IAIN STEWART: So last time we were talking about the subleading one of our MQ Lagrangian HQET. Hopefully more people remember that we have a makeup lecture today, or will not be stopped by the snow. So today we're going to continue this discussion of power suppressed Lagrangians. And I'll explain to you why we think of these as giving power corrections to observables. And we'll talk about a couple different observables where these specific operators are playing an important role.

And then we're going to turn to the topic of renormalons, which is fun stuff. How many people here know what a renormalon is?

AUDIENCE: [INAUDIBLE]

IAIN STEWART: Yeah. By the end of the lecture, you all know what a renormalon is. OK, so this is where we left off. We were talking about this symmetry reparameterization invariance. And what we showed last time is that the Wilson coefficient of this operator is 1 to all orders in perturbation theory because of that symmetry. The Wilson coefficient of this operator is not. And I told you that at leading log order, it would be given by some expression like this.

But in general, this is just the lowest order expression. And it gets perturbatively corrected. It's actually known at three loops. So let's continue. We've talked about reparameterization invariance for the Lagrangian. Reparameterization invariance also has important consequences for operators.

So just like we wrote down sub-leading Lagrangians, we should also write sub-leading currents. So there's 1 over mq supressed currents. And if you look at those currents and you construct them, then you find, again, that reparameterization invariance plays an important role. And again, there's relations between Wilson coefficients.

And what happens in this case is actually that the leading order operator had a Wilson coefficient, the lowest order operator. So the Wilson coefficients of the subleading operators get related to that of the leading operator. And I may give you a problem on that on your next problem set. But I haven't quite made up my mind yet.

So rather than go through that, I want to talk about some observables where these operators played a role. So we've seen one example of reparameterization invariance. And I think from that one example, you would be able to do other examples. Let's talk about masses. These operators here actually have an impact on the masses of hadrons.

If we think about the mass of a heavy meson, h, there's a heavy quark in that heavy meson. So we can pull out that heavy quark. And then there's some remainder. And we can characterize higher order terms in this formula. So let me first explain what this lambda bar is. Our Lagrangian here, if we're working at lowest order in the theory, our Lagrangian would be L0 of HQET plus, for the light quarks, they're basically just a QCD Lagrangian, so for the light quarks and the gluons.

So for the heavy quark, we have HQET, and then there would be higher order terms that are 1 over m suppressed, which are the ones up there. So the question is, if we have this Lagrangian, what is the mass of a heavy state like the b meson or the d meson or the d star? And we can characterize that by knowing the structure of the Lagrangian.

So first of all, what about this term, this lambda bar? That term actually comes from these parts of the Lagrangian. So if we take these parts of the Lagrangian, we can calculate the Hamiltonian. Let me call that H0. And the definition of this lambda bar is the thing that you get by taking that Hamiltonian and letting it act on the state.

So you can think that this is-- if you like, you can think that this state is just an eigenstate of the Hamiltonian. You get the energy. But since the Hamiltonian has no heavy quark mass in it, what you're getting is some parameter, which is traditionally called lambda bar, that doesn't have a heavy quark mass in it. And so there's no heavy quark mass in here. So it's independent of mq. It's also independent of spin.

So B versus B star and flavor B versus D. So whether it's a B meson, D meson, as long as we're in this limit where both the bottom mass and the char mass are heavy, if we're thinking of this expansion, then this is a universal thing just determined by these Lagrangians. It's traditionally called lambda bar. And it's telling you that if you think about this thing as having an order mq piece, this is the order mq to the 0 piece. The symmetry of the theory is telling us that this is a universal constant.

It does depend on the state here, although that state could be connected to other states by spin symmetry and flavor symmetry. There was this quantum number we talked about last time, which we called SL for the spin of the light degrees of freedom, and then pi for the parity of the light degrees of freedom. So it's the same within every multiplet that we get from a fixed SL pi. But there's a different value for each multiplet.

So you'd have a different value of lambda bar for the B, D, B star B than for the lambda B and the lambda C, for example. So that's this term. We can also characterize what's going on with this term by using this Lagrangian here, the L1. So at order 1 over MQ, we can figure out what kind of contributions to the mass we're getting. h1 is just minus L1. You can think of it now perturbatively. So that just flips the sign of what we had before.

And we get two parameters from taking matrix elements here. They get the following names. And let me just do everything here in terms of-- it's not totally crucial, but it's a little easier to interpret certain things if we work in the rest frame. So I'll just work in the rest frame. So if you work in the rest frame, you should just think of this as like giving some kind of measure of the kinetic energy that the heavy quark gets wiggled by. And that's called lambda 1.

