RTG Seminar
October 16, 2006
Martin Rocek

Gabriel C. Drummond-Cole

October 16, 2006

I guess I'm supposed to talk about principal gauge invariants, connections, things like that, classical field theory. Should I wait a minute or should I start?

So, um, when we considered mechanics, in the first set of lectures, the things that were discussed in all the previous lectures, we had coordinates and phase space and Lagrangians and Hamiltonians and all sorts of things. The physical interpretation of coordinates were the positions of particles. Sometimes it's better to think about the density of distributions, in which case we talk about fields.

There are many examples that one could write down, but I'll restrict myself to the most useful in pushing us in the right direction. A scalar field φ can be thought of as a map φ from space-time to the reals $\mathbb{R}$. So space-time is something like $\mathbb{R}^{3,1}$, it could be more complicated. So we write $S = \int d^4x \varphi(x)(\Box - m^2)\varphi(x)$, where $\Box$ is $-\partial_\mu\partial^\mu = -\partial_t^2 + \partial_i\partial^i$. So φ is the field and $S[\varphi]$ is proportional to this integral.

I want this integral to make sense, but I'm going to extremize it so it's a formal equation. If I want to quantize and think about path integrals, it's important to think about these things, but to get the field equation that's not important.

So I get the equation $(\Box + m^2)\varphi = 0$. This looks like a harmonic oscillator where the mass is the spring term. This is why the oscillator is so important. This is a free massive scalar field but it looks like an oscillator.

This action is chosen because its extrema give the field equation.

[When you quantize this gives you amplitudes.]

That explains why classical physics connects to quantum physics. The classical solutions give you the first approximation to the saddle points of the quantum theory.

In field theory there are well-known places where you can use the action even classically.

Even there it can be important to consider the sum over paths, for example it's one way of understanding optics.

Okay, so this is a free scalar field.

[If you know this equation, you don't know which one is the minimum, do you? In Donaldson theory, Yang Mills solutions correspond to minimums.]

For simple things, like this with boundary conditions, it becomes a well-posed problem to find solutions to this.

[Normally when you have a variational principle, you come up with some sort of natural boundary principles.]

In $\mathbb{R}^{3,1}$ we usually assume we have something like compact support. You have to be careful if you have wave solutions if you do that, and you're right, there's some degree of sloppiness there.

Any other questions about simple scalar field theory?

We can generalize this to a manifold in the domain, and that's not what I want to do right now. Instead I want to move from a real scalar field to a complex scalar field. I should have mentioned that I could include a potential $U(\phi)$. Then the field equation has the term $\frac{dU}{d\varphi}$.

I can put the mass into U as well. So then if $U = \frac{1}{2}m^2\varphi + \frac{1}{2}\varphi^2 + \frac{g}{5!}\varphi^3 + \frac{\lambda}{4!}\varphi^4$, this is called a $(5, 4)$ theory.

So now we go to a complex scalar field, and we'll have a Hermitian, so the action will be

$$\int d^4x[\bar{\varphi}(x)(\Box - m^2)\varphi(x) - U_I(\varphi\bar{\varphi})]$$

The important thing about the complex scalar field is that this thing has a symmetry. If I replace φ with a phase $e^i\alpha$ times φ, then $\bar{\varphi} \mapsto \bar{\varphi}e^{-i\alpha}$. It's strange that this is constant. One principle which is important in physics, solutions to the wave equation have a locality property. Taking a little slice, these only affect things in the light cone. It's strange that this symmetry acts on the whole space at the same time.

So this is strange, unnatural. Physicists call this global, but since this has a mathematical meaning, they've been calling it rigid.

So how can we make this local so that α can depend on the position?

[Could it mean that this has no physical meaning? If you make this global change of phase, that just doesn't have physical sense?]

Precisely for this reason, this is not the way you want it. I can sit here and set my scalar phase, and over here I can set up my apparatus and I have a different phase. But the fact that theree's a phase difference won't go away. So since physics is local, the only way that can make sense is if I can check the phase.

One example is lepton number. In principle we can measure the number of electrons plus the number of neutrinos. That number is really there.

This is a natural question which leads to the next stage. If we gauge this symmetry, make it local, make a new theory where it becomes local. So you make $\alpha \mapsto \alpha(x)$.

