
by Roger Penrose
Mathematical Institute, Oxford

Abstract Original motivations are recalled, for the introduction twistor theory, as a distinctive complex-geometric approach to the basic physics of our world, these being aimed at applying specifically to (3+1)-dimensional space-time, but where space-time itself is regarded as a notion secondary to the twistor geometry and its algebra. Twistors themselves may be initially pictured as light rays—with a twisting aspect to them related to angular momentum. Twistor theory provides an economical conformally invariant description of quantum wave functions for massless particles and fields, best understood in terms of holomorphic sheaf cohomology, subsequently leading to a non-linear description of anti-self-dual (“left-handed”) gravitational (and Yang-mills) fields. Attempts to remove this anti-self-dual restriction (the googly problem) led to a 40-year blockage to the development of twistor theory as a possible overall approach to fundamental physics. However, in recent years, a more sophisticated approach to this problem has been developed—referred to as palatial twistor theory—whose basic procedures are described here, where a novel generating-function approach to Λ-vacuum Einstein equations is introduced.

CONTENTS

Part A. Early motivations
A1. Geometrical background: two roles for a Riemann sphere
A2. The 2-spinor formalism
A3. Zero rest-mass fields

Part B. The emergence of twistor theory
B1. Robinson congruences
B2. Twistors in terms of 2-spinors
B3. Minkowski space compactified, complexified, and its conformal symmetry
B4. The basic twistor spaces
B5. Helicity and relativistic angular momentum
B6. Description under shift of origin

Part C: Fields, quantization and curved space-time
C1. Twistor quantization rules
C2. Twistor wave functions
C3. Twistor generation of massless fields and wave functions
C4. Singularity structure for twistor wave functions
C5. Čech cohomology
C6. Infinity twistors and Einstein’s equations

Part D: Palatial twistor theory
D1. Basic ideas of palatial twistor theory
D2. The spaces of momentum-scaled and spinor-scaled rays
D3. A palatial role for geometric quantization
D4. Palatial generating functions and Einstein’s equations
Part A. Early motivations

A1. Geometrical background: two roles for a Riemann sphere

The basic geometrical proposal underlying twistor theory effectively came together in early December 1963, when I was on a 9-month appointment at the University of Texas in Austin [1]. Various motivational notions had been troubling me for several years previously, concerning what I had felt to be a need for a novel approach to foundational physics, in which concepts from both quantum mechanics and relativity theory had significant roles to play. These were interrelated via the theme of complex analysis and complex-number geometry, areas of mathematics that had impressed me deeply from around 1950, during my time as an undergraduate in mathematics at University College, London. These ideas had then featured strongly in my mind in the early 1960s. The thought I had in late 1963 was the initial stage of the proposal that, a little later, I indeed referred to as “twistor theory”, owing to a key role that the twisted configuration of interlocking circles shown in figure 1 (a stereographically projected family of the Clifford parallels on a 3-sphere) had played for me. The reader might well ask what such an intriguing configuration might have to do with a basic theory of physics. We shall see later that this configuration represents the angular momentum of a massless particle with spin, but in order to explain this, it is necessary first to outline some of the various ideas that had been troubling me earlier. I shall come to the specific role of the configuration of Fig.1 in §B1, §B4. and B5, particularly t the end of that section.

Fig. 1:

A picture representing a non-null twistor: stereographic projection—to a Euclidean 3-space $E$—of Clifford parallels on a 3-sphere. The tangent directions to the circles point in the direction (projected into $E$) of the rays of a Robinson congruence. By continually reassembling itself, the entire configuration travels with the speed of light, as $E$ evolves in time, in the direction of the large arrow at the top right. The configuration represents the angular momentum structure of a massless particle with spin.
One of my main motivations had arisen from my feeling that there was a need for a formalism that was geared to that specific dimensionality of space-time structure that we directly perceive around us. This line of thinking was very unlike that of various other ideas for an underlying physics of the world that later became popular, e.g. string theory [2]. I had earlier become convinced that what was needed would be a formalism that should be very specific to the number of space and time dimensions, namely 3 and 1, respectively, that macroscopically present themselves to us, and I took the view that this should be central to the scheme. This indeed goes very much in opposition to the role of space-time dimensionality underlying many of the current trends, most particularly string theory, where extra space dimensions (and even an extra time dimension, in the case of “F-theory”) are regarded as essential ingredients of these various theories [2], taken to be serious proposals for the overall space-time geometry of the physical world that we inhabit. It also contrasts with the very natural and commendable desire, in pure mathematics, for formalisms that can be applied, generally, to any spatial dimensionality whatever, but the aims of theoretical physics are very different from those of pure mathematics, even though much of theoretical physics depends vitally on the latter.

Another of my basic motivations had been for a formalism that was essentially complex in the sense that it would be able to take advantage of what I had regarded, ever since my days as a mathematics undergraduate, as the “magic” of complex analysis and holomorphic (i.e. complex analytic) geometry. I had learnt that the complex number system has not only a profoundly deep power and elegance, but that it had also found a basic realization in its underlying role in the formalism of quantum theory. I later began to study quantum mechanics in a serious way, and was particularly impressed by the superb course of lectures given by Paul Dirac, when I was a graduate student (in algebraic geometry), and subsequently a Research Fellow, at St John’s College Cambridge. I became fascinated by the quantum description of spin, and how the complex numbers of quantum mechanics were directly related to the 3-dimensiality of physical space, via the 2-sphere of spatial directions being appropriately identified as a Riemann (or Bloch) sphere of the ratios of pairs of complex numbers (quantum amplitudes) where, in the case of a massive particle of spin $\frac{1}{2}$ such as an electron (see figure 2), we can think of these as being the complex components of a 2-spinor. Moreover, I had realized that in the relativistic context, there was another role for the Riemann sphere, this time as the celestial sphere that an astronaut in space would observe. The transformation of this celestial sphere to that of a second astronaut, moving at a relativistic speed while passing nearby the first would be one that preserves the complex structure of the Riemann sphere (i.e. conformal without reflection). The special (i.e. non-reflective) Lorentz group is thus seen to be identical with these holomorphic transformations of this Riemann sphere (Mobius transformations). Again this was clear from the 2-spinor formalism, this time in the relativistic context (see [3]).
The Riemann sphere (here in its role as a Bloch sphere) projects stereographically from its south pole S to the complex (Wessel) plane, whose unit circle coincides with the equator of the sphere. A general spin state $|\Psi\rangle = w|\uparrow\rangle + z|\downarrow\rangle$, of a spin-$\frac{1}{2}$ massive particle is represented by the pint $Z$ on the Wessel plane denoting the complex number $u = zw$, which is the stereographic image of $Z'$ on the sphere (so $S$, $Z$, and $Z'$ are collinear). The spin direction $\Psi$ is then $OZ$, where $O$ is the sphere’s center.

**A2. The 2-spinor formalism**

This dual role for the Riemann sphere, one fundamentally to do with quantum mechanics in the case of 3 spatial dimensions, and the other fundamentally to do with macroscopic relativity, in (3+1)-dimensional space-time, struck me as being no accident, but something that linked together these two great revolutions of 20th century physics—of the small and of the large—via the magic of complex numbers. I felt that this might represent a definite clue to a deep unifying relation between the two. Both could be seen as a feature of the 2-spinor calculus, as introduced be Cartan [4] and van der Waerden [5], and which I had learnt how to use from Dirac (see [6]), in an unexpected deviation from his normal Cambridge course on quantum mechanics.

I liked to think of a 2-spinor (often referred to by physicists as a “Weyl spinor”) in a very geometrical way, and I realized that, up to an overall sign, a non-zero 2-spinor can be represented as a future-pointing null vector (a vector pointing along the future null cone), referred to as the “flagpole”, together with a “flag plane” direction through that flagpole [7], [8]. The flag plane would be a null half-plane bounded by the flagpole. This flag geometry can be thought of in the following way. Imagine the Riemann sphere $\mathcal{S}$ of null (i.e. lightlike) directions at some point $O$ in space-time. (See figure 3.) We are thinking of the geometry in the tangent 4-space of the point $O$. The flagpole direction is represented by some point $P$ on a sphere of cross-section of the future null cone of $O$, which we identify with $\mathcal{S}$, and we choose a point $P'$ on $\mathcal{S}$ infinitesimally separated from $P$. The straight line extended out from $P$ in the direction of $P'$, when joined to $O$, defines the required flag half-plane. We note that as the point $P'$ rotates about $P$, the flag plane rotates about the flagpole. The spinor itself is defined only up to
sign by this geometry, but we must take note that if $P'$ rotates continuously around $P$ through $2\pi$, the spinor becomes replaced by its negative. To reach the original 2-spinor by this procedure, the rotation of the flag plane would have to be through $4\pi$.

I had found that 2-spinor methods were surprisingly valuable in giving us insights into the formalism of general relativity that were different from those that the standard Lorentzian tensor framework readily provides. Most immediately striking was the very simple-looking 2-spinor expression for Weyl’s conformal curvature [9] (see also [10]). Whereas the usual Weyl-tensor quantity $C_{abcd}$ has a somewhat complicated collection of symmetry and trace-free conditions, the corresponding 2-spinor is simply a totally symmetric complex 2-spinor quantity $\Psi_{ABCD}$.

Some comments concerning the 2-spinor index notation being used here are appropriate. Capital italic Latin index letters $A$, $B$, $C$, … refer to the (2-complex dimensional) spin space if they are upper indices, and to the dual of this space if lower ones; primed such letters $A'$, $B'$, $C'$, … refer to the complex-conjugate spin space. The tensor product of the spin space with its complex conjugate is identified with the complexified tangent space to the space-time, at each of its points, here the real tangent vectors arise as the Hermitian members of this tensor product. In general, I shall take these as abstract indices, in the sense described in my book with Wolfgang Rindler, Spinors and Space-Time, volume 1 [8], so that no coordinate system is implied, either for the space-time or to define a basis for the spin-space. This is notationally very handy, because the space-time indices $a$, $b$, $c$, … can then be thought of as “shorthand” for the spinor index pairs:

$$a = AA', \quad b = BB', \quad c = CC', \ldots$$

The spin-space (and hence also its dual and complex conjugate) has a symplectic structure defined by the skew-symmetric quantities

$$\varepsilon_{AB}, \quad \varepsilon^{AB}, \quad \varepsilon_{A'B'}, \quad \varepsilon^{A'B'},$$

these being used for lowering or raising indices, (where we must be a little careful about signs and index orderings):

$$\kappa_B = \kappa^A \varepsilon_{AB}, \quad \kappa^A = \kappa_B \varepsilon^{AB}, \quad \eta_{B'} = \eta^A \varepsilon_{A'B'}, \quad \eta^A = \eta_{B'} \varepsilon^{A'B'}$$
so that on terms of components,

\[ \kappa_1 = \kappa^0, \quad \kappa_0 = -\kappa^1, \quad \eta_1 = \eta^0, \quad \eta_0 = -\eta^1, \]

where the component form of each of the epsilons is

\[
\begin{pmatrix}
0 \\
1 \\
-1 \\
0
\end{pmatrix}.
\]

The metric tensor, in abstract-index form is

\[ g_{ab} = \varepsilon_{AB} \varepsilon_{A'B'}, \]

and the abstract-index form of the Weyl conformal curvature tensor for space-time is

\[ C_{abcd} = \Psi_{ABCD} \varepsilon_{A'B'} \varepsilon_{C'D'} + \varepsilon_{AB} \varepsilon_{CD} \Phi_{A'B'C'D'}. \]

Here, I have allowed for the case of a complex metric \( g_{ab} \), both \( \Psi_{ABCD} \) and \( \Phi_{A'B'C'D'} \) being totally symmetric, where \( \Psi_{ABCD} \) describes the anti-self-dual (left-handed) Weyl curvature and \( \Phi_{A'B'C'D'} \), the self-dual (right-handed) part. In the case of a real Lorentzian space-time metric (\( \varepsilon_{AB} = \varepsilon_{AB} \)) and \( \Phi_{A'B'C'D'} \) is the complex conjugate of \( \Psi_{ABCD} \): \( \Phi_{A'B'C'D'} = \overline{\Phi}_{A'B'C'D'} \),

but it will be important for what follows that we consider the complex case also, as we shall be concerned with self-dual (complex vacuum) space-times, for which \( \Psi_{ABCD} = 0 \) and anti-self-dual ones, for which \( \Phi_{A'B'C'D'} = 0 \), later (these complex fields being regarded as wave functions).

