II_1 Factors, subfactors and Jones' index

We begin by taking a closer look at the notion of *dimension* and suggest a reformulation of the notion that will be suitable for generalization. To start with, the space we live in is 'three-dimensional real space'; the analytic formulation of this space $\mathbb{R}^3$ is as the set of ordered triples $\mathbf{x}=(x_1,x_2,x_3)$ of real numbers, endowed with certain operations called *scalar multiplication* and *vector addition* which capture the 'linear structure' of $\mathbb{R}^3$. (This means that we can, given vectors $\mathbf{x},\mathbf{y}$ in $\mathbb{R}^3$ and a real number λ, form vectors denoted $\lambda\mathbf{x}$ $(=(\lambda x_1,\lambda x_2,\lambda x_3))$ and $\mathbf{x}+\mathbf{y}$ $(=(x_1+y_1,x_2+y_2,x_3+y_3))$, and these operations satisfy certain 'natural' compatibility relations.) Exactly the same prescriptions, with 3 replaced by an arbitrary positive integer n, result in n-dimensional real space $\mathbb{R}^n$. An age-old dictum of abstract mathematics is that all the information contained in a set endowed with a 'structure' is contained in the collection of those transformations of that set which 'preserve' that structure, the collection of such transformations inheriting a natural algebraic structure. In the case of $\mathbb{R}^n$, this leads us to look at the *so-called linear transformations* of $\mathbb{R}^n$, viz. those transformations—such as projections (onto subspaces), reflections and rotations—which fix the origin and map any three collinear points to three collinear points. It is not hard to prove that such transformations are described by $n\times n$ real matrices. (An $n\times n$ real matrix is an array $x=((x_{ij}))$ of n^2 real numbers arranged in a *square array* with n rows and n columns, with the above notation signifying that the real number x_{ij} appears at the intersection of the ith row and jth column of the matrix x.)

The collection $M_n(\mathbb{R})$ of all $n\times n$ real matrices has the structure of an *involutive algebra*; in addition to being able to multiply a matrix by a real number as well as to add two matrices, we can also multiply two $n\times n$ real matrices as well as form the 'adjoint or transpose' x^* of a matrix x, these operations being 'compatible' in a natural manner. (We do not go into more detail; it suffices to know formally that $M_n(\mathbb{R})$ admits such a 'structure'.) The final bit of structure on the matrix algebra is that it admits a unique 'normalized trace; i.e. there exists a unique mapping $x\to \mathrm{tr}\,x$, from $M_n(\mathbf{R})$ to $\mathbf{R}$, satisfying, for all matrices x,y and all real numbers λ,μ, the conditions: (i) $\mathrm{tr}\,(\lambda x+\mu y)=\lambda\,\mathrm{tr}\,x+\mu\,\mathrm{tr}\,y$, (ii) $\mathrm{tr}\,xy=\mathrm{tr}\,yx$, and (iii) $\mathrm{tr}\,1=1$, where the 1 on the left denotes the so-called identity matrix. (The normalized trace of matrix

$$x=((x_{ij}))\text{ is given by tr }x=\frac{1}{n}\,\Sigma_{i=1}^{n}\,x_{ii}.)$$

To each 'subspace' of $\mathbb{R}^n$—meaning a line or a plane or a possibly higher dimensional 'hyperplane' through the origin (or more precisely, a subset of $\mathbb{R}^n$ that is 'closed' under scalar multiplication and vector addition)—is associated a canonical matrix, viz. the matrix that represents the linear transformation corresponding to 'perpendicular projection' onto that subspace. Such projections are algebraically characterized as being elements p of $M_n(\mathbb{R})$ which satisfy $p=p^*=p^2$. It is an easy fact from basic linear algebra that if p denotes the projection onto the subspace $\mathscr{M}$, then $\mathrm{tr}\;p=k/n$, where k is the dimension of the subspace $\mathscr{M}$. (It goes without saying that as $\mathscr{M}$ varies, the corresponding dimension k ranges over the integers $0, 1, \ldots, n$.)