[See Blaine's latest paper.]

So once you have a local symmetry, it means the two configurations are locally indistinguishable.

[Is there a reason you wrote $e^{i\alpha}$ on the left and and on the right in the different parts?]

Yes, these will be matrices.

So the problem is that before we change anything, $\partial_\mu \varphi(x) \mapsto e^{i\alpha}\partial_\mu \varphi(x)$. If we could preserve this property, we would have this local symmetry. If I look at

$$\partial_\mu(e^{i\alpha}\varphi) = \underbrace{e^{i\alpha x}\partial_\mu \varphi(x)}_{\text{good}} + \underbrace{i(\partial_\mu \alpha(x))e^{i\alpha(x)}\varphi(x)}_{\text{bad}}.$$

Well, we take $\partial_\mu \varphi \to D_\mu \varphi(x) = \partial_\mu \varphi(x) - iA_\mu \varphi(x) \mapsto -iA_m u e^{i\alpha}\varphi - i(\partial_\mu(\alpha))e^{i\alpha}\varphi$.

So I'm letting $\varphi \mapsto e^{i\alpha}\varphi$ and $A_\mu \mapsto A_\mu \mapsto A_\mu + \partial_\mu \alpha$.

This could be thought of in terms of bundles and connections. So A is a U_1 connection.

Physicists allow singularities sometimes and other times just look at bundles with transition functions.

[What is A_μ?]

A_μ is a real valued one-form. It's the coefficients of a real valued one form. If I choose φ and A, and transform them in this way, that is a completely equivalent system. Now you need to put $\Box_A$. So this is $S[\varphi, A]$, or rather the scalar part of it. If we extremize this with respect to A, we get a statement, a constraint on the scalar fields. Well, actually, in this case since it's second order in the A we get a little, but basically we get a constraint on the scalar fields. But we want to put in some dynamics, like the curvature squared term. We want to add something to get something that looks like a wave equation (for A) to make it dynamical. We're often interested in when A is not dynamical.

[What does dynamical mean? I've been asking this for years.]

There should be a second order differential operator acting on the fields. It should be wavelike. In practice dynamical means a second order something for bosons, a first order something for fermions.

If we write this out, what do we have? We have two terms, $(\partial_\mu \bar{\varphi} + A_\mu \bar{\varphi})(\partial_\mu \varphi - iA_\mu \varphi)$. These are both third order or higher.

So I can write $\bar{\varphi}\vec{D}\varphi = 0$, which is $\bar{\varphi}D_\mu \varphi - (D_\mu \bar{\varphi})\varphi = 0$, where $D_\mu = (\partial_\mu - A_\mu)\varphi$. We want to get rid of the bad part and only have a second order part.

All I'm trying to say is that

$$D_\mu(e^{i\alpha(x)}\varphi(x)) - i(A_\mu + (\partial_\mu\alpha(x)))e^{i\alpha(x)}\varphi(x) = e^{i\alpha(x)}(\partial_\mu\varphi(x) - iA_\mu\varphi(x)).$$

So A is another type of field, one that transforms in this way. There is a deep issue here which is, physicists don't start thinking we have a bundle, we start with, say, Einstein's equations, and solve them which give us a metric and a geometry. We don't start with a geometry and then try to put a metric down on it. The topology and the local geometry are both dynamical, one can say.

[I think it's a reasonable question, why do you know these exist?]

When you make a gauge change, you can write that down, and if you write a second transformation, you have to check that the composition is appropriate.

What Yang and Mills did was to change $\mathbb{C}$ to an arbitrary vector space, well, I should talk about the curvature.

So then you ask, the curvature is not exactly unique, but sort of. So $D_\mu\varphi$ transforms covariantly, so φ is a representation and so is $D_\mu\varphi$. This doesn't mean covariant in the mathematical language. It means there are no inhomogeneous terms. This doesn't mean it's invariant, independent of the choice of coordinates. So covariant means, in the mathematical terminology, either covariant or contravariant. You can even put a q, a charge, into the phase change. Here I've chosen charge one.