**A3. Zero rest-mass fields**

We find that in the case of a (real Lorentzian) vacuum metric (with or without cosmological constant), the Bianchi identities become

\[ \nabla^{AA'} \Psi_{ABCD} = 0 \]

which may be compared with the Maxwell equations in charge-free space-time

\[ \nabla^{AA'} \phi_{AB} = 0, \]

where \( \phi_{AB} \) relates to a (possibly complex) Maxwell field tensor \( F_{ab} \) in the same way as \( \Psi_{ABCD} \) relates to \( C_{abcd} \), namely

\[ F_{ab} = \phi_{AB} \varepsilon_{A'B'} + \varepsilon_{AB} \Phi_{A'B'}, \]

where \( \phi_{AB} \) describes the anti-self-dual (left-handed) part of the field and \( \Phi_{A'B'} \), the self-dual (right-handed) part. For a real Maxwell field, they are complex conjugates of each other:
\[ \bar{\phi}_{A'B'} = \bar{\phi}_{A'B'}. \]

I had become interested in the issue of finding solutions of the general equation

\[ \nabla_{AA'} \phi_{ABC\ldots E} = 0 \]

in (conformally) flat space-time, \( \phi_{ABC\ldots E} \) being symmetric in its \( n \) spinor indices, the equation being the (conformally invariant) free-field equation for a massless field of spin \( n/2 \) [6], [11], [12]. This equation (together with the wave equation in suitably conformally invariant form, which includes an \( R/6 \) term, \( R \) being the scalar curvature) had a particular importance for me, and I believed it to have a rather basic status in relativistic physics. For I had come to the view that nature might have a “massless” structure at its roots, mass itself being a secondary phenomenon. In around 1961 (see [13]) I had found a formula for obtaining the solution of this field equation from general data freely specified on a null initial hypersurface. I had formed the view that this formula had a certain kinship with the Cauchy integral formula for obtaining the value of a holomorphic function at some point of the complex plane in terms of the function’s values along a closed contour surrounding that point. I had felt that, in some sense, this massless field equation might be akin to the Cauchy-Riemann equations. There had to be in some unusual “complex” way of looking at Minkowski space, I had surmised, in which the massless field equations were simply a statement of \textit{holomorphicity}—but in what sense could this possibly be true?

There was one remaining feature that I felt sure must be represented, as part of this mysterious “complex” way of looking at space-time. This arose from a discussion that I had had with Engelbert Schücking when I shared an office with him in the spring of 1961 at Syracuse University in New York State. Engelbert had persuaded me of the key importance to quantum field theory of the splitting of field amplitudes into positive and negative frequency parts. I was not happy with the standard procedure of first resolving these amplitudes into Fourier components and then selecting the positive ones, as not only did this strike me as too “top-heavy”, but also the Fourier analysis is not conformally invariant—and I had come to believe that this conformal invariance, being a feature of massless fields, was important (again, something that had been stressed to me by Engelbert).

I had become aware that for complex functions defined on a line (thought of as the time line) we may understand their splitting into positive- and negative-frequency parts in the following way. We view this time line as being the equator of real numbers in a \textit{Riemann sphere} which, as before, is the complex plane compactified by the single point labelled by “\( \infty \)”, but where the sphere is now being oriented somewhat differently from that of figure 2, with the real numbers now featuring as the equator (increasing as we proceed in an anti-clockwise sense on the horizontal plane), rather than the unit circle. Functions defined on this equatorial circle which extend holomorphically into the southern hemisphere (with usual conventions) are the functions of positive frequency, and those which extend holomorphically into the northern hemisphere are those of negative frequency. An arbitrary complex function defined on this circle can be split into a function extending globally into the southern hemisphere and one globally into the northern hemisphere—uniquely except for an ambiguity with regard to the constant part—and this provides us with the required positive/negative frequency split, without any resort to Fourier analysis. I wanted to extend this picture into something more global, with regard to space-time, and I had in mind that my sought-for “complex” way of looking at Minkowski space should exhibit something strongly analogous to this division into
two halves, where the boundary between the two could be interpreted in “real” terms, in some direct way. This had set the stage for the emergence of twistor theory.

Part B. The emergence of twistor theory

B1. Robinson congruences

A colleague of mine, Ivor Robinson, who had taken up a position at what later became the University of Texas at Dallas, had been working on finding global non-singular null solutions of Maxwell’s free-field equations in Minkowski space-time $\mathbb{M}$, where “null” in this context means that the invariants of the field tensor $F_{ab}$ vanish, i.e. $F_{ab}F^{ab}=0=\ast F_{ab}F^{ab}$ where $\ast F_{ab}$ is the Hodge dual of $F_{ab}$. Equivalently, in 2-spinor terms, $\varphi_{AB}\varphi^{AB}=0$, which tells us that

$$\varphi^{AB} = \kappa^A \kappa^B,$$

for some $\kappa^A$. It is not hard to show that the Maxwell source-free equations then imply that the flagpole direction of $\kappa^A$ points along a 3-parrameter family—a congruence—of null straight lines, which turn out to be what is called “shear-free”, which means that although the lines may diverge, converge, or rotate, locally, there is no shear (or distortion) as we follow along the lines.

Although, not relevant to the discussion at the moment, it is worth noting that the study of shear-free congruences of rays in curved space-times has a considerable historical significance—where I use the term “ray” simply to mean a null (i.e. lightlike) geodesic in space-time. In particular, the well-known Kerr solution [14], [15] of the Einstein vacuum equations for a rotating black hole possesses a shear-free ray congruence, and this played a key role in its discovery, as it did also in Newman’s generalization to an electrically charged black hole [16], and also in the Robinson-Trautman gravitationally radiating exact solutions [17], among other examples. As in the case of Minkowski space $\mathbb{M}$, as described above, it is also true that for any null solution $\varphi^{AB}$ of Maxwell’s equations in curved space-times, the flagpole directions of the $\kappa^A$-spinors point along a shear-free family of rays.

A simple example of a shear-free ray congruence in $\mathbb{M}$ is obtained from any fixed choice of a ray $L$ in $\mathbb{M}$, where the family of all rays that meet $L$ provides a shear-free ray congruence. I refer to such a congruence as a special Robinson congruence, and this includes the limiting case when $L$ is taken out to infinity, so our congruence becomes a family of parallel rays in $\mathbb{M}$. Ivor Robinson had developed ways of producing null solutions of the Maxwell equations, starting from any given shear-free null congruence, but when applied to the special congruences just described, he found that singularities would arise along the line $L$ itself (except in the otherwise unsatisfactory case where $L$ is a is at infinity). Desiring a singularity-free Maxwell field, he provided the following ingenious trick. Consider, instead, solutions of Maxwell’s equations in the complexified Minkowski space-time $\mathbb{C}\mathbb{M}$, and displace the line $L$ in a complex direction, so that it lies in $\mathbb{C}\mathbb{M}$, but entirely outside its real part $\mathbb{M}$. Complex analytic solutions of Maxwell’s equations, based on the complex “special Robinson congruence” defined by the displaced $L$ need not now be singular within $\mathbb{M}$, and the flagpoles of the $\kappa^A$-spinors within $\mathbb{M}$ now point along an entirely non-singular shear-free ray congruence in $\mathbb{M}$, which I later named a (general) Robinson congruence.
I became highly intrigued by the geometry of general Robinson congruences, and I soon realized that one could describe them in the following way. Consider an arbitrary spacelike 3-plane $E$ in Minkowski 4-space $\mathbb{M}$. $E$ has the geometry of ordinary Euclidean 3-space, and each ray $N$ of the congruence will meet $E$ in a single point, at which we can determine the location of that ray within $\mathbb{M}$ by specifying a unit 3-vector $n$ at that point, pointing in the spatial direction that is the orthogonal projection into $E$ of the null direction of $N$ there. Thus we have a vector field of $n$s within $E$ to represent the Robinson congruence. After some thought I realized what the nature of this vector field must be. The $n$-vectors are tangents to the oriented circles (together with one oriented straight line) obtained by stereographic projection of a family of oriented Clifford parallels on a 3-sphere. See figure 1, in §A1, for a picture of this configuration, and reference [18] for a detailed derivation. The large arrow at the top right indicates the direction in which the configuration appears to move with the speed of light by continually reassembling itself in that direction, as $E$ moves by parallel displacement into the future.

By examining this configuration, and counting the number of degrees of freedom that such configurations have, I realized that the space of Robinson congruences must be 6-dimensional. Moreover, it was reasonably clear to me that by its very mode of construction, this space ought to have a complex structure, and so must be, in a natural way, a complex 3-manifold. Within this space would lie the space of special Robinson congruences, each of which would be determined by a single ray (namely $L$). The space of rays in $\mathbb{M}$ is 5-real-dimensional, and it divides the space of general Robinson congruences into two halves, namely those with a right-handed twist and those with a left-handed twist. The complex 3-space of Robinson congruences, which came to be known as “projective twistor space” appeared to be just what I believed was needed, where the “real” part of the space (representing light rays in $\mathbb{M}$, or their limits at infinity) would, like the “real” equator of the Riemann sphere described at the end of §A2, divide the entire space into two halves. This, indeed appeared to be exactly the kind of thing that I was looking for!

**B2. Twistors in terms of 2-spinors**

To be more explicit about things, and to understand precisely how the space of Robinson congruences does indeed provide a compact complex 3-manifold divided in two by the real 5-space of special Robinson congruences, let us turn again to the relativistic 2-spinor formalism of §A2. We shall see how this allows us to provide a very neat description of individual rays in $\mathbb{M}$. In §B4, we see how this generalizes to describe general Robinson congruences. The physical interpretation in terms of relativistic angular momentum of massless particles will emerge in B5.

Consider some ray $Z$ in $\mathbb{M}$, and let us assign a strength to this ray in the form of a null 4-momentum convector $p_\alpha$, where the vector $p^\mu$ points along $Z$ at each of its points, parallel-propagated along $Z$. In fact, let us go a little further than this by assigning a (dual, conjugate) 2-spinor $\pi_A$, parallel-propagated along $Z$, where

$$p_\alpha = \bar{\pi}_A\pi^{A'},$$

so that in addition to having $\pi_A$’s flagpole pointing along $Z$, we also have $\pi_A$’s flag plane (and spinor sign) assigned to $Z$, and which is to be parallel-propagated along it. This will be referred to as a spinor scaling for the ray $Z$. 
We need to choose a space-time origin point $O$ within $\mathbb{M}$, so that any point $X$ of $\mathbb{M}$ can be labelled by a position vector $x^a (= x^{A'})$ at $O$. Then if $X$ is any point on the ray $Z$, we can define a 2-spinor $\omega^A$ by the equation,

$$\omega^A = i x^{AA'} \pi_A'$$

and we find that $\omega^A$ remains unchanged if $X$ is replaced by any other point on the ray $Z$, such a point having a position vector of the form

$$x^{AA'} + k \pi^{A'} \pi_A' ,$$

where $k$ is any real number (since $\pi^{A'} \pi_A' = 0$). The pair $(\omega^A, \pi_A)$, serves to identify the ray $Z$, together with a spinor scaling for $Z$.

The 2-spinors $\omega^A$ and $\pi_A$ are the spinor parts (with respect to the origin $O$) of the twistor $Z^{a}$, which represents the spinor-scaled ray $Z$, and often one simply writes

$$Z^a = (\omega^A, \pi_A).$$

However, for a ray, there is a particular equation that must hold between the spinor parts, namely

$$\omega^A \overline{\pi_A} + \pi_A \overline{\omega^{A'}} = 0$$

which follows from the fact that the vector $x^a$ is real, so that $x^{AB'}$ has the Hermitian property $x^{AB'} = x^{BA'}$. The above equation can be rewritten as

$$Z^a \overline{Z_a} = 0$$

where $\overline{Z_a}$, the complex conjugate of $Z^a$

$$\overline{Z_a} = (\overline{\pi_a}, \overline{\omega^{A'}}),$$

(and note the reverse order of the spinor parts) is a dual twistor. When $Z^a \overline{Z_a} = 0$, we refer to $Z^{a}$ as a null twistor, so it is that the null twistors represent (spinor-scaled) rays in $\mathbb{M}$—or rays at $\mathbb{M}$’s infinity.

The above equation

$$\omega^A = i x^{AA'} \pi_A'$$

is referred to as the incidence relation between the space-time point $X$ and the twistor $Z^{a} = (\omega^A, \pi_A)$. We may also be interested in this incidence relation when $X$ is allowed to be a complex point. Likewise, for a dual twistor

$$W_a = (\lambda_a, \mu^{A'}).$$
incidence with a (possibly complex) point $X$ is expressed as

$$\mu^A = -i x^{AA'} \lambda_A.$$ 

It is useful to get a picture of the geometrical role of the 2-spinor $\omega^A$, in addition to $\pi_A$, in the case of a general null twistor $Z^a = (\omega^A, \pi_A)$. Figure 4 shows this, where $O$ is the origin the point $Q$ is the intersection of the ray $Z$ with the light cone of $O$. The null vector $\overrightarrow{OQ}$ has index form $q^a$ and is proportional to the flagpole of $\omega^A$ where

$$q^{AA'} = \omega^A \overline{\omega}^{A'} (i \overline{\omega}^B \pi_B)^{-1}.$$ 

This expression fails only when $\overline{\omega}^B \pi_B = 0$ (but holding in a certain limiting sense) which occurs when the ray $Z$ lies in a null hyperplane through $O$, and the point $Q$ lies at infinity.

