What is a Twistor?

Christopher Alexander

June 11, 2018

The setting for twistors and indeed any local physical system is *Minkowski spacetime*. Minkowski spacetime is the pair $(\mathbb{R}^4, \eta_{ab})$, consisting of the vector space $\mathbb{R}^4$ and the 4×4 diagonal matrix η_{ab} given by:

$$\eta_{ab} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

providing the basis for $\mathbb{R}^4$ is an orthonormal one. Minkowski spacetime is a simplification of a more general and complicated setting known as a *Lorentzian manifold*, although an understanding of Lorentzian manifolds will not be required for now. The $(-+++)$ signature of η_{ab} is what distinguishes the first coordinate of $\mathbb{R}^4$ from the other three, regardless of which orthonormal basis is used. Let:

$$V^a = (V^0, V^1, V^2, V^3)$$

represent a vector in $\mathbb{R}^4$ in the standard basis:

$$\text{span}\{e_0, e_1, e_2, e_3\}$$

where:

$$\begin{aligned} e_0 &= (1, 0, 0, 0) \\ e_1 &= (0, 1, 0, 0) \\ e_2 &= (0, 0, 1, 0) \\ e_3 &= (0, 0, 0, 1) \end{aligned}$$

We may think of the first coordinate V^0 as representing some notion of time and the other three coordinates as representing some notion of space. The quantity:

$$\sum_{a=0}^{3} \sum_{b=0}^{3} V^a V^b \eta_{ab} = -(V^0)^2 + (V^1)^2 + (V^2)^2 + (V^3)^2 \tag{0.0.1}$$

is known as the *spacetime interval* and is invariant under a group of transformations known as the *Lorentz transformations*. From this point onwards, if the same letter is given as a superscript and a subscript, then a summation will be assumed over the appropriate range of values, this is called the *Einstein summation convention*. For example, (0.0.1) will be written as $V^a V^b \eta_{ab}$ from now on. η_{ab} can be thought of as acting as a sort of inner product and V^a can be thought of as a *four-velocity*, that is, a vector representing the velocity of a particle in both space and time. A particle whose four-velocity is non-zero but yet has no spacial components is not moving through space but is still moving through time. A vector V^a is said to be:

1. *timelike* if $V^a V^b \eta_{ab} < 0$,
2. *spacelike* if $V^a V^b \eta_{ab} > 0$ and
3. *lightlike* or *null* if $V^a V^b \eta_{ab} = 0$.

To give this machinery a physical interpretation, recall the postulate made by Einstein that states the speed of light must remain unchanged in all frames of reference, that is, invariant under certain change of bases of $\mathbb{R}^4$. Scientists had known that the speed of light was constant, and to great precision, before Einstein began to think about such. The motivation for Einstein's postulate was the application of the absence of absolute motion, that is, that one can only determine a quantity called *velocity* when comparing the motion of oneself to another object. To Einstein, the absence of absolute motion means one could continue to expect to view oneself in a mirror regardless of how fast they are travelling in relation to another observer. For the person looking in the mirror, should they experience a delay in their reflection, they would know they were moving, without reference to any other object, a violation of the absence of absolute motion. Einstein saw that the only remedy to a constant speed of light and lack of absolute motion was his postulate. Denoting the speed of light by c, this yields the condition:

$$(V^1)^2 + (V^2)^2 + (V^3)^2 \leq c^2$$

on any four-velocity in any orthonormal basis of $\mathbb{R}^3$. However, spacetime is four-dimensional and physical systems do not change based on which frame they are viewed in; this is where η_{ab} comes in. An *event* is a point in spacetime with respect to some frame of reference, events are invariant of frame and so will have different coordinate values in different coordinate frames. Suppose however that the standard Euclidean metric on $\mathbb{R}^4$ were to be used instead of η_{ab}, that is, the distance between two events X^a and Y^a is defined to be the quantity:

$$(X^a - Y^a)(X^b - Y^b)\delta_{ab} = (X^0 - Y^0)^2 + (X^1 - Y^1)^2 + (X^2 - Y^2)^2 + (X^3 - Y^3)^2$$