So you can ask are there other covariant things you can construct. Physicists did this by hand, whereas this is obvious to mathematicians. So it's the commutator. So

$$[D_\mu, D_\nu]\varphi = (D_\mu D_\nu - D_\nu D_\mu)\varphi = -i \underbrace{(\partial_\mu A_\nu - \partial_\nu A_\mu)}_{\text{Field strength tensor } F_{\mu\nu}} \varphi.$$

I just want to be sure about this arrow, when you do the arrow, the gauge transformation changes the φ and the A at the same time. The arrows always go together. The φ, the $\bar{\varphi}$, and the A all transform, it's *the* gauge transformation acting on all of them simultaneously. So you might say g is in a group and it acts as $\begin{cases} \varphi \to g\varphi \\ A_\mu \to^i A_\mu + g\partial_\mu g \end{cases}$

This describes electromagnetism. At the time that Yang Mills, they came up with a beautiful mathematical structure that didn't apply to the thing they were thinking about. It was noticed that there was an approximate symmetry, that if you ignored the charges of protons and neutrons, there was a rotation that moved them into each other. These acted like $\mathbb{C}^2$, since there are anti-protons and anti-neutrons, and this was an $SU(2)$ action, a generalized gauge transform.

In fact, for the dynamical part, once we have the $F_{\mu\nu}$ we add in $\frac{-1}{4e^2}F_{\mu\nu}(A)F^{\mu\nu}(A)$. If I want to build something depending on A, quadratic, gauge invariant, that's the leading term, the only one that I could write down.

I could write down something of higher order, and those are interesting. I could write down $\sqrt{\det \eta_{\mu\nu} + F_{\mu}\nu}$. You can see to leading order the leading piece is the quadratic piece. Then you get higher order terms. This is the free part, the leading part, there's nothing eles you can write down.

The free part means quadratic in fields, so linear in fields at the level of the field equation.

So the Yang Mills story, we run everything through in the same way, the only thing that changes, okay, let me back up. A_μ was a collection of real numbers so we stuck i in. So we could think of iA_μ as an antiHermitian. So thinking in that language of the connection as antiHermitian, we think $D_\mu\varphi = \partial_\mu\varphi + A_\mu\varphi$ in the $SU(2)$ case we would get $\varphi \to g\varphi$ and $A_\mu \to g(\partial_\mu + A_\mu)g^{-1} = g\partial_\mu g^{-1} + gA_\mu g^{-1}$, so g takes $D_\mu\varphi$ into $gD_\mu\varphi$.

In the $SU(2)$ case it's not invariant, but covariant. It transforms in the same way in the adjoint representation. Yang Mills used precisely this machinery, although not exactly this language. Then the curvature term can be added.

The $SU(2)$ associated with rotating protons into neutrons is not a gauge $SU(2)$, but there is an $SU(2)$ in nature, the weak interaction, which is spontaneously broken and a $SU(3)$ action which is the strong action.

A gauge theory would mean that things would be physically the same, and protons and neutrons are not the same, so it didn't apply to what they were talking about. They were trying to understand, they knew these issues. They created a theory with too precise a symmetry so that they could break it to get a real model.

Spontaneous breaking generates terms without making the mistake of things that fail to transform.

I'd like to back up and discuss symmetries in a broader context. This pencil is symmetric with respect to rotation around its axis, and so is the world, approximately. The physical laws have a symmetry, but the lowest energy state does not have that, the pencil falls, but the direction doesn't matter. A perfectly symmetric equation is not stable.

In physics there's a notion of a lowest energy solution, the vacuum. If the lowest energy solution doesn't have a symmetry, we say it's spontaneously broken. So the weak interaction has spontaneously broken symmetry. So you have $SU(2)$ symmetry, but the vacuum state does not have that symmetry. The Higgs mechanism and all that is a fairly long story.

[Can you say some of this good stuff the next time, because I have to leave?]

It's a fairly long story, especially if I make sense of it. The calculations are easy to do, but to give it a mathematical interpretation will be hard.

I think of a connection on a bundle, and an associated vector bundle on which the group acts. I get an expectation value on the associated bundle and the gauge field becomes massive. Describing this mathematically is pretty sophisticated.

[So starting with electromagnetism can you get mass?]

