MICHAEL JAMES DAVID POWELL

29 July 1936 — 19 April 2015

MICHAEL J.D. POWELL

29 July 1936 — 19 April 2015

Elected FRS 1983

by M.D. Buhmann(*), R. Fletcher FRS(†), A. Iserles(‡) & P. Toint(¶)

(*) *Mathematisches Institut, Universität Gießen, Heinrich-Buff-Ring 44, 35392 Gießen, Germany*

(†) *Division of Mathematics, University of Dundee, Nethergate, Dundee, DD1 4HN, Scotland, UK*

(‡) *Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Rd, Cambridge CB3 0WA, UK*

(¶) *Department of Mathematics, The University of Namur, 61, rue de Bruxelles, B5000, Namur, Belgium*

Michael James David Powell was a British numerical analyst who was among the pioneers of computational mathematics. During a long and distinguished career, first at the Atomic Energy Research Establishment Harwell and subsequently as the John Humphrey Plummer Professor of Applied Numerical Analysis in Cambridge, he has contributed decisively toward establishing optimization theory as an effective tool of scientific enquiry, replete with highly effective methods and mathematical sophistication. He has also made crucial contributions to approximation theory, in particular to the theory of spline functions and of radial basis functions. In a subject that roughly divides into practical designers of algorithms and theoreticians who seek to underpin algorithms with solid mathematical foundations, Mike Powell refused to follow this dichotomy. His achievements span the entire range from difficult and intricate convergence proofs to the design of algorithms and production of software. He was among the leaders of a subject area which is at the nexus of mathematical enquiry and applications throughout science and engineering.

1 Early days

Mike Powell was born in Kensington, London on 29th July 1936 to a professional family which, while abounding in solicitors and Church of England, has had also a mathematical connection in Baden Powell FRS (1796–1860), a Savillian Professor at

Oxford, a renowned theologian and the father of Robert Baden-Powell, the founder of Boy Scouts. His childhood – war years – was spent in London and Sussex.

Mike's father William was in the forces during much of Mike's early childhood and the main burden of his upbringing fell on his mother Beatrice. She was a warm and engaging woman and took great interest in Mike's education. At the age of seven he was sent to Frensham Heights School in Farnham, Surrey. Frensham Heights School was a progressive establishment and Mike started there to discover the joys of mathematics, while at home much of his early preoccupation was with mathematical games and puzzles. Seven years later, though, and reflecting the unhappiness of his more traditional parents with the progressive education at Frensham Heights, Mike transferred to Eastbourne College. This was a much more traditional, middle-ranking Public School. Having moved to Eastbourne in the middle of his studies, Mike needed to catch up socially, but this was assisted by his very active and successful performance in a wide range of sports. Much of Mike's mathematical education in Eastbourne and his commitment to a mathematical career are due to the interest and efforts of his mathematics tutor there, Paul Hirst.[1] Interestingly enough, three future professors of Numerical Analysis (Beresford Parlett from University of California at Berkeley, Peter Graves-Morris from University of Bradford and Mike himself) studied at roughly the same time at the fairly small Eastbourne College.

In 1956, having completed his National Service as a junior officer in Royal Artillery, Mike won a scholarship at Peterhouse, Cambridge. He completed his undergraduate studies at the Faculty of Mathematics within two years but, while widely predicted a First, achieved a Second in the Finals. A likely reason is that he had had a heavy cold during the Tripos exams but, one way or the other, it was a setback. This caused Mike to abandon (so he thought) an academic career and, instead of embarking on doctoral studies, complete a one-year Diploma in Computer Sciences, focussing on numerical calculations and with an eye on an industrial career. There might be another reason, though. As a young child, Mike has had a bout of encephalitis. This involved hospital stay and had had some ongoing after effects. It might have also influenced the trajectory of his eventual career, well after the effects of illness have completely dissipated: as a child, he was told by somebody (in a unique mixture of ignorance and irresponsibility) that one effect of encephalitis is that his brain is likely to deteriorate in his late teens – this unfounded and stupid rumour might have eventually contributed to his decision not to follow his undergraduate studies by a research degree. A Fellowship of the Royal Society, a Foreign Membership of US National Academy of Sciences, a Corresponding Membership of

[1] Mike next encountered Paul Hirst in Cambridge in 1970ties, where the latter was Professor of Education.

Australian Academy of Science, a Cambridge professorship, a long list of prestigious prizes and a mathematical œuvre of lasting impact: some deteriorating brain!

Mike met Caroline Henderson, his future wife, in Cambridge. A romance blossomed and they were married in 1959. This was a partnership and love that lasted a lifetime and was always at the very core of Mike's life.

2 The Harwell years

The Harwell years span the period of 1959–76 when Mike was employed at AERE Harwell, a government funded research laboratory founded in 1945 to promote the peaceful use of atomic energy, and in particular to support the new generation of nuclear power stations. Mike was a member of Theoretical Physics Division. Within TP there was what would now be called a Numerical Analysis group, whose remit was to carry out fundamental research in numerical analysis, to write computer subroutines to allow existing and new techniques to be available to other researchers and technical staff at AERE, and to assist in promoting the most effective use of this methodology.

The Numerical Analysis group was initially led by Jack Howlett, along with Alan R. Curtis, with interests in the numerical solution of differential equations, and Mike, whose interests were in methods for optimization and approximation theory. There were also support staff whose main remit was to maintain a Fortran library of subroutines and advise on its use. Roger Fletcher with interests in optimization joined the group in February 1969 and later left in October 1973. John K. Reid with interests spanning differential equations, linear algebra and Fortran standards joined the group in 1969.

In the first few years at Harwell Mike was not free of frustrations, not least with the UK computing facilities, and at a certain stage contemplated seriously abandoning his job in favour of a career as an actuary. Fortunately, Walter Marshall (later Lord Marshall of Goring, FRS), the head of the TP Division, stepped in to prevent this nonsense.

Mike's early work was in a range of calculations in atomic physics and chemistry, but it is for his contributions to optimization that he became justly famous. Not just for his theoretical contributions, for he was an outstanding mathematician and proved many important results, but primarily for his insight in developing many new and effective numerical methods and their implementation as Fortran subroutines. This attitude towards doing useful and relevant research pervaded the entire group, largely due to Alan and Mike's influence.

To tell the optimization story, one has to understand the stage of optimization provision in the 1950s. There were two main strands. Commercial and defence interests developed the idea of *linear programming* **(PJD)** and related topics, involving the minimization of a simple linear *objective function* subject to side conditions (*constraints*) such as linear equations and simple bounds on the variables. The other strand was *nonlinear optimization* involving nonlinear functions, both in the objective function to be minimized, and in the constraint equations. Both strands have had their own challenges, albeit of a different nature. Mike's interests were in nonlinear optimization, and initially in *unconstrained optimization*. The earliest methods had their inception in the physics community. One method, based on the iterative use of second order Taylor series expansions, is now universally referred to as Newton's method. But it was not popular at that time, mainly because of the need to derive and evaluate both first and second derivatives of the objective function. Another method, now called *Steepest Descent,* required only first derivatives, but was usually ineffective. Around the 1950's industrial companies such as ICI and Rolls Royce were becoming interested in minimizing what could be highly nonlinear functions, often for which derivatives were difficult or even impossible to obtain (*derivative-free optimization*). Such companies began to set up their own research groups to develop suitable methods, often of an heuristic nature.

Mike first presented a method for derivative-free optimization that improved on a method of Smith based on the concept of *conjugate directions.* Although no longer used, being superseded by other methods that Mike and others developed much later in his career, it helped focus the community attention on the concept of conjugacy as a way of building effective methods, notably the idea of *conjugate gradients* which was becoming known at about that time. He was also involved in developing methods for nonlinear least squares data fitting, for which there was considerable demand at Harwell and elsewhere. These included a derivative-free method [2, 3, 14] and somewhat later a 'dog-leg trajectory' [9, 10] that enabled the construction of a convergence proof of the so-called *Gauss–Newton method*.

