

Theorem 6*

If s is such that no map in s can have more than n mutually adjacent regions, then the range of the minimum number n' of distinct sets of mutually non-adjacent regions into which the regions of any map in s can be apportioned is $n' \leq n$.

Theorem 6 has already been decided** in the affirmative for all maps in surfaces with $c > 1$, and it is the proof of this last case, for $c = 1$, to which we shall now turn our attention.

In this case I shall effectively prove

Theorem 7*

Every standard map M in the plane is factorable into ≤ 2 maps M_1, M_2 so that both of M_1, M_2 are nodeless.

By factorable I mean that if M_1, M_2 are each drawn in separate transparent plane cells, the two transparencies can be amalgamated (what I call *idemposed*) to appear as M. The four-color map theorem follows as a corollary.

** by Heawood. See his brilliant and astonishingly beautiful paper of 1890. The reader might think, mistakenly, that my proof that $n \leq 4$ in the plane completes a proof of the pack-space theorem, but of course it does not, since the more highly-connected heawood surfaces suffer greater constraints than my ideal spaces, where the only constraint is how many objects in them can be mutually related, in respect of some given relationship r. I do consider the general pack-space theorem to be true, and it should not be beyond the wit of some ingenious contemporary to prove it again? I would put it forward as a conjecture, but it seems out of place to conjecture what I have already proved and then forgotten how I did it. The Heawood paper has a dream-like quality, since it is quite evident he did not reach his formulas by rational thinking. Fortunately he did not forget the dream, and so was able to present them as if he did. No proof of any but a trivial theorem has ever been arrived at by an act of will. Like a poem, either it comes or it doesn't. The difference is, in mathematics one may struggle for years before a proof comes: in poetry there is no struggle, because there is no aim.

Part II. The formation

We note that color is a physical not a mathematical term, and in mathematics we may give distinct values to distinct spaces and these values may be applied to the proper coloration of a map. We can easily prove

Theorem 8 (Tait)

A 3-coloration of the links of a planar 3-cast C is equivalent to a 4-coloration of the regions of a map M when M is an imbedding of C in the plane.

We note first that the *absence* of any mark in a mathematical space s can indicate a distinct value in s provided the spaces adjacent to s are all marked, so that to indicate a total of t values, not more than t - 1 distinct marks are required.

We note next that a combination of two distinct marks is itself a mark distinct from either, so that n marks considered primary can distinguish up to 2^n differently-marked spaces, the distinct markings ranging from no mark to all n primary marks.

Finally we note that in the application a different color can be assigned to each marking, so that the number of markings in the mathematics can always be equated with the same number of colors in the application.

It will sometimes be convenient in the representation of the mathematics to use actual colors as tokens of marks, in which case no color applied to a given space s may be considered nonetheless a coloring of s.

Proof of Theorem 8

Proofs of theorem 8 can be composed for casts of any valency, but it is sufficient for our purpose to compose a proof for 3-casts.

Since $4 = 2^2$ we see that to four-color a plane map ≤ 2 primary marks, say a, b, are needed. The 4 colors now identify with no mark, a, b, and $ab = ba = c$.

Suppose the regions of a plane map M are marked according to this scheme. The regions marked a, where a is solo or in combination, may be inclosed in a set S_1 of simple closed paths along the borders of M, and the regions marked b may be inclosed in such another set S_2. Illustrate the inclosure of the regions of S_1 by a broken line - - - - and that of the regions of S_2 by a continuous line ______ .

The borders terminating at each node in the cast C of M will now appear with one each of three kinds of marking, a, b, and ab. (Figure 3.)

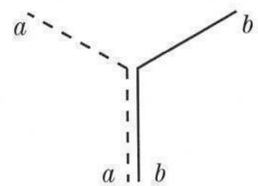


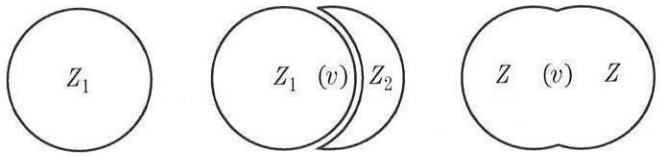
Figure 3

The converse is equally obvious, so the proof of theorem 8 is complete.

These a- and b-lines are considered to be in the links of C and are called a *formation* of, or over, or covering C. Thus when an a-line and a b-line are both in the same link they are both in precisely the same mathematical space, and when this is so we say they are *idemposed*. In the representation, they are drawn a short distance apart so that their composition can be distinguished.

For the purpose of illustration and calculation it is quicker and more striking to use different colors, say red for a and blue for b, and the reader will find it convenient to carry a red and a blue marker for these purposes. Also we may call a = red, b = blue, and $c = ab$ = purple.

Consider a color zone Z_1, say it is blue, in a map M. Suppose we idempose in M a second zone Z_2, of the same color, and bordering on Z_1. In the borders (v), say, where the inclosure-marks of Z_1, Z_2 idempose, the two zones Z_1, Z_2 are by formation indistinguishable and so, in the cast-formation of M, we may draw no mark between them (figure 4).



4.1 4.2 4.3

Figure 4

In this case we need to be more precise than ordinary language allows, for in saying we may draw no mark we do not wish to say we may-not draw a mark. Figures 4.2 and 4.3 indicate a mathematical equivalence where, in one expression, (v) is marked with two identical marks, and in the other expression, (v) is marked with none.

This observation supplies all we need to know, at this stage, about the mathematics of idempositions: *the idemposition of two identical formation-marks cancels both of*

them. It is the first and most fundamental principle of all formal mathematics, and we call it the *axiom of idemposition*.

We may note in passing that this one principle supplies *both* the initial equations of the primary arithmetic (p 10) in *Laws of Form*: applied over a partial bound of each mark, it supplies initial 1; applied over the whole bound of each mark, it supplies initial 2.

All future reference to the coloration of a *map* will mean the coloration of its *regions*, and to the coloration of a *cast*, since a cast has no regions, will mean the coloration of its *links*.

With the proof of theorem 8 we can forget about the map and attend to the formation over its cast. In what follows when we speak of a map M we shall be referring to its cast, and when we speak of its formations (F_i) we shall be referring to link- or border-formations, deficient, complete, or redundant over this cast.

We may now restate the four-color theorem as follows.

Every planar 3-cast is 3-colorable.

This is readily seen to be equivalent to theorem 7.

If F has n primary marks we say it is a formation of the n^{th} order, abbreviated n^r. The four-color theorem now becomes, every planar 3-cast can be covered by a 2^r formation.

If all 3-casts can be 3-colored then there is nothing more to prove. But by the theorem of Shannon (1949) we may need ≤4 colors to color a 3-cast, and the maximum of 4 colors is realized critically in what was known as the petersen graph (1891), which may be considered as a cast (figure 5).

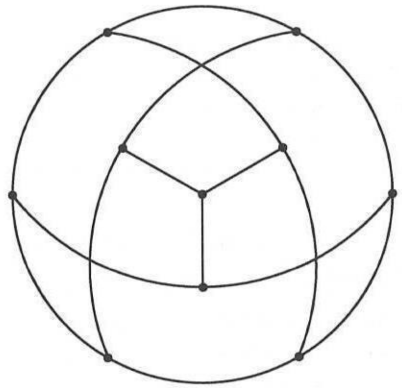


Figure 5. The petersen cast.

Over a cast C, call a formation F n-deficient if F covers all but n links of C, complete if F covers exactly all the links of C, and n-redundant if F covers C with n more links added. Normally we do not need to consider formations deficient or redundant except in non-adjacent links, so n can in these cases be equated with the number $\pm u_n$ of draft-strokes that must be subtracted from or added to C so that $C + u_n$ is exactly covered by F.

Call an entity E *critical* in respect of a property p of E if E is a *limiting case* of an E-type entity with p. For example, the petersen cast is a smallest (it is in fact *the* smallest) non-3-colorable 3-cast, and so is critical in respect of non-3-colorability. From this we know that the petersen cast, say P, can be covered with a 2^r-formation that is anywhere 1-deficient. Alternatively, we can say that $P - x$, where x is any link in P, is 3-colorable.

From this point we shall consider only $\leq 2^r$-formations, and unless the order of a formation is otherwise specified, it should generally be taken to be 2.

We represent borders in M or links in C by the final lower-case italic literals $p, q, \ldots, z$, and formation-elements over borders or links by the initial lower-case italic literals a, b, c, where in 2^r-formations $c = ab$.

These literals represent constants not variables in the calculi in which they are elements, and are thus representative of *arithmetical* (as distinct from algebraic) values. To represent *algebraic* values we may use the literals $d, e, \ldots$ and call them variables in the usual way, so that any variable in an algebraic formula can be replaced by an arithmetical constant and the formula is still true. We sometimes reserve the literal d for color-deficiency, so that dx will denote that x is a deficient or dead link in the formation.

To indicate a formation consisting of elements a, b, c over links x, y, z respectively, we write $x = a, y = b, z = c$.

Thus although cast-elements $x, y, \ldots$ are arithmetical in the sense of being constantly placed, and formation-elements $a, b, \ldots$ are also arithmetical, the former represent relatively algebraic variables that may take any of a number of values indicated by the latter. Once we accept the axiom of idemposition, we determine that a 3-cast cannot be formally presented except with at least 2 distinct elements, so in this respect the cast and its formation are identical.

Hence the 4-color theorem is more properly expressed, without any reference to color, as a general factorization theorem about planar casts.

The notations are summarized as follows.

$p, q, \ldots$	links of C
$a, b, c = ab$	elements of F
$d, e, \ldots$	algebraic variables in expressions representing the general properties of arithmetical elements of F in combination.

I = axiom of idemposition, notably,

I $\qquad ee = e^2 = \qquad = ɔ$.

To say a link of C is not formated, or is deficient of formation, is equivalent to saying it is formated with the blank or unmarked value. Although this value has strictly no name and is best represented by nothing, it is sometimes convenient to let it be called *ock* and noted in calculation by a broken circle like an inverted c, and on the understanding that anywhere it appears, ɔ can be removed from the mathematics by the equation ɔ = without in any way changing the value of the expression from which it is removed.

As a consequence of I, we can find the following elementary formulas in the arithmetic of formations.

$a = bc$
$b = ac$
$c = ab$
$ab = c$ $\qquad\qquad$ etc
$abc = \qquad = ɔ$.

Call them *equations of formation*.

Although a, b are considered primary and c composite, the mathematical relationship between a, b, c is symmetrical, so that any two of the three may be taken as primary and as the generators of the group in which they appear.

The idemposition-group of order 2^2 is a klein 4-group.

		a	b	c
		a	b	c
a	a		c	b
b	b	c		a
c	c	b	a	

Since in formal arithmetics there is no plus or minus, this group is in fact the formal analogue of the hamilton group of quaternions, with a, b, c replacing Hamilton's i, j, k. In formal arithmetics not only do addition, subtraction, multiplication, and division all condense into one operation, idemposition, but real and imaginary become interchangeable. I choose here what is invisible to be imaginary, which is opposite to the general convention in number theory, in which the real number 1 under multiplication is operationally invisible, i.e. is the ock of multiplication, although it is of course the visible building unit of addition, for which zero is the ock. My use of 'imaginary' in color-formations is thus different again from its use in Chapter 11 of *Laws*, where it is used to denote a paradoxical reentrant circuit. All 3-casts can be built from ≥ 2 sets of nonparadoxical reentrant circuits: the four-color theorem asserts that, in the plane, they can be built from just 2 such sets.

An idemposition-group of any order may be represented numerically by what I call a *reduction group*. This is the group of integers in which the exponents of all the primes are reduced to modulus 2, which has the effect of reducing all squares to unity. There is no difference between multiplication and division, and the product-quotient of any two numbers is called their *reduct*.

Interesting though these considerations are (they indicate, for example, a generic connexion between the four-color theorem and Goldbach's conjecture), they are not necessary to what follows and their analysis is reserved for later publications.

Suppose a formation F over a 3-cast C has primary elements a, b, and composite element $c = ab$. We have already seen that the primary elements will form simple closed paths in F, which we call *bands*. But examining the a-bands and the b-bands throughout F over C will show them to be composed of alternations of a, c and b, c elements respectively. Call such alternations *circuits*, and designate the type, measure, or *modulus* of each circuit by the missing element prefixed with a minus sign.

Thus we have *elements* (arranged in *bands*) a, b, c, and *moduli* (arranged in *circuits*) $-a$, $-b$, $-c$. If a, b are primary, a $-b$-circuit follows the exact path of an a-band, and a $-a$-circuit that of a b-band, so we can without confusion speak of the $-b$-circuit as the a-modulus, and the $-a$-circuit as the b-modulus.

Alternatively we use lower case Greek literals α, β, γ for moduli as tabled below.

elements	*moduli*
a	$-b = a = \alpha$
b	$-a = b = \beta$
c	$-c = -c = \gamma$

The importance of circuits and their moduli stems from the following theorem.

Theorem 9*

If two links w, y are in the same circuit K then a third link x running from w to y may be added to or taken from the formation.

Proof

If x is formated in the element e (say) missing from K, then by the equations of formation the e-band employed to formate or disformate x may close via x and either arc of K without canceling the formation over any link in K.

If w, y are stationed in a circuit K, and K is of the modulus μ, we may write $w - y\mu$. We

may omit the modulus and write $w - y$ if all we wish to state is that w, y are connected by some circuit-path and its modulus is obvious or immaterial. Similarly we write $w \mid y$ or $w \mid y\mu$ if w, y are not connected (via modulus μ).

If a border or link x originates and terminates at borders or links w, y, we call w, y *original* borders or links of x.

An operation Φ on a formation F is said to be of order zero if ΦF covers exactly the same cast-structure as F. Φ is called of order n if ΦF has just n links that either are not formated in F and formated in ΦF, or are formated in F and not formated in ΦF.

Elementary operations of order 0 all consist of idemposing with a circuit in F a band of the complementary color. No link is formated or disformated, and the color-markings of the circuit are transposed. If the circuit K in the formation F so operated on lies in the link x, and the modulus of K is $-e$, then we designate the operation on F by

$$Fe(x).$$

Elementary operations of order 0 are called *simple*, the rest *complex*, and all operations are designated by indicating certain of the links through which the operative band Φ must pass, and assuming Φ follows the appropriate simple circuit-path through any link that is not so indicated. Thus an indication of one link, together with the appropriate element of operation, is sufficient to define any simple elementary operation.

Reference to more than one link is necessary to define unambiguously a complex operation.

To summarize: an *operation* consists of the *idemposition* of a number of *formations* using the *equations of formation* to express the *resultant* or *consequent formation* in its minimal form. Typically, two formations are idemposed, and it depends on the viewpoint and motivation of the mathematician to decide which is the operator and which is the operand. Mathematically there is no distinction between them.

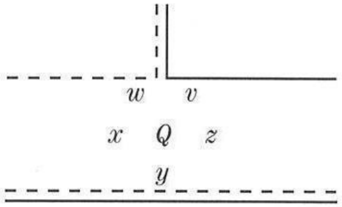
In an *elementary* operation the operator is one band of an element a, b, or c. If an operation Φ on F is *simple*, ΦF and F cover an identical cast-structure. Otherwise Φ on F is called *complex*.

A complex operation has a *real* and an *imaginary* component, the real being the simple component, and the imaginary component being where the cast-structure of the operand is changed.

In the *designation* of an elementary operation, say Φ on F, the element of the operator band is given, followed (optionally in brackets or parentheses) by indicators of the cast-structure where it is to be idemposed. Where no indicator is given, the operator is deemed to follow the unlabeled simple path of its complementary circuit.

In general we write A $\leftarrow$ B to denote that A has, or is a member of the set designated by, property B. In words, we can read 'A has B' or 'B belongs to A'.

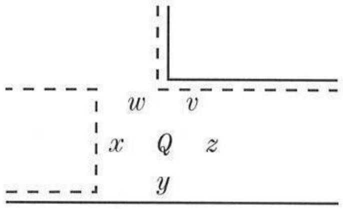
Suppose Q is a 5-region $vwxyz$ in a plane map whose cast is formated with a 2-deficient formation F, deficient over x, z in Q (figure 6.1). Suppose F has the property that v, y lie in the same circuit of modulus β. To indicate this we write $F \leftarrow v - y$. We do not in this case need to specify the modulus since there is only one modulus by which v, y could possibly be connected.



$$F = F \leftarrow v \text{-} y$$

Figure 6.1

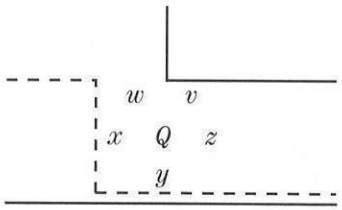
Choose $\Phi = a(vwxy)$ and now $F.a(vwxy) = \Phi F$ indicates the command to do Φ on F and also indicates the resultant expression (figure 6.2) when the operation is complete.



$$F.a(vwxy)$$

Figure 6.2

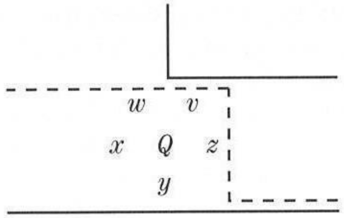
Should we wish to indicate that the simple component of the operation takes the alternative arc of the β-circuit we write $F.a(wx)$ and the resultant is as in figure 7.



$$F.a(wx) = F'$$

Figure 7

Figure 8 shows how, in a 5-region, 1-deficiency and 2-deficiency are interchangeable.



$$F'.aQ = F.a(wx).a(vwxyz) = F.a(vyz)$$

Figure 8

A formation is *trivially completable* if its deficiency lies entirely in a link or set of nonconsecutive links of a circuit R (or a series of circuits (R_i)), and the formated links of R (or each successive member of (R_i)) do not use all three marks. When a formation is trivially completable over C, we say the coloring of C is *solved*.

Theorem 10

A 1- or 2-deficient formation over a plane map is trivially completable, or can simply be made so, if its deficiency lies in a <5-region.

The proof is self-evident. 1-deficient <4-regions and 2-deficient 4-regions can all be trivially completed in their perimeters, and a 1-deficient 4-region that is not trivially completable can be made so by a simple operation through either origin of the deficient border.

Theorem 10 can be generalized for all casts, not necessarily planar, as follows.

Theorem 11*

A 2^r-formation over any 3-cast is trivially completable, or can be made so simply, if its deficiency lies in a <5-ring.

Thus no critical 4-colorable 3-cast can have a <5-ring. (Note that the petersen cast has 12 5-rings, and nothing smaller.) We can also prove

Theorem 12*

No critical 4-colorable 3-cast can be <5-connected.

This was first proved, for planar casts only, by Birkhoff (1913). He considered the plane map formed by imbedding the cast, and worked on a 4-marking of the regions instead of a 3-marking of the borders, thus at the same time making his proof less general and multiplying his labor by an unnecessary factor of 4. Astonishingly this method, which originated with Kempe (1879), was copied by subsequent investigators, Bernhart, Franklin, Ore, Winn, *et al*, and with this self-imposed handicap it is remarkable they managed to prove as much as they did.

I give a proof that is formally similar to Birkhoff's, which will serve at this stage to familiarize the reader with the usages of the formal mathematics I have introduced, as well as to show how very much simpler all these proofs become, and how much more general and obvious, when we employ the modular method.

Proof of theorem 12

Suppose the theorem to be untrue, and there exists a critical <5-connected 3-cast C that is not 3-colorable. Thus $C \leftarrow$ a 2-bridge, $C \leftarrow$ a 3-bridge, or $C \leftarrow$ a 4-bridge. We dispose of the cases where $C \leftarrow$ a 2-bridge or $C \leftarrow$ a 3-bridge first.

Procedure. Cut the links of the bridge and rejoin them to make 2 separate casts, C_1, C_2 as in figures 9.1, 9.2. Each of C_1, C_2 is evidently 3-colorable or smaller than C and thereby 3-colorable, and thus 2^r-formatable, by hypothesis. Formate C_1, C_2. Clearly if the color formations over x and x', and xy and $x'y'$, can be made to match, then C can be 3-colored. In either case, if they do not already match, they can be made to do so by a permutation. Whether or not C is planar is evidently irrelevant. (*Proof.* If C is nonplanar, any intersections of the links of the bridge when it is diagrammed in the plane can be removed until they are out of the diagram.) Thus in either case C is 3-colored and the proof is complete in case C is 2- or 3-connected.

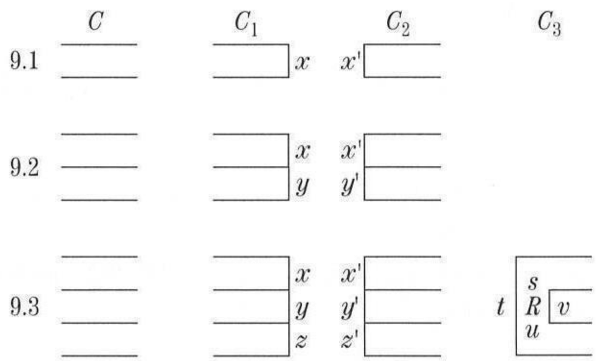


Figure 9

We now consider the remaining case, where C is 4-connected and so $\leftarrow$ a 4-bridge.

Again we cut the links of the bridge, rejoining them in the left-hand portion of the cast to make C_1, and rejoining them on the right-hand side in 2 different ways to make C_2, C_3 as diagrammed in figure 9.3.

Each of C_1, C_2, C_3 can evidently be 2^r-formated or is smaller than C and can by hypothesis be 2^r-formated. Formate each of C_1, C_2, C_3.

If the formation over x, y, z in C_1 can be made to match the formation over x', y', z' in C_2, then C can be formated. The only essential way they can differ is for (say)

$$\begin{array}{ll} x = a & x' = a \\ y = b & y' = b \\ z = a & z' = c. \end{array}$$

Clearly the formation over C_1, call it C_1F_1, must have $x - z\alpha$, and that over C_2, call it C_2F, must have $x' - z'\alpha$, otherwise $C_1F_1.b(z)$ will match C_2F or $C_2F.b(z')$ will match C_1F_1.

But if $C_1F_1 \leftarrow x - z\alpha$ we can disformate y (theorem 9) by $C_1F_1.b(y) = C_1F_2$ (say). Then $C_1F_2 \leftarrow x = a, y = ɔ, z = a$.

Consider now the formation over C_3, and call it C_3F. Suppose $R = stuv$ in C_3 with C_3F deficient over s, u.

