On the Origins of Twistor Theory. - Roger Penrose

Castalia Francon

Roger Penrose

On the Origins of Twistor Theory

Gravitation and Geometry, a volume in honour of I. Robinson, Biblipolis, Naples 1987

1. INTRODUCTION

It is an especial pleasure for me to have this opportunity to pay my respects to my friend and colleague Ivor Robinson. I have chosen to hold forth on the origins of twistor theory for two reasons. The first is that Ivor's 60th birthday very nearly coincides with what is (for me) the 20th birthday of the theory, so this seemed to be a reasonable point at which to examine how it stands today, in relation to its original aims and aspirations. But also Ivor himself played a direct and important part in those origins. In fact, I can think of several essentially independent influences that Ivor had, one of which was quite crucial. I hope the reader will bear with me and forgive me for presenting an account which, to a large extent, consists of personal or technical reminiscences. But I hope, also, that there will actually be some scientific value in these ramblings.

I should first make clear what I mean, here, by the "origins" of twistor theory. I am referring to the origins of my own rather specific approach to a physical theory. I appreciate that many of the ideas go back very much farther than twenty years. Most particularly, Felix Klein put forth his correspondence between the lines in complex projective 3-space and a general quadric in projective 5-space as long ago as 1870 (Klein 1870, 1926), this correspondence being based on the coordinates of Julius Plücker (1865, 1868/9) and Arthur Cayley (1860, 1869), (or of Hermann Grassmann, even earlier). Sophus Lie had noted essentially the key "twistor" geometric fact that oriented spheres in complex Euclidean 3-space (including various degenerate cases) could be represented as lines in complex projective 3-space (contact between spheres represented as meeting of lines) already in 1869 (cf. Lie & Scheffers 1896) as was pointed out to me by Helmuth Urbantke some years ago. The spheres may be thought of as the t = 0 representation of the light cones of events in Minkowski space, so the Lie correspondence in effect represents the points of (complexified compactified) Minkowski space by lines in complex projective 3-space, where meeting lines describe null-separated Minkowski points - the twistor correspondence! The local isomorphism between the "twistor group" SU(2,2) and the connected component of the group 0(2,4) was explicitly part of the Cartan's (1914) general study and classification of Lie groups. The physical relevance of 0(2,4) in relation to the conformal motions of (compactified) Minkowski space-time had been exploited by Paul Dirac (1936 b) and the objects which I call twistors (namely the spinors for 0(2,4)) had been explicitly studied by Murai (1953, 1954, 1958) and by Hepner (1962). (See also Gindikin 1983 for a discussion of these matters.)

Moreover, as Ivor Robinson pointed out to me some ten or so years ago, a certain line-integral expression for representing the general (analytic solution of the wave equation in terms of holomorphic functions of three complex variables was known to Bateman in 1904 (see Bateman 1904 and 1944, p. 96), this having arisen from a similar expression due to Whittaker (1903) for solving the three-dimensional Laplace equation, and Bateman also gave a similar line-integral expression for solving the free Maxwell equations (Bateman 1944, p. 100). By a simple transformation of variables, these become the helicities zero and one cases of the basic contour integral formula (Penrose 1968, 1969a) giving the linear field case of the so-called "Penrose transform" of twistor theory. The Radon transform (Radon 1917, Gel'fand Graev & Vilenkin 1966) and its generalizations may also, from a different angle, be regarded as providing models for (and generalizations of) this twistor expression. In addition, the classic Weierstrass (1866) construction (cf. Darboux 1914) (which was known to me!) had provided a paradigm for the explicit solution, in terms of free holomorphic data, of an important non-linear problem (Plateau's problem). This may be regarded as a direct antecedent of the later non-linear twistor constructions for (anti-) self dual gravitational (Penrose 1976; cf. also Hitchin 1979) and Yang-Mills fields (Ward 1977, Atiyah & Ward 19771 Aityah, Hitchin, Drinfeld & Manin 1978).

Much of this previously existing material was not known to me twenty years ago. But the Klein correspondence was something I had been well acquainted with since my undergraduate days. So some might argue that there was not a great deal left to be original about in the basic twistor scheme. Nevertheless I do feel that I have a good claim to some sort of originality! This - if we discount a fair number of (non-trivial) later mathematical developments - lies primarily in the essential "physical idea" that the actual space-time we inhabit might be significantly regarded as a secondary structure arising from a deeper twistor-holomorphic reality. The basic idea, it could be argued, pro vides little more than a shift in viewpoint, but it is this shift that provides crucial motivation and it also gives, in a sense, whatever physic al content the theory has had, so far. This viewpoint has guided us in certain unexpected and often fruitful directions, providing some surprising mathematical insights and descriptions of basic physical fields and concepts. It has enabled us to achieve results that had not seemed possible to achieve by more conventional procedures. (For accounts of some of these, see Hodges et al. 1980). Nevertheless, twistors do not, as yet, provide a new physical theory in the usual sense that predictions - different from those given by conventional procedures are yet forthcoming. However, in order that the above "physical idea" should have genuine physical content, it must at some stage lead to a successful physical theory in this sense - or else be consigned to the dustbin (1) of scientific history!