However the big breakthrough in his career, and for the subject in general, came in 1962. Bill Davidon, a mathematical physicist at Argonne National Laboratories, had written a somewhat hard-to-read report about a new 'variable metric' algorithm **(WCD)**, of which Mike had obtained a copy. He was able to see through the complexities of the presentation and understand why the basic idea would have immense promise. At about this time he was scheduled to give a seminar at Leeds University to talk about his derivative-free method, but changed his title at the last minute to talk about Davidon's work. On arrival at Leeds he found that Roger Fletcher had also obtained a copy of the report and was writing a paper for the

Computer Journal, presenting some results showing the effectiveness of the method. So they pooled their resources and submitted a joint paper. The method became known as the *DFP method*, and caused a significant improvement in both problem size and reliability that could be obtained. Mike would relate a story that he was presenting a paper at a meeting in London where participants were speaking of the 'curse of dimensionality' that precluded the solution of problems with as few as ten variables, whereas he was able to present results in which problems with hundred variables were solved in quick time. This was an important advance given the limited capabilities of computers at that time. As time has progressed, this and similar methods are now able to accurately solve much larger problems. The methods only require first derivatives to be available, and approximate either the matrix of second derivatives (the Hessian matrix) required by Newton's method or its inverse, using information derived from differences in first derivatives between successive iterates. At each iteration, a search direction is computed by applying a variant of Newton's method using the approximate Hessian. A *line search* then follows, where, as the name indicates, an improvement in the function to optimize is sought along this direction. In the DFP method, the inverse Hessian approximation at a given iteration is expressed as a low-rank modification of the previous inverse approximation and is always positive-definite (resulting in a local convex quadratic approximation). This last property defines a new metric on the space of variables, which is why methods of this type have come to be known as *variable-metric methods*. Algorithms where positive-definiteness of the Hessian approximation is not mandatory (as in [10]) are called *quasi-Newton methods*. This class of methods has proved significant for many other aspects of optimization and has generated a large body of literature.

Another important advance from Mike came in the late 1960s. This related to optimization with nonlinear constraints. One common approach has been to convert the problem into an unconstrained problem, encouraged by the improving effectiveness of such methods. At that time the idea of a *penalty function* was usual, in which an increasingly large penalty for violating any constraint is added into the objective function. Mike saw that the correct solution could be obtained more effectively and reliably by a certain shift of origin in the constraint function. Based on his experience as a gunnery officer in national service he drew the (not entirely correct) analogy of getting a projectile to reach the target by shifting the angle of the gun barrel, rather than by increasing the initial thrust of the projectile. The shift of origin was shown to be related to the classical concept of Lagrange multipliers, and such methods became known as *augmented Lagrangian* methods. The concept is still very relevant at the present day.

Around the 1970s there was intense interest in the community in elucidating

Figure 1: DFP: Bill Davidon, Roger Fletcher FRS and Mike, Trinity Hall, Cambridge 1981 (photograph by Kunio Tanabe).

the theoretical properties of algorithms for the unconstrained optimization of non-quadratic functions, and Mike played an important part in this research. Practical experience showed that the performance of existing derivative-free algorithms was disappointing and could be bettered by using finite-difference versions of gradient algorithms. Indeed, Mike constructed an ingenious example [16] which showed that the early coordinate search algorithm could potentially fail to converge. So most of the interest was focussed on two main types of algorithm which require the gradient vector to be calculated. One type comprises the conjugate gradient algorithms, the attractive feature of which is that they avoid the need to store a Hessian approximation matrix, and so are potentially applicable to solve extremely large problems. And on the other hand there are quasi-Newton algorithms which do store a Hessian approximation, but are seen in practice to converge much more rapidly. Mike showed that conjugate gradient algorithms could never converge superlinearly, unless periodic restarts along the steepest descent direction were made. He also suggested [20] a restart technique which performed better than restarting along steepest descent. But restarts are ineffective for very large problems, and arguably this is because of such results that there is now little interest in the conjugate gradient algorithm for non-quadratic minimization, being supplanted by the more recent Barzilai–Borwein algorithm and generalisations thereof.

The possibility of proving convergence of quasi-Newton algorithms was a serious challenge, and Mike contributed a great deal to this study over the years. His proof for the DFP algorithm with an exact line search [11] was the first such (highly non-trivial) proof, and was a significant technical achievement. It was built upon in subsequent research, including other contributions from Mike [17, 22]. But perhaps his most influential theoretical contribution was to analyse a very simple two-variable quadratic case with an inexact line search, showing why the DFP algorithm could perform significantly worse than the competitor BFGS algorithm, a result which had been observed in practice. This has contributed to the now almost universal acceptance of the BFGS method as the method of choice in general. Mike also proposed [10] another formula for updating the Hessian approximation, which became known as the *PSB update.*. Mike also played a significant part in researching Newton and quasi-Newton methods for solving systems of nonlinear equations. A product of heated discussions in the Harwell canteen was the Curtis–Powell–Reid method [15] for constructing a sparse Jacobian estimate by finite differences.

Mike's interest and activity in approximation theory started in Harwell, fostered by practical problems arising with optimization algorithms. For the sake of an uninterrupted narrative, they are surveyed in Subsection 4.2.

3 The Cambridge years

In 1976 Mike Powell was appointed to the John Humphrey Plummer Chair in Applied Numerical Analysis at the Department of Applied Mathematics and Theoretical Physics in Cambridge.

These dry facts require some elaboration. The John Humphrey Plummer Foundation has been established in 1929 at the University of Cambridge to dispense funds left by a Mr Plummer of Southport, Lancashire, "in perpetuity for the promotion of education in Chemistry, Biochemistry, Physical Science or other allied subjects in the University of Cambridge". The Foundation endowed a small number of Professorships, each of an annual value of £1200 (fortunately, pegged to inflation), based in diverse science or medicine departments. Each appointment was for a single tenure and, upon being vacated, the professorial chair would be typically reincarnated in a different department, following the usual unfathomable rules of Cambridge politics.

Sir James Lighthill FRS was the prime mover in establishing the chair and in persuading Mike to come to Cambridge. The chair was set up with either James Wilkinson FRS or Mike in mind. Jim did not want to leave the National Physical Laboratory, so Mike was elected.

The Department of Applied Mathematics and Theoretical Physics has been established in 1959 by George Batchelor FRS, who remained its Head until his retirement in 1983. Consistently with its name, it brought together applied mathematicians and theoretical physicists, although it is only fair to point out that during Batchelor's headship 'applied mathematics' meant mostly 'applied fluid mechanics'. George was a strong individual, whose vision what applied mathematics is all about and what applied mathematicians are supposed to do brooked little dissent. He built a department of great international distinction and excellence, yet focussing on just one part of the wide contemporary spectrum of applied mathematics.[2] His idea of the new applied numerical analysis professor was of an individual who will help local fluid dynamicists in their 'numerical simulations'. He could not have been more wrong.

Mike Powell was singularly uninterested in fluid dynamics (computational or otherwise) or in numerical solution of differential equations. Neither was he interested in George Batchelor's plans or visions. As at Harwell, his dream scenario was to sit in his office and busily scribble on lined paper, or alternatively run FORTRAN programs in the computer room. By his own admission, he was a loner: "I collaborate far less than other people. It is very helpful sometimes to be able to talk things over with somebody else, but my main interest in mathematics is the problems themselves and in a way I'd quite like to crack them myself." **(PJD)**. Mike was never interested in building a large group and truly loathed any time-wasting on administration and departmental duties, and this made him largely immune to George Batchelor's sticks and carrots.

In fairness, the department was much smaller in late 1970s and the number of professors smaller still – this was in the old days, when the standard 'career grade' in Cambridge was a University Lectureship. This meant that Mike had the clout to shape his working life according to his own priorities, while George was realistic enough to recognise and respect an immovable object.

This working life was spent mostly behind a closed door, in the sanctity of S2, his office, sitting behind a large desk with an immaculately empty black top, except for few sheets of lined paper, covered with dense scribbles and corrections in his characteristic long hand. The group was small, few research students and the occasional visitor: "...I certainly don't want to talk to people and say how are you getting on

[2]Of course, there was always at DAMTP a strong theoretical physics group, of great international distinction.

more often than maybe once a month, something like that. My preference is to find a nice piece of mathematics or an idea for a new algorithm and to get stuck into it, and I like my visitors to do that too." **(PJD)**. After two years the team was joined by Arieh Iserles, who was elected to a Junior Research Fellowship at King's College, and whose interests in numerical differential equations complemented Mike's.

