TOPICS IN CONTEMPORARY APPLIED MATHEMATICS

For a round table — Ferrara, 5 March 2001

Big math or soft math?

The main problem in big math is that almost necessarily it has to be very specialised, and only rarely the big results have practical applications. Mathematicians are gratified if once in five years it is discovered that one of the hundred thousand theorems they produce has applications, which seems ridiculous in the eyes of an engineer or biologist. A recent overview of the big math of the 20th century has been given by Odifreddi, mostly from the standpoint of a pure mathematician, he too complaining about the explosion of the number of publications leading to a hyperproduction of low level papers, swelling only the power of journal publishers and bringing scientists in the sphere of control of politics and commerce.

The American Mathematical Society is a well organized and clever organization, but still rather conservative in its main exponents, as becomes apparent from the interview with Hyman Bass expressing rather prudent and traditional views on novelties: "From my point of view, I don't see the need for major new initiatives by the AMS, but rather the need for sustaining the progress that's been achieved ... We need to be attentive to the ways in which the discipline has changed, to the presence of technology, to appropriate ways of presenting mathematical ideas in the classroom, and to contemporary understanding of instruction and student learning."

One of the most momentous mathematical techniques in applications has been the fast Fourier transform algorithm introduced in 1965 by Cooley and Tukey; it is not a piece of big mathematics, but has changed the course of engineering in all realms of information technology. A similar role shall probably have other soft mathematical techniques as genetic algorithms, neural nets, cellular automata or fuzzy logic, which go on almost without the participation of mathematicians.

H. Bass: Presidential views - an interview. Notices AMS 48 (2001), 312-315.
P. Odifreddi: La matematica del Novecento. Einaudi 2000.

An interdisciplinary profession

"The theorems of mathematics only rarely are directly applied: in general only the definitions are really useful." (David Sharp, cited in Rota, p. 159, translated)

"I also believe that changing fields of work during one's life is rejuvenating. If one stays too much with the same subfields or the same narrow class of problems a sort of self-poisoning prevents acquisition of new points of view and one may become stale. Unfortunately, this is not uncommon in mathematical creativity." (Ulam, p. 290)

One of the psychological reasons for the different behaviour of mathematicians is that in other sciences (physics, medicine, genetics, engineering) exchange of information between scientists is essential; mathematicians usually develop methods and intellectual skills they often are not so inclined to reveal. It is also seldom in mathematics that large bodies of knowledge (say databases of differential equations or algebraic curves) are collected.

I think that today one of the most important tasks for mathematicians is to create model systems and concepts for the emerging new developments in genetics and medicine. The strength of the mathematician is that he is at the same time a specialist and a generalist: a specialist of precise reasoning, a generalist, because the formal and abstract laws of mathematics have universal validity. But in order that this strength becomes fertile, the mathematician must also aquire knowledges in the fields of applications; without that he cannot perceive, he cannot communicate, he cannot find the fitting models.

G. Rota: Pensieri discreti. Garzanti 1993.
S. Ulam: Adventures of a mathematician. Univ. of California UP 1991.

Contents

Lost fields

"As we mathematicians concentrated on our own discipline, connections with other fields diminished. Statistics separated from mathematics; in many universities, the two fields are now distinct departments. Computer science was largely shunned by mathematicians in the 1970s; it too developed outside of mathematics departments. Academic bean-counting often ensures that once disciplines are divided, walls go up. Why hire an applied mathematician if the physics or computer science departments can be convinced to do so instead? ... In splitting, we have lost a real resource for mathematics ... Without question, we cannot be good mathematicians without an appreciation and understanding of abstraction. But the inward focus that developed in university mathematics departments has not always served mathematics well, and a broadening of the definition of what constitutes mathematical research, and what a mathematician is, is in order." (Susan Landau)

There are two main reasons for mathematicians losing fields:

1. They are often not interested in the applications and have not enough knowledge to appreciate achievements in other sciences.

2. Once there are no more new theorems to prove, mathematicians withdraw from the field instead of participating in the applications.

S. Landau: Something there is that doesn't love a wall. Notices AMS November 1995.

The cycle indicator

Let (X, G) be a finite transformation group and $|X| = n$, $z_1, \ldots, z_n$ indeterminates.

For every $g \in G$ the permutation of X induced by g is a product of disjoint cycles, say of b_1 cycles of length 1, b_2 cycles of length 2, ..., b_n cycles of length n. Clearly $b_1 + 2b_2 + \cdots + nb_n = n$. The monomial $\zeta_g := z_1^{b_1} z_2^{b_2} \cdots z_n^{b_n}$ is called the cycle indicator of the element g; it depends of course not only on g, but also on the operation.

The cycle indicator of a permutation of the form

$(\cdot)(\cdot)(\cdot\cdot)(\cdot\cdot)(\cdot\cdot)(\cdot\cdot\cdot)(\cdot\cdot\cdot)(\cdot\cdot\cdot\cdot)$

is therefore $z_1^2 z_2^3 z_3^2 z_4$.