TOP

2. SOME BACKGROUND IDEAS

Let me try to set in perspective my own state of mind some twenty years ago, and to explain some of the reasons why I felt that a different viewpoint with regard to space-time structure, of the kind provided by twistor theory, was needed. I had, for a good many years earlier, been of the opinion that the space-time continuum picture of reality would prove inadequate on some small scale. I do not propose to discuss all the reasons for this - and in any case it is a view that is hardly original with me. Indeed, that the quantum nature of reality should affect the very structure of space-time at some scale is now a more-or-less accepted viewpoint among those physicists who have examined this question in some depth (cf. Schrödinger 1952, Wheeler 1962). But I think that most physicists would believe that such effects should be relevant only at the absurdly small quantum gravity scale of 10^{-33} cm. (or smaller). My own attitude was somewhat different from this. While it might be that only at 10^{-33} cm is it necessary to invoke a description of space-time radically removed from that of a manifold, my view was (and still is) that even at the much larger levels of elementary particles, or perhaps atoms, where quantum behaviour holds sway, the standard space-time descriptions have ceased to be the most physically appropriate ones, and some other picture of reality, though at that level equivalent to the space-time one, should prove to be the more fruitful. The very fact that quantum behaviour is so hard to picture in the normal way had seemed to me to argue strongly that the normal space-time picture of things, even at that level, is inappropriate physically. Indeed, there was nothing really new in this either, inadequacy of space-time pictures being very much a part of the standard quantum-mechanical philosophy. However, I felt that one ought to try to be more positive than this, in actually providing a picture of objective reality, albeit one perhaps radically different from the usual one.

Space-time descriptions of the normal kind can, of course, be used at the atomic or particle level provided that the quantum rules are correctly applied, and they have implications that are extraordinarily accurate. Thus, this new geometrical picture must, at that level, be mathematically equivalent to the normal space-time picture - in the sense that some kind of mathematical transformation must exist between the two pictures. However, the new description ought to incorporate quantum behaviour more readily and naturally than the old. Moreover, at the quantum gravity level of 10^{-33} cm, or at the level of space-time singularities, it ought to provide an essentially different and more accurate picture of things.

TOP

3. MOTIVATION FROM SPIN-NETWORK THEORY

These considerations were among those that motivated the twistor approach, but the extent to which these particular aims have been actually fulfilled remains somewhat problematical as of now. Another initial motivation whose present status is unclear is that from spin network theory (originating in about 1958 and published later: Penrose 1969b, 1972b; cf.Moussouris 1983). I had had the idea, from the time I was a Research Fellow at Cambridge in the late 1950's, that the discrete combinatorial rules obeyed by quantum-mechanical (total) angular momentum could be used as a starting point for budding up space time structure. Though grounded in accepted quantum-mechanical principles, spin-network theory provided a picture of space which is entirely combinatorial in nature, so long as only a finite number of particles (or "units") are involved. Only in the limit, when this number becomes infinite, did a continuous picture of space arise. (These ideas were stimulated to a large extent by long discussions in Cambridge with Dennis Sciama about Mach's principle: "What happens to the concepts of space and direction if all the matter in the universe is removed save a small finite number of particles?") The spin-network scheme worked well enough (and in some respects surprisingly well) for a non-relativistic scheme in which directional concepts - though not spatial displacements - arise. The scheme was based on the representation theory of SO(3) (or, more correctly, of SU(2)), and for a more complete relativistic scheme, which included also spatial displacements, it had seemed that the (restricted) Poincaré group should replace SO(3) in the discussion. I had been unable to overcome certain obstacles to this (my former student John Moussouris having only quite recently been able to deal promisingly with them, cf. Moussouris 1983) and I had largely set the question of a direct generalization of spin-network theory aside.

However, in a slightly indirect way, spin-network theory did have a conceptual influence on the development of twistors. Moreover, there is perhaps a little irony in a certain "false start" which occurred in the winter of 1959/60, while I was at Princeton. I had felt that, in the context of the development of a suitable relativistic spin-network theory, it might prove fruitful to examine the total momentum-angular momen tum structure of a relativistic physical system in the cases where the 4-momentum was a null vector, the view being that this was more primitive than the standard timelike case. It seems that I had effectively worked out what was needed, and though the resulting objects are actually essentially twistors, the proper realizations had not then come to me and I did not develop the ideas further at the time. Only towards about 1970, long after twistors had been subsequently developed by quite another route, did I realize that they had also this null momentum angular momentum interpretation (cf. Penrose & MacCallum 1972).

As things have worked out so far, twistor theory has not moved much in the "combinatorial" direction of spin-networks. Instead, the (seemingly) very different complex-analytic aspects of twistors have been the ones that have proved to have greatest importance. The one place where the possibility of a connection with spin-network theory remains fairly strong is in twistor diagram theory (Penrose & MacCallum 1972, Penrose 1975a, Sparling 1975, Hodges & Huggett 1980, Hodges 1983, 1984) and in a certain sense it has been SU(2,2) rather than the Poincaré group which has so far replaced SO(3) in the discussion. Much more work will be needed to discern whether these relationships are more than superficial.

TOP

4. HOLOMORPHICITY IN CLASSICAL SPACE-TIME STRUCTURE

Other motivations than these played much more direct roles in the development of twistor theory. Spin-networks notwithstanding, the role of complex numbers in quantum theory had long struck me as a quite crucial one. If the "correct" geometry for the world is to be a closely quantum one, then these same complex numbers must be an essential part of this geometry. My training as a (largely pure) mathematician had taught me something of the power, subtlety and elegance of com plex (holomorphic) geometry. It had seemed fitting that this might be the geometry most basic to the structure of the physical world. Yet in its most obvious manifestations, physical geometry seems to be geometry over **R**, not **C**.

Nevertheless certain hints of a complex underlying structure had been apparent to me for some time. The fact that the null directions at a point have the holomorphic structure of a Riemann sphere (cf. Penrose 1959) had long impressed me, this fact being closely related to the complex nature of Lorentzian spinors. It had, for a long time, seemed to me that spinors, and particularly Lorentzian 2-spinors, are more fundamental than Minkowskian world-vectors, and that the latter should be regarded as derived from the former (2). (Compare Rzewuski 1958). In addition to this, complex numbers often have significant roles to play in solutions of Einstein's vacuum equations. I had been particular ly impressed by the nature of the plane-fronted waves that Ivor had introduced me to (these being due originally to Brinkmann 1923, but later rediscovered by Ivor Robinson, cf Robinson 1956) and also of their later generalizations to spherically fronted waves - due to Robinson and Trautman (1962). For these waves the behaviour along the null geodesics of propagation is fixed and there is completely arbitrary varia tion from wave-front to wave-front. But within each wave-front the strength and polarization of the wave is governed by a single arbitrary holomorphic function. What had struck me was the direct appearance of a free holomorphic function in the solution, the modulus and argument of this function both playing a direct role (as strength and polarization, respectively). This kind of feature had led me to believe in some role for holomorphic structure lying "behind the scenes" in solutions of Einstein's (vacuum) equations.

