Econometric Theory, **12**, 1996, 155–185. Printed in the United States of America.

THE ET INTERVIEW: PROFESSOR A.T. JAMES

Interviewed by Grant H. Hillier
University of Southampton
and
Christopher L. Skeels
The Australian National University

Alan T. James

Alan James (circa 1944)

Alan James (circa 1960)

Alan Treleven James retired from the Chair of Statistics at the University of Adelaide in 1989, becoming Emeritus Professor, but has by no means retired from his research at the interface of biochemistry and statistics. Alan's statistical research, which began in Australia's Commonwealth Sci-

We thank Kellie Maloney, Tracy Miners, and Clarence Tiong for their invaluable assistance in the difficult task of translating the spoken word to finished manuscript.

 0266-4666/96 $9.00 + .10

entific and Industrial Research Organisation (CSIRO) under Alf Cornish, has always been strongly motivated by real scientific problems. Nevertheless, it is almost certainly for his pathbreaking theoretical work in multivariate analysis that Alan James is best known to statisticians.

Alan's first published paper (1954), part of his Princeton doctoral thesis, has become a classic, introducing new and powerful algebraic tools into the analysis of distribution problems. The impact of that paper alone on distribution theory over the intervening 40 years has been immense (see, e.g., Farrell, 1976, 1985), but within the next 10 years he had produced two more papers that have had a profound influence on quite separate branches of multivariate analysis—the relationship algebra paper (1957), and the survey paper on the role of zonal polynomials in multivariate distribution theory (1964). In fact, as the interview that follows will reveal, there was an intimate connection between the zonal polynomial work and that on experimental design—a connection that took 20 years to come to light! (in 1982). The story should provide food for thought for contemporary journal editors!

Alan James' work on zonal polynomials opened the way for a systematic development of noncentral multivariate distribution theory, enabling the generalization of analogous univariate results. Because most interesting econometric models are both multivariate and entail nonzero means, this work, and its extensions by Davis (1979) and Chikuse and Davis (1986), has emerged as the key to a general distribution theory for both static structural models (where, under Gaussian assumptions, results are exact), and dynamic structural time series models [where it yields a nonstandard asymptotic theory (see Phillips, 1989; Choi and Phillips, 1992)]. To those familiar with this work, the interview that follows should provide some fascinating insight into its genesis; to others, it will surely provoke an interest in the deep algebraic structures that are at its roots.

Between 1958 and 1965, prior to his return to Adelaide to take up the Chair of Statistics, Alan James held the positions of Visiting Lecturer, Associate Professor of Mathematics and Professor of Mathematical Statistics at Yale University, where he was a founder-member of the Department of Statistics. In 1965 he was also Visiting Overseas Fellow at Churchill College, Cambridge. He has been an Associate Editor of the *Annals of Mathematical Statistics* and of the *Journal of Multivariate Analysis*. He has also been elected Fellow of the Institute of Mathematical Statistics and of the American Statistical Association. In 1992, he was awarded the prestigious Pitman Medal by the Statistical Society of Australia in recognition of his distinguished contributions to multivariate analysis (*Australian Journal of Statistics* 35, 1-4).

The interview that follows took place on September 16, 1994, in the Department of Statistics at the Australian National University. The interview was conducted by Chris Skeels and the questions largely posed by Grant Hillier. We hope that it will be of interest to both econometricians and statisticians.

Can we begin with your early mathematics training at Adelaide University—do any teachers stand out as being particularly influential over that period?

Yes, there was one, Hans Schwerdtfeger; I even had him in first year, and he came as a refugee from Nazi Germany. He had been at the Mathematisches Institut at Göttingen and brought to Adelaide, at the end of the earth, the very latest mathematics from what, at the time, was the leading school of Hilbert. In third year he taught a long course of analysis and matrix algebra and opened our eyes to the latest mathematics that was going on in the world. In particular, he had a great interest in Herman Weyl's work on the theory of groups and had also spent a year with Elie Cartan. So he brought to us, in his third year and honors courses, the work of these two men. Issai Schur in Berlin was a strong indirect influence. Herman Weyl (1946) dedicated his *Classical Groups* to Schur, and Bochner, who was at Princeton, had been a student under Schur.

From Adelaide you went on to work as an assistant research officer at CSIRO under Alf Cornish. Could you tell us a little bit about Cornish and his role in your interest in statistics?

Cornish was invited by Professor Sanders at Adelaide to set up a CSIRO Section of Mathematical Statistics in the Mathematics Department at Adelaide in 1944. They made rooms available to him for this section, although they were desperately short of space at that stage, and he began teaching a course of statistical methods in second year and a course of mathematical statistics in third year. Cornish had become interested in statistics and was made statistician to the Waite Institute in Adelaide. He saw Fisher's methods from Rothamsted as the answer to their problems. As a consequence, the University of Adelaide gave him leave to study for 2 years with Fisher. Naturally Cornish often spoke at great length about Fisher and introduced me, as other members of the Section of Mathematical Statistics, to the basic ideas of Fisher on statistical inference. I was most fortunate to work under Cornish because he gave me a remarkable degree of freedom and allowed me, all the time I was with him, a very free rein in carrying out my research, which I think, in retrospect, was really necessary to do original work.

Part of the time that you were working at CSIRO, I believe 1947 to 1948, you were sent up to Canberra. That's before Pat Moran, Ted Hannan, Joe Gani, and people like that appeared on the scene here. Was there anyone that you talked to very much up here?

Not at all in Statistics or Mathematics but in my spare time from the consulting work at CSIRO, as relieving biometrician to the Divisions of Plant Industry and Economic Entomology, I worked on my M.Sc. thesis, "The Geometrical Interpretation of the Analysis of Variance."

Now you submitted that in 1949 and then subsequently you went over to Princeton as a Ph.D. student where you were, as a student, from 1950 to 1952. That must have been moderately unusual for an Australian in those days when there were still very strong links with Britain. What drew you to Princeton in particular?

Dr. Cornish had secured this CSIRO studentship for me, and he felt that, although Fisher was undoubtably in his view the best statistician in the world. Fisher had become very interested in genetics and I might not pick up much on statistics from him. Also, he strongly believed that regression theory should be put into matrix algebra. This was one of his pet projects, although Yates, who had succeeded Fisher at Rothamsted, was still crossing out matrices from the papers of his staff when he vetted them for publication, until about the year 1960, when they became needed for computerization. Also, Cornish had been tutored in mathematics by Professor Wilton and had picked up the analysis of Cambridge of about 50 years beforehand but not the modern mathematics like Schwerdtfeger taught. This mathematics that Cornish had picked up was useful because he then had a similar background in mathematics to Fisher and they could understand each other very well. I take it that Cornish looked around the world for the best place to learn mathematics, especially algebra, and concluded that Princeton would be the best because, following the demise of the Mathematisches Institut at Göttingen from the attack on its faculty by the Nazis, the leading people had moved to Princeton University and the Institute of Advanced Studies at Princeton. I think that the recommendation of Schwerdtfeger must have been very influential in my gaining admission to Princeton. Students are something like plants, when you transplant them they take a while to put their roots down before they really flourish, but, in the case of Princeton, my training under Schwerdtfeger, and then later under George Szekeres, who came when Schwerdtfeger went to Melbourne, gave me a background of mathematics and an attitude very similar to the one I found at Princeton, so I found I could pick up the work at Princeton very quickly.

As a graduate student then at Princeton, did you see yourself primarily as an algebraist or a statistician?

I think my dominant interest was in mathematics but, as I had been sent there by the CSIRO Section of Mathematical Statistics, I felt I should pay considerable attention to statistics. In an orientation session in which the chairman of the department at Princeton [Lefschetz] had all the new students sitting on the floor of their Sanctum Sanctorum in Fine Hall, I put up my hand and said that I would like to concentrate on statistics. Lefschetz's reply was "James, at Princeton you are mathematician first and a statistician second."

Princeton was exceptionally strong in both mathematical statistics with Tukey, Wilks, and Feller and in pure mathematics with Bochner, and with Weyl at the Institute for Advanced Study. Was there much interaction between these groups (no pun intended), and how did these various members influence your own research?

Well Bochner tended to be something of a loner, somewhat off on his own in his analysis and differential geometry, but Tukey, who had been a topologist of course, got on very well with the mathematicians and Wilks made a point of doing so as well. In respect of Feller, I felt that there were people that were so good on the stochastic processes that he taught that I would prefer to work in a different field. Anyhow, Cornish had put great emphasis on the algebra of regression theory, which was my real interest. I had been trying to read the book of Weyl, *Classical Groups*, but found it very difficult to get beyond about the first two chapters until I took some lectures in algebra by Tate, who was an offsider of Artin, and from that I picked up the theory of algebras and that opened up the way to continue reading Weyl's *Classical Groups*.