The flagpole directions of the spinor parts of a general null twistor $Z^a = (\omega^A, \pi_A)$ are depicted, where $Q$ is the intersection of the ray $Z$ with the light cone of the origin $O$. [NEED lettering L, O, Q, $\omega$, $\pi$]

**B3. Minkowski space compactified, complexified, and its conformal symmetry**

At this juncture it would be helpful to clarify the nature of “infinity”, with regard to Minkowski space $\mathbb{M}$. We recall that when a ray $L$ is characterized in terms of the null congruence of rays that intersect $L$, we were led to consider the ray congruences that consist entirely of parallel rays, arising when $L$ is moved out to infinity. There is a whole 2-sphere’s-worth of such systems of parallel rays one for each null direction. Thus the family of limiting rays $L$ at infinity generates a kind of “light cone at infinity”, frequently denoted by the script letter $\mathcal{I}$ (and pronounced “scri”). We can regard $\mathcal{I}$ as being the identification of $\mathbb{M}$’s future conformal boundary $\mathcal{I}^+$ with its past conformal boundary $\mathcal{I}^-$ (see [12], [18]). This identification also incorporates the single point $i$ (the vertex of $\mathcal{I}$) where, which is the identification of the three points $i^+, i^0$, and $i^+$, respectively representing past infinity, spacelike infinity, and future infinity. See figure 5. This provides us with the picture of compactified Minkowski space $\mathbb{M}^b$ (whose turns out to have topology $S^1 \times S^3$) where figure 5a indicates the future and past null boundaries of $\mathbb{M}$, and figure 5b shows how these two conformal boundaries $\mathcal{I}^+$ and $\mathcal{I}^-$ are to be identified as $\mathcal{I}$ where future and past end-points of any ray in $\mathbb{M}$ are identified. This provides us with the highly symmetrical compact Lorentzian-conformal manifold $\mathbb{M}^b$. Every ray within $\mathbb{M}$ being compactified by a single point to become a topological circle.
we saw in §A2 (see figure 3) that a 2

where we think of this spinor scaling as parallel

also a spinor scaling, defined

Minkowski space

understood explicitly in terms of the geometrical representation of a null twistor

obtained simply by “unwrapping” the

fold cover of

us that this SU(2,2) actually acts on a 2

SO(2,4)

unitary group SU(2,2). It is, indeed, one of Cartan’s specif

so the group of (complex

than this. The pseudo

be

6

the null cone

C

4 connected components since it allows for reversals of time and space orientations. I shall be

frequently

referred to as the conformal group of flat 4-dimensional space-time. This group has 4 connected components since it allows for reversals of time and space orientations. I shall be

concerned here only with the connected component of the identity, referring to this group as

C(1,3).

Another way of understanding \(\mathbb{M}^\#\) is that it represents the family of generator lines of the null cone \(\mathcal{K}\) of the origin \(O^{2,4}\) (i.e. of entire rays through \(O^{2,4}\)) in the pseudo-Minkowskian 6-space \(\mathbb{M}^{2,4}\), whose signature is \((+ + - - - -)\). See [18]. If we consider these generator lines to be oriented, (or as being null half-lines, starting at \(O^{2,4}\)), then we get a 2-fold cover \(\mathbb{M}^{\#2}\) of \(\mathbb{M}^\#\), since an action of the pseudo-orthogonal group \(\text{SO}(2,4)\) on \(\mathcal{K}\) can continuously rotate any particular one of its oriented generators into itself but with opposite orientation (i.e. reflected in the origin \(O^{2,4}\)). Thus, the family of oriented rays through \(O^{2,4}\) provides us with a realization of the 2-fold cover \(\mathbb{M}^{\#2}\) of \(\mathbb{M}^\#\).

The symmetry group of the vector space \(\mathbb{T}\) of twistors \(Z^\alpha = (\omega^A, \pi_A)\) goes a step further than this. The pseudo-Hermitian form \(|Z| = Z^\alpha \bar{Z}_\alpha = (\omega^A \bar{\pi}_A + \pi_A \bar{\omega}^A)\) has split signature \((+ + - - - -)\), so the group of (complex-)linear transformations of \(\mathbb{T}\) that preserve the norm is the pseudo-unitary group \(\text{SU}(2,2)\). It is, indeed, one of Cartan’s specific local isomorphisms (see [18]) that \(\text{SO}(2,4)\) locally isomorphic to \(\text{SU}(2,2)\), the latter being a 2-fold cover of the former. This tells us that this \(\text{SU}(2,2)\) actually acts on a 2-fold cover \(\mathbb{M}^{\#4}\) of \(\mathbb{M}^{\#2}\). The space \(\mathbb{M}^{\#4}\) is therefore a 4-fold cover of \(\mathbb{M}^\#\). Topologically, we, we can understand such an \(n\)-fold cover \(\mathbb{M}^{\#n}\), of \(\mathbb{M}^\#\), as obtained simply by “unwrapping” the \(S^1\) of \(\mathbb{M}\)'s topology \(S^1 \times S^3\) to the required degree \(n\).

In fact, this strange-looking 4-fold cover \(\mathbb{M}^{\#4}\) of compactified Minkowski space can be understood explicitly in terms of the geometrical representation of a null twistor \(Z^\alpha\) in Minkowski space-time terms. We recall that a null twistor describes not just a ray \(Z\) in \(\mathbb{M}\), but also a spinor scaling, defined by \(\pi_A\), assigned to the null direction at each point of the ray \(Z\), where we think of this spinor scaling as parallel-transported along the ray \(Z\) within \(\mathbb{M}\). Now, we saw in §A2 (see figure 3) that a 2-spinor has a \(U(1)\) phase that is geometrically described
by a null flag half-plane, where if the flag plane is rotated about the flagpole through $2\pi$, the spinor changes sign. However, when we think of this flag half-plane as being parallel-transported all the way from $\mathcal{S}^-$ to $\mathcal{S}^+$ along $Z$, we find that if we were to try to match $\mathcal{S}^+$ directly to $\mathcal{S}^-$, as indicated in figure 5b, then we would find a discrepancy of a rotation through $\pi$, i.e. the flags would point in the opposite directions from one another across $I^+$. (This geometry is explained explicitly in [18], §9.4; see particularly Fig. 9-11.) Since a rotation of a spinor flag-plane by $2\pi$ results in a change of sign for the spinor, we need $4\pi$ to get it back to its original value. The rotation through $\pi$ that we find when we pass across $I^+$, represents a discrepancy of $i$ in the geometrical description of a null twistor in the space $\mathbb{M}^\#$. Moreover, the problem is not removed if we consider the flag-plane interpretation within just the 2-fold cover $\mathbb{M}^\#_2$, since we still have a sign discrepancy. Only when we pass to the 4-fold cover $\mathbb{M}^\#_4$ do we get a fully consistent picture of a null twistor—and, indeed, of a non-null twistor (see [18]), whose interpretation we turn to next.

B4. The basic twistor spaces

Let us now consider how to represent a non-null twistor $Z^\alpha$ in a geometrical way. It is best to think in terms of the family of null twistors $Y^\alpha$ that are orthogonal to $Z^\alpha$ in the sense that

$$Z^\alpha \bar{Y}_\alpha = 0$$

(or, equivalently $Y^\alpha \bar{Z}_\alpha = 0$). If $Z^\alpha$ were a null twistor—where $Y^\alpha$ is given as a null twistor—these respectively representing rays $Z$ and $Y$, then this vanishing of their scalar product asserts that these rays intersect (perhaps at infinity). Accordingly, if $Z$ is fixed, then this condition on $Y$ tells us that the $Y$ belongs to the special Robinson congruence defined by the ray $Z$. Now, let $Z$ be a fixed non-null twistor (but where $Y$ remains null). Then the congruence of $Y$-rays subject to orthogonality with $Z$ will provide a general Robinson congruence. See [18] for details.

As noted in §B3, the space $\mathbb{T}$ of all twistors $Z^\alpha$ is a 4-dimensional complex vector space, with pseudo-Hermitian scalar product $(Z^\alpha \bar{Y}_\alpha)$ of split signature $(++--)$. Geometrical notions are often best expressed in terms of the projective twistor space $\mathbb{P} \mathbb{T}$ of twistors up to proportionality, this being a complex projective 3-space $\mathbb{C} \mathbb{P}^3$. This compact complex manifold $\mathbb{P} \mathbb{T}$—or, more strictly, in accordance with the above discussion, the $\mathbb{C} \mathbb{P}^3$ of dual projective twistors $\mathbb{P} \mathbb{T}^*$—can indeed be identified with the space of Robinson congruences referred to above. The dual twistor space $\mathbb{T}^*$ is identified with the complex conjugate space $\bar{\mathbb{T}}$ of $\mathbb{T}$ via this pseudo-Hermitian structure. The points of the dual projective space $\mathbb{P} \mathbb{T}^*$ represent the complex projective planes within $\mathbb{P} \mathbb{T}$. The complex projective lines within $\mathbb{P} \mathbb{T}$ correspond to points of the complexified compactified Minkowski space $\mathbb{C} \mathbb{M}^\#$.

Whereas, generally speaking, it is the projective twistor space $\mathbb{P} \mathbb{T}$ that is useful to us if we are thinking of geometrical matters, the space $\mathbb{T}$ is appropriate if we are concerned with the algebra of twistors. For a non-zero twistor $Z^\alpha$, we can have three algebraic alternatives. These are:

$$Z^\alpha \bar{Z}_\alpha > 0,$$

for a positive or right-handed twistor $Z^\alpha$, belonging to the space $\mathbb{T}^+$,
$Z^\alpha\bar{Z}_\alpha<0$, for a negative or left-handed twistor $Z^\alpha$, belonging to the space $\mathbb{T}^\pm$.

$Z^\alpha\bar{Z}_\alpha=0$, for a null twistor $Z^\alpha$, belonging to the space $\mathbb{N}$.

The entire twistor space $\mathbb{T}$ is the disjoint union of the three parts $\mathbb{T}^+$, $\mathbb{T}^-$, and $\mathbb{N}$, as is its projective version $\mathbb{P}\mathbb{T}$ the disjoint union of the three parts $\mathbb{P}\mathbb{T}^+$, $\mathbb{P}\mathbb{T}^-$, and $\mathbb{P}\mathbb{N}$ (see figure 6).

The way that the various parts of twistor space $\mathbb{T}$ relate to their various projective counterparts of $\mathbb{P}\mathbb{T}$.

Each point of $\mathbb{P}\mathbb{T}$ represents a 1-dimensional vector subspace of $\mathbb{T}$. The points of $\mathbb{C}\mathbb{M}^\#$ are thus described by 2-complex-dimensional subspaces of $\mathbb{T}$ (complex lines in $\mathbb{P}\mathbb{T}$). The points of $\mathbb{M}^\#$ are described by 2-complex-dimensional subspaces in $\mathbb{N}$, i.e. by complex lines in $\mathbb{P}\mathbb{N}$. Robinson congruences are represented by the intersections of complex projective planes in $\mathbb{P}\mathbb{T}$ with $\mathbb{P}\mathbb{N}$. In the case of a special Robinson congruence, of rays meeting a particular ray $L$, the complex plane in $\mathbb{P}\mathbb{T}$, the complex plane has contact with $\mathbb{P}\mathbb{N}$ at a point representing the ray $L$ in $\mathbb{M}^\#$.  

**B5. Helicity and relativistic angular momentum**

It is the space $\mathbb{P}\mathbb{N}$ that has the most direct physical interpretation, since its points correspond to world-lines of free classical massless particles, which we can think of as the classical histories of (point-like) photons in free motion, though possibly at infinity, as a limiting case in Minkowski space-time $\mathbb{M}$; see figure 7. As stated above, not only are the points of *complexified* Minkowski space $\mathbb{C}\mathbb{M}$ represented as (complex projective) lines in $\mathbb{P}\mathbb{T}$, but so also are all the points of the complexified *compactified* Minkowski space $\mathbb{C}\mathbb{M}^\#$. Those lines that lie in $\mathbb{P}\mathbb{N}$, represent points of the real space-time $\mathbb{M}$ (possibly at infinity), but since these lines are still complex projective lines, they are indeed *Riemann spheres*, in accordance with the ambitions put forward in §A2; see figure 7.
The most immediate part of the twistor correspondence: a ray $Z$ in Minkowski space $\mathbb{M}$ corresponds to a point in $\mathbb{P}\mathbb{N}$; a point $x$ of $\mathbb{M}$ corresponds to a Riemann sphere $X$ in $\mathbb{P}\mathbb{N}$.

