IRISH MATHEMATICAL SOCIETY BULLETIN

The aim of the Bulletin is to inform Society members about the activities of the Society and about items of general mathematical interest. It appears twice each year, in March and December. The Bulletin is supplied free of charge to to members by Local Representatives, or by surface mail abroad. Libraries may subscribe to the Bulletin for IR£20 per annum.

The Bulletin seeks articles of mathematical interest written in an expository style. All areas of mathematics are welcome, pure and applied, old and new. The Bulletin is typeset using TEX. Authors are invited to submit their articles in the form of TEX input files. Articles submitted in the form of typed manuscripts will be given the same consideration as articles in TEX.

Correspondence concerning the Bulletin should be addressed to:

Irish Mathematical Society Bulletin
Department of Mathematics
University College
Galway
Ireland

Correspondence concerning the Problem Page should be sent directly to the Problem Page Editor at the following address:

Faculty of Mathematics
Open University
Milton Keynes, MK7 6AA
UK

The Irish Mathematical Society acknowledges the assistance of EOLAS, The Irish Science And Technology Agency, in the production of the Bulletin.

CONTENTS

THE IRISH MATHEMATICAL SOCIETY

OFFICERS AND COMMITTEE MEMBERS

President	Prof. Seán Dineen	Department of Mathematics University College Dublin
Vice-President	Dr. Fergus Gaines	Department of Mathematics University College Dublin
Secretary	Prof. A.G. O'Farrell	Department of Mathematics Maynooth College Maynooth
Treasurer	Dr. Gerard M. Enright	Department of Mathematics Mary Immaculate College Limerick

Committee Members: M. Brennan, N. Buttimore, R. Critchley, B. Goldsmith, D. Hurley, T. Hurley, R. Ryan, R. Timoney.

LOCAL REPRESENTATIVES

Cork	RTC	Mr. D. Flannery
	UCC	Dr. M. Stynes
Dublin	DIAS	Prof. J. Lewis
	Kevin St.	Dr. B. Goldsmith
	NIHE	Dr. M. Clancy
	TCD	Dr. R. Timoney
	UCD	Dr. F. Gaines
Dundalk	RTC	Dr. E. O'Riordan
Galway	UCG	Dr. R. Ryan
Limerick	MICE	Dr. G. Enright
	NIHE	Dr. R. Critchley
	Thomond	Mr. J. Leahy
Maynooth		Prof. A. O'Farrell
Waterford	RTC	Mr. T. Power

IRISH MATHEMATICAL SOCIETY

FIRST SEPTEMBER MEETING

September 9, 1988

The Irish Mathematical Society held an Ordinary Meeting in Trinity College, Dublin, from 10 am to 5:30 pm on Friday, September 9 1988. Thirty-six members and eight non-members attended.

1. The President, Professor S. Dineen, took the chair, and opened the meeting. He noted that this was the Society's first September meeting, for the purpose of hearing scientific papers, and that it was hoped to continue it in future years. He invited proposals in time for the Christmas meeting from schools that would consider hosting the meeting in September 1989.

2. The minutes of the meeting of March 30 1988 were read, approved, and signed. There were no matters arising.

3. Dr. Stephen Gardiner, of University College, Dublin, was introduced by the President and presented an invited paper on *Integrals of subharmonic functions*.

4. At 11 am, the meeting adjourned for tea.

5. At 11:30, Dr. T.W. Korner, of Cambridge University, was introduced by A.G. O'Farrell, and presented an invited talk on *Molehills out of Mountains*.

6. At 12:30, the meeting adjourned for lunch.

7. At 14:00, the meeting divided into two groups, to hear contributed short talks.

   Group A, chaired by Dr. John McDermott, heard the following:

   Professor Bryan Cain, of Iowa State University and University College, Dublin, on *Sylvester's theorem extended.*

   Professor T. Laffey, of University College, Dublin, on *Rings with a finite* *maximal subring.*

Dr. D. Lewis, of University College, Dublin, on *The Condition of an eigenvalue.*

Group B, chaired by Dr. R.M. Timoney, heard the following:

Dr. C. Nash, of Maynooth, on *Sheaf cohomology and functional integration.*

Dr. D. Wilkins, of Trinity College, Dublin, on *Finite Gaussian curvature.*

Dr. D. O'Donovan, of Trinity College, Dublin, on *Diagonalising a real symmetric matrix.*

8. At 15:00, the meeting reassembled in the Geography Theatre, with the President in the chair, for the panel discussion on *The impact of computers on the mathematical curriculum.* Short position papers were read by M. Klimek, A. Wickstead, and T. Murphy, and there was an animated discussion.
9. At 16:00, the meeting adjourned for tea.
10. At 16:30, Professor W.K. Hayman F.R.S., of York University, was introduced by Professor B. Twomey, and presented an invited lecture on *Bases of positive continuous functions.*
11. Professor Twomey expressed the meeting's gratitude to the speakers and the organisors, and the meeting closed at 17:30.

A.G. O'Farrell,
Secretary

Support for Conferences

The Irish Mathematical Society can provide limited support for conferences held in Ireland. Application should be madein advance of the conference to the Committee through the Treasurer.

NEWS

Personal Items

- **Professor Mario Matos** from the University of Campinas, Brazil, will be visiting the Mathematics Department of UCD during the months of January and February 1989.
- **Professor Jose Isidro** from the Universidad de Santiago de Compostela will be visiting the Mathematics Department of UCD from March to June 1989.
- **John Miller** of TCD and **Eugene O'Riordan** of Dundalk RTC are invited keynote speakers at the Conference on Discretization Methods in Flow Problems to be held in Mägdesprung, East Germany, April 3–7, 1989.
- **Michael Clancy** is on sabbatical leave from NIHE Dublin for 1988/89. He is spending the year at the University of Notre Dame.
- **Robin Harte, Siobhán Vernon** and **Con O'Leary** of the Mathematics Department UCC have all availed themselves of the early retirement scheme.
- **Brendan McCann** has taken up a one-year apppointment in the Mathematics Department at UCC.
- **Don Barry** of the Statistics Department at UCC is on leave of absence at the Statistics Department, Yale University until January 1989.
- **Seán Tobin** has taken sabbatical leave from the Mathematics department of UCG for the academic year 1988/89. He is presently visiting the Mathematics Department of the University of Manitoba.
- **Paddy Quinlan** of the Mathematical Physics Department at UCC has retired. He was appointed Professor of Mathematical Physics at UCC in 1951.

- **George Kelly** has been appointed temporary acting head of the Mathematical Physics Department at UCC.

Honorary Degree Awarded to J.G. Clunie

On May 17 1988, at a special conferring ceremony in University College Cork, Professor Jim Clunie was conferred with an honorary Doctorate of Science by the National University of Ireland. He was introduced by Professor P. Barry, who gave the following address:

Professor James Gourlay Clunie is a Scottish mathematician of the first rank. He was born in St. Andrews on 26 October 1926 and attended Madras School in St. Andrews: he was DUX of the school in 1944. He entered the University of St. Andrews in 1945 having won that University's Bursary Competition: he was ranked number one in the competition for all the faculties. He graduated in 1949 with First Class Honours in Mathematics. He took his Ph.D. in 1952 at Aberdeen University under the supervision of the late A.J. Macintyre.

He was appointed to a lectureship at the University of North Staffordshire, Keele in 1952 and taught there until 1956 when he joined the Professor Walter Hayman FRS at the expanding Department of Mathematics at the Imperial College of Science and Technology, London. He taught at Imperial College until 1980 being appointed Professor of Mathematics in 1964. From 1980 to 1985, he was Senior Research Felow at the Open University, Milton Keynes, and he has been Honorary Research Associate at the University of York since 1985. In 1959-60 he visited Massachusettes Institute of Technology.

During his twenty-five years at Imperial College, the Mathematics Department there gained world-wide recognition as one of the foremost centres for complex analysis. It attracted scholars from far and near, among them several Irish mathematicians who learned their trade from the masters Hayman and Clunie.

Professor Clunie has published more than 65 research papers, a little more than half of these in collaboration with seventeen or more mathematicians, ample testimony to his openess to new ideas, wide-ranging interests and versatility. These show him to be a mathematical analyst of real power and resourcefulness. He soon became an authority on the Wiman-Valiron method. He worked generally on power series, univalent functions, entire and meromorphic functions. The quality of his output can be judged by the high opinion

held of it by co-workers in his field, and the number of references to it. In so many of his papers he has contributed an original technique, later used by others; 'by Clunie's method' is a frequently used phrase. Recently, he has proved a conjecture of Polya's on the final set of an entire function, open since 1942.

He supervised over ten research students, particularly notable among them being Milne Anderson, David Brannan, Q.I. Rahman, Terry Sheil-Small, Derek Thomas and our own Brian Twomey.

Professor Clunie took a full share of mathematical administration. The London Mathematical Society has an even wider standing than its name would suggest. He served on its council for several years, was a Vice-President in 1967-68 and shared in the editing of its journals. He co-organised major international conferences in England and co-edited the ensuing proceedings.

His association with Irish Mathematicians began in the mid-fifties when he struck up a friendship with the late Paddy Kennedy, whose memory we cherish. Since then his circle of Irish friends has expanded, to the undoubted benefit of mathematics in Ireland. He has served the NUI in several different capacities. He was the Extern Examiner in Mathematics for the periods 1974-76 and 1983-85, and acted as a substitute in 1968 and 1986. As well, he has served on appointments and promotions boards in several of the NUI colleges.

He has watched with interest the growth of the Irish Mathematical community and through his formal and informal work, greatly aided its development.

Always a friendly face at converences and rational in discussion, he has been unfailingly helpful and generous to younger colleagues. He has a gift for patient encouragement. He has been particularly helpful to the growing mathematical community in Ireland. He has had a truly distinguished record as a mathematician. We honour him for his work, and the manner of its doing.

Millenium Scholarship

Brendan Boulter of NIHE (Dublin) was recently awarded a Millennium Science and Technology Scholarship, worth £7000, to undertake research leading to a Ph.D. The scholarship scheme was established by the Minister for Science and Technology and support includes £28,000 from the private sector, £70,000 from the Irish- American Partnership and £63,000 from the Minister's Science and Technology budget. Brendan's research will be primarily concerned with the development of parallel algorithms for initial and boundary value problems, with particular emphasis on exploiting the advantages of supercomputer architecture. Brendan, a graduate of the DIT, has just completed an M.Sc. at NIHE (Dublin) under the supervision of Dr. John Carroll.

Computer Algebra at UCG

Recently UCG completed the signing of a joint contract with the EC for a project on *Intelligent Computer Algebra Systems.* This project aims to bring together research workers in the University of St Andrews, Scotland, in Technisches Hochschule Aachen, West Germany, and at University College Galway, under the EC Stimulation Action programme popularly known as "twinning".

Computer Algebra may simply, but not exclusively, be described as the symbolic manipulation of algebraic Mathematical expressions—compare word-processing or data-base management which manipulate words and characters. It has had applications in such diverse areas as coding theory, data encryption, communication network design, crystallography and solid state physics as well as within Mathematics itself.

The three Colleges have common interests in Algebra and have come together for this project. The project directors are respectively Dr Edmund Robertson at St Andrews, Professor Joachim Neubuser at Aachen and Dr Ted Hurley at Galway. This whole area will revolutionise the teaching and power of Mathematics and is making accessible to research workers in many diverse areas hitherto unworkable Mathematical algorithms.

The programs, written in the C language, will be developed on the Digital VAX machines in Galway and St Andrews and on a MASSCOMP in Aachen. Rapid communication over the electronic computer networks has been established between the Colleges so that results or software developed in one can be instantly communicated to the others.

Dr. Fred Klotz

Fred Klotz, who died in a tragic accident earlier this year, was well known to many of our members. He was a lecturer in Mathematics in St. Patrick's College, Dublin. Fred was the instigator and joint coordinator of the LOGO courses for mathematically gifted children in Ireland. He is commemorated by the Fred Klotz Memorial Trophy, which is awarded at the Irish National Logo Contest.

Irish Girl Wins International Computing Contest

The Internationl Problem Solving Contest (ICPSC) is now in its eighth year. In 1987 an elementary LOGO division was introduced and the winner was John Farragher from Limerick. This year, the winner was Anne Chazarreta, a twelve year old sixth class pupil from Scoil an Spioraid Naoimh (Girls), Bishopstown, Cork.

For the last three years, experimental courses using the computer language LOGO for mathematically able children have been conducted at various centres throughout Ireland. The Cork centre is at Coláiste an Spioraid Naoimh, Bishopstown and is directed by Michael Moynihan and Declan Donovan.

IMS MEMBERSHIP

Ordinary Membership of the IMS is open to all persons interested in the activities of the Society. Application forms are available from the Treasurer and from Local Representatives. Special reciprocity rates apply to members of the IMTA and of the AMS.

Institutional Membership is a valuable support to the IMS. Institutional members receive two copies of each issue of the Bulletin and may nominate up to five students for free membership.

Subscriptions rates are listed below. The membership year runs from 1st October to 30th September. Members should make payments by the end of January either direct to the Treasurer or through Local Representatives. Members whose subscriptions are more than eighteen months in arrears are deemed to have resigned from the Society.

Ordinary Members	IR£5
IMS–IMTA Combined	IR£6.50
Reciprocity Members from IMTA	IR £1.50
Reciprocity Members from AMS	US$6
Institutional Members	IR£35

Note: Equivalent amounts in foreign currency will also be accepted.

ARTICLES

Edge Sums of Hypercubes

Niall Graham   Frank Harary

Introduction

A *hypercube* may be defined recursively in terms of the cartesian product of graphs as defined in [1, p.23] and in [2]:

$$Q_n = \begin{cases} K_2 & n = 1 \\ Q_{n-1} \times Q_1 & n \geq 2 \end{cases} \tag{1}$$

The hypercube Q_n of *dimension* n may equivalently be defined as the graph of 2^n nodes such that each node is uniquely labelled with a number expressed as an n-digit binary string and two nodes are adjacent whenever their labels vary in exactly one binary digit. The dimension of the edge uv, where $u = (u_1, \ldots, u_n)$, $v = (v_1, \ldots, v_n)$ is k if $u_k + v_k = 1$ and $u_i = v_i$ for $i \neq k$.

Given such a labeling of Q_n, we form the *network* $N(Q_n)$ by assigning integer weights to the edges of Q_n as follows. For two adjacent nodes i and j, let w_{ij} be the weight on edge $ij \in E(Q_n)$, defined by

$$w_{ij} = i + j \tag{2}$$

In words, each edge is assigned as its weight the sum of the two nodes with which it is incident. In Figure 1, the networks formed by labelling Q_1, Q_2 and

Q_3 in this way are shown.

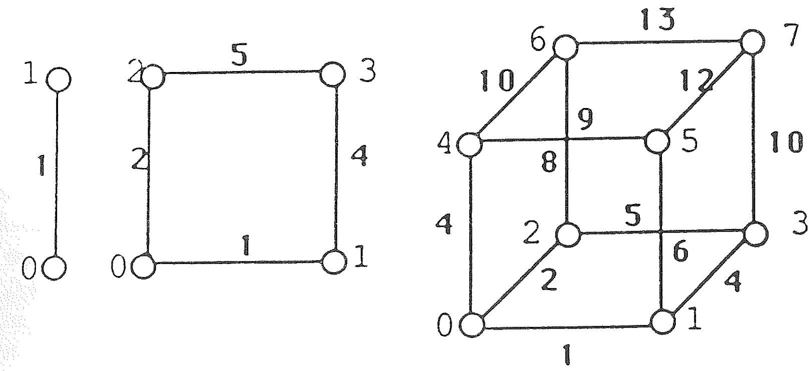


Figure 1.
The labelled hypercubes of dimensions 1,2 and 3 with edge weights.

With each hypercube Q_n we may now associate a set L_n and a multiset M_n of the weights on the edges of $N(Q_n)$. Thus $L_2 = M_2 = \{1, 2, 4, 5\}$ and $L_3 = \{1, 2, 4, 5, 6, 8, 9, 10, 12, 13\}$ while $M_3 = \{1, 2, 4, 4, 5, 6, 8, 9, 10, 10, 12, 13\}$.

Note that the numbers from 1 to 13 which are not in the set L_3 are 3, 7 and 11. When expressed in binary form, these numbers are 11, 111 and 1011. The fact that they all end in 11 is no coincidence, as we shall see.

Our purpose is to characterize the set

$$L_\infty = \bigcup_{n=1}^{\infty} L_n \tag{3}$$

that is, the set of all integers that are the weights of an edge in some $N(Q_n)$. We will also determine the limiting multiset

$$M_\infty = \lim_{n \to \infty} M_n \tag{4}$$

Of course L_∞ can also be written as the limit of the sets L_n.

Edge Sums

To characterize those numbers which are edge sums for some hypercube, we require the operation of the sum of two multisets of numbers. For multisets $S_1, S_2 \subset \mathrm{Z}^+$, define $S_1 + S_2 = \{w_1 + w_2 : w_i \in S_i\}$. When S_1 is a singleton w_1 we abuse the notation by writing $S_1 + S_2 = w_1 + S_2$.

To derive a recurrence relation for the sets M_n, we appeal to the recursive definition (1) of Q_n. Note that Q_n may be constructed from two copies of Q_{n-1}. One copy of Q_{n-1} has the usual integer values on its nodes, and thus has the multiset M_{n-1} of weights on its edges. The other copy is labelled similarly with node labels all greater by 2^{n-1}. Thus its edge weights are all greater by $2^{n-1} + 2^{n-1} = 2^n$. A pair of nodes, one from each copy of Q_{n-1}, are adjacent whenever their labels differ by 2^{n-1}, as the edges joining them lie in the nth dimension of Q_n. Consequently, M_n may be written as the union of M_{n-1} with the elements of $2^{n-1} + M_{n-1}$, along with the weights on the edges joining the two copies of Q_{n-1}. Thus, since the base case is $M_1 = \{1\}$ as seen in Figure 1, we find that

$$M_n = M_{n-1} \bigcup (2^n + M_{n-1}) \bigcup \{2^{n-1} + 2k : 0 \le k < 2^{n-2}\} \tag{5}$$

As L_n is the set of the multiset M_n, it follows that

$$L_n = M_{n-1} \bigcup (2^n + L_{n-1}) \bigcup \{2^{n-1} + 2k : 0 \le k < 2^{n-2}\} \tag{6}$$

for all $n \ge 1$. Let Z^+ be the set of all positive integers and write $\mathrm{Z}_0 = \mathrm{Z}^+ \cup \{0\}$. Using this notation we are now able to state our main result.