Choose an imaginary mass, the potential, well, first of all the potential is two dimensional. I'm just drawing the complex φ. So I am drawing $U(\varphi)$. The radial direction is the magnitude and the angular direction is the phase of φ. If I think $U = -m^2\varphi\bar{\varphi} + \frac{\lambda}{4}(\varphi\bar{\varphi})^2$. So there's a rotational symmetry at 0, but this is not the lowest energy state. So the expectation value $\langle\varphi\bar{\varphi}\rangle \neq 0$.

This produces the mass term for A_μ, and this breaks symmetry, and this is called the Higgs mechanism. The known spectrum of things we see, we have a massless photon theory, with very good evidence, light from very distant explosions comes at the same time independent of wavelength, very sensitive experiments, and then we have three vector fields, vector bosons with the weak interaction, and those have measured masses attributed to measured symmetry breaking.

[So when you write the potential, you want the radius to be what you calculated in the laboratory.]

The standard model has like twenty terms of that order. The self-consistency is well-tested.

So m would be the mass if m^2 were positive. But the actual mass will be m^2 down at the vacuum. The action puts imaginary mass in the formula, and you apply this argument, the symmetry breaking, to come up with real mass.

[Is that just an ad hoc mechanism to get the real mass or is it a mass sitting on the imaginary axis?]

I don't think we talk about the phase of the mass. We might also hope to see things for $SU(3)$ for the strong force. We have the third magical thing, called [unintelligible]. This idea we had of talking about dynamics by talking about the wave operator is just wrong. We need to talk about the full nonlinear picture which is nothing like the classical picture. In general that would be completely intractable, but two tools have shown us we can do things about this. The first is asymptotic freedom, which says at short distances things are weakly coupled and you can look at perturbative changes of the free model. The other thing that has come out of string theory is the ADS-CFT correspondence which has given us a weakly coupled gravitational dual picture. That came out of left field and was not expected.

Classical gravity, in an appropriate way, can tell us about strongly coupled quantum Yang Mills theory. Classical nonlinear gravity tells us about strongly coupled gauge theories, like the strong Yang Mills theory. This works only supersymmetrically, and only for $SU(n)$ as n becomes large, but it turns out that three is pretty large.

There is a third way I should have said that people tried for years, a very complicated way of studying this, which is putting it on a computer and discretizing, lattice gauge theory, which has told us some things.

The issues often are, what are the justifications for using it, you need some argument why the things you're doing should make sense. Even if string theory doesn't give us anything, it's given us huge insight into strongly coupled systems that we didn't have before. The details of applying it, we've been able to calculate quantities that we had no idea how to calculate before. It's giving an algorithm for things that you can test and experiment. It's working better than we have any hope for. Well, sometimes we can explain it with supersymmetry.

Should we have a break?

So Blaine suggested I work out the U_1 case for electromagnetism. I have

$$S_\varphi[\varphi, A] = \int d^4x[\bar{\varphi}(x)(\Box_A - m^2)\varphi(x) - U_I(\bar{\varphi}\varphi) - \frac{1}{4e^2}F_{\mu\nu}(A)F^{\mu\nu}(A)]$$

So U_I is something like $\frac{\lambda}{4}(\bar{\varphi}\varphi)^2$. where $\varphi, \bar{\varphi}$, and A_μ are the fields which transform as $\varphi \to e^{i\alpha\varphi}$, $A_\mu \to A_\mu + \partial_\mu\alpha$, and $D_\mu\varphi = \partial_\mu\varphi - iA_\mu\varphi$.

I'll treat φ and $\bar{\varphi}$ as independent things. Believe me that I can integrate by parts, but I don't have to. So taking $\frac{\partial S}{\partial\varphi} = 0$ gives us

$$\underbrace{(\partial^\mu - iA^\mu)(\partial_\mu - iA_\mu)}_{\Box_A}\varphi - m^2\varphi - U_I'\varphi = 0.$$

So $U_I'\varphi$ will be like $\frac{\lambda}{2}\bar{\varphi}\varphi^2$ I get this by considering $S[\bar{\varphi} + \delta\bar{\varphi}, A] - S[\bar{\varphi}, A]$, and I write this as

$$\int \bar{\varphi} \underbrace{(\quad)}_{\text{something}} + O(\delta\bar{\varphi}^2).$$

Then you take the complex conjugate of this equation to take the variation with respect to φ.