The cycle indicator ζ von G is then, as an element of the power series ring $\mathbb{Q}[z_1, \ldots, z_n]$, the averaged sum of the cycle indicators of the elements of G:

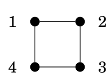
$$\zeta := \frac{1}{|G!} \sum_{g \in G} \zeta_g$$

Example 1: In the Klein four group, considered as a permutation group of $\{1, \ldots, 4\}$, the identity has the cycle indicator z_1^4, each one of the other three elements the cycle indicator z_2^2; der cycle indicator of the group is therefore

$$\frac{1}{4}(z_1^4 + 3z_2^2)$$

Example 2: The cycle indicator of S_3, considered as a permutation group of $\{1, 2, 3\}$:

$$\frac{1}{6}(z_1^3 + 3z_1 z_2 + 2z_3)$$

Example 3: 

The dihedral group D_2, considered as the group of symmetries of the square (labeled as in the figure), consists of the following 8 elements with the cycle indicators given on the right:

identity	z_1^4
(12)(34)	z_2^2
(14)(23)	z_2^2
(13)	$z_1^2 z_2$
(24)	$z_1^2 z_2$
(1234)	z_4
(13)(24)	z_2^2
(1432)	z_4

The cycle indicator of the permutation group is therefore

$$\frac{1}{8}(z_1^4 + 2z_1^2 z_2 + 3z_2^2 + 2z_4)$$

Example 4: We consider S_n as a permutation group of $\{1, \ldots, n\}$ and denote with ζ_n its cycle indicator. For $b_1 + \cdots + b_n = n$ then there exist exactly $\frac{n!}{1^{b_1} b_1! 2^{b_2} b_2! \cdots n^{b_n} b_n!}$ elements with the cycle indicator $z_1^{b_1} \cdots z_n^{b_n}$. Since S_n has $n!$ elements, we obtain $\zeta_n = \sum_{b_1 + 2b_2 + \cdots + nb_n = n} \frac{z_1^{b_1} \cdots z_n^{b_n}}{1^{b_1} b_1! 2^{b_2} b_2! \cdots n^{b_n} b_n!}$.

With a new indeterminate t we collect this in

$$\sum_{n=0}^{\infty} \zeta_n = e^{\sum_{k=1}^{\infty} \frac{z_k}{k} t^k}$$

The combinatorics of combinatorial chemistry

Combinatorial chemistry originated in solid phase peptide synthesis, which subsequently has been refined and generalized and is nowadays widely adopted for the generation of synthetic peptide combinatorial libraries for research and drug discovery, providing very large numbers of novel molecules. Both the planning and the description of such libraries invite to mathematical modelling by statistical and combinatorial techniques.

Going back to Cayley and Sylvester, algebraic-combinatorial description and counting of isomers has been founded by Redfield and some years later by Pólya and elaborated to considerable depth and completeness by Adalbert Kerber's school in Bayreuth.

These techniques rely on the Cauchy/Frobenius or Burnside lemma: Let X be a finite set and G a finite group acting on X. For $x \in X$ and $g \in G$ we define

$G_x := \{g \in G | gx = x\}$

$X_g := \{x \in X | gx = x\}$.

Then

$$|G||X/G| = \sum_{g \in G} |X_g| = \sum_{x \in X} |G_x|$$

The importance of group theory to combinatorics comes from the fact that any equivalence relation on a set X can be thought of as determinated by a *group* acting on X, e.g. the group of all permutations of X which leave each equivalence class invariant. The theorem of Pólya which we shall now explain is a powerful method for counting such structures.

Let C (the "set of colours") be another finite set. Each mapping $\varphi : X \longrightarrow C$ is considered as a colouring of X, C_X is then the set of all colourings. G, operating on X, operates also on C_X, if we define $(g\varphi)(x) := \varphi(g^{-1}x)$. The orbits $G\varphi$ of this operation are called *patterns* in this context.

A *weight* on the colours is a mapping $w : C \longrightarrow A$, where A is a $\mathbb{Q}$-algebra (essentially a ring in which division by integers $\neq 0$ is possible). For such a weight w and a colouring φ we define $w^*(\varphi) := \prod_{x \in X} w(\varphi(x))$. It is immediate that $w^*(g\varphi) = w^*(\varphi)$ for every $g \in G$, thus we can define $w^*(G\varphi) := w^*(\varphi)$.

For transparency we use, for a weight w, the following abbreviations:

$$[w] := \sum_{a \in A} w(a)$$

$$[w^2] := \sum_{a \in A} w^2(a)$$

$$\ldots$$

$$[w^*] := \sum_{\alpha \in C^X/G} w^*(\alpha)$$

Theorem (Pólya):

$$[w^*] = \zeta([w], [w^2], \ldots, [w^n])$$

where ζ is the cycle indicator of the permutation group (X, G) as introduced on the right. The book by Kerber contains many applications and generalizations.