We can make $C_3F \leftarrow t = a$, and now $v = a$, or $v = b$, or $v = c$.

If $v = a$, $C_3F.aR$ matches C_1F_2.
If $v = b$, $C_3F.a(v).cR$ matches C_1F_1.
If $v = c$, $C_3F.cR$ matches C_1F_1.

Thus C can be 3-colored in any case, and the contradiction completes the proof.

I have starred theorems 11 and 12 because hitherto they had been proved only for planar casts and the proofs had never been extended to 3-casts in general, whether planar or otherwise, so as to render their properties mathematically predictable without the need to consult particular cases. In fact, as we can readily see, all of the so-called 'reduction' proofs devised for plane maps can be extended, as I have extended the Birkhoff proof, to 3-casts irrespective of planarity, and it becomes a matter of practical interest to consider why, over a period of more than a century, otherwise competent mathematicians failed to see this.

The simple answer is that, in the absence of a suitable discipline, and a proper notation with which to communicate it, it becomes impossible to state theorems, much less to prove them, although their truth, on occasion, can still be recognized intuitively.

It may be a source of surprise to some readers to learn that Gauss (*Disquisitiones arithmeticae*, Leipzig 1801, article 16) was the first to state and prove the fundamental theorem of arithmetic, though earlier mathematicians must have recognized its truth. Euclid, whom we might otherwise have expected to prove it, did not possess a notation with which he could easily have stated it.

Formal mathematics had received no systematic treatment, or appropriate notation, before the first publication of *Laws of Form* in 1969. Without the principles, and the notation, enunciated there, and augmented here, many of the theorems, that are stated and proved quite easily in this communication, would be hardly conceivable.

In the absence of appropriate formalities, what has been presented in map theory hitherto is necessarily of an *ad hoc* nature, and what little had been proved is couched in theorems that are complicated and special, rather than simple and general. For example, a majorly restrictive concept was embodied in the term 'ring' (of regions), usurping the simpler, more appropriate, and more general term 'bridge' (of links or borders). Rings of regions are confined to maps. Bridges, in the formal mathematics of the networks in respect of which maps comprise but a specialized part, are universal. We may then reattribute the term 'ring' more simply to any connected path of links that would inclose a region in the appropriate imbedding-space.

One of the artificial difficulties created by such restrictive terms has been occasioned by the fact that, if the 4-color theorem is true, as I shall prove it is, then critical 4-colorable 3-casts do not exist in the species of casts that are planar, so we have no practical means of checking, if we confine our terms of reference to planar examples, any of the mathematical properties that we can prove such casts must possess.

Once freed from this restriction, we can point with delight to the petersen cast which, as I have already demonstrated, must exhibit *all* the properties that, by reason of its criticality, a critical 3-cast, planar or otherwise, can be shown to possess. We can thus employ it to illustrate past findings (such as theorems 11 and 12) and to suggest future lines of investigation.

It becomes clear that criticality, for discrete entities such as maps and map casts, is a duality-concept, and must provide at least two representatives in repect of each property p, one for p and another for its image or complement, *not-p*.

The petersen cast P is the smallest 3-cast that is not 3 colorable. Remove a link x from P and we have a largest 3-cast that is certainly 3-colorable. We may thus say that P itself is downcritical 4-colorable, whereas $P - x$ is upcritical 3-colorable.

It is evident that if the 4-color theorem were false, its falsity would have definite implications for the properties of upcritical 3-colorable planar 3-casts. I shall demonstrate exactly what these properties must be, and I shall prove that a planar cast can not have them.

Suppose we write A, B, or Γ to indicate, respectively, commands to perform simple operations on all the α-circuits, all the β-circuits, or all the γ-circuits of some formation F. As a result of these operations the changes to F will be only nominal, that is to say they will consist merely of permutations of the formation-marks that do not differ in any essential way, and whose formations are thus not mathematically distinct, from the original formation F. Thus the group formed by these operators will be a permutation group.

		A	*B*	*Γ*	*AB*	*BA*
		A	*B*	*Γ*	*AB*	*BA*
A	*A*		*AB*	*BA*	*B*	*Γ*
B	*B*	*BA*		*AB*	*Γ*	*A*
Γ	*Γ*	*AB*	*BA*		*A*	*B*
AB	*AB*	*Γ*	*A*	*B*	*BA*	
BA	*BA*	*B*	*Γ*	*A*		*AB*

As with all groups the identity operator is ock, we do not need to mark it and (see p 68) it is in some cases essential to leave it unmarked.

The permutations are tabled below.

operator	*elements exchanged*	*moduli exchanged*
A	$a \rightleftharpoons c$	$-a \rightleftharpoons -c = \beta \rightleftharpoons \gamma$
B	$b \rightleftharpoons c$	$-b \rightleftharpoons -c = \alpha \rightleftharpoons \gamma$
Γ	$a \rightleftharpoons b$	$-a \rightleftharpoons -b = \beta \rightleftharpoons \alpha$
AB	$a \rightharpoonup b \rightharpoonup c \rightharpoonup a$	$-a \rightharpoonup -b \rightharpoonup -c \rightharpoonup -a =$ $\beta \rightharpoonup \alpha \rightharpoonup \gamma \rightharpoonup \beta$
BA	$a \leftharpoondown b \leftharpoondown c \leftharpoondown a$	$-a \leftharpoondown -b \leftharpoondown -c \leftharpoondown -a =$ $\beta \leftharpoondown \alpha \leftharpoondown \gamma \leftharpoondown \beta$

Several implications follow. If F has no more than 3 circuits, one of each modulus, then a simple operation on F can have no effect whatever on its mathematical properties. The effects of such operations we call *nominal*, and the operations themselves we call *trivial*.

Each 2^r line-formation F is in general mathematically equivalent to each of its 6 distinct mark-permutations, which constitute no more than notationally different ways of representing F.

If formations F, G are mathematically so equivalent, we may write,

$$F \equiv_{\mathrm{m}} G.$$

Theorem 13*

If F has a total of k circuits of a given modulus $-d$ (say), and Φ is a set of simple elementary operations (d) applied to any n of them, and Ψ is the set of similar operations applied to the other $k - n$, then

$$\Phi F \equiv_{\mathrm{m}} \Psi F.$$

Proof

We introduce the sign T (say total) whose use is analogous, in formal mathematics, to that of Σ, Π in metrical mathematics. If a mathematical structure has n elements of a given kind c, which we can number from 1 to n, we write

$$\overset{n}{\underset{i=1}{\mathrm{T}}}\, c_i$$

to refer to all n of them.

Call the modulus $\text{-}d = \delta$ and say Δ indicates the set of simple elementary operations applied to some indicated total of δ-circuits.

Consider all k δ-circuits in F and idempose d-bands with n of them, which are then numbered $\delta_1, \ldots, \delta_n$. Number the remainder $\delta_{n+1}, \ldots, \delta_k$. Then

$$\Phi F \leftarrow \overset{n}{\underset{i=1}{\mathrm{T}}}\, \Delta\delta_i + \overset{k}{\underset{j=n+1}{\mathrm{T}}}\, \delta_j$$

Thus

$$\Delta\Phi F \leftarrow \overset{n}{\underset{i=1}{\mathrm{T}}}\, \Delta\Delta\delta_i + \overset{k}{\underset{j=n+1}{\mathrm{T}}}\, \Delta\delta_j = \overset{n}{\underset{i=1}{\mathrm{T}}}\, \delta_i + \overset{k}{\underset{j=n+1}{\mathrm{T}}}\, \Delta\delta_j$$

$$\Psi F \leftarrow$$

But

$$\Delta\Phi F \equiv_{\mathrm{m}} \Phi F$$

so

$$\Phi F \equiv_{\mathrm{m}} \Psi F$$

and the proof is complete.

Further use is made of T in expressions such as

$$\underset{\alpha}{\mathrm{T}}\, F = \tau\alpha F$$

to indicate the total number of α-circuits in F, or

$$\underset{\alpha,\beta\gamma}{\mathrm{T}}\, F = \tau F$$

to indicate the total number of circuits, of every kind, comprising F. We can also write $F \leftarrow \tau\alpha = k$, or $\tau\alpha = k$ with the F implied, to indicate that F has a total of k α-circuits. In general the lower-case Greek literal is employed unless we wish to refer, as in the previous theorem, to a subset. For $\tau\alpha$, etc, say 'tally of α, etc'.

Consider a map M in a 2-space s with $c \neq 1$. We may extend the concept of space-connectivity c to include zero and negative connectivity, so that spaces with $c < 1$ are disconnected, that is to say they have, when cleared of draft-strokes, a number >1 of

undefined regions. Call s fully utilized if M contains c - 1 draft-strokes that do not define a region. Suppose M fully utilizes s.

Draw a second map M' disconnected from M and wholly within a defined region of M. How many different colorings shall we need to make a complete formation over M'?

The answer is that we do not know, until we have decided the truth or otherwise of the 4-color theorem for a map in the plane. For whatever we decide here will be true locally in any other map, including maps in surfaces with a connectivity $c < 1$. Thus, *in so far as we do not know the properties of the plane, we do not know the local properties of any surface whatever.*

Write $a \sim n$ to denote the complete set of ordered terms from and including the lower (upper) bound a to and including the upper (lower) bound n. Write $\sim n$ to indicate all terms from and including some natural bound through n. For example, if n is a positive integer, $\sim n$ denotes all the integers $1, 2, \ldots, n$.

Theorem 14*.

If F is a first-order formation imbedded in the plane s and $F \leftarrow z, z'$, two arcs that can be idemposed by an operation Φ on F so that the resultant ΦF is also in s, then

$$\tau\Phi F = \tau F - 1 \textit{ if } F \leftarrow z \mid z'$$

and

$$\tau\Phi F = \tau F + 1 \textit{ if } F \leftarrow z - z'.$$

Proof

In a l^r-formation the bands are all of the same element, say e, and, since there can be no visible circuit $\text{-}e$, circuits are wholly imaginary. Thus – and | can refer only to connexion and disconnexion of bands.

By the planar property, each band separates s into two zones.

Suppose $F \leftarrow z \mid z'$. Imagine a 4-region $\mathrm{R} = tzt'z'$ deficient over t, t'. Then $F.eR$ displaces the deficiency to z, z' and the formerly separate bands are joined through t, t'. Clearly eR is a way of expressing Φ, so

$$\tau\Phi F = \tau F - 1 \text{ if } F \leftarrow z \mid z'.$$

In $F.eR$ exchange $t \rightleftharpoons z$, $t' \rightleftharpoons z'$. Call $F.eR = F_1$, for which rename $t \rightharpoonup t_1$, $t' \rightharpoonup t'_1$, $z \rightharpoonup z_1$, $z' \rightharpoonup z'_1$. Call eR = Φ and reapply Φ to give

$$\tau\Phi F_1 = \tau F_1 + 1 \text{ if } F_1 \leftarrow z_1 - z'_1.$$

The proof is complete.

For illustration, see figure 10. Notice from illustration 10.3 that theorem 14 is false unless the band-paths to be redistributed are actually imbedded in the plane.

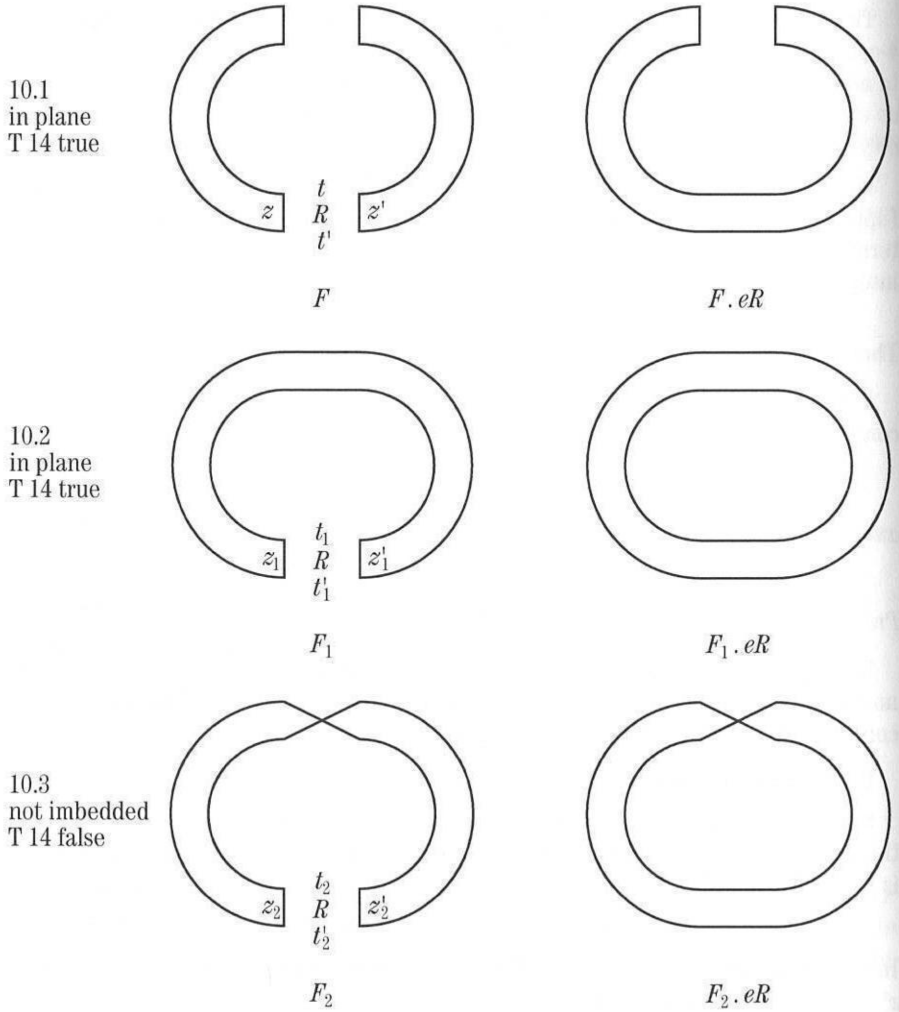


Figure 10

Theorem 15*. First order difference

If the operator-band Φ can be wholly imbedded in the plane s so that Φ stands in s as an extra band to its operand F, and F is a first-order formation in s having n distinct arcs standing immediately next n distinct arcs of Φ, and supposing the implementation of the operation ΦF consist of the idemposition of each of the n arcs of Φ with its corresponding arc of F, then

$$\tau\Phi F - \tau F \equiv n - 1 \ (mod\ 2).$$

Proof

Label the arcs of F in any order $z_1, \ldots, z_n$ and the corresponding arcs of Φ $z'_1, \ldots, z'_n$. Call $\Phi = \Phi_1, \ldots \Phi_n$ where each Φ_i stands for the idemposition of z_i with z'_i.

Now perform Φ on F in stages $\Phi_1, \ldots, \Phi_n$ to give $F.\Phi_1, F.\Phi_1.\Phi_2, \ldots, F.\Phi_1 \sim \Phi_{n-1}.\Phi_n = \Phi F$.

Clearly the extreme values for τΦF and τF are $k + n$ and $k + 1$, in either order, with k the constant number of bands untouched by Φ. Each application of a Φ_i is an application of theorem 14, and since Φ_1 reduces the total number of bands in $\Phi + F$ to τF, the congruence follows.

The proof is illustrated in figure 11.

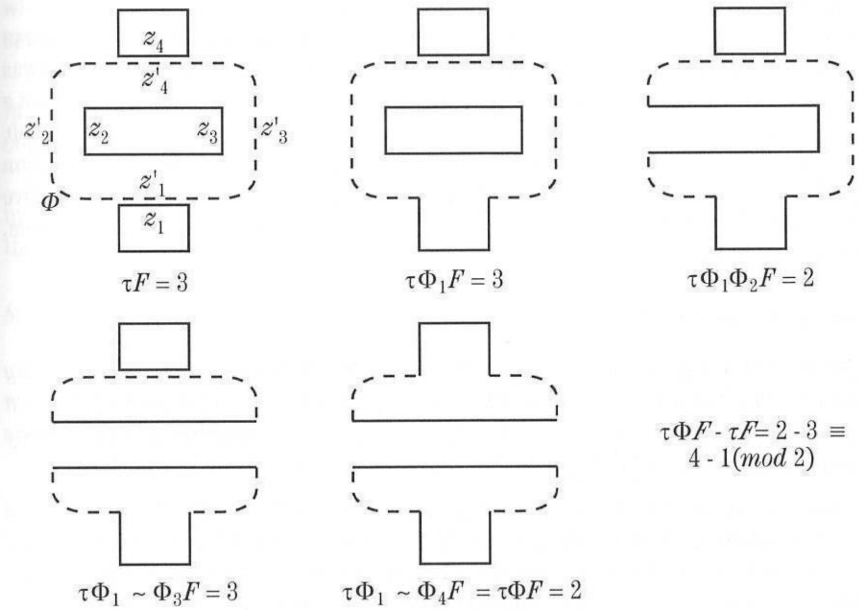


Figure 11

We see from the proof of theorem 14 that it is in general about the redistribution of links in the plane, however such redistributions may be brought about. With theorem 15 it is couched in terms of operations known or defined to bring about redistributions wherever they are applied, but the significance of either theorem extends beyond this.

First, by theorem 15, we see that the resultant of a multiple redistribution can be calculated by doing each of its constituent redistributions separately. The answer, as is intuitively obvious and as I shall show in the next proof, must be independent of the order in which its constituent computations are carried out.

Secondly we see that if Φ cannot be imbedded in the plane with F at any stage of the operation, then its application may achieve a re-routing of the band or circuit paths in F that is not a redistribution, so that theorem 15 might not hold in such a case.

All operations *re-route* the band or circuit paths. A *redistribution* is defined when, as a result of re-routing, bands or circuits are joined together or split apart.

Theorem 16*

The resultant of an operation is unique.

Proof

For every element e in the operator Φ executed on the formation F, each e-segment in F in the path of Φ is re-routed so that an e-segment is absent in ΦF where one was present in F, and present in ΦF where one was absent in F. Imagine, for each element e in Φ, we remove the present segments, suppose there are n, of e in F over the path of Φ, and store them elsewhere. The segments of any given element are by definition indistinguishable, so when we come to return them to ΦF we find exactly n alternative places for the n indistinguishable segments. There is thus no choice where to put them, and this completes the proof.

Theorem 17*. Indivisibility

Set a sole color-deficient link x between two b-colored bands of origin b_1, b_2. Call any possible set of a-colored bands in F of C - x a trail t projected from x. If t, when projected from some x in C, can have two or more seperate factors, of the same modulus, and C is uncolorable, then C is not a smallest uncolorable of its kind.

Comment. We need consider only nontrivial (i.e. $\geq$ 5-connected) casts. The 1-connected dumbbell configuration in figure 12.1 is not trivial, since it is the basis of all uncolorables, but it is not a cast, since it does not define a map in any space. Call the trail-factors touching both of the original blue bands of x, *connective.* If C is a cast, then at least one trail-factor must be connective.

Theorem 17 is evidently correct in case C is nonplanar, since the petersen P is known to be smallest and unique. No nontrivial operation on P can break t into two or more distinct factors of the same modulus. Operating trivially on both origins of x merely changes the coloration of t. See theorem 13. We discount all instances of this kind, calling the trails 'false' since they are merely complements of the true trails.

In the wording of theorem 17, the phrase 'of its kind' refers to the imagined "possibility" of a smallest planar uncolorable. We aim to establish the truth of theorem 17 in general, whether 'of its kind' means 'planar' or 'nonplanar'.

Call a lone trail-factor *divisible*, *reducible*, or *composite*, if, from some projection in C, it can be broken into two or more trail-factors, not false-trail factors, of the same modulus, covering the same number of nodes, not counting the origins of x. Call it *indivisible*, *irreducible*, or *prime*, if it can not be so broken.

First proof of theorem 17

Theorem 17 is axiomatic.

That is to say, any least factor of any particular kind (the kind not itself a determinant of divisibility) in any arithmetical system whatever, numerical or (as in this case) otherwise, must self-evidently be a prime factor. The reason is, that if not, there must be a smaller factor of that kind, contradicting the original requirement.

The theorem is adequately evident in the case of number systems. It is equally evident in the case of color-circuit systems. Clearly if a smallest uncolorable of any kind has a connective factor, say K, then K is also smallest of its kind and thus prime.

Although we can see that it is correct without proof, we can in fact *prove* theorem 17 by a rather unusual form of *reductio*, and in view of the central importance of the theorem, we will proceed to do so.

Second proof of theorem 17

For this we invert the statement of the theorem to say that no two or more seperate whole factors, of the same modulus, can together make a true trail of a smallest uncolorable cast of any kind.

Call a connective trail-factor that, by itself, does not, after any manipulation, supply a modulus to color x, a stop-factor, or zero-factor, designated by a figure zero. Call a trail-factor that, by itself, after manipulation if necessary, does supply a modulus for x, a go-factor or one-factor, designated by a figure one. Nonconnective factors, sometimes called *weak*, do not fall in either category, but do fall in the category of factors that are not 0-factors.

Call a trail t *intrinsic*, or *true*, if it contains a prime 0-factor, alone or in composition, outlined in red. Call it *extrinsic*, or *false*, if such a factor is outlined as a γ-modulus – i.e. if the red covers the blue-complement of this factor. Notice that if t defines no 0-factor, either intrinsically or extrinsically, this distinction is arbitrary and immaterial.

My original plan was to decide an arbitrary standard coloring from which to project the trail, with both origins of dx in pure blue. It was only after discovering that a prime 0-factor could not always be factored out in red unless a not-0-factor was draped in red over one or both of the x-origins, that I was forced to change the plan for formations with a 0-factor in residence, and thus for the subclass of all uncolorable casts. For the colorables, and the subclass of these that are planar, when we cannot relate them to a resident 0-factor, we are free to standardize projections of them as we wish, and it is convenient to standardize them by my original plan, i.e. make both origins pure blue, as in figures 13.1, 13.2.

Call a trail-factor *pure* if it contains all of a given kind of factor, 0 or not-0. Otherwise call it *mixed*. In the equations that follow, the 1's and 0's in the left-hand side refer to pure factors, but the 0's in the right-hand side must sometimes be mixed, referring to the stop- or go-quality of the resultant trail as a whole. And the '=' in these equations does not mean 'is equivalent to', but means 'can be equivalent to'.

Redefine pure 0-factors and pure 1-factors by inverting the original definition. The original definition looks at what effect they have, as seen from themselves: either stopping or being unable to stop a modulus for x. The inverted definition will note what effect they might have, looked at from elsewhere.

