Physical Applications of Geometric Algebra

Handout 13

Gauge Theories

The fundamental forces of nature can all be described in terms of *gauge theories*. In the early part of this century physicists noticed that electromagnetic interactions arise from demanding invariance of quantum wave equations under local changes of phase. There the position remained until the fifties, when Yang and Mills showed how to construct theories based on more complicated, non-commuting groups. This is the basis for the *standard model* of the electroweak and strong interactions. In the years since, many physicists and mathematicians have attempted to establish that *general relativity* (GR) is also a gauge theory. These attempts have met with mixed success. By the sixties it was established that GR *could* be formulated as a gauge theory, but the equations obtained always ended up looking extremely complicated. Certainly more so than those from the traditional view of gravity arising from spacetime curvature.

Geometric algebra provides a solution to this problem. Utilising the full structure of the spacetime algebra (STA), it is possible to construct gravity as a gauge theory in a formalism that is actually *easier* to understand and work with than the curved-space viewpoint. This is the subject of the final four lectures.

1 Electromagnetism as a Gauge Theory

The simplest example of a gauge theory is electromagnetism, so we start by analysing this in its STA form. Consider the free-particle Dirac equation,

$$\nabla\psi I\boldsymbol{\sigma}_3 = m\psi\gamma_0. \tag{1.1}$$

Since $I\boldsymbol{\sigma}_3$ commutes with γ_0, a *global symmetry* of this equation is the transformation

$$\psi \mapsto \psi' = \psi\,\mathrm{e}^{I\boldsymbol{\sigma}_3\theta}, \tag{1.2}$$

where θ is a constant. This is a symmetry because if Eq. (1.1) holds for ψ, it is also holds for ψ'. The symmetry is 'global' because θ has the same value everywhere in space and time. The quantity $\exp(I\boldsymbol{\sigma}_3\theta)$ is the STA version of a a phase factor. It can also be viewed as a rotor, corresponding to rotations in the $\gamma_2\gamma_1$ plane through angle 2θ. We write this rotor as R.

Now what if θ is not a constant, but depends on spacetime position x, $\theta = \theta(x)$? In this case ψ' will no longer be a solution of the equation if ψ is, since

$$\nabla\psi' = (\nabla\psi)R + (\nabla\theta)\psi RI\boldsymbol{\sigma}_3 \tag{1.3}$$

and so $\nabla\psi' \neq m\psi'\gamma_0$. Hence Eq. (1.2) is not a *local* symmetry of equation (1.1) as θ cannot be varied arbitrarily from point to point. So why do we want Eq. (1.2) to work as a local symmetry? The answer lies in the structure of the of physical statement that can be extracted from the Dirac theory. There are two main types:

1. The values of *observables*, formed from inner products between spinors,

$$\langle\psi|\phi\rangle = \langle\tilde{\psi}\phi\rangle_q = \langle\tilde{\psi}\phi\rangle - \langle\tilde{\psi}\phi I\boldsymbol{\sigma}_3\rangle I\boldsymbol{\sigma}_3. \tag{1.4}$$

2. Statements about the equality of two spinor expressions, for example

$$\psi = \psi_1 + \psi_2. \tag{1.5}$$

   This might decompose ψ into two orthogonal eigenstates of some operator.

In both cases, if all spinors pick up the same locally-varying phase factor (rotor) then the physical predictions are unchanged.