TOP

5. THE POSITIVE-FREQUENCY CONDITION

All this had concerned a possible role for "holomorphic structure" governing space-time even at the classical level. But I was searching for something in which quantum mechanical ideas arose in a unified way in relation to space-time geometry. It had been Ivor's- (and my) good friend Engelbert Schücking who had impressed upon me, at Syracuse N.Y. in the spring of 1961, the fundamental importance of the positive frequency condition in complex solutions of field equations, as the hallmark of quantum field theory. He had also persuaded me of the significance of conformal invariance, in relation, most particularly, to the massless fields of each spin (cf. Cunningham 1910, Bateman 1910, McLennan 1956; fields defined by Dirac 1936a). In my attempt to come to terms with the positive-frequency concept, especially in the context of general relativity where Fourier analysis does not find a comfortable home, I had been led to appreciate the value of formulations where this concept is expressed in terms of analytic continuation into complex regions of space-time. An added bonus was that this type of formulation can readily be manifestly conformally invariant, whereas the momentum-space approach afforded by Fourier analysis is particularly obscure in this respect.

TOP

6. MASSLESS FIELDS

I had been struck by what had tantalizingly seemed to be possibly a similarity (or relationship) between the analytic continuation properties needed for the positive-frequency condition and a certain generalization, to massless fields of arbitrary spin, of the Kirchhoff-D'Adhemar integral expression for solving the wave equation. I had found this generalization a short time earlier, probably while at Princeton in 1960 (cf. Penrose 1980, Newman & Penrose 1968, Penrose & Rindler 1984) and its conformal invariance properties had, I think, become apparent to me before the autumn of 1963.

I think that I had very much come around to the view that massless particles and fields were to be regarded as more fundamental than massive ones. My own reasons for believing this were probably very much bound up with the idea that 2-spinors should be regarded as more fundamental than Minkowski world-vectors, since it is a null vector (the 4-momentum of a massless particle), rather than a timelike one, which arises naturally from a single spin-vector of the 2-spinor formalism (cf. Penrose & Rindler 1984). Moreover, I had been particularly struck by the elegant properties of the massless free fields of arbitrary spin (in Minkowski space **M**), the conformal. invariance properties that Engelbert had drawn my attention to, having contributed considerably to my appreciation of this elegance.

It had also seemed to me that massive particles and fields should probably be, in some sense, built up from massless ones. The van der Waerden form of the Dirac equation is rather suggestive of the mass playing a role as a coupling between two Dirac-Weyl neutrino-type fields (cf. Penrose 1968 - but this idea is undoubtedly not original with me). In the summer of 1961 I had generalized my Kirchhoff-D'Adhernar type expression to handle the Dirac equation looked at in this way, and in 1962, the Maxwell-Dirac equations. The formulae were all sums of integrals over spaces of zig-zag and forked collections of null straight segments (Penrose & Rindler 1984). Thus, in this approach, all fields, whether massless or massive, are viewed as propagating along fight cones, and I liked to take the view that, in some sense, only the null directions were really "there"!

In accordance with this, it had seemed to me that it was important to search for some kind of background formalism, in which massless fields subject-to the positive-frequency condition of quantum field theory would play some sort of primitive role.

Moreover, the geometric framework of such formalism ought to be understood.

TOP

7. COMPACTIFIED MINKOWSKI SPACE (COMPLEXIFIED)

I had already thoroughly explored the structure of conformally compactified Minkowski space in the spring of 1962 (cf. Penrose 1965, the much earlier work of Bôcher 1914, Coxeter 1936, Kuiper 1949, Rudberg 1958 and, no doubt, Lie, Möbius, Cartan and others, having been, in detail, unknown to me at the time, though certain crucial ideas of "inversive geometry" had filtered through to me) and I had become convinced that some form of unified (but convoluted?) complex view point should exist, whereby the analytic extensions needed for the positive-frequency condition, the generalized Kirchhoff-D'Adhemar formula, null lines, spinors and the holomorphic structure of the space of null directions, the appearance of holomorphic functions in solutions for (linearized) gravity, the conformal invariance properties and the compactification, would all be among its manifestations. Apparently Engelbert Schücking had himself been thinking along somewhat related lines since he had repeatedly hinted to me, in the autumn of 1963 in Austin Texas (where he, I, Roy Kerr, Ray Sachs, Jerry Kristian and others had joined Alfred Schild's Relativity Center) that some kind of alternative "complex view" of space-time might prove fruitful.

Of course the possibility of simply describing things in terms of complexified (compactified) Minkowski space **CM** had occurred to me but - for reasons which are still not entirely clear to me - I had (correctly) *rejected this as insufficiently subtle for Nature. I think that one reason for being unhappy with **CM** as playing a primary role in physics was that the complexification is far too gross. As many additional "unseen" dimension (namely four) would need to be adjoined as are already directly physically interpretable (3).