As seen from itself, and by itself, a 0-factor can stop a coloration of x, a 1-factor can not.

As seen from elsewhere, a 0-factor can forbid the entry of either a 0-factor (figure 15) or a 1-factor (figures 15 and 14.1) if C is forbidden to go from colorable to uncolorable. And a 1-factor in these circumstances can forbid the entry of only a 0-factor (figure 14.1).

Since some two connective trail-factors, of whatever kind, can always = 1, the figure 1 on the right-hand side is irrelevant, and not marked there. Our inverse distinction, that must exist, between 0-factors and 1-factors, can be determined only by = 0. We have seen

1	$0 + 0 = 0$	(Figure 15, $F - o + o = U - x$)
2	$0 + 1 = 0$	(Figure 15, $F - q + q = U - x$)
3	$1 + 0 = 0$	(Figure 14.1, $F - o + o = J_5 - x$).

Suppose the fourth equation

4	$1 + 1 = 0$

were a possibility? If so, we would lose the distinction between 0, 1 in terms of what they forbid. For now each would forbid both.

The first three equations cannot be denied, because examples of them all exist. The fourth must be denied, it is impossible because it would contradict the inverse distinction that must exist between 0-factors and 1-factors.

This completes the proof.

Corollary to theorem 17*

Every least 0-factor is irreducible.

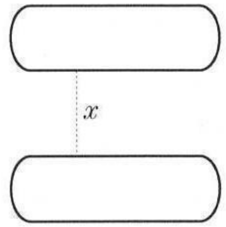


Figure 12.1. The dumbbell uncolorable.

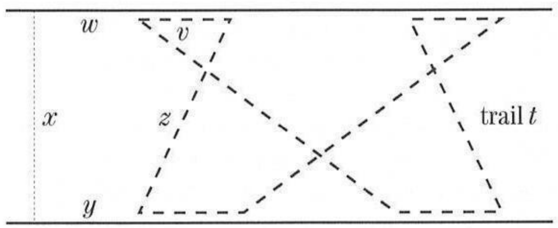


Figure 12.2. The petersen uncolorable.

In the petersen uncolorable, with trail t as defined in figure 12.2, there is no projection from which t can be split into separate factors, so we say that t is a least 0-factor o. If we call o_0 the absence of any trail whatever, as in figure 12.1, we can call the trail in figure 12.2, o_1, suggesting there might be other irreducible 0-factors o_2, etc. Whether or no such other irreducible 0-factors (o_n) exist, is immaterial to the proof I shall present here, since theorems 17, 23 remain valid however many kinds of irreducible (o_n) there might be.

[In a later communication (the Ross Ashby Memorial Lecture, University of Vienna, April 10th 1996) I show that o_1 is unique, and thus the sole relevant factor of all uncolorables, which leads to a proof of the color theorem that is almost disappointingly short. But that proof can scarcely be appreciated without this earlier account, first written and published in 1979, of the methods and theorems that preceded it, and led to the original proof I announced in 1976. – Author, editing this document in 1996]

Theorem 18*. The planar parity theorem

If F is a second-order planar formation, 3-cast, and Φ is simple, then

$$\tau F \equiv \tau \Phi F \ (mod\ 2).$$

Proof

Since the properties of elemental marks, with respect to one another, are symmetrical, it is sufficient to prove the theorem in the case of a simple elementary operation Φ over one circuit K, of any modulus, in F, and from this case the double generalization must follow.

Suppose K is of modulus γ. In the notation we have adopted, the other two moduli will now appear symmetrical with respect to one another, and it will be recalled that α-, β-circuits follow respectively the paths of a-, b-bands in any F.

Consider first the case where K is the only γ-circuit in F. In this case, by theorem 13, the application of Φ merely exchanges the α-circuits with the β-circuits (and thereby the a-bands with the b-bands) throughout F, so $\tau\Phi F = \tau F$ and the congruence is true trivially.

We note in this case that the subsequent identity of the paths of the β-circuits in ΦF with those of the α-circuits in F, and of the α-circuits in ΦF with those of the β-circuits in F, is sufficient to determine that an α-circuit is redistributed when, and not unless, a β-circuit is redistributed. For consider a further application of Φ to ΦF. In ΦF all the α-circuits of F have become the β-circuits of ΦF and all the β-circuits of F have become the α-circuits of ΦF. Since $\Phi\Phi F = F$, the further application of Φ to ΦF must do to the β-circuits of ΦF exactly what the original application of Φ did to the α-circuits of F, and to the α-circuits of ΦF exactly what the original application of Φ did to the β-circuits of F, since the paths of the one in the one case have become the paths of the other in the other.

Thus whatever redistributions Φ makes to α, in either case, it makes *in reverse* to β.

Now consider the case where K is not the only γ-circuit in F. The local properties in the neighbourhood of K remain exactly as in the previous case, and since F is planar, theorem 14 must apply to each redistribution of the a-bands and the b-bands considered separately as 1^r-formations in the plane, and so the redistributions made by an application of Φ in this case correspond exactly to redistributions made by an application of Φ in the previous case. The only difference is that there may no longer be strict reciprocity between joining (disjoining) in the one case and disjoining (joining) in the other, so the following possibilities are now determined.

	Resultant of Φ	
on α*-paths*	*on* β*-paths*	*on* τ
not redistributed	not redistributed	+ 0
disjoined	disjoined	+ 2
disjoined	joined	+ 0
joined	disjoined	+ 0
joined	joined	– 2

Thus Φ can never change τ except by a multiple of 2, and the proof is complete.

This beautiful elementary theorem is the key to the solution of the color problem, since it demonstrates, historically for the first time, a relevant and crucial difference between planar and nonplanar formations.

Before proceeding, it is appropriate, at this point, to present an analysis, in formal terms, of just what was wrong with Kempe's original and defective solution.

Kempe's mistake is to this day still mysteriously referred to as a 'Kempe-type error', with no consciousness whatever, on the part of the mathematician who so describes it, of precisely what mathematical canon was broken. I explain below just what a Kempe-type error is.

Kempe (1879) considered an expression (which happened to be a formation), call it E, and envisaged two different operations, say Φ and Ψ, each of which, when applied to E, resulted in a certain desired consequence c. From this he falsely concluded that the application of Ψ to ΦE would also result in c. In signals

$$\Phi E \leftarrow c$$

and

$$\Psi E \leftarrow c.$$

Therefore (*non sequitur*)

$$\Psi\Phi E \leftarrow c \qquad \text{(false conclusion).}$$

Example. Set $a = 1/2$ and $b = 1/6$ and consider $E = 30$. Write ΦE to be the product of E by a and ΨE to be the product of E by b. A Kempe-type error would be to suppose that, since both ΦE and ΨE are integers, then $\Psi\Phi E$ must also be an integer.

Translated thus into metrical arithmetic, the mistake is absurdly obvious. But Kempe had strayed into formal arithmetic, where nobody knew the rules, or even that such a discipline might exist.

Theorem 19*

The parity-consequence of a complex operation on a planar formation is determined by the situation of its imaginary component.

This is no more than an evident definition of fact, since the real component of the operation is by theorem 18 not parity-changing. Thus whether the operation as a whole is parity-changing or not must depend on the imaginary component.

Though obvious, the theorem is of great practical usefulness, since it gives rise to the following set of corollaries.

Theorem 20*

The parity-consequence of all examples of a given description of complex operation on a planar formation is determined by the parity-consequence of any particular example.

This is obvious, again by definition, since a class of description is decided by its complex operations.

Readers may compare theorem 16 in *Laws of Form*, where it was first stated and proved, for all algebras, that expressions are equivalent if and only if they take equal values in every case of one distinct variable. In either instance, a property of the set is determined by finding that property in one of its members.

Thus all we need to do to ascertain the effect of any complex operation on parity is to construct an easy and convenient example of it, and the answer in this particular case is the answer in the general case. There follow subcorollaries

Theorem 21*

If the only imaginary component of Φ is one by which an unformated link is formated between differently-marked origins in F, and both F and ΦF are planar, then

$$\tau F \equiv \tau\Phi F \ (mod\ 2).$$

That is to say, in this case the operation is not parity-changing.

Theorem 22*

If the only imaginary component of Φ is one by which an unformated link is formated between similarly-marked origins in F, and both F and ΦF are planar, then

$$\tau F \not\equiv \tau\Phi F \ (mod\ 2).$$

That is to say, in this case the operation is parity-changing.

Theorem 23*

If Q is a 5-ring vwxyz in a planar cast C and F of C is 1-deficient at x, then the number of factors in the trail t can be changed, or C can be colored.

Proof

Set $y = b, z = a, v = c, w = b, x$ uncolored. Do $\Phi F = F.a(xyz).b(xy).c(wxy).a(z).c(w)$.

The Φ-component $a(xyz)$ is parity-changing (an elementary experiment on any appropriate F will demonstrate this). The remaining four components are chosen to restore the original coloring of Q without a further parity-change, and in such a way that the parity-change of the first component will manifest, if Φ can be completed, only in t. It is evident that any joining or disjoining of circuits that prevents the completion of Φ will provide an immediate modulus for coloring Q, and if no such modulus at any stage exists (as it cannot if C is uncolorable), then Φ must be completable.

An illustration of the operation of Φ is provided in the Notes.

Theorem 23 is seen to contradict the corollary to theorem 17, that any least 0-factor must be prime: for if such a factor f were to exist in the plane, either C would be colorable or f would be further factorable.

Thus uncolorability in the plane is self-contradictory, and this first proof of the four-color map theorem is complete.

A second and more elegant proof, that does not require the theorem 17 corollary, is achieved by the method of descent. By theorem 23 the trail of an uncolorable planar C with respect to x can be broken, so by theorem 17, C cannot be a smallest such uncolorable. Eliminate redundant factors to make a smaller uncolorable, call it C', and then repeat the operation of theorem 23 on C' to break the trail again. Carry on until a cast is reached such that it and every smaller cast is obviously colorable, and the contradiction completes the proof.

Part III. Notes

The first recorded reference to theorem 0 is in a letter dated 23d October 1852 from Augustus De Morgan to Sir William Rowan Hamilton.

Kempe's 1879 proof of theorem 5 is complicated by the fact that he did not standardize the map. Nobody thought of standardization until eleven years later, when Heawood (1890) proved that the problem could be confined to maps with connected 3-casts.

Tait (1880) gave a complicated proof of theorem 8 and thought he had proved theorem 0. He assumed the false hypothesis that every planar cast has a circuit that traverses each node just once.

Although solutions to the great unsolved problems of mathematics are attempted by hundreds of amateurs every year, in only one case that I know of has a professional (Lindemann, in relation to a theorem of Fermat) ever risked his credibility by publicly claiming a defective solution to any such problem, other than the four-color problem, which alone has ruined more professional reputations than all the other problems put together. The following professionals since Kempe have all publicly claimed proofs that are either incorrect or incomplete. Tait, 1880 (false assumption of the existence of a hamilton circuit), Haken, Appel, and Koch, 1976-7 (failure to prove completeness of 'unavoidable' set of configurations to which they added more members in later drafts. Nowhere in their long and often irrelevant account do they provide the evidence that would enable the reader to check what they say. It may, or may not, be "possible" to prove the color theorem the way they claim. What is now certain is that they did not do so). The distinguished graph theorist G A Dirac told me that Haken, with Shimamoto in 1971, had also made an earlier false claim to a proof. There may be others, but even so, to find no less than seven false claims, or more if we include the multiplicity of claims made by Tait, in less than a hundred years, among six mathematicians, to the solution of one problem, shows just how unique this problem appears to be, in its capacity to lure almost everyone, at some time, into thinking he has a proof when he hasn't. (Commendibly, John Koch withdrew his name from the 1977 claim.)

A proof must supply clear instructions to the reader and predict correctly what will happen when he, or she, follows them. Euclid gave clear instructions to you and me, and correctly predicted what would happen when we followed them. Kempe gave clear instructions to Heawood, but failed to predict correctly what Heawood would find. HAK supply no instructions and no prediction, thus failing to meet either of the necessary criteria. You might ask, how could such a thing happen? Well, it happens all too frequently, but usually not so publicly. Of course they did not set out with the dishonest

intention of claiming a proof they did not have. No one but a fool would do that. They must have felt certain that what remained to be done would be easy, and that they could do it between the announcement and the publication. And when they found it wasn't and they couldn't, there was no way to save face except to brazen it out. Don't take my word for it. Have a look and see for yourself. You might have thought, somewhere in that mass of obscurely-written jargon, there might be a proof lurking. I am afraid you will be disappointed. There is nothing sensible there, other than unsubstantiated claims and non-sequiturs, and it does not take a very great intelligence to detect this. Somebody will have to take a look sometime, and discover the embarrassing truth, that not only is no proof to be found in what they published, but there is not anything that even begins to look like a proof. It is the most ridiculous case of "The King's New Clothes" that has ever disgraced the history of mathematics: the imaginary "proof" has been pumped out with propaganda and millions of dollars of misguided public money, and none of those involved now dares to admit that nothing of any value has resulted. Dirac's summary report should serve as the last word: 'Not a proof, even if the computer-work were correct, which it isn't.'

Kempe (1879) also proved (in effect) theorem 10. In my notation, theorems 8 and 10 are both self-evident. HAK stated, incorrectly, that Kempe standardized the map, thus exposing their illiteracy in not having read either Kempe or Heawood.

When I first devised the operation Φ in theorem 23, called a parity-pass *pp*, I knew of no formation on which it could be completed. I later discovered it could be done on Errera's famous map, call it E, in which all Kempe chain operations fail from the 5-region he chose. A parity-pass on E is illustrated in figure 13. Of course E solves simply by $E.c(w).b(wvz)$ leaving x on a β-modulus.

The full panoply of the parity-pass in theorem 23 is not strictly necessary, since the first component of Φ is parity-changing and recreates a 1-deficient Q. This component can in fact be seen as a Bernhart transformation on $C - x$, notably the transposition of a pair of nodes. But the *pp* offers us the bonus of coloring Q when it can not be completed, which Bernhart (1947) does not.

The definition of 'trail' in theorem 17 is adequate for our purposes there, but for greater precision it is necessary to state it in terms of circuits rather than bands. For complete generality, we may define the trail as the set of circuits in excess of the number required to avoid the possibility of x being on a modulus. If x is on similar colors there must be a minimum excess of 1, otherwise x would be just a dumbbell link, so in minimal cases the two definitions will define an identical situation. In general the number N of trail-factors is given by the equation $N = \tau F - 4$.

A proof procedure can be recounted as a formal dialogue between an instructor and a student. Convention has it that the author of the proof must play the part of instructor, the reader that of student. Although any proof is the recognition of what we knew already from the terms in which the theorem is stated, the play determines that, while the author (or instructor) is supposed to point this out, the reader (or student) is supposed to counter with all possible incorrect modes of thought, called 'objections', which the author is then in turn supposed to counter one by one.

Even Gauss (*Disquisitiones*, English edition 1966) became impatient with this charade, and remarked (*p*. 45), 'It is beyond the purpose of our investigations . . . to explain one by one the particular artifices that become familiar to anyone working in this field.' In short, since the number of ways of thinking incorrectly is infinite, a proof would never end if the reader could not at some stage be relied upon to experiment with the system in which the proof is made manifest, and discover what any normal person who so experiments will come to see.

Considered more directly, a proof is a way of making what appears less evident, appear more so. A certain standard is required for what we call validity, notably that, however it appeared before the proof, the evidence after the proof must appear absolute, like that of a self-evident proposition or axiom.

Thus the nearer to an axiom a proposition is, the harder it is to prove. And when, in the limit, it *is* an axiom, i.e. totally self-evident without proof, proof becomes impossible. Theorem 17 was hard to prove, because the theorem was so close to axiomatic already.

It follows that a perfect proof must make the required-to-prove proposition so abundantly evident as not to require proof. It must thus destroy the problem, and thereby the necessity for proving the proposition. A perfect proof is thus self-destructive in the sense of rendering itself unnecessary. If you come to see a theorem as self-evident, even if you needed a proof to do so, and do not then hark back to the incorrect modes of thought that blinded you to its self-evidence, you will forget why you ever needed a proof, and thus eventually forget what the problem or the proof was.

In "normal" mathematics, such as Euclid's proof of the infinity of primes, the answer, though certain, still does *not* appear *entirely* axiomatic, although, to a being with perfect vision, it must. The reason we can perform it again and again, like a favorite piece of music, and still feel edified, is because we constantly slip back into the habit of looking at the question incorrectly. After all, if a fact is so by definition (and all mathematical facts are so by definition), there must be something wrong with us if we cannot immediately see it so. Thus "normal" mathematics, like "normal" music, continues to 'make sense' because this wrongness of vision persists, and we derive pleasure by flipping between the 'wrong' and the 'right' appearances, between the tension and its resolution.

With *perfect* mathematics we make the flip irreversible. For example, reexpressing it in my notation, Tait's theorem becomes so obvious that proof is unnecessary. By

changing the notation, I have eliminated the flip between right and wrong. We flip over to 'right', and then we forget about ever flipping back to 'wrong', and Tait's and all other proofs of his theorem become unnecessary and we forget them.

"Perfect" music would resolve all the musical tensions before the piece began, thus spoiling the pleasure by rendering the piece as traditionally performed unexciting and unnecessary. The object of music is not to get rid of itself, but to create an imbalance and prolong the pleasure of fixing it. "Normal" or "traditional" mathematics has all so far followed exactly the same course as "normal" music, prolonging the pleasure of performance by not making the final resolution either so obvious or so final that we lose all desire and reason to perform it again.

In times of crisis it may be necessary to resolve mathematical tensions, i.e. arrive at correct answers, quickly by by-passing the prolonged pleasure of getting there, as I do with Tait's theorem. Now the original pleasure of Tait's proof is completely destroyed, there is no longer point in performing what has become meaningless gibberish, just to see how we used to convince ourselves of what we can now see perfectly without such a performance.

It is sensible, if a correct proposition has resisted proof for more than a century, to think that we might do better to find a way of making the whole set of supporting lemmata more evident, if the most difficult of them are not to appear totally recalcitrant. What happened here was, to make the later theorems manageable, I corrected the notation in a way that so clarified the earlier theorems that they became self-evident. After this the later theorems, such as the indivisibility and parity theorems, whose existence was impossible even to *see* before the change of notation, became, if not easy, at least clearly visible enough to prove.

The way I prove theorem 17 is unusual, and the only similar proof I can recall is my proof of theorem 16 in *Laws of Form*. Both of these theorems are hard to prove because each is within a whisker of being completely self-evident, that is, employable as an axiom. In fact theorem 16 of *Laws* was in practise taken as axiomatic before I supplied a proof, and theorem 17 of *Maps* may be similarly so taken. My proofs in either case merely confirm beyond any trace of doubt that, in seeing them as axiomatic, we were seeing them correctly.

We can summarize the method of proof in each case, showing clearly that each proof employs a transmission principle.

T 16 Laws

How does what we change in s_n transmit itself to what appears as a change of value to s_0?

T 17 Maps

How does what we change in t transmit itself to what appears as a change of colorability to dx?

My proofs of both theorems are based on the fact that what is one side of a bound in mathematics is formally identical to what is the other side, since the formality is the common bound they share. We might call it the *axiom of formation*: *What a thing is is formally identical to what it is not.*

Thus when we shape a bound "from the inside" it is bound [*sic!*] to show the shape we gave it in the "effect" it has on the things seen "from the outside". Thus a bound, such as a factor f in a trail t, fashioned "from the inside" so that it can or cannot provide a modulus for dx in C, must exhibit this very fashioning in the way it is seen to relate to the things in the mathematical half-universe outside of f, which consists of the remainder of the formation F over $C - x$. I prove the indivisibility theorem by showing that if it were false, we would have made a definite property of f that failed to show up in the world around f. Since this would contradict the axiom of formation, it is impossible, and theorem 17 must therefore be true.

In the old theorem 25 (now recast as theorem 23) I used a different parity-changing procedure, called a parity-mill pm, that works on a 2-deficient 5-region. This, although more spectacular, because it shatters circuits all over the place irrespective of color, is not quite so suitable as the parity-pass pp, which confines its attentions more demurely to the red bands in the trail t.

I have always regarded a proof by descent – the method was invented, as far as I am aware, by Pierre de Fermat – as one of the rarest and most beautiful forms of proof, so I was delighted when it turned out to be possible to prove the four-color theorem this way. We can quite evidently do so by proceeding as far as theorem 23, which demands a proof that some sort of parity-changing operation must in all cases be possible to a supposedly incompletable formation in the plane.

Perhaps the reader will have already sensed that we do not need even to proceed this far, because the fact that parity can *not* be changed by a simple operation in the plane can also supply the contradiction we require.

Consider the petersen trail o_1. Do a simple operation c on the γ-circuit in a link y adjacent to the deficient link x. Before applying $c(y)$, $\tau F = 5$. After applying $c(y)$, $\tau F.c(y) = 4$. The formation, being nonplanar, was not obliged to obey the parity law, and has swallowed up one its circuits. This was the only way it could behave to leave no excessive factor, and it duly did so, thus confirming the theorem 17 corollary.

Suppose now we are in the plane, having replaced the nonplanar o_1 with what we hope is a least planar 0-factor, let us say t. Call the new formation F', and the new cast C'. Now do c on γ at y. Before we did it, F' of $C' - x$ must have had exactly five circuits, two for each of the x-origins and one for the trail. Any greater number than five would constitute a factorization of t, contravening the mathematical necessity that t be irreducible. We are in the plane, remember, so parity must be conserved. τ clearly cannot go down to

three, so $\tau F.c(y)$ must either remain at five or go higher, in either case constituting a factorization of t. For with x now on different colors we do not need a trail at all, but the planar situation decides that in this case we must still have at least one excessive factor, so after the operation of c on y, either a red band or a blue band must be split. By judiciously choosing the situation of y, we can split the red band that, in the previous formation, was the supposedly irreducible factor t.

Thus in all cases we arrive back at the planar motto: either C can be colored or t can be factored.

Another way of stating the theorem 17 corollary is to say that the essential zeroness of any 0-factor in t is inviolable, and thus not divisible between factors of t. Thus if t can be factored, the principle in t that made the cast C uncolorable, will be found to exist complete and unchanged in one of the factors.

This is what enabled me, in recent studies, to prove more general theorems that imply the four-color theorem by exclusion. The approach in these cases is via the concept of uncolorability. The approach to the present proofs is via the concept of planarity. Both approaches are equally effective, and what I prove in this, the planarity approach, can be summarized quite simply.