In regard to Feller, although I did not undertake research in his stochastic processes, nevertheless I found him an excellent example of how a real mathematician goes about his business and found terrific help in how to approach mathematical problems from Feller. At his first lecture Feller said to the class, "Are there any topics in probability theory that you would particularly like me to deal with?" So I went to his office and said I was very interested in geometrical probability. This was really outside his field of interest but he did appreciate a student going to him and asking and he was always very helpful to me later on. In fact, when I started to write up the first paper from my thesis the editor sent it back and said it looked interesting but that it was too badly written. Feller offered to read what I wrote and criticize it and helped me with this very considerably too. There was also the differential geometer Albert Nijenhuis visiting and he helped me too. I always remember once when I took Feller something I had written, he wrote on it, "Don't be so stingy." Another time when I had started off in the conventional way about sigma-algebras, measures, and so on. Feller crossed all these out and said, "This is irrelevant in your particular instance, the people who know this can supply it themselves if they want it and the people that don't know won't be any worse off."

I attended lectures by Weyl at the institute during my third year at Princeton and presented a letter of introduction to him from Schwerdtfeger. I showed him my thesis and he returned it with a note expressing interest in it.

The paper that you mentioned just then is the classical 1954 paper "Normal Multivariate Analysis and the Orthogonal Group," which was part of your Ph.D. thesis. Can you tell us a little bit about the genesis of that paper and your thesis?

Well it really came about in connection with the M.Sc. thesis where I became interested in the simultaneous distribution of non-orthogonal sums of squares, that is to say, of two chi-squares that were not independent. When I was at Canberra, CSIRO could not find me any accommodation, all the hotels were full. So I went down to Sydney for a few days and I was in a little hotel in Sydney, thinking about this problem, and worked until about 2 A.M. and in effect rediscovered Jordan's results on critical angles between subspaces. At that stage, having solved the problem, I turned off the light and got into bed thinking I would go to sleep but I had a terrible time – I lost control of my thoughts. I don't understand psychology, but somehow one's will can direct one's thoughts in the direction one wants to go. In this case, I lost control of it completely and spent a very uncomfortable 2 hours tossing about in bed in a very strange mental state. Eventually I came to, enough to turn on the light, and as soon as I became conscious I went straight to sleep and woke up in the morning with the light on. Having found this result on critical angles, I recognized then that it was also related to Hotelling's canonical correlation and Cornish showed me grey notes of Wilks (1943) that later became his book and, in particular, the last chapter on multivariate analysis, which, in effect, had been written by Ted Anderson. So I read through these and became interested in the distribution of the canonical correlation coefficients. Then it occurred to me that one could have distributions of random planes and uniform or invariant distributions of random planes. When I mentioned this to Schwerdtfeger, he immediately pointed to Herglotz, who had produced invariant differential forms that could express this and these had been written up by Blaschke (1935). So I picked this up and then used these two as a new derivation of the null distribution of the canonical correlation coefficients and this, in effect, became the first part of my thesis at Princeton. In fact, I went over to Princeton with these results already. Wilks wanted me to write this up as a thesis and said this would be satisfactory, but Tukey said no, it would not be, it was not enough, it was not original enough. Of course several other people had derived them, both Fisher (1939) and Hsu (1939) simultaneously and, in fact, Mood at Princeton had also done so. Someone said that one should have seen Mood's face when he heard that Fisher and Hsu had scooped him in this! Anyhow, I then studied a book by Appell and Kampé de Fériet (1926) on hypergeometric functions and became interested in differential equations and the Laplace operator, and from that I found a way to get a power series expansion for the noncentral Wishart distribution, which became the second part of my thesis. The addition of this material to my thesis satisfied Tukey. Then in 1952 I heard that Fisher was visiting Storrs, Connecticut, so I got in the car and drove several other students up to spend a few days there with R.A. Fisher. We were just staying at the same college very informally and in a queue for dinner, Fisher made some remark, and I got into an argument with him, which he won of course. Then I made myself known and told him about this thesis and he offered to communicate it to the Royal Society of London, and so that is how the second half of the thesis became published on the noncentral Wishart distribution (James, 1955).

The "orthogonal groups" paper is one of the first in statistics that used the idea of decomposing matrices into the group of orthogonal matrices and the space of positive-definite symmetric matrices.

Corresponding to the matrix analog of decomposing vectors into polar coordinates, I decomposed the volume element by introducing exterior differential forms for invariant measures on the orthogonal group and its coset spaces, the Grassmann and Stiefel manifolds. This sufficed to give new derivations for central or null distributions where the multinormal density was invariant under the group.

This is just the matrix analog of decomposing vectors into polar coordinates. At a deeper level, invariance arguments are used to define equivalence classes, or orbits, and so you are able to integrate functions with respect to the invariant measure by first evaluating the integral for a given orbit and then averaging across all orbits, with the former step involving the idea of integrating over group manifolds.

For the noncentral or nonnull distributions, the multinormal density was not invariant and I had to use group representation theory to evaluate the integrals over the orthogonal group.

I believe that Stein and Karlin were working on techniques of orbits and invariants at Stanford at about the same time and, of course, the techniques are now widely used, particularly in invariance and ancillarity arguments. Was this something you could foresee at the time, and were you aware of Stein's and Karlin's work?

No, Karlin had been at Princeton but I was not aware of Stein's or Karlin's work at all. Rolf Bargmann remarked that I was doing a Hurwitz in reverse. Hurwitz (1897) developed the idea of generating invariants by integrating over the orthogonal group and, for this purpose, he used a parameterization to prove the existence of the integral. Haar (1933) generalized this to prove that there always existed an invariant measure on a locally compact topological group. By the existence of this measure, Weyl was able to establish much of the theory of group representations. Computation of integrals of specific functions was, however, very difficult, because any parameterization of the orthogonal group is highly unsymmetrical. As Bargmann pointed out, I used the theory the other way round and used knowledge of structure of groups to evaluate the integrals, thereby avoiding the complications of introducing a parameterization of the orthogonal group.

Your papers during the 1950's on the noncentral Wishart and related distributions are a little bit like a murder mystery, taking the reader from the initial formulation of the problem (which was your 1955 paper on the noncentral Wishart distribution), through to dénouement with the introduction of zonal polynomials in your 1960 paper "The Distribution of the Latent Roots of the Covariance Matrix" and your 1961 paper "Zonal Polynomials of the Real Positive Definite Symmetric Matrices." How much of the final answer were you aware of in 1954, and what were the highlights in your attack on the problem over this period?

I had no idea of zonal polynomials at that stage; I had just got on to the idea of differential equations based on the Laplace operator for obtaining recurrence relations to calculate the sequence of coefficients. It was rather surprising at that time that I was very close to later developments of differential equations satisfied by hypergeometric functions but Bochner rather put me off the track by pointing to matrices of determinantal operators developed by Lars Gårding, whereas, in fact, ultimately Angus Hurst of Adelaide put me on to the Laplace–Beltrami operator, which is used very much by physicists, which was really the basic operator.

When I came back to Adelaide, I was joined in a subsection of theoretical research by Graham Wilkinson and also Graham Constantine, who was appointed from Western Australia. Constantine was very strong mathematically—he had just worked through Montgomery and Zippin's (1955) book outlining how they had solved the Hilbert Fifth problem.

At the same time, I was still working on how to try to evaluate these integrals over groups, and I think that I'd had a lot of encouragement in continuing that work because there had been great interest in the statistics and mathematics departments at Princeton. Ted Anderson and others had been pushing problems very hard but had evidently needed a new approach to it. I thought the approach that I could follow then was through group theory, and so I continued to work on this. I started to see the Schur theory of the finite-dimensional representations of the linear group that related it to the symmetric group was related to the idea of group theory and experimental designs, so working through this to the idea of analysis of symmetric experimental design. I eventually worked through, in about 2 or 3 years, to producing the zonal polynomials.

How exactly did you find the zonal polynomials?

Wilkinson discovered that the matrix of a set of missing value equations for an incomplete block design has a pattern, because each element depends only on the relationship between the two plots to which it pertains, a relationship such as same block or same treatment. When he tried to use this pattern to obtain the inverse of the matrix in order to solve the missing value equations, I suggested that patterned matrices often belong to an algebra of low dimensions to which the inverse belongs. A basis of the algebra can be chosen from all products of subbasis elements given by the relations such as same block or same treatment. This led to the idea of "The Relationship Algebra of an Experimental Design" (James, 1957). I think perhaps that should have been published jointly with him really; anyhow, we did publish a paper later, James and Wilkinson (1971), which led to the idea of canonical efficiency factors.

The algebra has three bases. A relationship basis consisting of the matrices that generate the pattern, a statistical basis of the matrices of the quadratic forms, which are the sums of squares a statistician would naturally compute in analyzing the design and a canonical basis. If the algebra is commutative, the canonical basis consists of the idempotent matrices of the quadratic forms that appear in the analysis of variance table. A noncommutative algebra has multiplicities with intertwining operators in the canonical basis. This theory has been used by Brien et al. (Brien, Venables, James, and Mayo, 1984; Brien, James, and Venables, 1988) to analyze patterned correlation matrices.