invariant of orthonormal change of basis of $\mathbb{R}^4$. Then the time and space components would be completely indistinguishable, this would immediately contradict Einstein's postulate as the following example shows. Suppose in some frame of reference that event Y^a is related to event X^a by:

$$Y^a = X^a + T^a$$

where:

$$T^a = (T^0, 0, 0, 0)$$

and $T^0 > 0$. Then an observer sitting stationary in this frame at:

$$(0, X^1, X^2, X^3)$$

will experience the event X^a at a coordinate time of X^0 and then event Y^a at a coordinate time of $X^0 + T^0$. However, an appropriate rotation of $\mathbb{R}^4$ could give events X^a and Y^a the same time coordinate. Since rotations correspond to orthonormal changes of basis, it would then be possible to find a frame whereby the two events occur simultaneously; this is not in of itself a problem, the problem is that an observer in this frame would observe the sitting observer to pass between events instantaneously, since the sitting observer is present at both events. This motivates the use of η_{ab} and the Lorentz transformations that preserve the spacetime interval $V^aV^b\eta_{ab}$. If events X^a and Y^a satisfy the condition:

$$(X^a - Y^a)(X^b - Y^b)\eta_{ab} < 0$$

which is an invariant quantity, then a particle travelling at less than the speed of light may travel from one event to the next. If X^0 and Y^0 are written as $c\tilde{X}$ and $c\tilde{Y}$ respectively, then $\tilde{X} - \tilde{Y}$ would correspond to the *time* between events with respect to this coordinate frame; this is where the speed of light constant comes into the model. Timelike vectors such as $X^a - Y^a$ correspond to the four-velocities of particles with mass. In the same way, particles of light and other massless particles have four-velocities that are null. The set of Lorentz transformations form a group with matrix multiplication as the binary operation, this group is called the *Lorentz group* and is denoted by L. We have that $L = O(1,3)$, that is, the set of endomorphisms of $\mathbb{R}^n$ (or more generally some four-dimensional vector space V) preserving η_{ab}:

$$\Lambda^a{}_b \in L \iff \Lambda^a{}_b\Lambda^c{}_d\eta_{ac} = \eta_{bd}$$

Note that the determinant of a Lorentz transformation is either 1 or -1. The Lorentz group can be split up the into four components:

$$L = L_+^\uparrow \cup L_+^\downarrow \cup L_-^\uparrow \cup L_-^\downarrow$$

where $\pm$ indicates the sign of the determinant and $\uparrow$ and $\downarrow$ indicate ${\Lambda^0}_0 > 0$ and ${\Lambda^0}_0 < 0$ respectively. The component $L_+^\uparrow$ contains the identity and is referred to as the *proper orthochronous Lorentz group*. The Lorentz transformations with negative determinants will change the orientation of a vector and the Lorentz transformations with ${\Lambda^0}_0 < 0$ will change a future pointing timelike vector to a past pointing timelike vector and vice versa. For example, η is a time-reflection and:

1. $\mathrm{diag}(1, -1, 1, 1)$,
2. $\mathrm{diag}(1, 1, -1, 1)$ and
3. $\mathrm{diag}(1, 1, 1, -1)$

are examples of space-reflections. Some of the following material closely follows the arguments laid out in Chapter 4 of [1] and will require an understanding of differential geometry and sheaf cohomology. Let $V^a \in V$, where V is a four-dimensional real vector space and V^a has components:

$$V^a = (V^0, V^1, V^2, V^3)$$

In this light define:

$$\Psi(V^a) = \frac{1}{\sqrt{2}} \begin{pmatrix} V^0 + V^3 & V^1 + iV^2 \\ V^1 - iV^2 & V^0 - V^3 \end{pmatrix} = \begin{pmatrix} V^{00'} & V^{01'} \\ V^{10'} & V^{11'} \end{pmatrix} = V^{AA'}$$