1.1 Covariant Derivatives

Now that we have understood the motivation, we must find how to modify Eq. (1.1) in order that phase changes become a local symmetry. We first rewrite ∇ as

$$\nabla = \partial_a a\cdot\nabla, \tag{1.6}$$

so as to clearly separate out its vector and derivative characteristics. (If you find the $\partial_a\, a\cdot\nabla$ construction a bit too abstract, just think of it as $\gamma^\mu\partial_\mu$.) The equation for ψ' now includes the term

$$\nabla\psi' = \partial_a\left(a\cdot\nabla\psi R + \psi a\cdot\nabla R\right). \tag{1.7}$$

We clearly need to modify the ∇ operator to be able to cancel out the term in the derivative of R. We therefore define a new, 'covariant' derivative operator D by including an extra piece in ∇,

$$D\psi = \partial_a\left(a\cdot\nabla\psi + \tfrac{1}{2}\psi\Omega(a)\right) \tag{1.8}$$

(the factor 1/2 is inserted for later convenience). Here $\Omega(a)$ is a multivector field of some kind, whose nature and transformation properties we have to determine.

The behaviour we require is that under a local rotation, D should transform in such a way that ψR is still a solution of the modified equation. So, with D transforming to D', we require that

$$D'(\psi R) = (D\psi)R \tag{1.9}$$

for any R. We expect that D' should have the same functional form as D, so we also have

$$D'\psi = \partial_a \left(a\cdot\nabla\psi + \tfrac{1}{2}\psi\Omega'(a)\right). \tag{1.10}$$

Eq. (1.9) therefore gives

$$D'(\psi R) = \partial_a \left(a\cdot\nabla\psi R + \psi a\cdot\nabla R + \tfrac{1}{2}\psi R\Omega'(a)\right) = \partial_a \left(a\cdot\nabla\psi + \tfrac{1}{2}\psi\Omega(a)\right) R. \tag{1.11}$$

From this we can read off that

$$a\cdot\nabla R + \tfrac{1}{2}R\Omega'(a) = \tfrac{1}{2}\Omega(a)R, \tag{1.12}$$

hence

$$\Omega'(a) = \tilde{R}\Omega(a)R - 2\tilde{R}a\cdot\nabla R. \tag{1.13}$$

Now R is a rotor so satisfies $\tilde{R}R = 1$, so $\tilde{R}a\cdot\nabla R$ is equal to minus its own reverse and is therefore a *bivector*. That is, $\Omega(a)$ is a bivector-valued field, which is linear function of a, and a general function of position. We sometimes write this

$$\Omega(a) = \Omega(a;x) \tag{1.14}$$

if we want to make the position-dependence manifest. Usually we can drop the x label and just carry round the a, which records that fact that $\Omega(a)$ is linear on a. The bivector field $\Omega(a)$ is what we must introduce in order to make rotation by R a local symmetry of the Dirac equation.

The key point now is that we have only used the form of the $-2\tilde{R}a\cdot\nabla R$ term in (1.13) to say what type of object $\Omega(a)$ is — we are *not* asserting that $\Omega(a)$ is equal to $-2\tilde{R}a\cdot\nabla R$. On the contrary, as will become apparent later, if $\Omega(a)$ was given by the gradient of a rotor like this it would give rise to a vanishing field strength and therefore be of no physical interest. This step, of taking a term arising from a derivative (like $-2\tilde{R}a\cdot\nabla R$ here), and generalizing it to a field *not* in general derivable from a derivative, is the essence of the gauging process.

Our new derivative D_a, with

$$D_a\psi = a\cdot\nabla\psi + \tfrac{1}{2}\psi\Omega(a), \tag{1.15}$$

is called a *covariant* derivative, and the $\Omega(a)$ term is called a *connection*. The fact that the connection is a bivector field relates it directly to the underlying symmetry group. (In general, connections take their values in the Lie algebra of the associated symmetry group.)