A planar map cast either can be colored or the process of eliminating factors can be continued without limit. And since the latter is evidently impossible, the former must be the case.

A final word about factorization is in order. Factors can be of four kinds, best illustrated by examples.

1. Cast or graph factors: these are how any cast C, considered as a graph G, can be partitioned.

2. Operational factors: the resultant of the idemposition of two formations F_1, F_2 can be considered in this sense as having F_1, F_2 as factors. Thus o_1 can be 'factorized' operationally into 2 quadrilaterals, although in the third and fourth senses below, o_1 must be regarded as irreducible. Operational factors are in effect *fractions*. Analogously in number theory, 2 1/4 × 5 7/9 are *operational* factors of the prime number 13. In order to define primes in either system, we must confine the elements, that may be entered or eliminated, to whole numbers and whole circuits.

3. Formation factors: in this sense o_1 is irreducible, because there is no operation Φ that can be applied to F, over the relevant cast C, that can make a formation ΦF, such that $\Phi F \not\equiv_m F$, and such that ΦF has resolved o_1 into seperate circuits of the same modulus. (If we operate trivially on both origins of dx, we merely change the color of o_1.)

Each cast C can be thought of as a distinct arithmetical frame in which the circuits of F over C can dispose themselves. The concept of factoring a circuit is similar to that of factoring a number. If, in the arithmetical frame in which the circuit (or the number) exists, it can be broken into seperate circuits (factors), then it is reducible (composite). If not, then it is irreducible (prime).

We may subclassify formation factors into social or S-factors, consisting of the greatest number of a given modulus that can coexist in some maximal F over a given C, and lone or L-factors, into which an S-factor in the subfactor of F with the other S-factors eliminated, can sometimes be split further. In figure 14.1, o, p, q constitute a best-possible S-factorization, so are all S-irreducible. o is also L-irreducible, but each of p, q is L-reducible. I have shown elsewhere [e.g. Chicago 1990, Vienna 1996] that S-factorization alone is sufficient to isolate o in any uncolorable C.

It is evident that any projection of F from a dx in which not all factors, either S or L, can be isolated, they can all nevertheless be identified as operational factors.

4. Trail factors: these belong to the class of factors of a formation F projected from a deficient link dx in C, and consist of the subclass of such factors in excess of the minimum number required for x not to be on a modulus. We can further classify them as strong or weak in respect of whether they touch both, or less than both, of two original circuits of identical modulus traversing each of the seperate origins of x.

A trail factor f is considered irreducible if it is real and strong, and constitutes an excess of 1, and there exists no operation Φ such that $\Phi F = F'$, with both F and F' mathematically nonequivalent and 1-deficient over C, that can break f into several components of the same modulus. In other words, a strong connective factor f is considered irreducible if and only if it can not be L-factored.

Trail factors, strictly speaking, belong only to uncolorable casts: for in colorable casts there is always some operation that will place dx, in some situation, on a modulus. When this happens there is no trail and what were its factors can no longer be called trail factors, though we can still call them connective. It is only by keeping x off a modulus, in colorable cases, for as long as we can, that we can partially illustrate the inconsistent nature of the phenomena that would necessarily take place if it were possible for an uncolorable cast to exist in the plane.

The following four theorems are new and true, but I will not record my proofs of them here.

Theorem 24*. A mending of Kempe's algorithm

Unless a 5-region Q in a plane map M is at a center of radial symmetry (call it a polar region) there exists a set of color exchanges round the borders of Q that will allow Q to be colored.

Comment. Surprisingly, Heawood's famous map (1890) is not a counter to Kempe's method. It is easily colored if we allow all possible exchanges, not just those detailed by

Kempe. Errera's polar map can of course be easily colored in many ways, and can be Kempe-colored, in the extended sense of allowing all possible perimeter exchanges, from any but the polar 5-regions. More surprisingly, it appears to be unique. There are an infinite number of maps with a polar 5-region, and I have found none, except Errera's, that can not be Kempe-colored from a polar region. The reader should look for one, or prove they can not exist, a proof that has eluded me so far, although I can list convincing data suggesting they must be impossible.

Theorem 25*

A parity-pass pp can not be completed with dx in any but a border of a polar 5-region in M a plane map.

Comment. Here again the theorem appears too weak. Errera's is the *only* known plane map in which a parity-pass can be completed. In other polar 5-regions, in which a trail is possible, *pp* can not be completed, thus providing a repetitive algorithm, suitable for electronic computers, for coloring M, even from a polar region. Here again I suspect, but have not yet proved, that Errera's map is unique in allowing *pp* to be completed.

Thus theorems 24, 25 appear to be different ways of stating an identical fact: parity-pass completable is equivalent to not Kempe-colorable.

Theorem 26*

If a plane map M does not solve simply with dx in a 5-region Q, then M will solve simply if the deficiency at x is displaced to y, y a link adjacent to x in Q.

Comment. Theorem 26 is easier to use for hand-coloring, but demands a thought-out choice of what operation to do next. Theorem 25 is easier for programming computers, because although it may require more operations, they are entirely repetitive and demand no conscious choice.

Theorem 27*. An axiomization of Heawood's formula

If M is a nodeless map with not more than d divisions to each region in a surface s of connectivity c, then the number n of mutually contingent regions possible to M is in the range

$$n \leq [1/2\ (2d + 1 + (4\ c\ + (2d + 1)^2 - 12)^{1/2})]$$

where $[x]$ denotes the integral part of x.

Comment. Heawood's formula $n \leq [1/2\ (6d + 1 + (24\ c\ + (6d + 1)^2 - 72)^{1/2})]$ for noded maps does give the correct answer 4 when $c = d = 1$. And it does so because my formula for nodeless maps gives the correct answer 2. In the latter case there is no difficulty in seeing that the answer is correct. (*Proof.* Choose any region r in M and color it with one of the colors. Color all regions at an even depth from r with this color, and all regions at an odd depth with the other color.)

A number of short proofs of the four-color theorem for noded maps now become evident. In the case of nodeless plane maps, there is no way we can begin to color incorrectly and then have to go back to where we went wrong and start again. When we amalgamate two nodeless maps to make a noded map (theorem 7) we are faced with exactly this difficulty, which is the only reason that the four-color theorem ever appeared to be a problem.

I suggested as early as 1961 that theorem 7, reducing the four-color problem to a two-color problem, was a more-correct way to regard it, and was surprised to learn, fifteen years later, that no one had thought of this way before.

It is obvious enough that all plane maps can be made from two sorts of distinction, if we are allowed to dispose examples of them any way we please. All we have done is written up the calculus of indications twice, first with red crosses and next with blue, and colored the intervening spaces more simply, so that they behave as precursors of maps rather than propositions. Since each write-up is plane-filling, there can be no chance of either going wrong, and no chance of the resultant of their idemposition requiring more colors than we began with. Whichever way we color the expressions, as maps or as propositional precursors, the exigencies of Chapter 11 apply, and we cannot mismatch the colors unless we subvert the plane. Thus if the four-color theorem were false, so that we *could* mismatch the colors in the plane, the whole of *Laws of Form* would be invalid. But since *Laws of Form* clearly is valid, the four-color theorem must be true.

This is so evident, even without further elaboration, that it was considered an adequate proof of the color theorem by all the major mathematicians to whom I published it in the early 1960's, Bertrand Russell, JCP Miller, and DJ Spencer-Brown. And if that publication could be considered not general, the first publication of *Laws* in 1969 certainly was, and in effect recounts the same proof. Here again we may ask, Why was there a problem? The answer is because the map was not seen this way. Instead of being seen as an idemposition of two colorings, which together define, in a choice of ways, a border-structure that need not be further marked, the map was wrongly seen as a border-structure without the colors, which, being not how the map is really made, at once creates a problem.

The axiom of idemposition operates in every construction throughout the universe, and it was only by being blind to this axiom for centuries that mankind could think maps might possibly be an exception to the rule. But since maps are not an exception, the four-color theorem was proved the moment we thought of applying the axiom.

The artificial problem, being given the border-network, to find a pair of nodeless maps

that, among many such pairs, will match it, is another problem altogether. Any successful solution to this problem is what is called a coloring algorithm, and I set myself to discover such algorithms in 1976, beginning with a rehabilitation of Kempe's algorithm in theorem 24. All such algorithms are hard to find and difficult to prove. But they are mere practical details, of little relevance to the main question. I must leave the reader something to do, and he, or she, can have a wonderful time exploring theorems 24 sq, and, if possible, making satisfactory proofs.

Another shorter proof, already mentioned, is to see that o_1 is a largest prime connective factor and the *only* prime 0-factor. And since it is nonplanar, the four-color theorem is proved by exclusion*.

A third short proof is to look for prime planar connective factors. The paradigm of such factors is a red quadrilateral, 2-connected to top and bottom blue origins of dx, and all such factors are obviously 1-factors. Since any such prime connective factor can be not more than 4-connected to the rest of the map, it follows that in a 5-connected plane map, required by theorem 12, all connective factors must be composite, a fact I proved at greater length in the main text above.

When the reader has become fluent in the calculus for maps, he, or she, will no doubt find pleasure in constructing other short proofs. I have made some twenty-seven to date. They are proofs of the most powerful kind, showing that we thought of maps in a wrong way: and that by correcting the way we think of them, the so-called "problem" ceases to exist.

Finally, we should be aware that calculi, by themselves, do not do anything. It is us that do the doing. However good the calculus, in the hands of a wrong member of us, will achieve nothing spectacular: and however bad it is, in the hands of a right member of us, will achieve the spectacular. Ramanujan, in making his incredible summation formulas for π, was obviously expressing his results in what, to him, was the wrong calculus. He was clearly using, as indeed Fermat must have been to prove his 'last' theorem, some much more direct way of thinking, that he was aware of and we are not.

George Spencer-Brown
Cambridge England September 1979
[Existing text in the Library of the Royal Society edited and shortened by the Author, London February 1996 and February 1997]

* Cf. Spencer-Brown, George, Uncolorable Trivalent Graphs, *Cybernetics and Systems* 29 (4) (June 1998) 319-344. This is the Ross Ashby Memorial Lecture mentioned on p 159..

In the following figures the origins of x are supposed to be parallel straight lines, so only the trails need to be shown.

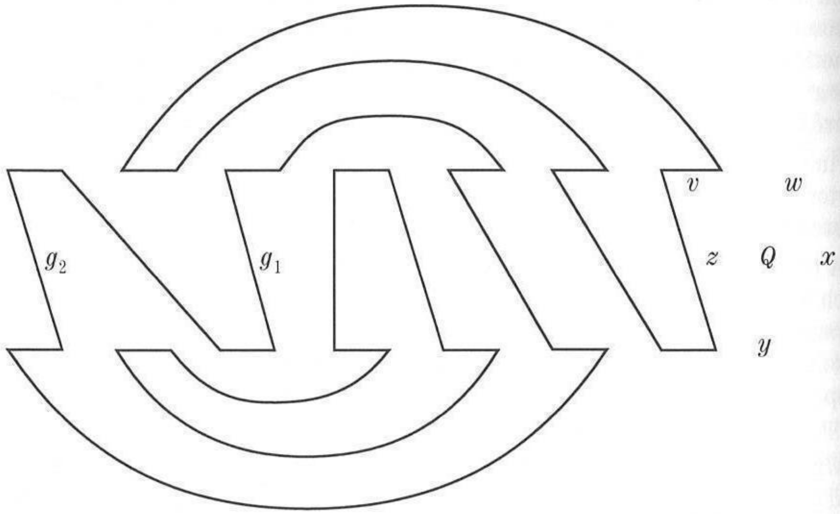


Figure 13.1. Errera's map E. The displacement of the color-deficiency to one of the (g_n) can be done in such a way as to leave g_n on a modulus.

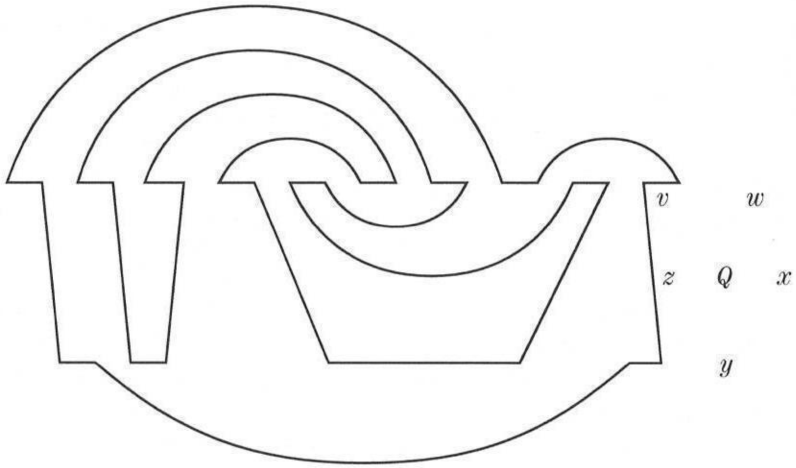


Figure 13.2. $\Phi E = E.a(xyz).b(xy).c(wxy).a(z).c(w) = E'$. Notice that E' has one fewer factors in the trail. Φ can be repeated four times to make $E'''' = E$. The seperate calls of Φ can also be applied in reverse, say Φ', as in $\Phi' E = E'''$.

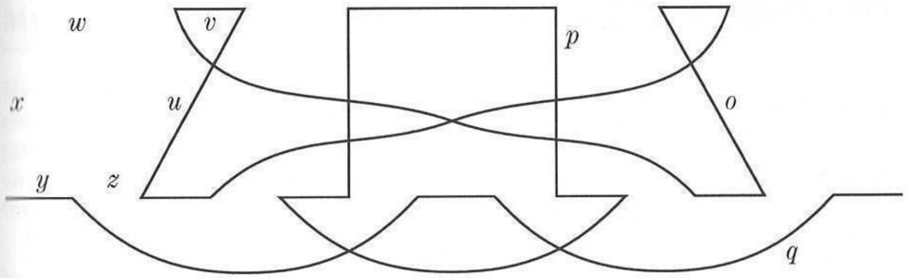


Figure 14.1. Factorized trail of the Isaacs J_5 uncolorable. Factors p, q, by themselves, can be split further, but o is irreducible.

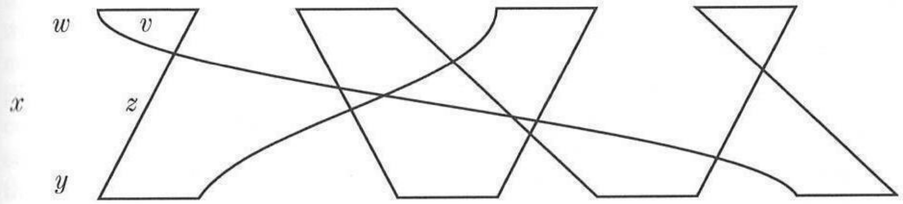


Figure 14.2. Trivial uncolorable T, with two prime stop-factors that can be factored out. To find one, do $T.a(y)$. Take it out of the picture to make T', and then do $T'.a(y)$ to find the other.

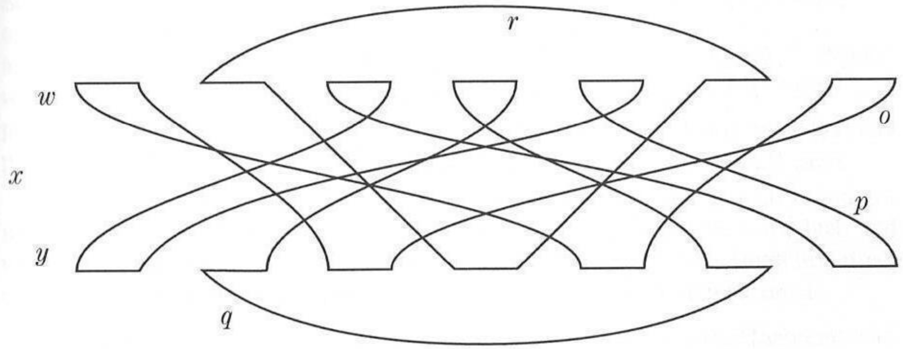


Figure 15. The Isaacs Unicorn, misnamed by him "Double-Star". My name for it is an elision of Unique Ornament. Note that the elimination of either of the stop-factors o, p leaves x colorable. See theorem 17.

Part IV. References

Bernhart, Arthur, Six-rings in minimal five-color maps,
Am J Math 69 (1947) 391-412.

Birkhoff, George David, The reducibility of maps,
Am J Math 35 (1913) 115-128.

Errera, Alfred, Du Coloriage des Cartes et de Quelques Questions d'Analysis Situs,
doctorate thesis, University of Brussels, 1921.

Gauss, Karl Friedrich, *Disquisitiones Arithmeticae*,
Leipzig 1801, English translation by A A Clarke SJ, New Haven and London 1966.

Heawood, Percy John, Map-colour theorem,
Quart J Math, Oxford series, 24 (1890) 332-338.

Isaacs, Rufus, Infinite families of nontrivial trivalent graphs which are not tait colorable,
Am Math Monthly 73 (1975) 221-239.

Kempe, Alfred Bray, On the geographical problem of the four colours,
Am J Math 2 (1879) 193-201.

Petersen, Julius, Die Theorie der Regulären Graphs,
Acta Math, Stockholm 15 (1891) 193-220.

Shannon, Claude Elwood, A theorem on coloring the lines of a network,
J Mathematics and Physics 28 (1949) 148-151.

Spencer-Brown, George, *Laws of Form*,
London 1969.

Tait, Peter Guthrie, On the colouring of maps,
Proc. Royal Society of Edinburgh 10 (1880) 501-503.

Appendix 6

My simplest proof* of the four-colour map theorem

My *Concise Oxford* defines a theorem as a 'proposition to be proved by a chain of reasoning'. In mathematics the term is sometimes confined to propositions for which a proof is generally known. This I shall suggest is inconvenient. What is convenient in my view, is to restrict the term 'theorem' to propositions that are seen to *require* proof, as distinct from propositions that are evident without proof, which are sometimes called axioms.

Why I think it is inconvenient to restrict the word 'theorem' to propositions with known proofs, rather than to propositions merely claimed by some authority to be true, is that if such a proposition, unproved perhaps for many years, is suddenly proved, its status must equally suddenly be changed to 'theorem' without any such change to its truth-value. (If a proposition is eventually proved true, it is then supposed to have been true before it was proved.) Furthermore, the usage is not applied consistently by those who think it should be. Fermat's "last theorem" was always called a theorem even though no proof was known until recently. And I can mention an even worse example: not all of Gauss's "theorems" in the *Disquisitiones* are true. The class of false ones, with obvious counter-examples, cannot even be properly described as 'conjectures', because they are simply mistakes, and their so-called "proofs" meaningless gibberish. But Gauss calls them theorems, and there is no convenient way we can refer to them otherwise.

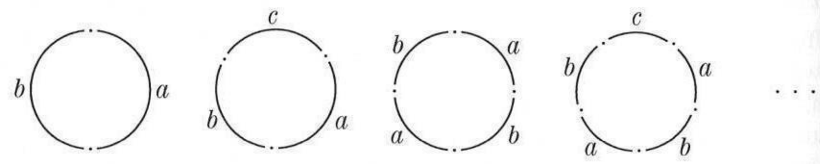
Having decided on the word 'theorem' let us turn to the word 'proof'. In mathematics we mean a published chain of reasoning that is (a) valid, and (b) can be checked by the reader to be so. Any claim to have published a proof would be at least mistaken if it did not fulfil criterion (a), and fraudulent if it did not allow criterion (b). Appel, Haken, and Koch claimed to have proved the 4-colour theorem (say 4CT) in 1976. Their claim might or might not have been true, according to whether or not it fulfilled criterion (a). But any claim by them to have *published* a proof would have been misleading, because there was no way what they did publish could be said to fulfil criterion (b). It is also very probable that they failed to fulfil criterion (a), because the details of what they did publish were changed in successive drafts.

Also in 1976 I naively claimed to have proved the 4CT without being ready to present my proof in writing. The math community is rightly unforgiving of such conduct, and when later I did present a written account of my proof it was ignored. One's first proof of anything is often long and complicated, and I later made many simpler proofs.

Let me present what I think is my simplest proof, uncluttered by other theorems with which it is presented in the footnote reference above. For this we must go back to

* Extracted from Spencer-Brown, George, Uncolorable Trivalent Graphs, *Cybernetics and Systems*, 29 (4) (June 1998) pp 319-344

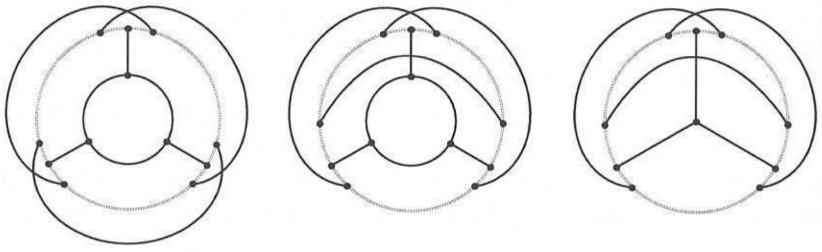
Chapter 11 in the *Laws*. Here I point out that if we make a reentrant chain of lines separated by markers, so that each marker determines a change of value that can be pictured in a change of colour, it can be done with two colours if the chain has an even number of lines, but we require one more colour if the chain has an odd number of lines. In the latter case we call the chain paradoxical, because whatever of two colours we select for the first line, we always find it must be coloured differently when we get back to it after going round the circuit.



1 2 3 4

A standardized plane map is made up of units of 3 lines meeting at a point, say a Y-shape. Each line in the Y-shape must be coloured differently, so we shall need at least 3 colours. The smallest odd number of such units that, connected any way we please to form a reentrant circuit of such units, is 3. We are not allowed to connect any two outer limits of any given Y-shape directly to each other, because if we did we would not be able to make a proper circuit of such units.

There being a limited number of original Y-shapes (only 3) there must be a limited number of different ways their outer points can be connected to form a reentrant circuit, and we may try each in turn. The 4CT will be true (by Tait's theorem) *either* if we can find no way of connecting the Y-shapes paradoxically, i.e. they can all be coloured with the same 3 colours a, b, c, *or* if we do find a way of connecting them paradoxically, so that a fourth colour d must be used to resolve the paradox, but we also find that such a connexion cannot be imbedded (i.e. drawn with no lines crossing) in the plane.