Arieh's arrival has demonstrated two features of Mike's character which coexisted, sometimes in a measure of tension, with his wish to be just allowed quality time with his theorems, algorithms and programs: his responsibility and generosity. Once Mike believed that it was his obligation to do something – be it teaching, administration or taking care of his group – he would suppress his grumbles and discharge his responsibilities in a truly exemplary fashion, never taking shortcuts. Thus, having realised that, while Arieh displayed some potential to become a useful mathematician, his level of mathematical exposition left much to be desired, Mike declared that he wished to read and comment upon everything Arieh has written up. Only much later Arieh realised what a painful exercise it was for Mike Powell to read sheaves of (badly written) research on computational differential equations! Yet, he read them carefully, with his copious comments filling the margins and every empty space.

Mike displayed similar commitment to his research students (with the caveat that at least the subject matter of their work interested him). His research time was precious – but never too important to listen to them, discuss their ideas and read their work. While Mike could be abrupt with colleagues, his patience with students was exemplary. Moreover – and this is crucial in understanding the huge devotion of Mike's students and colleagues – while he was a self-declared loner at work, he always had all the time in the world for them outside working hours. At Mike and Caroline's home his students, colleagues and visitors were always made to feel members of the extended Powell family and were always welcomed with great warmth and hospitality. Arieh Iserles remembers the moment, at a conference dinner in United States, when his wife exclaimed "I recognise almost everybody here and we met them all at Mike and Caroline!".

Insofar as departmental responsibilities – whether teaching or administration – were concerned, Mike's basic attitude was that they are a form of cruel and unusual punishment, to be avoided whenever possible. At the same time he recognised his obligation to share the burden and, once he undertook any departmental responsibility, he discharged it in an outstanding, almost obsessively exemplary fashion. His undergraduate lectures were always the epitome of logic, precision and clarity and their syllabi form to this day the spine of the schedules of numerical analysis courses in Cambridge Mathematical Tripos. These lectures were accompanied by meticulously written lecture notes (two sides of hand-written A4 per lecture) which were the embodiment of lucidity. His two-year Chairmanship of the Faculty Board of Mathematics displayed exemplary leadership and diligence.

Soon after arrival in Cambridge, Mike was considered for a Professorial Fellowship at Peterhouse, his old college, but this did not work out mostly because the sole mathematician on the committee – little in eminence but great in 'purity' – declared that Mike is "not a real mathematician". Soon thereafter this had a happy ending, Mike having been offered a Fellowship at Pembroke College and in 1979 became a Professorial Fellow there. Mike enjoyed his time at Pembroke and the friendly atmosphere there immensely. Playing bowls after lunch on Pembroke's lawn became a constant feature of his daily routine.

In 1979 Mike was awarded by University of Cambridge a Doctorate of Science (Sc.D.) in recognition of his published work: "the Master of Pembroke suggested that I should follow the procedure for becoming a Master of Arts. Rather than expressing my reservations about it, I offered to seek an ScD degree instead, which required an examination of much of my published work. Thus I became an academic doctor in 1979." **(LNV)**.

On 27th July 1983 Mikes son David was killed in an accident while on a holiday in France. Mike was devastated: this was a difficult and painful period in his life. David was an undergraduate mathematics student at University of Bath and Mike had established a student prize in numerical analysis at University of Baths School of Mathematics in Davids memory.

While Mike regarded all non-research duties of academical life as a distraction from the research he loved, this is true *a forteriori* with regard to obligations outside the department, be it college, university, research council or international committees and bodies. These he simply avoided as much as possible. "I think if all mathematicians would do what they can do best and try to make valuable contributions then that would help the subject enormously." **(PJD)** – for Mike "what they can do best" self-consciously did not include committee work and he shied away from deciding for others how to do their work or which research areas should be supported.

The one exception was editorial work, in particular his enormous investment of time and effort as (in tandem with Bill Morton) the founding Editor-in-Chief of *IMA Journal of Numerical Analysis.* His attention to detail was exemplary (and time consuming), Associate Editors were kept on a fairly short leash and Mike's ethos of quality and relevance as the sole determinants for the acceptance of papers established IMAJNA in short order as one of world's leading numerical analysis journals.

Mike displayed an enduring commitment to free dissemination of his work and software. He felt that nobody deserves the ownership of what, after all, is a collective long-term enterprise of the entire mathematical community. Decades before this has been endorsed by governments and funding agencies, he made his scientific output, inclusive of technical reports and software, freely and easily accessible to all.

Mike had a highly personal attitude to gadgets and accoutrements of modernity. At Harwell he was the last to use a Facit hand calculator and in Cambridge he was the last DAMTP user of punched cards: for two years the punching-card reader in the-then computing room was maintained for his exclusive use. However, lest this implies hard-core traditionalism, he was also the first individual at DAMTP to acquire a workstation (characteristically, not for himself but for the use of all group members). This was the CAMNUM SUN3 workstation which became an object of DAMTP lore. After a while, workstations proliferated in the Silver Street building, CAMNUM moved to Mike's office – and there it stayed forever, surviving the move to the Centre for Mathematical Sciences at Wilberforce Road. Parts have fallen into disuse and were replaced, old workstations were cannibalised, but CAMNUM, by then a contraption of many origins rather than a SUN product, went on. And there, tapping laboriously with two fingers, to his last day he used the antique UNIX `vi` utility to key-in his programs and correspondence. For Mike's attitude had nothing to do with fashion, everything with functionality. Change required good argument, but once that argument has been made, he was all in favour. And, as always, he cared little for what others thought.

The story of Mike in Cambridge would be incomplete without mentioning his dedication to sport. Mike was physical but also *very* competitive and sport was for him an ideal pastime and hobby. His favourite sport, from his schooldays onwards, was hockey. In Eastbourne, Harwell and Peterhouse he played hockey (and played it well), while at Harwell he had passed many a night rally driving and navigating and numerous evenings playing bridge. In his early forties he moved from the first to the second hockey team in Cambridge, finally giving up in favour of golf. Together with Caroline, they joined the Gog Magog Golf Club, and soon it became a focus for much of Mike's weekend (and the occasional weekday) activity. He played golf very competitively with a highly unorthodox swing that brought him success and with great skill. The 'Gogs' was also at the centre of Mike and Caroline's social activities and they have made there many enduring friendships. Ultimately, he went on to become the club Captain in 2005.

4 Research in Cambridge

The two threads running through Mike Powell's research in Cambridge, like in Harwell, were optimization and approximation.

4.1 Optimization

Arriving in Cambridge in the fall of 1976 , Mike was already an internationally leading figure in optimization. Both his theoretical and algorithmic contributions were widely cited among the researchers in the field. His main interests, at this time, were variable-metric methods, see [19], nonlinear conjugate gradients [20] augmented Lagrangian methods [21].

It is this line of research that he actively pursued after his arrival in Cambridge, as he was investigating another forward step in methods for solving nonlinearly constrained problems, based on sequential quadratic programming. While not the originator of the idea (Han and Biggs have had already made some inroads), Mike provided, in his seminal papers [22, 23], a clarification and implementation as well as a convergence theory for the method which very quickly made it mainstream research worldwide. Following his words, his proposal is "a variable metric method for constrained optimization", in the sense that it approximates Newton's method applied to the first-order optimality conditions of the constrained problem in way similar to that used by variable metric for unconstrained problems to approximate Newton's method on the same conditions for the unconstrained problem. The main idea is that the first-order conditions for the constrained problem can be viewed as the optimality conditions of a specific quadratic programming problem. This problem is defined, at a given iterate of the method, as that of minimizing a quadratic approximation of the Lagrangian function subject the locally linearized constraints. The remarkable feature of the method is that it provides not only a new approximation of the problem's solution, but also a new approximation of the associated Lagrange multipliers. Moreover, since it mimics variable metric methods for the unconstrained case, it also enjoys a similarly fast local convergence rate.

It is difficult to overestimate the influence, both in the short and long terms, of this important contribution. The class of methods introduced (later known as SQP, for Sequential Quadratic Programming) still forms the backbone of today's methodology for solving nonlinearly constrained optimization problems, and the ideas developed in [22, 23] have resonated in the international optimization community for very many years, giving birth to countless variants, reinterpretations and computer implementations.

It is interesting to note that Mike's papers on SQP emphasized the imperative to construct convex approximations of the Lagrangian function by insisting on choosing a positive-definite matrix to define the underlying quadratic optimization subproblem. As for the unconstrained case, this in turn requires that the change in first-order derivatives can be interpreted (along the step taken) as a change in a convex function. Unfortunately, this is not always possible in the presence of constraints, and a special technique (sometimes called *partial updating*) was designed by Mike to enforce the positive-definite nature of the updated approximation to the Lagrangian's Hessian matrix. As the name 'partial updating' suggests, this implies that it may be impossible to fully incorporate the new curvature information at a given iteration. This small blemish was clearly recognized by Mike, but his faith in the advantages of maintaining convex approximations was unabated, even in later years when SQP methods using true second derivatives were introduced. Moreover, developing robust SQP codes implied efficiently solving quadratic programs, a subject he embraced as a challenging necessity [33, 35].