1.2 The Minimally Coupled Dirac Equation

Returning to electromagnetism, we are concerned with the restricted class of rotations which take place wholly in the $\gamma_2\gamma_1$ plane. In this case, writing $R = \exp(I\boldsymbol{\sigma}_3\theta)$, we have

$$-2\tilde{R}a\cdot\nabla R = -2\,\mathrm{e}^{-I\boldsymbol{\sigma}_3\theta}\,a\cdot(\nabla\theta)\,\mathrm{e}^{+I\boldsymbol{\sigma}_3\theta}\,I\boldsymbol{\sigma}_3 = -2a\cdot(\nabla\theta)I\boldsymbol{\sigma}_3. \tag{1.16}$$

So in generalizing to $\Omega(a)$, we can see that this must take the form

$$\Omega(a) = -\lambda a\cdot A\, I\boldsymbol{\sigma}_3, \tag{1.17}$$

where A is a general 4-d vector, and λ is some coupling constant. If A was in fact the gradient of a scalar, then we would expect the field strength to vanish. Having reached this point we are back on familiar ground of course, since this is just the statement that $\nabla\wedge A$ vanishes if $A = \nabla\phi$.

We are now in a position to reassemble our full, covariant Dirac equation. We have

$$D\psi = \partial_a(a\cdot\nabla\psi + \tfrac{1}{2}\lambda a\cdot A\, I\boldsymbol{\sigma}_3) = \nabla\psi + \tfrac{1}{2}\lambda A\psi I\boldsymbol{\sigma}_3, \tag{1.18}$$

where we see that that connection reassembles with the ∂_a term to give a vector A multiplying ψ from the left. The Hamiltonian from this operator contains a new term in $\lambda\gamma_0 A/2$, and the scalar part of this is $\lambda V/2$. It is clear that for an electron we require $\lambda = 2e$, so the '*minimally coupled*' Dirac equation is

$$\nabla\psi I\boldsymbol{\sigma}_3 - eA\psi = m\psi\gamma_0. \tag{1.19}$$

The equation is *minimally coupled* because by adding an interaction term solely in A we are making the simplest possible modification to the original equation. We could, for example, add further terms in F, or F^2 multiplying ψ, and the equation would still be gauge invariant. It appears, however, that nature does not employ this possibility. Why this should be so is far from clear.

2 Gauge Principles for Gravitation

Having successfully derived electromagnetism, we now turn our attention to gravity. We first need to be clear about our aim. This is to model gravitational interactions in terms of (gauge) fields defined in the STA. Already, this is a radical departure from GR. The STA is the geometric algebra of *flat* spacetime, and the introduction of fields cannot alter this basic property. What then are we to make of the standard arguments that spacetime is curved? The answer is that all of these arguments involve light paths, or measuring rods, or such like, and all of these processes are also modeled by fields defined in the STA. Since all physical quantities correspond to fields, the *absolute* position and orientation of particles or fields in the STA is not measurable. The only predictions that can be extracted are relative relations between fields. Ensuring that this property is true locally means there is no conflict with any of the principles by which one is traditionally led to GR.

The preceding considerations become clearer if we consider relations between quantum fields. Suppose that $\psi_1(x)$ and $\psi_2(x)$ are spinor fields. A physical statement could be a simple relation of equality,

$$\psi_1(x) = \psi_2(x). \tag{2.1}$$

But all this statement says is that at a point where one field has a particular value, then the second field has the same value. This statement is completely independent of where we choose to place the fields in the STA. And, more importantly, it is totally independent of where we choose to locate other values of the fields. We could equally well introduce two new fields

$$\psi_1'(x) = \psi_1(x'), \quad \psi_2'(x) = \psi_2(x'), \tag{2.2}$$

where x' is an arbitrary function of position x. The statement $\psi_1'(x) = \psi_2'(x)$ contains precisely the same physical content at the original equation.

The same picture emerges if both fields are acted on by a spacetime rotor, giving rise to new fields

$$\psi_1' = R\psi_1, \quad \psi_2' = R\psi_2. \tag{2.3}$$

Again, the statement $\psi_1' = \psi_2'$ has the same physical content as the original equation. Similar considerations apply to the observables formed from ψ, such as the vector $J = \psi\gamma_0\tilde{\psi}$. Replacing ψ by ψ' produces the new vector $J' = RJ\tilde{R}$. Invariance of the equations under this transformation ensures that the absolute direction of vectors in the STA is not measurable, only the relative orientation of two physical vectors is measurable. We now have a clear mathematical statement of the invariance properties we want to establish. The next task is to study the form of the gauge fields needed to enforce this invariance.