In figure 8 this picture is extended to include a physical interpretation of non-null twistors, where points of $\mathbb{P}\mathbb{T}^+$ and $\mathbb{P}\mathbb{T}^-$ are represented, in Minkowski space, as though they are light rays with a twist about them. This is schematic, but indeed these points can be regarded as representing massless particles with spin. In relativistic physics, if a massless particle has a non-zero spin, the “spin-axis” must be directed parallel or anti-parallel to the particle’s velocity. We say that the particle has a helicity $s$, that can be positive or negative (or zero, for a spinless massless particle). If $s \neq 0$, then the particle’s space-time trajectory is not precisely defined (in a relativistically invariant way) as a world-line, but can be specified in terms of its 4-momentum $p_a$ and 6-angular momentum $M^{ab}$ about some chosen space-time origin point $O$. These must be subject to

$$p_a p^a = 0, \quad p_0 > 0, \quad M^{(ab)} = 0, \quad \frac{1}{2} \varepsilon_{abcd} p^b M^{cd} = s p_a$$

(curved or square brackets around indices respectively denoting symmetric or anti-symmetric parts), where $\varepsilon_{abcd} = \varepsilon_{abcd}$ is the Levi-Civita tensor fixed by its component value $\varepsilon_{0123} = 1$ in a right-handed orthonormal Minkowskian frame (with time-axis basis vector $\delta_0^A$, so $p_0$ is the particle’s energy in units where the speed of light $c = 1$). Note that $\frac{1}{2} \varepsilon_{abcd} M^{cd} = \ast M_{ab}$ is the Hodge dual of $M^{ab}$, so the second displayed equation becomes $\ast M_{ab} p^b = s p_a$. The connection between these quantities and twistor theory is that if

$$Z^a = (\varpi^A, \pi_A)$$

then we can make the interpretation

$$p_{\lambda A} = \pi_A \bar{\varpi}^A, \quad M^{AB, A'B'} = i \varpi^{(A} \bar{\varpi}^{B)} e^{A'B'} - i \bar{\varpi}^{(A} \varpi^{B)} e^{A'B'},$$

and all the above conditions are automatically satisfied, provided that $\pi_A \neq 0$. Conversely, the twistor $Z^a$ (with $\pi_A \neq 0$) is determined, uniquely up to a phase multiplier $e^{i \theta}$, by $p_a$ and $M^{ab}$, subject to these conditions. The helicity $s$ finds the very simple (and fundamental) expression

$$2s = \varpi^A \bar{\varpi}_A + \pi_A \varpi^A = Z^a \bar{Z}_a.$$
There is, however, the subtlety referred to above that in this interpretation of a non-null twistor, when the helicity \( s \) is non-zero, there is no actual world-line that can describe the particle’s location in a relativistically invariant way, the world-line being, in a sense, “spread out” in accordance with the configuration depicted in figure 1, as we shall see at the end of §B5. This issue is an important undercurrent to the application of twistor ideas in general relativity as discussed in D3 and D4.

Fig. 8:

Classical massless particles with positive (right-handed) helicity can be represented as points of \( \mathbb{PT}^+ \) and those with negative (left-handed) helicity, as points of \( \mathbb{PT}^- \).

B6. Description under shift of origin

Under a displacement of the origin \( O \) to a new point \( Q \) of \( \mathbb{M} \),

\[
O \mapsto Q,
\]

where the position vector \( \overrightarrow{OQ} \) is (in abstract-index form) \( q^a \), the spinor parts of the twistor \( Z^\mu=(\omega^A, \pi_A) \) undergo

\[
\omega^A \mapsto \omega^A - i q^{AA'} \pi_{A'}, \quad \pi_{A'} \mapsto \pi_{A'}.
\]

For a dual twistor \( W_\alpha=(\lambda_A, \mu^A) \), we correspondingly have

\[
\lambda_A \mapsto \lambda_A, \quad \mu^A \mapsto \mu^A + i q^{A\alpha'} \lambda_{\alpha'}.
\]

This turns out to be consistent with the standard transformation of \( M^{\mu\nu} \) (and \( p_a \)) under origin change, where the position vector \( x^\alpha \) of a space-time point \( X \) correspondingly undergoes

\[
x^\alpha \mapsto x^\alpha - q^\alpha.
\]

At this juncture, it is worth pointing out that whereas there may be a temptation to identify a twistor, as represented as a pair of 2-spinors \( (\omega^A, \pi_A) \), with the 2-spinor form of a Dirac 4-spinor (see [11], [19]), which it superficially resembles, this would be entirely inappropriate. It is certainly true that the behavior of twistors and Dirac4-spinors is basically the same under Lorentz transformations (leaving the origin fixed), as is defined by their spinor-index structures. But the above behavior of a twistor under shift of origin (in effect, under translation) has no place in Dirac’s electron theory. As we have seen in §B3, twistors provide the basic (finite-dimensional) representation space for the (restricted) conformal group \( C(1,3) \),
whereas Dirac spinor fields provide an infinite-dimensional representation space for the (restricted) Poincaré group. The normal Dirac equation describes a massive particle, and is not conformally invariant. The Dirac-Weyl equation for a massless neutrino is, however, and its relation to twistor theory is contained in the discussion given in §C3. Massive particles can also be handled with twistor theory, but such descriptions normally require more than one twistor. See [18], [20], [21], [22], [23]; see also the remarks given at the end of §D4.

There is a connection between the above direct physical interpretation of a twistor in terms of angular momentum—particularly a non-null twistor $Z^a$—and the Robinson congruence defined by $Z^a$. This congruence is provided by the family of rays defined by the null (dual) twistors $W^a=(\lambda^A, \mu^A)$, satisfying

$$Z^a W_a = 0.$$  

To see the connection with angular momentum, let us examine this relation at an arbitrary point $Q$ of $\mathbb{M}$, where we now take $Q$ as a (variable) origin point. We are interested in the ray $W$ of the congruence which passes through $Q$. With respect to $Q$, as origin, $W_a$ then takes the form

$$W_a = (\lambda_A, 0)$$

($\mu^A$ being zero, since $W_a$ is now incident with the origin point $Q$; see §B2, and also figure 4 in complex-conjugate form). Accordingly, the relation $Z^a W_a=0$ now becomes

$$\omega^A \lambda_A = 0,$$

at the point $Q$. This tells us that the flagpole direction of $\omega^A$ is the same as that of $\lambda^A$, namely the direction of the ray $W$. Tus, the angular momentum $M^{ab}$ of the spinning massless particle determined by $Z^a$ has a structure that is characterized by the flagpole directions of its two spinor parts with respect to $Q$. We may refer back to figure 1, to see the curious spatial geometry of all this, where the flagpole directions of $\omega^A$ (small arrows in figure 1) twist around in this complicated (Robinson congruence) way, while that of $\pi_A$ simply points in the direction of motion of the configuration (large arrow at the top right of figure 1). It may perhaps be mentioned, that the choice of letters “$\pi$” and “$\omega$” come from the normal usage of “p” for momentum, and “omega” for angular momentum.
C1. Twistor quantization rules

Up to this point, we have been considering twistor theory only in relation to classical physics in flat space-time geometry. Quantum twistor theory—and, indeed, as we shall be seeing later (in Part D), also space-time curvature— involves considering twistors (and dual twistors) as non-commuting operators, satisfying certain commutation laws:

\[ Z^\alpha \tilde{Z}_\beta - \tilde{Z}_\beta Z^\alpha = \hbar \delta_\beta^\alpha \]

and, as far as our current considerations go,

\[ Z^\alpha Z^\beta - Z^\beta Z^\alpha = 0, \quad \tilde{Z}_a \tilde{Z}_b - \tilde{Z}_b \tilde{Z}_a = 0 \]

[18], [24]. Now, the twistors are taken to be linear operators generating a non-commutative algebra \( \mathbb{A} \), whose elements are taken to be acting on an appropriate quantum “ket-space” […] of some kind [25], but it is best not to be specific about this, just now. We could alternatively think of our operators as dual twistors, subject to the commutation laws

\[ W_a \tilde{W}^\beta - \tilde{W}^\beta W_a = -\hbar \delta_\beta^\alpha \]

and

\[ W_a W_\beta - W_\beta W_a = 0, \quad \tilde{W}^a \tilde{W}_\beta - \tilde{W}_\beta \tilde{W}^a = 0, \]

which is the same thing as before, but with \( \tilde{Z}_a \) re-labelled as \( W_a \).

These commutation laws are almost implied by the standard quantum commutators for 4-position and 4-momentum

\[ p_a x^b - x^b p_a = i \hbar \delta^b_a, \quad x^a x^b - x^b x^a = 0, \quad p_a p_b - p_b p_a = 0, \]

but there appears to be an additional input related to the issue of helicity. By direct calculation, we may verify that the twistor commutation laws reproduce exactly the (more complicated-looking) standard commutation laws for the \( p_a \) and \( M^{ab} \) as defined in §B5, which arise from their roles as translation and Lorentz-rotation generators of the Poincaré group (see [18]). In this calculation, there is no factor-ordering ambiguity in the expressions for \( p_a \) and \( M^{ab} \) in terms of the spinor parts of \( Z^\alpha \) and \( \tilde{Z}_a \) (owing to the symmetry brackets). Yet, the calculation for the helicity \( s \) (writing the operator as \( s \)) yields:

\[ s = \frac{1}{4} (Z^\alpha \tilde{Z}_a + \tilde{Z}_a Z^\alpha) \]
**C2. Twistor wave-functions**

If we are to express wave-functions for massless particles in twistor terms, to be in accordance with standard quantum-mechanical procedures we need functions of $Z^\alpha$ that are “independent of $\bar{Z}^\beta$”. This means “annihilated by $\partial/\partial \bar{Z}^\beta$”, i.e. holomorphic in $Z^\alpha$ (Cauchy–Riemann equations). Thus, a twistor wave-function (in the $Z^\alpha$-description) is holomorphic in $Z^\alpha$, and the operators representing $Z^\alpha$ and $\bar{Z}^\alpha$ act as:

$$Z^\alpha \mapsto Z^\alpha \times, \quad \bar{Z}^\alpha \mapsto -\hbar \frac{\partial}{\partial \bar{Z}^\alpha}.$$  

Alternatively, we could be thinking of functions of $\bar{Z}^\alpha$ that are “independent of $Z^\beta$”, i.e. anti-holomorphic in $Z^\alpha$. Here it would be better to re-name $\bar{Z}^\alpha$ as $W^\alpha$ and consider functions holomorphic in $W^\alpha$. Accordingly, in the dual twistor $W^\alpha$-description, a wave-function must be holomorphic in $W^\alpha$ and we have the operators representing $\bar{W}^\alpha$ and $W^\alpha$, again satisfying the required commutation relations, but now with:

$$\bar{W}^\alpha \mapsto \hbar \frac{\partial}{\partial W^\alpha}, \quad W^\alpha \mapsto W^\alpha \times.$$  

To ask that our wave-function describe a (massless) particle of definite helicity, we need to put it into an eigenstate of the *helicity operator* $s$, which, by the above, is

$$s = -\frac{1}{2} \hbar (Z^\alpha \frac{\partial}{\partial Z^\alpha} + 2)$$

in the $Z^\alpha$-description, and

$$s = \frac{1}{2} \hbar (W^\alpha \frac{\partial}{\partial W^\alpha} + 2)$$

in the $W^\alpha$-description. These are simply displaced *Euler homogeneity operators*

$$Y = Z^\alpha \frac{\partial}{\partial Z^\alpha} \quad \text{or} \quad \bar{Y} = W^\alpha \frac{\partial}{\partial W^\alpha},$$

so a helicity eigenstate, with eigenvalue $s$, in the $Z^\alpha$-description requires a holomorphic twistor wave-function $f(Z^\alpha)$ that is homogeneous of degree

$$n = -2s - 2,$$

where I henceforth adopt $\hbar=1$. Then $2s$ is an integer (odd for a fermion and even order for a boson). In the $W^\alpha$-description, the dual twistor wave-function $\tilde{f}(W^\alpha)$ is homogeneous of degree $\tilde{n}$ where

$$\tilde{n} = 2s - 2.$$
C3. Twistor generation of massless fields and wave-functions

In ordinary flat space-time terms, the position-space wave-function of a massless particle of helicity $2s$ ([6], [11], [12]) satisfies a field equation, this being expressible in the 2-spinor form

$$\nabla^{AA'}\psi_{AB\ldots E} = 0, \quad \square \psi = 0, \text{ or } \nabla^{AA'}\tilde{\psi}_{A'B'\ldots E'} = 0,$$

for the integer $2s$ satisfying $s<0$, $s=0$, or $s>0$, respectively, these equations having been already considered in §A3, but where the scalar case $s=0$ is included also, involving the D’Alembertian

$$\square = \nabla_u \nabla^u.$$

We have total symmetry for each of the $|2s|$-index quantities

$$\psi_{AB\ldots E} = \psi_{(AB\ldots E)} \quad \text{and} \quad \tilde{\psi}_{A'B'\ldots E'} = \tilde{\psi}_{(A'B'\ldots E')}.$$

What is the connection between the holomorphic twistor wave-function $f(Z')$, or dual twistor wave-function $\tilde{f}(W_z)$, with these space-time equations? In most direct terms this is given by contour integrals [21], [24], [26], generalizing earlier expressions found by Whittaker [27] and Bateman [28], [29].

$$\psi_{AB\ldots E}(x) = k \int_{\omega = i\pi} \frac{\partial}{\partial \omega} \frac{\partial}{\partial \omega} \ldots \frac{\partial}{\partial \omega} f(\omega, \pi) \delta\pi, \quad \text{if } s \leq 0;$$

$$\tilde{\psi}_{A'B'\ldots E'}(x) = k \int_{\omega = i\pi} \pi_A \pi_B \ldots \pi_E f(\omega, \pi) \delta\pi, \quad \text{if } s \geq 0.$$

Here $AB\ldots E$ or $A'B'\ldots E'$ are $|2s|$ in number, and the 1-form $\delta\pi$ is

$$\delta\pi = \varepsilon^{FG} \pi_F d\pi_G,$$

and where $k$ and $k'$ are suitable constants. I have taken the liberty of writing $x^a$, $\omega^A$, and $\pi_A$ without their abstract indices in places here, and using bold-face upright type instead. The contour, for these integrals, lies within the Riemann sphere, in $\mathbb{P}^T$, of twistors $Z^\alpha=(\omega^A, \pi_A)$ satisfying the incidence relation $\omega^A=x^{AA'}\pi_A$ (written $\omega=i\pi$, below the integral sign), which removes the $\omega^A$-dependence and introduces $x^A$-dependence, and then the contour integration itself removes the $\pi_A$-dependence, leaving us with just $x^A$-dependence. Satisfaction of the field equations is an immediate consequence of these holomorphic expressions. The 2-form $d\pi_0 \wedge d\pi_1 = \frac{1}{2} d\delta\pi$ is sometimes more appropriate to use, rather than $\delta\pi$, the contour then being 2-dimensional, lying in $\mathbb{T}$ rather than $\mathbb{P}^T$. In the dual twistor description, we have corresponding expressions.