Mike's interests for problems involving many variables, already manifest in his papers about nonlinear conjugate gradients [21] and later [38], led him to suggest this research topic to a young PhD student, Philippe Toint, who arrived in Cambridge from Belgium on Royal Society's funding in January 1977, after waiting a year during Mike's move from Harwell to Cambridge. The idea was to extend the variable metric ideas to structured problems with structurally sparse Hessian matrices. His interactions with that student were constant, at a time where Mike, despite having got his degree in Peterhouse, was not a member of any college. While the subject of sparse quasi-Newton methods was eventually developed mostly by Toint, it also involved Mike in three interesting contributions [24, 26, 31], the first of which is an extension to symmetric matrices of his famous paper with Curtis and Reid on estimating *sparse Jacobians* [15].

From then on, Mike's research in optimization focused mostly on two main themes. The first was to establish sound convergence theory for a variety of methods of interest among researchers in optimization as well as users. This includes global convergence (in the sense of convergence of a method's iterates to a first-order critical point from any starting point, not to be confused with the convergence to a global minimizer) for the trust-region method (which he pioneered in [10]) in a series of papers [36, 53, 68, 73, 79, 81, 83] scattered from 1984 to his very last paper in 2015. Some of these papers were in collaboration with his then-PhD student Ya-xiang Yuan.

Yuan arrived in Cambridge in 1983 and in 1988, having completed his doctoral studies under Mike's supervision, followed by Rutherford Research Fellowship at

Fitzwilliam College, returned to China. He became there arguably the most influential expert in optimization, a founder of the Chinese school in optimization. In 2011 he was elected to the Chinese Academy of Sciences. He was often visited in Beijing by Mike.

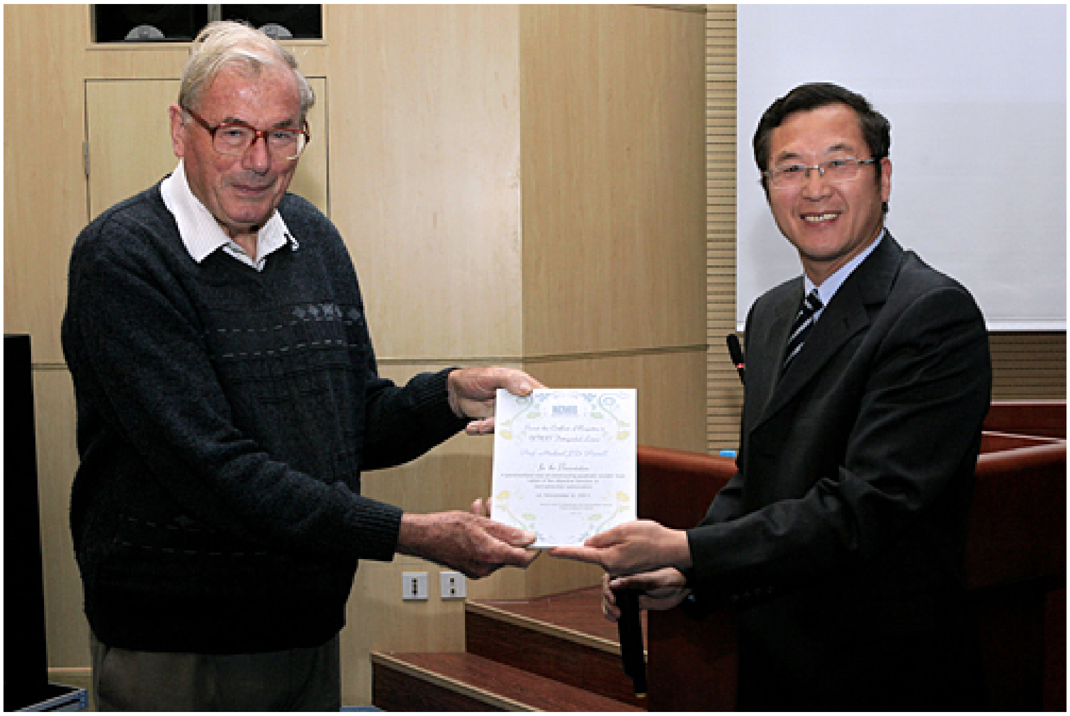

Figure 2: Mike and Ya-Xiang Yuan – the latter no longer a research student but an Academician of Chinese Academy of Sciences, in November 2011. The occasion was a Distinguished Lecture on "A parsimonious way of constructing quadratic models from values of the objective function in derivative-free optimization" at at the National Center for Mathematics and Interdisciplinary Sciences, Beijing.

Further Mike's papers in this area are dedicated to the analysis of local rates of convergence of variable metric matrices [36] and iterates [39, 71]. The penultimate of these papers is quite well known because it displays Mike's remarkable talent for producing intriguing counter-examples to sometimes widely held beliefs (another example of this skill can be found in [28]).

The second main theme, and probably that closest to his heart for many years, was the development of robust and efficient software for solving optimization problems without using derivatives of the involved (typically user-supplied and possibly expensive) nonlinear functions. Such problems occur very frequently in practice (often under the name of "black box optimization"), and Mike's many contacts with practitioners had convinced him of the high practical value and relevance of the topic. While he had already considered the topic in his Harwell days [2, 4], Mike's new idea was to build a model of the function to be minimized, the work on the model being cheaper and not involving expensive evaluations of the problem's true functions. He first attempted to build such models using radial basis functions (possibly sparking his interest in this field in which he later made substantial contributions, see Subsection 4.2), but this did not bring, in the optimization context, the results he had hoped for. He then turned his attention to models constructed by multivariate polynomial interpolation, generating a long stream of famous papers and conference talks on various algorithms for unconstrained, bound-constrained and linearly constrained optimization [42, 50, 64, 69, 72, 73, 74, 75, 76, 78, 81, 82]. A recurring theme of many of these papers is a continuously improving view of how to construct second-order information in objective function models. Giving full attention to every detail of the evolving algorithms and tirelessly improving the associated (freely available) codes, Mike became the most prominent advocate of the topic and one of its key actors. Maybe as a side effect of this remarkably strong focus, his involvement (and interest) in more theoretical questions in optimization slowly decreased over the years, in favour of a staunch defence of the more practical aspects of the algorithms and their computer implementations.

Of course there were distractions from this main path of research. One of the most noticeable came when Karmarkar presented his famous algorithm for linear programming in the International Symposium on Mathematical Programming in Boston, July 1985, creating the entire field of interior-point methods. Mike could not resist giving some thought to the subject, in particular showing in a famous example (once more) that the complexity of Karmarkar's methods was not always as good as expected [51, 59, 61]. He also contributed to the nascent field of interior-point method (more specifically, log-barrier methods) in [63, 66], as well as to various other optimization topics, such as linear algebra questions of interest in the optimization [30, 42, 50] exact penalty methods [40] and nonlinear equations [37].

This overview of Mike's activity as an optimization expert during his years in Cambridge would be incomplete if one forgets his impressive and continuous pedagogical effort. Although, quite surprisingly and due to the vagaries of Cambridge Mathematical Tripos, he never gave a course on optimization in Cambridge itself, he wrote, for the general international community a very large number of review and survey papers on various aspects of the field [12, 25, 32, 34, 43, 44, 46, 49, 57, 69, 78, 80].

These have helped and shaped many generations of young researchers, and remain, even after many years, truly informative and remarkably insightful.

4.2 Approximation

We now outline the vast contributions of Mike's to approximation theory both in Harwell and Cambridge. Mike has worked most influentially on approximation with uni- and multivariate polynomial splines, very profound and ground-breaking on radial basis functions (themes included convergence, fast algorithms for their computation or evaluation, Krylov-space methods, thin-plate splines), rational functions, ℓ_1-approximation, best Chebyshev approximations and convex or monotone approximations. As was typical for him – we have remarked that before – Mike's interests particularly extended to convergence questions, efficient and accurate numerical computations and implementations.

Mike has written a first-class textbook on approximation theory (*Approximation Theory and Methods*, Cambridge University Press, [27]), based on his Cambridge Mathematical Tripos lectures.