A. Beck-Sickinger/P. Weber: Kombinatorische Methoden in Chemie und Biologie. Spektrum 1999.

K. Burgess/A. Liaw/N. Wang: Combinatorial technologies involving reiterative division/ coupling/recombination - statistical considerations. J. Med. Chem. 37 (1994), 2985-2987.

B. Déprez a.o.: Orthogonal combinatorial chemical libraries. J. Am. Chem. Soc. 117 (1995), 5405-5406.

K. Jacobs: Einführung in die Kombinatorik. De Gruyter 1983.

A. Kerber: Applied finite group actions. Springer 1999.

H. Maehr: Combinatorial chemistry in drug research from a new vantage point Bioorg. Med. Chem. 5 (1997), 473-491.

H. Maehr/R. Yang: Optimization of a leukotriene D4 antagonist by combinatorial chemistry in solution. Bioorg. Med. Chem. 5 (1997), 493-496.

E. Martin a.o.: Measuring diversity - experimental design of combinatorial libraries for drug discovery. J. Med. Chem. 38 (1995), 1431-1436.

N. Terrett: Combinatorial chemistry. Oxford UP 1998.

T. Wieland: Combinatorics of combinatorial chemistry. J. Math. Chem. 21/2 (1997), 141-157.

Combinatorics on words

Aldo De Luca, a pupil of Caianiello, has written as a young man an advanced treatise on neural nets and related topics in cybernetics. Since many years he is, together with Jean Berstel and Dominique Perrin, one of the leaders in the combinatorics of texts, a field of pure mathematics at the border between combinatorics and formal language theory with strong schools in Italy and France (and partly in Germany), but altogether unnoticed by the big math.

The ingenious techniques in some also very well written new papers (e.g. Carpi/De Luca) allow to reconstruct a word from its parts, a task very important in modern genetics.

Semigroup theory has always been helpful in the field, symbolic dynamics is emerging more recently as a rich collection of topological tools.

J. Berstel/D. Perrin: Theory of codes. Academic Press 1985.

A. Carpi/A. De Luca: Words and repeated factors. Sem. Lothar. 42 (1997).

C. Choffrut/J. Karhumäki: Combinatorics of words. Turku Centre CS TR 77 (1996).

A. De Luca: On the combinatorics of finite words. Theor. Comp. Sci. 218 (1999), 13-39.

A. De Luca/L. Ricciardi: Introduzione alla cibernetica. Franco Angeli 1971.

M. Lothaire (ed.): Combinatorics on words. Addison-Wesley 1983.

Drug discovery

An exciting field, where mathematicians could do much.

"A lead compound is only the beginning of the way to a drug. In a usually long iterative process one must optimize strength of efficacy, specificity and duration of efficiency, and minimize side-effects and toxicity ... The mathematicians in the last decades took over the principles of evolution to their toolbox. They embodied reproduction, mutation and crossover in their genetic algorithms. Who could ever admire how such an algorithm performs optimization tasks of a very complex type unerringly and in surprisingly short time, shall have no more doubts that also the evolution of the biological species run off in an analogous way." (Böhm a.o., pp. 147 and 231, translated)

H. Böhm/G. Klebe/H. Kubinyi: Wirkstoffdesign. Spektrum 1996.
A. Romani: Quando Linux incontra la chimica. Linux & C Giugno 2000, 17-21, Luglio 2000, 26-29.
L. Weber/S. Wallbaum/C. Broger/K. Gubernator: Optimization of the biological activiy of combinatorial compound libraries by a genetic algorithm. Angew. Chemie Int. Engl. Ed. 34 (1995), 2280-2282.

Order and topology in biochemistry

Let X be a finite topological space. Then each $x \in X$ has a minimal neighbourhood U_x and we may define a relation $\leq$ on X by $x \leq y :\iff U_x \subset U_y$. It turns out that $\leq$ is reflexive and transitive, i.e. a *quasiorder*.

Conversely let on a finite set X be given a quasiorder $\leq$. For $x \in X$ we put then $U_x := \{y \in X | y \leq x\}$. In this way we obtain the minimal neighbourhoods in a topology on X. The two constructions are one the inverse of the other, and a mapping between two finite topological spaces is continuous if and only if it is monotone with respect to the associated quasiorder. Therefore finite topological spaces and finite quasiorders are exactly the same thing. Since finite quasiorders are one of the most fundamental structures of combinatorial theory, finite topological spaces are not as trivial as one often thinks. A finite topological space is T_0 if and only if the associated quasiorder is antisymmetric, i.e. a partial order.