In his pioneering investigations into certain special algebras (that he called 'rings of operators' and which are more commonly known these days under the name of 'von Neumann algebras'), von Neumann identified certain 'building blocks' (that he called, and are still) called *factors*, and these he classified into three basic types labelled I, II and III. In a sense, the matrix algebra $M_n(\mathbb{R})$ is the prototype of a (finite) type I factor.

Our concern is with the (finite) type II factors, the so-called II_1 factors. Like the matrix algebra, these are involutive algebras (over complex rather than real numbers, due to certain technical reasons) that admit a unique normalized trace; the key difference is that in this case, as p ranges over projections in a II_1 factor—i.e., satisfy $p=p^*=p^2$—the numbers $\mathrm{tr}\;p$ range range over the continuum of all real numbers between 0 and 1 (inclusive). (This must be contrasted with the case of the matrix-algebra $M_n(\mathbb{R})$ where these permissible dimension-values range over the discrete set of multiples of a smallest possible dimension.) Another pleasing feature of II_1 factors is that, while the real (or complex) 'vector spaces of finite dimension' are parametrized by the non-negative integers, the 'modules of *finite dimension*' over a II_1 factor N (say) are parametrized by the non-negative real numbers; the number so associated to an N-module $\mathfrak{H}$ is usually denoted by $\dim_N(\mathfrak{H})$. (The difference between N-modules and real vector spaces (such as $\mathbb{R}^n$) is that now 'scalar' multiplication makes sense with the II_1 factor N playing the role of the real numbers.)

Jones considered nested pairs $N\subset M$ of pairs of II_1 factors. Since products x,y make sense for all x,y in M, we see, by forgetting that x can come from outside N, that M gives rise naturally to an N-module usually denoted by L^2 (M,tr). (Since modules over II_1 factors are usually Hilbert spaces, the module in question is 'the Hilbert space-completion $L^2(M,\mathrm{tr})$ of M' with respect to an inner product induced by the trace, whereby $(x,y)=\mathrm{tr}\,y^*x$ for x,y in M.) When this N-module 'is of finite N-dimension', let us follow Jones' terminology and say that the subfactor N has finite index in M, and (denote and) define this index by $[M:N]=\dim_N L^2(M,\mathrm{tr})$. Past experience with II_1 factors would seem to indicate that all real values (not less than 1—since the index can be seen, by definition, to be always at least one) might arise as possible 'index-values—i.e. numbers of the form

$[M:N]$ as above. However, in Jones' words, this effect cannot be turned on until after four! Thus, in sharp contradiction to the naive guess expressed earlier, we have the following striking result of Jones.

Theorem: (Jones)

If N is a subfactor of finite index in a II_1 factor M, then either $[M:N] \geqslant 4$ or $[M:N] = 4\cos^2 \pi/n$ for some positive integer $n > 2$.

Further, if λ is any real number either not less than four, or equal to one of the numbers $4\cos^2 \pi/n$, then $\lambda = [M:N]$ for some pair $N \subset M$ of II_1 factors. □

A key ingredient in the proof of the above theorem is the following: if $N \subset M$ is a pair of II_1 factors as in the theorem, then $L^2(N,\mathrm{tr})$ sits naturally as a subspace of $L^2(M,\mathrm{tr})$ and we may consider the *perpendicular projection* e_0 onto the subspace $L^2(N,\mathrm{tr})$; the key fact is that then, 'the von Neumann algebra generated by M and e_0' is again a II_1 factor M_1 that contains M as a subfactor with index equal to $[M:N]$. (The pair $M \subset M_1$ is 'dual' to the pair $N \subset M$ in a precise sense that contains the Pontrjagin duality theorem, at least for finite abelian groups, as a special case.)

Clearly we may start with the pair $M \subset M_1$ rather than the pair $N \subset M$, and the same game, as before, can be played with the 'dual pair' $M \subset M_1$ to produce a projection e_1 such that if M_2 is 'the von Neumann algebra generated by M_1 and e_1', then the pair $M_1 \subset M_2$ is 'dual' to the pair $M \subset M_1$ just as the pair $M \subset M_1$ was 'dual' to the pair $N \subset M$. Play this game *ad infinitum* to obtain a sequence $(e_n : n = 0,1,2,\ldots)$ and a tower $N \subset M \subset M_1 \ldots \subset M_n \subset \ldots$ of II_1 factors, in which each member of the tower has index $[M:N]$ in its successor.