So the next thing that we want is the equation for A. Note that these equations are not invariant but they are covariant.

Let me remind you that $F_{\mu\nu}$ is $\partial_\mu A_\nu - \partial_\nu A_\mu$. So then summing will get rid of the one fourth. So I'll get $\frac{\partial S}{\partial A_\nu} = 0$ which gives us

$$\frac{1}{e^2}\partial^\mu F_{\mu\nu}$$

plus the terms from $\Box$, which we'll work out carefully, integrating by parts.

So I get $-(\partial^\mu + iA^\mu)\bar{\varphi}(\partial_\mu - iA_\mu)\varphi$. So varying with respect to A_μ, I get in this equation

$$-i\bar{\varphi}D_\mu\varphi + i(D_\mu\bar{\varphi})\varphi.$$

So I get

$$\frac{1}{e^2}\partial^\mu F_{\mu\nu} - i\bar{\varphi}D_\nu\varphi + i(D_\nu\bar{\varphi})\varphi = 0.$$

Let's interpret this. I don't know how you remember Maxwell's equations. Usually there are four of them, but we'll see where they come from. We have two equations. The whole system is very nonlinear. So then I'll get Maxwell's equations with φ as a current. So $i(\bar{\varphi}\dot{\phi} - \dot{\bar{\varphi}}\varphi)$ will describe, well, ϕ will be the amplitude of the charge density.

So
$$-i\bar{\varphi}D_\nu\varphi + i(D_\nu\bar{\varphi})\varphi = 2i\rho^2(\dot{\theta} - A_0)$$
which corresponds to the charge density ρ_Q, and then a spacial current $\vec{j} = 2i\rho^2(\vec{D}\theta - A_i)$.

There in an equation that $F_{\mu\nu}$ obeys identically, the Bianchi identities, which are $\epsilon^{\mu\nu\rho\sigma}\partial_\mu F_\nu\rho = 0$. So $F_{0i} = E_i$, for $i \in \{1,2,3\}$. This is the time component. The other parts $F_{ij} = \epsilon ijkB_k$, the magnetic field. What do the Bianchi identities imply on this? So I get first the divergence of B vanishes,
$$\vec{\nabla}\cdot\vec{B} = 0.$$

What about the time derivative? I can get either the curl of the E field or the time derivative of the B field. So I get
$$\dot{\vec{B}} - \vec{\nabla}\times\vec{E} = 0.$$

These are the ones, in more advanced treatments, these are trivial. They are kinematical, since F is the curl of A. The others come from the field principle. So let's see how they work. If I take $\mu = 0$ I get
$$-\frac{1}{e^2}\dot{\vec{E}} + \vec{\nabla}\times B = j = 2i\rho^2(\vec{\nabla}\theta - \vec{A}) = i(\bar{\varphi}\vec{\nabla}_A\varphi - \vec{\nabla}\bar{\varphi}\varphi).$$

The next equation I get
$$\frac{1}{e^2}\vec{\nabla}\cdot E = \rho_a.$$

Usually we rescale by e. Now $\vec{j}$ is the flow of charge.

[The formula is incomprehensible.]

This is a complicated coupled theory.

[You're adding an interpretation to the formula.]

We know this has a particular form. We know with the regular Maxwell's equations if we have a static charge at a point, we get $\vec{j} = 0$, and then there's no magnetic field so A is zero.

I can take initial data with a blob of charge. The thing would flow apart and become diluted. We can solve that on a computer, solve this thing.

[This is a system, well, now $\vec{j}$ and ρ can be written in a way so that [unintelligible]]

[What were, how was this advertised while I was at Stop-N-Shop getting my flu shot?]

This is a charged scalar field coupled to electromagnetism.

[Where do you see those? Why is it charged?]

Because $D_\mu\varphi = \partial_\mu\varphi - iA_\mu\varphi$.

[Leon: Complex valued is what it means.]

I think it means it couples to the gauge field.

[Denny: φ seems to be an artificially introduced scalar field having something to do with an honest to God electric field.]

The field will be different because something is a spinor.

[What about photons?]

Photons are described by A_μ.