Finite topological spaces play a role in digital topology and some other fields of theoretical computer science, but are also often used by chemists. At the Institute of Theoretical Chemistry at the university of Wien, Peter Schuster and his collaborators are studying a variety of mathematical facets of molecular evolution theory, using for example finite topologies to obtain a description of the evolutionary dynamics of sets of secondary structures of RNA molecules or more general molecular quasispecies.

That the complete knowledge of the human genome is only the beginning of a long epoch of research becomes convincingly evident from the fact that one knows the exact genome of the worm *Caenorhabditis elegans*, a much studied organism much simpler than man but, according to Schuster, only 7 percent of its genes are genetically or biochemically understood. Therefore mathematical techniques, especially from those fields which are suited to supply principles of order and structure for understanding, comparison, prediction and information organization, are much needed in future functional genomic biochemistry and, for example, in an effective theoretical understanding of the mechanisms of emerging infectious diseases.

K. Baik: Towards a topological view of databases. Comp. Math. Appl. 18/5 (1989), 401-409.
J. Cupal/P. Schuster/P. Stadler: Topology in phenotype space. In R. Giegerich a.o. (ed.): Computer Science and Biology GCB '99, 9-15.
E. Domingo a.o.: Quasispecies structure and persistence of RNA viruses. Emerging Infectious Diseases 4/4 (1998).
W. Ebeling/R. Feistel: Physik der Selbstorganisation und Evolution. Akademie-Verlag 1986.
W. Hoppe a.o. (ed.): Biophysik. Springer 1982.
E. Khalimsky/R. Kopperman/P. Meyer: Computer graphics and connected topologies on finite ordered sets. Topology and Appl. 36 (1990), 1-17.
A. Rosenfeld: Digital topology. Am. Math. Monthly 86 (1979), 621-630.
P. Schuster: Molecular insights into life and evolution. Schrödinger lecture in Dublin, November 1998.
P. Schuster: Perspektiven einer computergestützten Molekularbiologie. Talk at the University of Wien, 2000.
P. Schuster: Molecular insights into evolution of phenotypes. Preprint Univ. Wien TBI pks-00-002 (2000).
H. Sharp: Quasi-orderings and topologies on finite sets. Proc. AMS 17 (1966), 1344-1349.
R. Stocsits/I. Hofacker/P. Stadler: Conserved secondary structures in hepatitis B virus RNA. In R. Giegerich a.o. (ed.): Computer Science and Biology GCB '99, 73-79.
R. Stong: Finite topological spaces. Trans. AMS 123 (1966), 325-348.

A most surprising theorem

Let A be a finite set. On the compact metrizable space $A^{\mathbb{Z}}$ consider the shift operator T defined by $(Tx)_n = x_{n+1}$. Symbolic dynamics, the study of this dynamical system (or, with a slight variation in generality and methods, topological dynamics or, when the measure-theoretic aspects are in the foreground, ergodic theory), is a beautiful rather old field of mathematics; it originated as a method in nonlinear ordinary differential equations, where it is still important. Less known are the implications in computer science. There are connections with Lindenmayer systems and other applications in the combinatorial theory of texts. Klaus Schmidt's theory of shifts in higher dimensions uses advanced techniques e.g. from algebraic topology and is surely a difficult part of pure mathematics, but it seems that it will have soon practical applications.

Here we state only a fact, which perhaps is one of the most surprising theorems in all of mathematics.

A subset X of $A^{\mathbb{Z}}$ is called *invariant*, if for $x \in X$ also $Tx \in X$. A closed (and therefore compact) invariant subset of $A^{\mathbb{Z}}$ is called a *symbolic dynamical system* and this is clearly the main subject of study in symbolic dynamics, as groups are for group theory, topological spaces for topology and so on.

We define now an apparently completely different concept, very natural in computer science. Denote, as usual, by A^* the set of words which can be formed by means of the alphabet A, and let F be a (not necessarily finite) subset of A^*. We think of F as a set of *forbidden words* and may consider the set D_F of all elements of $A^{\mathbb{Z}}$ which don't contain any of the forbidden words (this is important in physical data storage for example). Then D_F is closed and invariant, i.e. a symbolic dynamical system.

Conversely, if X is a symbolic dynamical system, there always exists (not uniquely defined) a subset $F \subset A^*$ such that $X = D_F$.

The proofs are rather easy, but the theorem says that symbolic dynamical systems (a genuine concept in pure mathematics) and sets of infinite texts described by sets of forbidden words (a genuine concept in computer science) are exactly the same thing.

W.H. Gottschalk/G.A. Hedlund: Topological dynamics. AMS 1955.
G.A. Hedlund: Endomorphisms and automorphisms of the shift dynamical system. Math. Systems Theory 3 (1969), 320-375.
D. Lind/B. Marcus: An introduction to symbolic dynamics and coding. Cambridge UP 1995.
K. Schmidt: Dynamic systems of algebraic origin. Birkhäuser 1995.