See figure 8 for the geometrical set-up, where “X” is the Riemann sphere representing the (possibly complex) point labelled $x^a$. In this particular picture, X represents a complex pint

---

1 The upper one of these two displayed integral expressions was put forward by Lane Hughston [21], as a complement to the lower one, which I had found earlier [24], [26]. The significance of having both ways of doing it was not recognized immediately, but it was later realized that both are required or the complete picture.
lying in a region of $\mathbb{CM}$ referred to as the *forward tube* $\mathbb{CM}^+$, this being given by complex position vectors whose imaginary parts $s$ are past-pointing timelike. In twistor terms such points are represented by lines lying entirely in $\mathbb{PT}^+$, and we see this depicted in figure 8. This particular arrangement is important for quantum mechanics, because it is a convenient characterization of the important requirement, for a wave function, of energy positivity (see [30] and see also the comments at the end of §A3).

Fig. 8: The geometry relevant to the twistor contour integral for a wave-function. The regions $Q_1$ and $Q_2$ of the text are the respective complements, within $\mathbb{PT}^+$, of the depicted regions $\mathcal{V}_2$ and $\mathcal{V}_1$. Here, the open sets $\mathcal{V}_1, \mathcal{V}_2$ provide a 2-set open covering of $\mathbb{PT}^+$ and $\mathcal{R}$ is their intersection.

**C4. Singularity structure for twistor wave-functions**

For such expressions to provide non-zero answers, the function $f$ must have appropriate singularities. The situation of specific interest to us here is the case of a wave-function for a free massless particle, although these formulae can also be used under many other circumstances, such as for real solutions of Maxwell’s equations in particular domains. Real solutions can clearly be obtained from the complex ones described here, by taking the real part, the equations to be satisfied being linear. Completely general solutions of the equations are obtained in this way provided that they are analytic. In fact, precursors of these equations were found long ago, for the Laplace equation by Whittaker [27] in 1903, and for the wave equation (in 1904) by Bateman [28] who later generalized it for the Maxwell equations in the 1930s, see [29].

As remarked at the end of §C3, for a wave-function, we require complex solutions of *positive frequency*, and here is where the important early motivation for twistor theory referred to at the end of §A4 was, in a sense, finally satisfied. But, as initially presented, this was only in a way that seemed somewhat odd. Eventually this apparent oddness was re-interpreted as something remarkably “natural” when properly understood, with potentially deep implications.

Let us see how this works. First, we take note of the fact, already noted in §C3 that the family of points of $\mathbb{CM}$ that constitute the sub-region $\mathbb{CM}^+$ known as the forward tube, corresponds to the family of lines that lie entirely in $\mathbb{PT}^+$. A complex function $\psi$, defined on $\mathbb{M}$, which extends smoothly to a holomorphic function throughout $\mathbb{CM}^+$ is indeed of *positive frequency* and conversely, positive frequency being a key requirement for a wave-function [30]. Thus, for our twistor wave-function $f$, we require “regularity” of an appropriate sort throughout the region $\mathbb{PT}^+$. Yet it would be far too restrictive to demand holomorphicity for $f$ over the whole of $\mathbb{PT}^+$ and, in any case, such a function would simply give the answer zero when contour integrated. What we seem to need is a function with two separated regions of singularity on each Riemann sphere (complex projective line) that corresponds to a point in
CM⁺, i.e. to a projective line in ℙ⁺, since then we could obtain a non-trivial answer to the contour integration, the contour being a closed loop on the Riemann sphere that separates the two regions of singularity on the sphere. The situation is depicted on the right-hand side of figure 8. This is achieved if the singularities of \( f \) are constrained to lie in two disjoint regions \( Q_1 \) and \( Q_2 \) (each region being closed in \( \mathbb{T}^+ \)) so our contour integrations can take place within the holomorphic region \( \mathcal{R} \) between them (figure 8). Our twistor wave-function \( f \) is thus taken to be holomorphic throughout the (open) region

\[
\mathcal{R} = \mathbb{T}^+ - (Q_1 \cup Q_2).
\]

This, appears to be a somewhat odd requirement for the twistor description of such a fundamental thing as a massless particle’s wave-function. Moreover, the region \( \mathcal{R} \) is very far from being invariant under the holomorphic motions of \( \mathbb{T}^+ \), some of these representing the non-reflective Poincaré (inhomogeneous Lorentz) motions of Minkowski space \( \mathbb{M} \). Any particular choice of the region \( \mathcal{R} \) clearly cannot take precedence over any other such choice obtained from the original one by such a motion, so there is clearly much non-uniqueness involved in the choices of \( \mathcal{R} \) and \( f \) in this description. This difficulty looms large if we try to add two twistor wave-functions which might have incompatible singularity structures. Linearity is, after all a central feature of quantum mechanics as we currently understand that subject, so how are we to deal with this problem?

C5. Čech cohomology

The resolution of these puzzling features, leads us to an understanding of what kind of an entity a twistor wave-function actually “is”. This lies in the notion of Čech sheaf cohomology. It is not appropriate that we go into much detail, here, but some indication of the issues involved will be of importance for us. What we find is that the twistor wave-function \( f \) is not really to be viewed as being “just a function” in the ordinary sense, but as representing an element of “1st cohomology” (actually 1st sheaf cohomology). I shall call such an entity a 1-function. An ordinary function, in this terminology, would be a 0-function. There are also higher-order entities referred to as 2-functions, 3-functions, etc., but we shall not need to consider these here.

An important aspect of 1-functions (or of \( n \)-functions, where \( n > 0 \)) is that they are non-local entities in an essential way (a feature of twistor theory which appears to reflect aspects of non-locality that occur in quantum mechanics). A good intuitive way of appreciating the idea of a 1-function is to contemplate the “impossible tribar” depicted in figure 9. Here we have a picture that for each local region, there is an interpretation provided, of a 3-dimensional structure that is unambiguous, except for an uncertainty as to its distance from the viewer’s eye. As we follow around the triangular shape, our interpretation remains consistent (though with this mild-seeming ambiguity) until we return to our starting point, only to find that it has actually become inconsistent! The element of 1st cohomology that is expressed by the picture is a measure of this global inconsistency [31].
An impossible “tribar”, as an illustration of the notion of (1\textsuperscript{st}) cohomology. There is an unambiguous interpretation of each local part, except for an ambiguity as to the distance from the viewer’s eye, but globally this ambiguity leads to a non-local inconsistency. The measure of this inconsistency is an element of 1\textsuperscript{st} cohomology. Twistor wave-functions exhibit a similar feature, where the rigidity of analytic continuation replaces the rigidity of a material body.

How might we assign such a measure to the degree of this impossibility? I shall not go into full details here, but the idea is to regard the object under consideration—here the tribar—as being built up from a number of regions (open sets) which together cover the whole object, but which are “locally trivial” in some appropriate topological (or differential) sense. In the case of the tribar, we might have a local picture of each vertex, say $\mathcal{V}_1$, $\mathcal{V}_2$, $\mathcal{V}_3$, where the three pictures overlap pairwise in smaller open regions $\mathcal{V}_i \cap \mathcal{V}_j$, somewhere along each relevant arm of the tribar, so that taken together they provide a picture of the entire tribar. On each overlap region $\mathcal{V}_i \cap \mathcal{V}_j$, we require some numerical measure $F_{ij}$ which describes the ratio of the displacement from the eye that needs to be made for the pictures to be considered to match, and since we need an additive measure we take $F_{ij}$ to be the logarithm of this ratio, and accordingly the $F_{ij}$ are anti-symmetric ($F_{ij} = -F_{ji}$), since the ratio goes to its reciprocal when taken in the opposite order. The triple $(F_{12}, F_{23}, F_{31})$, taken modulo the particular triples of the form $(H_1 - H_2, H_2 - H_3, H_3 - H_1)$ where $H_i$ refers to the freedom that is inherent in the interpretation of each particular vertex picture $\mathcal{V}_i$. The resulting algebraic notion gives us the required cohomology element, describing the degree of impossibility in the figure. This notion is what I am calling a 1-function. For further issues see [32].

This is just to give a little flavor of what sort of an entity a 1-function actually is. More specifically, in the context of twistor theory, we are concerned with complex spaces and holomorphic functions on them. Thus, in the case of a twistor wave-function there is the important subtlety, in that the global “impossibility” arises from the “rigidity” of holomorphic functions rather than that of the solid structures conjured up by the local parts of figure 9. But let us be a bit more general here, and imagine some complex manifold $\mathcal{K}$. We shall need a locally finite open covering $\mathcal{C} = (\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \ldots)$, of $\mathcal{K}$. To define a 1-function $f$, with respect to $\mathcal{C}$, we assign a holomorphic function $f_{ij}$ on each non-empty pairwise intersection:

$$f_{ij} = -f_{ji} \text{ is holomorphic on } \mathcal{V}_i \cap \mathcal{V}_j,$$

and on each non-empty triple intersection:

$$f_{ij} + f_{jk} + f_{ki} = 0 \text{ on } \mathcal{V}_i \cap \mathcal{V}_j \cap \mathcal{V}_k,$$

where the collection $\{f_{ij}\}$ is taken modulo corresponding collections of the form $\{h_i - h_j\}$, where each
is holomorphic on $\mathcal{V}_i$,

so that two 1-functions are considered to be equal if the difference between their \{\(f_{ij}\)\} representations is of the form \(h_i - h_j\). This defines a 1-function with respect to the particular covering $\mathcal{C}$. For the full definition, we would have to take the direct limit for finer and finer coverings. Fortunately, in the case of complex manifolds, as is being considered here, we are assured that provided that the sets $\mathcal{V}_i$ are of suitable type (e.g. Stein spaces; see [33]) then we gain nothing from taking such a limit, and the 1-function concept is already with us. Nevertheless, in order to add two 1-functions defined by different coverings, we do need to take their common refinement in order to perform this operation, which can be a little complicated in practice.

In the case of main interest here, namely $\mathcal{K}=\mathbb{P}T^*$, it will be adequate for our immediate purposes here simply to take a 2-set covering of $\mathbb{P}T^*$, namely $\mathcal{C}=\{\mathcal{V}_1, \mathcal{V}_2\}$, with open sets given by the complements, within $\mathbb{P}T^*$ of the respective singularity regions $Q_2$ and $Q_1$. Then we have our required covering $\mathcal{C}$ (not actually with Stein spaces, but that is not of great importance here)

$$\mathbb{P}T^+ = \mathcal{V}_1 \cup \mathcal{V}_2, \quad \mathcal{R} = \mathcal{V}_1 \cap \mathcal{V}_2;$$

(see figure 8), just as we had earlier. The family \{\(f_i\)\} consists of the single twistor function $f$, which, by an abuse of notation I may identify with the 1-function it determines.

This interpretation in terms of 1st cohomology finally fully realized the motivation described in the final paragraph of §A3. The division described there, of the Riemann sphere into southern and northern hemispheres by its equator (representing the real numbers), where positive-frequency functions extend holomorphically into the southern hemisphere and negative frequency ones into the northern hemisphere is precisely reflected by the division of $\mathbb{P}T$ into $\mathbb{P}T^+$ and $\mathbb{P}T^-$ by the “equatorial” $\mathbb{P}\mathbb{R}$. The only essential difference, apart from the increase of dimensionality from the Riemann sphere $\mathbb{CP}^1$ to projective twistor space $\mathbb{CP}^3$, is that the holomorphic functions (0-functions) on the Riemann sphere are replaced by holomorphic 1st cohomology elements (1-functions) on projective twistor space.

In the cases of homogeneity 0 or 2 (left-handed electromagnetism or left-handed linearized gravity, respectively), there are generalizations of the twistor-space contour-integral expressions that allow one to view the 1-function nature of a twistor function in a different light, in which non-linearities of general relativity and particle physics begin to play a significant role. To appreciate this, let us return to the general Čech descriptions given earlier, where a locally finite open covering $\mathcal{C}=(\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \ldots)$ of a complex space $\mathcal{K}$ was considered. In the specification of a 1-function $f$, in relation to this covering, we required a family of holomorphic functions \{\(f_i\)\} defined on the non-empty overlaps $\mathcal{V}_i \cap \mathcal{V}_j$. Here, the functions are entirely passive, being just “painted on” the space $\mathcal{K}$. However, we can consider a somewhat more active role for such a 1-function $f$, such as (a) specifying the generation of a bundle above $\mathcal{K}$, or (b) using $f$ to specify the generation of a deformation of $\mathcal{K}$ itself. In each case, the rules (see [33]) defining a 1-function are exactly what is needed to fulfil this purpose. However, in each case, this specification by a 1-function would only be as an infinitesimal generator of the bundle or deformed space (except for an Abelian group in case (a)) because of non-linearities.
Nevertheless, the general idea expressed in (a) and (b) still holds true; it is just that the linear nature of a 1-function ceases to hold. In effect, we have a kind of “non-linear 1-function”.