This, again, is a matrix element that doesn't know about the heavy quark mass. And the heavy quark mass in an explicit MQ. So this guy has dimension 2 if you work out the dimensions. And then you can do the same for the other operator. And here you have to think a little bit harder.

And if you think a little bit harder, you can characterize what the coefficients are for different spins. So there's some Wilson coefficient. And then we have this operator. And this here is like sigma dot B. I told you that before. And so you should think of what's going on here as you have sigma, which is giving the SQ, and then you have a B field. And then it's a question of what vector could be left over that the B field could know about. And it's basically SL.

So this guy here is leading to the SQ. And this guy here is leading to the SL. And so that's where this SQ.SL comes from. So you can think about other possibilities, but you have to have the right symmetry structure under time reversal and parity and things like that. And that forces you to use a spin here for the B and not something like the V.

So this SQ.SL is something that you can write out as, remembering the definition of it and its relation to J squared. You can write out that factor like that. And then we can derive here how this lambda 2 contributes to the differing states, because we have a different contribution for the B and the B star from the lambda 2 guy, because it was violating the spin symmetry. So the lambda 1 guy is going to violate the flavor symmetry. That's this 1 over MQ.

And so charm quarks-- charm mesons and bottom mesons will get different contributions from that. They have the same lambda 1, but they have a different connection because it gets suppressed by a different factor. And then again for the lamba 2, Lambda 2 is a little bit different because it actually has mu dependents.

That's not quite right. So it actually has MQ dependents because the coefficient here has mu dependents. So this operator doesn't have any MQ dependence, but the coefficient does, remember. The mu dependence cancels and there's a left over MQ dependence to the lambda 2, but it's logarithmic MQ dependence.

If I define it this way-- so that's why I wrote MQ there. So you could, as we talked about last time, you could sum up the logs of large logs and have this guy-- you could re-sum logs, if you want. So you could think of these as you can re-sum logs inside this lambda 2. So if we ignore those logs or we imagine that we re-sum them and we just look at the remainder, then we have the expectations based on power counting for how big these things are.

And we just [? expect ?] that they're both given by the dimension. And the only [? dimensionful ?] parameter around is lambda QCD because the MQs can only occur in logarithms, and the Wilson coefficient in nowhere else. And these are non-perturbative parameters. Lambda bar is also a non-perturbative parameter. These ones have a little more dynamics in them.

So then we could go out and write for our states the results by putting these things together. So we'd have M capital B plub M little B plus lambda bar lambda 1 over 2 and little B. And if we go through the spin structure there, we get a 3/2 and then lamba 2 logarithmic MB dependence over MB.

So when we start with the B star, the first three terms are the same. And then this guy comes in with a plus. And it's lambda 2 over MB over 2MB. And then the D states are similar in turn. Lambda bar lambda 1 over 2M term, and then these two terms again. Same spins, so it's lambda 2. And the only difference is, now we're evaluating the lambda 2 as logarithm and charm dependence,

OK. So there's some formulas that are correct, actually-- including all the way up to 1 over M corrections. And you can see that there's still sort of a structure to these things. This is a universal correction for all of them. These corrections are universal between B and B star. And so the splitting between B and B star is just given by this lambda 2. That's what causes the states to split. So up to that level, they're the same. And that's why the B and B star are very degenerate and the masses are very close.

So you can form combinations to cancel things out. So you can take a kind of spin average mass where you take 3 times the vector plus the scalar divided by 4. And if you do that, then you're canceling out the lambda 2. And then you can also form differences. So if you looked at MV star squared minus MB squared, then the only thing that's causing a difference is lambda 2.

Numerically, this is 0.49 GV squared. And from our formula up there, it's 4 lambda 2 of MB, so something like four times 0.12 GV squared. So that's the value of this lambda 2 is 0.12 GV squared, which is about lambda QCD squared. So our dimensional analysis is working. And then you can do the same thing for the D meson. Everything here works beautifully, actually.

Extract the value of lambda 2 and you see there's a slight difference because that difference can be attributed to the logarithms. So if you take lambda 2 MB over lambda 2 MC, then the only thing that differs is the Wilson coefficient. So that's something you can predict. So experimentally, if you take the ratio of those two things there, you get a prediction like this one. And if you plug in theory like leading log RGE, then you're getting alpha S of MB.

Oh, sorry. [INAUDIBLE]. And this is 1.17. So that's for three light flavors, which is the right thing to do for the beta function. So it's not so bad for a leading log prediction. That's kind of the accuracy you would expect. All right. So we're really understanding quite a bit from our [INAUDIBLE] [? instruction ?] of this effective theory about the states, and even make predictions for the ratios of the mass divided by that as something preservative, which is kind of non-trivial.