It had seemed to me that one needed to find some higher dimensional analogue of the way that the real axis divides the complex plane into two halves, functions holomorphic in one half providing the positive frequency functions on **R** and those holomorphic on the other half, the negative frequency functions. Minkowski space **M** forms a tiny 4-real-dimensional subspace of the 8-real-dimensional manifold **CM** and therefore is, by itself, incapable of dividing **CM** into two halves. I had been aware of the two "tube" domains $\mathbf{CM}^+$ (the forward tube - given by position vectors whose imaginary parts are past-timelike) and $\mathbf{CM}^-$ (the backward tube - imaginary parts future-timelike) into which the positive frequency and negative frequency fields, respectively, extend holomorphically. But **M** is not a boundary of these domains (in the ordinary sense) and there is also a large open region in **CM** (imaginary part spacelike) lying "between" $\mathbf{CM}^+$ and $\mathbf{CM}^-$. Fortunately I was not aware, at the time, of the concept of Shilov boundary - the sense in which **M** can be regarded as a kind of boundary of each of $\mathbf{CM}^+$ and $\mathbf{CM}^-$ - for if I had been, I might perhaps have been tempted, after all, in that kind of direction!

So I had this feeling that some sort of complex space was needed, which somehow fell naturally into two halves, the common boundary between these two halves being, like the real axis of the complex plane, the more readily physically identifiable "real" part of the space. In this way, a good part of the holomorphic structure that seemed already to be "visible" in relativistic physics might be interpretable directly in terms of the induced partial complex structure (or CR-structure, as I learned later) on this common boundary. I had no idea what this complex space was, nor, indeed, whether or not its very existence might be merely a product of wild fantasy.

TOP

8. ROBINSON CONGRUENCES AND TWISTORS

It was in such frame of mind that I recall being driven (4) from San Antonio to Austin Texas (by Pista Oszváth) following a weekend family outing with Rindlers and Oszváths, the others following in a later car. I began thinking about an ingenious construction that Ivor had, some while earlier, described to me. He had been endeavouring to find null (real) solutions of Maxwell's free-space equations which were everywhere non-singular and of finite total energy (5). His procedure was to base such a solution on a certain twisting shear-free congruence of rays (null straight lines) in **M** - now referred to as a Robinson congruence. To construct Ivor's congruence consider first the family (congruence) of rays x meeting a given ray q. This is a special Robinson congruence. The tangent directions to the various rays x at the various points of each x provide a shear-free geodetic field of null directions which is, however, singular along q. Next perform an arbitrary complex translation, or indeed any complex Poincaré transformation, to this configuration (complexified). At each point of **M** we now have a complex null direction. But a complex null direction determines two real null directions. In spinor terms (cf. Penrose & Rindler 1984), if the complex null direction is that of $a^{A}b^{A'}$, these two real null directions are those of $a^{A}a^{A'}$ and $b^{A}b^{A'}$. Select (say) the former, consistently. We thus have another real field of null directions on **M** which again turns out to be geodetic and shear-free. But now this field is (in general) everywhere non-singular (and twisting). The family of rays whose tangent directions constitute this field is a Robinson congruence.

It is reasonably clear from the construction that Robinson congruences constitute a holomorphic family (i.e. the manifold whose points are the different Robinson congruences is naturally a complex manifold) essentially because they arise from general holomorphic motions of **CM**. Among these congruences are the special Robinson congruences, each of which may be viewed as a way of describing a particular ray q. The general Robinson congruences describe, in a sense, "complexified" rays.

I had, somewhat earlier, worked out the geometry of a general Robinson congruence: in each time-slice t=const. of **M** the projections of the null directions into the slice are the tangents to a twisting family of linked circles (stereographically projected Clifford parallels on S^4 - a picture with which I was well familiar), and the configuration moves with the speed of light in the (negative) direction of the one straight line among the circles. (See fig. 1).

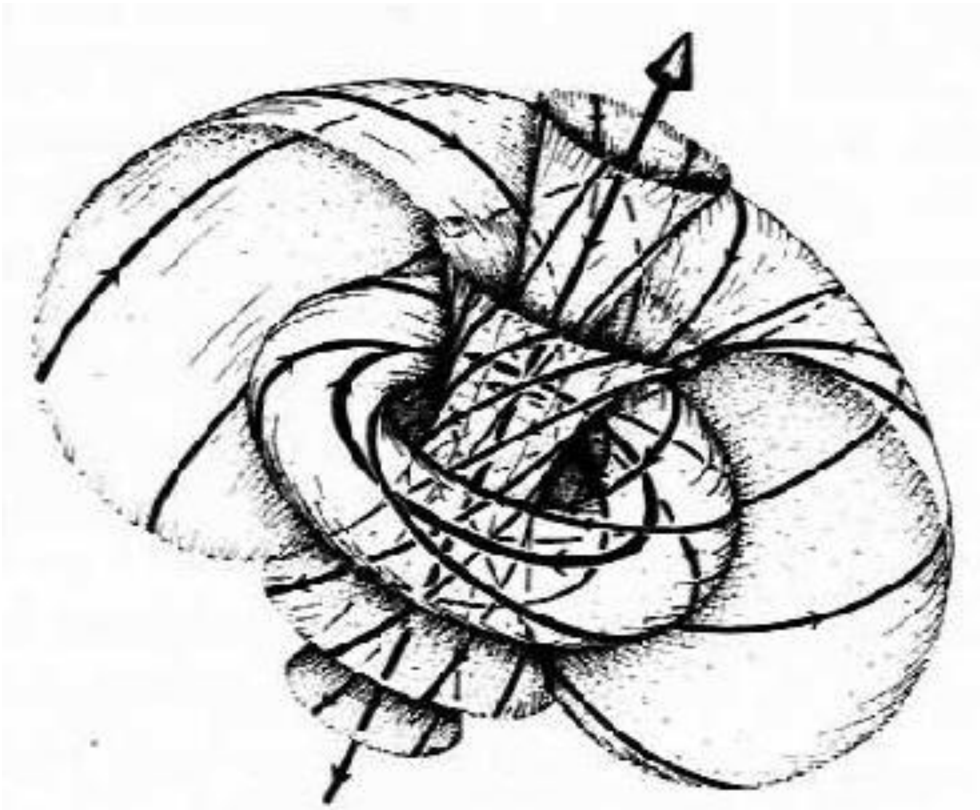

FIGURE 1: A time-slice (t=0) of a Robinson congruence.