Formal concept analysis

Let X and M be finite sets and $\tau \subset X \times M$ be a relation. Then the triple (X, M, τ) is called an *incidence structure* or, by Wille, a *formal context*. This fundamental structure in abstract geometry can now be used for a mathematical theory of concepts which seems to have many applications (the book of Bartel elaborates in great detail applications to the classification of molecules in chemistry). For $A \subset X$ define
$A' := \{m \in M | a\tau m \text{ for all } a \in A\}$,
similarly for $B \subset M$ let
$B' := \{x \in X | a\tau m \text{ for all } m \in B\}$. It is easy to see that in this way one obtains a Galois connection between $\mathcal{P}(X)$ and $\mathcal{P}(M)$.

A *formal concept* in the context (X, M, τ) is a pair (A, B) with $A \subset X$, $B \subset M$ such that $A' = B$ and $B' = A$.

If (A, B) and (C, D) are formal concepts, we write $A \leq B : \Longleftrightarrow A \subset C$ (which is equivalent to $D \subset B$). And now one shows that the set of formal concepts with this partial order is a complete lattice. The study of this lattice has been called *formal concept analysis*; it can be generalized (e.g. to many valued and fuzzy contexts) and has some intriguing applications (the booklet of Becker is a reformulation of known formal equivalences in algebraic geometry).

For example, perhaps one could try to classify the topologies of neural nets by this theory.

H. Bartel: Mathematische Methoden in der Chemie. Spektrum 1996.
T. Becker: Formal concepts analysis and algebraic geometry. Shaker 1999.
B. Ganter/R. Wille: Formal concept analysis. Springer 1999.
G. Stumme/R. Wille (ed.): Begriffliche Wissensverarbeitung. Springer 1999.

Random number generators

Sequences or tables of random numbers (or random vectors) are used in many problems of statistics, simulation, numerical integration and cryptography. There is still a great need for new and versatile random number generators and reliable techniques for the assessment of their properties, requiring usually very sophisticated methods from number theory and harmonic analysis.

Let $(x_1, \ldots, x_N)$ be a finite sequence of real numbers and denote by $\{a\}$ the fractional part of a real number a. We define the **-discrepancy* of the sequence by

$$D_N^*(x_1, \ldots, x_N) := \sup_{0 < v \leq 1} |\frac{1}{N} \sum_{n=1}^{N} (0 \leq \{x_n\} < v) - v|$$

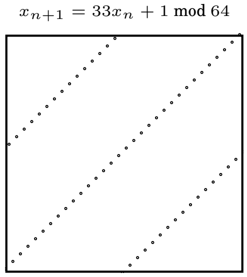
Then for every function $f : [0,1] \longrightarrow \mathbb{R}$ with finite total variation $V(f)$ the following inequality holds:

$$|\frac{1}{N} \sum_{n=1}^{N} f(x_n) - \int_0^1 f(u)du| \leq V(f) D^*(x_1, \ldots, x_N)$$

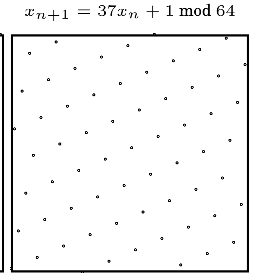
Once known the total variation, the error term depends therefore only on the *-discrepancy of the sequence. But finding low-discrepancy sequences and calculating their discrepancy is an all but trivial task.

That apparently simple situations can require attentive study is shown by two pictures which show the points (x_n, x_{n+1}) for the widely used linear congruential generator $x_{n+1} = ax_n + b \bmod m$. Both generators have maximal period, but the first one is a very bad choice when used as a generator of random points in the plane.

$x_{n+1} = 33x_n + 1 \bmod 64$



$x_{n+1} = 37x_n + 1 \bmod 64$



L. Afflerbach/J. Lehn (ed.): Zufallszahlen und Simulationen. Teubner 1986.
M. Drmota/R. Tichy: Sequences, discrepancies and applications. SLN Math. 1651 (1997).
L. Hua/Y. Wang: Applications of number theory to numerical analysis. Springer 1981.
D. Knuth: The art of computer programming II. Addison-Wesley.
L. Kuipers/H. Niederreiter: Uniform distribution of sequences. Wiley 1974.
H. Niederreiter: Quasi-Monte Carlo methods and pseudo-random numbers. Bull. AMS 84 (1978), 957-1041.
H. Niederreiter: Random number generation and quasi-Monte Carlo methods. SIAM 1992.
H. Niederreiter/P. Shiue (ed.): Monte Carlo and quasi-Monte Carlo methods in scientific computing. Springer 1995.

Data security

Number theoretical cryptography is very popular now. Less known is that finite projective planes can be used very effectively to construct reliable security mechanisms.