At the same time as I was working with Wilkinson, I noticed similar patterns in matrices that commute with tensor representations of the general linear group, according to the Schur theory as described in Weyl's *Classical Groups*. They were the relationship matrices of an experimental design relating to a tournament in which each "plot" consisted of a pairing of $2f$ players into f singles, giving $(2f)!/(2^f f!)$ plots. The reason is that a polynomial of degree f in the elements of a vector is a symmetric tensor of rank f invariant under $f!$ permutations of the f indices. If it is a polynomial in the elements of a symmetric matrix, it has rank $2f$ and is also invariant under 2^f transpositions of the indices of the symmetric matrices. Out of curiosity, I found the idempotents of the relationship algebra of this design. When translated, via the Schur theory, back to the tensor representations of the linear group, they were the zonal polynomials. The results were published about 20 years later in James (1982). In this paper, I discussed the analysis of tensors of a given symmetry, such as the preceding, into the irreducible representations of the symmetric group that take place in them, and I developed an algorithm for calculating the corresponding idempotents.

At the time I knew of no use for the zonal polynomials in statistics, but a simpler result of Cartan led me to conjecture that the distribution of the latent roots of the sample covariance matrix would be a sum of products of zonal polynomials, a zonal polynomial of the population roots being multiplied by the same zonal polynomial of the sample roots.

The attempts to integrate over the orthogonal group by a parameterization of it, which inherently must be unsymmetrical, were replaced by an algebraic theory of projections.

When I wrote out the formulae for the zonal polynomials up to degree 4 in a letter to Constantine, he substituted them in an expansion for the nonnull distribution of the canonical correlation coefficients, which we had previously published (Constantine and James, 1958), and about 45 terms condensed down to 5.

Where did the name "zonal" come from?

The Peter–Weyl theorem (Peter and Weyl, 1927) states that the functions in all the irreducible orthogonal representations of a compact group constitute a complete orthonormal basis of the functions on the group. Cartan (1929) extended the result to the functions on a space that a compact group transforms transitively, such as the sphere transformed by the rotation group. They are called the spherical functions. For any point in the space, such as the north pole of the 2-sphere, the subgroup that leaves it fixed is called the isotropy or stability subgroup. This subgroup divides the space into orbits. In the case of the 2-sphere, the orbits are the parallels of latitude between which are the zones of the sphere. Functions that are constant on the orbits are called zonal functions. For the 2-sphere, they are functions of the polar angle or colatitude, θ.

The Archimedes theorem states that the area of a zone of a sphere is the area of the corresponding zone of the circumscribing cylinder, hence the cosine of the colatitude has a uniform distribution on the interval $(-1,1)$. Now, functions on inequivalent irreducible representations are orthogonal under integration of their product with respect to the uniform distribution. Hence, the zonal spherical functions of the 2-sphere are the orthogonal polynomials of $\cos(\theta)$, which are the Legendre polynomials.

For the n-sphere, the zonal spherical functions must be orthogonal relative to the null distribution of the ordinary correlation coefficient. They are the Gegenbauer polynomials.

The Grassmann manifold is the set of p-subspaces of n-space. Its zones are determined by the set of critical angles between two subspaces apart from order. It followed from the theory of Cartan and Weyl that there must be zonal spherical functions of the Grassmann manifold, but mathematicians thought they would be impossibly complicated to compute explicitly. Statistics, however, gradually pointed the way.

Functions on groups determine measures, and these can be convolved with respect to the invariant measure. The points of a space on which a group acts transitively correspond to cosets of the isotropy group. Functions on the space therefore correspond to functions on the group which are constant on the cosets.

Under convolution, a zonal spherical function is either an idempotent which projects on the irreducible representation to which it belongs or, if there is multiplicity of a representation, an intertwining function which maps group-isomorphically from one irreducible representation to an equivalent representation.

The sphere and the Grassmann manifold have some equivalent representations of the orthogonal group in common. The intertwining function must be a function of the cosine of the angle, θ, between a unit vector on the sphere and the subspace element of the Grassmann manifold. Intertwining functions corresponding to inequivalent representations must be orthogonal relative to the distribution of the $\cos(\theta)$, and this is the null distribution of the multiple correlation coefficient. Such polynomials are Jacobi polynomials.

Now if we have two Grassmann manifolds of p-subspaces and q-subspaces of n-space, the distribution of the cosines of the critical angles between them will be the null distribution of the canonical correlation coefficients. Intertwining functions of inequivalent representations must be orthogonal relative to this distribution. In James and Constantine (1974), we call them the generalized Jacobi polynomials.

How then did statistics play a role in finding these polynomials?

Well, when $n \to \infty$, the beta distribution of the multiple correlation coefficient tends to the gamma distribution of chi-squared. Correspondingly, the Jacobi polynomials tend to Laguerre polynomials that are orthogonal relative to the gamma distribution. Multivariate-wise, the multivariate beta distribution tends to the Wishart distribution in which Carl Ludwig Siegel at Princeton was interested in connection with lattice point theory, and Ted Anderson had had some discussions with him about this. Carl Herz produced some generalized Laguerre polynomials orthogonal relative to the Wishart distribution by taking Hankel transforms of determinants, but Constantine (1966) pointed out that Hankel transforms of all the zonal polynomials, not just determinants, are required for a complete set. They could be expressed as confluent hypergeometric functions of matrix argument with partitioned index.

Since every representation of the orthogonal group belongs to a unique representation of the linear group with unit multiplicity, James and Constantine (1974) were able to prove that the generalized Jacobi polynomial must begin with a zonal polynomial of maximum order. The radial part of the Laplace–Beltrami operator for the Grassmann manifold was then found, that is the part depending on the critical angles that the variable p-space makes with a fixed q-space. From it, one could find recurrence relationships to find the coefficients of lower order zonal polynomials.

Why do zonal polynomials play a key role in functions of matrices?

In the theory of analytic functions of a single complex variable, z, the powers, z^n, of z are the irreducible rational representations of the multiplicative group of nonzero values of z. Correspondingly, a power series theory of scalar functions of matrices requires the irreducible rational representations. For symmetric functions of the eigenvalues of symmetric matrices, these are the zonal polynomials, perhaps divided by powers of the determinant for rational functions.

Carl Herz, whom you just mentioned, was another key player in the early work in this area, in particular, on hypergeometric functions of matrix arguments. His 1955 paper "Bessel Functions of Matrix Argument" was also, I believe, part of his Princeton Ph.D. thesis. Did you know Herz well?

Yes, we started at the same time and were often in the same courses, but he was working with Bochner. I was working with, I think really pretty much on my own, with my own ideas. We only discovered at the end that we were working on very similar problems, about toward the end of the second year that I was there; just as I was writing up a thesis and he was writing up his, we suddenly realized that we were dealing with much the same problems. At that stage then I started to talk to Bochner, and Bochner, for example, pointed out that all you could do with the real orthogonal group you could do with the unitary group, too.

One of the breakthroughs that came in this work was that initially integrating over the orthogonal group I was concentrating on the orthogonal group itself and not looking into the structure of it, but then later on I realized that the positive-definite symmetric matrices correspond to the elements of a coset space of the linear group over the orthogonal group, and that the problem really lay in the linear group not in the orthogonal group. It took a long time to come to that idea.

Part of the work on the hypergeometric functions really involved the development of a theory of symmetric functions of the latent roots of a matrix variable. This theory is really much simpler than, say, just the theory of functions of two complex variables because it is much more like the theory of analytic functions of one complex variable. When one goes from one complex variable to two complex variables, you are going from the domain of these functions from an algebraically closed field, with all that structure, down to a mere vector space. And of course, if you cut down the assumptions, the theory becomes very much more general, so that the theory of two complex variables is a very difficult, very advanced theory, dealt with by Sheaf theory and all sorts of things. With symmetric functions of the latent roots of a symmetric matrix, one has an ordering of them according to their weights and something like a division algorithm – it's a thing that sometimes in maths courses they seem to miss in teaching symmetric polynomials. Weyl in his *Classical Groups* tried to avoid them because Elie Cartan tended to use them quite a lot, although he did have to do it at one stage because there is something you can't do without weights. I once mentioned it to Bochner and he said that weights are very difficult – you have to write about three papers on them before you really understand them. Anyhow, with the theory of analytic functions, the physicists tend to come at it with differential equations because that explains basic physical processes and that is how they developed the theory of so-called special functions of a single variable. The mathematicians, like Riemann, came at it via power series, and the statisticians, of course, came via integrals to obtain the marginal distributions of their sampling statistics. And that is how these functions arose in statistics, from the integrals needed to get the sampling distributions in statistics. Well, in effect, what Constantine and I did was to supply the power series expansions for these functions. Later on, when I told Herz we had power series expansions, he didn't believe it for a start. The differential equations and the integrals are useful in getting asymptotic expansions because, of course, the power series expansions are fairly inefficient for purposes of computation. As I mentioned, in the differential equations I was rather misled by Bochner for quite a long time in going to these determinantal operators of Lars Gårding, but then later I got back to the Laplace–Beltrami operator and Robb Muirhead, a graduate student largely supervised by Graham Constantine, worked those out fairly systematically and has written the whole thing up in his book (Muirhead, 1982).