where A and A' range over the values $0, 1$ and $0', 1'$ respectively. Note that this gives a bijective correspondence between elements of V and 2×2 Hermitian matrices. Furthermore, define the map:

$$V \times \mathrm{SL}(2, \mathbb{C}) \to V$$
$$V^{AA'} \mapsto {t^A}_B V^{BB'} {\bar{t}^{A'}}_{B'}$$

which corresponds to the multiplication of the matrix $\Psi(V^a)$ on the left by an element of $\mathrm{SL}(2, \mathbb{C})$ and on the right by its Hermitian conjugate. Such a map will result in another Hermitian matrix whose determinant is equal to that of $\Psi(V^a)$. This map defines a linear transformation on V^a that preserves $V^a V^b \eta_{ab}$; these maps are the familiar Lorentz transformations:

$$V^a \to {\Lambda^a}_b V^b$$

Thus a map $\mathrm{SL}(2, \mathbb{C}) \to L$ has been established to which it can be shown to possess the following properties:

1. It is a group homomorphism.
2. It maps into $L_+^\uparrow$.
3. The kernel consists of I and $-I$ where I is the 2×2 identity matrix.

Now if V^a is a null vector, that is, $V^a V^b \eta_{ab} = 0$, then $\Psi(V^a)$ has rank one and can thus be represented as the outer product of a two-dimensional complex vector α^A and its conjugate:

$$V^{AA'} = \begin{pmatrix} V^{00'} & V^{01'} \\ V^{10'} & V^{11'} \end{pmatrix} = \begin{pmatrix} \alpha^0 \bar{\alpha}^{0'} & \alpha^0 \bar{\alpha}^{1'} \\ \alpha^1 \bar{\alpha}^{0'} & \alpha^1 \bar{\alpha}^{1'} \end{pmatrix} = \alpha^A \bar{\alpha}^{A'}$$

The α^A that correspond to null vectors V^a are called *spinors* and the two-dimensional complex vector space to which they live and acted on by $\mathrm{SL}(2, \mathbb{C})$ is called *spin-space*, denoted by S. From the spin-space, the complex conjugate spin-space $S' = \overline{S}$ along with their respective topological dual spaces S^* and S'^* can be defined. The introduction of such dual spaces allow for the construction of spinor tensors of arbitrary valence:

$$\Psi^{A\ldots BC'\ldots D'}{}_{E\ldots FG'\ldots H'} \in S\otimes\cdots\otimes S\otimes S'\otimes\cdots\otimes S'\otimes S^*\otimes\cdots\otimes S^*\otimes S'^*\otimes\cdots\otimes S'^*$$

Similarly to vector fields on a manifold, a *spinor field* on a manifold $\mathcal{M}$ is defined to be a section of a suitable vector bundle $\mathcal{S}$ over $\mathcal{M}$. See [2] or recall that a *vector bundle* of rank r on a manifold $\mathcal{M}$ is the triple $(\mathcal{S}, \mathcal{M}, \pi)$ consisting of:

1. The total space $\mathcal{S}$,
2. the base space $\mathcal{M}$ and
3. a continuous surjection $\pi : \mathcal{S} \to \mathcal{M}$ that is locally trivial of rank r.

A continuous surjection $\pi : \mathcal{S} \to \mathcal{M}$ is *locally trivial* of rank r if:

1. For each $x \in \mathcal{M}$ the *fibre* $\pi^{-1}(x)$ has the structure of an r-dimensional vector space and
2. there is an open neighbourhood U of x and a fibre-preserving diffeomorphism:

   $$\phi : \pi^{-1}(U) \to U \times \mathbb{R}^r$$

   such that for every $y \in U$ the restriction:

   $$\phi|_{\pi^{-1}(y)} : \pi^{-1}(y) \to \{y\} \times \mathbb{R}^r$$

   is a vector space isomorphism.