Theorem 1 *A positive integer z is the weight of some edge in $N(Q_n)$ for sufficiently large n if and only if $z \not\equiv 3 \pmod 4$. Thus $z \in \mathrm{Z}^+$ is not a hypercube weight just if $z = 4x+3$ for some $x \in \mathrm{Z}_0$, i.e., the binary expression of z ends in 11.*

Proof Let $\alpha 00 \beta$ and $\alpha 10 \beta$ be adjacent nodes in Q_n, where α and β are binary strings. The weight on the edge joining these nodes is then $\alpha 01 \beta 0$, found by base 2 addition. As α and β may both be null strings we see that the resultant string $\alpha 01 \beta 0$ can be any even number. An odd sum may only arise if the summands vary in their least significant digit, that is, between nodes with labels $\alpha 0$ and $\alpha 1$ for some binary string α. Obviously, their sum is $\alpha 01$ and odd numbers of the form $\alpha 11$ cannot be formed in this way. Thus the only admissible odd numbers are those of the form $4x+1$. Combining this with the fact that all even numbers are admissible, the result is established.

As well as characterizing all the integers in L_∞, we may also find the multiplicity of each integer in M_∞.

Corollary 1a *The multiplicity $f(y)$ of the integer y in M_∞ is given by*

$$f(y) = \begin{cases} \lfloor \log_2 y \rfloor & y = 2x \\ 1 & y = 4x+1 \\ 0 & y = 4x+3 \end{cases} \tag{7}$$

Proof We have already shown that $f(4x+3) = 0$. We now consider the other two cases. Without loss of generality let i and j be adjacent nodes in Q_n, with $i < j$. Clearly, $j = i + 2^k$ for some k with $0 \le k \le n$. Then the integer y found by adding i to j is

$$y = 2i + 2^k \tag{8}$$

Obviously, when y is odd we require that $k = 0$ and the solution, if any, is unique. We have already seen that when $y = 4x + 1$ for some integer $x \ge 0$ such a solution does exist. For even $y = 2m$ we see that $y = i + 2^{k-1}$, which has as many solutions as there are values of 2^{k-1} which are less than m, that is, $2^k < 2m = y$. All such solutions are admissible and thus the multiplicity of $y = 2m$ is given by the number solutions of the diophantine equation (8). The number of such solutions is obviously given by $\lfloor \log_2 y \rfloor$.

Using (6) and (7), we find L_4 and M_4:

$$\begin{aligned} L_4 &= \{1,2,4,5,6,8,9,10,12,13,14,16,17,18,20,21, \\ &\quad 22,24,25,26,28,29\} \\ M_4 &= \{1,2,4,4,5,6,8,8,9,10,10,10,12,12,13,14,16, \\ &\quad 17,18,18,20,20,20,21,22,22,24,25,26,26,28,29\} \end{aligned} \tag{9}$$

Note that (for $n \le 4$) the sets L_n and multisets M_n are symmetric about the value $2^n - 1$. We now show that this is true in general.

Corollary 1b *A number z is the weight of some edge of Q_n if and only if its reflection about $2^n - 1$, viz., $z + 2(2^n - 1 - z) = 2^{n+1} - 2 - z$, is an edge weight of Q_n.*

Proof Let i and j be adjacent nodes in Q_n expressed as binary strings and let i' and j' be the bitwise complements of i and j so that

$$i + i' = j + j' = 2^n - 1$$

As i and j are adjacent, the complementary nodes i' and j' are adjacent. Thus, for every edge whose weight is z there exists another edge whose weight is $2(2^n - 1) - z$.

References

[1] F. Harary, *Graph Theory*. Addison-Wesley, Reading, 1969.

[2] F. Harary, J.P. Hayes and J-H. Wu, *A survey of the theory of hypercube graphs*, Comput. Math. Appl., to appear.

Department of Computer Science
New Mexico State University
Box 3CU Las Cruces
New Mexico 88003-0099
USA

Algebraic Techniques Of System Specification

Mike Holcombe

1 Introduction

The design of complex software systems is a relatively new occupation and is still in its infancy. With the rapid growth in the applications of microprocessor technology more and more areas of life are being affected and in some of this activity there is serious cause for concern. Many manufacturers are using microcomputers to control safety-critical systems. Such systems are usually defined to be systems, the malfunctioning of which could lead directly to injury or death on a small (local) or large (global) scale. Examples of recent products and systems that have caused injury or death through inadequate software design are :

- Chemical processing plants,
- Washing machines,
- Car cruise controls,
- Intensive care systems,
- Industrial robots, etc.

In many of these systems there is a serious problem in the formal specification of the total system and its environmental interaction. Most interactive systems involve three important components :

- the system,
- the environment,
- the user.

Each component interacts and 'communicates' with all the others and it is therefore crucially important that we take this into account at all stages of the design process, from specification, design, validation, maintenance and disaster analysis.

Much activity currently centres around the development of rigorous techniques, often based on formal logics, for the verification of systems. For this to be practical it is essential that the system specification is based on a complete model of the environment and the user's behaviour. Any verification of the system can only be valid subject to a correct model of these two important aspects of the total situation.

The modelling of complex environments has been an important research activity for many years, involving, perhaps, thermodynamics, hydrodynamics, electromagnetics, materials theory etc. and many sophisticated models have been produced. However even these models are far from being complete but they are all that we have available. In the design of a safety-critical control system for some industrial process, a chemical plant or a nuclear power station, it is only possible to validate the software subject to the model of the environment being realistic.

The Bide Report has examined some of the problems facing the Information Technology industry following the UK Government's Alvey Initiative and has stressed how important the user interface is in any computer system. The report makes the point that, no matter how reliable and well designed a system is, if the user interface is not well designed and sympathetic to the user's needs then the success of the system as a whole is in serious doubt. This is especially true in the case of interactive, safety-critical systems. When we turn to the problem of modelling the actions of a user, which could be fundamental to the safety of the system, we have a serious problem. Although much experimental evidence has been amassed about user behaviour much of it is contradictory and there is no body of formal theory which could act as a basis for reasoning about such important matters. Several attempts are being made to develop rigorous design methodologies to take account of these problems. These methodologies require the development, as does any method which is trying to design the user interface, of a sensible series of models of user behaviour and belief. The construction of formal user conceptual models is an area of importance and these models must be based on some sort of foundational logic that is rich enough for the expression of possibly irrational and ill-defined beliefs about the system. The recent work on 'belief' logics looks very promising in this respect [4].

One basic problem with current specification and analysis methodologies is that they tend to be rather specialised and cannot always deal with different aspects of a system and its environment. We have developed a method, based on the theory of X-machines (see [8]), which enables the formal description and analysis of most aspects of a system and its environment in a unified way. Systems may involve concurrent or real-time processing and yet the X-machine model is sufficiently robust that it can be used to specify such systems. Analogue aspects of a real-time control system can be described using X-machines with a topological basis. At the heart of such machines are suitable models of data types and operations which can be expressed either in model-theoretic form, such as is used in VDM or Z, or in functional or algebraic paradigms (some elementary ideas from these approaches are discussed in §2).

2 The specification of data types

One very promising approach to the design of more reliable software systems is the formal specification of data types and operations. There are several approaches; the two most popular are algebraic specification and model-based specification.

The algebraic approach to data type specification involves the definition of data types in terms of universal algebras. Let us suppose, as before, that our system involves a collection of sets and functions or operators. There may be some sets that are constructed of products of other sets and so on. We specify first a collection of basic sets and operators. In our examples we will consider the specification of data types needed in the design of a simple word processor, since that is a system that many people may now be familiar with. The most important set is the set of finite sequences of letters and numerals which we call $seq[Char]$. Although this is a basic set it does involve some interesting mathematical problems. We will construct a specification of this data type from the more primitive type $Char$.

Let $Char$ denote the set of all possible symbols to be used for the construction of documents, so $Char = \{A, a, B, b, \ldots, Z, z, 0, 1, \ldots, 9, \Box\}$, where $\Box$ represents a blank space.

The main operations that we wish to carry out with sequences are

1. construct sequences,

2. combine sequences,

3. test to see if a sequence is the null sequence,

4. extract the leftmost symbol of a string,

5. delete the leftmost symbol from a string.

These operations will be defined using functions. We first identify the sets *Char*, *seq*[*Char*] and *Bool*, the 2-valued truth set. There are then some function declarations:

$$\begin{aligned} null &: \ \to seq[Char] \\ | &: \ Char \times seq[Char] \to seq[Char] \\ * &: \ seq[Char] \times seq[Char] \to seq[Char] \\ isnull &: \ seq[Char] \to Bool \\ head &: \ seq[Char] \to Char \\ tail &: \ seq[Char] \to Char \end{aligned}$$

Here we are postulating that a null sequence, denoted by $\hat{}$, exists and this is defined by the first function declaration. The next thing we can do is to generate sequences of length 1 using the second function and perhaps write $a|\hat{}$ as $< a >$ etc. Then $b| < a >$ would represent $< ba >$ and so on. Further applications of the functions described above could be

$$head(c|(a|\hat{})) = c, \qquad tail(c|(a|(b|\hat{}))) = a|(b|\hat{}))$$

and so on.

However we have not given a precise semantics for these functions and this is done using equations like the following :

$$\begin{aligned} head(x)|tail(x) &= x \\ x * (y * z) &= (x * y) * z \\ x * \hat{} &= \hat{} * x = x \quad \text{etc.} \\ isnull(\hat{}) &= T \end{aligned}$$

where $x, y, z \in seq[Char]$. The precise choice of the equations to describe the semantics of the data type is not uniqely determined as long as the algebraic model that these equations represent is consistent with the original system requirement.

To take a more abstract view we can postulate the existence of a set of 'sorts' that will, eventually, be replaced by explicit sets like *Char* and *seq*[*Char*]. Let us call these sorts $s_1, s_2, \ldots, s_n$. Then we define various operators $w_1 : \ \to s_2$, $w_2 : s_1 \times s_2 \to s_2$, $w_3 : s_2 \times s_2 \to s_2$, $w_4 : s_2 \to s_3$ etc. to represent null, |, $*$, *isnull* and so on.

We can now define the abstract concept of a ***data algebra***. A ***signature*** is a pair $\Sigma = (S, \Omega)$ where S is a set of sorts, and $\Omega = \{\Omega_{x,s}\}$ is a set of operators indexed by pairs of the form (x, s) where $x \in S^*$, $s \in S$,(S^* is the free semigroup generated by S, x is called the 'arity' of the operators in $\Omega_{x,s}$). Thus $w_2 \in \Omega_{s_1 s_2, s_2}$ etc. A Σ-***algebra*** is a pair $A = (S_A, \Omega_A)$ containing a family S_A of carrier sets s_A for each sort $s \in S$, and a family of operations $w_A : s_{1,A} \times \ldots \times s_{n,a} \to s_A$ for each operator $w \in \Omega_{s_1 \ldots s_n, s}$.

Now a *specification* consists of a pair $D = (\Sigma, E)$ where Σ is a signature and E is a set of Σ-equations. A D-algebra is any Σ-algebra which satisfies the set of equations E. A central result of the theory is that there exists an initial D-algebra (in the categorical sense). This initial D-algebra can then serve as a model for the specification of the system.

Such algebraic specifications can be 'implemented' using a language such as OBJ which is available for many mainframe computers. The standard reference for this work is now [9].

In model-based specifications the data types are defined in terms of sets and functions or operations defined on these sets with a semantics prescribed by a collection of predicate sentences or an explicit (possibly recursive) construction.

Example. Consider the possible fundamental data type associated with a simple word processor. We form the set, *seq*[*Char*], of all finite sequences or words from *Char*, including the empty word $\hat{}$, constructed above. Usually we will write a sequence in the form $< abcdefg >$.

The set *DOC* is defined to be the product

$$seq[Char] \times seq[Char]$$

and this represents the state of a simple document with a document of the form (α, β) representing the situation

$$\alpha \blacksquare \beta$$

that is, the string of symbols corresponding to α and the string corresponding to β with the cursor over the first symbol of β. It is possible to introduce a more realistic representation of a document broken up into lines, paragraphs pages, windows etc. at a later stage.

We will use the notation Z, see [5], and declare each function with its semantics given below it.

$move : DOC \nrightarrow DOC$ $delete : DOC \nrightarrow DOC$ $insert : DOC \nrightarrow Char \nrightarrow DOC$ $print : DOC \rightarrow seq[Char]$
$\mathrm{dom}\, move = \mathrm{dom}\, delete = \{\, l, r \mid l \neq \hat{}\,\}$ $(\forall (l, r) : DOC; a : Char)$ $move(l * < a >, r) = (l, < a > * r)$; $delete(l * < a >, r) = (l, r)$; $insert(l, r)\, a = (l * < a >, r)$; $print(l, r) = l * r$:

Notes. (1) The notation $f: A \nrightarrow B$ means that the function is partial and not necessarily completely defined.
(2) The notation ":" is often used in place of $\in$ and $A \rightarrow B$ means the set of all functions from A to B.
(3) We use $f\ a$ to represent $f(a)$.
(4) dom means domain.

Using these definitions we can describe more complex data types and functions and consequently build up a more detailed and realistic specification of the data types and operations associated with the system. For example we need to be able to move in the opposite direction to the way the function *move* works. This can be done by constructing a simple function that 'reverses' a string of characters and then apply this in composition with the existing *move* function (before and after) suitably adapted for the type *Doc*. We can then define higher level functions which include direction parameters '*right*' and '*left*'.

The approach taken in [5] takes this view. A less constructive approach which just describes the properties that a function must satisfy without actually describing how this function can be constructed is also used in practice. The book [6] describes some simple examples of this approach. VDM is another, similar, approach with a more structured implementation environment which is discussed in [7].

3 Dynamic system specification

Although data type specification is of great importance there are several aspects of a system that are better specified by a more 'dynamic' model. The use of various types of machine is becoming more widely used for the formal specification of systems.

We discuss the concept of an X-machine, which is a general model of computation with the intention of using this model in the specification of computer systems.

The main mathematical model of computation is the Turing Machine. Athough this has received much study in a variety of theoretical areas it is not used by software engineers for the specification of systems, the principle reason being that the model is based at a very low level of abstraction and is not very amenable to analysis and system development. Less general models, such as finite state machines, machines with stacks and/or registers and Petri nets, however, are the basis of many system specification and development methodologies.

The use of graphical elements in a specification methodology is attractive from the point of view of user understanding, conveying dynamic information, and system refinement. Since the Turing Machine model is impractical and the finite state machine model is too restrictive, it would seem that the graphical advantages possessed by these models are not going to be available for general system specification. However, there is a much more appropriate model of computation that can, when combined with suitable data type methods, provide us with an appropriate environment for the description and analysis of arbitrary systems. Since this model also has very promising capabilities for use in discussing concurrent systems, it seems worthy of further investigation.

We start with the definition of the *X-machine* and show how this definition relates to previously studied concepts such as Turing Machines, push down machines and finite state machines. Then we examine some elementary aspects of the theory of X-machines and conclude with a few examples. It should be remarked that although these machines were introduced in 1974 [9] they have not received much attention.

Let X be any non-empty set, henceforth referred to as the fundamental data type, and Φ a finite set of relations defined on X. Thus Φ consists of relations of the form $\phi: X \to X$. If one prefers we can regard each ϕ as a function, which is possibly incompletely specified, from the set X into the set $\mathcal{P}(X)$, the set of all subsets of X (also known as the power set of X).

Intuitively X represents the set of data to be processed and ϕ are the set of functions or relations that carry out the processing. In some cases the data type X can represent internal architectural details, such as contents of registers etc. and it is in this way that the model can assume its full generality.

Clearly we need to specify some relationship between the input and output information of the overall system and the data type X, especially when X contains information that is not directly involved with the system input and output. This is done by specifying two sets, Y and Z, to represent the input and output information respectively. In many cases, as in much processing, these sets are free semigroups or subsets of free semigroups (i.e., languages over some finite alphabet).

Two coding relations, $\alpha: Y \to X$ and $\beta: X \to Z$ describe how the input is coded up prior to processing by the machine, and how the subsequently processed data is then prepared (or decoded) into a suitable output format. Some examples will demonstrate how this works in a few basic cases.

Finally we need to describe some suitable control structure that will actually determine how the processing is performed. This structure is very similar to the state transition graph of a finite state machine and will appear familiar. However, this appearance masks a model of considerable computational power since much of the similarity with finite state machines is concerned with the *control* of the processing and not with the *type* of processing that the machine performs. Nevertheless, the similarities with finite state machines are extremely useful since they allow us, at times, to apply techniques for the analysis of machines that have proved to be tremendously successful.

The final ingredient is the *state space* of the machine, which consists of a finite set, Q, of states and a function

$$F: Q \times \Phi \to \mathcal{P}(Q)$$

called the *state transition function.*

For many purposes this state space can be described using a graph which has the elements of Q at the nodes (vertices) and for each $q, q_1 \in Q$, $\phi \in \Phi$ there is a labelled arc

$$q \xrightarrow{\phi} q_1$$

precisely if $q_1 \in F(q, \phi)$.

It is also necessary to identify a subset $I \subseteq Q$ of initial states and a subset $T \subseteq Q$ of terminal states. An initial state will be indicated in the state space by being the target of an unlabelled and sourceless arrow, e.g.,

$$\to q$$

whereas a final state will be described by being the source of an unlabelled and targetless arrow, thus:

$$q \to$$

An example of a state space is given in Fig 1.

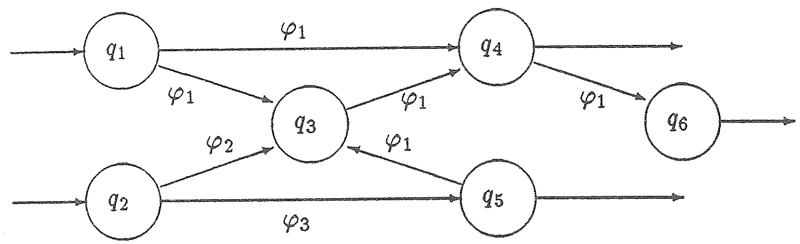


Fig.1 The state space of an X-machine.

In Fig. 1 states q_1 and q_2 are initial states and states q_4, q_5 and q_6 are terminal states. This example is of a ***non-deterministic*** machine; witness the two arrows leaving q_1 labelled with ϕ_1. It is also incomplete in the sense that no arrow labelled with ϕ_1 leaves state q_2.

The formal definition of an X-machine is presented in the following definition.

Definition. An X-machine is a 10-tuple:

$$M = (X, \Phi, Q, F, Y, Z, \alpha, \beta, I, T);$$

where

X, Y, Z are non-empty sets;
Φ is a set of relations on X;
Q is a *finite* non-empty set;
$F: Q \times \Phi \to \mathcal{P}(Q)$ is a, possibly partial, function;
$\alpha: Y \to X$ and $\beta: X \to Z$ are relations;
$I \subseteq Q$ and $T \subseteq Q$ are subsets.