The final ingredient from Jones' work on subfactors that is needed to complete the connection with knot theory is the following fact, also established by Jones.

Proposition: (Jones)

Let $N, M (= M_0)$, M_n, e_n be as above, and let τ denote $([M:N])^{-1}$. Then the sequence $(e_n : n = 0,1,2,\ldots)$ generates a II_1 factor R in which the following relations hold:

(a) $e_n e_m = e_m e_n$ if $m \neq n \pm 1$;
(b) $e_n e_{n\pm1} e_n = \tau e_n$ for all n; and
(c) $\mathrm{tr}\, w e_n = \tau\, \mathrm{tr}\, w$, for any 'word' w in $e_0, e_1, \ldots, e_{n-1}$.
(The trace 'tr' appearing above is the unique normalized trace on R.) □

The (one-variable) Jones polynomial invariant of links

The first step is to observe the formal similarity in the relations satisfied by the elementary braids (of Artin's theorem) and by the Jones projections (of the last proposition in the last section). This formal similarity is converted into a precise connection as follows.

Suppose then that $(e_n : n = 0,1,2,\ldots)$ and τ are as in the last-stated proposition above. (It is a fact converesely that if the e_n and τ are related as in conditions (a)–(c) of the proposition above, then the reciprocal τ^{-1} of τ must be one of the numbers ($4\cos^2 \pi/n$) if it is smaller than 4!). Define

$$g_i = \sqrt{t}\,((t+1)e_i - 1), \text{ where } t \text{ satisfies } \tau = \frac{t}{(t+1)^2}.$$

It can then be seen that for each n, $(g_i : 0 < i < n)$ satisfy the 'braid relations' and hence, by Artin's theorem, there exists a unique (homomorphism, meaning) product-preserving map $\pi^{(n)}$ from B_n into (the group of invertible elements in the II_1 factor) R such that $\pi^{(n)}(\sigma_i^{(n)}) = g_i$ for $0 < i < n$. The reason for this whole exercise is contained in the following:

Theorem: (Jones)

With the notation as above, the equation

$$V_{(\alpha^{(n)})^\wedge}(t) = \left(-\frac{t+1}{\sqrt{t}}\right)^{n-1} \mathrm{tr}_R\, \pi^{(n)}(\alpha^{(n)})$$

defines an invariant of the link $(\alpha^{(n)})^\wedge$.

(Here, tr_R denotes the unique trace on the II_1 factor R.) □

Remarks

(1) The content of the theorem is that if L is any link, and if $\alpha^{(n)}$ is any braid with closure equivalent to L, then the right side of the equation in the theorem depends only on the equivalence class of the link L (and hence independent of the choice of the representing braid α).

(2) The left side has been written to look like a function of the variable t; as such, it makes sense, as yet, only for some special values of t—those for which $(t+1)^2/t$ is one of the 'Jones' numbers, i.e. is a possible index-value. It is a remarkable fact that this function $V_L(t)$ actually behaves as a Laurent polynomial in $\sqrt{t}$—i.e. is a finite sum of powers of $\sqrt{t}$ and $\sqrt{t}^{-1}$; in fact, if K is a knot (=1-component link), then $V_K(t)$ is actually a Laurent polynomial in t itself.

(3) The constants occurring in the theorem are all forced by the requirements; thus, the elements $(\tilde{g}_i)$ defined by $\tilde{g}_i = c(\lambda e_i - 1)$ satisfy the braid relations if and only if $\lambda = t + 1$; and a prescription as in the theorem will be invariant under Markov moves of type II if and only if $\mathrm{tr}\, \tilde{g}_i = \mathrm{tr}\, (\tilde{g}_i)^{-1}$ which happens precisely for the choice we made for c in the theorem. □

As a concrete application of this polynomial invariant, it may be noted that the two trefoil knots have the presentations $T_{\pm} = (\sigma_1^{(2)\pm 1})^3$—where, of course σ^3 means $\sigma.\sigma.\sigma$. An easy computation then shows that $V_{T_+}(t) = t + t^3 - t^4$ while the invariant for T_- is obtained by replacing t by t^{-1} everywhere in the invariant for T_+. Also the unknot U_1 has the obvious presentation 1_1 and hence $V_{U_1}(t) \equiv 1$. Thus the three knots U_1, T_+ and T_- are indeed, pairwise inequivalent, as stated earlier.