I decided that the time had come to count the number of dimensions of the space *R* of Robinson congruences. I was surprised to find, by examining the freedom involved in fig. 1, that the number of real dimensions was only six (so of only three complex dimensions) whereas the special Robinson congruences, being determined by single rays, had five real dimensions. The general Robinson congruences must twist either right-handedly or left-handedly, so *R* had two disconnected components R^+ and R^-, these having a common five-dimensional boundary *S* representing the special Robinson congruences. The complex 3-space of Robinson congruences was indeed divided into two halves R^+ and R^- by *S*.

I had found my space! The points of *S* indeed had a very direct and satisfyingly relevant physical interpretation as "rays", i.e. as the classical paths of massless particles. And the "complexification" of these rays led, as I had decided that I required, to the adding merely of one extra real dimension to *S*, yielding the complex 3-manifold **PT** = $S \cup R^- \cup R^+$.

After returning home (in Austin) I was able to employ the spinor techniques that had long been familiar to me and it did not take me long to realize that **PT** was indeed a complex projective 3-space ($\mathbf{CP}^3$) the lines within which having **CM** as their Klein representation. This was perhaps a somewhat roundabout route to the picture of **CM** as a Klein quadric. I had been well aware for some time that **CM** was indeed a non-singular complex 4-quadric, and for considerably longer that non singular complex 4-quadrics are always Klein representations of lines in some $\mathbf{CP}^3$. But without such a roundabout route, would the "absurd" thought of combining these two ideas ever have occurred to me before? (Curiously, the answer to this question is "yes". Apparently I had done this about nine months earlier, but the odd-looking reality conditions on the Cayley-Plücker coordinates and lack of physical motivation at that time had put me off. So I completely forgot about it until going though my notebooks again not many days ago!)

The various physical motivations that I had collected together were, indeed, crucial for me, in order that the necessary ingredients of this strange idea should come to me and, more importantly, take hold.

TOP

9. THE KERR THEOREM

It is one of those odd coincidences, which seem to happen more frequently than they should, that a day or so later I overheard Roy Kerr (who had an office on one side of mine, Engelbert Schücking having had the office on the other side) explaining his method of obtaining all (analytic) shear-free geodetic null congruences in **M** to Ray Sachs. I had him explain it to me and was amazed to find that the coordinates that came up naturally in his construction were precisely those "twistor coordinates" which I had just found as the projective coordinates for **PT** (projective twistor space). So almost at once a new role for my complex space was at hand: the general (real-analytic) shear-free ray congruence in **M** is provided by the intersection of **PN** (= S, the space of rays in **M**) with a general complex-analytic (holomorphic) 2-surface X in **PT** (see Penrose 1967).

Indeed, this had further ramifications. It was not long before I realized that another of Ivor's results had an immediate interpretation in terms of **PT**. He had shown that the general (real-analytic) null solution of Maxwell's equations could be described in terms of a general (analy tic) shear-free geodetic congruence, with just the additional information of one free holomorphic function F, of two complex variables. In effect, the variables labelled the rays of the congruence, and in **PT** the function F turned out to be holomorphic on X. So one was presented with a very neat "holomorphic" description of null Maxwell fields in **PT**.

TOP

10. GENERAL MASSLESS FIELDS IN M

This elegant picture generalized in an obvious way to null massless fields of higher (integral or half-integral) spin and, in particular, as Ivor noted, to linearized Einstein fields. So at least in the linear limit, my motivation from the "holomorphic" structure of plane-fronted and spherically fronted waves (both of which are indeed null) was satisfied. But I needed more than this. My motivations required than general solutions of the massless free field equations be simply representable in terms of twistors.

I tried many formulations, but these, being effectively only straight forward transcriptions of the space-time fields and equations, lacked sufficient elegance and naturality. Functions of Cayley-Plücker coordinates presented some awkwardness because of the identical relations they satisfy - though I found a reasonable enough formulation in these terms. The alternative seemed to be two-point functions in **PT** and this, also, seemed not quite what was needed. It had occurred to me that the generalized Kirchhoff-D'Adhemar expression somewhat resembled a Cauchy integral formula, and I had it in mind that, in some sense, the massless free field equations ought to be the Cauchy-Riemann equations in disguise. "Would it not be most fitting" I had thought to myself "if a massless field could be described by a single free holomorphic function f on **PT**?" The freedom in the solutions that would occur should be just right. **PT** (for holomorpbic information) counts as three dimensional, which is the same dimensionality as an initial data surface in **M**, from which fields would uniquely propagate. The hope, then, was that the field equations should simply evaporate!

Perhaps this was just a pipe-dream - but in due course the thought occurred to me (not until 1966, in London, though!) that, in the null case, Roy Kerr's surface X in **PT** should be treated as a pole of the putative "f", where Ivor's holomorphic function F would be its residue, the field being obtained by contour integration of f. (It was the presence of certain singularities in the field, and the fact that these would duly result from pinching of such contours - owing to a tangency with X - which suggested this.) The formula (Penrose 1968, 1969a) then dropped out. It was many years later that Ivor (and Lane Hughston) pointed out to me that Bateman had, in effect, found essentially the same formula over sixty years earlier! My own formulation did, however, have the advantage that geometric and conformal transformation properties of the field were made manifest, and also connections with the (above) Robinson-Kerr theorem, etc., were exhibited.

TOP

11. COHOMOLOGY

But what about the positive-frequency condition? I had been motivated by the hope that positive frequency fields were to be associa ted with holomorphic extension into **PT⁺** (=*R*⁺, the space of right handed Robinson congruences), and negative frequency fields, extension into **PT⁻** (=*R*⁻, the space of the left-handed Robinson congruences). The condition turned out, instead, to be that f had separated singularity sets in **PT⁺**, or **PT⁻**, according as the resulting massless field was to be of positive or negative frequency. What had happened to my "crucial" motivation that **PT** needed to be naturally divided into two halves by **PN** (=*S*)?