Remark. The relations appearing in the definition are often functions or partial functions in many examples. The definition is presented here for the record in its most general setting. The set $\mathcal{P}(Q)$ denotes the power set (or set of subsets) of Q.

We call Y the *input type* and α the *input relation*. The set Z is the *output type* and β is the output relation.

The process of computation that this machine performs can be described by choosing an element $y \in Y$ from the input type and studying how this element is processed.

First the input relation is applied to the element y to produce an element or set of elements $\alpha(y)$ of X.

Next a path in the state space of the machine is selected that starts from a state in I and ends in a state from T. There may, in a non-deterministic or incomplete machine, be many or none. If a path is selected it will determine a sequence from Φ^* using the labels of the arcs of the path in order. If the labels of the arcs are ϕ_1, $\phi_2, \ldots, \phi_n$ then the word

$$\phi_1 \circ \phi_2 \circ \ldots \phi_n$$

defines a composite relation (or function) on the set (or type) X. (In this notation we apply the relation ϕ_1 then ϕ_2 and so on, which is a common practice in algebra but may seem unusual elsewhere!)

When this composite relation is applied to $\alpha(y)$ we obtain an element or subset of X and this yields an element or subset of the output type Z on applying β.

The result of the computation is thus

$$\beta((\phi_1 \circ \phi_2 \circ \ldots \phi_n)(\alpha(y)))$$

If at any stage we find that the result of a partial computation

$$(\phi_1 \circ \phi_2 \circ \ldots \phi_k)(\alpha(y))$$

is the empty set for some $k \leq n$ then we will regard that computation as halting and the output, if any, is obtained by applying β as before. Fig. 2 gives a diagrammatical interpretation of the process.

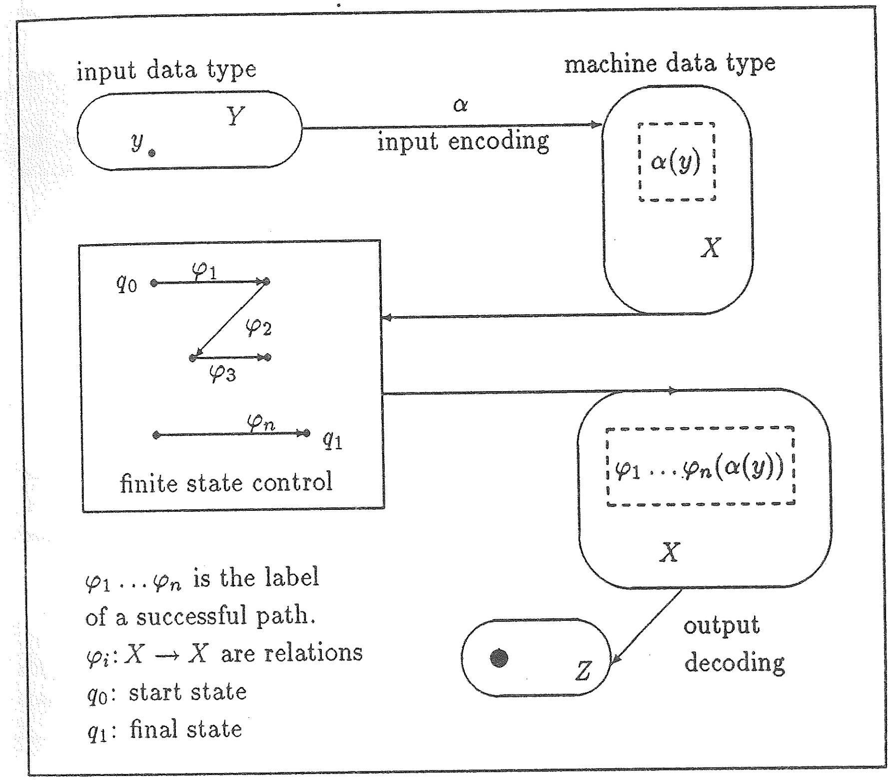


Fig. 2 An X-machine computation.

4 Some examples of X-machines

The most general model of computation so far investigated in any detail is the Turing machine model and its equivalent theories. There is, however, a newcomer to the scene that is claimed to be more general, namely the Quantum computer of Deutsch [10]. We do not intend to enter the controversy surrounding this new model and its relevance to computer science at this stage, merely note its existence. We will, however, demonstrate that the Turing machine is just a special case of the X-machine defined above.

Before we examine the connections between X-machines and other machines we need to introduce some terminology.

Let Σ be any non-empty set. Some relations will now be defined on the set Σ^* of all finite sequences or words in Σ. For any $\sigma \in \Sigma$ we define some fundamental relations:

$$
\begin{array}{ll}
L_\sigma : \Sigma^* \to \Sigma^* & (\forall x \in \Sigma^*)\ xL\sigma = \sigma x \\
L_\sigma^{-1} : \Sigma^* \to \Sigma^* & (\forall x \in \Sigma^*)\ xL_\sigma^{-1} = \{ y \in \Sigma^* \mid \sigma y = x \} \\
R_\sigma : \Sigma^* \to \Sigma^* & (\forall x \in \Sigma^*)\ xR_\sigma = x\sigma \\
R_\sigma^{-1} : \Sigma^* \to \Sigma^* & (\forall x \in \Sigma^*)\ xR_\sigma^{-1} = \{ y \in \Sigma^* \mid y\sigma = x \} \\
left : \Sigma^* \times \Sigma^* \to \Sigma^* \times \Sigma^* & \\
(a,b)left = (reverse(tail(reverse(a))), head(reverse(a)) * b) &
\end{array}
$$

(The purpose of the last string processing function will become clearer when we consider a later example, essentially it transfers the last symbol of the first word to the front of the second word. The standard functions reverse, head and tail are assumed to be defined already as is concatenation, $*$.)

The Turing machine model. The essential features of a Turing machine consist of an alphabet d, a finite set of states Q and a finite set of n-tuples ($n = 4$ or 5) which describe the behaviour of the machine under various circumstances. The set of 5-tuples that we will use here will be elements of the form

$$(q, q_1, \theta, \theta_1, d)$$

where $q, q_1 \in Q$; $\theta, \theta_1 \in \Sigma \cup \hat{}$ where $\hat{}$ denotes a blank; and either $d = L$ or $d = R$. The interpretation of such a tuple is that if the machine is in state q and the current symbol being scanned is θ then the next state is q_1, the symbol θ_1 is printed on the tape instead of θ and the read-write head is moved 'left' if $d = L$ and 'right' if $d = R$. Further details and examples of Turing machines will be found in many texts on the theory of computer science.

Added to this is a start state q_0 and a set $T \subseteq Q$ of terminal states. The initial tape contains a string of characters from the set Σ^* which is input to the machine in the state q_0. Processing consists of applying a sequence of appropriate tuples so that if at any stage the machine is in state q and is reading the tape symbol θ then any tuple of the form

$$(q, q', \theta, \theta', d)$$

where $q' \in Q$, $\theta' \in \Sigma \cup \{\hat{}\}$, $d \in \{L, R\}$ can be applied to yield the next state q', the symbol θ replaced by the symbol θ' and the tape head moved either left or right.

If the tape head moves left then the processing takes a tape of the form

$$[\sigma_1\sigma_2 \ldots \sigma_k, \sigma_{k+1} \ldots \sigma_n]$$

with the head reading the symbol σ_k and either produces a resultant tape of the form

$$[\sigma_1\sigma_2 \ldots \sigma_{k-1}, \sigma_k'\sigma_{k+1} \ldots \sigma_n]$$

where σ_k' is the new symbol printed on the tape after applying the tuple or

$$[\sigma_1\sigma_2 \ldots \sigma_{k-1}, \sigma_{k+1} \ldots \sigma_n]$$

For a right move the resultant tape is of the form

$$[\sigma_1\sigma_2 \ldots \sigma_k'\sigma_{k+1}, \sigma_{k+2} \ldots \sigma_n]$$

or

$$[\sigma_1\sigma_2 \ldots \sigma_{k+1}, \sigma_{k+2} \ldots \sigma_n]$$

In some cases the tuple may involve the replacing of a symbol on the tape by a blank.

In the context of an X-machine we first define the set X as

$$X = \Sigma^* \times \Sigma^*$$

The set of states is Q and the initial and terminal states as in the Turing machine case. For each tuple of the form

$$(q, \sigma, q', \sigma', L)$$

we insert an arrow from q to q' labelled by the relation

$$R_\sigma^{-1} \times L_{\sigma'}$$

on X. For each tuple of the form

$$(q, \sigma, q', \sigma', R)$$

we insert an arrow from q to q' labelled by the relation

$$\phi = (R_\sigma^{-1} \times 1) \circ (R_{\sigma'}^{-1} \times 1) \circ left$$

etc. The definition of the input and output relations for the X-machine are given next.

$$\alpha, \beta : \Sigma^* \to \Sigma^* \times \Sigma^*$$
$$(a)\alpha = (\hat{}\,, a)$$
$$(a, b)\beta = a$$

This interpretation is of a Turing machine that behaves as a function on Σ^*. If the machine halts during a computation this means that there is no arrow leaving the current state which has, as a label, an applicable relation. The result is then obtained by use of the decoding relation.

Finite state machines. The classical model of a finite state machine can be represented as an X-machine in the following way.

Let Q be a finite state set, Σ a finite input set and Ω a finite output set; then a finite state machine is a quintuple

$$A = (Q, \Sigma, \Omega, F, G)$$

where $F: Q \times \Sigma \not\to Q$ and $G: Q \times \Sigma \not\to \Omega$ are partial functions defining the next state and output functions.

The X-machine is defined as follows, The set $X = \Omega^* \times \Sigma^*$, the set of states is Q and the sets of final and initial states are also equal to Q. The set of relations Φ are defined as follows. If $q, q' \in Q$, $\sigma \in \Sigma$, $\theta \in \Omega$ are such that $F(q, \sigma) = q'$ and $G(q, \sigma) = \theta$ then we insert an arrow from state q to state q' labelled by the relation

$$\phi = R_\theta \times L_\sigma^{-1}$$

The input and output codes are given by

$$\alpha: \Sigma^* \to X \qquad \text{where } \alpha(a) = (\hat{}\,, a)$$

$\hat{}$ being the empty string, and

$$\beta: X \to \Omega^* \qquad \text{where } \beta(a, b) = a\,.$$

If it is necessary to only carry out computations starting from a given initial state we will define I to be the singleton set containing this state.

The X-machine computes exactly the same sequential function as does the original finite state machine.

In the previous section we gave the general definition of an X-machine and illustrated this with some examples to show that the concept is fully general. In this section we will briefly review some of the theory of X-machines, although at this time this theory is not as well developed as it might be. The definition of the *behaviour* of an X-machine can be made in terms of the function or relation that it computes or in terms of the language it recognizes.

Let $M = (X, \Phi, Q, F, Y, Z, \alpha, \beta, I, T)$ be any X-machine. If

$$c: q_0 \xrightarrow{\phi_1} q_2 \xrightarrow{\phi_2} q_2 \to \ldots \xrightarrow{\phi_n} q_n$$

represents a sucessful path in the state space of M, so that $q_0 \in I$ and $q_n \in T$, then the relation

$$|c| = \phi_1 \circ \phi_2 \circ \ldots \circ \phi_n : X \to X$$

will be called the *relation defined* by that *labelled path*. The *behaviour* of M is then

$$|M| = \bigcup |c| : X \to X$$

where the union is taken over all the successful paths in the state space.

The *relation computed* by the machine is then defined as

$$f_M = \alpha \circ |M| \circ \beta : Y \to Z$$

For the recognition of languages we define the output set to be $\hat{}$ and the output function $\beta: X \to Z$ yields a subset

$$A = \hat{}\, f_M$$

of Y.

The article [1] discusses some of the applications of this material. We can develop a methodology for the description of systems by a combination of the data type methods of the first sections with the machine based methods of the latter ones. In situations when architectural features of the system are important, these can be incorporated into the X-machine by defining the set X suitably, perhaps including models of registers etc.

References

[1] M. Holcombe, *X-machines as a basis for dynamic system specification*, Software Engineering Journal (in press).

[2] M. Holcombe, *Formal methods in the specification of the human-machine interface*, Int. CIS Journal 1(1) (1987), 24–34.

[3] M. Holcombe, *Goal-directed task analysis and formal interface specifications*, Int. CIS Journal 1(4) (1987), 14–22.

[4] R. Fagin and J.Y. Halpern, *Belief, awareness and limited reasoning*, Art. Intel. 34 (1988), 39–76.

[5] B. Suffrin, *Formal specification of a display oriented text editor*, Science of Comp. Prog. 1 (1982), 157–202.

[6] I. Hayes, *Specification case studies*, Prentice-Hall, 1986.

[7] C.B. Jones, *Systematic software development using VDM*, Prentice-Hall, 1986.

[8] S. Eilenberg, *Automata, languages and machines*, Academic Press, 1974.

[9] H. Ehrig and B. Mahr, *Fundamentals of Algebraic specification*, vol.1 EATCS Monographs, Springer 1985.

[10] D.A. Deutsch, *The quantum computer and the Church-Turing thesis* Proc. Royal. Soc. A 400 (1985), 97–117.

Department of Computer Science
University of Sheffield

Crossed Modules

Graham Ellis

Crossed modules were invented almost 40 years ago by J.H.C.Whitehead in his work on combinatorial homotopy theory [W]. They have since found important roles in many areas of mathematics (including homotopy theory, homology and cohomology of groups, algebraic K-theory, cyclic homology, combinatorial group theory, and differential geometry). Possibly crossed modules should now be considered one of the fundamental algebraic structures. In this article we give an account of some of the main occurrences and uses of crossed modules and we describe some recent developments in their theory.

Before presenting the definition of a crossed module, we shall consider several motivating examples. Throughout G denotes an arbitrary group.

Example 1 Let N be a normal subgroup of G. The inclusion homomorphism $N \to G$ together with the action ${}^g n = gng^{-1}$ of G on N is a crossed module.

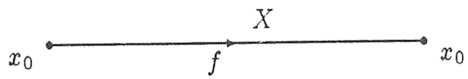
Example 2 If M is a $\mathbb{Z}G$-module then the trivial homomorphism $M \to G$ which maps everything to the identity is a crossed module.

Example 3 Let $\partial\colon H \to G$ be a surjective group homomorphism whose kernel lies in the centre of H. There is an action ${}^g h = \tilde{g}h\tilde{g}^{-1}$ of G on H where $\tilde{g}$ denotes any element in $\partial^{-1}(g)$. The homomorphism ∂ together with this action is a crossed module.

Example 4 Suppose that G is the group $Aut(K)$ of automorphisms of some group K. Then the homomorphism $K \to G$ which sends an element $x \in K$ to the inner automorphism $K \to K$, $k \mapsto xkx^{-1}$ is a crossed module.

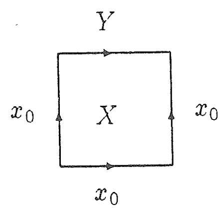
Each of these examples consists of a group homomorphism with an action of the target group on the source group. Before stating the precise algebraic properties needed by such a homomorphism for it to be a crossed module, let us consider some more substantial examples.

Example 5 Let X be a topological space in which a point x_0 has been chosen. Recall that the fundamental group $\pi_1(X, x_0)$ consists of homotopy classes of continuous maps $f : [0,1] \to X$ with $f(0) = f(1) = x_0$. (Two such maps are *homotopic* if one can be continuously deformed into the other in such a way that the image of 0 and 1 remains x_0 throughout the deformation.) We think of these maps as paths in X beginning and ending at x_0; the appropriate picture is

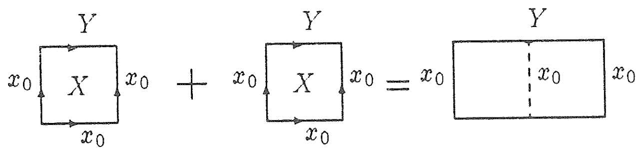


Composition of paths yields a (not necessarily abelian) group structure on $\pi_1(X, x_0)$.

Now if Y is a subspace of X containing the point x_0 then we can consider the second relative homotopy group $\pi_2(X, Y, x_0)$. This group consists of homotopy classes of continuous maps $g : [0,1] \times [0,1] \to X$ from the unit square into X which map three edges of the square onto the point x_0 and the fourth edge into Y. The appropriate picture of such a map g is



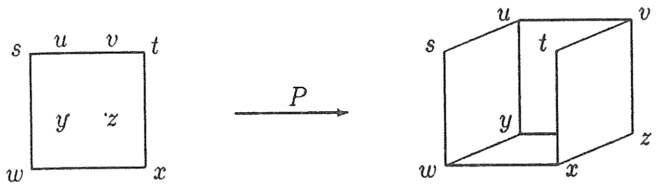
Juxtaposition of squares



yields a (not necessarily abelian) group structure on $\pi_2(X, Y, x_0)$.

By restricting to the fourth edge of the unit square we obtain a boundary homomorphism $\partial : \pi_2(X, Y, x_0) \to \pi_1(Y, x_0)$. Moreover there is an action of

$\pi_1(Y, x_0)$ on $\pi_2(X, Y, x_0)$ which makes the boundary homomorphism into a crossed module. To understand this action note that there is a continuous map



from the unit square onto four faces of the unit cube which sends s to s, t to t and so on. Now given a path $f : [0,1] \to Y$ representing an element of of $\pi_1(Y, x_0)$, and a square $g : [0,1] \times [0,1] \to X$ representing an element of $\pi_2(X, Y, x_0)$, we can construct a continuous map $^f g$ from the four faces of the unit cube to the space X by using g to map the face $uvyz$, and mapping each horizontal line in the remaining three faces by f. On composing $^f g$ with p we get a map which represents an element of $\pi_2(X, Y, x_0)$. It can be checked that the assignment $(f, g) \mapsto {}^f g \circ p$ induces an action of $\pi_1(Y, x_0)$ on $\pi_2(X, Y, x_0)$.

Example 6 Let M and N be normal subgroups of G. A non-abelian tensor product $M \otimes N$ has been introduced by R. Brown and J.-L. Loday [B-L]; it is the group generated by the symbols $m \otimes n$ ($m \in M$ and $n \in N$) subject to the relations

$$\begin{aligned} mm' \otimes n &= (mm'm^{-1} \otimes mn'm^{-1})(m \otimes n) \\ m \otimes nn' &= (m \otimes n)(nmn^{-1} \otimes nn'n^{-1}) \end{aligned}$$

for all $m, m' \in M$ and $n, n' \in N$. In general $M \otimes N$ is a non-abelian group. If however conjugation in G by an element of M (resp. N) leaves all the elements of N (resp. M) fixed then $M \otimes N$ is precisely the usual abelian tensor product of abelianised groups $M/M' \otimes N/N'$. For any normal subgroups M and N there is a homomorphism $\partial : M \otimes N \to G$ defined on generators by $\partial(m \otimes n) = mnm^{-1}n^{-1}$. There is also an action of G on $M \otimes N$ defined on generators by $^g(m \otimes n) = (gmg^{-1} \otimes gng^{-1})$. This homomorphism and action is a crossed module.