1 2 3

The three diagrams above show there is just one way of connecting the original 3 Y-shapes paradoxically, but (diagram 2) it cannot be drawn in the plane without at least

two crossings. Diagram 3 eliminates the 'reducible' configuration in this graph, the circuit of 3 lines in the middle. The resulting graph has no reducible configuration, i.e. is at least 5-connected and has no circuit of fewer than 5 lines. It is called Petersen's graph. It is also easy to see that any larger 3-graph that is paradoxical must contain* this graph, since there is no alternative.

But we have already seen that it, or any larger such graph, cannot be drawn in the plane without crossings, so that any graph that can be drawn in the plane without crossings cannot contain Petersen's graph and so must be line-3-colourable and therefore, by Tait's theorem, region-4-colourable. Quod erat demonstrandum.

I have always maintained that if a theorem can be proved, there will be a proof that a child of six can follow and understand. My simplest proof of the 4CT above is a case in point. It is easy enough for a six-year-old to experiment with pencil and paper to see that (s)he cannot draw a reentrant circuit of more than two original Y-shapes in the plane without crossing at least two lines.

* Just as, in the simpler case, the larger circuits that are not 2-colourable must contain circuit 2.

Appendix 7

The prime limit theorem

Riemann's Explicit Formula* for the growth of $\pi(n)$, the number of primes $\leq n$, may be written

(0)
$$R(n) = \sum_{k=1}^{\infty} \frac{\mu(k)\mathrm{Li}(n^{\frac{1}{k}})}{k}$$

where μ is the Möbius function. It first struck me in January 1998 that Riemann had missed a much simpler way of expressing this result. Call $\mathrm{DR}(n) = \mathrm{R}((n+1)^2) - \mathrm{R}(n^2)$. Then $\mathrm{DR}(n)$ must be asymptotic to $n/\log n$. In signs

(1)
$$\mathrm{DR}(n) \sim n/\log n.$$

What makes (1) so useful is the closeness of the approach through the whole range of n. For example when $n = 10^4$, $n/\log n = 1085.7362...$ and the more-laborious computation of $\mathrm{DR}(n)$ using Riemann's formula shows no difference before the third decimal place.

It is now evident that we can adapt my formula (1) to achieve more simply what Riemann's does. Thus

(2)
$$\mathrm{S}_t(n) = \frac{n}{(\sqrt{n}+t)^2}\left(\sum_{k=2}^{[\sqrt{n}+t]-1} \frac{k}{\log k} + \langle\sqrt{n}+t\rangle \frac{[\sqrt{n}+t]}{\log[\sqrt{n}+t]}\right)$$

where $\langle x \rangle$ stands for 'fraction part of x' , i.e. $x - [\, x]$. I have used the fraction part of $\sqrt{n} + t$ to interpolate when n is not a natural square or t is not a whole number. This can introduce a small continuity error in a remote decimal place, of no consequence in a formula designed to predict cardinal numbers of primes. It is easy to prove that $\pi(n) \sim \mathrm{S}_t(n)$ for any finite value of t, but here we aim merely to match the performance of (0). Setting $t = 2$ will match evaluations of Riemann's formula to within ± 2 up to $n = 10^7$, after which t must be progressively reduced to allow for the diminution of the difference in (1).

Set $t = 1 - (\log n)/2$ and call $\mathrm{S}_t(n) = \mathrm{S}'(n)$ in this case. Then

(3)
$$\mathrm{S}'(n) \overset{\frown}{=} \mathrm{Li}(n).$$

* Riemann, Bernhard, *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse.* Monatsberichte der Berliner Akademie, November 1859.

Thus although Riemann's formula cannot follow the curve of $\pi(n)$ when it crosses that of $\mathrm{Li}(n)$, mine can. A crossing is known to happen when n is around 6.2×10^{370}, so set $t = -426$ for this occasion. It is evident that the convergence of $\pi(n)$ to $S_t(n)$ must occur for any specific setting of t, positive or negative. Littlewood's theorem together with (3) assures us of covergence however large the negative value of t. For the positive side, choose t sufficiently large to make $S_t(n) = n/\log n$. Then $S_{t'}(n') = n'/\log n'$ when $t' \triangleq t\sqrt{n}'/\sqrt{n}$, and this together with (4) assures us of convergence whatever the positive value of t. QED

The structure of Riemann's formula assures us that $DR(n)$ must be almost identical with $\mathrm{Li}((n+1)^2) - \mathrm{Li}(n^2)$, so it remains only to show that this last expression is also strongly convergent to $n/\log n$, which is not difficult. Then the convergence to $\pi(n)$ of all the quantities listed here, including Riemann's, must follow.

My formula (1) implies that the function $n/\log n$ is an approximate indicator of the number of primes between n^2 and $(n+1)^2$ rather than between 0 and $n+1$ as was hitherto supposed. We can thus correct the Prime Number Theorem from

(4) $$\pi(n) \sim n/\log n$$

to the much more accurate

(5) $$\pi(n^2) \sim \sum_{k=2}^{n-1} \frac{k}{\log k} = S_0(n^2)$$

What is remarkable is that, in the last 200 years, Legendre, Gauss, Tschebycheff, Riemann, Hadamard, de la Vallée Poussin, and the lesser mathematicians of the Twentieth Century all misapplied $n/\log n$ to the wrong stretch of numbers.

This might have been because none of them could prove there were necessarily any primes between n^2 and $(n+1)^2$, so no one thought of asking how many there might be. Even I did not think of addressing this question until I had made an elementary proof, on 24th January 1998, of the existence of at least two primes between any two successive natural squares. After this the question of how many to expect became urgent, and I was able not only to prove that the expected number is near $n/\log n$, but also to ascertain sharp limits beyond which the actual number of primes between any n^2 and $(n+1)^2$ cannot deviate from $n/\log n$.

The total number $t(n)$ of primes between n^2 and $(n+1)^2$, where n is a natural number greater than 1, is within the limits

(6) $$A - (B-1) < t(n) < A + (B-1)$$

where A is $n/\log n$ and B is $A/\log A$.

Since my formula (6), which I call the Prime Limit Theorem, is not necessarily true unless n is one of the natural numbers 2,3,4,..., it further confirms the existence of a fundamental relationship between logarithms to the base e and the natural numbers, their squares, and the primes.

George Spencer-Brown, *1998 October 24*
Trinity College Cambridge

Correcting the Prime Number Theorem

After suggesting that the density of primes might be $1/\log x$, Gauss proceeded to make a conceptual blunder so enormous that it apparently stunned those who followed him into accepting it without question.

The chapter of horrors began when Gauss failed to say just *where* this hypothetical density was supposed to apply. The obvious place would have been at x, but instead Gauss proposed the theorem $\pi(x) \sim x/\log x$ (the present PNT), suggesting with uncharacteristic lack of precision that it must somehow apply indiscriminately to the whole stretch of positive numbers before x, resulting in the well-known answer that is, of course, absurdly too small.

The thought processes that led to Gauss's mistake are of some interest, since they were not confined to Gauss alone. But rather than consider why he made it, I propose, as simply as I can, to correct it.

The prime point density principle ppdp

Suppose the density of primes at the real number $x > 1$ to be exactly $\log^{-1} x$.*

I have defined density at a point, the real number x. To give it a verifiable meaning I shall have to expand the point into a restricted interval, say Δ, say of size $s \leq 2x$. We can now suppose prime density to be the ratio of primes to whole numbers in Δ.

Define $°\pi(\Delta)$, say the prime degree or content of Δ, to be the number of primes in the interval Δ. We can now define Δ in terms of s and x so that $\Delta s(x)$ (say delta s of x) is the interval of size s whose midpoint is x. Interpret $s\log^{-1}x$ to be an *expectation* of the number of primes in $\Delta s(x)$. Assuming the truth of the ppdp, certain consequences must follow.

* Even this supposedly "well known" theorem is new, because no one had the discrimination to use the word 'exactly' in this context. To say 'approximately', as previous authors stated or implied, is simply wrong. If the theorem is not explicitly stated to be exact, the important consequences I draw from it will not follow.

Consequence 1

There exists an interval of size $s \leq 2x$ in which $^\circ\pi(\Delta s(x))$ and $s\log^{-1}x$ differ by less than 1.

Consider $x = 500$, $s = 2x$. We establish an inequality, that for $x \geq 111\tfrac{1}{2}$, $s = 2x$, $^\circ\pi(\Delta s(x)) > s\log^{-1}x$, which manifests in this case as $^\circ\pi(\Delta 1000(500)) = \pi(1000) = 168 > 1000\log^{-1}500 = 160.911...$. My Consequence 2 below assures us that there exists a critical value of $s < 2x$ where this inequality will be reversed. Let us try chopping 125 units off each end of the 1000, reducing s to 750. This happens to be about right (we do not need to look for anything more exact), $750\log^{-1}500 = 120.683...$ and the prime count in this smaller interval is $^\circ\pi(\Delta 750(500)) = 120$. The expected reversal has just taken place, and we can call $s = 750$ a critical value for $x = 500$. Here we have what I call a perfect population of elements with a ratio of primes to integers that is within a fraction of the calculated density. Any further reductions of s will merely have the effect of taking samples from this perfect population. Since the procedure can be generalized for all x, I have, by this means, transformed what was, for 200 years, a notoriously inexact science into an unexpectedly exact one.*

Consequence 2

For every x, there exists a critical value of s where the curve of $^\circ\pi(\Delta s(x))$ either kisses or crosses that of $s\log^{-1}x$.

We demonstrate this simply by remarking that if it were not so, the ppdp would be false.

What is now becoming clear is that I have constructed a set of five consequential theorems, the ppdp and its four consequences, which we may call the *elementals* of the system, any one of which may be taken as the initial basis of the system, and from the truth of which will follow the truth of all the other elementals in the system. Once we find such a system, all will fall into place and everything will confirm everything else.

The elementals of any system have no natural ordinal sequence. They have to be stated in some order, but whatever order we choose we find we have to refer to those we have not yet mentioned to explain those we have.

Consequence 3

The density of primes at x^k is exactly k^{-1} of their density at x. (x and k real numbers, $x > 1$, $k > 0$.)

*For even greater accuracy, we can average the prime counts above and below the point where the curve of $^\circ\pi(\Delta s(x))$ first crosses that of $s\log^{-1}x$. We may even shift x to fractionally off centre so that prime counts do not jump by two at a time.

Corollary

All intervals $(\Delta ks(x^k))$ have the same prime expectation.

Call this the inverse power law. Having eliminated the natural logarithm, we can begin to see how the unifying principle of primality is of spectacular simplicity.

I decided to try out the ppdp in 1999 April 17, while rereading H M Edwards's book on Riemann's Zeta Function. The newness of the resulting mathematics, with the fact that I had to invent an unfamiliar technology to describe it, soon convinced me that nobody had seriously examined the idea before, even though it does seem such an obvious thing to do. The result has been that my C3 remained entirely unnoticed in all the previous literature. However obvious it may seem now that I have published it, it is certain no one was conscious of it before. I have known and worked with a dozen or more of the top number theorists for nearly half a century, and if any of us had thought of it he would have (a) mentioned it, (b) exploited it, and (c) published it. The law is far too *useful* to keep secret, once it has been seen.

Prime avenues

We can now apply the Corollary to my C3 to generate endless vistas of expanded intervals with the same prime expectation. Call them prime avenues.

x^k	Δ *of size ks*	$^\circ\pi(\Delta ks(x^k))$	x^k	Δ *of size ks*	$^\circ\pi(\Delta ks(x^k))$
350	$\pm$ 50	16	$350^{2.5}$	$\pm$125	14
$350^{1.5}$	$\pm$ 75	16	350^{3}	$\pm$150	18
350^{2}	$\pm$100	18	$350^{3.5}$	$\pm$175	17

Table 1. Prime avenue 350 : 100 (1 : 0.5 : 3.5). Expectation = $100\log^{-1}350 = 17.0708673798$
Average count per frame so far = 99/6 = 16.5

A colleague has just sent me an avenue 350 : 100 (1 : 1 : 50). The total count is 855 and the average is 17.1. The largest primes are 128-figure numbers and the nearest total to expectation would have been 854. It is not in fact necessary to calculate the expectation, because the prime counts will eventually supply it. Thus the natural logarithm can be eliminated not only practically, but also theoretically, from the theory of primes, since the primes themselves will always tell us what it is. In other words the primes, marshalled into avenues by my C3 Corollary, constitute yet another way of computing the number e.

Consequence 4

If m is the mean prime count in the avenue $x : s$, then $x^{m/s}$ is asymptotic to e.

This is the first time that any method has been devised for computing e from data supplied by the primes alone.* Of course, to get near the right answer, we would need to average hundreds more counts than my colleague did.

Supersaturated densities

We get prime densities > 1 when $x < e$, suggesting that in this region there are more primes than integers. When $x < 2$ there are apparently none of either, but the supersaturated densities my theorems demand can nevertheless be verified by raising x to a power k sufficient to make a number around which integers and primes can be counted. Do we regard the phenomenon as a computational convenience, or look for some other interpretation?

The accuracy of Riemann's formula

The known theorem that for $x \geq 11$, $\pi(2x) < 2\pi(x)$, implies that, for $x \geq 22^3/_2$ and $s = 2x$, $s\log^{-1}x < {}^{\circ}\pi(\Delta s(x))$. If this inequality were not to reverse itself for some $s < 2x$, the ppdp would be false (compare C2), so in general we must expect $ks\log^{-1}x^k$, when ks is small in comparison with x^k, to overestimate very slightly the expectation of primes in $\Delta ks(x^k)$, but that this discrepancy will approach zero as k increases.

We can use this theorem as a highly-sensitive test, adjudging Riemann's formula to be accurate if it can be shown to incorporate this discrepancy. To check this we need to construct avenues of Riemann differences in place of prime counts. My uncorrected† expectation for an interval of size 7 whose midpoint is 10 is $7\log^{-1}10 = 3.040061373...$, and $R(10 + 3^1/_2) - R(10 - 3^1/_2) = 2.7974...$, a little below this expectation, as anticipated. Repeating the subtraction with $10^3 \pm 10^1/_2$, $10^6 \pm 21$, and $10^9 \pm 31^1/_2$ yields successive differences 2.9818, 3.0384, and 3.0401, which is extremely close to what my theorem predicts. So Riemann's formula is confirmed to be nearly perfect, subject only to small corrections as detailed on p 184.

*I have never understood the fuss about the so-called Prime Number Theorem, which has been to prove that a horribly inexact and basically wrong approximation to $\pi(x)$ is nevertheless convergent to it. I can invent a dozen or more such approximations in an afternoon, all better than Gauss's, but what is the point? We are aiming to discover what familiar object, if any, the sequence of primes resembles. I have shown beyond all question that this object is e, in other words, that the primes themselves are a *manifestation* of this number. A wrong formula will use e (they all did) to get a bad answer for the primes, but will not reconstitute e from the primes. A correct formula will work in both directions, as mine does. This is the only theoretical prerequisite. It is independent of empirical evidence, which *must* follow provided we guessed right about the object the primes resemble (= reassemble). Gauss and Legendre both guessed right, but were hopelessly inadequate to the task of constructing the requisite arithmetic.

† See p 188.

What Riemann could have asked himself, if he had approached the problem in a less roundabout way, is, can his formula be reversed to compute e from the primes? The answer is yes, but not as simply or as accurately as mine can.

It is not generally known that Ramanujan was also clever enough to rediscover Riemann's formula. In 1913 he sent it to Hardy, who replied: 'As regards your work about primes, the result is certainly wrong. Of this there is no doubt whatsoever.' No mathematician feels the need to be this emphatic*, unless he has taken up a false position.

Twenty three years later, Hardy discovered his mistake, having at last read the Riemann paper. Now realizing Ramanujan's formula was identical to Riemann's, but derived more simply, he called it, in lectures, the Riemann-Ramanujan formula, too late to benefit Ramanujan. But Hardy never admitted in print that he had changed his mind about it, so very few people are aware that Ramanujan had anything useful to say on this subject. – Author, reviewing this document in 1999 May 27.

Note (June 2005). My formula (2) may be replaced by its simpler equivalent

$$\text{(2a)} \qquad SB_t(n) = \frac{n}{(\sqrt{n}+t)}\left(\mathrm{Li}((\sqrt{n}+t-\tfrac{1}{2})^2)-1\tfrac{1}{2}\right)$$

This is much quicker for large n, where the summation in (2) can take many hours on a pocket calculator. Using (2a) (But see p221.), $\mathrm{Li}(x)$ can be computed in a matter of seconds with Soldner's formula

$$\mathrm{Li}(x) = \gamma + \log\log x + \sum_{k=1}^{\infty}\frac{\log^k x}{k!k}$$

where γ is Euler's constant and when necessary the absolute value of y is used to compute $\log y$.

* We can contrast the comment of my own tutor, J C P Miller, to whom I sent a theorem that was wrong to the point of idiocy: 'I don't think we can use this.'

† Note to page 186

How to correct the raw expectation $E(x,s) = s\log_e^{-1}x$, is of astonishing beauty, and caused me an 18-month search to find the answer. Contrary to (or in refinement of) what I have suggested so far, the primes relate not exactly to e but to a function $\sigma(x)$ that converges to e more slowly than the historical definition $(1+x^{-1})^x$.

Define $\sigma(x)= (1+x^{-1/2})^{x^{1/2}}$. Then the corrected expectation of primes in the interval of size s whose midpoint is x is

$$C(x,s) = s\log_{\sigma(x)}^{-1} x,$$

provided s is subcritical to x.

To make sense of this we must consider the fact that the Eulerian number e is the primary value of the exponential function "at infinity", and this is of no interest to the primes in the region of some number x less than infinity. At first I thought that what might be of interest to these primes would be the partial computation of e at x by the usual formula $(1+x^{-1})^x$, but this gave entirely wrong answers. The final piece of the puzzle remained hidden until I realized that only the primes not greater than $\sqrt{x}$ have any bearing on the names and places of the primes in the immediate neighbourhood of x. Thus the only function of any interest to these primes will be the partial computation of e at $\sqrt{x}$, and this is my function $\sigma(x)$.

We can now compare my corrected expectations, which are exact, with the Riemann differences in the avenue 10 : 7 instanced above. It will be seen that the Riemann differences have a tendency to wander from the line of exact expectations, especially in the early stages.

x	10	10^3	10^6	10^9
Spencer-Brown exact expectations	2.6415	2.9930	3.0385	3.0400
Riemann differences	2.7974	2.9818	3.0384	3.0401

Appendix 8

Primes between squares

I shall employ the expression square segment, or simply *segment*, to indicate an interval between squares whose roots differ by 1, and write seg(n) when the roots are natural numbers, and seg(x) in the general case. In general seg(x) will indicate the stretch of numbers between x^2 and $(x + 1)^2$.

Call a square natural if its root is an integer other than zero. Such squares cannot be primes, so they may be considered as bounds to an infinite set of intervals that together include all the primes. Thus any two distinct such squares will define an interval that must in some case include a prime, and I shall prove that it does so in all cases.

seg(0) is the null case, there being no integer in the segment. Write $°\pi\text{seg}(x)$ to denote the number of primes between x^2 and $(x + 1)^2$. In general when x is a real number greater than the least absolute value of $(\sqrt{2} - 1)$, $°\pi\text{seg}(x) \geq 1$, and when x is a natural number n, $°\pi\text{seg}(x) \geq 2$. When $n \geq 6$, $°\pi\text{seg}(x) \geq 3$, and the minimum number of such primes, or more exactly its lesser bound, will increase regularly and predictably as n increases.[1]

Surprisingly none of these facts was known before 1998 01 24, when I completed proofs of a number of theorems implying the universal existence of primes between natural squares. The weakest of these is that $°\pi\text{seg}(n) \geq 1$. Since even this had defeated the attentions of all previous arithmeticians, I shall begin by proving it here.

Theorem 1

There is at least one odd prime in the square segment of every natural number.

Lemma 1

A number h is prime iff h is an integer and h does not make an integer quotient with any integer i in case $2 \leq i \leq \sqrt{h}$.

Proof

There is no such i unless $\sqrt{h} \geq 2$. If i divides h then $h/i = j$, say. Now either $i = j$, $i < j$, or $i > j$. In the first case $i = \sqrt{h}$, in the second $i < \sqrt{h}$, and in the third $j < \sqrt{h}$ and can replace i as the divisor. So effectively all proper divisors of h will have been tried by confining i in case $2 \leq i \leq \sqrt{h}$. QED*

It is evident we may further confine the trial divisors to primes, and since the primes $\leq \sqrt{h}$ alone will determine the names and places of all the other primes $\leq h$, we may call the primes $\leq \sqrt{h}$ the *prime generators* of the primes in h or, for short, the pg of h.

* The constraint on i ensures that i, $\sqrt{h}$, and h are all positive, but in fact primes, like moduli, are signless.

Imagine the system of integers, call it an arithmetic A, mapped onto a line of regularly-spaced indistinguishable marks or points stretching endlessly in either direction. Call these marks the elements of a system S.

Consider prime multisectors $d_1, d_2, d_3, \ldots$ that can be imbedded in every second, every third, every fifth, and so on of the elements in S. We may thus call $d_i = p_i$ as long as we recognize that the multisectors are not primes or prime divisors, but merely operators on elements of S that can be placed where we please to imbed with or *strike* certain of these elements.

Consider S after d_1 has struck every second element, and eliminate the elements that have been struck, leaving the remainder to form a system S', say. S' is now identical to S, and d_1 is the only multisector that makes such an identity. So we can save labour by mapping the odd numbers in A, call them A', onto the elements of S'.

We can employ the noun *strike* to indicate the appearance of a section of S after the operation of a set of odd d's. It will of course appear similar to a section of numbers in A in which a corresponding set of prime divisors have imbedded to make composites, leaving the untouched numbers therein as new primes. (Compare the sieve of Eratosthenes.*) Every possible arrangement of multisectors in S is matched by a similar arrangement of divisors somewhere in A, and the reason for separating the concepts of S and A is to be able to experiment with the multisectors in S, by shifting them from place to place, which we cannot do with the divisors in A. In other words, I have made in S a working model of A, with the aim of discovering sharp limits to what is possible in A by seeing what is possible and what is not, in S.

Write $p_{\sim j}$ (and similarly $d_{\sim j}$) to indicate the set of all primes (and similarly all prime multisectors) $\le p_j$ (d_j). In S' and the odd-number arithmetic A', p_1 (or d_1) has no place, so $d_{\sim j}$ in S' effectively means d_2 through d_j.