The answer had to wait ten years (until the spring of 1976), after 1 had acquired some appropriate re-education from Michael Atiyah! It turned out that f was to be interpreted as a representative cocycle for an element of a holornorphic (first) sheaf cohornology group. I call such an element a holomorphic l-function. Then, the 1-function does indeed extend globally to **PT⁺** or **PT⁻** according as the resulting massless field is of positive or negative frequency.

I have never been able to decide whether I was actually "insightful" here or just "lucky"! I had come to feel that there perhaps was something a little phoney about my "crucial" motivation, and had real ized that neither **PT⁺** nor **PT⁻** could globally admit non-trivial ordinary holomorphic functions (0-functions). I was unaware (or had forgotten what little I knew) of holomorphic sheaf cohomology and had learned to live with the concept of "separated singularities". But then, after thirteen years this motivation was finally satisfied after all, in a quite unexpected way!

TOP

12. SPACE-TIME CURVATURE

I suppose that it was the many attractive properties of twistor geometry which mainly held me to that picture for so long. Even so, I recall having had, back in 1964, some considerable difficulty in persuading my colleagues, (Engelberg Schücking apart) of the significance or possible fruitfulness of this odd viewpoint, regarding space-time, that I was attempting to put forward. I expect that part of the reason for this was that my colleagues were basically all general-relativists (though, no doubt, I should have perceived even less enthusiasm had they been particle physicists or quantum field theorists!), my ideas seeming to be completely tied to the budding up of flat (or at least conformally flat) space-time. Indeed, looking back on all this I find it difficult to see why I was not myself more disturbed by this lack of curvature than I seem to have been. For a good many years earlier I had been (and remain to this day) firmly convinced of the (conformal) curvature of physical space-time. I think that I had been able to keep my worries at bay, at first, by the thought that such twistor ideas referred only to the quantum level of things, and that at this level space-time could be treated as flat, for the most part, the curvature perhaps arising only in discrete "quantized" lumps. I had been somewhat influenced by some of Tullio Regge's ideas for budding effectively, curved spaces in polyhedral fashion out of flat pieces (Regge 1961, Penrose 1972a) and had imagined that twistor ideas might be carried over to such situations in a piecewise manner. Somewhat later, while at Cornell in 1967, I was actually able to follow up this kind of idea in a concrete way. By examining plane-fronted, and later spherically fronted, gravitational waves with delta-function curvature - so the space-time was flat on either side of the wave - it proved to be possible to provide a twistor description involving a "shift" in the structure of the twistor space (Penrose 1968, Penrose & MacCallum 1972; cf. Bondi, Pirani & Robinson 1959). The very holomorphic functions (referred to above) which describe the structure of the curvature along the wave-front, proved to have significance here and had valuable implications in later developments in twistor theory.

In more recent years I have become less happy about the above point of view whereby space-time curvature arises in some sort of "quantized" discrete bits. Fortunately some newer developments, not unrelated to those described above, have arisen. Rather than "piecing together" the space-time from separate flat pieces, it is the twistor space which must, in an appropriate sense, be so pieced together. The local holomorphic structure of twistor space allows just the right sort of flexibility to enable such piecings to encode, in the global structure of the resulting twistor space, the appropriate (local) curvature information for the resulting space-time (Penrose 1976, Penrose & Ward 1980; an important input had been the H-space ideas of Newman 1976). This curious encoding of local space-time information in global twistor structure has emerged as a characteristic feature of the representation of physical fields in twistor terms. Indeed, this is already a feature of the holomorphic 1-function description of massless linear fields in **CM**. And the transition functions for the above piecing together of twistor space portions provide, in effect, a kind of non-linear 1-function.

TOP

13. SELF-DUAL AND ANTI-SELF-DUAL FIELDS

In fact the above type of construction has been successful in producing the general anti-self-dual solution of Einstein's vacuum equations (Penrose 1976). A corresponding subsequent construction due to Richard Ward (1977) provided the general anti-self-dual solution of the Yang-Mills equations. More accurately, these are non-linear version not of my original 1-functions, but of a slight modification suggested (in effect in 1973) by Lane Hughston. This produced anti-self-dual linear fields (with standard twistor conventions) whereas my original formula tion produced self-dual fields. With the positive frequency condition these are the negative- and positive-helicity fields, respectively (cf. Pen rose & MacCallum 1972).

For many years now I have been struggling to find a non-linear version of my original 1-function description in the gravitational case (the "googly graviton" - which, as cricketers know, spins right handedly though the bowling action suggests the reverse). I believe that it is likely a full solution is close - though there remains the possibility that all will collapse. The hope is that combining these "googly" ideas with the earlier ("leg-break") ones will eventually result in a complete formulation of the Einstein equations. Indeed, I feel that such grandiose objectives must necessarily be achieved if twistor theory is ultimately to realise its aims whereby space-time concepts can all be supplanted by twistor-holomorphic ones. We live, as before, in hope.

The idea that massless fields can usefully be split into their self-dual and anti-self-dual parts is clearly a crucial one in all this. Where first did I hear of this concept? I think that it was on the steps of some budding in Amsterdam, at the International Congress of Mathematicians in 1954. Someone was explaining to me, with characteristic enthusiasm and flair, what had seemed to me at the time to be a very odd idea indeed: that one should add the Maxwell field tensor to i times its dual! That someone was Ivor Robinson(6).

TOP

14. OTHER DEVELOPMENTS

It is hard for me to assess the success, or lack of it, that twistor theory has had to date. Much has happened in twenty years, though it has seemed for most of the time that progress has been grindingly slow. The mathematics has turned out to be much more difficult than I had guessed, and many of the necessary ideas, more sophisticated. That is my excuse for the slowness. But the programme has not as a whole been turned back, and that is encouraging. Some parts of the programme, though initially promising, have not moved a great deal in recent years, perhaps because of a lack of some essential new insight. (The particle programme, cf Penrose 1975b, Hughston 1979, Perjés 1975, 1979, Perjés & Sparling 1979 seems to be one of these.) Others have begun to move again only very recently - and here I refer particularly to the theory of twistor diagrams, into which some very promising looking new ideas have been injected by Andrew Hodges (1985) within the last year or so. It will be exciting to see where these lead.