Example 7 Let Λ be an associative ring with identity, let $GL(\Lambda)$ be the general linear group, and let $E(\Lambda)$ be the subgroup of $GL(\Lambda)$ generated by the elementary matrices $e_{ij}(\lambda)$ with $i \neq j$ and $\lambda \in \Lambda$ (recall that $e_{ij}(\lambda)$ has 1's on the diagonal, λ in the (i,j) position, and 0 elsewhere). The group $E(\Lambda)$ is a normal subgroup of $GL(\Lambda)$, and the non-abelian tensor square $E(\Lambda) \otimes E(\Lambda)$ is known as the Steinberg group and denoted $St(\Lambda)$. (This definition of the Steinberg group is equivalent to the usual definition [B-L].) As a special case of Example 6 we have a crossed module $\partial : St(\Lambda) \to GL(\Lambda)$. It can be shown that $\partial(St(\Lambda)) = E(\Lambda)$. The groups $K_1(\Lambda) = \text{Coker}(\partial)$ and $K_2(\Lambda) = \text{Ker}(\partial)$ are known as the first and second algebraic K-theory groups of Λ.

The essential features of these examples are captured in the following definition.

Definition A *crossed module* consists of a group homomorphism $\partial : C \to G$ together with an action of G on C such that

(i) $\partial({}^g c) = g(\partial c)g^{-1}$,

(ii) ${}^{\partial c}c' = cc'c^{-1}$,

for all $c, c' \in C$ and $g \in G$.

If $\partial : C \to G$ and $\partial' : C' \to G'$ are crossed modules, then we say that a pair of homomorphisms $\varphi : C \to C$, $\psi : G \to G$ is a *morphism* of crossed modules if $\psi(\partial(c)) = \partial(\varphi(c))$ and $\varphi({}^g c) = {}^{\psi(g)}\varphi(c)$ for all $c \in C$ and $g \in G$.

An easy consequence of this definition is that for any crossed module $\partial : C \to G$ the group $\partial(C)$ is a normal subgroup of G; the quotient $G/\partial(C)$ is denoted by $\pi_1(\partial)$. Also it is easily checked that the action of G on C induces an action of $\pi_1(\partial)$ on $\text{Ker}(\partial)$, and that $\text{Ker}(\partial)$ is abelian; we denote the $\pi_1(\partial)$-module $\text{Ker}(\partial)$ by $\pi_2(\partial)$.

In all algebraic theories the notion of a free object is important. For a crossed module $\partial : C \to G$ the notion of "freeness" is made precise by saying that ∂ is *free* on a function $\eta : W \to G$ from some set W into G if:

(i) W is a subset of C;

(ii) η is the restriction of ∂ ;

(iii) for any crossed module $\partial' : C' \to G$, if $\nu : W \to C'$ is a function satisfying $\partial'\nu = \partial$, then ν induces a unique morphism $\varphi : C \to C'$, $\psi : G \to G$ of crossed modules with ψ the identity homomorphism.

Bearing Example 2 in mind, it is readily seen that free $\mathbb{Z}G$-modules are one instance of free crossed modules.

Another instance of free crossed modules arises from Example 5. For suppose that the space X can be constructed by choosing a point x_0 in X, then attaching copies of the unit interval $[0,1]$ to x_0 by gluing the end points 0 and 1 of each copy to x_0, and then finally attaching copies of the unit square $[0,1] \times [0,1]$ by gluing the edges of each copy along the various copies of the unit interval in some fashion. In other words, suppose that X is a ***reduced 2-dimensional CW-space***. The copies of the unit interval in X are called ***1-cells***, and the copies of the unit square are called ***2-cells***. Let Y be the subspace of X consisting of the 1-cells; in the jargon, Y is the ***1-skeleton*** of X. It was shown by J.H.C.Whitehead [W] that in this situation the boundary homomorphism $\partial : \pi_2(X, Y, x_0) \to \pi_1(Y, x_0)$ is a free crossed module. It is free on the function

$$\{\text{2-cells of } X\} \longrightarrow \pi_1(Y, x_0)$$

which sends each 2-cell to the element of $\pi_1(Y, x_0)$ represented by the boundary of the 2-cell.

To illustrate the above, suppose that X is the torus. Now the torus can be constructed by gluing together two 1-cells and one 2-cell. In this case we take Y to be the union of the two circles. Thus $\pi_1(Y) = F(a, b)$ is the free group on two elements a, b. The crossed module $\partial : \pi_2(X, Y, x_0) \to \pi_1(Y, x_0)$ is free on the function $\{w\} \to F(a,b)$, $w \mapsto aba^{-1}b^{-1}$.

Whitehead showed that the homotopy theoretic information contained in 2- dimensional reduced CW-spaces is completely captured in the algebra of free crossed modules. More precisely he showed that if X and X' are 2-dimensional reduced CW-spaces with Y and Y' their respective 1- skeleta, then the set of homotopy classes of continuous maps from X to X' is bijective with the set of (appropriately defined) homotopy classes of crossed module morphisms from $\pi_2(X, Y, x_0) \to \pi_1(Y, x_0)$ to $\pi_2(X', Y', x_0) \to \pi_1(Y', x_0)$. Using this bijection certain homotopy theoretic problems (such as the enumeration of the homotopy classes of maps from a compact connected closed surface to the projective plane) can be solved purely algebraically (cf. [E]).

The idea of studying 2-dimensional CW-spaces by means of their associated free crossed modules has applications to combinatorial group theory. Any presentation $< V : R >$ of the group G gives rise to a reduced CW-space with one 1-cell for each generator $v \in V$ and one 2-cell for each relator $r \in R$. The associated crossed module $\partial : C \to F(V)$ is free on the inclusion function

$R \hookrightarrow F(V)$ where $F(V)$ is the free group on V. Clearly $\pi_1(\partial)$ is isomorphic to G. The G-module $\pi_2(\partial)$ is known as *the module of identities*, and is a measure of the "non-trivial identities among the relations." A good introduction to this area can be found in [B-Hu].

A rather more algebraic use of free crossed modules is to do with the homology of groups. For suppose that $\partial : C \to G$ is a free crossed module, and let H denote the image of ∂ in G. It can be shown [E-P] that the commutator subgroup $[C, C]$ of C depends only on H. (In fact $[C, C]$ is isomorphic to the quotient of the non-abelian tensor product $H \otimes H$ by the subgroup generated by the elements $h \otimes h$ $(h \in H)$.) Moreover the intersection $[C, C] \cap \mathrm{Ker}(\partial)$ is isomorphic to $\mathrm{H}_2(H, \mathbb{Z})$, the second integral homology (or Schur multiplier) of H.

Crossed modules also have a role in the cohomology of groups. It has long been known that the second cohomology group $\mathrm{H}^2(G, A)$ of G with coefficients in a G-module A is bijective with the set of isomorphism classes of extensions of G by A. (Recall that a pair of group homomorphisms

$$A \xrightarrow{i} E \xrightarrow{p} G$$

is an *extension* of G by A if p is surjective, i is injective, $\mathrm{Im}(i) = \mathrm{Ker}(p)$, and the module action of $g \in G$ on $a \in A$ corresponds to conjugating by some $\tilde{g} \in p^{-1}(g)$.) In the mid 1970's various people (see [ML] for an incomplete list of references) discovered an analogous interpretation of the third cohomology group $\mathrm{H}^3(G, A)$ in terms of crossed extensions of G by A: a sequence of homomorphisms

$$A \xrightarrow{i} C \xrightarrow{\partial} N \xrightarrow{p} G$$

is a *crossed extension* of G by A if p is surjective, i is injective, $\mathrm{Im}(i) = \mathrm{Ker}(\partial)$, $\mathrm{Im}(\partial) = \mathrm{Ker}(p)$, and ∂ is a crossed module such that the resulting action of $N/\partial(C)$ on A corresponds to the module action of G on A. This interpretation has been used by Huebschmann [Hu] to obtain some new exact sequences in the cohomology of groups. In his recent book K.MacKenzie [MK] notes that the interpretation of $\mathrm{H}^3(G, A)$ carries over to the case of Lie groups and smooth morphisms. This leads him (via a more general result about Lie groupoids) to a reformulation of the Čech classification of principal bundles which works entirely in terms of abelian Čech cohomology.

One of the most fruitful areas in the theory of crossed modules stems from work by R. Brown and P.J. Higgins [B-Hi] on generalising to higher dimensions Van Kampen's famous theorem about the fundamental group of a space. This

theorem states that if a space X is the union of pathwise connected open subspaces U and V such that the intersection $U \cap V$ is pathwise connected and contains a point x_0 , then the fundamental group $\pi_1(X, x_0)$ is isomorphic to the amalgamated sum

$$\pi_1(U, x_0) \underset{\pi_1(U \cap V, x_0)}{*} \pi_1(V, x_0) ;$$

in other words the fundamental group construction preserves certain amalgamated sums. It has been shown by Brown and Higgins [B-Hi] that the crossed module construction on pairs of spaces given in Example 5 also preserves certain amalgamated sums. This new "2-dimensional Van Kampen theorem" is a useful tool in algebraic topology, and has led to several new results. Perhaps more importantly it has led to a successful search for an algebraic structure which will model n-dimensional homotopy theoretic phenomena, and which will satisfy some sort of Van Kampen theorem.

It has long been known that a crossed module $\partial : C \to G$ is equivalent to a set Q which possesses both a group structure and the structure of a category, the group multiplication being compatible with the category composition $\circ$ in the sense that $(x \circ y)(x' \circ y') = xx' \circ yy'$ for all $x, x', y, y' \in Q$ such that the left hand side of the equation is defined. As a group, Q is the semi-direct product $C \times G$. The category composition on Q is defined for those pairs of elements (c, g) and (c', g') satisfying $g' = \partial(c)g$, and is given by $(c', g') \circ (c, g) = (c'c, g)$. In [L] J.-L. Loday used this description of a crossed module to show that crossed modules are equivalent "up to homotopy" to connected CW-spaces X whose homotopy groups $\pi_i(X, x_0)$ are trivial for $i > 2$. He went further and showed that groups possessing n compatible category structures, which we now call *cat^n-groups*, are equivalent "up to homotopy" to connected CW-spaces X with $\pi_i(X, x_0) = 0$ for $i > n + 1$. His method was to assign to each space X a space W containing n subspaces $U_1, \ldots, U_n \subseteq W$, and then to construct from the $(n+1)$-tuple $(W, U_1, \ldots, U_n)$ a cat^n-group.

It has since been shown [B-L] that this construction of a cat^n-group from an $(n+1)$-tuple of spaces satisfies a Van Kampen type theorem (that is, it preserves certain amalgamated sums). The technicalities involved in using this n-dimensional Van Kampen theorem have lead to some interesting algebraic problems, such as the computation of amalgamated products of cat^n-groups. In [GW-L] it was shown that algebraic problems about cat^2-groups are often better reformulated using a non-trivial equivalence between cat^2-groups and algebraic structures known as *crossed squares*. (Intuitively a crossed square is a crossed module in the category of crossed modules. Thus it consists of a morphism of crossed modules together with an "action" of the target crossed module on the source crossed module, and certain algebraic conditions are satisfied.) More generally in [E-S] the notion of a *crossed n-cube* was introduced and shown to be equivalent to a cat^n-group. Since the publication of [L] in 1982 over 55 articles have been published on the subject of cat$n$-groups; a fairly comprehensive bibliography can be found in [B].

Finally we should mention that by imitating in other algebraic settings the equivalence between cat^1-groups and crossed modules, one arrives at the notion of a *crossed module* in these settings. Crossed modules of Lie algebras turn out to be useful in studying the cyclic homology of an associative algebra [K-L]. Crossed modules of commutative rings are useful in studying the Koszul Complex [P]. And many of the (topologically motivated) results on crossed modules of groups, such as the description of group cohomology, carry over to these other settings.

References

[B] R. Brown, *From groups to groupoids*, Bull. London Math. Soc. 19 (1987), 113-134.

[B-Hi] R. Brown and P.J. Higgins, *On the connection between the second relative homotopy groups of some related spaces*, Proc. London Math. Soc (3) 36 (1978), 193-212.

[B-Hu] R. Brown and J. Huebschmann, *Identities among relations*, Low Dimensional Topology, London Math. Soc. Lecture Note Series 48 (ed. R.Brown and T.L.Thickstun, Cambridge University Press, 1982), 153-202.

[B-L] R. Brown and J.-L. Loday, *Van Kampen theorems for diagrams of spaces*, Topology (3) 26 (1987), 311-335.

[E] G.J. Ellis, *Homotopy classification the J.H.C.Whitehead way*, Expositiones Mathematicae 6 (1988), 97-110.

[E-P] G.J. Ellis and T. Porter, *Free and projective crossed modules and the second homology group of a group*, J. Pure Applied Algebra 40 (1986), 27-31.

[E-S] G.J. Ellis and R.J. Steiner, *Higher dimensional crossed modules and the homotopy groups of (n+1)-ads*, J. Pure Applied Algebra 46 (1987), 117-136.

[GW-L] D. Guin-Walery and J.-L. Loday, *Obstructions 'a l'excision en K-théorie algébrique*, Evanston Conf. on Algebraic K-Theory 1980, Lecture Notes in Math. 854 (Springer, Berlin, 1981), 179-216.

[Hu] J. Huebschmann, *Group extensions, crossed pairs and an eight term exact sequence*, J. Reine und Angewandte Math. 321 (1981), 150-172.

[K-L] C. Kassel and J.-L. Loday, *Extensions central d'alg 'ebres de Lie*, Annales de L'institut Fourier, 32 (1982), 119-142.

[L] J.-L. Loday, *Spaces with finitely many homotopy groups*, J. Pure Applied Algebra 24 (1982), 179-202.

[MK] K. MacKenzie, *Lie Groupoids And Lie Algebroids In Differential Geometry*, Cambridge Univ. Press, 1987

[ML] S. MacLane, *Historical Note*, J. Algebra 60 (1979), 319-320.

[P] T. Porter, *Crossed module techniques in commutative algebra*, J. Algebra (to appear).

[W] J.H.C. Whitehead, *Combinatorial homotopy II*, Bull. American Math. Soc. 55 (1949), 453-456.

Department of Mathematics
University College Galway

Recent Computations of Pi

Donal Hurley

1 Early History

We first give a brief sketch of the history of computing π. Details can be found in [1].

Computations of the number π go back to the time of Archimedes (287–212 BC). He inscribed and circumscribed regular polygons on a circle with diameter 1. He began with hexagons and doubled the number of sides to get polygons of 96 sides which yielded the estimate

$$3\frac{10}{71} < \pi < 3\frac{1}{7}$$

By continuing to double the number of sides, one is in principle able to get as many decimal places of π as one desires. However, the convergence is slow, since the error decreases by about a factor of four per iteration. Until the discovery of calculus in the 17th century, efforts at calculating π relied on the method of Archimedes.

With the use of calculus, series were discovered for π. The formula of Leibniz

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$$

has a very slow convergence rate. Various other series and formulae were used in the computation of π, some of the more famous being

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots$$

due to James Gregory (1638–1675), and

$$\frac{\pi}{4} = 4\arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right)$$

due to John Machin (1680–1752). Machin substituted the Gregory formula for arctan into his formula to get 100 decimal places of π in 1706.

In 1844 Johann Dase (1824–1861) computed π correctly to 200 decimal places using the formula

$$\frac{\pi}{4} = \arctan\left(\frac{1}{2}\right) + \arctan\left(\frac{1}{5}\right) + \arctan\left(\frac{1}{8}\right)$$

and in 1853 William Shanks published 607 places, although the digits after the 527th place were incorrect. This error was not discovered until 92 years later when D.F. Ferguson produced 530 digits in one of the final hand computations. Two years later Ferguson used a desk calculator to get 808 digits.

The advent of digital computers saw a renewal of efforts to calculate even more digits of π. The first such computation was made in 1949 on ENIAC (Electronic Numerical Integrator and Computer) and 2037 digits were produced in 70 hours by John Von Neumann and his colleagues. In 1958 F. Genuys computed 10,000 digits on an IBM 704. In 1961 D. Shanks and J. W. Wrench Jr. calculated 100,000 digits in less than nine hours on an IBM 7090 [7]. The million-digit mark was set by J. Gilloud and M. Bouyer in 1973 in a feat that took under a day of computation on a CDC 7600. All these computations used series for arctan and identities such as Machin's.

Despite the increased speed of the computers, it was realised that there were limits to the number of digits which could be produced. An examination of the rate of convergence of the arctangent series shows that the arctangent method uses $O(n)$ full-precision operations to compute n decimals of π. By an operation we mean one of $+$, $\times$, $\div$, $\surd$. For example, the Shanks and Wrench computation of 100,000 decimals used 105,000 full-precision operations. Thus there are two basic time costs involved in doubling the number of digits; firstly, the number of operations increases by a factor of two, and secondly, the time for each full-precision operation is about twice as long. So doubling the number of digits lengthens computing time by a factor of four.

In 1975 Brent and Salamin [4,6], independently discovered an algorithm that dramatically lowered the time needed to compute large numbers of digits of π. The Brent-Salamin algorothm requires only $O(\log n)$ full-precision operations for n digits of π, and the ideas used go back to the work of Gauss and Legendre in the early part of the 19th century. The formula for the algorithm exploits the speed of convergence of the defining sequences for the arithmetic-geometric mean of two numbers.

2 The Brent-Salamin Algorithm

If a and b are two positive real numbers, with $a > b$, then we have the familiar arithmetic-geometric mean inequality

$$\frac{a+b}{2} \geq \sqrt{ab}$$

Thus, from two positive numbers a and b we get a second pair, $(a+b)/2$ and $\sqrt{ab}$. If we iterate this process we obtain sequences $\{a_n\}$ and $\{b_n\}$ defined by

$$a_0 = a\,, \quad b_0 = b\,, \quad a_{n+1} = \frac{a_n + b_n}{2}\,, \quad b_{n+1} = \sqrt{a_n b_n}$$

The sequence $\{a_n\}$ is strictly decreasing and bounded below, while the sequence $\{b_n\}$ is strictly increasing and bounded above. A simple computation, beginning with $a_{n+1}^2 - b_{n+1}^2$, shows that

$$a_{n+1} - b_{n+1} < \frac{1}{2}(a_n - b_n)$$

and so one concludes that the sequences have a common limit, which is denoted by $\mathrm{AG}(a,b)$. It can also be shown that

$$a_{n+1} - b_{n+1} < \frac{(a_n - b_n)^2}{8\mathrm{AG}(a,b)}$$

so that $a_n - b_n$ approaches 0 quadratically.