We examine what happens when a set of relatively prime multisectors strike over a consecutive set of points in S or S'. Observe in Figure 1 we can place inwards-facing mirrors at each end of any quadrant to get the full picture to infinity, so the quadrant is the *quantum*, or least stretch of space that contains all the visual information we require**. Figure 1 shows the repetitive arrangements, and the names I have given them, of a simple strike of two odd prime multisectors or divisors. It can of course be extended to any number of such multisectors or divisors.

* The true pg of n is *all* the multiplicative integers $\le \sqrt{n}$, because it demands no previous knowledge of primes.
** The mirror placings are of particular interest since they must be through the exact location of a point. Thus if these points are euclidean, having position but no size, they will be entirely obliterated by the mirror. To reconstruct them we must allow them an infinitesimal size, so that half the point will appear in the quadrant, and the other half as its image in the mirror.

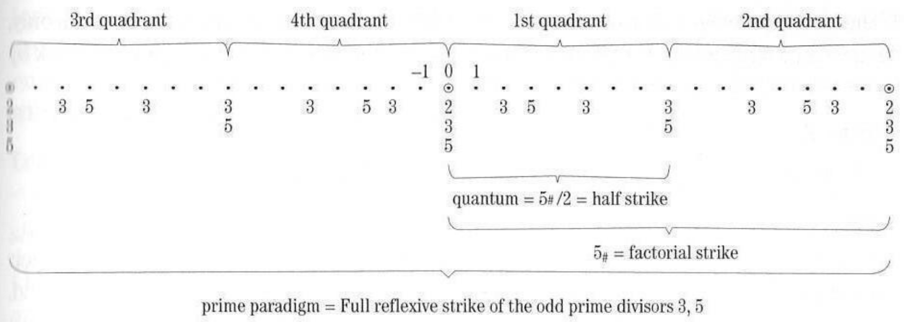


Figure 1. Full reflexive strike of the odd primes 3, 5. The ringed points are mirror-points at multiples of 2, which do not strictly belong to the odd-strike.

Write $x\#$, say x-hash or x-strike, for the prime factorial of a nonnegative real number x. $x\#$ is defined as the factorial $[x]!$ of the integer part of x, divided by the composite numbers $\le x$. For example $\sqrt{80}\# = 8.944...\# = 8!/8.6.4 = 7\# = 210$. Thus the primes ≤ 7 form the pg of 80, their strike covers 210 integers, and their quantum 105.

Since these facts are basic to my explication, they will bear further analysis. The primes form a self-generating construction. For example the four primes 2, 3, 5, 7 are also the pg of 100, that is to say, they are the sole determinants of all the primes < 100. All primes other than these have nothing to do with it. The primes up to 7 tell us everything there is to know about the primes up to 97.

In turn, the 25 primes ≤ 97 tell us all about the 1229 primes < 10000, and these in turn tell us all about the 5761455 primes $< 10^8$, and so on. Three iterations of this self-generating process has turned four primes into nearly six million. And we do not need four primes to begin with. Every fact, without exception, about every prime in the universe is contained in our knowledge of the situation of the first prime 2. Furthermore, we do not even need to know that 2 is prime, merely that it is the first multiplicative integer. 1 is nonexistent as a multiplier, divisor, exponent, or root.

Thus 2 can tell us the answer to every question that can possibly he asked about primes. It can say whether Goldbach's conjecture is true or false, and whether there are infinitely many prime twins. It can say whether the Riemann Hypothesis is true or false, and whether there are any more prime Fermat numbers. It can also say whether there is always a prime between successive natural squares, and to persuade it to reveal the answer, all we have to do is discover how to put the question.

I propose in this case to put the question in the following way. What is the maximum set of successive composites that can be generated by the prime divisors in the pg of n^2?

If the answer comes to less than $2n$, or less than n if we consider the odd numbers only, then we are done, since it will mean that not all the numbers in seg(n) can be struck by these prime divisors, so at least one of the numbers in seg(n) must be prime.

Lemma 2

There are exactly n odd numbers between n^2 and $(n + 1)^2$.

Proof

$(n + 1)^2 - n^2 = 2n + 1$. The number of integers in a segment is the difference between their bounds less 1, i.e. $2n$ in this case, and in any stretch of $2n$ integers n of them must be odd. QED

We shall need one further special term. Mnemonic: imagine a *set* as a Show of Elements Together. Then construct the term *sect* to denote a Show of Elements Consecutively Together. By a happy chance it extends the ordinary meaning of the word 'sect' and perfectly fits its etymology, which is from Middle English *secte*, from Old French, from Latin *secta*, from *sectus*, archaic past participle of *sequi* to follow.

From Figure 1 we see there are just two basic ways in which a maximum sect of prime divisors $\leq p$ might be constructed.

1. By marrying the two reflexive halves of the paradigm of prime divisors as they appear on either side of zero, and thus on either side of any multiple of p_j, and occupying the two holes corresponding to 1 and –1 with proper divisors, or (if possible) by some improvement on this recipe, and

2. By noting the natural sect that begins at 2 (or using the odd numbers, from 3 onwards) in the first quadrant, and ends just before the least prime $q > p$.

These two recipes are together exhaustive: indeed if we are allowed to search the entire quantum of the first quadrant, instead of just the beginning of it, to which we are confined by recipe 2, we shall be certain to find the maximum sect we are looking for, or its mirror image. The only reason for recipe 1 is that it might be easier to experiment with a given set of multisectors in S or S' over what would be, in A or A', the reflexive part of the prime paradigm, than to look for the answer in what might be a very long stretch of numbers in the first quadrant.

Call the sect determined by recipe 1 a *reflexive* or R-sect, and that determined by recipe 2 a *quadrantic* or Q-sect.

Confining our attention to the odd prime multisectors, in their natural sequence, we note that for up to four of them, the maximum R-sects and maximum Q-sects are identical, but by introducing the next prime multisector, $d_5 = 11$, the maximum R-sect exceeds the maximum Q-sect by 1, and that the inequality maximum R-sect $\geq$ maximum Q-sect,

tends to become more pronounced as additional multisectors are introduced. If we can prove this inequality is always so, we are half way there: all that remains is to prove that no odd sect can exceed the maximum odd R-sect by more than 2, or if it might, that it will have no bearing on the main question.

Lemma 3

The maximum possible odd R-sect $mR(d_{\sim j})$ of consecutive elements in S' that can be struck using all the odd prime multisectors d_2 through d_j is

(1) $mR(d_{\sim j}) = d_{j-1} - 1$,

and this sect, counting mirror images as identical, is unique.

A quick mnemonic is to note that the maximum odd R-sect always takes us to the brink of the next lesser prime than p_j, so if $p_j = p_9 = 23$, the maximum odd R-sect will be of $p_{j-1} = p_8 - 1 = 19 - 1 = 18$ consecutive elements.

Alternatively the maximum odd Q-sect will take us to the brink of the next greater prime than p_j, yielding the formula

(2) $mQ(d_{\sim j}) = (d_{j+1} - 1)/2 - 1$.

Adding the even multisector $d_1 = 2$ to any odd sect makes

(3) full sect = twice the odd sect + 1.

We see that $mQ(23) = (29 - 1)/2 - 1 = 13$, and $18 > 13$, so we shall have reached our goal if we can prove that formula (1), i.e. Lemma 3, is always true, and that no odd sect can exceed* it by more than 2, which will include a proof that the inequality (4) in

Lemma 4

(4) $d_{j-1} - 1 \geq (d_{j+1} - 1)/2 - 1$

is also always true.

Proof of Lemma 3

Suppose we have found Lemma 3 to be true (as it is) for all odd multisectors $d_{\sim j}$ when $j = 4$, and for all lesser values of j. We have

• • • •
3 5 7 3

This is the last case where the R-sect and the Q-sect are identical. As a Q-sect, it is merely the strike of odd prime divisors on the positive side of zero, notably over the numbers 3

* In fact Lemma 3 identifies the maximum possible sect for the pg of every n^2 except in case $n = 23$, for which (uniquely) there are six maximum odd sects of 19. But it will not be necessary to prove this, provided we can prove Lemma 4, when it will become evident that the maximum sect that can appear *in the segment*, for $n > 5$, must be increasingly shorter than the maximum sect that can appear *in the quantum*.

through 9. As an R-sect, we imagine both 3's in their paradigmatic places on either side of zero, and arbitrarily choose 5, 7 to fill the two holes in the strike left by the improper divisors 1 and –1.

Now consider an induction on j. To introduce the next multisector, in this case $d_5 = 11$, we take $d_3 = 5$ from its nonparadigmatic place between the 3's and replace it paradigmatically at both ends of the strike, using d_5 to fill the hole in the middle left by d_3. In each case we complete the strike, if we can, with the smaller d's already in use. Thus

$$\begin{array}{cccc} \bullet & \bullet & \bullet & \bullet \\ 3 & 5 & 7 & 3 \end{array}$$

is followed by

$$\begin{array}{cccccc} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ 5 & 3 & 11 & 7 & 3 & 5 \end{array}$$

is followed by

$$\begin{array}{cccccccccc} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ 3 & 7 & 5 & 3 & 11 & 13 & 3 & 5 & 7 & 3 \end{array}$$

and so on.

We say a multisector is in a *paradigmatic place*, or pp, if it imbeds in a point that can be identified with the associated prime in A'. Otherwise we say it is in a nonparadigmatic place np. In any R-sect, the two central multisectors begin by being np. Then by introducing the next larger multisector, the smallest of them goes from np to pp.

Because at every addition of 1 to j there is no alternative to what we may do, this validates the induction and proves the uniqueness of the result. QED

Furthermore, we are now aware that Lemma 3 will always indicate the maximum* possible sect that might appear in seg(n), provided Lemma 4 is true, in other words, provided the primes in the paradigm are not too far apart.

We now consider what we mean by 'not too far apart'. Suppose Tchebychef's theorem, that for $n > 3$ there is a prime between n and $2n - 2$, were only barely true. Then at some stage we might find a prime $p_i > 5$ with the next prime, p_{i+1}, of the form $2p_i - 3$. Suppose the next two primes > 23 were of this form, e.g. were 43 and 83, and associate all three with successive multisectors d_{j-1}, d_j, and d_{j+1} respectively. Now the maximum R-sect will be $p_{i-1} - 1 = 22$ and the maximum Q-sect will be a massive $(p_{i+1} - 1)/2 - 1 = 40$.

We need to show that this kind of occurrence is impossible, i.e. that Tchebychef's theorem, and the postulate of Bertrand it was designed to verify, is misleadingly weak. A sure way to do this is to prove Lemma 4, but I propose to defer this proof until I have completed a proof of the main theorem, assuming Lemma 4 to the true.

* Because for $n > 5$ the pg divisors cannot all be pp in the segment. For $n > 5$ the maximum sect in the segment will be shorter by an increasing margin than the maximum R-sect, provided the inequality in Lemma 4 is true by an increasing margin, as I shall prove it must be.

If Lemma 4 is true, as we are supposing, then Lemma 3 identifies the maximum possible sect that might apear in seg(n), as a result of some strike of all the prime multisectors $\leq n$, and the proof of Theorem 1 can now be completed without further difficulty. We must show that the inequality

$$n > mR(d_{\sim j})$$

is always true when j is the ordinal number of the greatest prime in the pg of n^2.

When $n = 1$ or $n = 2$, we cannot use just the odd numbers. There is no prime in the pg of 1^2, so all the integers in seg(1), 2 and 3, must be prime, and 3 is odd. There is only one prime, 2, in the pg of 2^2, so all the odd numbers, 5 and 7, in seg(2) must be prime. For larger n we can consider the odd numbers only. There can be no odd prime nearer to n on the small side than $n - 2$, and since by Lemma 3 $mR(d_{\sim j}) = d_{j-1} - 1$, the maximum odd sect must be of length $p_{i-1} - 1$, i.e. is always short of n by at least 3.* So the maximum successive strike of odd numbers by the odd primes in the pg of n^2 is at most $n - 3$, so at least one of the n odd numbers between n^2 and $(n + 1)^2$ must be an odd prime. QED

To validate this result we must now complete a proof of Lemma 4. To do this we construct a more-general theorem by substituting a continuous variable x for the stochastic variable n in Theorem 1.

Theorem 1A

If x is a real number ≥ 1, the next prime $> x$ is at a maximum distance of $2\sqrt{x}$ from x.

It will be seen that Theorem 1 merely recounts the instances of Theorem 1A where x is a perfect square.

Suppose we could show that Theorem 1A is true in all such instances of x after a certain point. How would this affect Lemma 4? First we construct a table of maximum distances between successive primes that would render Lemma 4 barely true.

* The special case of n=23 makes no difference, since 23, not being a lesser prime twin, has no effective strike in seg(n).

n of "p_n"	next "prime"	difference
0	[1]	
1	2	1
2	3	1
3	5	2
4	7	2
5	11	4
6	15	4
7	23	8
8	31	8
9	47	16
10	63	16

Table 1. Maximum distances between "primes" compatible with Lemma 4 being barely true. If n is even, the n^{th} "prime" must be $2^{n/2+1} - 1$, and if n is odd, the n^{th} "prime" must be the mean of its neighbours.

If the primes in real life were as far apart as the distances in Table 1, Lemma 4 would still be true, and the maximum quadrantic sect in these cases would always be equal to the maximum reflexive sect. Of course these distances would soon make Theorem 1 false, but they would mean we no longer need to worry about the quadrantic sect being greater than the reflexive sect, so we may take the latter as maximal.

Taking the latter as maximal, as we then may, we next construct a table of comparisons between $mR(d_{\sim j})$ and $mQ(d_{\sim j})$ supposing Theorem 1A to be barely true.

For "successive" "primes" p_{j-1}	p_j	p_{j+1}	actual maximum R-sect	theoretial maximum Q-sect	difference
11	17	25	10	11	–1
13	19	27	12	12	0
17	25	35	16	16	0
19	27	37	18	17	+1
23	31	41	22	19	+3
3001	3109	3219	3000	1608	1392
1000003	1002003	1004005	1000002	502001	498001

Table 2. Comparisons of maximum R-sects and maximum Q-sects for "successive" "primes" barely within the limits of Theorem 1A. While the R-sect remains at $p_{i-1} - 1$, the Q-sect approaches a limit of half this value. The p_{i-1} are real primes, but the p_i and p_{i+1} are the nearest odd numbers within the limits set by the theorem.

We can allow ourselves a permissible latitude, assuming that the nearest odd number

within the bare limits of Theorem 1A is "prime". If it happens to be composite, the next prime must be still farther within these limits than we had supposed.

Comparing where the tables become comparable, we see that, for $p_{i-1} = 23$, Table 2 makes a span 23, 31, 41, while Table 1 makes a wider span 23, 31, 47, so Table 2 from here on will always give values within the limits of Lemma 4 being true. In fact it begins to happen at $p_{i-1} = 13$, and from $p_{i-1} = 19$ the maximum R-sect is always greater than the maximum Q-sect. Thus if Theorem 1A is true, it locks Lemma 4 into inequality from $n = 19$ onwards. This in turn ensures that Lemma 3 must indicate the longest possible sect of the multisectors in the pg of n^2 that can appear in seg(n). And if Lemma 3 from then on always indicates the longest possible sect in seg(n) of the multisectors corresponding to the prime divisors in the pg of n^2, then Theorem 1A, and with it Theorem 1, must be true from then on. And if Theorem 1 is true from then on, Lemma 4 is true, and therefore irrelevant, from then on, therefore Lemma 3 indicates the maximum sect in seg(n) from then on, and so on *ad infinitum*.

It is a simple matter to ascertain that Theorem 1 is true for n up to 19, and is true for the intervening values of n to 23, and so on, so that Lemma 4 cannot be false for these values either. So Theorem 1, by proving its own lemma, has rendered itself true for all n, and no other argument is either necessary or relevant.

I suggest the reader pause for a moment to consider the astonishing nature of this argument. We have three theorems, Lemma 3, considered as indicative of the longest sect that can appear in seg(n), Lemma 4, and Theorem 1. They are all different, i.e. they are not equivalent, and each can vary within certain limits. None of them can be proved unless we assume first that one of them is true. We cannot prove that Lemma 3 produces the maximum sect that can appear in seg(n) unless Lemma 4 is true, we cannot prove Lemma 4 is true unless Theorem 1 is true, and we cannot prove Theorem 1 is true unless Lemma 3 produces the longest sect that can appear in seg(n). All these theorems stand or fall together. If one of them is true, they must all be true, and if one of them is false, they may all be false. Yet there is no way of proving any one of them is true without knowing that another of them is also true.

No such argument was ever envisaged by Aristotle, or even by Boole, because these two gentlemen limited themselves to a form of logic that is confined, in its algebraic representation, to equations of the first degree. The present argument is represented by a valid equation of the second degree, with real, i.e. self-confirming, roots. I first examined the mathematics of such equations in Chapter 11 of Reference 1 (all editions). Later, in Appendix 7 of the 1999 edition, I coined the new noun *elemental* to denote any of a set of consequential theorems that dovetail with each other in this way, so that they all stand or fall together. To prove that any of them is true, it is necessary first to suppose that one of them is, whereafter the resulting truth of the others confirms the truth of the initial one. It is a kind of contrary to an ordinary indirect proof, where we first have to suppose the initial proposition is false, and get a contradiction. Here we can suppose it is true, and get a confirmation.

To make their mathematical logic conform with the logic of Aristotle, Russell and Whitehead introduced an arbitrary principle, the so-called "Theory of Types", that effectively forbade all arguments that would have to be represented by equations of degree higher than unity.

In 1967 Russell, in my presence, abandoned his belief in the "Theory of Types", after studying a prepublication draft of Chapter 11 of Reference 1 in typescript. But although we both then agreed that a mathematical proof with a logical argument requiring an algebraic equation of higher degree was possible, neither of us could then imagine what such a proof might look like.

The validity of the argument here can be seen from the fact that it reenacts the formal nature of the primes, since the primes themselves comprise a recursive system. The properties of the primes up to a certain point c_1 say, fully determine the properties of the subsequent primes to c_2, and these in turn determine the properties of even more remote primes to c_3, and so on. We are interested in a particular property, notably that every prime p_{i+1}, less than c_1, is not more than $2\sqrt{p_i}$ away from p_i, which will determine that the primes up to c_2 must also possess this property, provided Lemma 4 is true. And once we have found that this property is also true of the primes up to c_2, the truth of Lemma 3 and Theorem 1 in this particular instance must lock Lemma 4 into 'true', from which point onwards none of the three theorems has any possible way of being falsified.

I have previously[1] suggested that Fermat might have used such an argument to prove his "last" theorem. The theorem is not at all obvious, and I see no reason why anyone who had not proved it should suppose it to be true. It is the only theorem about which Fermat expressed astonishment regarding the nature of the proof, which suggests that its logic might have appeared to him to be very unusual.

Irrespective of whether Fermat employed such an argument, it is certain that Newton did. His recursive formula for finding a root of an algebraic equation is equally non-aristotelian, and hunts out the root to any degree of exactitude after repeated applications. But Newton was primarily a physicist, and was more interested in the practical usefulness of his equations than in the mathematical principles they displayed. The fact that they work in his formula again assures us that these principles are correct, and they are exactly the principles I used to prove Theorem 1. In short, Theorem 1, like the root in Newton's formula, hunts out its own truth.*

* And so do Lemmas 3 and 4, since each elemental of an interlocking set is of equal status. Thus we need no longer consult Table 2 to see experimentally that $mQ(d_{-j})/mR(d_{-j}) \rightarrow 1/2$ as $j \rightarrow \infty$, because the interlocking truth of Lemma 4 with Theorem 1A now assures us *mathematically* that this must be so.

It is not the "logic" of the argument that finally settles the question: it is the evident behaviour of the arithmetic. Logic is merely the formal relationship of the propositions in a statement of proof. They have to be properly arranged, but this arrangement, however logical, is valueless if the propositions do not conform with the character of the subject matter under review. A common mistake of writers on chess is to suggest that 'logical thinking' is relevant to the game. It has no relevance whatever. To demonstrate a mate in a certain number of moves we have to envisage successive positions according to the rules of chess, which are quite arbitrary and have nothing to do with logic. Similarly how numbers behave is a matter of observation, not logic.

Now that we know for certain that every square segment has a prime occupant, we are free to consider questions that arise from it, such as how many primes can we expect to find there. What is again astonishing is that no such question has ever been asked before. It is as if our failure to prove the segments are occupied had cast a spell of invisibility over all matters concerning them. Once the spell is broken, it reveals a garden of delights, techniques so powerful we wonder why me never thought of employing them before. But we have one or two matters of unfinished business to attend to before we cross that threshold.

Returning to Theorem 1A, we can reexpress it in the form

$$(5) \qquad p_{i+1} \leq p_i + \tfrac{1}{k}\sqrt{p_i} \qquad\qquad (k = \tfrac{1}{2}).$$

The next question is, can we get a sharper value for k, and the answer is we can. We can substitute $k = \sqrt{7/4} = 0.6614378...$ when (5) becomes an equality for $i = 4$ and remains unequal for all other values of i. This depends on knowing that the jump from 7 to 11 is the biggest jump in the whole prime paradigm, when compared with the size of the take-off point.*

As soon as this jump has been achieved, k advances to the next ratchet point, which occurs at the jump of 14 integers from 113 to 127 making $k = \sqrt{113/14} = 0.7592961...$ A ratchet point is where k advances to a new value from which it cannot slip back. After 7, we know that 113 is the next ratchet point because the jump via 114 through 126 is over a maximum R-sect including 11^2. When this jump has been achieved, k advances to a value > 1, the next ratchet point being $\sqrt{1327/34} = 1.0714120...$

I can prove

Theorem 1B

The size of k in (5) *can be as large as we please, and* (5) *will be true of all numbers from some i upwards.*

* Interestingly, if we count 1 as a prime, as mathematicians born before 1900 did, $k = 1$, anticipating the truth of the Riemann Hypothesis, will just make the jump from 1 to 2.

Corollary

If $i \geq 32$, $(p_{i+1} - p_i)$ and $(p_i - p_{i-1})$ are both less than $\lfloor p_i^{1/2} \rfloor$.

We now enter the unexplored region of theoretically counting the primes in the segments. Once we see that a prime is merely a hole in the pg strike, it becomes quite easy to calculate sharp limits below and above which the prime count in any segment cannot stray. The nested application of $n/\log n$ in my Prime Limit Theorem[1] either hits these limits exactly*, or produces values so close to them that it makes no practical difference to any prediction of what we might find there.