There have also been some unexpectedly encouraging developments which I might refer to as "spin-off" rather than part of what I had thought of as the main objectives. I would include the solution of the Euclidean Yang-Mills "instanton" problem (Atiyah, Hitchin, Drinfeld & Manin 1978) and the construction of positive definite Ricci-flat self-dual 4-manifolds (cf. Hitchin 1979) among these; also "Killing spinors" (cf Walker & Penrose 1970), non-realizable CR-structures (cf. Penrose 1982a) and my own construction of a "quasi-local mass and angular momentum" complex, which provides a twistor-motivated definition of these quantities as surrounded by general spacelike topological 2-spheres in arbitrary space-times (Penrose 1982b cf. also Tod 1983, Shaw 1983, Penrose & Rindler 1985). The status of this definition is not yet completely clear, but at worst it seems to be a considerable improvement on all earlier suggestions, and some results provide good cause to be optimistic.

Whatever the future provides for twistor theory, it should be interesting.

TOP

15. NOTES AND REFERENCES

(1) Perhaps this seems unduly harsh. An idea - Hamiltonian theory, for example - may have immense utility and lead to new insights without, in this sense, having any new physical content. Thus in this sense it was, I suppose, the advent of quantum theory which saved the Hamiltonian viewpoint from the dustbin!

(2) It is not appropriate for me to dwell here at length on all the numerous other motivations, some vague and

some fairly clear-cut, which influenced the direction of the development of twistors. Among these was a desire for a formalism tailored to the four-dimensional (+---) structure of our space-time, rather than for something not so specific. The holomorphic nature of the space of null directions, in our particular dimension and signature, seemed to he a highly suggestive clue. Other motivations were provided by the experimental facts of left-right asymmetry and non-locality (cf. Lee & Yang 1956; Bohm 1951, Aharonov & Bohm 1959). Twistors have emerged as very compatible with these objectives.

(3) Perhaps in the present climate of eleven-dimensional generalized Kaluza-Kiein theories this objection would carry little weight with most people. However, to me it was, and still is, a fundamental drawback.

(4) Apparently on 1, December 1963 - for which date I thank Zsuzsi Ozsvath.

(5) He had not quite succeeded in satisfying this final requirement, but a slight modification of his solution (found a few months later by twistor-type methods, cf. Penrose 1965, provided what he had been seeking. We now refer to such solutions as elementary states (see Penrose 1975a) and they have importance in twistor theory.

(6) He was probably also explaining to me about self-dual null bivectors, but their relationship to spinors seems to have been something I learnt later!

REFERENCES

AHARONOV, Y. & BOHM, D. (1959), Phys. Rev. 115, 485.
ATIYAH, M. F., HITCHIN, N. J., DRINFELD, V. G. & MANIN, YU. 1. (1978), Phys. Lett., 65A, 185.
ATIYAH, M. F. & WARD, R. S. (1977), Commun. Math. Phys., 55, 111.
BATEMAN, H. (1904), Proc. Lond. Math. Soc. (2) 1, 451.
BATEMAN, H. (1910), Proc. Lond. Math. Soc., (2) 8,

223.
BATEMAN, H. (1944), Partial Differential Equations of Mathematical Physics (Dover, New York).
BOCHER, M. (1914), Bull. Amer. Math. Soc., 20, 185.
BOHM, D. (1951), Quantum Theory (Prentice-Hall, Englewood Cliffs).
BONDI, H., PIRANI, F. A. E. & ROBINSON, 1. (1959), Proc. Roy. Soc. (Lord.), A251, 519.
BRINKMAN, H. W. (1923), Proc. Nad. Acad. Sci. (U.S.), 9, 1.
CARTAN, A. (1914), Ann. Acole Norm. Super., 31, 263.
CAYLEY, A. (1860), Quart. J. of Pure and Appl. Math., 3, 225 (see Collected Mathematical Papers, 4, p. 447).
CAYLEY, A. (1869), Trans. Camb. Phd. Soc., 11 (2), 290 (see Collected Mathematical Papers, 7, p. 66).
COXETER, H. S. M. (1936), Ann. of Math., 37, 416.
CUNNINGHAM, E. (1910), Proc. Lond. Math. Soc., (2) 8, 77.
DARBOUX, G. (1914), Lecons sur la Theorie Generale des Surfaces, Pt. 1, 2nd edn. (Gauthier-Villars, Paris).
DIRAC, P. A. M. (1936a), Proc. Roy. Soc. (Lond.), A155, 447.
DIRAC, P. A. M. (1936b), Ann. of Math., 37, 429.
GEL'FAND, 1. M., GRAEV, M. 1. & VILENKIN, N. Ya. (1966), Generalized Functions, vol. 5: Integral Geometry and Representation Theory (Academic Press, New York).
GINDIKIN, S. G. (1983), Math. Intelligencer, 5, no. 1, 27.
HEPNER, W. A. (1962), Nuovo Cim., 26, 351.
HITCHIN, N. J. (1979), Math. Proc. Camb. Phd. Soc., 85, 465.
HODGES, A. P. (1983), Proc. Roy. Soc. (Lond.) A385, 207; A386, 185.
HODGES, A. P. (1985), Proc. Roy. Soc. (Lond), A 317, 41, 375.
HODGES, A. P. & HUGGETT, S. A.; HUGHSTON, L. P.; PENROSE, R.; TOD, K. P.; WARD, R. S. (1980), Surv. High Eenergy Physics, 1, 333, 313, 267, 299, 289.
HUGHSTON, L. P. (1979), Twistors and Particles; Lecture Notes in Physics 97 (Springer, Berlin).
KLEIN, F. (1870), Math. Annalen, 2, 198.
KLEIN, F. (1926), Vorlesungen über höhere Geometrie