Just out of interest, Herz doesn't really seem to appear in this literature after his 1955 paper. Do you know what became of him?

Well I spoke to him later on and he just said, "Oh I thought that was old hat."

After finishing your thesis at Princeton, you spent a further year at Princeton with an assistant researchship, and then you returned to Adelaide to go back to work at CSIRO. While you were at CSIRO you continued to do work on multivariate distribution theory. It was in this period that you wrote the "Relationship Algebra" paper that you mentioned earlier, a paper that also became a classic in quite a separate area of multivariate analysis. Can you tell us a little bit about this period at CSIRO and about the relationship algebra paper in particular?

Well, as I've mentioned earlier, the idea really arose from Wilkinson in connection with his M.Sc. thesis on analysis of orthogonal designs with missing values and his recognition that the coefficients depended on the relationship of the positions of the plots for which there were missing values, and that then immediately suggested to me the idea of algebras and relationships and so I developed this. At this stage, a paper I wrote on groups and analysis of variance was rejected by *Biometrika*—perhaps it was badly written. Later on when I developed these ideas further in connection with the analyses of variance determined by symmetry, using some fairly advanced algebra, the resulting paper was obviously not understood by the referees and they rejected it at both the *JRSS(B)* and the *Annals of Statistics*, and this was a very important link in the algebraic approach to zonal polynomials. And consequently, the absence of this paper made it very difficult for anyone else to follow how I had developed zonal polynomials from an algebraic point of view. Ultimately it was published in the Rao anniversary volume (James, 1982). By this stage the editor of *Multivariate Analysis* had felt that further papers on zonal polynomials weren't of much value and had a policy of not accepting them, but when I submitted this, as Rao had invited me, he was hardly in a position to refuse it and so the paper was ultimately published, then about 20 years later on. But I think the rejection of this paper certainly made it very difficult for anyone else to follow my train of thought.

In 1958 you left CSIRO and went across to the mathematics department at Yale, where you had students including Chris Bingham, George Anderson, and Graham Constantine, all of whom were to make important contributions to multivariate analysis. Do you have any particular memories of these and other students at Yale, and of Yale more generally?

Oh, oh, very well indeed. Kit Bingham, as he used to be called, was something of a protege of Chester Bliss. Bliss had a $\frac{9}{10}$ appointment at the Connecticut Agricultural Experiment station and a $\frac{1}{10}$ appointment at Yale University. And initially for his Ph.D. Bingham started to work on Fourier series analysis of weather data, particularly temperatures. But it rather appeared to me that one wouldn't be able to produce anything terribly original in this. So we went down to the geology department, which was just two doors down the road from the mathematics department, and found that they were very interested in the distributions of the layers of rock, of random planes, which was something of a variant from Fisher's idea of distributions on the sphere, and as a consequence Kit Bingham started to develop distributions of this form, which I think is now referred to as the Bingham distribution, and completed his Ph.D. on this topic. George Anderson worked on some asymptotic theory of the distribution of the latent roots of the variance matrix. Of course, Constantine had worked with me, along with Wilkinson, in the theoretical statistics group that Cornish had established at CSIRO. He came across to Yale about a year after me, in 1959. Although strictly I was his supervisor, we had been collaborators and had already had a joint paper published in the *Annals of Mathematical Statistics.* I should point out that the work of Graham Constantine, in developing series expansions for the hypergeometric functions of matrix argument, was helped tremendously by the work of Lowdenslager, who had produced a microfilm of the work of Hua Loo Keng. More generally, at Yale University I found the mathematical environment very stimulating with the colloquia that they had every week and also attended various of the advanced graduate lectures, like Einar Hille's lectures in hypergeometric functions and Kakutani's lectures in homogeneous chaos. Another experience I found at Yale University was the very relaxed atmosphere. For example, in CSIRO they have an elaborate internal refereeing system for papers, but when I went to Yale University I went to the chairman of the department and said that I had written up this paper and he just looked at me in surprise and said, "Why don't you submit it for publication?" Well, he wasn't interested, other than of course that Yale appears on the paper. And another time I went along and said that "I want to take two weeks off during the vacation." He just looked at me and said, "Well your time is your own."

You mentioned Kakutani. He was in the mathematics department at Yale then but is also familiar to econometricians in relation to fixed point theorems and the uniqueness of general equilibria. Did you have much contact with him?

I saw quite a lot of Kakutani. At one time he was something of a terror with the students. Once Einar Hille, who was director of graduate studies, asked me to examine a student in real variable because Kakutani had not turned up to the examination. I had almost passed him—he seemed to be doing alright as far as I was concerned—when Kakutani came along and took over and just tore him to pieces. I'm not sure that he failed him in the end, but he certainly asked him far more searching questions than I did. Another time, in a course of ergodic theory that Graham Constantine attended, he set lectures for the students to give and he set a terrifically difficult one for Graham Constantine. In two weeks he had to pick up the infinite-dimensional representations of groups, spectral decomposition of differential operators, and hyperbolic geometry. Well Graham Constantine worked day and night and put together a reasonable sort of talk but was fairly shaky on it. When he got up to give it, Kakutani just asked him terrifically searching questions and revealed that he didn't have a very deep understanding of these things, and poor old Graham had to go off and have about three double martinis afterward. But I always found Kakutani very friendly. When I first went there I shared an office with Oystein Ore and so I got to know him a little bit. He only needed a desk to put his lecture notes on to pick up; he had his main office away from the rest of the department.

In 1963 you and Frank Anscombe were the founding members of the new statistics department at Yale. Can you tell us how that came about?

Well, Yale wanted to develop its statistics from the late 50's and had a statistics committee; I believe at that stage Jim Tobin and Tjalling Koopmans in Economics were on that committee. They produced a position of Visiting Lecturer and inquired of people. Generally they didn't advertise the positions; they just inquired around and then invited people, and Sam Wilks recommended me. Jim Tobin thereupon rang Chester Bliss and said, "Do you know anything about this man James?" Well Chester Bliss didn't recognize me from having been at Storrs because he was too busy entertaining Fisher, and you could see the look of delight on Chester's face when he was entertaining Fisher, but Chester said, "Wait a minute, I'll look up my reprints," and thereupon he looked them up. The statistics secretary at Adelaide had dutifully sent off copies of reprints to various people, including Chester, and he suddenly saw two highly mathematical papers communicated by Sir Ronald Fisher. Chester immediately gave a very strong recommendation on this. I think he was afraid that people of the other school, Neyman–Pearsonites, might be appointed. So I was offered this position of Visiting Lecturer. I resigned from CSIRO to take this one-year appointment and set off with my wife and family for America. Then at a later stage, one day in the kitchen at home, I was busy preparing a lecture when the phone rang and a voice said, "This is the office of the President." I had never heard the term before and I asked who was speaking. This was Kingman Brewster and he was inviting me to serve on the statistics committee. At that stage I then recommended to them, and was backed up by Fisher, to set up a separate department from mathematics. They eventually appointed Frank Anscombe as chairman and I transferred from the mathematics department to the statistics department, where I was promoted from Associate Professor of Mathematics to Full Professor of Mathematical Statistics.

Just across the road from the statistics department was the Cowles Foundation, where there were a number of econometricians working on applications of multivariate methods through economic data. You obviously knew of Koopmans, but were you aware of this group and its statistical interests more generally?

Yes indeed. Tjalling Koopmans was chairman of the statistics committee and we collaborated very closely. In fact, Chester Bliss had started to set up a statistical lab in the Cowles Foundation, with desk calculators, which were used at that stage before computers. I immediately instituted a 2-hour statistical lab for my mathematics students, to whom I was teaching statistics, because I felt that they did not know anything about statistics unless they were able to do statistical calculations and analyze data. When Graham Constantine came over, he took up a position to supervise this, so we were in and out of the Cowles Foundation all the time and had very close relationships with Tjalling. One thing I remember of Tjalling Koopmans was at a party at Frank Anscombe's, just after the Three Mile Island nuclear accident. Frank Anscombe was a Bayesian, of course, and Tjalling Koopmans said with a quizzical smile on his face, "What was the prior probability, Frank, of this occurring?" Frank said that "Bayesianism does not always apply!"

Was there much academic interaction with the staff of the Cowles Foundation?

Not really. Tjalling Koopmans once tried to explain to me the importance of the work that Tobin was doing on the purchase of large items. Tobin's innovation, as I understood it, was to formulate a model specifying the probability at each time that the purchase would be made. Tobin was clearly held in very high regard by his colleagues. Tjalling Koopmans, who subsequently won the Nobel Prize in Economics, told me that Tobin was offered the Directorship of Cowles and invited to move from Yale to Chicago. He stayed at Yale and instead, a year or so later, Koopmans and the rest of the team moved from Chicago to Yale.