It was in 1977 that Richard Ward introduced the procedure indicated in case (a) above, first in the situation provided by the (left-handed) Maxwell equations, which allowed interactions of the field with charged particles to be considered. Almost immediately afterwards he showed how this procedure could be generalized to the (left-handed) Yang-Mills equations \[34\]. This turned out to have considerable importance in the theory of integrable systems (see, for example, \[32\], \[35\]). Shortly before all this, in 1976, procedure (b) had been introduced \[36\], to provide a twistorial representation of all conformally complex-Riemannian 4-manifolds which are anti-self-dual (i.e. \(\mathcal{P}_{ABC} = 0\); see end of §A3). When an additional simple condition is imposed, this provides not only a (complex) metric but automatically generates the general anti-self-dual solution of the Einstein vacuum equations, either without \[36\] or with a cosmological constant \(\Lambda\) \[37\].

### C6. Infinity twistors and Einstein’s equations

It is a fairly straight-forward procedure to generate the desired deformed twistor spaces satisfying the required conditions ensuring satisfaction of the Einstein \(\Lambda\)-vacuum equations (\(\Lambda\) being the cosmological constant, \(\Lambda=0\) allowed). Basically, what is required is to match appropriate portions of (non-projective) twistor space, while preserving the Euler operator

\[ Y = Z^\alpha \frac{\partial}{\partial Z^\alpha} \]

and the 2-form

\[ \Theta = I_{\alpha\beta} \, dZ^\alpha \wedge dZ^\beta \]

where the anti-symmetrical \(\Lambda\)-infinity twistor \(I_{\alpha\beta}\) (and its dual \(I^{\alpha\beta}\)) are given by

\[ I_{\alpha\beta} = \begin{pmatrix} \Lambda \varepsilon_{AB} & 0 \\ 0 & \varepsilon^{AB} \end{pmatrix}, \quad I^{\alpha\beta} = \begin{pmatrix} \varepsilon^{AB} & 0 \\ 0 & \frac{\Lambda}{6} \varepsilon^{AB} \end{pmatrix}. \]

We see that \(I^{\alpha\beta}\) and \(I_{\alpha\beta}\) are both complex conjugates and duals of one another:

\[ I_{\alpha\beta} = \overline{I^{\alpha\beta}}, \quad I^{\alpha\beta} = \overline{I_{\alpha\beta}}, \]

\[ I_{\alpha\beta} = \frac{1}{2} \varepsilon_{\alpha\beta\rho\sigma} I^{\rho\sigma}, \quad I^{\alpha\beta} = \frac{1}{2} \varepsilon^{\alpha\beta\rho\sigma} I_{\rho\sigma}, \]

where \(\varepsilon_{\alpha\beta\rho\sigma}\) and \(\varepsilon^{\alpha\beta\rho\sigma}\) are Levi-Civita twistors, fixed by their anti-symmetry and \(\varepsilon_{0123} = 1 = \varepsilon^{0123}\) in standard twistor coordinates. The preservation of \(Y\) and \(\Theta\) on the overlaps where \(\mathcal{V}_i\) as matched to \(\mathcal{V}_j\), is ensured, if we shift infinitesimally along the vector field

\[ I^{\alpha\beta} \frac{\partial f_{ij}}{\partial Z^\alpha} \frac{\partial}{\partial Z^\beta}, \]

where each \(f_{ij}\) has homogeneity degree 2 (i.e. \(Y f_{ij} = 2 f_{ij}\)), which corresponds to helicity 2). We can imagine exponentiating these infinitesimal deformations to obtain a finite one. In the case of a
2-set covering, we can achieve this explicitly by exponentiating the single function $f_{12}$, but with larger numbers of sets, we can encounter difficulties in satisfying the required condition on triple overlaps. A simpler procedure for satisfying the required condition of preserving $Y$ and $\Theta$ is to use generating functions, see [36]. This has particular relevance to the procedures of D4.

It is, however, not at all a direct matter to obtain the (complex) curved space-time $\mathcal{M}^C$ from the deformed twistor space $\mathcal{T}$, according to this construction. The points of $\mathcal{M}^C$ correspond to “lines” in $\mathbb{P}T$, that are completed Riemann spheres, stretching across from one patch to the other, or perhaps several patches if the covering involves more than two. These Riemann spheres are not easy to locate, in a general way, since they are determined by the global requirement that they be compact holomorphic curves within $\mathbb{P}T$ of spherical topology (and belonging to the correct topological family). The very existence of these “lines”, as I shall call the, together with the fact that they belong to a 4-parameter family of such lines (provided that the deformation from $\mathbb{P}T^*$, or from some other appropriate part of $\mathbb{P}T$, is not too drastic), depends upon key theorems by Kodaira and Kodaira-Spencer (see [38], [39]). The space whose points represent these lines is the required complex 4-manifold $\mathcal{M}^C$. Its complex conformal structure comes about from the simple fact that meeting lines in $\mathbb{P}T$ correspond to null separated points in $\mathcal{M}^C$, and the definition of its metric scaling comes about through use of the form $\Theta$. With this complex metric, the complex 4-manifold automatically satisfies the Einstein $\Lambda$-vacuum equations, and the construction provides the general anti-self-dual solution. This procedure has become known as the “non-linear graviton construction” [36], [37]. It has found numerous applications in differential geometry [32], [35].

At this point, it is worth emphasizing an essential but unusual feature of the non-linear graviton construction. This is that the “curvature” in the deformed twistor space (encoding $\mathcal{M}^C$’s actual curvature) is not local, in the sense that a small-enough neighbourhood of any point in the deformed twistor space is identical in structure with that of ordinary flat twistor space $\mathbb{T}$ (with the given $\Lambda$ assigned to it). The “curvature” in the deformed twistor space is a non-local feature of the space $\mathbb{P}T$, but in the construction of the “space-time” manifold $\mathcal{M}^C$, we consequently find genuine local curvature in the normal sense (Riemann curvature, Weyl curvature).

If we are to regard twistor theory as providing an overall approach to basic physics, however, then we must face up to the fundamental obstruction to progress that as confronted the phis programme for some four decades, namely what has become known as the “googly problem” (an apposite term borrowed from the game of cricket, for a ball bowled with a right-handed spin about its direction of motion, but bowled with an action that would appear to be delivering a left-handed spin). This comes about from the very nature of the construction. The points of $\mathbb{P}T$ have an interpretation within $\mathcal{M}^C$ as what are called “$\alpha$-surfaces” (totally null self-dual complex 2-surfaces) [36], and there would have to be a 1-complex-parameter family of such surfaces through each point of $\mathcal{M}^C$ (corresponding to the 1-parameter family of points on

---

2 The historical point must be made here that a key input to the development of the non-linear graviton construction was the introduction, in 1976, by Ezra T. Newman, of his notion of the $\mathcal{H}$-space, for an asymptotically flat space-time, as described initially in [40]; for more detail, see [41]. For some fascinating new developments of these ideas, relating $\mathcal{H}$-space to equations of motion, see [41].
each line of $\mathcal{PT}$). This would imply $\bar{\mathcal{P}}_{\mathcal{A}B\mathcal{C}D}=0$, i.e. $\mathcal{M}^c$ being conformally anti-self-dual. Thus, it is the very existence of points of the space $\mathcal{PT}$ that implies the existence of these troublesome $\alpha$-surfaces. This illustrates the fundamental underlying difficulty of the googly problem.

Nevertheless, it is clear that if twistor theory is to have any hope of providing a basis for fundamental physics, there needs to be a way around this problem. Many ideas for addressing it have been made over the years, often resorting to examining the twistor structure at infinity, where the geometry is simpler than that at finite regions (see, for example, [42]), but none has been able to achieve very much. The most successful approach has been that of ambitwistors [43], [44], that is based on complex null geodesics, modelled on twistor, dual twistor pairs $(Z^\alpha, W^\beta)$, subject to $Z^\alpha W^\alpha=0$. This enables complex-Riemannian 4-manifolds to be studied in relation to twistor-type ideas, and the Einstein vacuum equations to be examined in this light [43]. But it does not follow the twistor route of “non-linearizing” the 1-function description of quantum wave functions, where left- and right-handed helicities can be combined together to describe gravitational interactions and classical space-times in a way that directly relates to twistor wave functions.

It should be mentioned that even without progress on the googly problem twistor ideas have found many significant applications, mainly in pure mathematics (most particularly in certain areas of differential geometry, representation theory, and integrable systems (see, in particular, [32], [35]), but, relatively recently, also in physics, where great simplifications to calculations in high-energy scatterings have been obtained. The main impetus to this relatively recent work came from Edward Witten [45], who introduced several novel ideas, partly based on earlier work in publications by others [46], [47], [48], [49], [50], [51], [52], [53]), and which subsequently stimulated a great deal of further activity, e.g. [54], [55], [56], [57], [58], [59], [60], [61], which allowed twistor methods to help in enormously reducing the calculations of scattering amplitudes at very high energies (where all particles involved could be considered to be massless). Nevertheless, despite all this impressive work, relatively little of the deeper results of twistor theory has been incorporated, and it is my view that these are needed for twistor theory to realize its aims to become an underlying scheme for fundamental physics generally. In Part D, I outline what I now believe to be the appropriate direction for progress to be made in this regard.
Part D: Palatial Twistor Theory

D1. Basic ideas of palatial twistor theory

As has just been remarked upon (in §C6 above), it is the seeming need to have an unambiguous notion of a point in twistor space that appears to drive us inevitably to this incomplete and lop-sided approach to space-time in twistor theory, that an unresolved googly problem presents us with. It is fortunate, therefore, that there is a novel approach to generalizing the non-linear graviton construction, so that both helicities can be accommodated within the same general framework, and that classical conformal Lorentzian space-times should also come under the same umbrella. To understand the basic idea, let us first return to the procedure that we considered earlier, in the non-linear graviton construction of §C6, where to produce a suitably deformed twistor space, we “glue together” pieces of complex manifold, preserving the complex structure from patch to patch. The notion of “complex structure” can be encapsulated in terms of the algebra of holomorphic functions on each patch—or, technically, the “sheaf” of such functions, where we require holomorphicity throughout a small neighbourhood of each point. Now, we saw from the above that the basic conundrum was the existence of the actual points in each patch, since it was the interpretation of the points in \( \mathbb{P} \mathcal{T} \) that gave rise to the unwanted \( \alpha \)-surfaces. Thus, it would seem, we somehow need to find a way of matching the sheaves of algebras of holomorphic functions from one patch to another, without actually having “patches” that, in the ordinary sense, would consist of individual points. This will not do as it stands, however, because the function algebras already “know” the points, this being true so long as the algebras are, like the algebra of holomorphic functions on a region of twistor space, commutative.

This suggests that we take, instead, a holomorphic algebra that is non-commutative\(^3\) As we have seen in §C1, there is a natural non-commutative algebra in twistor theory, namely that generated (via complex linear combinations and products—these basic operations being allowed to be repeated many times—and the taking of appropriate limits) from the operators (see C2):

\[
Z^a \times \quad \text{and} \quad -\frac{\partial}{\partial z^a}.
\]

I shall refer to this algebra as \( \mathbb{A} \), the basic quantum twistor algebra for Minkowski space \( \mathbb{M} \). The idea is that, in some sense, we could patch together two (or more) “sub-regions” of the algebra \( \mathbb{A} \), analogous to the \( V_1 \) and \( V_2 \) of the non-linear graviton construction that, in an appropriate sense “cover” the entire algebraic structure of interest. This “patched” algebra \( \mathcal{A} \) is then to represent some open portion of the complexified space-time \( \mathcal{M}^\mathbb{C} \).

The essential idea is that the algebra \( \mathbb{A} \) can be thought of as a system of complex linear operators acting on holomorphic functions defined locally on some complex space which, initially, we think of as twistor space \( \mathbb{T} \). In Dirac’s quantum-mechanical terminology [25], \( \mathbb{T} \) is a “ket” space for the algebra \( \mathbb{A} \) of quantum operators. We are to think of \( \mathbb{A} \) as an abstract algebra that is not dependent upon this particular realization. For example, the same \( \mathbb{A} \) could also be thought of as the space of complex linear operators acting on holomorphic functions on

\(^3\) I am very grateful to Michael Atiyah for making me aware of this important requirement, in a brief conversation in 2013 (which happened to be at an occasion in Buckingham Palace—hence the name).
the dual twistor space $\mathbb{T}^*$, where the respective operators above would (as displayed in §C2) now be

$$\frac{\partial}{\partial W_\alpha} \quad \text{and} \quad W_\alpha \times$$

which satisfy the same commutation rules as before. Thus, in Dirac’s terminology, $\mathbb{T}^*$ would be an alternative ket space for $\mathbb{A}$.