A feature of this invariant has already been indicated: if a link $L^{\#}$ is obtained as a 'mirror-reflection' of a link L, then $L_{L^{\#}}(t) = V_L(t^{-1})$; thus the Jones invariant is quite good at detecting when a link is inequivalent to its reflection.

Much has been omitted in this terse, perhaps over-simplified, discussion of the Jones invariant. Thus, for instance, nothing has been said about the so-called *skein relations* satisfied by three links which 'look the same except at one crossing'—these relations being the basic method of proving various assertions about the invariant (including, for example, the polynomial nature of the invariant). Also, no attempt has been made to discuss the 'two-variable' Jones invariant (which subsumes, as special cases, the one-variable invariant discussed above, as well as one of the oldest knot-invariants, the Alexander polynomial). Furthermore, no mention has been made of Jones' other work in von Neumann algebras, or of the subsequent profound results due to Ocneanu, Popa, Wasserman and Wenzl that have been motivated by Jones' work and that strongly influence the trend of current research in operator algebras.

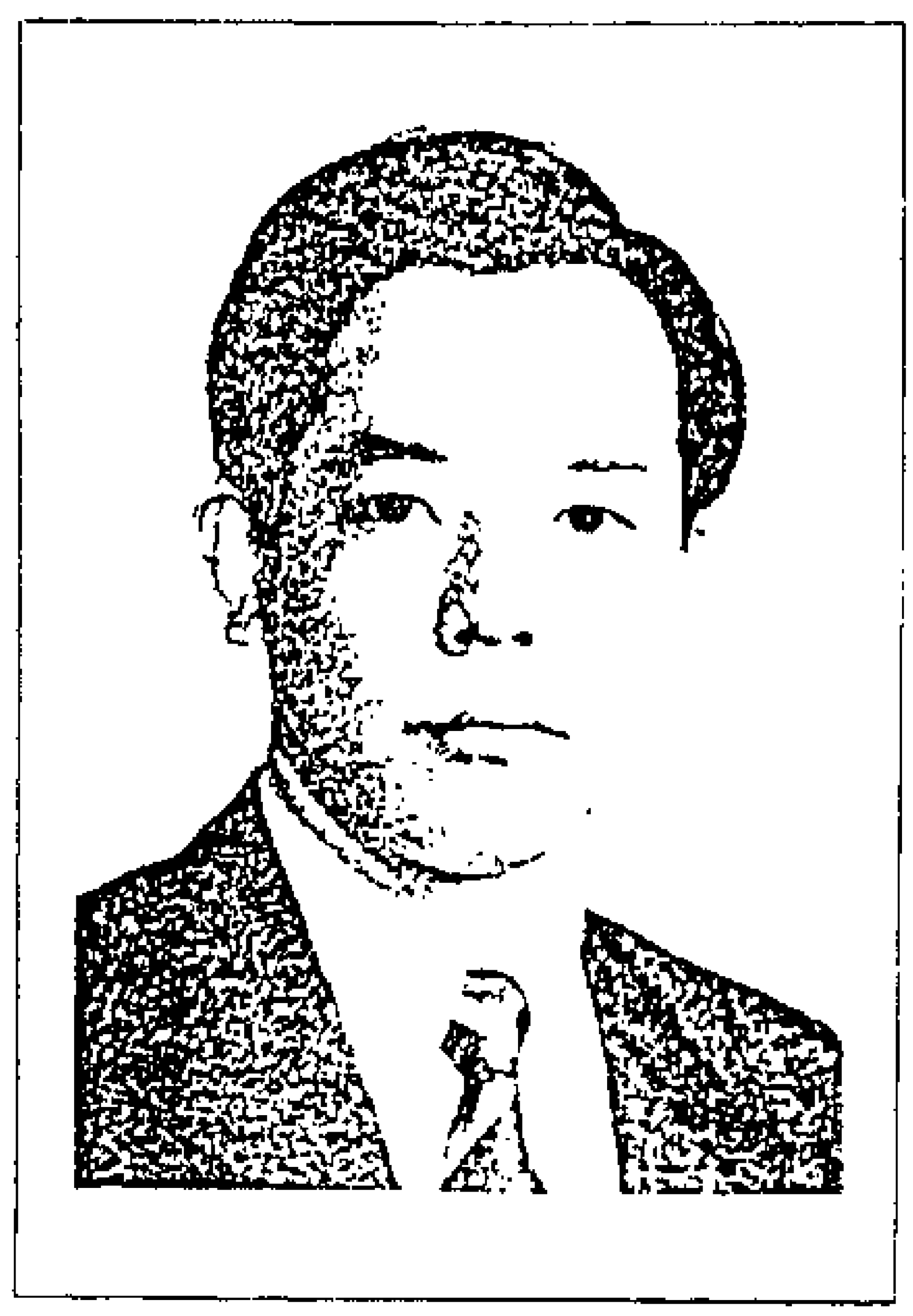
We conclude with a short bibliography; the reader interested in more rigour is directed to [H] which is an elaborate and detailed exposition of the (two-variable) Jones polynomial.