Gauss (1777–1855) studied these limits in his work on elliptic integrals. A complete elliptic integral of the first kind is given by

$$K(a,b) = \int_0^{\pi/2} \frac{d\theta}{\sqrt{a^2\cos^2\theta + b^2\sin^2\theta}}$$

The change of variable $t = a\tan\theta$ yields

$$K(a,b) = \int_0^{\infty} \frac{dt}{\sqrt{(t^2+a^2)(t^2+b^2)}}$$

A further substitution helps to make the connection between this integral and $\mathrm{AG}(a,b)$. If we put $u = (t - \beta^2/t)/2$ in the last integral, we get

$$\begin{aligned} K(a,b) &= \int_0^{\infty} \frac{du}{\sqrt{(u^2+a_1^2)(u^2+b_1^2)}} \\ &= K(a_1,b_1) \end{aligned}$$

Repeating this, we have

$$K(a,b) = K(a_1,b_1) = \ldots = K(a_n,b_n) = \ldots$$

By continuity of the integral K in its arguments, we have

$$K(a,b) = K\left(\mathrm{AG}(a,b), \mathrm{AG}(a,b)\right) = \frac{\pi}{2\mathrm{AG}(a,b)}$$

Thus

$$\mathrm{AG}(a,b)\int_0^{\pi/2} \frac{d\theta}{\sqrt{a^2\cos^2\theta + b^2\sin^2\theta}} = \frac{\pi}{2} \tag{1}$$

This was used to compute elliptic integrals by Gauss.

A second relation of Gauss relates the arithmetic-geometric mean $\mathrm{AG}(a,b)$ to complete elliptic integrals of the second kind. These integrals are

$$E(a,b) = \int_0^{\pi/2} \sqrt{a^2\cos^2\theta + b^2\sin^2\theta}\, d\theta$$

Since the elliptic integrals satisfy the homogeneity relations

$$K(\lambda a, \lambda b) = \frac{1}{\lambda}K(a,b)\,, \quad E(\lambda a, \lambda b) = \lambda E(a,b)$$

the variables can be normalised to $a = 1$. There is a relation between these due to Legendre (1752–1833). For $0 < x < 1$ and $0 < y < 1$, where $x^2 + y^2 = 1$,

$$K(1,x)E(1,y) + K(1,y)E(1,x) - K(1,x)K(1,y) = \frac{\pi}{2} \tag{2}$$

For a proof of this, see [2] and [3]. Using the relation (2), Gauss then proved the following:

$$E(a,b) = \left[a^2 - \sum_{n=0}^{\infty} 2^{n-1}\left(a_n^2 - b_n^2\right)\right] K(a,b) \tag{3}$$

The details of this are in [3].

Following the presentation in [5], we now derive a formula for π. If in (2), $x = y = 1/\sqrt{2}$, then with $K = K(1, 1/\sqrt{2})$ and $E = E(1, 1/\sqrt{2})$, we have

$$2KE - K^2 = \frac{\pi}{2} \tag{4}$$

In the Gauss relations (1) and (3), if $a = 1$ and $b = 1/\sqrt{2}$, then

$$K = \frac{\pi}{2\mathrm{AG}(1, 1/\sqrt{2})} \quad \text{and} \quad E = (1 - S)K \tag{5}$$

where

$$S = \sum_{n=0}^{\infty} 2^{n-1} \left(a_n^2 - b_n^2\right)$$

From (4) and (5)

$$2K^2(1 - S) - K^2 = \frac{\pi}{2}$$

that is,

$$\frac{\pi^2}{4\left(\mathrm{AG}(1, 1/\sqrt{2})\right)^2}(1 - 2S) = \frac{\pi}{2}$$

giving

$$\pi = \frac{2\left(\mathrm{AG}(1, 1/\sqrt{2})\right)^2}{1 - 2S} \tag{6}$$

But

$$\begin{aligned} 1 - 2S &= 1 - \sum_{n=0}^{\infty} 2^n \left(a_n^2 - b_n^2\right) \\ &= 1 - (1 - \frac{1}{2}) - \sum_{n=1}^{\infty} 2^n \left(a_n^2 - b_n^2\right) \end{aligned}$$

since $a_0 = a = 1$ and $b_0 = b = 1/\sqrt{2}$. Thus

$$1 - 2S = \frac{1}{2} - \sum_{n=1}^{\infty} 2^n \left(a_n^2 - b_n^2\right)$$

Substituting this into (6) we get

$$\pi = \frac{4\left(\mathrm{AG}(1, 1/\sqrt{2})\right)^2}{1 - \sum_{n=1}^{\infty} 2^{n+1}\left(a_n^2 - b_n^2\right)}$$

This was discovered by Salamin in 1973 [6] and independently by Brent at the same time [4].

If we now define

$$\pi_n = \frac{4a_{n+1}^2}{1 - \sum_{n=1}^{\infty} 2^{n+1}\left(a_n^2 - b_n^2\right)}$$

then from error analysis, it can be shown that π_n converges to π quadratically [2,6]. This means, roughly, that the number of correct digits doubles from one value of π_n to the next.

The Brent-Salamin algorithm was implemented in Japan in 1983 by Y. Kanada, Y. Tamura, S. Yoshino and Y. Ushiro to compute 16,000,000 digits in less than 30 hours.

In recent years the algorithm has been modified by the brothers Jonathan and Peter Borwein (natives of St. Andrews, Scotland, and both at Dalhousie University, Novia Scotia) to obtain iterative algorithms for computing π. Details of these are in [2] and [3]. These algorithms are now being implemented to compute π. In January 1986, D.H. Bailey of the NASA Ames Research Center produced 29,360,000 decimal places using one of the Borwein algorithms iterated 12 times on a Cray–2 supercomputer. A year later, Y. Kanada and his colleagues carried out one more iteration to obtain 134,217,000 places on a NEC SX–2 supercomputer. Earlier this year Kanada computed 201,326,000 digits on a new supercomputer manufactured by Hitachi, requiring only six hours of computing time.

3 Utility

One may ask what is the point of all of this, since about 40 decimal places is all one requires for any application imaginable. One use is that the calculation of π has become a benchmark in measuring the sophistication and reliability of the computers that carry it out. In addition, pursuit of more accurate ways has led researchers into intriguing and unexpected areas of number theory. Finally, a statistical analysis of the first 10,000,000 by Y. Kanada shows that the digits are distributed in a way that is expected from the conjecture that π is normal. This means that the frequency of appearance of each string s of digits of length m is asymptotically equal to 10^{-m} i.e.,

$$\lim_{n\to\infty} \frac{N(s, n)}{n} = 10^{-m}$$

where $N(s,n)$ is the number of occurrences of s in the first n digits of π. Because of this, the digits of π are sometimes used in algorithms to generate sequences of random numbers.

References

[1] P. Beckman, *A History of* π, The Golem Press (1970).

[2] J.M. Borwein and P.B. Borwein, *The arithmetic-geometric mean and fast computation of elementary functions*, SIAM Review 26 (1984), 351–366.

[3] J.M. Borwein and P.B. Borwein, *Pi and the AGM: A Study in Analytic Number Theory and Computational Complexity*, J. Wiley and Sons Inc., (1986).

[4] R.P. Brent, *Multiple-precision zero-finding methods and the complexity of elementary function evaluation*, in *Analytic Computational Complexity*, ed. J.F. Traub, Academic Press (1976), 151–176.

[5] H. Flanders, *Computing* π, College Mathematics Journal 18 (1987), 230–235.

[6] E. Salamin, *Computation of* π *using arithmetic-geometric mean*, Math. of Computation 30 (1976), 565–570.

[7] D. Shanks and J.W. Wrench, *Calculation of* π *to 100,000 decimals*, Math. of Computation 16 (1962) 76–99.

Mathematics Department
University College
Cork

Topological Equivalents of the Axiom of Choice

S.D. McCartan

Recall that, within the terms of Von Neumann-Bernays-Gödel set theory, one form of the axiom of choice (abbreviated AC) is stated as follows:

> *If* $\{X_i : i \in I\}$ *is a non-empty disjoint family of non-empty sets, then there exists a set* C *such that* $C \cap X_i$ *is a singleton for each* $i \in I$.

The axiom of choice has become virtually indispensable in mathematics since a large number of important results have been obtained from it in almost all branches of the subject without leading to a contradiction. However, although this axiom is consistent with, yet independent of, the other axioms of set theory, its status has long been a source of controversy and not all mathematicians are willing to accept it. Perhaps the principal appeal of the axiom of choice resides in the extensive list of its logical equivalents which exist in apparently disparate areas of mathematics. A fairly comprehensive dossier of these was compiled by the Rubins [4] in 1963.

Most topologists side with the majority of mathematicians, assume the axiom of choice, and do not hesitate to use it whenever necessary. Indeed some would argue that the following proposition (usually known as Tychonoff's theorem) constitutes the single most important result in general topology:

> *The product of a family of non-empty compact topological spaces is compact.*

The point here is that Tychonoff's theorem is logically equivalent to the axiom of choice (see [3]). In this note some other such topological equivalents are introduced.

Classically a topological space (X,τ) is said to be a T_0*-space* (T_1*-space*) if and only if for every pair of distinct points in X there exists a τ-neighbourhood of one which does not contain the other (exist τ-neighbourhoods of each which do not contain the other). Properties like T_0 and T_1, when possessed by a topological space, essentially express a degree of separation enjoyed by the points in the space. A non-empty subset Y of space (X,τ) is said to be *dense* (*codense*) if and only if there exists no non-empty τ-open (τ-closed) subset H of X such that $Y \cap H$ is empty. Let us call Y *thick* if and only if there exists no non-empty τ-open and τ-closed subset H of X such that $Y \cap H$ is empty. Evidently if Y is either dense or codense then it is thick.

Given some topological invariant property P, consider the following statements:

(MP) every topological space (X,τ) has a subspace $(Y,\tau|Y)$ (where $\tau|Y$ is the relativization of τ to Y), with property P, which is maximal (with respect to inclusion);

(DP) every topological space (X,τ) has a subspace $(Y,\tau|Y)$, with property P, which is dense (in (X,τ));

(CP) every topological space (X,τ) has a subspace $(Y,\tau|Y)$, with property P, which is codense (in (X,τ));

(TP) every topological space (X,τ) has a subspace $(Y,\tau|Y)$, with property P, which is thick (in (X,τ)).

It is clear that either of DP or CP implies TP. Schnare [5] showed that MT_0 and MT_1 are each equivalent to AC, and, here, his results are used to confirm that the same is true for DT_0 and CT_0.

Theorem 1 *The following statements are equivalent:*
(i) AC
(ii) DT_0
(iii) CT_0.

Proof (i) implies (ii). Let (X,τ) be any topological space so that, by hypothesis and [5], there exists a maximal T_0 subspace $(Y,\tau|Y)$. Then Y is dense in (X,τ), otherwise there exists a τ-open subset H (of X) which is disjoint from Y and contains a point x, so that, since $\{x\}$ is $\tau|Z$-open, the subspace $(Z,\tau|Z)$ is T_0, thereby contradicting the maximality of Y (where $Z = Y \cup \{x\}$).

(i) implies (iii). Replace the word "dense" by "codense" and the word "open" by "closed" throughout the argument above.

(ii) implies (i). Let $\{X_i : i \in I\}$ be a disjoint family of non-empty sets, let $X = \bigcup\{X_i : i \in I\}$, and consider the partition topology

$$\tau = \{G \subseteq X : G \cap X_i \neq \emptyset \text{ implies } X_i \subseteq G\}$$

If $(Y,\tau|Y)$ is a dense T_0 subspace of (X,τ), then, by its density, Y meets each X_i but, since Y is T_0, in exactly one point.

(iii) implies (i). Replace the word "dense" by "codense" throughout the argument above.

However, DT_1 and CT_1 are false.

Example Let X be the set of real numbers and consider the nested topology $\tau = \{G \subseteq X : G = (a,\infty), a \in X\} \cup \{\emptyset, X\}$ (where (a,∞) denotes the interval $\{x \in X : a < x\}$). It is immediate that any subspace of (X,τ) is nested, that any T_1 subspace is therefore a singleton and bounded, whereas any dense (codense) subspace is unbounded above (below). Observe that any singleton subspace is a maximal T_1 subspace which is neither dense nor codense.

In view of the equivalences obtained by Schnare, if T_α is any hereditary invariant property lying in logical strength between T_0 and T_1 (including those separation axioms discussed in [1] and [2]), it is tempting to conjecture that MT_α is equivalent to AC. So far this remains an open question. Although DT_α implies DT_0 and CT_α implies CT_0, so that, by Theorem 1, each implies AC, the example above seems to suggest (for all known such T_α which are strictly stronger than T_0) that DT_α and CT_α are false. Certainly (mindful that a space is called a T_{ES}*-space* if and only if every singleton subset is either open or closed), in the example, every T_{ES}-subspace is at most a doubleton while, indeed, every doubleton subspace is a maximal T_{ES}-subspace which is neither dense nor codense. That is, for instance, DT_{ES} and CT_{ES} are false.

On the other hand, we have:

Theorem 2. *The following statements are equivalent:*
(i) AC
(ii) TT_a

Proof Since AC implies MT_1, and TT_1 implies TT_α implies TT_0, it only remains to verify that MT_1 implies TT_1 and TT_0 implies AC.

MT_1 implies TT_1: Let (X,τ) be any topological space so that, by hypothesis, there exists a maximal T_1 subspace $(Y,\tau|Y)$. Then Y is thick in (X,τ), otherwise there exists a τ-open and τ-closed subset H (of X) which is disjoint from Y and contains a point x, so that, since $\{x\}$ is $\tau|Z$-open and $\tau|Z$-closed, the subspace $(Z,\tau|Z)$ is T_1 (where $Z = Y \cup \{x\}$), thereby contradicting the maximality of Y.

TT_0 implies AC: Repeat the argument of (ii) implies (i) in Theorem 1, with the word "dense" replaced by "thick".

Remarks It is interesting to contrast and compare MP, DP, CP and TP for a general invariant P. For example, if P is "connected", MP is true (since, as is well known, the maximal connected subspaces are the connected components), DP is false (since, as is well known, the closure of a connected subspace is connected), CP is false (since each connected subspace of a disconnected space, being contained in a component, is therefore disjoint from any other (closed) component) and TP is false (since each connected subspace of a locally connected disconnected space, being contained in a component, is therefore disjoint from any other (open and closed) component).

References

1 B. Johnston and S.D. McCartan, *Minimal T_F-spaces and minimal T_{FF}-spaces*, Proc. R. Ir. Acad. 80A (1980), 93–96.

2 S.D. McCartan *Minimal T_{ES}-spaces and minimal T_{EF}-spaces*, Proc. R. Ir. Acad. 79A (1979), 11–13.

3 M.G. Murdeshwar, *General Topology*, Wiley Eastern, New Delhi, 1983.

4 H. Rubin and J. Rubin, *Equivalents of the axiom of choice*, Studies in Logic, North Holland, Amsterdam, 1963.

5 P.S. Schnare, *The maximal T_0 (respectively, T_1) subspace lemma is equivalent to the axiom of choice*, Amer. Math. Monthly 75 (1968), 761.

Department of Pure Mathematics
Queen's University
Belfast

HISTORY OF MATHEMATICS

The Culmination Of A Dublin Mathematical Tradition

On The Maxwellian Struggle For A New Mathematical Physics And The Birth Of Relativity

N.D. McMillan

This paper is to celebrate the centenary of the Hertz Exporimatum Crucis that proved the FitzGerald electromagnetic theory of radio transmission.

Fitzgerald And The Electromagnetic Description of Light Propagation.

FitzGerald's chosen field of study in Dublin University for his Fellowship examinations in the period 1871–1877 was MacCullagh's mathematical researches. This study perhaps uniquely prepared him to comprehend the full significance of James Clerk Maxwell's development of an electromagnetic theory of light in 1865, which had until 1879 remained largely ignored, except for a handful of "electricians" from outside of the establishment of science and engineering.

FitzGerald and his uncle George Johnstone Stoney in Dublin, were the first mathematicians from the established universities to see Maxwell's work as the departure point for a programme of mathematical researches that would provide a unifying theory for physics. If successful of course, such a unified theory would have also established Cambridge and Dublin at the unchallenged head of developments in British science. There was at the time a determined ideological challenge to the scientific leadership based on the mathematicians in

This article is an abridged version of a longer, fully referenced article. Copies of the latter can be obtained from the author.

the established universities by "the practical men of science", and this threat was particularly keenly felt by FitzGerald in Trinity's Engineering School [1]. The Dublin Cambridge alliance had been developing since 1827 and this link was strengthened by Stoney's work in the 1860s, before Maxwell moved to Cambridge, in applying his kinetic theory to optical, astronomical and thermodynamic problems.

FitzGerald realised, perhaps in 1876, that MacCullagh's mechanical equations were the key to the problems of electrodynamics and worked on those also of Cauchy, Neumann and Green. It was the great treatise of Maxwell which sparked FitzGerald's first significant work. FitzGerald realised that MacCullagh's mechanical equations could be transposed into the new electrodynamic form to provide a full theoretical description of the reflection and refraction of light at a boundary, leading as did MacCullagh's equations earlier, to Fresnel's law for both polarizations and to the MacCullagh's equations for the amplitude of the refracted and reflected rays on refraction.

FitzGerald in 1879 produced three important papers on "Electromagnetic Theory of the Reflection and Refraction of Light" published by the Scientific Transactions of the Royal Dublin Society. These papers were significant in a number of ways. FitzGerald here produced what were the first real applications of the new Maxwellian theory. The papers were subsequently rewritten for the Proceeding of both the Philosophical Transactions of the Royal Society and Part II in the Philosophical Transactions of the Royal Society. This work establish FitzGerald as an important mathematical physicist in Britain. Finally, it was Maxwell himself who refereed the papers.