A quantum-mechanical way of thinking about this would be to assert that the commuting complex parameters $Z^0, Z^1, Z^2, Z^3$ constitute what we may call a complete set of commuting operators, in the sense that they, together with their respective partial derivatives $\partial \partial Z^0, \partial \partial Z^1, \partial \partial Z^2, \partial \partial Z^3$, generate, in an appropriate sense, the algebra $\mathbb{A}$. A different representation of the same algebra $\mathbb{A}$ would be obtained if we chose, instead, the complete set of commuting operators $W_0, W_1, W_2, W_3$, where the formal component replacements $W_\alpha \mapsto -\partial \partial Z_\alpha$ and, accordingly, $\partial \partial W_\alpha \mapsto Z_\alpha$ are made. In this sense, $\mathbb{T}$ and $\mathbb{T}^*$ would provide alternative ket spaces for the same algebra $\mathbb{A}$. There would be many other possible choices of ket space for the same algebra $\mathbb{A}$.

In order to obtain something analogous to the non-linear graviton construction we would seem to have to involve something analogous to a connected open subset of such a ket space $\mathbb{T}$. The general idea would be to patch together various different such “ket patches”, by analogy with the patching together of different open regions in a locally finite covering $\mathcal{C}=(\mathcal{V}_1, \mathcal{V}_2, \mathcal{V}_3, \ldots)$, if we are to build up a non-trivial (complex) manifold. In an appropriate sense, the corresponding algebras $\mathbb{A}_1, \mathbb{A}_2, \mathbb{A}_3, \ldots$ would have to “agree” on overlaps (locally isomorphic on the overlaps, in some suitable sense.), but the ket spaces would generally differ from patch to patch, so that the “patched-up” algebra $\mathbb{A}$ that we end up would have no overall ket space, and would therefore differ from $\mathbb{A}$, though agreeing with it in a local sense (see [62] and [63] for earlier tentative descriptions of this idea).

However, there are various thorny issues that need to be faced, with regard to this sort of “patching” if we are asking for a duly rigorous picture. We might, for example, consider some sub-region $\mathcal{X}$ of twistor space $\mathbb{T}$, which we propose to use as a ket-space “patch”. The operator $\exp(A^\alpha \partial \partial Z_\alpha)$ for constant $A^\alpha$, in particular, would appear to be harmless enough, but it could present us with difficulties, as a candidate for membership of the algebra whose ket space is to be $\mathcal{X}$. A problematic issue is that a holomorphic function $f$, defined on $\mathcal{X}$, whose analytic continuation from inside $\mathcal{X}$ to a point displaced by the vector $A^\alpha$ to a somewhere outside $\mathcal{X}$, where this analytic continuation of $f$ becomes singular, would exclude $\exp(A^\alpha \partial \partial Z_\alpha)$ from membership of the algebra, since its action on $f$ would be singular. On the other hand, the operation of multiplication by $\exp(A^\alpha W_\alpha)$, on a ket space that is any sub-region of $\mathbb{T}^*$ would be completely harmless. Such issues need to be better understood for a properly rigorous picture of this intended procedure to be obtained.

It is clear from all this that there is a considerable vagueness in this proposal, as put forward above. Most particularly, we do not have a clear notion of topological issues, such as “local” and “open set”, when it comes to these algebras. These difficult issues are not properly resolved as things stand, and in the following sections I shall adopt a policy of providing explicit procedures for the needed patching without worrying about topological issues, complete rigour, and so forth, leaving such matters to later consideration. Nevertheless, I
believe that we can go some way towards addressing these topological matters if we, in a sense “stay close to the space $\mathbb{P}\mathcal{N}$”, where $\mathbb{P}\mathcal{N}$ is the space whose points represent null geodesics in the space-time $\mathcal{M}$ that is intended that we are describing in twistor terms. This will be explored in the next section.

**D2. The spaces of momentum-scaled and spinor-scaled rays**

In accordance with this, let us indeed explore the ray-space $\mathbb{P}\mathcal{N}$, whose points represent individual rays in a space-time $\mathcal{M}$, where $\mathcal{M}$ is taken to be a smooth time-oriented Lorentzian globally hyperbolic 4-manifold. Thus $\mathbb{P}\mathcal{N}$ plays the same role for $\mathcal{M}$ as does $\mathbb{P}\mathcal{N}$ for Minkowski space $\mathbb{M}$. The global hyperbolicity of $\mathcal{M}$ (equivalent to $\mathcal{M}$ containing some global spacelike 3-surface which can act as a Cauchy hypersurface for physical fields within $\mathcal{M}$; see [64]) ensures that $\mathbb{P}\mathcal{N}$ is Hausdorff and that $\mathcal{M}$ contains no causal anomalies such as closed or “almost closed” rays [64]. We are to imagine that our proposed “palatial space” $\mathbb{P}\mathcal{T}$ can, in some appropriate sense, be viewed as being some sort of “extension” of $\mathbb{P}\mathcal{N}$, to include non-null twistors, though not in any direct sense as an ordinary manifold.

In addition, we shall be interested not only in the 5-manifold $\mathbb{P}\mathcal{N}$ of rays in $\mathcal{M}$, but also in the 6-manifold $\mathbb{P}\mathcal{N}$ of momentum-scaled rays in $\mathcal{M}$. Thus, each point of $\mathbb{P}\mathcal{N}$ represents a ray $\gamma$ together with a momentum co-vector $p_a$ which is parallel-propagated along $\gamma$, the future-null vector $p^a$ being tangent to $\gamma$. We shall be interested also in the smooth Hausdorff 7-manifold $\mathcal{N}$ representing the spinor-scaled rays $\gamma$, where, in addition to the momentum scaling $p_a$, we attach a spinor phase provided by the flag plane of $\pi_A$, where

$$p_a = \bar{\pi}_A \pi_{A'},$$

providing a spinor scaling for $\gamma$ (see §B2), these requirements being all conformally invariant (see [58] Chapter 7). Importantly, this is just the arrangement of spaces required for the procedure of geometric quantization as described (explicitly, in relation to twistor theory) in Woodhouse’s book [65], and we shall be seeing the significance for us of some of the ideas of geometric quantization in the next section, §D3. Explicitly: the 6-space $\mathbb{P}\mathcal{N}$ is a symplectic manifold, and $\mathcal{N}$ is a circle bundle over $\mathbb{P}\mathcal{N}$, the circles given by the phase freedom in $\pi_{A'}$.

I shall write the symplectic 2-form $\Sigma$ of the space $\mathbb{P}\mathcal{N}$ as

$$\Sigma = i \, d\zeta^\alpha \wedge d\bar{\zeta}_\alpha$$

in anticipation of a role for twistor notation for $\mathcal{N}$, where we take note of the fact that in the case of flat twistor space $\mathbb{T}$ we then find

$$\Sigma = dp_a \wedge dx^a$$

(by a brief calculation in the notation of §B2 for twistors in $\mathbb{M}$). This is the standard symplectic 2-form for the cotangent bundle of a manifold, where in this case we are thinking of the cotangent bundle of $\mathbb{M}$, symplectically reduced by the Hamiltonian $p_a p^a$ where we take $p_a p^a = 0$, so that the above indeed gives us $\Sigma$ as the symplectic structure of $\mathbb{P}\mathbb{M}$ (i.e. of $\mathbb{P}\mathcal{N}$ when $\mathcal{M} = \mathbb{M}$). See [18], [44] and [64] for details. This suggests that the twistor expression above, for the
2-form $\Sigma$ might perhaps also have meaning in the case of general Lorentzian-conformal 4-manifold $\mathcal{M}$. Indeed, a certain justification for this twistorial expression for $\Sigma$ is given in [18], §7.4, see fig 7-4, particularly.)

Up to this point, we have been fixing attention on what will be regarded as the region

$$Z^a \bar{Z}_a = 0,$$

in other words, the actual geometrically defined circle bundle $\mathcal{N}$ over the symplectic $\mathbb{P}\mathcal{N}$. We are now going to try to think of this region as somehow extendible as a manifold to where $Z^a \bar{Z}_a \neq 0$, but there cannot in general be any unique way of doing this. Nevertheless, there are indeed many way of doing it locally. The picture I am presenting is that we apply the ordinary notion of a locally finite open covering to $\mathbb{P}\mathcal{N}$, this extending to $\mathcal{N}$, by virtue of the spinor scaling, as described above. Let $(\mathbb{P}\mathcal{N}_1, \mathbb{P}\mathcal{N}_2, \ldots)$ be such an open covering of $\mathbb{P}\mathcal{N}$ extending, accordingly, to a locally finite open covering $(\mathcal{N}_1, \mathcal{N}_2, \ldots)$ of $\mathcal{N}-\{0\}$, where we shall require that in some sense—as yet to be specified more precisely—the individual spaces $\mathcal{N}_k$, together with their intersections $\mathcal{N}_i \cap \mathcal{N}_j$, are appropriately “simple” enough that the requirements below can be satisfied.

Whereas the patchings of these $\mathcal{N}$-regions are to be pointwise, in the ordinary sense, this cannot be expected to hold for the proposed “extensions” to $Z^a \bar{Z}_a \neq 0$. Instead, it will be that there is a (quantum) twistor algebra $\mathbb{A}_j$ defined for each patch $\mathcal{N}_j$, and it is these algebras that we shall require match on the overlaps between patches. A key issue is that we cannot necessarily expect the ket spaces for these algebras will match. For if they did, we would have a matching of actual spaces, which is just what we wish to avoid, for consistency with the aspirations expressed in §D1.

We shall be seeing shortly (in §B3) how the procedures of geometric quantization allow us, in a general way, to construct and match the required algebras $\mathbb{A}_j$, but before proceeding to this, we shall need a little more about how twistor ideas and notation can be used to describe the needed structure of these patches. For this, I shall ask that the geometric structure of $\mathcal{N}$ be analytic. This may not be essential, but it makes descriptions easier. Analyticity allows us to complexify $\mathcal{N}$ to a complex 7-manifold $\mathbb{C}\mathcal{N}$ (a complex “thickening” $\mathcal{N}$ which contains $\mathcal{N}$ as a real submanifold). Correspondingly, there will also be complex manifolds $\mathbb{C}\mathbb{P}\mathcal{N}$ and $\mathbb{C}\mathbb{P}\mathcal{N}$, where $\mathbb{C}\mathbb{P}\mathcal{N}$ is complex-symplectic and where $\mathbb{C}\mathcal{N}$ may be regarded as a holomorphic bundle over $\mathbb{C}\mathbb{P}\mathcal{N}$ whose fibre is an open annular region containing the unit circle in the complex (Wessel) plane.

We are to choose the open regions $\mathbb{C}\mathbb{P}\mathcal{N}_j$ small (i.e. “simple”) enough so that there can be no obstruction to mapping each $\mathbb{C}\mathcal{N}_j$ holomorphically to a subregion of the 7-complex-dimensional complexification $\mathbb{C}\mathbb{H}$ of $\mathbb{H}$, in a way that preserves the bundle over symplectic structure of each (there being no local notion of “curvature” for these local structures which could prevent this). This will allow us to use the ordinary twistor descriptions of Part B to describe the geometry and algebra of each $\mathbb{C}\mathcal{N}_j$, as inherited from standard twistor space $\mathbb{T}$. This mapping will be far from unique, of course, but the very freedom that is allowed by this non-uniqueness is an important issue for the construction.

Let us fix attention on one particular $\mathbb{C}\mathcal{N}_j$, and consider two alternative such twistor descriptions, which I denote by
We take $Z^\alpha$ and $W_\alpha$ to be independent twistors and dual twistors and, likewise, we take $Z^\alpha$ and $\mathcal{W}_\alpha$ to be independent. The idea is that whereas the descriptions to be given below should then lead us to the description of a complex space-time $\mathcal{M}^\mathbb{C}$, we can then specialize them, in a way that will be described in §D4, so as to describe real-Lorentzian space-times $\mathcal{M}$, by being able to revert to the required Hermiticity:

$$\mathcal{W}_\alpha = \mathcal{Z}_\alpha \overline{\alpha}$$

In the general picture, we are to regard one of these twistor descriptions $(Z^\alpha, W_\alpha)$ as holding on one patch, and the other description $(\mathcal{Z}_\alpha, \mathcal{W}_\alpha)$ as holding on another which overlaps it. We are interested in the quantum twistor algebras in each system and how to match these algebras from patch to patch.

### D3. A palatial role for geometric quantization

The one piece of clear geometry that we do wish to match between the two systems is the symplectic structure given by $\Sigma$:

$$\Sigma = i \, dZ^\alpha \wedge dW_\alpha = i \, dZ^\alpha \wedge d\mathcal{W}_\alpha.$$  

but even this is geometrically meaningful only on the region $\mathbb{C}N$. Outside this region, there is an arbitrariness involved in the extensions away from $\mathbb{C}N$. Nevertheless, I shall demand that the above relation continues to be maintained outside $\mathbb{C}N$, this still allows a very considerable freedom in the extension. In the picture that I am presenting, it will be legitimate to regard the 8-complex-dimensional space of pairs $(Z^\alpha, W_\alpha)$ to be identifiable, in some local region of each, with the 8-complex-dimensional of pairs $(\mathcal{Z}_\alpha, \mathcal{W}_\alpha)$, but any such identification is not to be regarded as meaningful in itself. It is only a means to an end, namely to identify the allowable quantum twistor algebras that are to be abstracted from this framework.