For any two maps $\pi : \mathcal{S} \to \mathcal{M}$ and $\pi' : \mathcal{S}' \to \mathcal{M}$ the map $\phi : \mathcal{S} \to \mathcal{S}'$ is said to be *fibre-preserving* if:

$$\phi\left(\pi^{-1}(x)\right) \subset (\pi')^{-1}(x) \ \forall\ x \in \mathcal{M}$$

The collection $\{(U, \phi)\}$, with the sets U forming an open cover of $\mathcal{M}$, is called a *local trivialisation* for $\mathcal{S}$. Given a vector bundle $(\mathcal{S}, \mathcal{M}, \pi)$ of rank r and a pair (U, ϕ_U) and (V, ϕ_V) over which the bundle trivialises, the composition function:

$$\phi_U^{-1} \circ \phi_V : (U \cap V) \times \mathbb{R}^r \to (U \cap V) \times \mathbb{R}^r$$

is well defined on the overlap and satisfies:

$$\varphi_U^{-1} \circ \phi_V(x, v) = (x, g_{UV}(x)v)$$

for some GL(r)-valued function:

$$g_{UV} : U \cap V \to GL(r)$$

The functions g_{UV} are called *transition functions* of the vector bundle and correspond to coordinate transformations on the vector bundle. A *section* of a vector bundle $(\mathcal{S}, \mathcal{M}, \pi)$ is a map $s : \mathcal{M} \to \mathcal{S}$ such that $\pi \circ s$ is the identity map on $\mathcal{M}$. A *frame* for a vector bundle $(\mathcal{S}, \mathcal{M}, \pi)$ of rank r over an open set U is a collection of sections $s_1, \ldots, s_r$ of $\mathcal{S}$ over U such that at each point $x \in U$ the elements $s_1(x), \ldots, s_r(x)$ form a basis for the fibre $\pi^{-1}(x)$. This definition of a frame actually corresponds to the physical notion of a frame of reference, as discussed previously. A *frame bundle* of the vector bundle $(\mathcal{S}, \mathcal{M}, \pi)$ is a disjoint union of frames over all of $\mathcal{M}$. Now to define the spinor bundle $\mathcal{S}$, the manifold $\mathcal{M}$ must be orientable, so that a global choice of orientation is possible, and also time-orientable, so that a global choice of future-directed vectors can be made. The orientability and time-orientability of $\mathcal{M}$ means that an orthonormal frame bundle B of $\mathcal{M}$ can be reduced to an $L_+^\uparrow$ bundle, that is, the transition matrices for B can be chosen to be in $L_+^\uparrow$. Many objects involving spinors are only defined up to sign, the kernel of the map $\mathrm{SL}(2,\mathbb{C}) \to L$ being a good example, this leads after having reduced B to an $L_+^\uparrow$ bundle, to find $\mathcal{S}$ as a double cover of B; for this to be possible, $\mathcal{M}$ must satisfy a topological technicality which will be discussed shortly. Starting with orientability first, consider a locally finite open cover $\{U_i\}_{i\in I}$ of $\mathcal{M}$ and a choice of orthonormal frame f_i over U_i. On the non-empty intersections $U_i \cap U_j$ the frames f_i and f_j will be related by a Lorentz transformation:

$$f_j P_{ij} f_i$$

The Lorentz transformations P_{ij} define the orthonormal frame bundle B and must satisfy:

$$\begin{aligned} P_{ji} &= P_{ij}^{-1} \\ P_{ij}P_{ki}P_{jk} &= I \end{aligned}$$

providing:

$$U_i \cap U_j \cap U_k \neq 0$$

This is where sheaf cohomology comes in, so a few definitions are first in order. As a more general object to the sheaf of holomorphic functions $\mathcal{O}$, a *sheaf* $\mathcal{S}$ over a topological space $\mathcal{M}$ is a topological space together with a mapping $\pi : \mathcal{S} \to \mathcal{M}$ such that:

1. π is a local homeomorphism,
2. the *stalks* $\pi^{-1}(x)$ are abelian groups and
3. the group operations are continuous.