In conversations we have had previously, you also mentioned contact that you had with the electrical engineers at Yale.

Yes, part of the proposal for the Visiting Lecturer was conduction of a seminar in statistics, as they put it. When I did this on regression theory, I was very surprised to find electrical engineers appear, particularly Peter Schultheis and an engineer Tuteur. Later on, we discovered we were dealing with the same sort of problems, only using different terminologies and approaches. They were interested in signal and noise theory, and I have found, subsequently, the engineering approach quite a useful one to regression and to some extent one sometimes talks of random error as noise. It took us some time to reconcile our terminologies, but we did have a fruitful collaboration.

At the end of 1964 you left Yale to return to Adelaide, but on the way home you spent time at Cambridge as a Visiting Fellow of Churchill College. Could you tell us a little bit about how that came about?

Well that was a invitation by David Kendall. I had met him at Princeton in 1952 when he came as a Visitor invited by Feller. We sat together in some of Feller's courses at that stage and got to know each other quite well. He once said to me, "Are you interested in coming to England, Alan, or do you want to become a cousin of Uncle Sam?" Although Kendall was a top mathematical probabilist, he had inherited a department with a very strong tradition of very practical statistics. For example, although it was a department of pure mathematics and statistics, there was really very little interaction between the mathematicians and the statisticians which was in direct contrast to what I had known at Yale. Consequently, I didn't have a lot of interaction with the people there.

In 1965, you then returned to the mathematics department at the University of Adelaide, where you've been ever since. Sir Ronald Fisher had spent his last years there but passed away in 1962, so you didn't overlap. Had you known Fisher well?

Well, as I mentioned, I felt very familiar with Fisher from all of Cornish's anecdotes about him and, in a sense, had been introduced to Fisher's views of statistics by Cornish. Cornish used to say at the end of the '40's, "Well, of course, there is all this new Neyman–Pearson theory and maybe we ought to look into this," and then he would shake his head and say, "But Fisher doesn't seem to think much of it." At Princeton University, Erich Lehmann was a Visiting Professor in the beginning of 1951 and he gave a beautiful exposition, in mathematical terms, of the Neyman–Pearson theory. Later on, Erich admitted that some of the Neyman approach was not appropriate, particularly missing the ideas of conditioning that is necessary in statistics. Nevertheless, some of the mathematical development was excellent, and he has also written this excellent book on nonparametric methods. He was certainly a very fine mathematician and a real gentleman.

Some of your later work indicates that, unlike many statisticians, you seem relatively at ease with Fisher's fiducial arguments.

I've found most of Fisher's work, particularly in the ideas of likelihood and maximum likelihood conditioning and so on, very valuable, and I think that they have proved very valuable in practice. But fiducial theory together with the Behrens problem seem very difficult, and I don't think that I've achieved very much from the efforts I've put in to it. I have really come to the conclusion that one can spend one's time much more profitably on other topics. But when one comes to Fisher's ideas of likelihood, there's a very interesting history there. Chester Bliss visited Fisher in about 1933, on his way to work in the Soviet Union, and brought problems of probit analysis of dosage mortality curves for insects and, in particular, asked Fisher the question, "What do you do when you test 10 insects and they all die, the probit is infinity?" And I think Chester may have had an idea in his mind that perhaps you just take nine or something like that. To his surprise Fisher said, "Oh, you just have to leave that out." Chester thereupon said, "Oh if you only had nine it would have strong weight, and leaving it out changes this very considerably—it doesn't make biological sense." Fisher and, "Well if a biologist says that, it's up to the statistician to come out with something sensible," and Fisher thought about this and produced the idea of a working value in probit analysis and then that was generalized further to producing working values anyhow, in effect to get the maximum likelihood solution of the probit analysis. Well then Chester never got the chance to write a book about this topic; he was rather scooped by Finney (1947), who took up this probit analysis and maximum likelihood very strongly. And then from there onward it was computerized by John Nelder as GLIM and has led to the very important extension of regression, of analysis of variance, into generalized linear interactive models. So it's an interesting thing. Oh, Fisher published his idea as an appendix to Bliss's paper (see Bliss, 1935; Fisher, 1935a).

Another idea that came from Fisher was that Chester Bliss took to him the problem of estimating of the ratio of two normal expectations in connection with slope-ratio assay, I believe. And Fisher produced what has now become known as the Fieller solution and Chester published this. Fieller then wrote a paper in the *Royal Statistical Society* acknowledging Bliss and generalizing it somewhat, and so this original solution, which was really due to Fisher but published by Bliss, became known as the Fieller solution. Well then, Miss Creasy introduced what she thought was the fiducial solution to this, which was a different one, and Fisher had the awkward problem of disagreeing with someone who thought she was supporting the Fisher position and this led to the Creasy-Fieller paradox. But in effect I mean the Fieller solution is in

Fisher's *The Design of Experiments* (1935b), and Fisher always backed the Fieller solution, which in fact came from him originally.

Your zonal polynomial results, and in particular your 1964 paper "Distributions of Matrix Variates and Latent Roots Derived from Normal Samples," generated great excitement and an enormous volume of research in multivariate analysis over the next 25 years or so. How do you view that work and subsequent development of it now?

Well, I was a little deflated when Constantine informed me that Fisher was never at all interested in it, and I haven't followed it up myself really. I think that Rob Muirhead, in his book, has sort of taken the thing up and given a fairly definitive exposition of this. I have felt, however, with the increasing power of computers and developments of asymptotic methods, that in effect it may be possible to use this theory rather more in practice. I've hoped to see that for a long time but, in my experience, it was perhaps of more theoretical than practical interest, though I believe some others, particularly some economists, have picked it up and made good use of it. Mathematically, however, a theory of these functions is needed and power series is an integral part of any mathematical theory of functions.

In 1965 Ted Hannan wrote a survey of some of the material on group theory and its role in statistics that first appeared in the *Journal of Applied Probability* and was subsequently published as a short monograph. Do you recall this paper, and do you have any comments on it?

Yes, I thought it was a very useful contribution, putting all this material together in a single article.

A lot of the results that have been used in econometrics involved the invariant polynomials with multiple matrix arguments, and there are uniqueness problems associated with them.

Yes. Now that's been picked up by Bill Davis very much (Chikuse and Davis, 1986). The so-called nonuniqueness is a multiplicity of the representations. What happens with the zonal polynomials is that they are the zonal spherical functions of the positive-definite symmetric matrices, and this is a symmetric space in the sense of Elie Cartan, one of very few symmetric spaces, the sphere, the Grassmann manifold, and the positive-definite symmetric matrices. The spherical functions are the irreducible representations of the functions that belong to the matrices of the irreducible representations of the group, and the zonal spherical functions are the functions that were invariant under the isotropy group. There is only one of these in each representation, and each representation only occurs once for a symmetric space. So one gets a very unique situation. Beyond the zonal spherical functions are the spherical functions, and at this stage one starts to get more multiplicities, and so the thing groupwise is more complex. But just talking about some of that multiplicity, of course, going back to the idea of experimental designs and the group symmetry, one does get zonal spherical functions of finite groups, like the symmetric group. And this has been developed in the paper by Brien et al. (1988), as mentioned earlier. In it we've analyzed correlation matrices, to test for structure within correlation matrices, and in effect one does get a similar algebraic theory there of zonal spherical functions. Oh, except it's in respect to a finite group and one gets some of the multiplicities in that case. I find in some of the journals, the leading journals, that they cut down one's space very much, and I felt to some extent in some of those papers that the mathematics had to be emasculated somewhat. For example, I didn't have room even to mention that in effect one had zonal spherical functions and so on, that there is more in it.

One thing that you have mentioned a little bit in our discussion so far is that the 1971 paper seems to have roots in your earlier work on zonal polynomials as well. Do you have any more to say about that?

Yes. Oh, well, that's closely related to the relationship algebra paper. The paper on the variance information manifold is related to the differential geometry of the space of positive-definite symmetric matrices, which is not euclidean geometry but a case where statisticians have to go to a Riemannian geometry, and with a metric differential form, and one can develop theories for this of the geodesics and so on that have a natural statistical application. I had one graduate student, Mara Lee McLaren, who elucidated quite a lot of this theory for the geometrical interpretation of the space of positive-definite symmetric matrices, together with some of the Lie group theory associated with this statistical implications of the Lie group theory of the orthogonal group.

There seem to be hints in some of these ideas that you are hankering after a "justification" for statistical methods that is more fundamental, less model dependent or less tied to distributional assumptions. Is this fair inference?