2.1 Displacements

We write $x' = f(x)$ for an arbitrary (differentiable) map between spacetime position vectors. The transformation we are interested in is where the field $\psi(x)$ is moved around to the new field

$$\psi'(x) \equiv \psi(x'). \tag{2.4}$$

The map $f(x)$ should not be thought of as a map between manifolds, or as 'moving points around'. The function $f(x)$ is just a rule for relating one position vector to another within a single vector space. It is then the fields that are transformed in this space. We need a good name for this operation of moving fields around. One possibility is '*translation*', but this suggests a rigid map where all fields are translated by the same amount. Mathematicians favour the term '*diffeomorphism*', but this is a bit unwieldy and any has some unwanted technical connotations. We prefer to use the term '*displacement*', which does correctly suggest the idea of moving the field around from one point to another.

As with electromagnetism, we now need to consider the behaviour of the derivative of ψ, $\nabla\psi = \partial_a\, a\cdot\nabla\psi$. If we form the derivative of the displaced field we find that

$$\begin{aligned} a\cdot\nabla\psi'(x) = a\cdot\nabla\psi[f(x)] &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(\psi[f(x+\epsilon a)] - \psi[f(x)]\right) \\ &= \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(\psi[f(x)+\epsilon\mathsf{f}(a)] - \psi[f(x)]\right). \end{aligned} \tag{2.5}$$

where

$$\mathsf{f}(a) = \mathsf{f}(a;x) = a\cdot\nabla f(x) \tag{2.6}$$

and we have Taylor expanded $f(x+\epsilon a)$ to first order. The function $\mathsf{f}(a)$ is linear on a, and also position dependent. We usually suppress this position dependence. We now have

$$a\cdot\nabla\psi'(x) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left(\psi[x'+\epsilon\mathsf{f}(a)] - \psi(x')\right). \tag{2.7}$$

But this is the vector derivative with respect to x' taken in $\mathsf{f}(a)$ direction. We therefore have

$$a\cdot\nabla\psi'(x) = \mathsf{f}(a)\cdot\nabla_{x'}\psi(x'), \tag{2.8}$$

where the subscript x' on $\nabla_{x'}$ records that the derivative is now with respect to the new vector position variable x'. Since

$$\mathsf{f}(a)\cdot\nabla_{x'} = a\cdot\bar{\mathsf{f}}(\nabla_{x'}) \tag{2.9}$$

and Eq. (2.8) is true for all a and for all fields ψ, we establish the operator relation

$$\nabla_x = \bar{\mathsf{f}}(\nabla_{x'}). \tag{2.10}$$

The function $\mathsf{f}(a)$ is a coordinate-free way of writing the Jacobian.

Now suppose, for example, that we had a physical relation equating the gradient of a scalar field $\nabla\phi$ to a vector field $A(x)$,

$$\nabla\phi = A. \tag{2.11}$$

This might correspond to the statement that an A-field is pure gauge in electromagnetism. If we now replace $\phi(x)$ by $\phi'(x) = \phi(x')$ and $A(x)$ by $A'(x) = A(x')$, the left-hand side becomes

$$\nabla\phi'(x) = \bar{\mathsf{f}}(\nabla_{x'})\phi(x') = \bar{\mathsf{f}}[A(x')] = \bar{\mathsf{f}}(A') \tag{2.12}$$

which is no longer equal to A'. It is clear that we must introduce a gauge field which assembles with the vector derivative to form an object which, under displacements, simply re-evaluates to the derivative with respect to the new position vector. We construct such an object by replacing ∇ with a new derivative $\bar{\mathsf{h}}(\nabla)$. Here

$$\bar{\mathsf{h}}(a) = \bar{\mathsf{h}}(a;x) \tag{2.13}$$

is an *arbitrary* function of position, and is a linear function of a. We again suppress this position dependence where clarity permits. We allow $\bar{\mathsf{h}}(a)$ to have arbitrary position dependence so that $\bar{\mathsf{h}}(a)$ cannot simply be gauged away.