Perturbative up to the order we were working. There would be corrections if we continued in our series here. There could be some non-perturbative corrections. But up to 1 over M, the ratio of this to that, which is what I'd wrote over there, is perturbative. All right. So that's some phenomenology. It's kind of baseline phenomenology that we can do with the Hamiltonian or with Hamiltonian derived from our subleading Lagrangian.

There's other phenomenology that we can do. And I'll mention some of the most important phenomenology. There's another class of predictions that we can make where we have a lot of predictive power by using the effective theory. And I want to talk a little bit about that.

So you can look at semi-leptonic decays. And there's two different types of semi-leptonic decays that you can look at-- so-called exclusive decays and inclusive decays. So exclusive is making transitions between the meson states and inclusive is making a transition where you allow any charm state, not just the lowest order ground state charm, but you could have in this state here, XE, you could have a D pi or a D star pi pi or other things like that. So it doesn't even have to be a single hadron state.

So exclusive refers to a particular channel. Inclusive refers to some overall channels. So the theory in these two is quite different. In this one, you would have form factors for the current between the states. And I already mentioned to you that heavy quark symmetry reduces the number of form factors. You get this single Isgur-Wise function. And so heavy quark symmetry is very powerful here.

And you can actually also work out 1 over MQ corrections. And people have done that. And it turns out there aren't any.

So there's something called Luke's theorem, which actually was a result derived, I think, when Luke was a graduate student that proved there was no 1 over m Q corrections here. And then you can work out alpha s corrections. And so this kind of-- you can just keep going. People have worked out 1 over m Q squared corrections. And this kind of formalism is used to measure V c b because these decay rates would depend on V c b, and that's kind of what you're after if you're doing this. So you'd like to parameterize and figure out the hadronic physics as much as possible, and aid the experiments in getting V c b, and exactly this framework is used.

People nowadays-- what they do is, the 1 over m Q squared corrections need to be computed on the lattice. So lattice QCD computes those corrections. The perturbative corrections need to be computed by continuum people like me, so those are computed. And they put these things together to get V c b from these decays.

You could also do it with this inclusive guy. And here it's more interesting because you can basically do everything with pen and paper. So here, you can use an Operator Product Expansion, OPE. And actually, HQET constrains the form of the operator product expansion, as well.

And indeed, when you work out the leading power corrections, they turn out to not enter until two orders down. So again, you're protected from first-order corrections, so you just get corrections that occur at 1 over m Q squared. And furthermore, they depend only on the two parameters we already defined.

So it's not like you even get any new nonperturbative matrix comments. You just get the two we were talking over here, were the masses. OK? So there's perturbation corrections, and there's power corrections, but you're actually not even getting any new information relative to the masses, any new matrix elements. OK? So we'll talk a little bit more about this second one with the OPE.

There's a lot of work that's gone into this one, both from the theory side, as well as from experiment. So when you think about the decay rate for this B to X sub c l nu bar, it's a doubly differential decay rate in Q squared over the lepton pair, so that's Q mu, E l-- because you can pick it as a second variable-- and then you don't know the mass of the state X sub c because it could be different things. It could be a D. It could be a D star. It could be a D pi.

So that's another variable. So you have a triply differential spectrum, and you can compute it with an operator product expansion. I'm not going to go through the details of carrying out the operator product expansion. If you want to do reading about that, there's some supplemental reading that I haven't assigned to you where you can read about that in the book *Heavy Quark Physics* by Manohar and Wise.

So for our purposes here, let's think of what operator product expansion means. It is simply that we want to carry out an expansion in lambda QCD over m Q. And if you go into the details, there's an important role played by the fact that you're summing over all states, and that's allowing you to connect the partonic calculations to the hadronic calculations, basically by probability conservation, but I'm not going to go so much into that.

So when you do this operator product expansion, you can think of it in terms of diagrams. And usually, you draw these diagrams as forward diagrams. So here's kind of the matrix elements squared. Here is the final state. There's the b quark going in and another b quark going out, so that's like a matrix element squared. The amplitude squared.

And you think about doing an operator product expansion for that. And you match onto operators in the effective theory. Just to give you a schematic structure for those operators. So there's a leading-order operator. There's a subleading-order operator that has an extra covariant derivative in it. That's kind of what you would do.

And then you have Wilson coefficients. The leading-order operator here is just b bar b. That's what it turns out to be. This operator here would look familiar because it's D transverse squared, so that's like our lambda 1 operator, and that's where the lambda 1's are going to come from. And then there would be a lambda 2 term, and I just didn't write it. So there's a magnetic guy here, too.

So when you look at the leading-order operator, b bar b, that counts the number of b quarks, OK? That's just a number operator. So to all orders in perturbation theory, b bar b is 1. Yeah?