It is worth noting that FitzGerald's papers did not really require any great new development of mathematical method, as he explained himself: "Following a slightly different line from his (Maxwell) I obtained the same results as to Wave propagation, reflection and refraction as to those obtained by MacCullagh" [2]. FitzGerald had however been forced to pay careful attention to the physical interpretation of the mathematical equations. He had clearly been helped in this respect by the heroic, but tragic, efforts of MacCullagh to find a physical interpretation for his mechanical equations that required a very unusual medium, the ether, through which the wave could propagate. Unfortunately for MacCullagh, this ether, if mechanical in its operation, had to be an elastic solid of a type physically unknown, in that it had a strain energy dependent on the rotation of the volume element, rather than that observed in solids, in which the strain energy is dependent upon the deformation of the volume element. MacCullagh had introduced into the mathematical description of light the curl, by in essence, solving the problem of the form of the equations for light in a medium. For the new electrodynamic theory such physical requirements could be explained without contradiction, and perhaps it was this contradiction which led to MacCullagh's suicide.

FitzGerald's methods in this work were essentially mathematical and he inclined in any case strongly to the view at this time that a description of optical and electrical phenomena in terms of an elastic solid ether was inconsistent with physical requirements, since this would need to be soft enough to allow the free motion of planets yet be "The means by which tramcars are driven by shearing stresses" [3]. On the other hand the ether could not be "as thin as jelly "as it is the possession of properties analogous to rigidity that require explanation" [4]. Significantly, it was at this time when FitzGerald was unfettered with any of the burden of model building that he made his greatest mathematical advances, which began with this work of subsuming MacCullagh into the body of Maxwellian theory, while he was in revolt against "the thraldom of the material ether" [5]. He never was however able to completely break with the mechanical notions.

The Maxwellian Programme in Dublin

For the Maxwellians to establish themselves at the leadership of world science, first and foremost they required a consistent mathematical theory, which could address any practical problems in electrodynamics or optics. They also sought to extend their domain into the whole body of physics and chemistry, and in particular in this respect, thermodynamics. It is essential to understand these points, if the Dublin mathematicians concern with ether modelling is to be fully appreciated. It is also necessary to see that Stoney and FitzGerald in the 1880s inaugurated an entire programme of research and educational reform.

Stoney's mathematical researches really prepared the ground for the development of the Maxwellian programme in Dublin and he was a major influence on his younger relative FitzGerald. Stoney's early researches integrated into the Dublin programme, once this took final shape after 1879. Stoney had begun in 1861 a geometrical study of the examination of the conditions of propagation of undulations of planewaves in media. This interest in producing MacCullagh type geometrical procedures to generalize the treatment of optics, continued through his long and active life, with particularly important developments of these methods as late as 1897. Stoney's objective here was incredibly ambitious, since he sought to produce an analytical method for optics by which he resolved any wave front into flat wavelets, and thereby to do for optics what the calculus had done for geometry. In his last really major paper in 1896, Stoney extended his methods by resolving the wavefronts into spherical wavelets, to demonstrate the mechanism of image formation by interference.

The mathematical techniques evolved by Stoney in his optical researches, were subsequently applied to his major work on spectroscopy. His other major interest of thermodynamics began by his application of Maxwell's estimate of mean free path of a molecule, to obtain an estimate of the number of molecules present in a unit volume, and then in 1858 he demonstrated that inherent in Boyle's law was a model of a gas as an assembly of particles in constant motion and that such a gas could not be a continuous homogeneous substance. A decade later Stoney demonstrated that the motion of the gas molecules was related to the emitted radiation wavelengths using a comparative study of the relative magnitudes of the mechanical and optical components of the kinetic model. He then applied the kinetic theory to the interpretation of emissions from the sun and stars. and also to the atmosphere of planets and satellites. These researches aimed to integrate optical and thermodynamic principles, the essential foundation for a unifying mechanistic ether theory.

Stoney and FitzGerald both saw the discovery of the Crooke's force as a vital experimental test of theories linking radiation and kinetic theories. Stoney developed a kinetic-radiation theory of Crooke's radiometer in 1876 and two years later FitzGerald improved the mechanical theory of Crooke's Force. In these researches Stoney showed there was a distinction between the translatory motion of the molecule, that determines its temperature, and other internal motions within the molecule to "occasion the spectral lines". Stoney's "molecular" model of 1868, was of course non-electromagnetic and entirely mechanical, with the action resulting from mechanical vibrations of molecules producing a series of waves in the ether. Stoney extended this model in 1871 to propose an explanation for the harmonic sequences of spectral lines. Stoney subsequently applied his theory to hydrogen molecules, to obtain good theoretical match with Angstrom's measurements.

Both Stoney and Fitzgerald carried out an impressive range of thermodynamic researches, which are outside the scope of this discussion, except to point out that they related in the most part to attempts to establish the relation of kinetic principles to the spectroscopic and other optical and electromagnetic phenomena such as Fluorescence.

The reforming Maxwellians of course required to present a rationalized new view of science and there is considerable evidence that Stoney, and subsequently FitzGerald with Trouton, spent considerable efforts in attempts to introduce radical reforms in scientific terminology and concepts. Stoney was a member of the BAAS 1862 Committee to consider the standard of resistance, and this committee eventually sat for fifty years with Stoney being a member for all but the last ten. Almost from the outset in 1863 Stoney took a lead with radical proposals suggesting a qualitative vocabulary for dimensional analysis of lengthine, massine, timine, forcine, velocitine and so forth, and very significantly for charge, electrine. In 1873, when the BAAS committee recommended a complete new system of units, Stoney submitted a minority report. Stoney proposed a totally radical unified electrostatic and electromagnetic system of units. This proposed system was based upon fundamental quantities in nature. He consequently proposed for the unit of charge, the "atomic" charge of the electron, which Stoney named and was the first to obtain a value for this fundamental charge. He brilliantly derived the value of charge from Faraday's Law of electrolysis. Stoney's system aimed to remove the necessity of establishing connecting co-efficients between quantities in a system of units.

Trouton and FitzGerald carried out studies on Ohm's Law in Electrolysis between 1886 and 1888 for the BAAS and this work made the distinction between ionization of solutions, in which electrical seperation of charges occurs completely, and ordinary dissociation, in which it does not. It was in this work FitzGerald introduced the term *ionization*. FitzGerald's notebooks are indeed quite full with his notes on ideas for systems of units and nomenclature. His grasp of these questions led to FitzGerald's suggestion one month after J.J.Thomson's discovery of the Cathode ray particles, or corpuscles as he termed these, that the cathode particle was in fact a free electron. Furthermore he proposed the very useful, but unused term *electronization*, for the process of molecular decomposition involved in the formation of cathode streams.

The use of the term electron, itself inexplicably had only been adopted by the Maxwellians in 1894, following the growing collaboration of FitzGerald, Lodge, Heaviside and Larmor, in which they employed rather indiscriminately the term "ion" in their correspondence. This situation changed from 19th July 1894 when FitzGerald wrote to Larmor, "Johnstone Stoney was here just now and he will send a copy of his paper on the double line and c, Stoney was rather horrified at calling these ionic charges 'ions'. He or somebody called them 'electrons' and the ions is the atom and not the electric charge".

The discovery and naming of the electron [6], was according to Joly [7] the most important service Stoney rendered science. Stoney had another claim to fame, which is most appropriate for the man whose vision presaged the modern SI system. Stoney introduced the term "oscillator frequencies" in his pioneering work on spectroscopy [8], which were the reciprocal wavelength and became known in due course as wavenumbers.

FitzGerald's interest in mechanical models no doubt can be originally traced to Stoney's passion for molecular and atomic theory. FitzGerald, certainly did his most important work from 1879 using only mathematical paradigms. FitzGerald's soft quarto notebook contains work on ether modelling and it appears such models assumed some importance in his thinking from 1881. This notebook is undated until the 1887 Honours Lectures, but the first notes in the volume concern calculations on the electromagnetic action of charging spheres. This would date his first serious studies in ether modelling about 1881, and demonstrate that in this book, these attempts continued in a substantial way up until 1887(?), by which time he had converted to the liquid vortex-sponge ether model, which had first been proposed by Kelvin in 1867.

From about 1884 FitzGerald's mind was crystallizing with respect to his prefered model of the ether, he knew that he was against a solid ether, the jelly ether, and the stagnant ether, which was wholely unable to account for electromagnetic phenomena. He had been impressed by J.J. Thomson's 1883 study which had shown that the simple vortex theory predicted that the enertia of atoms ought to increase, and their velocity increase, as their temperature rises, a fact at variance with observations. FitzGerald was to declare that "to suppose atoms to be simply ring vortices in a perfect liquid can hardly be an adequate theory". He was inclined towards a moving liquid model , and in 1885 he published his well known "bands-and-wheels" model of the ether. FitzGerald's rectangular array of spinning wheels on vertical axes connected to their neighbours by rubber bands around their rims in his ether model, did impressively mimic Maxwell's electromagnetic field. Light propagation was demonstrated by an impulsive angular displacement given to a wheel which caused the disturbance to move out from this centre as a wave of tightened stretched rubber bands. The purpose of this model, was explained in an unpublished paper , but its main purpose was to show how Maxwell's "electric displacement" could in fact be represented by a change of structure (tightening bands) and thus to demonstrate that Maxwell's theory was not "unmechanical". This point was of particular importance to FitzGerald, who despite his philosophical idealism was in physics a convinced materialist. From 1885 FitzGerald became increasingly convinced of the truth and reality of the

vortex-sponge model, which he believed was a "likeness" of the ether and not simply an analogue like his model. For FitzGerald ether modelling was a search for the ultimate questions of science for he declared, "with the innumerable possibilities of fluid motion it seems almost impossible but that an explanation of the properties of the universe will be found in this conception" [9].

The ultimate development of the ether model produced by FitzGerald was perhaps that given in his letter to Heaviside on the 23 August 1893: "If the ether is ultimately an incompressible liquid, and I can't conceive any simpler hypothesis, there must be actions between different things (whirls, vortices, what not) in it like the action between two vortex rings in a perfect liquid. Effects propagated but immediately presents each vortex ring in itself really infinite, each atom of matter is infinite, the most probable. Thus gravity." This model was really close to Hamilton's Boscovichean atomic model [1] and had developed a considerable distance from those FitzGerald had inherited from MacCullagh.

The most important Dublin atomic model was the famous Stoney Atomic Model, which he proposed in his classic paper *An Analysis of the Spectrum of Sodium* in 1891. This was an energy model of the atom, in which the accelerated motions of the electrons in the atom or molecule, was resolved by Fourier's theorem and resulted in the splitting of the spectroscopic lines into doublets and triplets. This was a Maxwellian electromagnetic atomic model. The idea of the model that the electron from its own elliptical (apsidal) motion would produce emissions with split states. These ideas are very close to the modern concept of two "spin" states of the electron, while the idea of apsidal motion was of course later taken up by Sommerfeld and the Irishman Orr. Also, it may be noted, Larmor later developed his concept of precession from his development of Stoney's ideas.

There were other aspects of the Maxwellian programme in Dublin which deserve a mention. FitzGerald in his central position in Irish science at Trinity, was able to establish a dominant influence in an Irish Astronomical Network [11], which carried out very important observational, photographic and photometric astronomical measurements. FitzGerald also inspired and supervised the production of two classic Maxwellian textbooks , by his former student Thomas Preston *On Light* and *On Heat*. Stoney inspired the first textbook *On the Electron Theory*, and a number of others such as Minchin in the Dublin circle produced impressive mathematical reform texts. The exposition of the electromagnetic paradigms was vital contribution of the Dublin Maxwellian programme.

The Dawn Of The Theory Of Relativity

In the years 1881 to his death in 1901, FitzGerald's fertile mind was working almost incessantly on the conundrum of the ether, which of course as we now know with the benifit of hindsight, was a struggle to establish a relativistic basis for physics. The ideas which dominated his thinking were profound and most contradictory to common sense, and classical physics. FitzGerald's inner mental struggle disrupted his sleep and he was tormented by insomnia. He appeared to the contemporary scientific community as an electromagnetic crank, but he was a prophet of a new but only partially crystallized world view.

In the light of this situation, and the fact that FitzGerald started to teach the contraction hypothesis in Dublin University about 1881, it was quite significant that he composed a verse for the BAAS Meeting that year on J.J. Thomson which ends on very personal note, "Feels that fools be but am willing to play the part".

It was at these yearly meetings in particular that Fitzgerald found a forum to sound his bizarre ideas. Larmor was to explain, "he was the life and soul of the debate, he was always ready with some semi-paradoxical but wholly suggestive idea" [11]. The most outlandish of these ideas was the contraction hypothesis.

FitzGerald's electromagnetic departures, which led to his great discovery had began with the reworking of the MacCullagh equations, that contained an inversive geometrical analysis of space. This work was done using Hamiltonian algebra that itself incorporated a vectorial view of space and time. Trinity also produced during the second half of the nineteenth century, a number of important geometers, but in particular George Salmon's researches were of importance to the theories of FitzGerald. Salmon worked principally on the theory of invariants and covariants of algebraic forms to the geometry of curves and sufaces, and in this research collaborated with the "invariant twins" Cayley and Sylvester in Cambridge.

Hamilton had produced a non-commutative algebraic description of space involving a time dimension. The formal break with Euclidean geometry was however left to Riemann to achieve in 1854. In 1868 Plücker published *A New Geometry of Space*, which took up Hamilton's idea of space being defined by a set of lines, "a cosmic haystack of infinitely thin, infinitely long straws". Later Lie and Klein extended and unified these two mathematical developments and showed that Euclidean space is related to four dimensional space

by transformational groups.

FitzGerald was attempting to produce an ether theory which would unify these geometric theories and electromagnetism. Following Stoney's discovery of the electron this modelling centred on the electron theory. With the discovery by spectroscopists of intra-atomic processes, FitzGerald had to refine his conceptions as he explained most candidly to his collaborator Heaviside in 1889, "I admire from a distance those who contain themselves till they worked to the bottom of their results but as I am not in the very least sensitive to having made mistakes I rush out with all sorts of crude notions in the hope that they may set others thinking and lead to some advance" [14].

The epoch making breakthrough in thought had been made by FitzGerald's introduction of the first *relativistic principle*, but he failed to identify this as the key to progress on the exposition of space-matter theories in physics. This failure caused him to historically lose the credit for his great discovery and this misfortune was compounded by some very harsh quirks of fate, as will be explained below.

FitzGerald's involvement with the Dutchman Heinrich Antoun Lorentz (1853–1928) arose from his first major paper *The Electromagnetic Theory of the Reflection and Refraction of Light*. In the review of this paper by Maxwell for the Philosophical Transactions the great man noted that the paper related to the work of Lorentz. It was in 1882 that FitzGerald first published a paper *On Electromagnetic Effects due to the Motion of the Earth*, in the Transactions of the R.D.S., and showed that he had at this early date developed a theoretical Maxwellian based notion of this important question, which had a central bearing on the famous Michelson-Morley experiment (1887).

It seems that the later development of Relativity has distorted the perception of the contemporary importance of this experiment, but FitzGerald certainly believed, as has been fully explained in the ether model discussion, that "gross matter" was held together by electrical forces and was to be explained by ether theories. FitzGerald had discussed this hypothesis with Lodge in 1892 and his tentative claim to priority has recently been dramatically borne out [15], and his priority established by the discovery of his lost publication in the journal Science. The FitzGerald-Lorentz correspondence which is preserved in the Lorentz collection at the Algemeen Rijksarchief, The Hague, reveals the events which unfolded in 1894. Lorentz wrote on November 10th, 1894 to FitzGerald mentioning Lodge's comments in Aberration problems of FitzGerald's hypothesis on the negative result of Mr. Michelson's experiment. Lorentz sent him a number of the proceedings of the Dutch Academy of Sciences (1892), in which he considered the subject of aberration on the basis of his development of the theory of the refraction of light and he asked, "you would oblige me very much by telling me if you have published your hypothesis. I have been unable to find it, and yet I should want to refer the reader to it". FitzGerald's reply dated 14th November 1894, is centrally important and will be quoted in full,

> My dear sir,
> I have been for years preaching and lecturing on the doctrine that Michelson's experiment proves, and is one of the only ways of proving, that the length of a body depends on how it is moving through the ether. A couple of years after Michelson's results were published, as well as I recollect, I wrote a letter to "Science" the American paper that has recently become defunct, explaining my view, but I do not know whether they ever published it, for I did not see the journal for some time afterwards. I am pretty sure that your publication is then prior to any of my printed publications for I have looked up several places where I thought I might have mentioned it but cannot find that I did. I certainly never wrote any special article about it as I ought to have done for the information of others besides my students here. I am particularly delighted to hear that you agree with me, for I have been rather laughed at for my view over here. I could not even persuade my own pupil Mr. Preston to introduce this criticism into his book on Light published in 1890 although I pressed upon him to do so and it was only after reiterated positiveness that I induced Dr. Lodge to mention it in his paper; but now that I hear you as an advocate and authority I shall begin to jeer at others for holding any other view. Thank you very much for your papers. I can make out their general drift and wish I were able to reciprocate by replying to you in Dutch.
> Yours most sincerely,
> Geo. Fran. FitzGerald.

Lorentz generously gave FitzGerald credit for independently establishing the hypothesis in his 1895 paper, and it was from this that Einstein developed the terminology "Lorentz-FitzGerald contraction", which was the only reason that the Irishman's name became associated with this relativistic hypothesis. The 1889 letter to Science, however, properly establishes FitzGerald's priority as this sets forward clearly this hypothesis and provides a physical meaning of the Michelson-Morley experiment. Lorentz's contribution to this contraction idea however is real and significant, because he set out to discover the conditions that the Maxwellian laws of Electromagnetism should be invariant, that is, have the same form in a moving and stationary frame to use the modern relativistic concept, and it was he therefore, and not FitzGerald, who generalized Hamilton's algebra of time and set Einstein on his path of discovery.

In 1881 J.J. Thomson began a new departure in theoretical electrodynamics in his investigation which aimed to determine directly the effects produced by moving charged bodies, by means of Maxwell's equations of the electric field combined with the appropriate conditions at the surface. The theory probed the behaviour of a charged body as to whether it carried along with it an electric field. An important result of the paper was that the magnetic field thus produced, possessed a kinetic energy which carried along by an electron involves an addition to its effective mass. In his paper on J.J. Thomson's experiment FitzGerald pointed out that the analysis offered did not give the correct magnetic force and that Thomson had added a term to make a vector potential circuital and, "thus on closer examination, each portion of the electric charge is found to act independently; and so far from being able to exclude the electric charges from view by merging them in interfacial conditions, it turns out that their convection is the sole cause of phenomena of the electric field. When considered for the point of view of the aether, the hypothesis is that the total current is circuital, which lies at the very root of Maxwell's theory, involves and is equivalent to the magnetic influence of moving charges, which was verified experimentally before this time by Rowlands, though doubt still occasionally arises on the part of unsuccessful experimenters" [16,17].