Under displacements the gauge field $\bar{\mathsf{h}}(a)$ must transform such that

$$\bar{\mathsf{h}}'(\nabla\phi') = \bar{\mathsf{h}}'[\bar{\mathsf{f}}(A')] = \bar{\mathsf{h}}[A';x']. \tag{2.14}$$

Suppressing the position dependence, we can see that the basic requirement is that

$$\bar{\mathsf{h}}'(a) = \bar{\mathsf{h}}\bar{\mathsf{f}}^{-1}(a), \tag{2.15}$$

which must hold for an arbitrary vector a. We can now systematically replace every occurrence of ∇ with $\bar{\mathsf{h}}(\nabla)$, and all our equations will be invariant under arbitrary displacements. So, for example, the Dirac equation is now

$$\bar{\mathsf{h}}(\nabla)\psi \boldsymbol{I\sigma}_3 = m\psi\gamma_0. \tag{2.16}$$

The introduction of the $\bar{\mathsf{h}}$-field ensures that derivatives of fields can also be moved around arbitrarily. The $\bar{\mathsf{h}}$-field is not a connection in the conventional Yang-Mills sense. The coupling to derivatives is different, as is the transformation law (2.14). It is clear however, that the $\bar{\mathsf{h}}$-field embodies the idea of ensuring that a symmetry is local, so can sensibly be called a gauge field. Since $\bar{\mathsf{h}}(a)$ is an arbitrary, position-dependent linear function of a, it has $4 \times 4 = 16$ degrees of freedom.

2.2 Rotations

The second symmetry we require is that our wave equation should be invariant under the transformation

$$\psi \mapsto \psi' = R\psi, \tag{2.17}$$

where R is an arbitrary, position-dependent rotor in spacetime. We refer to the rotor R as generating rotations, understanding that boosts are now a special case of a rotation. Now we are back in the familiar territory of Section 1. We first write

$$\bar{\mathsf{h}}(\nabla)\psi = \bar{\mathsf{h}}(\partial_a)\, a\cdot\nabla\psi. \tag{2.18}$$

To make (2.17) a symmetry we need to modify the $a{\cdot}\nabla$ directional derivatives by adding a bivector connection $\Omega(a)$. We define

$$D_a\psi = a\cdot\nabla + \tfrac{1}{2}\Omega(a)\psi \tag{2.19}$$

where $\Omega(a)$ has the transformation law

$$\Omega(a) \mapsto \Omega'(a) = R\Omega(a)\tilde{R} - 2a\cdot\nabla R\tilde{R}. \tag{2.20}$$

Since R is an arbitrary rotor there is now no constraint on the blades that $\Omega(a)$ can contain, so $\Omega(a)$ has $6\times 4 = 24$ degrees of freedom.

Our equation now reads

$$D\psi I\boldsymbol{\sigma}_3 = \bar{\mathsf{h}}(\partial_a)D_a\psi I\boldsymbol{\sigma}_3 = m\psi\gamma_0. \tag{2.21}$$

If we now replace ψ by ψ' and $\Omega(a)$ by $\Omega'(a)$, we find that the left-hand side becomes

$$\bar{\mathsf{h}}(\partial_a)D'_a(R\psi)I\boldsymbol{\sigma}_3 = \bar{\mathsf{h}}(\partial_a)RD_a\psi I\boldsymbol{\sigma}_3 \tag{2.22}$$

whereas the right-hand side is simply $mR\psi\gamma_0$. In order for the equation to remain invariant we also need to transform the $\bar{\mathsf{h}}$-field as

$$\bar{\mathsf{h}}(a) \mapsto \bar{\mathsf{h}}'(a) = R\bar{\mathsf{h}}(a)\tilde{R}. \tag{2.23}$$