We hope to observe the Higgs field in the next few years.

Okay, so $\varphi = \rho e^{i\theta}$, and

$$\vec{j} = i(\bar{\varphi}\vec{\nabla}_A\varphi - \vec{\nabla}_A\bar{\varphi}\varphi) = -\rho^2(\vec{\nabla}\theta - vecA)$$

with

$$\rho_Q = -\rho^2(\dot{\theta} - A_0).$$

So these are equations for φ and A rewritten with ρ and $\vec{j}$. So this is $\frac{1}{e^2}\partial^\mu F_{\mu\nu} = -\rho^2(\partial_\nu\theta - A_\nu)$.

Next we could plug in and solve the φ equations for constraints on ρ and θ.

One solution is $\varphi = 0$ and light is propagating. Another would be with $\theta = 0$ and $A_\mu = 0$ and ρ a plane wave, that would also be a solution, those are decoupled solutions.

[What about, you have E and B which are real things. But are ρ and j real things? They have A in them.]

But they don't, they also have θ in them which is not real, but $\nabla\theta - A$, if you change θ by α and A by $\nabla\alpha$.

[Denny: I think Dennis is wrong, I think E and B are not real things, they change under Lorentz transformations.]

[Dennis: We're talking at the level of gauge.]

If I hit the one side $\partial^\mu F_{\mu\nu}$ with another ∂_ν I get zero. So I get two terms when I do this on the right side. I get two types of terms. The ones where they hit the φ and $\bar{\varphi}$ cancel with one another, but then the other half relies on the original $\frac{\partial S}{\partial\varphi}$ equation to be identically zero.

So these are $\dot{\rho}_Q - \vec{\nabla}\cdot\vec{j} = 0$, which is conservation of charge.

We started out with two φ and four A, and then we came out with six equations, which are the correct numbers.

[On the face of it there are nine equations and ten variables.]

There is another constraint on the dynamics. These give you the conservation and also the dynamics of the matter. I don't know how to rewrite this in some nice way. I want to see how to write it in terms of j and ρ_Q.

If j is zero, what does that say? This tells you how to make a magnet, radio transmitters, antennas, if j is zero, and B is a flow, then E is frozen in the moton of B. That's how a vector field acts on a vector field, apply the curl.

If you set j to zero you get an E field and a perpendicular B field.

[Dan: For the $SU(2)$ case, what does φ mean?]

None of these things will be directly physical observables. Everything will be strongly coupled. In the Yang Mills case F is already nonlinear. Quantum corrections also change the behavior for the non-Abelian case.

[Blaine: let me be a bonehead, what is going on here in the big picture, is this a replacement for classical electromagnetics?]

In classical electrodynamics j and ρ can be anything that obey the conservation equations. But here j and ρ evolve in time, and this gives a particular model which is guaranteed to be self-consistent. By deriving everything from an action principle I get everything in the right way.

[Let me be really bonehead. Back when we were sticking things in the right hand side, well, how do you find your φs? Do you think of it as having any physical significance in terms of the classical theory?]

In the classical theory we do particles for the charges and field theory for the fields. That's a very bad thing when we want to quantize. But the charges we see in daily life are fermions, which don't clump, and so you get individual charges as a good model for fermions. Particularly if you have a little charge on a big blob of matter, then we, so, that's in real life, we're used to this hybrid picture. Well, fermions are not so nice to treat at a field level.

[Dennis: this is how I would answer the question if I knew what you knew. This doesn't describe reality. But if there was a world where this described what happened, I could, say, go measure the electric and magnetic field. By some stroke of genius we write E and B as F and then write F as dA. Now there's j and ρ and we introduce another field called φ. What is φ, in this other world where it exists?]

φ is the charge field.

[But it's not a real thing.]

Well, ρ_Q is the charge density, and you can write φ in terms of it. It's the wave function for the charge.

[So φ is like the wave function for the charge?]

[That's why it's complex and not real?]

I suppose if you tried this with the real φ, it wouldn't work out.

You can make a real version of this story by freezing ρ. You get some sort of model, not very exciting.

[When was this done?]

The forties, fifties?

[Oh, so it's only fifty years old, we don't have to understand it, it's only been fifty years.]

Thank you very much.