I failed to notice in Appendix 7 of Reference 1 that my upper limit to the possible number of primes in seg(n)

(6) $\quad A + (B - 1) \qquad (A = n/\log n, B = A/\log A)$

is also an upper limit to $\pi(n)$.

By associating primes with empty spaces in the strike, we can produce calculated theorems much stronger than Theorem 1. Because $n\#$ grows so much faster than n^2, we see that $n = 5$ is the last time that $n\#$ or any multiple of it can occur within seg(n) with all the pg divisors paradigmatically placed. In this segment we have exactly two primes, 29 and 31, corresponding to the two unoccupied central points in the R-strike of 2, 3, 5, all of which are pp. From here on at least one of the pg divisors will have to be np, making at least one extra hole in the strike. In practice each np divisor leaves an average of about one extra hole, so an approximate answer for the number of primes in a segment is 2 plus the number of np divisors in the pg.

I will close with a more-difficult theorem, that I first proved in the summer of 1998.

Theorem 2

*For all natural squares, there is a prime in every half segment.***

In simple arithmetic, it means there is a prime between n^2 and $(n + 1/2)^2$, and another prime between $(n + 1/2)^2$ and $(n + 1)^2$.

George Spencer-Brown
England 0422 14 06 2000

* I call it 'exactly' when the approximating function comes to within a fraction below or above the verified upper or lower limit of the prime count, which must of course be an integer.
** Oppermann's (hitherto unproved) conjecture of 1862 is incomplete, because its wrong notation disallows the listing of 2 between 1^2 and $(1 + 1/2)^2$. See Dickson's *History* Vol I p 435. I can thus claim that this delightful theorem is mine.

Notes

It should be pointed out that any reader seeking previous literature on primes between squares will draw a blank. This is presently the only work on the subject in the entire history of mathematics.*

My policy with respect to any unsolved problem in mathematics is always to solve it first, and then read the literature to see how other people attempted to solve it and failed. It was not until I had written this paper that I began a serious search for previous literature and found none. Only then did I become fully aware of how uniquely distinct this paper is.

I cannot believe that none of the ablest of my predecessors ever made any attempt to solve this problem, so why the complete lack of literature? The only explanation I can think of is this: it is an all-or-nothing problem. Either you solve it or you don't. And if you don't solve it, what you will have to say about it will obviously be entirely worthless, so you don't say it. I had thought about the problem on and off ever since I knew it existed, but there was nothing sensible I could say about it until I had solved it.

My theorem that Lemma 3 always denotes a maximum sect has this single exception at d_{-9}. The fact never ceases to astonish me, because it means that the primes ≤ 23 can be rearranged to strike more efficiently then they do in the paradigm. It would seem *prime facie* to be impossible, but the fact that it can happen once makes it even more surprising that it can never happen again.** Fortunately I do not have to prove this, it being easier to prove that for all $n > 5$, the maximum sect that can appear in seg(n) must always be shorter than that prescribed by Lemma 3. The anomaly at 23 reappears in Rosser and Schoenfeld's 1962 inequality[2] for Euler's $\varphi(n)$, the number of positive integers not greater than n and relatively prime to n, which is that

$$\varphi(n) \geq e^{-\gamma} \frac{n}{\log\log n + \dfrac{5}{2e^{\gamma} \log\log n}}$$

for all $n \geq 3$, with the single exception of $n = 23\#$, for which 5/2 must be replaced by 2.50637. (γ is Euler's constant).

The fact that there is even *one* prime between successive natural squares is much stronger than anything hitherto proved about the distances between primes, and will render obsolete most of what is in the present text books.

Formula (6). The best upper limit to $\pi(x)$ hitherto was Sylvester's 1892 refinement of

* The subject was so forbidden, and the outcome so much in doubt, that no one had dared even to *conjecture* that Theorem 1 is true.
** The anomaly at 23 deserves a paper all on its own..

Tchebychef's result, notably that

$$0.95695 \frac{x}{\log x} < \pi(x) < 1.04423 \frac{x}{\log x}$$

for all sufficiently large x. We may ignore the lower limit since it is easy to prove that $n/\log n < \pi(n)$ for all $n \geq 17$. Sylvester's upper limit is too small until n is of the order of 10^{11}, whereas my upper limit is true for all n.

George Spencer-Brown
England 0723 06 07 2000

References

1. Spencer-Brown, George, *Laws of Form,* second German edition, Lübeck and London 1999, Appendix 7 (in English) The Prime Limit Theorem, pp 195-202, and Chapter 11 Gleichungen zweiten Grades, pp 47-59.

2. Rosser, J B, and Schoenfeld, L, *Approximate formulas for some functions of prime numbers*, Illinois J. Math, 6, 1962, 64-94.

Appendix 9

A proof of Riemann's hypothesis via Denjoy's equivalent theorem

Description of the proof

The Riemann hypothesis is true if and only if the numbers of positive and negative signs of $\mu(n)$ are asymptotically equal, and I prove it by showing that every other disposition of the signs is impossible.

Prolegomena

On 2006 06 27 I announced on the net a theorem very much stronger than Riemann's. Write $R_2(n)$ for $\text{li}\, n - 1/2\text{li}\, n^{1/2}$, and let $x^{1/2}$ denote the positive square root of x. Then for all $n>1$

(1) $\quad R_2(n) - 1/2(R_2(n))^{1/2} < \pi(n) < R_2(n) + 1/2(R_2(n))^{1/2}$.

My theorem in (1), for which I suggested an inadequate proof on the net, is about as much stronger than the RH as the RH is stronger than the PNT. In working up a more-general account of this theorem, with rigorous proofs of its validity, I discovered a neat proof, this time of Riemann's hypothesis only, on entirely different lines.

For this second proof we refer to what is called Denjoy's probabilistic interpretation, notably that the RH is equivalent to the proposition that any square-free number, taken at random, has an equal probability of containing an odd or an even number of (different) prime divisors. The description above is equivalent to this.

Legendre's formula for $\pi(n)$

Legendre's formula (*Essai* 2nd edition, Paris 1808, pp 412sq) is a recipe for calculating the exact number $\pi(n)$ of primes $\leq n$ without identifying them all. It can be written

(2) $\pi(n) = \pi(n^{1/2}) + (\Sigma\, \mu(d)\, [n/d]) - 1$

where μ is the Möbius function, and the denominators (d) are all the natural numbers that have no large prime p in their decomposition. A prime p is 'large' (in relation to n) if $p > n^{1/2}$.

Analysis. The formula in (2) works correctly because its summation term yields the number of numbers $\leq n$ that are not struck out by the Eratosthenes procedure of striking out those of them that are divisible by a prime q that is 'small' in relation to n, i.e. is such that $2 \leq q \leq n^{1/2}$. The procedure will obviously still work if we redefine one or more of the large primes as 'small'. The unstruck numbers include 1, which is not nowadays classed as a prime. Students of arithmetic born before 1900 were taught that 1 is the least prime, making Goldbach's conjecture apply to all even numbers including 2. Present-day arithmeticians find it more convenient to exclude 1 from the class of prime numbers, making it the unique natural number whose number of distinct prime divisors is zero.

The function $\mu(d)$ can now be defined as equal to zero if d has a repeated divisor (other than 1), and otherwise to +1 if the number of its (different) prime divisors is even, and to –1 if it is odd. Since 1 is not struck out by the sieve of Erastosthenes and is also included in the count of "primes" calculated by the section $\Sigma\mu(d)\ [n/d]^*$, the count must be reduced by one in either case, and then to get the complete answer the number of small primes (q) used as strikers must be added to the total.

Illustration of Legendre's formula with n = 20

d	$f(d) = \mu(d)\ [n/d]$
1*	+20
2*	–10
3*	–6
5	–4
6*	+3
7	–2
10	+2
11	–1
13	–1
14	+1
15	+1
17	–1
19	–1
	Σ +1

$\Sigma\mu(d^*)\ [n/d^*] = 7$

$7 - 1 + \pi(n^{1/2}) = 8 = \pi(20)$

The (d*) are the denominators with no large prime in their decomposition. The small primes 2, 3 must be known explicitly, then the number of large primes 5, 7, 11, 13, 17, 19 can be calculated without any of them being identified.

In the table above we see an illustration of the use of Legendre's formula to calculate $\pi(n)$ for n = 20. The starred terms, with no large prime divisors of d, are used to calculate the number of large primes $\leq n$. Notice I have used all the (d) that yield an $f(d)$ other than zero, and the sum of these, for any n, must always be 1, since only one number, 1 itself, remains unstruck if we use all the primes.

The stage is now set for my proof of Denjoy's equivalent to Riemann's hypothesis.

First we get rid of the 1, which is the only number left standing after my extension of Legendre's procedure. To do this we remove it at the beginning, quite legitimately,

* The fact that, with d unrestricted, the formula $\Sigma\mu(d)\ [n/d] = 1$ is true for all n was first noted by Meissel in *Observationem quaedam in theoria numerorum*, Berlin 1850, and proved by Sierpiński in *Elementary theory of numbers*, Warsaw 1964, pp 180 and 181.

because it is neither prime nor composite, and so does not belong to either of the two complementary classes, composites and primes, to which we reduce the number system in accordance with modern practice.

We thus further rectify the procedure by making the following change:

use $f(d) = \mu(d)\ [(n-1)/d]$ for d=1

and use $f(d) = \mu(d)\ [n/d]$ for all other values of d.

Now rework n=20 using the new procedure

d	$f(d)$		
1	+19		
2	–10		
3	–6		upper section
5	–4		
6	+3		
7	–2	sum to half way	
10	+2	+2	
11	–1		
13	–1		
14	+1		lower section
15	+1		
17	–1	sum in 2nd half of n	
19	–1	–2	
	0	sum complete.	

I have divided the terms into two sections, in the upper of which each $f(d)$ consists of $\mu(d)$ multiplied by some positive number >1, and in the lower each $f(d)$ is simply $\mu(d)$. (We should note that every value of $\mu(d)$ other than the first must appear in the lower section of one or more natural n.) In the upper section the $f(d)$ sum need not be exactly the sum of the $\mu(d)$ pluses and minuses to this point, but it is obviously positively correlated to it. (For example if all the terms were negative the answer would be negative, and vice versa.) In the lower section the sum of the terms is exactly the sum of the $\mu(d)$ in the section. Call a series of such terms an LSB series.

Recall that, by Denjoy's equivalent, the Riemann hypothesis is true if and only if the algebraic sums of the pluses and minuses of $\mu(d)$, taken progressively at unit increments of d, vary asymptotically around zero*. Suppose it is untrue. This can only mean that the sums must vary around some number other than zero**, or around no number.

* This means we can get it as close as we like to an average of zero difference between the two.

** This means we cannot get the average difference as close as we like to zero after n has reached a certain size, but can get it as close as we like to some other number.

Suppose they vary around a number such that the upper sections of all LSB series, taken progressively, vary in aggregate around +2, which we recognize is the sum of the upper-section terms for $n = 20$, so we may take this n as a typical example. Now the sum of the lower-section terms for this n must be exactly -2 to compensate the upper section. So double n *to* $2n = 40$. The aggregate of pluses and minuses in the upper section of $2n = 40$ is exactly what it was in the whole of $n = 20$. But it contains two more minus signs than did the upper section of $n = 20$, so its sum is likely to be reduced towards or beyond zero. Suppose by an unlikely chance it is still +2. The lower section of $2n = 40$ must again be -2 to compensate this, so repeat the procedure by doubling $2n$ to $4n = 80$. Now the upper section of this new number $4n = 80$ must contain four extra minus signs, leaving it equally biased towards zero.*

These unlikely chances cannot continue for ever, because every time we doubled the argument we would have to add an average of two more minus signs to the upper-section terms of the new doubled argument, so there must come a time when the sum of the upper-section terms of the new doubled argument is reduced to or beyond zero. Suppose it is reduced to zero. Then the sum of the lower-section terms for this argument will also be zero, and there will be no tendency in either direction when it is doubled again. But suppose the sum in the upper section is reduced beyond zero to a negative value. Now the lower-section sum for this argument must be positive, and the whole process must play itself out again, this time in the opposite direction.

Recall, finally, that the lower sections of every argument consist of increasingly protracted sets of consecutive terms of $\mu(d)$, and that all values of $\mu(d)$ except the first must be presented, either explicitly or implicitly, in the lower sections of the LSB series for one or more natural n.

Because of the negative feedback between the two sections, any difference between them, however small, can be seen to be self-annihilating.

I have thus demonstrated, by an easily-generalized example, that the average differences between the numbers of the positive and the negative signs of $\mu(n)$ must approach zero asymptotically. And because they have been shown to approach a specific number (i.e. zero) they cannot be said to approach some other number or no number.

It follows that the numbers of positive and negative signs of $\mu(n)$ must be asymptotically equal and the Riemann hypothesis must therefore be true. [2008 04 15]

* We must distinguish, in the upper sections, the aggregate of the plus and minus *signs* from the aggregate of the $f(d)$ *terms*, which in the upper sections are not necessarily the same. In the lower sections they must be the same, and this is what allows the proof. For example, when $n = 20$ the aggregate of the upper-section *terms* is +2, but the aggregate of the upper-section *signs* is -1, i.e. there is a total of one more minus signs than plus signs. In the upper section for $n = 40$ there are three more minus signs than plus signs, i.e. two more than before, as predicted.

Abbreviations

PNT = prime number theorem, notably that $\pi(n)/(n/\log n) \to 1$ as $n \to \infty$

RH = Riemann('s) hypothesis that

$\zeta(s) = \Sigma n^{-s} = \Pi(1-p^{-s})^{-1}$

in which n runs through the natural numbers 1, 2, 3, ... and p through the primes 2, 3, 5, ... cannot be equal to zero for nonreal s other than of the form $s = 1/2 + iy$ with $i = \sqrt{-1}$ and y real.

LSB series = Legendre/Spencer-Brown series as explained in the text.

This proof by Professor George Spencer-Brown of Georg Friedrich Bernhard Riemann's hitherto unproven hypothesis, originally proposed in 1859, was first published on the internet in 2008 04 24, and in hard copy in this book in 2008 09 15.

An analysis of the proof

From the proof we can see that the successive Möbius values are not at all random, as many commentators have mistakenly supposed, but follow what is called by unsophisticated gamblers a "maturity of chances" hypothesis. This is a supposition that, for example, after a run of successive black numbers at the roulette table, the probability of a red number appearing next is increased to "redress the balance". But the (to some extent) empirically verified hypothesis of probability determines axiomatically that the result of the next spin will be independent of previous results, so the player loses from the inclusion of a zero number that renders the probability of red or black slightly less than 1/2. But if the casino were naïve enough to offer evens against a + or − appearing in any continous set of consecutive nonzero values of $\mu(n)$, the player could win a fortune by betting against the trend.

Offhand I can think of no series of naturally-produced numbers, other than $\mu(n)$, for which a maturity-of-chances hypothesis happens to be true.* Can any of my readers?

An unintended confirmation of this is provided by John Derbyshire in his book *Prime obsession* (New York 2003), which is the best and most complete account of the Riemann hypothesis I have seen. It contains, moreover, fewer serious mistakes than any other account I have read, though it is of course impossible to write a book of this size (422 pages) without including some mistakes. He makes the common mistake of suggesting that a series of consecutive values of $\mu(n)$ might be random in respect of their + and −

* Of course it is also true of primes and composites, but these categories are not independent of $\mu(n)$.

signs (p 322). In a previous page (250) he quotes sets of Mertens's function (cumulative $\mu(n)$) that he says tell us very little except that their absolute value increases as n does. In fact they tell us a great deal, and if he had noticed it he might, admittedly with a fair amount of further detective work, have discovered my beautiful proof of Riemann's hypothesis several years before I did.

On page 322 he correctly points out that the average difference between n randomly-produced +1's and −1's is $\sqrt{n}$. To convert Mertens's function into a corresponding function for square-free (n) we must multiply each n by $6/\pi^2$, or about 0.608. From his second set of figures we find arguments

1000	2000	3000	4000	5000	6000	7000	8000	9000	10000

conversions

608	1216	1824	2432	3040	3648	4256	4864	5472	6080

roots of conversions

25	35	43	49	55	60	65	70	74	78

Mertens's values for original arguments

2	5	−6	−9	2	0	−25	−1	1	−23.

It is evident that the final set of values (considered absolutely as differences) are ridiculously below what they should be if the original sets of +1's and −1's were randomly produced. I was going to do a table of his third set of arguments, in millions, whose values are equally impressive, but the set I have tabulated all denote unrandomness so obviously that I will leave the tabulation of his third set to the reader, for the good feeling of being part of the research. What they show is that the successive Möbius +1's and −1's are unrandom to an enormous degree, being hugely biased towards a maturity-of-chances hypothesis.

My strong theorem at the beginning of this memoir shows that the primes are similarly unrandom, in the sense of being much more evenly-spread than if they had been randomly placed in the sequence under review.

I had experimented with Legendre's method of counting primes for many years, convinced that it could lead to a very elementary proof of the PNT, but ironically could not see how to do it until faced with the more heroic prospect of proving the Riemann hypothesis. In this case it seems to be impossible to prove the one without simultaneously proving the other.

In common with other experts in the field, I had begun to suspect that the RH is a problem in elementary arithmetic and not in analysis, as many of us, probably including Riemann, had previously thought. [2008 04 18]

My Denjoy proof of Riemann's hypothesis looks so simple that there seems to be little we can say about it except to reenact it. But one or two things are worth remarking.

Consider the natural counting numbers from 1 up to n and decompose each of them into its prime components. Add up all the exponents of the prime components of each of them. A well-known equivalent to Riemann's hypothesis is the proposition that, for any one of these numbers selected at random, the sum of the exponents of its prime components is equally likely to be even or odd.*

In short, the RH states that (3) for any fixed $\varepsilon>0$ $\lim\limits_{n\to\infty} \frac{\lambda(1)+\lambda(2)+\ldots+\lambda(n)}{n^{1/2+\varepsilon}} = 0$,

where λ is Liouville's function (=+1 if the sum of the exponents is even, −1 if it is odd) whereas the PNT is the much weaker theorem obtained by substituting n for $n^{1/2+\varepsilon}$.

What Professor Denjoy noticed is that we need not consider any of the numbers $\leq n$ with a prime exponent of two or more. In short, if all square-free numbers (d), say, could be shown to be equally likely to have an even or an odd number of prime divisors, the truth of the Riemann hypothesis would follow. Hence all that is required is to show that the average of the differences between the numbers of plus and minus terms of $\mu(d)$ varies around and is asymptotic to zero as n increases without limit.

We should next note that we can reclassify one or more of the large primes $\leq n$ as "small" without affecting the result. Suppose for $n=20$ we use, as we must, the small primes 2 and 3, and add to them the extra primes 7 and 13. For the relevant $f(d) \neq 0$ we now have

d	$f(d)$
1	+19
2	−10
3	−6
6	+3
7	−2
13	−1
14	+1
	Σ +4

$\Sigma f(d)=4$. Add to this the redefined "small" primes 2,3,7,13 gives $4+4=8=\pi(20)$.

It is further evident that the use of the extra "small" primes as strikers in the sieve of Eratosthenes will make no difference to the result, apart from the fact that they will strike out themselves, so provided we add them back to the total the result will be the same as that of the Legendre/Spencer-Brown procedure, say LSB.

Now notice what I did to prove the Riemann hypothesis. By using all the square-free $(d) \leq n$ I in effect reclassified *all* the primes $\leq n$ as "small". This ensures that whenever

* Cf Borwein and others, *The Riemann hypothesis*, Burnaby 2008, pp 6 and 7.

I use the rectified Legendre procedure to count the remaining large primes $\leq n$, I shall get the answer zero. I thus made the LSB procedure useless for counting primes, but extremely useful for proving the Riemann hypothesis, since with an $f(d)$ total of zero for every n, I can split the series of $f(d)$ terms anywhere I choose into two sections, and whatever the total $+t$ (say) in one section will be balanced against a total of $-t$ in the other.

The most instructive place to split the series is at the point where $(n/d)<2$. We then get a lower section of $f(d)$ terms all of value ± 1. They comprise all the values of $\mu(d)$ for square-free (d) from this point up to n. The upper section of the series will now consist of the early terms of $\mu(d)$ for square-free (d) magnified by a factor ranging from 2 up to $n-1$.

The fact that the plus and minus values of $\mu(d)$ are magnified in the upper sections ensures that any excesses of positive or negative terms of $\mu(d)$ that appear in the lower sections become overcorrected in the upper sections, to which they are transferred as n increases. Effectively what happens is that both sections correct each other towards an asymptotic average of zero.

The Riemann hypothesis, as we have noted, must be true if the pluses and minuses of $\mu(d)$ are merely equiprobable. But the fact that any excess of one sign over the other that begins to appear in a lower section gets magnified when this part of the lower section gets incorporated in an upper, ensures that the difference between them stays closer to zero than it would if the successive plus or minus signs of $\mu(d)$ were merely randomly distributed like successive falls of a coin.

Consider the last two terms of my rectified $(f(d))$ for $n=20$, notably with $d=17$ and $d=19$. Both of these terms must, with complete certainty, be -1. In a random sequence of terms, the value of each term within the range considered must be completely unaffected by the values of previous terms. But in this case, as we see, the values of these terms are entirely determined, and therefore completely predictable, by the values of the previous terms. This means that whatever patterns the values of $(\mu(d))$ can display, they cannot be random.

In particular it means that both of the arguments 17, 19 for $n=20$ can have only an odd number of prime divisors, and since the cube root of 20 is less than three and both of these numbers are odd, they must both be prime. Thus in general, once we have ascertained the number of prime divisors of $d=1$ and found it to be the even number zero, making $\mu(1) = +1$, all subsequent values of $\mu(d)$ for square-free d can be found by a simple elementary algorithm without having to know any divisor of any d whatever. *

* The fact that the prime parity of 17 is odd, for example, completely determines the fact that the prime parities of 237 618 987 and 1 009 003 027 are both even. Although this is unavoidable when we think about it, the way we present it to ourselves seems to invest it with an aura of incredibility. Certainly this astonishing fact was neither known nor suspected before this publication.