Here is where the procedures of geometric quantization come into play—though, technically, it is only the more primitive procedure of geometric “pre-quantization” that will be called upon here. The idea will be that we require a bundle-connection on $\mathbb{C}N$, determined by a 1-form $\Phi$, whose curvature is the 2-form $\Sigma$ (apart from the factor of $i$ introduced here). The various possible such connections will give the different possible realizations of the quantum twistor algebra. These different possibilities come about from the different ways that the 2-form $\Sigma$ is expressed as the exterior derivative of a 2-form $\Phi$ (here with a factor of $i$):

$$\Sigma = i d\Phi.$$  

The bundle connection is then given, in the $(Z^\alpha, W_\alpha)$ system by

$$\left( \frac{\partial}{\partial Z^\alpha} + P_\alpha, \frac{\partial}{\partial W_\alpha} + Q_\alpha \right)$$

where
Here,

$$P_\alpha = -\frac{\partial E}{\partial Z^\alpha} \quad \text{and} \quad Q^\alpha = \frac{\partial F}{\partial W^\alpha},$$

where \( E + F = Z^\alpha W^\alpha \), \( E \) and \( F \) being (holomorphic) functions of \( Z^\alpha \) and \( W^\alpha \).

Using this bundle connection, we can realize the operations of the non-commutative quantum twistor algebra \( \mathbb{A} \). In the description given at the beginnings of §C1 and §C2 (with \( \hbar = 1 \)), we get the canonical case where the commutation relations can be written:

$$[Z^\alpha, Z^\beta] = 0, \quad [-\frac{\partial}{\partial Z^\alpha}, -\frac{\partial}{\partial Z^\beta}] = 0, \quad [Z^\alpha, -\frac{\partial}{\partial Z^\beta}] = \delta^\alpha_\beta$$

and we can directly realize these in our bundle connection if we take \( P_\alpha = 0, Q^\alpha = Z^\alpha \) (i.e. \( E = 0, F = Z^\alpha W^\alpha \)), so that our connection is represented by

$$\left( \frac{\partial}{\partial Z^\alpha} - W^\alpha, \frac{\partial}{\partial W^\alpha} + Z^\alpha \right).$$

Here we take our ket space to be given by holomorphic functions of \( Z^\alpha \), where \( W^\alpha \) is to be taken constant, so that the first term in the above gives us \( \partial/\partial Z^\alpha \) and the second gives us \( Z^\alpha \). The same algebra, but taken with functions of the dual twistors as the ket space would be obtained by taking \( P_\alpha = -W^\alpha, Q^\alpha = 0 \) (i.e. \( E = Z^\alpha W^\alpha, F = 0 \)), so that our connection is represented by

$$\left( \frac{\partial}{\partial Z^\alpha} - W^\alpha, \frac{\partial}{\partial W^\alpha} \right),$$

where the commutation relations can be written:

$$[\frac{\partial}{\partial W^\alpha}, \frac{\partial}{\partial W^\beta}] = 0, \quad [W^\alpha, W^\beta] = 0, \quad [\frac{\partial}{\partial W^\alpha}, W^\beta] = \delta^\alpha_\beta$$

and we can directly realize these in our bundle connection by taking \( P_\alpha = -W^\alpha, Q^\alpha = 0 \) (i.e. \( F = 0, G = 0 \)), so that our connection is represented by

$$\left( \frac{\partial}{\partial Z^\alpha} - W^\alpha, \frac{\partial}{\partial W^\alpha} \right).$$

Now, we can take our ket space to be given by holomorphic functions of dual twistors \( W^\alpha \), where we take \( Z^\alpha \) to be constant.

Clearly, by making other ways of splitting \( Z^\alpha W^\alpha \) into a sum of functions \( E \) and \( F \) we can arrive at many different realizations of the algebra \( \mathbb{A} \), at least in some local sense. Yet, the meaning of the word “local”, in this context is clearly something that also needs attention, but, as indicated in §D1, I am here leaving aside the thorny issues as are raised by topology and related matters, concentrating primarily on purely formal matters. Nevertheless, such topological issues must be dealt with at some stage, in order to address the important matters
corresponding to those treated §C6. In relation to this, we need to recall that the considerations of this section, following the initial paragraph, have been expressed entirely in the \((Z^\alpha, W_\alpha)\) system, whereas in order to express the relations in different patches, we need to relate the \((Z^\alpha, W_\alpha)\) system of one patch to the \((Z^\alpha, W_\alpha)\) system of another. As things stand, in addition to topological matters, the actual carrying out of the transformations involved would be likely to get exceedingly complicated in practical calculations and generally unilluminating. In the following section I shall show how these twistor transformations between patches can be greatly facilitated by means of generating functions.

**D4. Palatial generating functions and Einstein’s equations**

We have seen in §D3 that although the patching together of different regions is to be understood in terms of the quantum twistor algebras assigned to the various regions, the construction of these algebras can apparently be achieved in a directly geometrical way. Most particularly, the relation between one patch and another patch overlapping it can be described in terms of respective complex symplectic structures that we may assign to each patch as defined by the different \((Z^\alpha, W_\alpha)\) and \((Z^\alpha, W_\alpha)\) systems with the same symplectic 2-form \(\Sigma\). We want the \(\Sigma\) s to match from path to patch, but the 1-forms \(\Phi\) would be expected to differ, as would the bundle connection that \(\Phi\) defines. This connection provides the needed algebra \(A_j\) for the \(j\)th patch and the specific \((Z^\alpha, W_\alpha)\) or \((Z^\alpha, W_\alpha)\) system that is used to define this algebra, like a coordinate system in ordinary manifold construction, would not be taken to have significance.

There will be no loss of generality if we adopt the convention that we choose the ket-space description for that algebra to be the space of holomorphic functions in the \(Z^\alpha\) variables rather than the \(W_\alpha\) variables (and the \(Z^\alpha\) variables rather than the \(W_\alpha\) variables, etc.). In this way, we can regard the choice of \(\Phi\) in the \((Z^\alpha, W_\alpha)\) system as actually determining the quantum twistor algebra, together with its ket space. But the ket space itself would not be part of the palatial structure, nor would \(\Phi\), and certainly not the specific \((Z^\alpha, W_\alpha)\), since these structures are not determined by the algebra. All we ask is that the algebras agree from patch to patch, being the same on the overlaps.

Yet, it needs to be pointed out that, in all this, I am being deliberately a bit vague in using terms like “the same as” rather than “isomorphic to”, for the reason that I am not at all sure what the precise term should be. The global natures of the algebras might be different in situations where in some more local sense the structure of the algebras ought to be judged as “the same”. Again, this raises an issue of rigour that I am leaving aside in this article.

The most explicit way to relate one \((Z^\alpha, W_\alpha)\) system to another one \((Z^\alpha, W_\alpha)\) while preserving \(\Sigma\) is by means of a generating function, which chooses one set of half the variables from one system and the other half from the other one:

\[ G(Z^\alpha, W_\alpha). \]

The remaining variables are then provided by

\[ Z^\alpha = \frac{\partial G}{\partial W_\alpha} \quad \text{and} \quad W_\alpha = \frac{\partial G}{\partial Z^\alpha}. \]
which by direct calculation immediately gives the required
\[ \Sigma = i \, dZ^\alpha \wedge dW_\alpha = i \, dZ^\alpha \wedge dW_\alpha. \]

For the purposes of palatial twistor theory, we require this to have a total homogeneity of 2 in all its variables. In terms of Euler homogeneity operators, this can be expressed as
\[ (Z^\alpha \frac{\partial}{\partial z^\alpha} + W_\alpha \frac{\partial}{\partial w_\alpha}) G = 2 \, G \]
from which, we obtain
\[ G = \frac{1}{2} (Z^\alpha W_\alpha + Z^\alpha W_\alpha). \]

Having an explicit expression for the patching, in terms of the given generating function, and therefore of the twistor algebras and their respective ket spaces, we can envisage piecing together an entire quantum twistor algebra $\mathcal{A}$ for a complex space-time $\mathcal{M}^c$. As was envisaged in §D1. But how are we to identify the points in $\mathcal{M}^c$? From general considerations, it can be seen that the points ought to correspond to completely commutative 4-dimensional sub-algebras of $\mathcal{A}$, although this does not appear to be a sufficient characterization. Yet, in this kind of way, the notion of points being null separated would have a simple interpretation in terms of sub-algebras so that the identification of $\mathcal{M}^c$ as a complex-conformal manifold ought to be derived from the structure of the algebra. Clearly much clarification is needed, and not the least of these problems would appear to be a need for theorems analogues to those of the Kodaira and Kodaira-Spencer [38], [39] that were so important for the original non-linear graviton construction of §C6.

Even if all this works out well, there would remain at least four important issues for this programme to have importance for physics: (1) How do we incorporate the conformal scaling that will actually provide a space-time metric? (2) How do we express Einstein’s equations for this metric? (3) how do we ensure that we obtain a real-Lorentzian space-time $\mathcal{M}$, rather than just a complex one $\mathcal{M}^c$? (4) How do we generalize this to apply to the Ward construction for Yang-Mills fields [34], and hence to particle physics?

It is perhaps very remarkable that it appears to be possible to address all three of (1), (2), and (3) with a single generating function of a particular type, which I shall come to very shortly. With regard to (4), I see no reason why these palatial procedures should not apply also to the Ward construction and perhaps give some new insights into particle physics. However for this to be realistic, one needs to understand the proper way to introduce mass. Earlier proposals for this involved functions of several twistors [20], [22], [23], and it might be well worth while to re-open these discussions in the light of the above palatial ideas. It would be interesting to see whether this leads to any new insights.

To end this article I provide a method, using a curious kind of generating function, which preserves not only the 2-form $\Sigma$, but also another 2-form
\[ \Theta = dZ^\alpha \wedge Z^\beta I_{\alpha \beta} + dW_\alpha \wedge dW_\beta I^{\alpha \beta} \]
where $I_{\alpha \beta}$ and $I^{\alpha \beta}$ are the $\Lambda$-infinity twistors introduced in §C6. From general considerations of twistor theory, it can be inferred that the preservation of $\Theta$ defines for us not only a metric scaling (over and above the conformal structure that is inherent in the twistor formalism), but also the Einstein $\Lambda$-vacuum equations (Ricci scalar equals $\Lambda$), although I do not have a direct argument for this that can be presented succinctly here. Bearing this in mind, if we can find a generating function that preserves both $\Sigma$ and $\Theta$, then our palatial procedure should provide us with the general $\mathcal{M}^c$ satisfying the $\Lambda$-vacuum equations. Somewhat remarkably this can be achieved by a generating function

$$\Gamma(\lambda Z^0, W_1, Z^1, \lambda W_2; W_0, \lambda Z^1, \lambda W_3, Z^2),$$

which is homogeneous of total degree 2, where

$$\lambda = i \frac{\sqrt{\Lambda}}{\sqrt{6}}$$

and we impose the symmetry that $\Gamma$ is unchanged if

- the 1st and 2nd entries are interchanged, together with the 3rd and 4th being interchanged

or else

- the 5th and 6th entries are interchanged together with the 7th and 8th being interchanged.

This should give us a general $\mathcal{M}^c$ satisfying the $\Lambda$-vacuum equations. If, in addition, we impose the Hermiticity relation that $\Gamma$ becomes its complex conjugate under the complete reversal of the order of the first 4 entries, together with the complete reversal or the order of the last four entries, then we should obtain a real Lorentzian $\mathcal{M}$ satisfying the $\Lambda$-vacuum equations. If desired, we can use $\Gamma$ first to obtain the required relation between the $(Z^\alpha, W_\alpha)$ and $(Z^{\alpha'}, W^{\alpha'})$ systems and then reconstruct the original type of generating function $G$ used earlier, from the relation $G=\frac{1}{2}(Z^\alpha W_\alpha+Z^{\alpha'} W^{\alpha'})$ and proceed as before.

It should be re-iterated that although ideas from quantum mechanics are crucially incorporated into this construction, this is still just classical Einstein general relativity. There is no role here for the Planck length, for example. To extend these ideas for a quantum gravity theory, it would appear that the commutation rules of §C1 should be extended to include

$$Z^\alpha Z^\beta - Z^\beta Z^\alpha + W_\rho W_\sigma \epsilon^{\alpha \beta \rho \sigma} = \zeta I^{\alpha \beta},$$

$\zeta$ being some constant connected with the Planck length (and we may note that he above formula is actually symmetric with respect to interchange of $Z^\alpha$ with $W_\alpha$ because of an identity involving the Levi-Civita tensor $\epsilon^{\alpha \beta \rho \sigma}$).
Acknowledgement: I am deeply grateful for financial assistance through a personal endowment from J.P. Moussouris.

REFERENCES