The book includes new results (or new, highly accurate, streamlined proofs of previous results), very well thought-out problems (we always remember his emphasis on considering and solving *problems,* this interest having started with him even as a boy), and it pays attention to computational considerations in Mike's typically precise and clear manner.

It covers mostly univariate approximation theory, and within that characterisation theorems for best approximation, polynomial interpolation and approximation, rational approximation, splines, interpolation, error estimates and convergence results. It is written in Mike's typical most readable style, more explanations and less formulæ to make for a particularly fluent and clear text. Speaking about clarity and style, many of us will always remember not only the quality of Mike's writing but also his remarkable talks using coloured hand-written slides in his typical long-hand writing with carefully and systematically colour-coded texts and formulæ.

Some of the early Mike's research contributions in univariate approximation theory covers best uniform approximation and the so-called exchange algorithm. A few years later, rational Chebyshev approximations are considered using the ordinary differential correction (ODC) algorithm. Another use of rational approximations is studied by Mike with Arieh Iserles and it regards the approximation theory underlying A-stable methods for numerical ordinary differential equations. There, a rational approximation of the same (polynomial) form is sought as the one mentioned earlier, approximation the exponential function. Such approximation is called 'maximal' if its error function has $m+n+1$ zeros with non-negative real part. If its error is $O(x^{m+n+1-k})$, the approximation is by definition in the set $E_{n,m,k}$. The two authors establish in [29] that a maximal approximation is A-acceptable if and only if it is either in $E_{n,n,k}$ for k at most two, or in $E_{n,n-2,0}$, or in $E_{n,n-1,k}$ for k at most one.

A further contribution to approximation theory in the ℓ_1 norm concerns the discrete test choice for functions from a space A of continuous approximations. Here, the conditions for the solution of the equivalent linear programming problem (referring to the previous subsection, again a nice link between approximation theory and optimization) are expressed in terms of the data on a finite set [21].

Mike wrote another article simultaneously related to optimization and approximation theory with Ya-Xiang Yuan in 1984 [37] that concerns ℓ_p-approximations for p either one or infinity. Here solutions are sought by iterative methods providing superlinear convergence of overdetermined nonlinear functional equations.

The trust region algorithms (see Subsection 4.1 here) incorporate a step length restriction by radii and decide on increasing or decreasing the radius by comparing the actual reduction of the objective function with the predicted reduction. Using the ratio of these as a criterion for changing the radii – with the goal of achieving global convergence – is a method introduced by Mike. Moreover, see also the previous subsection, a trust region scheme modifies the Hessian in the iteration step by adding a multiple of the identity matrix, giving a bias towards the steepest descent direction.

With co-author research student Ioannis Demetriou in the articles [54, 55], least-squares smoothing of univariate data was carried out to achieve piecewise monotonicity and to minimize the sum of squares of the required changes to those data providing non-negative second divided differences. A dynamic programming procedure is given for minimising the global ℓ_2-error. The principal advance in this work is that a recursive scheme has been found to reduce efficiently the number of data to be considered when finding the optimal knots which are integers.

Much of Mike's important work within approximation theory was concerned with splines: univariate and then mostly bivariate, radial function spline-like functions etc. In [6] for example, he considered norms of univariate polynomial spline operators and the deterioration of the Lagrange functions' localisation along the real line when their polynomial degrees grow. In [7], again localisations of spline approximations were analysed, but now in the context of weighted least-squares approximations.

The localisation can even be improved when the weighted least-squares norm of the error is augmented by the discrete ℓ_2-norm of the coefficients, again weighted suitably. Divided differences – these being incidentally one of the approaches taken often from Mike's toolbox – of the approximation are taken into account too.

In [8], a weighted sum of squares of the discrete error function of an approximation plus a smoothing term is considered. It is to be minimized again by a cubic spline, except that now it has free knots. Splines of arbitrary degree and using a number of knots are analysed, for approximations with respect to standard norms without weights in [5], where they should minimize the continuous Euclidean error between the approximation and a given bounded square-integrable approximated f, defined over an interval. The error functional is minimized – here of course we see Mike's expertise in optimization again – by a minimization method from [1].

So-called optimal interpolation is the subject-matter of a paper [18], where a minimal $c(x)$ for all x is sought such that for a given non-negative integer k and for a function f which is not a polynomial of degree less than k the pointwise error is at most $c(x)\|f^{(k)}\|_\infty$, the approximation being required as a linear combination of the values of f at the data points of which we have at least k.

We next turn our attention to multivariable approximations. In [19] a review of bivariate approximation tools is given, e.g. tensor-product methods derived from univariate schemes, considering interpolation and other linear conditions, i.e. least-squares. Both equally-spaced and scattered data are addressed. Piecewise polynomials of degree one and higher are discussed, as well as Shepard's method, moving least-squares and natural neighbourhood interpolations.

In a joint work [20] with Malcolm Sabin, Mike sought a globally continuous piecewise quadratic two-dimensional spline approximation, given function values and first partial derivatives at each vertex of a specified triangulation of a domain. This ended up in the celebrated Powell–Sabin split, where each such triangle is divided into either six or twelve sub-triangles with interior continuous differentiability conditions. The latter are exactly suitable (nine conditions, the right number of conditions for the degrees of freedom) for the aforementioned six sub-triangles, otherwise normals on edges are required and they may be computed by linear interpolation. The edges of the interior triangles go from the midpoints of the 'big' edges to the circum-centre of the triangle. The triangles are always acute. As a result the midpoint in the 'big' triangle is the intersection of the normals at the midpoints of the big edges. Therefore this midpoint of the big triangle is in the plane spanned by the points between it and the edge's midpoint and that gives the required continuous differentiability. The Powell–Sabin split is used most frequently in finite elements and CAGD.

In [70], again in the shared space of optimization and approximation theory, Mike considered an optimal way of moving a sequence of points onto a curve in two-dimensional space. The idea is to identify the smoothest function that maps a two-dimensional picture, for example, on a square into another while preserving some fixed points.

The next subject-matter are multivariate approximations using radial basis functions. Although subjected to a great deal of research by others, Mike was at that time one of the main persons in that area beginning with the review paper [41] that already contained new proofs of important results. As Yuan Xu said in an article in the Journal of Approximation Theory there was "a flow of preprints [on radial basis functions] pouring out of Powell's group in Cambridge".

The paper [41] addresses properties of radial basis function interpolation matrices, in particular non-singularity of the interpolation problem at pairwise distinct points, where for example the 'radial basis function' ϕ could be the identity $\phi(r) = r$ or a multiquadrics function with a real parameter c or the Gauss-kernel $\phi(r) = \exp(-c^2r^2)$ with positive real parameter c. All those matrices are indeed nonsingular if there are at least two points used, therefore admitting unique existence of real coefficients λ_j such that the interpolant from the linear space spanned by the shifts of the radial basis functions satisfies interpolation conditions for any given right-hand sides. Usage of the radial symmetry here is relevant, as for instance in the Chebyshev- or ℓ_1-norm the interpolation matrices could be singular for some of those simple kernel functions in spite of the points being distinct. It is interesting to note that one of the reasons why Mike was interested in the radial basis function (multivariate) interpolation or more generally approximation from the linear spaces spanned by the kernel functions was mentioned much earlier in this article; it was that they might be used for local approximation required in optimization methods. Indeed, much later his PhD student, Hans Martin Gutmann, addressed such questions with success. We saw in the previous section that initially at that time Mike found out by numerical tests that the radial basis function approximations are not so useful for the local approximations required in optimization after all, but he then developed a great interest in their properties independently of that. Hans Martin's work started much later after Mike's initial interest, spurred by a conversation with Carl de Boor in the mid-80s, in radial basis functions.

Motivated by the remarkable non-singularity results of Micchelli and others for radial basis function interpolation in many variables, Mike summarised and improved several further results in [45]. This concerns *inter alia* the question whether polynomials are contained in the aforementioned linear spaces spanned by the (now infinitely many) shifts of the radial basis function kernels. In the article [48], the Lagrange functions for the interpolation problem with such gridded centres and the appropriate basis functions were computed for finite grids. Therefore a matrix is to be inverted to allow for the calculation of the coefficients of the cardinal functions. For that, a Gauss–Seidel-type iteration is applied to a preconditioned matrix replacing the interpolation matrix, the centres x_i coming from the cardinal grid and the preconditioned kernel being linear combinations of shifts of the initial homogeneous radial basis function.