This is sensible if we recall that the equation $\bar{\mathsf{h}}(\nabla\phi) = A$ was invariant under displacements. This will also be invariant if both vectors are rotated, and the rotation of the $\bar{\mathsf{h}}(\nabla\phi)$ term must be driven by transforming $\bar{\mathsf{h}}$. With these considerations, we now see that Eq. (2.21) is invariant under both rotations and displacements. This has been achieved at the cost of introducing two new gauge fields, the $\bar{\mathsf{h}}(a)$ field for displacements and the $\Omega(a)$ field for rotations. In the next Lecture we will see what equations these new fields satisfy.

2.3 Covariant Derivatives for Observables

Having seen what the covariant derivative of a spinor looks like, it is a simple matter to establish a formula for the derivative of the observables formed from a spinor. In general, these observables have the form

$$A = \psi\Gamma\tilde{\psi}, \tag{2.24}$$

where Γ is a constant multivector formed from combinations of γ_0, γ_3 and $I\boldsymbol{\sigma}_3$. The observable A inherits its transformation properties from the spinor ψ, so under displacements it transforms as

$$A(x) \mapsto A'(x) = A(x') \tag{2.25}$$

and under rotations it transforms as

$$A \mapsto A' = RA\tilde{R}. \tag{2.26}$$

Multivectors with these transformation properties are said to be *covariant*.

If we now form the directional derivative of A we get

$$a\cdot\nabla A = (a\cdot\nabla\psi)\Gamma\tilde{\psi} + \psi\Gamma(a\cdot\nabla\psi)\tilde{}. \tag{2.27}$$

This immediately tells us how to construct a covariant derivative for A. We simply replace spinor directional derivatives with their covariant version and form

$$\begin{aligned}(D_a\psi)\Gamma\tilde{\psi} &+ \psi\Gamma(D_a\psi)\tilde{} \\ &= (a\cdot\nabla\psi)\Gamma\tilde{\psi} + \psi\Gamma(a\cdot\nabla\psi)\tilde{} + \tfrac{1}{2}\Omega(a)\psi\Gamma\tilde{\psi} - \tfrac{1}{2}\psi\Gamma\tilde{\psi}\Omega(a) \\ &= a\cdot\nabla(\psi\Gamma\tilde{\psi}) + \Omega(a)\times(\psi\Gamma\tilde{\psi}).\end{aligned} \tag{2.28}$$

We therefore define the covariant derivative $\mathcal{D}_a$ by

$$\mathcal{D}_a A \equiv a\cdot\nabla A + \Omega(a)\times A. \tag{2.29}$$

This is the form appropriate for acting on covariant multivectors, including observables formed from spinors. There are two important features about the bivector commutator appearing here. This first is that it is grade preserving, so the full $\mathcal{D}_a$ operator preserves grade. The second is that

$$\Omega(a)\times(AB) = (\Omega(a)\times A)B + A(\Omega(a)\times B). \tag{2.30}$$

This ensures that $\mathcal{D}_a$ is a *derivation*, that is, it satisfies Leibniz' rule

$$\mathcal{D}_a(AB) = (\mathcal{D}_a A)B + A(\mathcal{D}_a B). \tag{2.31}$$

Both of these are necessary for $\mathcal{D}_a$ to be a suitable generalisation of a directional derivative. We assemble a full, covariant version of the vector derivative by writing

$$\mathcal{D} = \bar{\mathsf{h}}(\partial_a)\mathcal{D}_a. \tag{2.32}$$

This raises and lowers grade by one, so we also write

$$\mathcal{D}A = \mathcal{D}\cdot A + \mathcal{D}\wedge A = \bar{\mathsf{h}}(\partial_a)\cdot(\mathcal{D}_a A) + \bar{\mathsf{h}}(\partial_a)\wedge(\mathcal{D}_a A). \tag{2.33}$$