This statement is extremely important in understanding the relationship between FitzGerald's work and that of Stoney. It is also the basis of Larmor's claim that MacCullagh originated electronic theories. MacCullagh worked back from the experimental description of "crystalline reflexion" by Brewster and Seebeck to discover his function of the Lagrangian type, but he was unable to conceive of any kind of material elastic medium, as ordinarily understood, whose properties were represented by his equations, because internal stress forces could not produce the unbalanced torque his treatment demanded, and "that a mechanical theory of this type can subsist only if there are polar forces (e.g. quasi-magnetic) capable of compensating the torque, or if there is a kinetic reaction torque arising from a distribution of gyrostatic rotations forming a part of the constitution of the medium" [17].

FitzGerald did not for a long period do any further work on this question, as he was occupied in other related studies of the generation and detection of electromagnetic radiation. In 1893 he produced a significant paper *On the Period of Vibration of Disturbances of Electrification of the Earth.* Then in 1900 with the growing acceptance of his theories, he was inspired to write four papers on the topic and he suggested an experiment to test the nature of the ether. This experiment, like the Michelson Morley experiment before it, yielded a null result. The experiment was shown eventually by Lorentz to be equivalent to the earlier null result, but this was after FitzGerald's death.

Summary

Stoney and FitzGerald were both leader of the Unionist faction of the Irish scientific community whose power base in Dublin was the R.D.S. The great Protestant mathematical tradition in Dublin, which culminated in a glorious fashion with their work, was eclipsed following the victory of Nationalism early in our own century. Their influence continued least in part, in the work of other Irish mathematicians working in the Royal College of Science, the National University and the Queen's Colleges, and more recently the Dublin Institute for Advanced Studies.

FitzGerald was the theoretical father of modern radio and the initiator and prophet for a new relativistic electromagnetic world view. Stoney was the theoretical father of electron theories, and therefore of modern theoretical chemistry. The Dublin mathematicians were consequently the most significant axis of the Maxwellians, in their very significant theoretical battle for the new physics and chemistry.

The very important scientific developments described above can only be properly understood in the context of Irish history, but only then within the general context of the international developments of mathematical physics.

FitzGerald philosophically was an idealist and follower of George Berkeley. This Trinity philosophical commitment was a considerable inspiration to him in proposing his non-common sense theories. His famous hypothesis was not in any way a flight of absolute idealism, and we can see now with historical research that this was no ad hoc hypothesis, as so frequently suggested by those ignorant of the facts. FitzGerald, paradoxically for an absolute idealist, in his work in physics, demonstrated theoretically the inextricable relationship between theory and practice, and the impossibility of ever absolutely separating the action of the mind from the material world.

Stoney's discovery of the electron inaugurated a new era in theoretical chemistry and atomic physics, but his identification of its charge as the first quantized quantity in physics marks the experimental birth of quantum theory. Stoney's consequent exposition of electron theories prepared the way with FitzGerald's discoveries, for our own modern quantum mechanical view which has arisen in our own century to replace older philosophical tenents such as those of Berkeley.

References

[1] B. Hunt, *Practice vs Theory*, Paper written March 26, 1982, John Hopkins University (unpublished).

[2] G. F. FitzGerald, *On the Electromagnetic Theory of the Reflection and Refraction of Light*, Proc. Roy. Soc. 1879, received Oct. 26, 1878.

[3] G. F. FitzGerald, *The Relation between Ether and Matter*, Nature, July 19, 1900.

[4] G. F. FitzGerald, *Some Properties of the Ether*, Scient. Proc. R.D.S., January 19, 1885.

[5] op. cit., Note 2, submitted in 1878 and refereed by Maxwell, but not printed until 1880.

[6] The general adoption of the name electron came after the 1902 Lorentz Nobel Lecture On the theory of electrons and the propagation of light.

[7] op.cit., Note 54. Von Helmholtz in his Faraday Memorial Lecture, Fellows of the Chemical Society, Royal Institution, 5 April, 1881, independently drew attention to "definite elementary charges which behave like atoms of electricity".

[8] G. J. Stoney, *An Analysis of the Spectrum of Sodium*, Trans. R.D.S., Vol.4, 1891.

[9] op.cit., F.P. p.239 (1888).

[10] R. Kargon, *William Rowan Hamilton, Michael Faraday, and the Revival of Boscovitchean Atomism*, American J. of Physics 32, 792–795, 1964.

[11] N. McMillan, *The Irish Astronomical Network in the 19th century*, Irish Astronomical Journal (to appear).

[12] op.cit., F.P., p.xx.

[13] ibid, p.590.

[14] FitzGerald to Heaviside, February 4, 1889.

[15] A. M. Bork, *The 'FitzGerald' Contraction*, Isis, (1966), 57: pp. 192–207. The view that FitzGerald's claim to priority despite being of the most good natured and modest kind was presented here as being unfounded. In S. G. Brush, Notes on the History of the FitzGerald-Lorentz Contraction, Isis, (1967), 58, pp. 230–232, the published priority of FitzGerald is succinctly demonstrated in a devastating way to the Bork thesis.

[16] O. J. Lodge, Obituary Notice of the Royal Society, 1901.

[17] J. Larmor, obituary, F.P. p.xliv (emphasis added).

[18] ibid, p. xlliii.

School of Science
Regional Technical College
Carlow

NOTES

Diagonalising a Real Symmetric Matrix and the Interlacing Theorems

Donal P. O'Donovan

In most linear algebra textbooks the diagonalisation of a real symmetric matrix is accomplished by first proving that the eigenvalues are real and then proceeding to the orthogonal diagonalisation. Anton's book [1] notes in the preface that the first part of this can require an excursion into the theory of complex vector spaces. The purpose of this note is to show that a more direct route is possible if one proves the realness of the eigenvalues and the orthogonal diagonalisation simultaneously.

In itself this would be of very little interest, at least for mathematics students, who usually handle $\mathbf{C}^n$ as readily as $\mathbf{R}^n$. However what one is led to, is something much more, namely the Cauchy inequalities, between the eigenvalues of any finite dimensional self adjoint operator and its compressions [3], and also to the Courant-Fisher min-max formulae for the characteristic numbers [2[, and these are topics not usually found in Linear Algebra textbooks. So, many mathematicians must be unaware of them. Yet the interlacing that one finds is both elegant and useful. For example it gives in several lines, the proof, that a symmetric matrix is positive if the principal minors are all positive.

I find linear operators a better setting for diagonalisations than matrices, so we work with them. Recall that if U is any subspace of an inner product space V, and $T : V \to V$ is a linear operator then the compression of T to U is just the operator $P_U T : U \to U$, where P_U is the orthogonal projection onto U. For students who prefer matrices, if an orthonormal basis $u_1, ..., u_r$ for U is expanded to an orthonormal basis $u_1, ...u_r, u_{r+1} ..., u_r$ for V then the matrix for $P_U T$ is just that block of the matrix of T whose entries are in both the first r rows and first r columns. For those who like pictures.

$$[T] = \begin{pmatrix} t_{11} & \dots & t_{1r} & \dots & t_{1n} \\ \vdots & & \vdots & & \vdots \\ t_{r1} & \dots & t_{rr} & & \\ \vdots & & \vdots & \vdots & \\ t_{n1} & \dots & \dots & & t_{nn} \end{pmatrix} \Rightarrow [P_U T] = \begin{pmatrix} t_{11} & \dots & t_{1r} \\ \vdots & & \vdots \\ t_{r1} & \dots & t_{rr} \end{pmatrix}$$

Recall also that the characteristic numbers are just the eigenvalues of T, repeated according to multiplicity, arranged in descending order

$$\lambda_1(T) \geq \lambda_2(T) \ldots \geq \lambda_n(T),$$

as we shall write them. We will use $C_T(\lambda)$ to denote the characteristic polynomial for T, that is det $(T - \lambda I)$.

Theorem *Let T be a symmetric linear transformation on a finite dimensional real inner product space V, then its eigenvalues are real and T can be diagonalised with respect to an orthonormal basis.*

Proof We proceed by induction on n, the dimension of V. If $n = 1$, the statements are trivial, so suppose both statements are proven for $n-1$. Let V have dimension n and U be any $n-1$ dimensional subspace. Let S be the compression of T to U. Then we may apply the induction hypothesis to S, obtaining real characteristic numbers, $\lambda_1(S), \ldots, \lambda_{n-1}(S)$, and corresponding orthonormal eigenvectors $u_1, \ldots, u_{n-1}$. Choose u_0 to complete the orthonormal basis $u_0, u_1, \ldots u_{n-1}$, and consider the matrix representation for T.

$$T = \begin{bmatrix} t_0 & t_1 & . & . & . & t_{n-1} \\ t_1 & \lambda_1(S) & . & . & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ t_{n-1} & 0 & . & . & . & \lambda_{n-1}(S) \end{bmatrix} \qquad (*)$$

It is easy to see that

$$C_T(\lambda_1(S)) = -t_1^2 \prod_{j=2}^{n-1} (\lambda_j - \lambda_1)$$

So

$$C_T(\lambda_1(S)) = \begin{cases} \leq 0 & \text{if } n \text{ is even} \\ \geq 0 & \text{if } n \text{ is odd} \end{cases}$$

But $C_T(\lambda) \to +\infty$ or $-\infty$ when $\lambda \to +\infty$ according as n is even or odd. So in both cases the graph of $C_T(\lambda)$ crosses or at least touches the λ - axis. This gives a real root λ_0 (which must be $\geq \lambda_1(S)$). Now one solves for an eigenvector v_0 and applies the induction hypothesis to the compression of T to the orthogonal complement of v_0, giving the desired result.

We noted above the extra bit of information, namely that $\lambda_0 \geq \lambda_1(S)$. In fact we can deduce easily that if S is any compression of T to an $n-1$ dimensional subspace then

$$\lambda_1(T) \geq \lambda_1(S) \geq \lambda_2(T) \geq \ldots \geq \lambda_{n-1}(S) \geq \lambda_n(T)$$

which we will refer to as "interlacing".

Again we proceed by induction. For $n = 2$ we have

$$\begin{pmatrix} t_0 & t_1 \\ t_1 & \lambda_1(S) \end{pmatrix}$$

If $t_1 = 0$ the result is immediate. If $t_1 \neq 0$, then $C_T(\lambda_1(S)) = -t_1^2 < 0$, but $C_T(\lambda) \to +\infty$ as $\lambda \to \pm\infty$, so $\lambda_1(S)$ lies between the roots.

For arbitrary n, we have as before the matrix $(*)$ for T. First if any $t_i = 0$ then $\lambda_i(S)$ is an eigenvalue for T, and so are the eigenvalues of the matrix gotten by ignoring the i^{th} row and the i^{th} column. By the induction hypothesis $\{\lambda_j(S)\}_{j \neq i}$ interlace the eigenvalues of this second matrix. A moments thought shows that $\{\lambda_j(S)\}$ then interlace the set of all eigenvalues of T.

Next one sees that

$$\begin{aligned} C_T(\lambda_1(S)) &= -t_1^2 \prod_{j=2}^{n-1} (\lambda_j(S) - \lambda_1(S)) \\ &= -t_1^2 (-1)^{n-2} \prod_{j=2}^{n-1} |\lambda_j(S) - \lambda_1(S)| \end{aligned}$$

and

$$
\begin{aligned}
C_T(\lambda_2(S)) &= -t_2^2(\lambda_1(S)-\lambda_2(S))\prod_{j=3}^{n-1}(\lambda_j(S)-\lambda_2(S)) \\
&= -t_2^2|\lambda_1(S)-\lambda_2(S)|(-1)^{n-3}\prod_{j=3}^{n-1}|\lambda_j(S)-\lambda_2(S)|
\end{aligned}
$$

and so on.

Thus we see that if the $\lambda_i(S)$ are all distinct then the signs of $C_T(\lambda_i(S))$ alternate, and the roots of $C_T(\lambda)$ interlace the $\lambda_i(S)$. If some $\lambda_k(S) = \lambda_{k+1}(S)$, writing

$$C_T(\lambda) = t_0\prod_{j=1}^{n-1}(\lambda_j(S)-\lambda) - \sum_{i-1}^{n-1} t_i^2 \prod_{j\neq i}(\lambda_j(S)-\lambda)$$

we have $\lambda_k(S)-\lambda$ as a factor of $C_T(\lambda)$. Hence $\lambda_k(S)$ is an eigenvalue for T, and it is immediate from the form of $C_T(\lambda)$ above, that

$$\frac{C_T(\lambda)}{\lambda_k(S)-\lambda} = C_R(\lambda)$$

where R is the $(n-1)\times(n-1)$ matrix

$$
\begin{pmatrix}
t_0 & t_1 & \cdot & \cdot & \sqrt{t_k^2+t_{k+1}^2} & t_{k+2} & \cdot & t_{n-1} \\
t_1 & \lambda_1(S) & \cdot & \cdot & 0 & 0 & \cdot & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \\
\sqrt{t_k^2+t_{k+1}^2} & 0 & \cdot & \cdot & \lambda_k(S) & \cdot & \cdot & 0 \\
t_{k+2} & 0 & \cdot & \cdot & \cdot & \lambda_{k+2}(S) & \cdot & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
t_{n-1} & 0 & \cdot & \cdot & 0 & 0 & \cdot & \lambda_{n-1}(S)
\end{pmatrix}
$$

Applying the induction hypothesis to R gives the desired interlacing.

If $(t_{ij})_{i,j=1,n}$ is an $n\times n$ matrix then the principal minors $\Delta_k(T)$ are defined as

$$\Delta_k(T) = \det(t_{ij})_{i,j=1,k}$$

We want to show that if $\Delta_k(T) \geq 0$ for all $1 \leq k \leq n$ for an $n\times n$ symmetric matrix T then T is positive. This follows easily if we show all the eigenvalues are positive. Again we proceed by induction. The case $n = 1$ is clear. Now applying the induction hypothesis to $(t_{ij})_{i,j=1,n-1}$ we have all its eigenvalues positive. Then by the interlacing $n-1$ of the eigenvalues of T are positive. Then $\Delta_n(T) \geq 0$ shows that all of them are positive.

Finally, if $T^{(r)}$ is the compression of T to an $n-r$ dimensional subspace, we may successively invoke the interlacing r times to obtain the Cauchy inequalities

$$\lambda_i(T) \geq \lambda_i(T^{(r)}) \geq \lambda_{i+r}(T)$$

It follows that

$$\lambda_i(T) \leq \lambda_1(T^{(i-1)})$$

But by diagonalisation, equality can be achieved, so

$$\lambda_i(T) = \min \lambda_1(T^{(i-1)})$$

where the min is over all $n-i+1$ dimensional compressions, and of course

$$\lambda_1(T^{(i-1)}) = \max\{(T^{(i-1)}v, v) : ||v|| = 1\}$$

which gives the desired min-max characterisation.

We can also note that the interlacing result and its consequences are obviously also true for a self adjoint operator on a finite dimensional complex inner product space.

References

[1] H. Anton, *Elementary Linear Algebra*, Wiley, New York, 1984.

[2] R. Bellman, *Introduction to Matrix Analysis*, McGraw-Hill, New York, 1970.

[3] M. Marcus and H. Mink, *A Survey of Matrix Theory and Matrix Inequalities*, Allyn and Bacon, Boston, 1964.

School of Mathematics
Trinity College Dublin

CONFERENCE REPORTS

Cork Operator Conference 1988

The 3rd annual conference in Operator Theory and Operator Algebras took place in UCC from 29 June to 2 July 1988. It was supported by the European Office of Aerospace Research and Development, the Royal Irish Academy, the Irish Mathematical Society and UCC. There were 33 participants from 9 countries (Canada, France, Ireland, Poland, Sri Lanka, U.K., U.S.A., West Germany, Yugoslavia). The principal speakers were Joachim Cuntz on *Differential structures on C*-algebras* and *KK-theory and cyclic cohomology* and Larry Brown on *Semicontinuity and closed faces in C*-algebras* and *Existence of projections in multiplier algebras.*

The organizers wish to thank the participants and sponsors. Next year's conference will probably be in May.

Gerard Murphy

Groups In Galway 88

The tenth anniversary of Groups in Galway was celebrated by a three-day meeting (May 26-28, 1988) which was attended by some two dozen participants. There were twelve very interesting talks (details below) and the usual generous breaks for coffee etc. Moreover, the traditional pilgrimage to the Weir was not forgotton! It is a pleasure to thank the participants, the speakers, and the sponsors (Irish Mathematical Society, Royal Irish Academy and University College Galway) for their support.
The speakers were: G. Ellis (Galway): *Groups with category structures*, N. Gilbert (Bangor): *Factor stabilisers in the automorphism group of a free product*, T. Laffey (U.C.D.): *Order-transitive groups*, Dr. MacHale (Cork): *Automorphisms of groups sending many elements to their nth powers—a survey of results*, M. Newman (Canberra): *Groups of exponent four*, G. Sherman (Terre Haute): *What is the probability that a subgroup is normal?*, E. Ormerod (Canberra): *The Wielandt subgroup of a metacyclic p-group*, B. McCann (Galway): *Examples of normal products*, B. Goldsmith (D.I.T.): *Maximal order abelian subgroups of symmetric groups*, E. O'Brien (Canberra): *Algorithm for the determination of finite p-groups*, S. Andreadakis (Athens): *Residually finite*

HNN extensions of abelian groups, L. Kovacs (Canberra): *Classification of varieties of groups.*

In 1989, Groups in Galway reverts to its usual two-day format—the meeting will take place on Friday/Saturday, May 12-13. Further details should be available early in 1989, from the organiser: John McDermott, Mathematics Department, University College, Galway.

John McDermott

Fourth Dublin Conference on Matrix Theory and Its Applications

The conference was help in UCD on March 10 and 11, 1988. There were thirty participants, coming from Ireland, England, France, Holland, West Germany Portugal, the United States, Canada and Japan. The Oxford University Press put on a display of books on linear algebra and related topics. The following lectures were presented:
C. Johnson (College of William and Mary): *Precise intervals for the eigenvalues of an hermitian and a positive definite matrix*, B. Reichstein (Catholic University of America): *On a problem of expressing a cubic form as a sum of cubes of linear forms*, T.T. West (TCD): *Linear operators with finite nullity and defect*, L. Beasley (Utah State University): *Linear transformations on Boolean matrices*, G.N. de Oliveira (University of Coimbra): *Invariant polynomials of partitioned matrices*, R. Westwick (University of British Columbia): *Spaces of matrices of finite rank*, R. Thompson (University of California at Santa Barbara): *The Schubert calculus and spectral inequalities*, R. Harte (UCC): *Companion matrices revisited*, T.J. Laffey (UCD): *Some questions on integer matrices*, T. Furuta (Hirosaki University): *Matrix inequalities*, S. Barnett (University of Bradford): *Routh's array and euclidean remainders for polynomials from an observability matrix*, H. Bart (Erasmus University): *Complementary triangularization of pairs of matrices*, L. Cummings (University of Waterloo): *A characterization of the permanent by the Cauchy-Binet formula*, R. Gow (UCD): *Skew-symmetric matrices in characteristic two*, R. Puystjens (University of Gent): *A generalization of the kernel theorem for Moore-Penrose invertibility of morphisms*, R. Grone (San Diego State University): *The Laplacian of a graph*
Synopses of the conference lectures will appear in a report in Linear Algebra and its Applications.