The astronomer August Ferdinand Möbius was born in Schulpforta in 1790 11 17, and his famous number-theoretic function predates Riemann's paper by at least two decades, and the use of it by others, e.g. Euler, by much longer. That the mathematical profession has to this day failed to notice the most significant property of this function, and has compounded the failure by wrongly supposing its terms to be in some respect "random", defies belief.*

So we can discover the parities of the numbers of primes in all square-free numbers simply by considering the parities of the numbers of primes in lesser numbers hitherto thought to be unrelated to them, something that Hardy and Wright supposed could not be done.

I remember in the 1950's Lord Cherwell, then a colleague of mine at Christ Church Oxford, and I used to write to the surviving author inclosing various formulas to predict prime numbers larger than the last one determined, suggesting he incorporate them in the next edition of the book. He always refused to do so, we thought wrongly.

None of the suggestions was as ingenious as the one I have just proved, notably that the successive nonzero values of $\mu(d)$ behave like the falls of a magic coin that, from the moment it is struck, remembers exactly how many times it has fallen with one side or the other uppermost, and whenever one side exceeds the other, biases itself towards the other until the excess is eliminated.

This self-correcting property is more typical of living organisms, and thus surprising when we find it in what we thought of as an inanimate number system. It of course makes no difference to the truth of the Riemann hypothesis, which would still be true if the successive values of $\mu(d)$ for square-free (d) merely behaved like the falls of an ordinary unbiased coin, instead of like a being that watched what it was doing and modified its behaviour accordingly. But the fact that the differences between the plus and minus values for square-free (d), because of this self-correcting tendency, stay *closer* to zero than would the differences between the two sides of successive falls of an unbiased coin, suggests that the Riemann hypothesis might be in some way *more than* true.

This could be interpreted as saying that Riemann's 1859 proposition might be too weak, and that my stronger propositions below might more nearly represent the true state of affairs. In short, that a stronger theorem than Riemann's might also be true.

By Professor Denjoy's equivalent theorem, the Riemann hypothesis is seen to be true if and only if the successive +1's and –1's in $\mu(d)$ are equiprobable. Since it is

* This terrible mistake stems from the prevailing myth that the primes are somehow "randomly" distributed, if not wholly so, than at least partially. On the contrary, the primes are an example of the most beautiful and paradoxical form of order imaginable, a perfectly ordered series, i.e. completely unrandom, that never repeats itself. Every point in it signposts the way to every other point, and no two points are confused. And as with this, so with August Möbius's beautiful function that perfectly reflects it.

meaningless to select a term at random from an infinite set, this can only mean that their differences over the number of LSB terms displayed, which we may call their average differences, tend towards zero as n increases without limit. It does not necessarily mean that the nonzero terms have to be randomly distributed, like the falls of an ordinary unbiased coin, and evidently, as I have proved, they are not.*

This indicates that not only the Riemann hypothesis itself was unclear in the minds of its previous investigators, but that other things about it were unclear too, since if this associated lemma had been clarified, it is hardly likely that such an obvious clue** would not have led to a speedy proof of the hypothesis. But this was not the case, and I cannot recall any mathematical problem I have solved where the muddle in the minds of most of its previous investigators was more extensive or complete.***

For example I have seen no previous account of the Riemann hypothesis and its environs that does not make the mistake of suggesting that the nonzero terms of $\mu(n)$ must be randomly distributed, with probabilities of a half each, for the RH to be true. All that is required is that they be equiprobable. That they be equiprobable *and* random is clearly not required, and is equally clearly not true. By proving them to be equiprobable I proved the RH, but by simultaneously proving them to be non-random in the particular way they are, I proved something that suggests Riemann's original guess was not strong enough, and that the real theorems associated with this branch of arithmetic are in fact much more constraining than he supposed. This turns out to be the case, as we shall presently see.

It also highlights another example of the muddled thinking of previous investigators. Although not explicitly stated, it is nevertheless implied in all the accounts I have seen, that Riemann's guess must be the holy grail of all numeric theorems, and therefore that it must impose the narrowest possible limits on the range of the prime count. In fact these limits are much narrower than Riemann's guess requires them to be.

What mathematicians tend to do when they cannot prove or disprove a proposition is to invent equivalent statements that they think might be easier to decide. Denjoy's probabilistic interpretation of Riemann's guess, that I found easier to prove, is a case in point.

* Much of the confusion here springs from the fact that the Möbius +1's and −1's comprise neither proper numbers nor proper signs, but are merely convenient ways of saying whether the prime parity of a number n is even or odd. Boole made a similar mistake, using numbers and signs to represent truth values, which I corrected in the calculus of indications by eliminating both. We could do the same here, substituting 'even' for +1 and 'odd' for −1, but even this is too specific, since in a calculus of only two values all we need to do is to distinguish them without saying which is which. Thus we can generalize $\mu(n)$ to $\sigma(n)$, say, with a single ambiguity that we can resolve at any point, conveniently at n=1. Then if $\sigma(1)$=+1, we have $\mu(n)$, and if $\sigma(1)$=−1 we have $-\mu(n)$. In either case the values denote prime parities for which we need not factorize n.
** In fact there was another obvious clue in the surprisingly small sizes of the values of Mertens's function, but it too was overlooked.
*** Of course I do not discount explorers such as von Koch, Denjoy, and Littlewood, without whose preliminary findings my task would have been much harder.

Another such interpretation, by von Koch, concerns constraints on the errors of the prime count from some proven asymptote to it, such as $n/\log n$ or li n. von Koch proved in 1901 that Riemann's conjecture is equivalent to the errors of the prime count $\pi(n)$ being constrained, for all n above some large-enough value, to within the range of li n $\pm K\sqrt{n}\log n$ with K constant.

Improving on von Koch's (and therefore Riemann's) extremely gross limits to the prime count, Littlewood, in 1907, correctly concluded that 'if the P.N.T. were true 'with error about $\sqrt{n}$', the R.H. would follow.' (*Miscellany**, Cambridge 1953.) He means here by P.N.T. the proposition that $\pi(n)/\mathrm{li}\, n \to 1$ as $n \to \infty$, rather than the more-usual $\pi(n)/(n/\log n) \to 1$. Both propositions have been proven true, but the former converges faster than the latter.

Littlewood made no sustained attempt to investigate his or von Koch's not-too-difficult restatement of Riemann's hypothesis, either theoretically by trying to prove it, or empirically by comparing it with known prime-counts. Even more astonishingly, for the next 99 years, nobody else looked for an empirical verification of von Koch's equivalent until I announced on the net, in 2006, the stronger theorem that

(4) $\mathrm{li}\,n - (\mathrm{li}\,n)^{1/2} < \pi(n) < \mathrm{li}\,n + (\mathrm{li}\,n)^{1/2}$

is true for all prime counts when $n > 1$ and $x^{1/2}$ is the positive square root of x. (Or we can be more sophisticated and apply the negative square root to the left-hand side of the inequality and the positive square root to the right.)

This is increasingly stronger than Littlewood's lemma, and therefore stronger than Riemann's hypothesis, because it substitutes $(\mathrm{li}n)^{1/2}$ for Littlewoods's $n^{1/2}$, and in the range considered, li$n$ is greater than 1 and lesser than n.

The reason I decided to publish first my Denjoy proof of Riemann's much weaker theorem (formerly known as his 'hypothesis'), is because it requires no principle or axiom that was not available to Euclid, and therefore makes no demands on the reader other than those that have been traditional and available in school text books for the last two thousand years.

My proofs that do make further demands, are of much stronger theorems than Riemann's. Calling

1. $m(n) = \mathrm{li}n - 1/2(\mathrm{li}n^{1/2})$, and
2. $m'(n) = S_d(n)$ with d=0 (see p 221),

and calling $r(n) = m(n)$ or $m'(n)$, I can prove that

(5) & (6) $r(n) - 1/2(r(n))^{1/2} < \pi(n) < r(n) + 1/2(r(n))^{1/2}$

assuming, as usual, that $x^{1/2}$ denotes the positive square root of x.

* For consistency I exchanged n for x in the text.

These I call my strong theorems because, as is evident from a cursory glance, they are very much stronger than anything Riemann conjectured, and also stronger than my theorem in (4). They indicate that the denominator in (3) can be reduced, possibly to $1/2\ (n^{1/2+\varepsilon})$.

I will not prove them here, because there is no need to burden the reader with additional new ideas until a later date, when he, or she, will have become familiar with at least one proof of Riemann's hypothesis and is ready for more.*

My Denjoy proof, in summary, runs as follows.

1. What Professor Denjoy showed is that the RH is equivalent to the proposition that the number of primes in a square-free $d \leq n$ of any size is even or odd with equal probability.

2. I rectify Legendre's method of counting large primes in n and then corrupt it to give the answer zero for all (n).

3. I split the rectified Legendre terms into two sections, upper and lower, so that the lower sections eventually include all values of $\mu(d)$ for square-free $(d) > 1$.

4. I show that the average algebraic sum, i.e. the sum divided by the number of LSB terms displayed, in each section varies around and is asymptotic to zero as n for the $f(d)$ terms increases without limit.

5. Since the lower sections eventually include the values of $\mu(d)$ for all square-free $(d) > 1$, and their signs are by the previous proposition equiprobable in the limit, the RH, quod erat demonstrandum, must be true.

6. In addition, since the upper sections eventually contain all the values of $\mu(d)$ for square-free (d), but magnified by various factors ranging from 2 up to $n-1$, that are independent of the signs of $\mu(d)$, and the average differences between the plus and minus values of these magnified terms also tend to zero as n increases, this fact constitutes a second proof of Riemann's hypothesis, since if an average of a set of magnified differences tends to zero, then the average of the same set of differences unmagnified must also tend to zero.

It should be noted that in proving the RH in this elementary way, I have also made a simple elementary proof, much simpler than Selberg's, of the prime number theorem, since the one implies the other. I have furthermore decided other propositions, such as a version of Goldbach's conjecture, that Riemann's hypothesis implies. I can also show that my strong theorems imply the truth of all forms of Goldbach's conjecture, and of other previously undecided propositions about prime numbers.

* It will also give us something further to discuss when I am invited to talk about my proofs to academic audiences.

Spencer-Brown's cascade

We can redescribe the Legendre/Spencer-Brown procedure LSB to be

(7) $\mathrm{LSB}(n) = n - 1 + \Sigma\mu(d)[n/d]$ with various restrictions on d.

(7.1) Restricting the (d) to primes not greater than $|n^{1/2}|$ and their mutual multiples yields the number of large primes not greater than n.

(7.2) Derestricting the (d) to all integers > 1 makes the $\mathrm{LSB}(n)$ equal to zero for all (n).

The fact that (7.2) = 0 for all (n) allows us to demonstrate one of the most astonishing facts of arithmetic, notably that we can know how any natural number will factorize without factorizing it, and by an algorithm so simple that a child of six can do it.

It is done by what I call a cascading algorithm. This is a spinoff from the procedure I adopted in my Denjoy proof of Riemann's hypothesis.

From our knowledge that $\mu(1) = +1$ because 1 has zero prime divisors and zero is an even number, we proceed to discover $\mu(n)$ for every subsequent number n without having to factorize any such n, as follows.

Take $n = 2$. The LSB terms for this n will be

d	$n-1$ $+f(d)$	running total
	+1	+1
2	−1	0

Since the running total, by (7.2), must reach zero for every n, the last value of $f(d)$ must in this case be −1 to reach this total, so $\mu(2) = -1$ and 2 therefore has an odd number of prime divisors.

Take $n = 3$. Now the LSB series will be

d	$n-1$ $+f(d)$	running total
	+2	+2
2	−1	+1
3	−1	0

so $\mu(3) = -1$ and so 3 has an odd number of prime divisors.

Take $n = 4$

d	$n-1$ $+f(d)$	running total
	+3	+3
2	−2	+1
3	−1	0
4	0	0

since the penultimate total is already zero, $\mu(4) = 0$ and so 4 has a square divisor. Therefore we can ignore 4 in subsequent cascades.

Take $n = 5$

d	$n-1$ $+f(d)$	running total
	+4	+4
2	−2	+2
3	−1	+1
5	−1	0

so $\mu(5) = -1$ and so 5 has an odd number of prime divisors.

Take $n = 6$

d	$n-1$ $+f(d)$	running total
	+5	+5
2	−3	+2
3	−2	0
5	−1	−1
6	+1	0

so $\mu(6) = +1$ and so 6 has an even number of prime divisors.

Take $n = 7$

d	$n-1$ $+f(d)$	running total
	+6	+6
2	−3	+3
3	−2	+1
5	−1	0
6	+1	+1
7	−1	0

so $\mu(7) = -1$ and so 7 has an odd number of prime divisors.

Take $n = 8$

d	$n-1$ $+f(d)$	running total
	+7	+7
2	−4	+3
3	−2	+1
5	−1	0
6	+1	+1
7	−1	0
8	0	0

so $\mu(8) = 0$ and so 8 has a square divisor and can be ignored in subsequent cascades.

Take $n = 9$

d	$n-1$ $+f(d)$	running total
	+8	+8
2	−4	+4
3	−3	+1
5	−1	0
6	+1	+1
7	−1	0
9	0	0

so $\mu(9) = 0$ and so 9 has a square divisor and can be ignored in subsequent cascades.

Take $n = 10$

d	$n-1$ $+f(d)$	running total
	+9	+9
2	−5	+4
3	−3	+1
5	−2	−1
6	+1	0
7	−1	−1
10	+1	0

so $\mu(10) = +1$ and so 10 has an even number of prime divisors.

Notice we are entirely unconcerned whether the d-arguments are prime or not, or whether or not the divisions n/d are exact. The cascade does not require this information. All it requires is what the previous cascades have told it. We also see that there is no need to list the final term for any n, since the answer must be the penultimate term in the running total with the sign reversed.

It is easy to see how our six-year-old will continue to find the values of $\mu(n)$ for every subsequent value of n without ever having to factorize, or even to find one divisor of, any n whatever.

Whenever I make a discovery like this that the best of the rest appear to have overlooked for the past eight thousand years or so, I find it difficult not to suspect that somebody else might have thought of it before. I usually consult an "authority" to make sure. One such "authority" is H M Edwards, *Riemann's Zeta Function*, New York 1974, a detailed account in 336 pages of how not to prove the Riemann hypothesis.

In p 268 he remarks that it is 'plausible to say that successive evaluations of $\mu(n)$ are "independent" since knowing the value of $\mu(n)$ for one n would not seem to give any information about its value for other values of n.'

This tells us that my law of succession of ($\mu(n)$), how to find its next value from its previous values, was not known before I announced it. Unfortunately nearly all such useful "authorities" are secondary sources, they copy uncritically what other authors have written, including mistakes, and do no experiments of their own that might have led to an independent observation. Consequently they are useful only for verifying historical facts, since whenever they venture an opinion it is never based on arithmetical evidence.

Mr Edwards's opinion above is both wrong and nonsensical, wrong because it is untrue (he suggests there is no law of succession for $\mu(n)$, and I have just demonstrated a very simple one), and nonsensical because it is arithmetically evident that we must always be able to decide what will come next in a calculus whose elements are 1, 2, 3, 4, ..., and in any definite function of these elements, so it must be possible, at least in principle, to find the next value of $\mu(n)$ from its previous values, and what I did was discover a way of doing so that is simple enough for a child of six to operate.

It is nice because it gives us a way to find the value of $\mu(n)$ for any n merely from its previous values, instead of from factorizing n and checking to see which of its factors is prime and then checking again to see if any of them is repeated, and if not counting the prime divisors to see if their cardinal number is even or odd.

It is important because it shows that apparently complicated properties of n, whether it has a repeated divisor or if not, an odd or an even number of prime divisors, previously thought to be determinable only by factorizing n and counting its various kinds of factors, are actually determined by its ordinal place in the system, and its divisors must for this reason fall into the appropriate category without having to be counted or even observed at all.

G Spencer-Brown
England 2009 04 17 the 40th anniversary of the publication of this book.

References

Denjoy, Arnaud, Probabilités confirmant l'hypothèse de Riemann sur les zéros de $\zeta(s)$. C. R. Acad. Sc. Paris, t. 259 (9 novembre 1964) pp 3143 - 3145.

Koch, Helge von, Sur la distribution des nombres premiers, *Acta Math.*, 24(1901), 159 - 182.

The prime sum theorem PST

The most astonishing fact about the prime sum theorem is that nobody noticed it before. It is that the sum of the primes up to a given prime p is asymptotically equal to the prime count at p^2. In signs

$$(8) \qquad \sum_{i=1}^{n} p_i \sim \pi(p_n^2).$$

For example the sum of the primes to 23 is 100, and the prime count to $23^2=529$ is 99. When the sum exceeds the count, as in the example, it is called a Spencer-Brown or brownian crossing.

Excersise 1. Prove that the number of brownian crossings is infinite.

Excersise 2. Prove that for all primes except 3 the PST will yield a closer approximation to the prime count than Gauss's guess of $n/\log n$, or $p^2/\log(p^2)$ in this case.

Closing remarks

Publishing a book is a venture in which persons other than the author take part, and it is usual for the author to acknowledge their contributions where the list of those mentioned does not do so. My chief thanks for this edition are due to Thomas Wolf for typesetting it. He also typeset his translation in the German edition.

I am additionally indebted to James Flagg, who informed me that I had in effect proved Riemann's hypothesis in Appendix 7, and supplied me with the necessary works in analytic number theory to make the fact apparent.

My original purpose in Appendix 7 was to make a simpler formula than Riemann's for approximating the prime count at a given n. The formula I produced there is not entirely satisfactory because the correction factor t is not quite constant. It took me five more years to remedy this with the formula

$$S_d(n) = \sum_{k=1}^{[n^{1/2}]-1} \frac{2k+d}{\log(k+1)^2} + \frac{<n^{1/2}>(2[n^{1/2}]+d)}{\log([n^{1/2}]+1)^2} + \frac{1}{2}$$

where $[x]$ is the integer part and $<x>$ the fraction part of x, and $n^{1/2}$ denotes the positive square root of n. The correction factor d is constant and can be set for any particular large n and will still give acceptably accurate answers for smaller n.

In the table that follows, my specification $j = 11$ records how high we have taken k in Riemann's explicit formula (Appendix 7, formula (0)) to calculate the result in question, and $d = 0.06886$ is set to agree with Riemann's answer (using this j) for $n = 3\text{x}10^6$.

The arguments up to 10^8 were programmed by me in a hand-held calculator limited to 12 digits. The argument 10^{12} is beyond the accurate range of this calculator and was programmed in a personal computer by Mr Chen Kupperman of Haifa.

n	$\pi(n)$	$S_d(n)$	difference	$R_j(n)$	difference
50 000	5 133	5 133.691	+1	5 133.401	0
100 000	9 592	9 587.759	–4	9 587.434	–5
150 000	13 848	13 844.326	–4	13 843.985	–4
200 000	17 984	17 981.877	–2	17 981.534	–2
250 000	22 044	22 035.144	–9	22 034.818	–9
300 000	25 997	26 023.687	+27	26 023.336	+26
350 000	29 977	29 959.891	–17	29 959.550	–17
400 000	33 860	33 852.387	–8	33 852.051	–8
450 000	37 706	37 707.487	+1	37 707.114	+1
500 000	41 538	41 529.822	–8	41 529.508	–8
550 000	45 322	45 323.293	+1	45 322.978	+1
600 000	49 098	49 090.841	–7	49 090.533	–7
650 000	52 831	52 834.941	+4	52 834.644	+4
700 000	56 543	56 557.665	+15	56 557.737	+15
750 000	60 238	60 260.737	+23	60 260.465	+22
800 000	63 951	63 945.694	–5	63 945.415	–6
850 000	67 617	67 613.777	–3	67 613.519	–3
900 000	71 274	71 266.167	–8	71 265.905	–8
950 000	74 907	74 903.826	–3	74 903.572	–3
1 000 000	78 498	78 527.635	+30	78 527.402	+29
1 050 000	82 134	82 138.422	+4	82 138.184	+4
1 100 000	85 714	85 736.853	+23	85 736.626	+23
1 150 000	89 302	89 323.591	+22	89 323.367	+21
1 200 000	92 938	92 899.203	–39	92 898.986	–39
1 250 000	96 469	96 464.208	–5	96 464.011	–5
1 300 000	100 021	100 019.121	–2	100 018.925	–2
1 350 000	103 544	103 564.355	+20	103 564.170	+20
1 400 000	107 126	107 100.335	–26	107 100.153	–26
1 450 000	110 630	110 627.424	–3	110 627.250	–3
1 500 000	114 155	114 145.981	–9	114 145.811	–9
3 000 000	216 816	216 816.291	0	216 816.326	0
10^8	5 761 455	5 761 548.223	+93	5 761 551.872	+97
10^{12}	37 607 912 018	37 607 910 550.8	–1467	37 607 910 542.9	–1475

Table of comparisons of $\pi(n)$ with Spencer-Brown's $S_d(n)$ and the Riemann/Spencer-Brown $R_j(n)$ for small round values of n, using $d = 0.06886$ and $j = 11$. No estimate is wrong by more than about one tenth of one percent, and the last estimate is out by a factor of only 3.91×10^{-8}. Riemann's estimate is so complicated that we cannot predict exactly what will happen to it when n gets very large, but $S_d(n)$ is so simple that we can predict it will get better and better as n gets larger and larger.

This new edition of G Spencer-Brown's all-time classic comes with previously unreleased work on prime numbers as well as on the four-colour theorem.

"Spencer-Brown was Einstein's successor as a Research Fellow of Christ Church Oxford, and I think it will eventually be recognized that Spencer-Brown is the greater of the two. Einstein merely changed the laws of physics, but what Spencer-Brown changed were the laws of arithmetic."
– *J C P Miller,* Lecturer in Mathematics in the University of Cambridge

"His proof of Riemann's hypothesis is correct. I have read it five times, and there is no mistake."
– *James Flagg*

"The world's most difficult problem has been solved with the world's most elegant solution by the world's greatest arithmetician. What more is there to say?"
– *Thomas Wolf*

"I suspect I am reviewing a work of genius."
– *Stafford Beer,* Nature

"Not since the publication of Euclid's *Elements* have we seen anything like it."
– *Bertrand Russell*

This crucial recommendation was not achieved without intrigue, and required me (not unwillingly) to sleep with one of Russell's granddaughters, who asked me in the morning,

'What exactly do you want from Bertie?'

'To endorse what he said about the book when he first read it in typescript', I told her.

'He never will!' she exclaimed. 'You'll have to twist his arm, you'll have to blackmail him. How can I help?'
– *from the Preface*

Fourth printing

ISBN 978-3-89094-580-4

9 783890 945804

ISBN 978-3-89094-580-4