A. Beutelspacher/U. Rosenbaum: Projektive Geometrie. Vieweg 1992.

Optimization theory

In the last twenty years optimization theory has become a very difficult field of mathematics, spanning from graph theory, convex geometry, 4-dimensional topology and the geometry of numbers to constraint logic programming, so that it is difficult for small teams of 2-3 persons, which are typical in Italian universities, to compete with the large groups of 50 people in other countries.

C. Großmann/J. Terno: Numerik der Optimierung. Teubner 1997.
P. Gruber/J. Wills (ed.): Handbook of convex geometry. 2 volumes. North-Holland 1993.
A. Joereßen/H. Sebastian: Problemlösung mit Modellen und Algorithmen. Teubner 1998.
D. Jungnickel: Graphen, Netzwerke und Algorithmen. Bibl. Inst. 1994.
K. Marriott/P. Stuckey: Programming with constraints. MIT Press 1998.
A. Schrijver: Theory of linear and integer programming. Wiley 1990.

Functional programming

Few, if any, mathematicians work in functional programming, which could be a very attractive mathematical part of computer science. A computer program can be considered as transforming an input x in an output y, i.e. as a mapping. In traditional programming languages this mapping is described by algorithms governed by case distinction and iteration. In functional programming languages this is achieved by describing the mapping as a composition of simpler mappings, exactly as mathematicians do when they say that $e^{x^2 \cos(x+y)}$ as a function of x and y can be considered as a composition of simpler functions, viz. $(x, y) \mapsto x + y$, $(x, y) \mapsto xy$, cos, $x \mapsto x^2$, $x \mapsto e^x$. In this sense algorithms are replaced by mathematical constructions, often leading to clean programs and easy mathematical proofs of correctness.

M. Erwig: Grundlagen funktioneller Programmierung. Oldenbourg 1999.
L. Paulson: ML for the working programmer. Cambridge UP 1996.
P. Pepper: Funktionale Programmierung. Springer 1999.
P. Thiemann: Grundlagen der funktionellen Programmierung. Teubner 1994.
J. Ullman: Elements of ML programming. Prentice Hall 1998.

Mathematical methods of bioinformatics

Bioinformatics is a mainly mathematical discipline, which provides algorithms and statistical techniques for the comparison of DNA sequences, the prediction of the spatial structures of proteins from their sequence of amino acids (e.g. via logic programming), phylogenetic trees, the assessment of probable biochemical or pharmacological activities. Methods of linguistics could, perhaps in an abstract mathematical form, be applied to the study of genes and amino acid sequences.

The complete availability of the genome of man and many other species will probably initiate a new era in biochemistry and pharmacology; it is significant than one of the most renowned leaders of human genome informatics is Eric Lander, once a mathematician working in combinatorial theory.

The comparison of genomes can allow pharmacologists to find genes for enzymes necessary to a parasite or other pathogenic agent, but not to man, and then it can be possible to develop drugs which inhibit that enzyme, though being innocent to man. A first famous example is the recent discovery of new antimalarial drugs by Jomaa and coll. who found a mevalonate-independent pathway of the biosynthesis of isoprenoids (as sterols and ubiquinons) and therefore, since in all mammals these molecules are synthesized on the mevalonate pathway, a target for new drugs.

H. Bandelt/A. Dress: Reconstructing the shape of a tree from observed dissimilarity data. Adv. Appl. Math. 7 (1986), 309-343.
H. Bandelt/A. Dress: A canonical decomposition theory for metrics on a finite set. Adv. Math. 92 (1992), 47-105.
J. Dopazo/A. Dress/A. Haeseler: Split decomposition - a technique to analyze viral evolution. Proc. Nat. Ac. Sci. USA 90 (1993), 10320-10324.
R. Hofestädt (ed.): Molekulare Bioinformatik. Shaker 1998.
R. Hofestädt (ed.): Bioinformatik 2000. Biocom AG 2000.
R. Hofestädt/H. Lim (ed.): Molecular bioinformatics. Shaker 1997.
H. Jomaa a.o.: Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalaria drugs. Science 3 September 1999, 1573-1576.
D. Sankoff/J. Kruskal (ed.): Time warps, string edits, and macromolecules. Addison-Wesley 1983.
S. Schulze-Kremer: Molecular bioinformatics. De Gruyter 1996.
J. Setubal/J. Meidanis: Introduction to computational molecular biology. PWS 1997.
M. Waterman: Introduction to computational biology. Chapman & Hall 1996.

Computational geometry and image processing

A very popular field with applications in medicine, geography, quality control and molecular modelling, involving many people.

H. Bungartz/M. Griebel/C. Zenger: Einführung in die Computergraphik. Vieweg 1996.
H. Edelsbrunner: Algorithms in combinatorial geometry. Springer 1987.
G. Farin: Curves and surfaces for computer aided geometric design. Academic Press 1993.
J. Foley a.o: Computer graphics. Addison-Wesley1996.
H. Handels: Medizinische Bildverarbeitung. Teubner 2000.
H. Hejmans: Morphological image operators. Academic Press 1994.
A. Janser/W. Luther/W. Otten: Computergraphik und Bildverarbeitung. Vieweg 1996.
J. Serra: Image analysis and mathematical morphology. Academic Press 1988.