1. [H] Pierre de la Harpe, l'Enseignement Mathematique, 1988, **32**, 271.
2. [J1] Vaughan Jones, *Invent. Math.*, 1983, **72**, 1.
3. [J2] Vaughan Jones, *Ann. Math.*, 1987, **126**, 335.
4. [J3] Vaughan Jones, Notes on Subfactors and Statistical Mechanics, preprint.

The work of S. Mori

V. Srinivas

It is a pleasure to write an article about the work of Shigefumi Mori, one of the four Fields medallists this year at the International Congress[1] of Mathematicians held at Kyoto, Japan in August. Mori's research is in the field of algebraic geometry, which is my own primary field of interest.



The Fields medals play something of the role of a Nobel prize in mathematics. They are awarded every four years at the International Congress to two to four mathematicians by a panel of renowned mathematicians appointed by the International Mathematical Union. One difference from the Nobel prize is that there is an age limit: only mathematicians below 40 qualify!

The names of the Fields medallists (and those of members of the Fields medal selection committee) are a closely guarded secret before each Congress, and there is often some suspense as to who the winners of medals will be. On the first day of the Congress, after the opening ceremonies, the names of the winners (and those of the members of the Fields committee) are read out, along with a mention of the major research works of each of the winners. The same afternoon, expository lectures are given explaining the works of the awardees.

V. Srinivas is in the Tata Institute of Fundamental Research, Bombay 400 005.

This year's Fields committee was chaired by L. D. Faddeev, and consisted of M. Atiyah, J. M. Bismut, E. Bombieri, C. Fefferman, K. Iwasawa, P. D. Lax, I. R. Shafarevich and J. G. Thompson. It is perhaps of interest to observe that four of the committee members are former Fields medallists (Atiyah, Bombieri, Fefferman and Thompson). The Fields prize is usually awarded for work in 'pure' mathematics; there is a similar award (the Rolf Nevanlinna Prize) for 'applied' mathematics, which is decided by the same committee. Interestingly, three of the Fields medallists this year had worked in areas related to mathematical physics, while the winner of the Nevanlinna prize is a theoretical computer scientist.

Shigefumi Mori was born in Nagoya, Japan on 23 February 1951. He received his Ph D from the Kyoto University, under the direction of Masayoshi Nagata, the famous algebraist and algebraic geometer. He is presently at the Research Institute for Mathematical Sciences of the Kyoto University.

The results cited for Mori's award are the following: (i) the proof of Hartshorne's conjecture, that the projective space is the only non-singular projective algebraic variety which has an *ample* tangent bundle, (ii) the classification of three-dimensional Fano varieties with second Betti number $b_2 > 1$, and (iii) his work on the theory of *minimal models* for three-dimensional algebraic varieties. I will try to explain more about this work in the course of this article. The lecture at the Congress explaining Mori's work was given by Hironaka, a former Fields medallist from Japan; in fact there have been three Japanese Fields medallists in all—Mori, Hironaka and Kodaira (in reverse chronological order), and incidentally, all of them have been algebraic geometers.