Fergus Gaines

BOOK REVIEWS

INVERTIBILITY AND SINGULARITY FOR BOUNDED LINEAR OPERATIONS by Robin Harte
Marcel Dekker, 1987, 528 pp, $119.50, ISBN 0-8247-7754-9

Reading the preface I was struck by the thought that this book would be something of a journey into the author's mind. The prospect filled me with some trepidation. In fact the journey was both better and worse than I had expected. Let me explain. Early in the preface the author states

> We have tried to write an introduction to operator theory accessible to students meeting the definitions for the first time.

Good, I thought. A little later I met

> The reader will also probably find our obsession with incomplete spaces tedious.

This sounded warning bells. In fact it transpired that the word accessible is given a breath of understanding not usually accorded to the word. Of this the author is clearly aware, and hence the honesty of the second statement.

This book really contains two books. One of these is a compendium of the spectral theory (in all shapes and sizes) of linear operators on Banach spaces. The second is an exploration of what happens if the space is simply a normed linear space and not necessarily complete. If the reader's interest is in the first book, then tedious is probably too mild a word to use. However, if one has a thorough understanding of the material of the first book, then the second book has quite a lot of interest. But not as part of "an introduction to operator theory". Intertwining the two books means that the entire is not, as a single entity, suitable for learning the subject. There are other drawbacks. In the first five chapters the basic results of elementary Banach space theory are covered. These include the big three, the open mapping, the uniform boundedness and the Hahn-Banach theorems. As in most of the book it is clear that the author has put a lot of thought into the means of presentation, with creditable results. However there is also a lot of less

important material interspersed, much of it from the second book. There is not sufficient discrimination between the really basic material and that which would be consigned to the exercises and notes section of other texts. This evenhanded treatment persists throughout the book so that the multitude of notions and notations lacks a hierarchy of importance.

Speaking of notations, there are three and a half pages of special symbols at the end of the book. And even this is not enough. On page 40, one encounters "Strictly speaking, we should write something like $l_p(\Omega, T)$ for the restriction of the mapping T^Ω: we shall however continue to write T^Ω." Then three lines later one meets "$l_p(\Omega, T)$" and two lines later "$C_\infty(\Omega, T)$", at whose meaning one must guess, as it is not among the aforementioned three and a half pages.

Additionally, it is confusing to have $x \in X$ on one line and $x \in X^\Omega$, a function space, two lines later. Items such as these and the many typographical errors make the technical material more difficult. In the preface the author admits to many "silly mistakes—left as a series of unspoken exercises for the reader". Unfortunately silly mistakes make silly exercises.

Another example of the lack of a firm editor is the use of the term Hilbert Algebra, confessedly "in total defiance of standard terminology" to describe a C^*-algebra, whilst even acknowledging the fact that Hilbert algebra is already in general use for something else.

In general the book seeks to illuminate even slight nuances of difference and to treat things in great generality. Thus sequences are not confined to being indexed by the natural numbers N, but are indexed by any bornological space Ω (Chap 1, §8), to what gain is not clear. And yet in other places distinctions are blurred. "We can treat a polynomial as a mapping" (page 21).

The pity is that these and other instances distract the reader from the many fine sections, for example the canonical factorisation of Theorem 2.3.3.

I have commented on chapters 1–5 which are "Normed Linear Spaces", "Bounded Linear Operations", "Invertibility and Singularity", "Banach Spaces and Completeness", and "Linear Functionals and Duality". I should mention also that the reader is introduced here to "Enlargements" §1.9, 2.7, 5.7, which allow one to replace statements about an operator on a space X, with statements about an "enlarged" operator on an "enlarged" space, and to "Composition Operations" in §2.9, and §5.6. These also recur throughout the book but one is left to wonder about their importance to the subject.

Chapter 6 is "Finite Dimensional Spaces and Compactness". This runs from defining linear independence to almost upper/lower semi-Fredholm. It includes a nice treatment of the equivalence of norms on finite dimensional spaces.

Chapter 7 is "Operator Algebra and Commutativity". This contains a lot of basic functional analysis, for example the "Stone-Weierstrass Theorem". Chapter 8 is "Inner Products and Orthogonality". The title says it all— again basic functional analysis. Chapter 9 is "Liouville's Theorem and Spectral Theory". Another good basic chapter, including the beginnings of the theory of C^*-algebras and their representations.

Chapter 10 is "Comparison of Operators and Exactness". The introduction to this chapter states "The various kinds of invertibility have "relative" analogues, in which one operator is compared to another. If we mix both left and right comparisons and then specialize we come down to concepts of "exactness"". Enough said.

Chapter 11 is "Multiparameter Spectral Theory". This contains the Taylor spectrum, an idea toward which much of the book seems aimed. There is also useful material on the Silov boundary and Tensor Products.

A final section is a collection of "Notes, Comments, and Exercises" for each chapter. It is clear from these that the author has researched his subject with diligence and thoroughness, and this section adds greatly to the value of the book as a compendium of results. In fact one is tempted to suggest that the book might have been called "everything you ever wanted to know about Spectral Theory, but were afraid to ask—in case you were told". The one thing that could be found missing is some mention of the many areas in which Spectral Theory finds its applications, and which provide it's Raison d'Etre.

Much of the above comment may seem negative, so let me hasten to add that this is a book that I am glad to have on my shelves. It has appeal on three levels. Firstly, the standard introductory results of functional analysis and operator theory are all there. Secondly, it collects many of the more esoteric notions of Spectral Theory, and finally it contains the authors own ruminations on completeness and the lack of it.

Donal O'Donovan
School of Mathematics
Trinity College Dublin.

PROBLEM PAGE

Editor: Phil Rippon

Here are a couple of attractive problems which fall into the category of 'geometric doodling'. The first one appears in Coxeter's book 'An Introduction to Geometry', but I heard it first from my school maths teacher.

21.1 *What is the minimum number of (strictly) acute angled triangles into which a square can be partitioned?*

The next problem was asked recently by an OU maths student at summer school. It has a very neat solution and I'd be interested to hear of any references to it.

21.2 *Find a configuration of finitely many points in the plane such that the perpendicular bisector of each pair of the points passes through at least two of the points.*

Finally, a wonderful sequence problem due to John Conway, who offered a prize of \$1000 in July this year for a solution (his audience thought he had offered \$10,000!).

21.3 *Let $a(n)$, $n = 1, 2, \ldots$, be defined as follows: $a(1) = 1$, $a(2) = 1$, and*

$$a(n+1) = a(a(n)) + a(n+1-a(n)), \qquad n = 2, 3, 4, \ldots$$

Thus the sequence begins:

$$1, 1, 2, 2, 3, 4, 4, 4, \cdots.$$

The problem is to determine an integer N such that

$$\left|\frac{a(n)}{n} - \frac{1}{2}\right| < 0.05 \qquad \text{for } n > N.$$

A solution was given three weeks later by Colin Mallow, a mathematician working at Bell Labs. In September, a British newspaper offered a magnum of champagne for a solution and awarded two prizes a week later, one of which went to a pupil from St. Paul's School, London. The Problem Page has no prizes to offer, but anyone who takes a closer look at this sequence will be amazed by its properties!

Next, here are the solutions to the problems which appeared in December 1987.

19.1 *To each vertex of a regular pentagon, an integer is assigned in such a way that the sum of all five integers is positive. If three consecutive vertices are assigned the numbers x, y, z respectively and $y < 0$, then the following operation is allowed: the numbers x, y, z are replaced by $x + y$, $-y$, $z + y$ respectively. Such an operation is performed repeatedly as long as at least one of the five integers is negative. Determine whether this procedure necessarily comes to an end after a finite number of steps.*

I am grateful to Tom Laffey for sending me the background to this problem, which appeared in the International Mathematical Olympiad at Warsaw in 1986, including a booklet on the Olympiad published by the Australian Mathematics Competition organisation. The U.S. student Joseph Keane was awarded a special prize for his solution, which begins as follows:
"As with so many problems of this type the key to a solution is the discovery of a function whose value decreases when the given operation is performed but which is always a whole number."

Keane goes on to show that if the five integers are v, w, x, y, z, then the expression:

$$\begin{aligned}&|v| + |w| + |x| + |y| + |z| +\\ &|v + w| + |w + x| + |x + y| + |y + z| + |z + v| +\\ &|v + w + x| + |w + x + y| + |x + y + z| + |y + z + v| + |z + v + w| +\\ &|v + w + x + y| + |w + x + y + z| + |x + y + z + v| +\\ &|y + z + v + w| + |z + v + w + x|\end{aligned}$$

is decreased by $|s - y| - |s + y|$, where $s = v + w + x + y + z$, when the operation is applied to the triple x,y,z. Since $s > 0$ and $y < 0$, this expression has the required property, showing that the operation can be performed only a finite number of times.

A similar solution, provided by the proposer of the problem, uses the quadratic

expression

$$\frac{1}{2}\left[(v - x)^2 + (w - y)^2 + (x - z)^2 + (y - v)^2 + (z - w)^2\right],$$

which decreases by $-sy$ when the operation is applied to x,y,z. My own solution uses the expression

$$\begin{aligned}&v^2 + w^2 + x^2 + y^2 + z^2 +\\ &(v + w)^2 + (w + x)^2 + (x + y)^2 + (y + z)^2 + (z + w)^2 +\\ &(v + w + x)^2 + (w + x + y)^2 + (x + y + z)^2 + (y + z + v)^2 + (z + v + w)^2\end{aligned}$$

which decreases by $-sy$ also.

This approach to the problem generalises to n integers placed at each vertex of a regular n–gon, as does a remarkable alternative solution due to J.M. Campbell of Canberra, which proves in addition that the final configuration is independent of the order in which the operations are performed.

To explain this solution, we let a_1, a_2, a_3, a_4, a_5 denote the integers in order around the pentagon, and σ_j denote the operation which reverses the sign of a_j and adds a_j to its pentagon neighbours. The operation σ_j has a very simple effect on the sequence of progressive sums:

$$\ldots, -a_4 - a_5, -a_5, 0, a_1, a_1 + a_2, a_1 + a_2 + a_3, \ldots$$

defined by $s_0 = 0$, and $s_i = s_{i-1} + a_i$, $i \in \mathbb{Z}$, where a_i is defined by taking the subscript mod 5. Indeed, one easily checks that σ_j swaps any two consecutive sums of the form s_{j-1}, s_j (taking j mod 5) and leaves all other s_i unchanged. Also, the operation σ_j is allowed, that is $a_j < 0$, if and only if $s_j > s_{j+1}$, and so the operations σ_j can be performed until the sequence s_i is sorted into increasing order.

The fact that s_i is in increasing order after only finitely many operations follows from the observation that

$$s_{i+5} = s_i + s, \qquad i \in \mathbb{Z},$$

where $s = a_1 + a_2 + a_3 + a_4 + a_5 > 0$. This implies that the number N_i of terms less than s_i, which lie to the right of s_i, is finite. Note that $N_{i+5} = N_i$, $i \in \mathbb{Z}$, and also that, in sorting the sequence, the term s_i moves precisely

N_i places to the right. Hence the total number of operations required is $N_1 + N_2 + N_3 + N_4 + N_5$ and the final configuration is indeed independent of the order in which the operations σ_j are performed. Moreover, Campbell is able to give an explicit formula for the N_i and for the final numbers around the pentagon.

The next problem is known in Maths Education circles as the Krutetskii Problem (see page 150 of 'The Psychology of Mathematical Abilities in Schoolchildren', by V.A. Krutetskii, The University of Chicago Press) and is attributed to Lovász by Ross Honsberger in an article in 'The Mathematical Gardner' (a volume dedicated to Martin Gardner). It is followed by an intriguing variant proposed by my colleague John Mason.

19.2 *A finite number of petrol dumps are arranged around a racetrack. The dumps are not necessarily equally spaced and nor do they necessarily contain equal volumes of petrol. However, the total volume of petrol is sufficient for a car to make one circuit of the track. Show that the car can be placed, with an empty tank, at some dump so that, by picking up petrol as it goes, it can complete one full circuit.*

19.3 *The petrol dumps are arranged as in* **19.2**, *but this time the total volume of petrol is sufficient for two circuits of the track. Can two cars be placed with empty tanks at the same dump so that, by picking up petrol as they go, they can each complete one full circuit in opposite directions? (The cars may cooperate in sharing petrol from the dumps.)*

There are various ways to solve **19.2**, but the following approach has the advantage that it can be used to solve **19.3**. Indeed, I came across it while working on **19.3**.

Suppose that p_k, $k = 1, 2, \ldots, n$, denote the volumes of petrol at the dumps D_k, $k = 1, 2, \ldots, n$, in order anticlockwise around the circuit, and that d_k, $k = 1, 2, \ldots, n$, denote the distances from D_k to D_{k+1}. If p_k and d_k are measured in comparable units, say gallons, then the assumption is that:

$$p_1 + p_2 + \ldots + p_n \geq d_1 + d_2 + \ldots + d_n. \tag{1}$$

Let us take as inductive hypothesis the statement that a solution is always possible with n dumps. Certainly, this is true for $n = 1$. Now consider

$n + 1$ dumps $D_1, \ldots, D_{n+1}$ with associated petrol $p_1, \ldots, p_{n+1}$ and distances $d_1, \ldots, d_{n+1}$, such that

$$p_1 + p_2 + \ldots + p_{n+1} \geq d_1 + d_2 + \ldots + d_{n+1}.$$

If $p_i \geq d_i$, for each i, then the car can start from any dump and complete the circuit. Otherwise $p_i < d_i$, for some i, and we consider a new configuration with n dumps, in which D_i is removed and the petrol p_i is added to D_{i-1}. Since (1) holds for this configuration, a solution is certainly possible, by the inductive hypothesis.

Following this solution, the car leaves D_{i-1} with an unknown volume P of petrol such that

$$P \geq d_{i-1} + d_i.$$

Since $p_i < d_i$, we deduce that $P - p_i > d_{i-1}$. It follows that the car could have completed the original $n+1$ dump circuit, with the same starting point, by picking up p_i gallons of petrol at D_i instead of at D_{i-1}. Thus we have a proof by induction.

The set up is similar in **19.3** with dumps D_1, D_2, ..., D_n, distances d_1, d_2, ..., d_n and petrol p_1, p_2, ..., p_n, but now we assume that

$$p_1 + p_2 + \ldots + p_n \geq 2(d_1 + d_2 + \ldots + d_n). \tag{2}$$

Once again the inductive hypothesis is that a solution is always possible with n dumps, and this clearly holds for $n = 1$. Now consider $n + 1$ dumps $D_1, \cdots, D_{n+1}$ with associated petrol $p_1, \ldots, p_{n+1}$ and distances $d_1, \ldots, d_{n+1}$, such that

$$p_1 + p_2 + \ldots + p_{n+1} \geq 2(d_1 + d_2 + \ldots + d_{n+1}).$$

If $p_i \geq d_{i-1} + d_i$, for each i, then the two cars can begin at any dump and complete opposite circuits by picking up at each dump exactly enough petrol to reach the next dump. Otherwise $p_i < d_{i-1} + d_i$, for some i, so that

$$p_i = q_{i-1} + q_i, \qquad \text{where } 0 < q_{i-1} < d_{i-1}, \quad 0 < q_i < d_i,$$

and we consider a new configuration with n dumps, in which D_i is removed and the petrol q_{i-1}, q_i is added to D_{i+1}, D_{i-1}, respectively. Since (2) holds for this configuration, a solution is certainly possible, by the inductive hypothesis.

Following this solution, the 'anticlockwise' car leaves D_{i-1} with an unknown volume P of petrol such that

$$P \geq d_{i-1} + d_i.$$

Since $q_i < d_i$, we deduce that $P - q_i > d_{i-1}$. It follows that the 'anticlockwise' car could have completed the original $n+1$ dump circuit, with the same starting point and without altering the petrol rations of the 'clockwise' car, by picking up q_i gallons of petrol at D_i instead of at D_{i-1}. Since a similar argument applies to the clockwise car, we again have a proof by induction.

Remark Problems **19.2** and **19.3** are in fact special cases of a more general problem in which the two cars have different rates of petrol consumption. The above argument needs only slight modification to deal with this more general problem.

Phil Rippon
Faculty of Mathematics
Open University
Milton Keynes MK7 6AA, UK

Second Dublin Differential Equations Meeting

The Second Dublin Differential Equations Meeting will be held in NIHE Dublin on May 22–25, 1989. Invited speakers include S.S. Antman (Maryland), J. Carr (Heriot-Watt), W.N. Everitt (Birmingham), J.K. Hale (Georgia Tech.), R.E. O'Malley (Rennselaer) and V. Moncrief (Yale). The programme will include sessions of contributed talks and workshops on both theory and applications. Possible subjects for workshops include bifurcation theory, singular perturbations and gelation. Financial support has been received from the Irish and London Mathematical Societies. Those interested in participating are invited to write to Dr. D.W. Reynolds, School of Mathematical Sciences, N.I.H.E., Dublin 9.

INSTRUCTIONS TO AUTHORS

Authors may submit articles to the Bulletin either as TEX input files, or as typewritten manuscripts. Handwritten manuscripts are not acceptable.

Manuscripts prepared with TEX

The Bulletin is typeset with TEX, and authors who have access to TEX are encouraged to submit articles in the form of TEX input files. Plain TEX, AMS-TEX and LATEX are equally acceptable. The TEX file should be accompanied by any non-standard style files which have been used.

The input files can be transmitted to the Editor either on IBM or Macintosh diskette, or by electronic mail to the following Bitnet or EARN address:

MATRYAN@CS8700.UCG.IE

Two printed copies of the article should also be sent to the Editor.

Authors who prepare their articles with word processors other than TEX can expedite the typesetting of their articles by submitting the input file in the same way, along with the printed copies of the article. The file should be sent as an ASCII file.

Typed Manuscripts

Typed manuscripts should be double-spaced, with wide margins, on numbered pages. Commencement of paragraphs should be clearly indicated. Handwritten symbols should be clear and unambiguous. Ilustrations should be carefully prepared on separate sheets in black ink. Two copies of each illustration should be submitted: one with lettering added, and the other without lettering. Two copies of the manuscript should be sent to the Editor.