It is important to note that the sections of $\mathcal{S}$ over an open subset of $\mathcal{M}$ form an abelian group $\mathcal{S}(U)$. Let $\{U_i\}$ be an open cover of $\mathcal{M}$, then a *p-cochain* is a collection of sections:

$$s_{i_0\ldots i_p} \in \mathcal{S}\left(\bigcap_{k=1}^{p} U_{i_k}\right)$$

that is, one for each non-empty $p+1$-fold intersection of the sets U_i, which are completely skew-symmetric:

$$s_{i_0\ldots i_p} = s_{[i_0\ldots i_p]}$$

This p-cochain will be denoted as $\{s_{i_0\ldots i_p}\}$. For example, a 1-cochain is a collection of sections:

$$s_{ij} \in \mathcal{S}(U_i \cap U_j)$$

such that:

$$s_{ij} = -s_{ji}$$

The set of all p-cochains has an abelian group structure, inherited from the sheaf, and is denoted by $C^p(\{U_i\};\mathcal{S})$. The *coboundary map*:

$$\delta_p : C^p(\{U_i\};\mathcal{S}) \to C^{p+1}(\{U_i\};\mathcal{S})$$

will now be defined. Given a 0-cochain s_i, a 1-cochain s_{ij} can be defined as follows:

$$s_{ij} = \rho_j s_i - \rho_i s_j = 2\rho_{[j} s_{i]}$$

where $\rho_i s_j$ is the restriction of s_j to the set $U_i \cap U_j$. The coboundary map on $C^0(\{U_i\};\mathcal{S})$ is then defined to be:

$$\delta_0(\{s_i\}) = \{2\rho_{[j} s_{i]}\}$$

with the general coboundary map being defined analogously:

$$\delta_p(\{s_{i_0 \dots i_p}\}) = \{(p+1)\rho_{[i_{p+1}} s_{i_0 \dots i_p]}\}$$

It is easy to see that:

$$\ker(\delta_0) = \mathcal{S}$$

Moreover, there are special notations for the kernels and images of such coboundary maps, they are:

$$Z^p(\{U_i\};\mathcal{S}) = \ker(\delta_p)$$
$$B^p(\{U_i\};\mathcal{S}) = \mathrm{im}(\delta_{p-1})$$

with the elements of $Z^p(\{U_i\};\mathcal{S})$ being called p-cocycles and the elements of $B^p(\{U_i\};\mathcal{S})$ being called p-coboundaries. $Z^p(\{U_i\};\mathcal{S})$ and $B^p(\{U_i\};\mathcal{S})$ are abelian groups and since:

$$\delta_{p+1} \circ \delta_p = 0$$

then $B^p(\{U_i\};\mathcal{S})$ will be a normal subgroup of $Z^p(\{U_i\};\mathcal{S})$. The sheaf that is considered in the construction of the vector bundle $\mathcal{S}$ is $\mathbb{Z}_2$. Recalling the Lorentz transformations P_{ij}, define:

$$\tau_{ij} = \det(P_{ij})$$

Then τ_{ij} is an assignment of ± 1 to every non-empty intersection $U_i \cap U_j$ and is symmetric. Thus τ_{ij} defines a 1-cochain $\tau \in C^1(\{U_i\};\mathbb{Z}_2)$ and furthermore:

$$\tau_{ij}\tau_{ki}\tau_{jk} = 1$$

so that τ is actually a cocycle, that is, $\tau \in Z^1(\{U_i\};\mathbb{Z}_2)$. Now a change in the orientation of the f_i corresponds to a zero cochain $\omega \in C^0(\{U_i\};\mathbb{Z}_2)$ as so:

1. $\omega_i = 1$ if the orientation of f_i is unchanged,

2. $\omega_i = -1$ if the orientation of f_i is changed.