Broadly speaking, algebraic geometry[2] is the study of the geometry of the spaces of solutions of systems of *polynomial equations*. It may be said to have originated in the work of the Greeks on conic sections. Closer to modern times, one may identify the 'real' beginning of algebraic geometry with the introduction of coordinate geometry by Fermat and Descartes; it began to come into its own in the last century with the theory of algebraic curves, studied by Galois, Abel, Jacobi, Riemann and others.

The first object of interest in algebraic geometry is a *plane algebraic curve*[3], which is the solution set of a single polynomial equation

$$f(x, y) = \sum_{m,n} a_{m,n} x^m y^n = 0.$$

One usually considers the solutions $x = a, y = b$ of this equation where a, b are *complex numbers*; further, one augments the curve C of these solutions by a finite set of *points at infinity*. One reason to do this is to have a principle of *continuity of number*, which states roughly that if C_t is a continuous family of curves, the number of points of intersection of each C_t with a fixed curve D is a constant (independent of t). This is already false for lines in the plane, because of parallel lines, showing the necessity of points at infinity. Similarly, if we count the number of points of intersection of each of the lines.

$$x = t,$$

for varying t, with the circle

$$x^2 + y^2 = 1,$$

we see that there are always two points of intersection if we consider points with complex coordinates, and count the tangential intersection for $t = 1$ 'with multiplicity 2'; if we work only with points with real number coordinates, then for $t > 1$ there are no points of intersection.

The problem of parallel lines suggests how to introduce 'points at infinity'—associated to any line in the plane, one has a 'point at infinity'; in order to give truth to the statement that 'any two distinct lines have exactly one point of intersection', we are forced to define the points at infinity on two lines to be the same precisely if the two lines are parallel. One thus obtains a new space called the *projective plane*, first introduced by the French geometer Poncelet early in the last century. The picture is rounded out by making the set of all points at infinity into a line; the fact that each 'finite' line has exactly one point at infinity becomes a particular case of the statement that any two lines meet in a point, namely the case when one of the lines is the line at infinity.

More generally, given an algebraic plane curve C defined by a polynomial equation $f(x, y) = 0$, define the *degree* of C to be the largest value of the sum $m + n$ such that f has a monomial $x^m y^n$ with a non-zero coefficient. If C has degree d, there is a simple procedure to 'enlarge' C by adding d points at infinity to it, determined[4] by the polynomial f. The resulting object is called a *projective plane curve*.

One has the famous theorem of Bezout from the last century: if C and D are two projective plane curves of degrees m and n which intersect in a finite set of points, then the number of points of intersection 'counted properly' is mn ('counted properly' refers to the fact that tangential intersections are counted as more than one point). Indeed, we have already seen an instance of this—the number of points at infinity of a plane curve C of degree d is just the number of intersections of the corresponding projective curve $\bar{C}$ with the line at infinity (which has degree 1).

The above notions are generalized in several directions:

(i) by considering sets of common solutions of several equations in an arbitrary number of variables, called *algebraic varieties*;

(ii) by considering projective space of any dimension and were reworked and complemented by more recent authors).

The most important application of Mori's theory is in the construction of minimal models for function fields of 3-folds (three-dimensional varieties), which is the work of several mathematicians[13]. Roughly speaking, one would like to have a theorem of the following form: if X is an arbitrary smooth projective variety, and $\kappa_X(n)$ is the dimension of its space of n-canonical forms (sections of the nth tensor power of the bundle of d-forms, where $d = \dim X$), and $\kappa_X(n)$ grows like n^r for large positive n, then there is a 'naturally defined' mapping to an r-dimensional variety Y with fibres of a 'simple' type, if $r < \dim X$; or else, $r = \dim X$, and there is a *minimal model* for the field of meromorphic functions, and the given X is obtained from the minimal model by a sequence of simple operations.