Geomathematics

This is an emerging, beautiful, difficult and not overcrowded field of applied mathematics. We give here only a bibliography and recommend the home page at *www.mathematik.uni-kl.de / wwwgeo /* of the strong geomathematics group in Kaiserslautern, from which we list some of the many challenging topics they study: special functions of mathematical geophysics, spherical harmonics, pseudodifferential operators of mathematical geodesy, multivariate approximation methods, applications of splines, wavelets, finite elements in mathematical geodesy, determination of the gravitational field of the Earth, deformation analysis of the Earth's surface, atmospheric refraction effects, determination of the magnetic field of the Earth via satellite data and vectorial wavelets.

H. Engl/A. Louis/W. Rundell (ed.): Inverse problems in geophysical applications. SIAM 1997.
W. Freeden/T. Gervens/M. Schreiner: Constructive approximation on the sphere. With applications to geomathematics. Oxford UP 1998.
W. Freeden/U. Windheuser: Combined spherical harmonic and wavelet expansion - a future concept in Earth's gravitational determination. Appl. Comp. Harmon. Anal. 4 (1997), 1-37.
E. Groten: Geodesy and the Earth's gravity field. 2 volumes. Dümmler 1980.
S. Heitz: Coordinates in geodesy. Springer 1988.
L. Hörmander: The boundary problems of physical geodesy. Arch. rat. Mech. Anal. 62 (1976), 1-52.
M. Hotine/J. Zund: Differential geodesy. Springer 1991. A collection of papers by Hotine, commented by Zund.
A. Marussi: Intrinsic geodesy. Springer 1985.
J. Zund: Foundations of differential geodesy. Springer 1994.

Statistics

It would be easy for mathematicians to claim that statistics is, especially in its foundations, a mathematical discipline. Of course a statistician is destinated to become an applied mathematician and has to acquire knowledge in his fields of application, but the basic courses on probability theory, harmonic analysis, stochastic processes and time series, combinatorial theory (for skills in elementary probability theory and e.g. for the planning of experiments) are mathematics courses which should be taught in a mathematics department. But *"in mathematics ... research appears to be disproportionally driven by the internal development of mathematics itself ... Most statisticians understand that applied research is applied to something; this is not traditional usage in mathematical circles."*, as Moore/Cobb remark. Largely unexplored is the geometry of multidimensional statistics; cluster analysis is another important field where mathematicians could more participate.

There are so many applications of statistics, in medicine, bioinformatics, pharmacology, financial mathematics and linguistics, that statistics can attract many abstractly inclined students which should be in the first part of their education mathematicians.

In some countries the situation is different. In his book on medical statistics Horst Fassl complains several times about mathematicians in Germany being involved even too much in the planning and organization of medical activities; in fact, in Germany every university has a big institute of medical informatics or medical statistics with 20-50 people, many of them being mathematicians.

H. Bock: Automatische Klassifikation. Vandenhoeck 1974.
R. Farrell: Multivariate calculation. Springer 1985.
H. Fassl: Einführung in die medizinische Statistik. UTB 1999.
D. Moore/G. Cobb: Statistics and mathematics - tension and cooperation. Am. Math. Monthly 107 (2000), 615-630.

Lindenmayer systems

Lindenmayer systems can be considered as a part of symbolic dynamics. Start with the word *0* and rewrite this according to the rules *0* ⟶ *01* and *1* ⟶ *10* to obtain along *0* ⟶ *01* ⟶ *0110* ⟶ *01101001* ⟶ ... the famous Morse sequence.

Lindenmayer, a botanist, used such rules to model the development of plants, by interpreting the letters in such evolving sequences of words as structural elements of plants. One obtains today surprisingly realistic pictures. Also many fractals can be constructed this way. Modifications can be introduced for example via genetic algorithms.

O. Deussen/B. Lintermann: Computerpflanzen. Spektrum der Wiss. Februar 2001, 58-65.
C. Jacob: Principia evolvica. dpunkt 1997.
P. Prusinkiewicz/A. Lindenmayer: The algorithmic beauty of plants. Springer 1990.
G. Rozenberg/A. Salomaa: The mathematical theory of L-systems. Academic Press 1980.