Such a change in orientation modifies the P_{ij} and so will modify the τ_{ij} according to:

$$\tau_{ij} \to \omega_i \tau_{ij} \omega_j$$

that is, τ changes by a coboundary. Using the cochain ω, the frames f_i can be modified so that all the τ_{ij} become $+1$; this is fixing a choice of orientation. A similar argument allows $\mathcal{M}$ to oriented in time such that the matrices P_{ij} are all in $L_+^\uparrow$. Now to construct the spin bundle $\mathcal{S}$, a choice of matrix $\sigma_{ij} \in SL(2,\mathbb{C})$ is needed where σ_{ij} is one of the two inverse images of P_{ij}, the other being $-\sigma_{ij}$. These can evidently be chosen to satisfy:

$$\sigma_{ji} = \sigma_{ij}^{-1}$$

but on the non-empty triple intersections:

$$\sigma_{ij}\sigma_{ki}\sigma_{jk} = z_{ijk}I$$

where I is the 2×2 unit matrix and $z_{ijk} = \pm 1$. To be able to construct $\mathcal{S}$, the σ_{ij} needs to be chosen so that all the z_{ijk} are $+1$. As like before, the z_{ijk} define a cochain but now in $C^2(\{U_i\};\mathbb{Z}_2)$, which again is actually a cocycle. A 1-cochain ω_{ij} can be defined by changing the choice of σ_{ij}, that is, $\omega_{ij} = -1$ if the choice of $-\sigma_{ij}$ is taken and $\omega_{ij} = 1$ if the choice of σ_{ij} is taken. This changes z_{ijk} by a coboundary:

$$z_{ijk} \to z_{ijk}\tau_{ij}\tau_{ki}^{-1}\tau_{jk}$$

and so the cocycle τ_{ij} can then be used to change the signs of the σ_{ij} so that they satisfy:

$$\sigma_{ij} = \sigma_{ji}^{-1}$$
$$\sigma_{ij}\sigma_{ki}\sigma_{jk} = I$$

and can be used to build the bundle $\mathcal{S}$, up to a technicality. Let γ be a path in the fibre of the frame bundle that is not homotopic to zero in that fibre. If γ is still not homotopic to zero when arbitrarily deformed in the whole frame bundle, then $\mathcal{S}$ satisfies the technicality and will be the desired spinor bundle. Local sections of $\mathcal{S}$ are the unprimed spinor fields $\pi_A(x)$. Note that, like before, the conjugate, dual and conjugate dual bundles $\mathcal{S}'$, $\mathcal{S}^*$ and $\mathcal{S}'^*$ respectively can also be constructed from $\mathcal{S}$. Now that the appropriate machinery is in place, the Levi-Civita connection ∇_a of $\mathcal{M}$ can be extended uniquely to a connection $\nabla_{AA'}$ on the spin bundles, further details on this can be found in [1]. Finally, a *twistor* is defined as a spinor field $\Omega^A(x)$ in Minkowski spacetime $\mathcal{M}$ satisfying the *twistor equation*:

$$\nabla_{A'}{}^{(A}\Omega^{B)} = 0$$

which is equivalent to the equation:

$$\nabla_{AA'}\Omega^B = -i\delta_A^{\ B}\pi_{A'}$$

for some other spinor field $\pi_{A'}$ and where $\delta_A^{\ B}$ is the identity spinor. *Twistor space* is then the four-dimensional complex vector space of solutions to the twistor equation. The range of applications of twistors is enormous, particularly in the field of mathematical physics. Indeed the whole machinery of differential geometry, the bedrock of spacetime, can be completely reformulated in terms of spinor fields. One may think that stopping at the definition of a twistor is of great injustice to such a magnificent field, and they would be right. Thus the reader is referred to [1] and [3]-[7] to continue the journey.

References:

1. *An Introduction to Twistor Theory* by S. Huggett and K. Tod.

2. *An Introduction to Manifolds, Second Edition* by Loring Tu.

3. *Twistor Theory - Lecture Notes in Pure and Applied Mathematics* by S. Huggett.

4. *Twistors in Mathematics and Physics* by T. Bailey and R. Baston.

5. *Twistor Geometry and Field Theory* by R. Ward.

6. *Spinors and Space-time, Volume 1* by R. Penrose and W. Rindler.

7. *Spinors and Space-time, Volume 1* by R. Penrose and W. Rindler.