There is even a link again in this work with an optimization program Mike wrote [47] – in line with many other important links between different aspects of Mike's research – because the sums of moduli of the off-diagonal elements of the mentioned matrix was minimized in order to improve convergence of the method (where Gauss–Seidel iteration is applied) subject to a normalisation condition. Furthermore, there are coefficient constraints which give the needed algebraic decay of the ψs. The first step from gridded data to scattered data was again due to Mike [52], an article whose careful analysis of multiquadrics kernels' translates led to quasi-interpolants in one dimension when centres are scattered. Focussing further on interpolation with multiquadrics, Mike studied in [56] the uniform convergence of approximation and its rate when m equally spaced centres are used solely on the univariate unit interval and the multiquadrics constant c is the spacing between the centres.

In [58], one of Mike's *opera magna* (as he called it himself), he summarises and explains many recent developments, including nonsingularity theorems for interpolation, polynomial reproduction and approximation order results for quasi-interpolation and Lagrange interpolation on cardinal grids for classes of radial basis functions, including all of the ones mentioned above as well as thin-plate splines, inverse (reciprocal) multiquadrics and several others. The localisation of Lagrange functions for cardinal interpolation is considered in great detail and several best possible improvements of known approximation order results are given. Much like his earlier review papers and his book, this work does not just summarise his and other authors' work, but offers simplifications, more clarity in the exposition and improvements of proofs and theorems, often to the best possible way, typical for Mike's search for true optimality.

Further beautiful connections between approximation and optimization techniques can be found in [67], where approximations in two dimensions and mapping to two dimensions as well are considered in the form of componentwise thin-plate splines. The goal is to find a mapping between two regions in the two-dimensional plane, where certain control points and control curves are mapped to prescribed positions. Projecting control points to points with the thin-plate splines method as such is not hard, but a curve must be discretised and it is not initially clear whether the discretisation is the same in the image region even though the curve retains its shape. Because thin-plate splines yield the interpolation of minimal second derivatives in the least-squares sense, there is already one optimising feature in that approach. In this article, Mike uses once more the universal algorithm from [47] to determine the optimal positions of the discrete points on the curve in the image. The idea is to minimize again the square of the semi-norm of the interpolant which consists of the sum of the square-integrals of its second partial derivatives but now with respect to the positions of the points on the discretised curve.

In [65], the most general (with respect to the choice of the domain of convergence) results with regard to the convergence of thin-plate splines are obtained. There are several prior articles about convergence of thin-plate spline interpolations to scattered data on domains in the two-dimensional plane, but these domains have always been required to have at least Lipschitz continuous boundaries. Mike succeeds in maximally generalising and proving convergence within *any* bounded domain. The speed of convergence shown is within a factor of $\log h$ (h being the largest minimum distance between interpolation points and any points in the domain), the same as the best of the earlier results that required additional conditions on the domains where the function being approximated is defined. On top of all this, he gets the best multiplicative constants for the error estimates for interpolation on a line or within a square or a triangle, i.e., when we measure the error of thin-plate spline interpolation between two, three or four data points, where in the latter case they form a square. The $\log h$ term is due to the fact that the point x where we measure the error need not to be in the convex hull of centres (though it does need to be in their h-neighbourhood, due to the definition of h).

One of Mike's latest works on radial basis functions considers the efficient solution of the thin-plate spline interpolation problem for a large volume of data [77].

A closely related problem is the efficient evaluation of a large number of given linear combination of translates, an application for that being the rendering on a computer screen. These two issues are related because the conditional positive definiteness of the interpolation matrix makes the conjugate gradients algorithm a suitable tool to solve the interpolation equations for the coefficients. And, of course, the conjugate gradient algorithm needs many function evaluations. One approach for evaluating the approximating functions uses truncated Laurent expansions [60, 62], of the thin-plate splines and collecting several terms that are shifted thin-plate spline kernels for data points far away from the argument into one expression in order to minimize the number of evaluations of the logarithm, a computationally expensive task.

After completing the articles mentioned in the last two paragraphs, Mike explicitly stated to one of the authors (Martin Buhmann) that he, from now on, continues to work on optimization only – another of the clear decisions that were so typical for Mike.

5 Epilogue

In 2001 Mike retired from his professorship at DAMTP. As a matter of fact, he retired two years before the obligatory age of 67, the main reason being to escape all the facets of academic life that often irked him – administration, teaching, ... – while maximising the aspects he loved: research and travel.

This indeed was a pattern of Mike's retirement. A few months each year, typically escaping the winter gloom, were spent at City University of Hong Kong, where Mike held a part-time research position. Mike and Caroline delighted in the many pleasures of Hong Kong, Mike hiking in the New Territories and Caroline studying Chinese painting, and both enjoying local food and scenery. They also travelled elsewhere for extended periods, from Christchurch (New Zealand), Minneapolis (Minnesota), Stellenbosch (South Africa) to Victoria (British Columbia) and Mike (often with Caroline) remained very active on the conference circuit, where (as evidenced in Fig. 3) he was universally feted as one of the pioneers and leaders of optimization and approximation theory.

Back in Cambridge Mike followed the usual routine, except that he was exempt from the university chores he so disliked. So, he was to be found either on the fairways of the Gog Magog Golf Club, or pottering in the garden but, most frequently, in office F204 in the department or at a desk in his living room, patiently working on his research or writing and polishing computer programs.

By this stage Mike's research concerns were fully formed. His main focus was on optimization, revisiting the themes of his former work on variable-matrix and trust region methods and improving them with added insight and experience. He also spend a great deal of time and effort writing a new generation of software for large-scale optimization which, as was his custom throughout his working life, he made freely available to all and sundry.

In Autumn 2014 Mike and Caroline went again for several weeks to Hong Kong and there it became obvious that something was badly wrong with Mike's health. They returned to Cambridge and Mike visibly lost weight and was unwell. Medical checks discovered cancer of the oesophagus which, by that stage, had spread and was inoperable.

Faced with the option of painful chemotherapy to prolong his life for few more months, Mike resolutely said "no". He lived his life on his own terms and he would go on his own terms, clear of mind, free of pain and free of mind-numbing drugs. He patiently put all his affairs in order, whether at home or in mathematics, lend finishing touches to software and completed mathematical papers.

The end came swiftly. When it was clear that it was a matter of days, two of

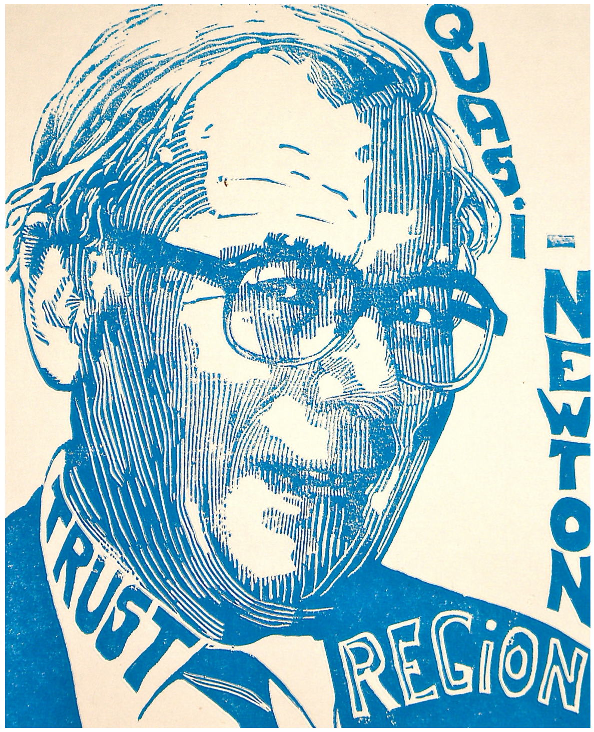

Figure 3: Mike post-retirement with his enduring research concerns (linocut by Henk van der Vorst, 2011).

us (Martin Buhmann and Arieh Iserles) arranged to see Mike on the weekend – this was an opportunity to say the last goodbye. Mike passed away peacefully on Sunday morning, our noontime visit was to console Caroline and their daughters.

Acknowledgements We wish to thank Caroline M. Powell for her help, comments and a permission to use Fig. 1 from her collection, Naomi Buhmann for a permission to use a photograph of Mike and Prof. Henk van der Vorst for a permission to use the linocut.