A result of the above type was one of the goals of the theory of algebraic surfaces, which began with the work of the Italian algebraic geometers in the early part of this century, led by Enriques, Castelnuovo and Severi, and was completed by more recent authors, notably Kodaira.

There are new complications in dimension 3. One more or less had control of the case when $\kappa_X(n)$ was at most quadratic. However, in case $\kappa_X(n)$ grows like n^3, there were examples known to show that there cannot exist a *smooth* projective minimal model as desired. One of the insights of the classification theory of 3-folds is that one needs to allow singular models, and to recognize precisely which singularities are acceptable — the class of so-called *terminal singularities*. Nicely enough, the notion of a terminal singularity is defined in terms of canonical divisors, and for surfaces, only smooth points are 'terminal singularities'!

Again, unlike in the case of surfaces, one has examples to show that there is no chance of reconstructing any X from a minimal model $\bar{X}$ (if one exists) just by blow-ups of points and curves — a new operation called a *flip* is also required; it is a blow-up of a rational curve, and then a blow-down of the resulting 'exceptional surface' over this rational curve to another rational curve, in a 'different direction'. So as a point set, the flip does not do anything, it seems! Of course the topology has changed. In the reverse direction, if we want to construct a minimal model from a given X, one needs to 'undo' the effects of flips. The inverse to a flip also looks like a flip (i.e. the situation between a variety and the 'flipped' one is symmetric), so one needs to show that for a given X, one can either (i) blow down something, keeping only terminal singularities, or (ii) find a finite sequence of flips, so that we will be able to carry out step (i) on the new variety. This decisive final step was also accomplished by Mori, by a very complicated argument[14].

1. More information about the International Congress can be found in Albers, D. J., Alexanderson, G. L., Reid, C., *International Mathematical Congresses: An illustrated history, 1893–1986*, Springer-Verlag, New York 1987.
2. Two basic works on algebraic geometry, very different in spirit, are Hartshorne, R., *Algebraic Geometry*, Grad. Texts in Math. No. 52, Springer-Verlag, New York, 1978; Griffiths, P. A. and Harris, J., *Principles of Algebraic Geometry*, Wiley-Interscience, New York, 1978.
3. A down-to-earth introduction to the theory of curves can be found in Fulton, W., *Algebraic Curves*, Benjamin, New York, 1969.
4. Let $f_d = \Sigma_{m+n=d} a_{m,n} x^m y^n$ be the 'highest degree' part of f; this factors as a product of d linear factors, corresponding to d lines; the points to be added to C are the points at infinity of these lines.
5. This allows one to obtain a sort of 'bridge' between algebraic geometry over fields of positive characteristic and over **C**, which is exploited by Mori.
6. In the simplest case of a plane curve, defined by a polynomial equation

$$f(x, y) = \sum_{m,n} a_{m,n} x^m y^n = 0,$$

one can show using the *implicit function theorem* that C is non-singular if for each point $x = a$, $y = b$ of C, one of the partial derivatives $\partial f/\partial x$, $\partial f/\partial y$ is non-zero when we set $x = a$, $y = b$. In a similar way, there is a sort of implicit function test for the non-singularity of any algebraic variety.

7. Proofs of the statements made about Riemann surfaces may be found in Forster, O., *Lectures on Riemann Surfaces*, Grad. Texts in Math. No. 81, Springer-Verlag, New York, 1981, and the references given there.
8. If k is algebraically closed, rational curves over k are precisely the curves of genus 0.
9. Mori, S., *Ann. Math.*, 1976, 110.
10. Mori, S. and Miyaoka, Y., *Ann. Math.*, 1986, 124.
11. Mori, S., *Ann. Math.*, 1982, 116.
12. Mori, S. and Mukai, S., *Manuscr. Math.*, 1981/82, 36.
13. Kollar, S., *Bull. Am. Math. Soc.*, 1987, 17.
14. Mori, S., *J. Am. Math. Soc.*, 1988, 1.

ACKNOWLEDGEMENTS. I would like to thank a number of my colleagues at the Tata Institute for their comments on a preliminary version of this article: H. Janwa, V. B. Mehta, M. S. Raghunathan, T. R. Ramadas and S. Ramanan.