Well my philosophy—for example in regard to regression theory, which has been one of my primary interests coming, as it were, from Fisher through Cornish—has been that the assumption of normality, homoskedasticity, linearity, etcetera, is an important case to study to develop the ideal theory in the simplest situation and then that makes a paradigm for what one then hopes to achieve where in practice these assumptions are violated. But how I rather picture the thing, the way one does this analysis on the computer now with any set of data, is that one makes all these assumptions and runs it through the computer, something like getting into a smart sports car and charging straight down the straight and narrow and, of course, very soon finding oneself in a ditch, because the assumptions have not held. Thereupon some people advocate that you should transfer to a more robust tracked vehicle, but my philosophy, which I hope reflects Fisher, is that there is no point in grovelling about in the mire of low likelihood; one gets out one's little spotlights, the regression diagnostics, and spies where the high likelihood ground is and then finds one's way back onto high likelihood ground. What one finds is that the high likelihood road has curved, but on the computer you tend to do things step by step; you do the obvious and when it goes wrong your diagnostics tell you how to correct this and progressively you do this. But I rather adhere to the idea of trying to obtain high likelihood inference via modifications of the specification to meet the exigencies of the data. And in regard to the Gauss theorem, that least squares gives the minimum variance linear unbiased estimator, this is the useful exact criteria to apply to much data which is not too far from normal so that least squares gives a good approximation to optimal analysis. When things go far enough from normal, though, with serious, say, asymmetry or outliers, then in that case linear estimates are in no way best and they just won't work in practice. I had been interested in a very speculative sort of idea of whether one could show that in a regression situation if data isn't normal then the best estimate isn't linear, but that seems a fairly difficult sort of thing to formulate exactly and prove.

On a related point, I noticed that you've been a regular discussant over the years of Royal Statistical Society invited papers on what might be called the philosophy of inference. Is this something that you have thought about a great deal?

Yes, I have thought a lot about inference, very largely under stimulation of Graham Wilkinson, who has been very interested in problems of inference. One of the ideas that has occurred to me that I did put in one of these discussions is that for precise statistical analysis one must answer a properly posed question; therefore, in the process of analysis of data, one has to work one's way through to the appropriate precise question and then a precise answer is possible as the correct statistical inference. But unless the right question is asked, and it may take some time to formulate the right question, the statistical inferential answer cannot be exactly the correct one.

Your successor at Yale in 1964, Jimmie Savage, was of course famously involved in these issues. Did you have the chance to discuss the foundations of statistics with him?

Well, just one Saturday morning we had a debate for quite some time and a very interesting debate. I found in informal debate that he was much more flexible and less didactic than when he was taking a public position. I think it was important to him to be very definite when he was making highly public pronouncements or he would be giving away his position as leader of the Bayesian school. But I wonder if Chester Bliss and perhaps myself, this is just very highly speculative, had some influence on him, in getting him to go back and take another look at Fisher, which he did in his very celebrated address to the Institute of Mathematical Statistics, "On Rereading Fisher" (Savage, 1976). Of course, as the publication of Fisher's correspondence shows, Jimmie Savage had been in correspondence with Fisher earlier, particularly about the problem of the Nile and questions of inference. But Fisher rather had a habit of answering mail fairly quickly, and if he didn't have time to answer a very difficult question, rather than simply being discourteous and leaving the mail unanswered, he'd pick a more superficial answer and, as the correspondence shows, this at some stage upset Jimmie Savage, who thought Fisher was being flippant with him. The attitudes of these people show up very interestingly in some of this correspondence that now is published (Bennett, 1990). Another time I had suggested a method that I had picked up from David Duncan, of regressing the value of means, like varieties, on the observations. It involves the problem that if you pick the top variety, in a varietal trial, the one that yields most, then there's a bias in that you don't expect that to yield as much the following harvest, and David Duncan had produced a suggestion for handling this that amounted, in a sense, to something like shrinkage. I told this to Chester Bliss and he was recounting it to Jimmie Savage. Jimmie Savage said, "That's Bayesianism," and Chester Bliss retorted, "If that's Bayesianism, I'm for Bayesianism!"

Although much of your published work is highly theoretical, your CV contains regular contributions to a diversity of much more applied problems. How important to your theoretical work have your applied interests been?

Oh, I think that applied work is as necessary to the rigor of statistics as theory is to the rigor of statistical applications. The applied problems define a very specific situation for statistical inference, in that of the myriad theoretical possibilities a certain finite set of problems is defined by that situation, and a lot of other problems can be ignored in that data. If it's departing from the simplest assumptions, it will be evident that it departs in one way and not another way. So, therefore, a specific practical statistical problem defines a unique situation for the theoretical statistician to analyze properly. And there's proof of the importance of the things that have been done because it has arisen in practice; one isn't merely inventing the sort of problems one can solve and producing rather artificial theory in that way.

I know that you have very strong views about the importance of the role of a statistician as part of a research group.

Yes, I think that, in many cases at least, the statistician should be a principal investigator within a research group, helping to define the aims of the investigation rather than simply being hired as a statistical consultant. I think that when statisticians can assume such a position of an equal, with other scientists, I think they can contribute much more to the result.

And you've expressed that through your interests in biochemistry.

Yes, yes.

Although your work is regarded by many as one of the key developments in distribution theory, you have never published a textbook on what many, even now, would regard as a difficult topic. You seem to have been happy to leave the task to others, like Farrell (1976, 1985) and your student Rob Muirhead (1982). Was this a deliberate decision on your part?

No. I meant to do it and, in fact, at Adelaide I did give a course over two terms on this theory and got out a set of notes. But then I was really slowed in going back to Adelaide and taking charge of statistics there, and fighting to get a separate department, and I never had the chance to develop it further. There's a lot of work in going from a set of notes to a properly produced textbook.

The course you taught at ANU has really stressed the role of linear algebra and group theory and being able to exploit structure in problems. Would you like to say more on that?

Well, in one sense, I felt that Cornish sent me to Princeton to learn linear algebra to do regression properly; this is one of the things that he foresaw and he wanted me to do. I'm often very slow to do things and it's taken about 40 years. What I've seen, in relation to some of the more advanced parts of regression, is that it requires a deeper mathematical foundation to handle it, rather than trying to gerry-build the mathematics along with the statistics as one goes along. And, in particular, abstract linear algebra and group theory are necessary to handle the more complicated formulae that start to arise, which emphasize the role that vectors and matrices are playing, rather than looking upon them as mere columns and arrays of numbers. When one starts to see the role that these vectors and matrices are playing, the formulae just start to fit together according to rules that are dictated by the basic mathematics, particularly the underlying group theory. In fact, all the matrix multiplications are tensor contractions of indices and the one index is a covariant index to be transformed cogrediantly and the index with which it contracts must then be a contravariant index, to be transformed contragrediantly. And the requirement of this symmetry means that these are the only contractions that make any mathematical, practical, or statistical sense. When one sees this mathematical structure, then the formulae almost assemble themselves in a sort of self-assembly, as if they were like proteins. Instead of one having to remember or deduce the whole thing, one can see in a sense that there is only one way the symbols can all fit together to make mathematical sense. So I think that it's important then for statisticians, when they start to attack more advanced topics in regression, to go and strengthen their mathematics with a little abstract linear algebra and group theory.

I think that statistics could profit considerably from group theoretical calculations to formulate better statistical models, more sensitive tests, and a clearer view of the structure of their problems. The mathematicians who have pioneered group theory and geometry have no motivation to go into detail, but statisticians have concrete problems arising from data that can be elucidated this way.

One thing that you were showing me the other day was the use of color as a way of displaying structure. Do you think that there is much scope for that as a device?

It seems to work fairly well. The mathematical physicists work out this system very largely of covariant and contravariant tensors by means of superscripts and subscripts, and for a mathematical physicist colleague I once wrote regression theory out in that form. It looked very curious, but it didn't seem satisfactory to me from the point of view of statistics. We need our matrices so that we can print out our arrays and look at them. And therefore, I have looked at the idea of using color instead to distinguish these, and it's very important in regression theory, for example, to be able to distinguish the different types of vectors that occur. For example, for the parameter vector, one may use one color for this and then the score that is a linear functional of it, a dual of it, and so one needs another color and these two colors will pair off in any formulae. And then for sample space, and for contrast space, and so in that way by representing these vectors in different colors students can grasp the different role that they are playing and come to recognize what role they have. Linear transformations of them then, of course, have to follow the rules of contraction, so unfortunately some of the matrices had to have two colors, one on the left and one on the right. But it seems to me a possibility to emphasize the inherent mathematical structure of this system.

Using colors in that sort of way is an obvious example of some fairly lateral thinking, and I know you have some views on the role of lateral thinking in science, as well.

Yes. I think that I have inherited limited mathematical ability but a considerable ability for lateral thinking. Lateral thinking requires independence, and this is a big problem for the scientific administrator, who is especially responsible for spending public money on the scientists, that they want to move them in the direction of important goals but at the same time if they deny them freedom then they won't do their lateral thinking needed to achieve scientific goals. I rather believe that this is a Hegelian dialectical dilemma for administrators of science. I was very fortunate indeed, too, for the most part, to have terrific freedom, especially as it was given to me by Cornish, to do this.