Honours

1979	Doctorate of Science (Sc.D.) of University of Cambridge
1983	Fellow of the Royal Society
2001	Foreign Member of the United States National Academy of Sciences
2007	Corresponding Fellow to the Australian Academy of Science

Awards

1982	George Dantzig Prize
1983	LMS Naylor Prize and Lectureship
1996	IMA Gold Medal
1999	LMS Senior Whitehead Prize
2007	IMA Catherine Richards Prize

In 2001 Mike Powell was awarded a Honorary Doctorate of Science by University of East Anglia.

Postscript

The second author, Roger Fletcher FRS, sadly passed away on 15th July 2016, soon after completing the draft of Section 2. Roger was a longstanding collaborator of Mike Powell and their work together at Harwell has laid the foundations to modern theory of optimization. He was also one of Mike's oldest friends.

References to other authors

(WCD) William C. Davidon, "Variance algorithm for minimization", *Comput. J.* **10** (1967/1968) 406–410.

(PJD) Philip J. Davis, "An interview with Michael J.D. Powell", Computer History Museum (2005).

(LNV) Luis Nunes Vicente, "An interview with M.J.D. Powell", *Bull. International Center for Mathematics* **14** (2003).

Cited M.J.D. Powell's papers

References

[1] "A rapidly convergent descent method for minimization" (with R. Fletcher), *Computer Journal* **6**, 163–168 (1963).

[2] "An efficient method for finding the minimum of a function of several variables without calculating derivatives", *Computer Journal* **7**, 155–162 (1964).

[3] "The volume internal to three intersecting hard spheres", *Molecular Physics* **7**, 591–592 (1964).

[4] "A method for minimizing a sum of squares of nonlinear functions without calculating derivatives", *Computer Journal* **7**, 303–307 (1965).

[5] "On best spline approximations", in *Numerische Mathematik, Differentialgleichungen, Approximationstheorie,* Birkhäuser Verlag (Basel), 317–339 (1968).

[6] "A comparison of spline approximations with classical interpolation methods", in *Proceedings of the IFIP 1968 Congress*, North-Holland (Amsterdam), 95–98 (1969).

[7] "The local dependence of least squares cubic splines", *SIAM J. Numer. Anal.* **6**, 398–413 (1969).

[8] "Curve fitting by splines in one variable", in *Numerical Approximation to Functions and Data* (J.G. Hayes, ed.), Athlone Press (London), 65–83 (1970).

[9] "A Fortran subroutine for unconstrained minimization requiring first derivatives of the objective function", AERE Report No. R.6469, Harwell Laboratory (1970).

[10] "A new algorithm for unconstrained optimization", in *Nonlinear Programming* (J.B. Rosen, O.L. Mangasarian and K. Ritter, eds), Academic Press (New York), 31–65 (1970).

[11] "On search directions for minimization algorithms", *Math. Programming* **4**, 193–201 (1973).

[12] "A Fortran subroutine for plotting a cubic spline function" (with S. Marlow), AERE Report No. R.7470, Harwell Laboratory (1973).

[13] "Numerical methods for constrained optimization", *IMA Bulletin* **9**, 326–327 (1973).

[14] "Unconstrained minimization algorithms without computation of derivatives", *Boll. della Unione Matematica Italiana* **9**, 60–69 (1974).

[15] "On the estimation of sparse Jacobian matrices" (with A.R. Curtis and J.K. Reid), *J. Inst. Maths. Applics.* **13**, 117–119 (1974).

[16] "On the convergence of cyclic Jacobi methods" (with K.W. Brodlie), *J. Inst. Maths. Applics.* **15**, 279–287 (1975).

[17] "Some global convergence properties of a variable metric algorithm for minimization without exact line searches", in *Nonlinear Programming SIAM–AMS Proceedings, Vol. IX* (R.W. Cottle and C.E. Lemke, eds), American Mathematical Society (Providence), 53–72 (1976).

[18] "Optimal interpolation" (with P.W. Gaffney), in *Numerical Analysis Dundee 1975*, Lecture Notes in Mathematics No. 506 (G.A. Watson, ed.), Springer-Verlag (Berlin), 90–99 (1976).

[19] "Numerical methods for fitting functions of two variables", in *The State of the Art in Numerical Analysis* (D.A.H. Jacobs, ed.), Academic Press (London), 563–604 (1977).

[20] "Piecewise quadratic approximations in triangles" (with M.A. Sabin), *ACM Trans. Math. Software* **3**, 316–325 (1977).

[21] "Algorithms for nonlinear constraints that use Lagrangian functions", *Math. Programming* **14**, 224–248 (1978).

[22] "The convergence of variable metric methods for nonlinearly constrained optimization calculations", in *Nonlinear Programming III* (O.L. Mangasarian, R.R. Meyer and S.M. Robinson, eds), Academic Press (New York), 27–63 (1978).

[23] "Variable metric methods for constrained optimization", in *Computing Methods in Applied Sciences and Engineering 1977 (1)*, Lecture Notes in Mathematics No. 704 (R. Glowinski and J.L. Lions, eds), Springer-Verlag (Berlin), 62–72 (1979).

[24] "On the estimation of sparse Hessian matrices" (with Ph.L. Toint), *SIAM J. Numer. Anal.* **16**, 1060–1074 (1979).

[25] "Optimization algorithms in 1979", in *Optimization Techniques*, Lecture Notes in Control and Information Sciences No. 22, (K. Iracki, K. Malanowski and S. Walukiewicz, eds), Springer-Verlag (Berlin), 83–98 (1980).

[26] "A note on quasi-Newton formulae for sparse second derivative matrices", *Math. Programming* **20**, 144–151 (1981).

[27] *Approximation Theory and Methods*, Cambridge University Press (Cambridge) (1981).

[28] "An example of cycling in a feasible point algorithm", *Math. Programming* **20**, 353–357 (1981).

[29] "On the A-acceptability of rational approximations to the exponential function" (with A. Iserles), *IMA J. Numer. Anal.* **1**, 241–251 (1981).

[30] "An upper triangular matrix method for quadratic programming", in *Nonlinear Programming* (O.L. Mangasarian, R.R. Meyer and S.M. Robinson, eds), Academic Press (New York), 1–24 (1981).

[31] "The Shanno–Toint procedure for updating sparse symmetric matrices" (with Ph.L. Toint), *IMA J. Numer. Anal.* **1**, 403–413 (1981).

[32] "Algorithms for constrained and unconstrained optimization calculations", in *Current Developments in the Interface: Economics, Econometrics, Mathematics* (M. Hazewinkel and H.G. Rinnooy Kan, eds), Reidel (Dordrecht), 293–310 (1982).

[33] "VMCWD: A Fortran subroutine for constrained optimization", Report No. DAMTP 1982/NA4, University of Cambridge (1982).

[34] "Variable metric methods for constrained optimization", in *Mathematical Programming: The State of the Art, Bonn 1982* (A. Bachem, M. Grötschel and B. Korte, eds), Springer-Verlag (Berlin), 288–311 (1983).

[35] "ZQPCVX: A Fortran subroutine for convex quadratic programming", Report No. DAMTP 1983/NA17, University of Cambridge (1983).

[36] "On the global convergence of trust region algorithms for unconstrained minimization", *Math. Programming* **29**, 297–303 (1984).

[37] "Conditions for superlinear convergence in and solutions of overdetermined nonlinear equations" (with Y. Yuan), *IMA J. Numer. Anal.* **4**, 241–251 (1984).

[38] "Nonconvex minimization calculations and the conjugate gradient method", in *Numerical Analysis Proceedings, Dundee 1983*, Lecture Notes in Mathematics No. 1066 (G.A. Watson, ed.), Springer-Verlag (Berlin), 122–141 (1984).

[39] "The performance of two subroutines for constrained optimization on some difficult test problems", in *Numerical Optimization 1984* (P.T. Boggs, R.H. Byrd and R.B. Schnabel, eds), SIAM Publications (Philadelphia), 160–177 (1985).

[40] "A recursive quadratic programming algorithm that uses differentiable exact penalty functions" (with Y. Yuan), *Math. Programming* **35**, 265–278 (1986).

[41] "Radial basis functions for multivariable interpolation: a review", in *Algorithms for Approximation* (J.C. Mason and M.G. Cox, eds), Oxford University Press (Oxford), 143–167 (1987).

[42] "Updating conjugate directions by the BFGS formula", *Math. Programming* **38**, 29–46 (1987).