Mathematical methods in chemistry

This is a vast field, which is *"central to rational drug design, ... contributes to the selection and synthesis of new materials, and ... guides the design of catalysts"*, as the National Academy report (150 pages) states, still in many parts unexplored. *"In contrast [to physics], there has been relatively little interaction between mathematicians and chemists ... a large number of diverse areas in chemistry ... would benefit from input from mathematicians. A small subset of these areas includes: strategies [via statistical prediction] for the design [and assessment] of new drugs and agricultural chemicals, development of molecular dynamics algorithms, optimization problems, folding of proteins, transport across biological membranes, coiling and uncoiling of DNA, the structure of crystals and quasicrystals, the relationship between quantum mechanics and simpler approximate models, and a wide range of numerical analysis problems"*, writes George Hagedorn.

Other research areas in which mathematicians could contribute to chemistry mentioned in the NAS report are distance geometry (applied to the study of the 3-dimensional structure of molecules, for example in docking problems), topology and graph theory (with many old and new applications, e.g. to the theory of fullerenes, to the representation of chemical reactions, or to the organization of the chemical literature or of chemical databases), connections with number theory via quasicrystals, combinatorics, molecular diversity and combinatorial chemistry in drug discovery (including cluster analysis and genetic algorithms), quantitative structure-activity relationship and similar methods, fast Fourier transform methods for applications in molecular spectroscopy, and the many aspects of computational quantum chemistry (including quantum Monte Carlo solutions of the Schrödinger equation, where the statistical properties of pseudorandom number generators are needed).

It may be interesting that the book by Bartel is almost entirely dedicated to methods from discrete mathematics, usually not so well known to chemists: sets, mappings and relations; graph theory; algebraic structures (mainly groups) and orders with an exposition of the algebraic structure of quantum theory; formal concept analysis; cluster analysis.

"Neither the chemist nor the mathematician is generally the optimal person to construct a mathematical model, as the model by its very nature lies at the interface between theory and observation. To build the model, an iterative process of refinement is required ... It is exactly this need for iterative model construction that may motivate the collaboration of mathematicians and chemists, against the self-referential and conservative tendencies of each discipline ... although theoretical chemists understand sophisticated mathematics ... they have typically not involved mathematicians directly ... theoretical chemists have become accustomed to self-reliance in mathematics ... both fields are affected by the value system of academia ... For mathematicians, the potential career damage of collaboration rises when it involves work in a field seen as peripheral to mathematics. In some instances, interdisciplinary work may be regarded by one's mathematical colleages as not real mathematics or as less valuable than traditional mathematics ... For both fields, the difficulty of interdisciplinary collaboration is exacerbated by the lack of a well-established network of contacts between mathematicians and chemists ... chemists may regard mathematicians as unapproachable or uninterested in chemistry problems; mathematicians may not realize that chemistry problems contain interesting and novel mathematics ... Part of the gap specifically between mathematics and chemistry can be explained by long-standing pedagogical practices in mathematics. Much of classical applied mathematics is based on constructions associated with mechanics and physics: every student of mathematics studies the heat equation, elastic rods, electrical networks, and fluid flow. However, no problems explicitly associated with chemistry are widely taught to or recognized by mathematicians." (NAS report)

H. Bartel: Mathematische Methoden in der Chemie. Spektrum 1996.
G. Hagedorn: Crossing the interface between chemistry and mathematics. Notices AMS March 1996, 297-299.
National Academy of Sciences: Mathematical challenges from theoretical/computational chemistry. National Academy Press 1995. Available at www.nap.edu/readingroom/books/mctcc/.

Caesars in the villages?

In his book on genes and languages Cavalli-Sforza comments on one of Leopardi's *Operette morali* which cites Plutarch's report on Caesar saying to his officers, while crossing the Alps in the journey to his province and passing a small village of barbarians, that he had rather be the first man among those fellows than the second man in Rome. Let us abuse this famous incident to place some philosophy on the difference between politicians and scientists.

In fact, the assertion referred to is so typical for politically inclined people, that the question *"Would you prefer to be the first man in a village or in a factory etc. or rather to participate in a general human enterprise as a second or third man, without commanding at all, but only warranted to be allowed to contribute to the enterprise?"* could be a valid test to discriminate among similarly talented and ambitioned people between politicians and scientists.

L. Cavalli-Sforza: Gènes, peuples et langues. Odile Jacob 1996.

Teaching mathematics

There are different levels and different ages at which mathematics can be taught, and probably the most important technique in teaching mathematics is to bear in mind the needs, ambitions, possibilities and foreknowledges of the students. Teaching to 10 year old boys must be different from teaching in technical high schools, and at the university level teaching to mathematics or physics students shall be different from courses to economists or biologists.

Also for mathematics students at the university a good organization is important. One should not underestimate the students, one should also attract good students to mathematics. This requires a careful, but versatile and differentiated planning of curricula, effective exercise assignments and interesting topics.

Beauty is not enough today, if this means that occupying one's time with abstract mathematics is beautiful; there are too many other much more beautiful distractions for young people. Attractiveness must therefore come from the students aquiring applicable knowledges which give them self-confidence to compete with engineering and computer science students, with openings to future interdisciplinary activities.