One person who has obviously been important in your professional life was Graham Wilkinson; we have mentioned him a number of times during the conversation. Looking at your CV, Bill Venables is another one whose name occurs fairly frequently. Do you have any thoughts you want to share there or on anybody else who you think is important who we've missed?

Well. Bill Venables came to Adelaide in 1966 as a Ph.D. student of mine. He took up a position of statistical consultant at a tutorship level, which gave him a very sound grasp of practical statistics. At the same time, he developed outstanding ability in statistics and computing, combined with considerable strength in mathematics. He has since developed an extremely critical mind and maybe he is too critical of himself for his own good. But I find him absolutely invaluable; in fact, sometimes I feel as I discuss my scientific ideas with him as if the roles have reversed, and he's now my supervisor criticizing my ideas. But I find the criticism very valuable because he often anticipates the criticism that I would otherwise get from referees.

Another notable statistician whom I have known since 1951 is Ingram Olkin, who tackled the same problem of transformation of multivariate distributions that I was working on via the wedge product, by Jacobians of matrix transformations (Olkin, 1953).

Finally, we spoke the other day about the limits of genius, and Fisher and fiducial arguments is the example that you used.

Yes. Well there's the Peter-Paul Principle that people will be promoted or work up to a problem that ultimately defeats them. Riemann hit his famous Riemann hypothesis, which has never been proved yet or disproved. I think that it's now recognized that it's fairly dangerous to be supervised by a genius because they are liable to give you a problem that they can't solve, an impossible problem, and somehow or other these very difficult problems have some attraction for them so that they tend to gravitate to them. There's only one other point, slightly unrelated to this, and that was that there's a very interesting book by Hadamard (1945), *An Essay on the Psychology of Invention in the Mathematical Field.* He discusses what makes real mathematical ability, and he dismisses memory and many other things, and comes down ultimately to an aesthetic sense. And I have felt this myself, that somehow where there's something possible in a line of investigation it has a mysterious attraction and that if one has the freedom to follow this then sometimes it will lead one to very interesting yet unexpected results. I might mention that the determinantal equations that appear in multivariate analysis seemed to have a terrific fascination for me, and I could feel that there was something very interesting in that and behind it. It sort of led me on to discover the zonal polynomials and the differential geometry and the existence of these things. Of course, some mathematicians knew of this differential geometry of positive-definite symmetric matrices, but they didn't realize there was any use for it – it was just a mathematical curiosity that they didn't bother to say much about.

Are there any other comments you would like to make?

Well, I just might mention one other thing in my life and that is my forebears were all self-employed, and I felt that in a sense I was the first member of the family who ever had to work for a boss. I couldn't have had a better one than Cornish, or as far as that goes the chairmen at Yale University, but it is difficult for a lot of people who formally could be self-employed, that they all have to fit within this hierarchy of society. Personally, it has seemed a very important thing. I was always brought up with the idea that the greatest thing in life was independence – it was worth more than riches or fame or anything else. More recently, I have realized that independence is a gift of society and that one should always endeavor to do something in return, though personally, I cannot claim to have given as much as I have received.

REFERENCES

Appell, P. & J. Kampé de Fériet (1926) *Fonctions Hypergéométriques et Hypersphériques*. Paris: Polynomes d' Hermite.

Bennett, J.H. (1990) *Statistical Inference and Analysis: Selected Correspondence of R.A. Fisher*. Oxford: Oxford University Press.

Blaschke, W. (1935) *Integralgeometrie. Actualités Scientifiques et Industrielles*, vol. 252. Paris: Herman.

Bliss, C.I. (1935) The calculation of the dosage-mortality curve. *Annals of Applied Biology* 22, 134–167.

Brien, C.J., A.T. James, & W.N. Venables (1988) An analysis of correlation matrices: Variables cross-classified by two factors. *Biometrika* 75, 469–476.

Brien, C.J., W.N. Venables, A.T. James, & O. Mayo (1984) An analysis of correlation matrices: Equal correlations. *Biometrika* 71, 545–554.

Cartan, E. (1929) Sur la determination d'un systeme orthogonal complet dans une espace de Riemann symetrique clos. *Rendiconti Circolo Matematico di Palermo* 53, 217–252. (Reprinted, in *Oeuvres Complete*, pt. 1, vol. 2, pp. 1045–1080, 1952. Paris: Gauthier Villars.)

Chikuse, Y. & A.W. Davis (1986) A survey on the invariant polynomials with matrix arguments in relation to econometric distribution theory. *Econometric Theory* 2, 232–248.

Choi, I. & P.C.B. Phillips (1992) Asymptotic and finite sample distribution theory for IV estimators and tests in partially identified structural equations. *Journal of Econometrics* 51, 113–150.

Constantine, A.G. (1966) The distribution of Hotelling's generalized T_o^2. *Annals of Mathematical Statistics* 37, 215–225.

Constantine, A.G. & A.T. James (1958) On the general canonical correlation distribution. *Annals of Mathematical Statistics* 29, 1146–1166.

Davis, A.W. (1979) Invariant polynomials with two matrix arguments extending the zonal polynomials: Applications to multivariate distribution theory. *Annals of the Institute of Statistical Mathematics* 31 (A), 465–485.

Farrell, R.H. (1976) *Techniques of Multivariate Calculation*. New York: Springer-Verlag.

Farrell, R.H. (1985) *Multivariate Calculation: Use of Continuous Groups*. New York: Springer-Verlag.

Finney, D. (1947) *Probit Analysis*. Cambridge: Cambridge University Press.

Fisher, R.A. (1935a) The case of zero survivors in probit assays. *Annals of Applied Biology* 22, 164–165.

Fisher, R.A. (1935b) *The Design of Experiments*. Edinburgh, London: Oliver & Boyd.

Fisher, R.A. (1939) The sampling distribution of some statistics obtained from non-linear equations. *Annals of Eugenics* 9, 238–249.

Haar, A. (1933) Der Maßbegriff in der Theorie der kontinuierlichen Gruppen. *Annals of Mathematics* 34, 147–169.

Hadamard, J.S. (1945) *An Essay on the Psychology of Invention in the Mathematical Field*. New York: Dover.

Hannan, E.J. (1965) *Group Representations and Applied Probability*. London: Methuen & Co.

Herz, C.S. (1955) Bessel functions of matrix argument. *Annals of Mathematics* 61, 474–523.

Hsu, P.L. (1939) On the distribution of the roots of certain determinantal equations. *Annals of Eugenics* 9, 250–258.

Hurwitz, A. (1897) Über die Erzeugung der Invarianten durch Integration. *Nachrichten von der königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische*, pp. 71–90. (Reprinted 1933, in *Mathematische Werke von Adolf Hurwitz*. Band II: *Zahlentheorie, Algebra und Geometrie*, pp. 546–564, Basel: Emil Birkhäuser & Cie.)

James, A.T. (1954) Normal multivariate analysis and the orthogonal group. *Annals of Mathematical Statistics* 25, 40–75.

James, A.T. (1955) The noncentral Wishart distribution. *Proceedings of the Royal Society (London), Series A* 229, 364–366.

James, A.T. (1957) The relationship algebra of an experimental design. *Annals of Mathematical Statistics* 27, 993–1002.

James, A.T. (1960) The distribution of the latent roots of the covariance matrix. *Annals of Mathematical Statistics* 31, 151–158.

James, A.T. (1961) Zonal polynomials of the real positive definite symmetric matrices. *Annals of Mathematics* 74, 456–469.

James, A.T. (1964) Distributions of matrix variates and latent roots derived from normal samples. *Annals of Mathematical Statistics* 35, 475–501.

James, A.T. (1982) Analyses of variance determined by symmetry and combinatorial properties of zonal polynomials. In G. Kallianpur, P.R. Krishnaiah, & J.K. Ghosh (eds.), *Statistics and Probability: Essays in Honour of C.R. Rao*, pp. 329–341. Amsterdam: North-Holland.

James, A.T. & A.G. Constantine (1974) Generalized Jacobi polynomials as spherical functions of the Grassmann manifold. *Proceedings of the London Mathematical Society* 28, 174–192.

James, A.T. & G.N. Wilkinson (1971) Factorization of the residual operator and canonical decomposition of nonorthogonal factors in the analysis of variance. *Biometrika* 58, 279–294.

Montgomery, D. & L. Zippin (1955) *Topological Transformation Groups*. New York: Interscience.

Muirhead, R.J. (1982) *Aspects of Multivariate Statistical Theory*. New York: John Wiley & Sons.

Olkin, I. (1953) Note on the Jacobians of certain matrix transformations useful in multivariate analysis. *Biometrika* 40, 43–46.

Peter, F. & H. Weyl (1927) Die Vollständigkeit der primitiven Darstellungen einer geschlossenen kontinuierlichen Gruppe. *Mathematische Annalen* 97, 737–755.

Phillips, P.C.B. (1989) Partially identified econometric models. *Econometric Theory* 5, 181–240.

Savage, L.J. (1976) On rereading Fisher. *Annals of Statistics* 4, 441–483 (with discussion, pp.483–500).