[43] *The State of the Art in Numerical Analysis* (editor with A. Iserles), Oxford University Press (Oxford) (1987).

[44] "Methods for nonlinear constraints in optimization calculations", in *The State of the Art in Numerical Analysis* (A. Iserles and M.J.D. Powell, eds), Oxford University Press (Oxford), 325–357 (1987).

[45] "Radial basis function approximation to polynomials", in *Numerical Analysis 1987* (D.F. Griffiths and G.A. Watson, eds), Longman Scientific and Technical (Burnt Mill), 223–241 (1988).

[46] "A review of algorithms for nonlinear equations and unconstrained optimization", in *ICIAM '87: Proceedings of the First International Conference on Industrial and Applied Mathematics* (J. McKenna and R. Temam, eds), SIAM Publications (Philadelphia), 220–232 (1988).

[47] "TOLMIN: A Fortran package for linearly constrained optimization calculations", Report No. DAMTP 1989/NA2, University of Cambridge (1989).

[48] "Radial basis function interpolation on an infinite regular grid" (with M.D. Buhmann), in *Algorithms for Approximation II* (J.C. Mason and M.G. Cox, eds), Chapman and Hall (London), 146–169 (1990).

[49] "Algorithms for linearly constrained optimization calculations", in *Proceedings of the Third IMSL User Group Europe Conference* (M. Vaccari, ed.), IMSL (Houston), paper A-1, 1–17 (1990).

[50] "The updating of matrices of conjugate directions in optimization algorithm", in *Numerical Analysis 1989* (D.F. Griffiths and G.A. Watson, eds), Longman Scientific and Technical (Burnt Mill), 193–205 (1990).

[51] "Karmarkar's algorithm: a view from nonlinear programming", *IMA Bulletin* **26**, 165–181 (1990).

[52] "Univariate multiquadric approximation: reproduction of linear polynomials", in *Multivariate Approximation and Interpolation* (W. Haussmann and K. Jetter, eds), Birkhäuser Verlag (Basel), 227–240 (1990).

[53] "A trust region algorithm for equality constrained optimization" (with Y. Yuan), *Math. Programming A* **49**, 189–211 (1991).

[54] "Least squares smoothing of univariate data to achieve piecewise monotonicity" (with I.C. Demetriou), *IMA J. Numer. Anal.* **11**, 411–432 (1991).

[55] "The minimum sum of squares change to univariate data that gives convexity" (with I.C. Demetriou), *IMA J. Numer. Anal.* **11**, 433–448 (1991).

[56] "Univariate multiquadric interpolation: some recent results", in *Curves and Surfaces* (P.J. Laurent, A. Le Méhauté and L.L. Schumaker, eds), Academic Press (New York), 371–382 (1991).

[57] "A view of nonlinear optimization", in *History of Mathematical Programming: A Collection of Personal Reminiscences* (J.K. Lenstra, A.H.G. Rinnooy Kan and A. Schrijver, eds), North-Holland (Amsterdam), 119–125 (1991).

[58] "The theory of radial basis function approximation in 1990", in *Advances in Numerical Analysis II: Wavelets, Subdivision Algorithms and Radial Basis Functions* (W. Light, ed.), Oxford University Press (Oxford), 105–210 (1992).

[59] "The complexity of Karmarkar's algorithm for linear programming", in *Numerical Analysis, 1991* (D.F. Griffiths and G.A. Watson, eds), Longman Scientific and Technical (Burnt Mill), 142–163 (1992).

[60] "Tabulation of thin plate splines on a very fine two–dimensional grid", in —em Numerical Methods in Approximation Theory IX (D. Braess and L. Schumaker, eds), Birkhäuser Verlag (Basel), 221–244 (1992).

[61] "On the number of iterations of Karmarkar's algorithm for linear programming", *Math. Programming* **62**, 153–197 (1993).

[62] "Truncated Laurent expansions for the last evaluation of thin plate splines", *Numerical Algorithms* **5**, 99–120 (1993).

[63] "Log barrier methods for semi-infinite programming calculations", in *Advances on Computer Mathematics and its Applications* (E.A. Lipitakis, ed.), World Scientific (Singapore), 1–21 (1993).

[64] "A direct search optimization method that models the objective and constraint functions by linear interpolation", in *Advances in Optimization and Numerical Analysis* (S. Gomez and J.-P. Hennart, eds), Kluwer Academic (Dordrecht), 51–67 (1994).

[65] "The uniform convergence of thin plate spline interpolation in two dimensions", *Numerische Mathematik* **68**, 107–128 (1994).

[66] "Some convergence properties of the modified log barrier method for linear programming", *SIAM J. Optimization* **5**, 695–739 (1995).

[67] "A thin plate spline method for mapping curves into curves in two dimensions", in *Computational Techniques and Applications: CTAC95* (R.L. May and A.K. Easton, eds), World Scientific (Singapore), 43–57 (1996).

[68] "Trust region calculations revisited", in *Numerical Analysis 1997* (Dundee), Longman (Harlow), 193–211 (1998).

[69] "Direct search algorithms for optimization calculations", *Acta Numerica* **7**, Cambridge University Press (Cambridge), 287–336 (1998).

[70] "An optimal way of moving a sequence of points onto a curve in two dimensions", in *Computational Optimization – A Tribute to Olvi Mangasarian, Part II*, *Computational Optim. Appl. 13*, 163–185 (1999).

[71] "On the convergence of the DFP algorithm for unconstrained optimization when there are only two variables", in *Studies in Algorithmic Optimization*, *Math. Program.* **87**, 281–301 (2000).

[72] "UOBYQA: unconstrained optimization by quadratic approximation", in *Math. Program.* **92**, 555–582 (2002).

[73] "On trust region methods for unconstrained minimization without derivatives. New trends in optimization and computational algorithms", *Math. Program.* **97**, 605–623 (2003).

[74] "On the use of quadratic models in unconstrained minimization without derivatives. The First International Conference on Optimization Methods and Software", *Optim. Methods Softw.* **19**, 399–411 (2004).

[75] "Least Frobenius norm updating of quadratic models that satisfy interpolation conditions", *Math. Program.* **100**, 183–215 (2004).

[76] "The NEWUOA software for unconstrained optimization without derivatives", in *Large-Scale Nonlinear Optimization*, Springer (New York), 255–297 (2006).

[77] "Fast evaluation of polyharmonic splines in three dimensions" (with R.K. Beatson and A.M. Tan), in *IMA J. Numer. Anal.* **27**, 427–450 (2007).

[78] "A view of algorithms for optimization without derivatives", in *Math. Today* **43**, 170–174 (2007).

[79] "On the convergence of a wide range of trust region methods for unconstrained optimization", *IMA J. Numer. Anal.* **30**, 289–301 (2010).

[80] "On nonlinear optimization since 1959", in *The Birth of Numerical Analysis*, World Sci. Publ. (Hackensack), 141–160 (2010).

[81] "On the convergence of trust region algorithms for unconstrained minimization without derivatives", *Comput. Optim. Appl.* **53**, 527–555 (2012).

[82] "Beyond symmetric Broyden for updating quadratic models in minimization without derivatives", *Math. Program.* **138**, 475–500 (2013).

[83] "On fast trust region methods for quadratic models with linear constraints", *Math. Program. Comput.* **7**, 237–267 (2015).

Martin D. Buhmann is Professor of Numerical Analysis at the Justus-Liebig University of Giessen. His main area of interest is approximation theory and he is a leading expert in multivariate approximation, in particular the theory and applications of radial basis functions.

Roger Fletcher FRS was a Professor of Mathematics at University of Dundee and, together with Mike Powell, among the pioneers of optimization theory. His name is immortalised *inter alia* in the DFP (Davidon–Fletcher–Powell) and BFGS (Broyden–Fletcher–Goldfarb–Shanno) algorithms, both in wide use worldwide.

Arieh Iserles, who recently retired from the Chair in Numerical Analysis of Differential Equations at University of Cambridge, is an expert in numerical ordinary and partial differential equations, geometric numerical integration, highly oscillatory phenomena and orthogonal polynomials. He is the Managing Editor of *Acta Numerica* and past chair of Society for Foundations of Computational Mathematics.

Philippe Toint is a Professor of Mathematics at the University of Namur and its past Vice-Rector for Research and IT Policy, as well as Honorary Professor at University of Edinburgh. He is widely known for his contributions to the theory and algorithms of optimization and is past chairman of Mathematical Optimization Society.