(Springer, Berlin), pp. 80, 262.
KUIPER, N. H. (1949), Ann. of Math., 50, 916.
LEE, T. D. & YANG, C. N. (1956), Phys. Rev., 104, 254.
LIE, S. & SCHEFFERS (1896), Berührungstrsfn. 453.
McLENNAN, Jr., J. A. (1956), Nuovo Cim., 10, 1360.
MOUSSOURIS, J. (1983), Quantum Models of Space-Time Based on Recoupling Theory (D. Phil. thesis, Oxford).
MURAI, Y. (1953/4), Progr. Theoret. Phys., 9, 147; 11, 441.
MURAI, Y. (1958), Nucl. Phys., 6, 489.
NEWMAN, E. T. (1976), Gen. Rel. Grav., 7, 107.
NEWMAN, E. T. & PENROSE, R. (1968), Proc. Roy. Soc. (Lond.), A305, 175.
PENROSE, R. (1959), Proc. Camb. Phil. Soc., 55, 137.
PENROSE, R. (1965a), in Proceedings of the 1962 Conference on Relativistic Theories of Gravitation, Warsaw (Polish Acad. Sci. Warsaw).
PENROSE, R. (1965b), Proc. Roy. Soc. (Lond.), A284, 159.
PENROSE, R. (1967), J. Math. Phys., 8, 345.
PENROSE, R. (1968), Int. J. Theor. Phys., 1, 61.
PENROSE, R. (1969a), J. Math. Phys., 10, 38.
PENROSE, R. (1969b), in Quantum Theory and Beyond (ed. E. Bastin; Cambridge Univ. Press, Cambridge).
PENROSE, R. (1972a), in General Relativity (ed. L. O'Raifeartaigh; Clarendon Press, Oxford).
PENROSF, R. (1972b), in Magic without Magic (ed. J. R. Klauder; Freeman, San Francis co).
PENROSE, R. (1975a), in Quantum Gravity, and Oxford Symposium (eds. C. J. Isham, R. Penrose & D. W. Sciama; Clarendon Press, Oxford).
PENROSE, R. (1975b), in Quantum Theory and the Structures of Time and Space (eds. L. Castell, M. Dreischner & C. F. von Weizsacker; Carl Hanser Verlag, Munich).
PENROSE, R. (1976), Gen. Rel. Grav., 7, 31, 17 1.
PENROSE, R. (1979), in Advances in Twistor Theory, Research Notes in Mathematics 37 (eds. L. P. Hughston & R. S. Ward; Pitman, San Francisco).
PENROSE, R. (1980), Gen. Rel. Grav., 12, 225.
PENROSE, R. (1982a), Bull. (New Set.) Amer. Math.

Soc., 8, 427.
PENROSE, R. (1982b), Proc. Roy. Soc. (Lond.), A381, 53.
PENROSE, R. & MACCALLUM, M. A. H. (1972), Phys. Repts., 6C, 241.
PENROSE, R. & RINDLER, W. (1984), Spinors and Space-Time; vol. 1, Two-Spinor Calculus and Relativitic Fields (Cambridge Univ. Press, Cambridge).
PENROSE, R. & RINDLER, W. (1985), Spinors and Space-Time; vol. 2, Spinor and Twistor Methods in Space-Time Geometry (Cambridge Univ. Press, Cambridge).
PENROSE, R. & WARD, R. S. (1980), in General Relativity, One Hundred Years after the Birth of Albert Einstein (ed. A. Held; Plenum, New York).
PERJÉS, Z. (1975), Phys. Rev., 111), 2031.
PERJÉS, Z. (1979), Phys. Rev., 201), 1857.
PERJÉS, Z. & SPARLING, G. A. J. (1979), in Advances in Twistor Theory, Research Notes in Mathematics 37 (eds. L. P. Hughston & R. S. Ward; Pitman, San Francisco).
PLÜCKER, J. (1865), Phil. Trans. t., 155, 725.
PLÜCKER, J. (1868/9), Neue Geometrie des Raumes gegründet auf die Betrachtung der geraden Linic als Raumdement (ed. F. Klein, Leipzig).
RADON, J. (1917), Ber. Verb. Sächs. Akad., 69, 262.
REGGE, T. (1961), Nuovo Cim., 19, 558.
ROBINSON, I. (1956), Report to the Eddington Group, Cambridge.
ROBINSON, I. (1961), J. Math. Phys., 2, 290.
ROBINSON, I. & TRAUTMAN, A. (1962), Proc. Roy. Soc. (Lond.), A265, 463.
RUDBERG, H. (1958), dissertation, Univ. of Uppsala, Uppsala, Sweden.
RZEWUSKI, J. (1958), Nuovo Cim., 9, 942.
SCHRÖDINGER, E. (1952), Science and Humanism (Cambridge Univ. Press, Cambridge).
SHAW, W. T. (1983), Proc. Roy. Soc. (Lond.), A390, 191.
SPARLING, G. A. J. (1975), in Quantum Gravity, an Oxford Symposium (eds. C. J. Isham, R. Penrose & D. W. Sciama; Clarendon Press, Oxford.)

TOD, K. P. (1983), Proc. Roy. Soc. (Lond.), A388, 457.
WALKER, M. & PENROSE, R. (1970), Commun. Math. Phys., 18, 265.
WARD, R. S. (1977), Phys. Lett., 61A, 81.
WEIERSTRASS, K. (1866), Monatsberichte der K. P. Akademie, 612 (collected works, vol. 3).
WHITTAKER, E. T. (1903), Math. Annalen, 57, 333.
WHEELER, J. A. (1962), Geometrodynamics (Acad. Press, New York).
TOP