Weyl, H. (1946) *The Classical Groups, Their Invariants and Representations*. Princeton, NJ: Princeton University Press.
Wilks, H. (1943) *Mathematical Statistics*, lithograph. Princeton, NJ: Princeton University Press.

BIBLIOGRAPHY OF ALAN TRELEVEN JAMES

1954

James, A.T. (1954) Normal multivariate analysis and the orthogonal group. *Annals of Mathematical Statistics* 25, 40–75.

1955

James, A.T. (1955) The noncentral Wishart distribution. *Proceedings of the Royal Society (London), Series A* 229, 364–366.
James, A.T. (1955) A generating function for averages over the orthogonal group. *Proceedings of the Royal Society (London), Series A* 229, 367–375.

1957

James, A.T. (1957) The relationship algebra of an experimental design. *Annals of Mathematical Statistics* 27, 993–1002.

1958

Constantine, A.G. & A.T. James (1958) On the general canonical correlation distribution. *Annals of Mathematical Statistics* 29, 1146–1166.

1959

Warner, P. & A.T. James (1959) The dose-response relationship of the Ehrlich Ascites tumour. *British Journal of Cancer* 13, 288–301.
James, A.T. & B. Shorr (1959) Discriminant Functions: A Proposed Technique in Statistical Weather Prediction. Report under contract AF19(604)-3877 by Travelers Weather Research Center, Hartford, CT, on *Multivariate Statistical Analysis of Atmospheric Processes*, pp. 3–25.
James, A.T. & B. Shorr (1959) Canonical correlation as a tool in meteorological analysis. Report under contract AF19(604)-3877 by Travelers Weather Research Center, Hartford, CT, on *Multivariate Statistical Analysis of Atmospheric Processes*, pp. 26–51.

1960

James, A.T. (1960) The distribution of the latent roots of the covariance matrix. *Annals of Mathematical Statistics* 31, 151–158.

1961

James, A.T. (1961) The distribution of noncentral means with known covariance. *Annals of Mathematical Statistics* 32, 874–882.
James, A.T. & H.P. Papazian (1961) Enumeration of quad types in diploids and tetraploids. *Genetics* 46, 817–829.
James, A.T. (1961) Zonal polynomials of the real positive definite symmetric matrices. *Annals of Mathematics* 74, 456–469.

1964

James, A.T. (1964) Distributions of matrix variates and latent roots derived from normal samples. *Annals of Mathematical Statistics* 35, 475–501.

1966

Bliss, C.I. & A.T. James (1966) Fitting the rectangular hyperbola. *Biometrics* 22, 573–602.

James, A.T. (1966) Inference on latent roots by calculation of hypergeometric functions of matrix argument. In P.R. Krishnaiah (ed.), *Multivariate Analysis*, pp. 209–235. New York: Academic Press.

1967

Dyen, I., A.T. James, & J.W.L. Cole (1967) Language divergence and word retention rate. *Language* 43, 150–171.

1968

James, A.T. (1968) Calculation of zonal polynomial coefficients by use of the Laplace–Beltrami operator. *Annals of Mathematical Statistics* 39, 1711–1718.

1969

James, A.T. (1969) Tests of equality of latent roots of the covariance matrix. In P.R. Krishnaiah (ed.), *Multivariate Analysis II*, pp. 205–217. New York: Academic Press.

1971

James, A.T. & G.N. Wilkinson (1971) Factorization of the residual operator and canonical decomposition of nonorthogonal factors in the analysis of variance. *Biometrika* 58, 279–294.

1973

James, A.T. (1973) The variance information manifold and the functions on it. In P.R. Krishnaiah (ed.), *Multivariate Analysis III*, pp. 157–169. New York: Academic Press.

1974

James, A.T., G.N. Wilkinson, & W.N. Venables (1974) Interval estimates for a ratio of means. *Indian Journal of Statistics* 36, 177–183.

Parkhurst, A.M. & A.T. James (1974) Zonal polynomials of order 1 through 12. In H.L. Harter & D.B. Owen (eds.), *Selected Tables in Mathematical Statistics*, vol. 2, pp. 199–338. Providence, Rhode Island: American Mathematical Society.

James, A.T. & A.G. Constantine (1974) Generalized Jacobi polynomials as spherical functions of the Grassmann manifold. *Proceedings of the London Mathematical Society* 28, 174–192.

1975

James, A.T. (1975) Special functions of matrix and single argument in statistics. In R.A. Askey (ed.), *Theory and Application of Special Functions*, pp. 455–463. New York: Academic Press.

1977

James, A.T. (1977) Tests for a prescribed subspace of principal components. In P.R. Krishnaiah (ed.), *Multivariate Analysis IV*, pp. 73–77. Amsterdam: North-Holland.

1978

Venables, W.N. & A.T. James (1978) Interval estimates for variance components. *Canadian Journal of Statistics* 6, 103–111.

James, A.T. (1978) Assessing the accuracy of the maximum likelihood estimator: Observed versus expected Fisher information. Comments on paper by B. Efron & D.V. Hinkley. *Biometrika* 65, 484–485.

Wilkinson, G.N., W.N. Venables, & A.T. James (1978) The error in the Behrens–Fisher fiducial solution and its resolution. Conference Proceedings, International Statistical Institute, Manila, pp. 563–567.

1980

James, A.T. (1980) Liver redox resistance: A dynamic model of gluconeogenic lactate metabolism. *Journal of Theoretical Biology* 83, 623–646.

James, A.T. & W.N. Venables (1980) Interval estimates for a bivariate principal axis. In P.R. Krishnaiah (ed.), *Multivariate Analysis V*, pp. 399–411. Amsterdam: North-Holland.

1981

James, A.T. (1981) Tangential and curvilinear coordinates in nonlinear regression. In L. Endrenyi (ed.), *Kinetic Data Analysis: Design and Analysis of Enzyme and Pharmaco-kinetic Experiments*, pp. 51–57. New York: Plenum.

1982

James, A.T. (1982) A note on the linear relation between lactate redox potential and the hydrogen shuttle flux. *Journal of Theoretical Biology* 94, 129–133.

James, A.T. (1982) Analyses of variance determined by symmetry and combinatorial properties of zonal polynomials. In G. Kallianpur, P.R. Krishnaiah, & J.K. Ghosh (eds.), *Statistics and Probability: Essays in Honour of C.R. Rao*, pp. 329–341. Amsterdam: North-Holland.

1983

James, A.T. (1983) On the wedge product. In S. Karlin, T. Amemiya, & L. Goodman (eds.), *Studies in Econometrics, Time Series and Multivariate Analysis*, pp. 455–463. New York: Academic Press.

1984

Brien, C.J., W.N. Venables, A.T. James, & O. Mayo (1984) An analysis of correlation matrices: Equal correlations. *Biometrika* 71, 545–554.

1985

James, A.T. & R.A.J. Conyers (1985) Estimation of a derivative by a difference quotient: Its application to hepatocyte lactate metabolism. *Biometrics* 41, 467–476.

1987

James, A.T. & R.A.J. Conyers (1987) Response to a reader reaction to Conyers and James (1985). *Biometrics* 43, 245–248.

1988

Brien, C.J., A.T. James, & W.N. Venables (1988) An analysis of correlation matrices: Variables cross-classified by two factors. *Biometrika* 75, 469–476.

1989

James, A.T., J.T. Wiskich, & R.A.J. Conyers (1989) Statistical modelling of mitochondrial power supply. *Journal of Mathematical Biology* 27, 29–48.

Field, J.B.F., A.T. James, & W.N. Venables (1989) The influence of assumptions implicit in a model on parametric inference. *Journal of the American Statistical Association* 84, 101–106.

1990

James, A.T., J.T. Wiskich, & I.B. Dry (1990) A squared Michaelis Menten curve of substrate concentration for plant mitochondrial respiration. *Plant Physiology* 92, 265–267.

Soole, K.L., I.B. Dry, A.T. James, & J.T. Wiskich (1990) The kinetics of NADH oxidation by complex I of isolated plant mitochondria. *Physiologia Plantarum* 80, 75–82.

1993

James, A.T., J.T. Wiskich, & R.A.J. Conyers (1993) t-REML for robust heteroscedastic regression analysis of mitochondrial power. *Biometrics* 49, 339–356.

James, A.T. & W.N. Venables (1993) Matrix weighting of several regression coefficient vectors. *Annals of Statistics* 21, 1093–1114.

1994

James, A.T., W.N. Venables, I.B. Dry, & J.T. Wiskich (1994) Random effects and variances as a synthesis of nonlinear regression analyses of mitochondrial electron transport. *Biometrika* 81, 218–235.

James, A.T. (1994) Normal multivariate analysis of families of regression coefficient vectors. In T.W. Anderson, K.T. Fang, & I. Olkin (eds.), *Multivariate Analysis and Its Applications (IMS Lecture Notes—Monograph Series)* 24, 281–295.