AUDIENCE: How do you cut these-- like, you're actually calculating [? B to A, right? ?] [INAUDIBLE]

IAIN STEWART: Right. So you should think of actually looking-- so this is a charm quark. And so you would write down that propagator, and then you have a large injection of momentum from the b quark, right? Use momentum conservation and expand the propagator. Thinking about the propagator, if you like, thinking about this guy here having-- so you could write this guy's momentum as m b v plus k, all right? And you would expand in k.

So at first order, you just drop all the k's, and that's basically giving you this term. At second order, you'd keep the k's, so the k squared would be identified with this D t squared.

AUDIENCE: Right, but how do you cut the effective [INAUDIBLE]?

IAIN STEWART: The effective theory diagram, it's already been cut, if you like. So this is just a real thing, and I've already taken the imaginary part. So this is all real. There's no cut to make in the effective theory. The thing I'm integrating out here-- so this is a good comment. The thing I'm integrating out here is hard and off-shell.

So when I go over to the effective theory and I get rid of the off-shell stuff, I have a real part here. There's nothing to cut in the effective theory. Sometimes, as you know, there are things to cut in the effective theory, but here there's nothing to cut on this side.

OK, so this is one tall order. So that actually means that we don't even have a nontrivial matrix element here. We just have a Wilson coefficient. So everything in the first-order term is completely calculable. No nonperturbative parameter because of symmetry. And this guy here is the power correction.

And if you look at what the Wilson coefficient is-- so it's a function of alpha s. You could think of it like this, if you want. And all the kinematic variables, like kinematic variables like these guys-- and it is equal to the free b quark decay, including the loop corrections, of course. So not only the tree-level result, but adding loop corrections to the right-hand side here.

If you add the loop corrections and you drop all the k squared's, then you get just the free b quark decay, no nonperturbative matrix elements, so you can just calculate that first term, order by order. Just doing a [? partonic ?] calculation gives you the right thing to describe this decay. And the OPE is telling you that that formally is exactly the right thing. Even if you just started to do it, it's actually technically the right thing to do here. It wouldn't be the right thing to do for the exclusive decays, but it is the right thing to do here for the inclusive.

So what is the range of validity of this sort of-- is there any restrictions on these kinematic variables? And there is. We're treating the kinematic variables as if they're hard. And that means basically that we're thinking that they scale like powers of the heavy quark mass. So there can be regions of phase space where that wouldn't be true, but as long as we stay away from those corners, edges, then what I said is right.

And it's a little more powerful than that. So stay away from edges. Or there's another way of thinking about it, and that is that you can integrate over regions of the Dalitz plane for the phase-space variables. That includes all the way up to the edges. And as long as the size of your integration region is order m Q, then you're also fine. So think of, like, something with, like, m B squared up here, d m x squared, for example. And you can-- it doesn't have to be exactly m B squared. It could be m B squared over 2, m B squared over 4, something that you're counting as order m B squared. And again, you would be fine with doing this type of operator product expansion.

So what happens is, if you restrict yourself to be close to the edges, then restricting yourself to be close to the edges introduces new scales in the problem. And if there's new scales in the problem, just having a power counting that separates out lambda, QCD, and m B would not be enough. You would have to do more detail. And there is actually some interesting things that happen there, and we probably will talk about at least one example later on when we talk about [? SCT. ?]

OK, so I already said it, that there's a nontrivial fact happening at NLO. Well, I already implied that there's no 1 over m b corrections, and so that's an important outcome from this result. If you want to derive this, you need the effective theory. You basically use the equation of motion to get that to be true. So this is NLO in the power corrections, and then NNLO as I said, is just lambda 1 and lambda 2 at order lambda QCD squared over m b squared.

So phenomenologically, this is actually wildly successful. This is-- as far as I know, this is the case in QCD or in any theory where people have actually carried out the operator product expansion in the most detail. So people have gone up to 1 over m Q to the 4th. So they calculated at least two loops for everything and maybe even three loops for the [? NF ?] pieces. And experimentally it's been explored to death.

You have these three variables, and they've constructed of order 80 moments from these three different kinematic variables, sliced and diced the decay rate in all sorts of imaginative and crazy ways. And they've really tested that this OPE [? always ?] works beautifully, OK? So you can think of that you're getting a consistent picture, 80 different observables, just from a few simple predictions and your perturbative results.

You also get v c b. So everything fits together with the framework agreement I've discussed with you, and you get a result for v c b, which I'm not going to write on the board. OK? So that gives you-- I didn't go through the details. I will give you some reading to learn more about this OPE if you're interested, and sort of [? seize ?] how some of the things I mentioned, like using the equation of motion, how that actually works out. It's all done very nicely in this book, *Heavy Quark Physics.*

But I just wanted to give you a flavor with this effective theory that there's something very useful you can do with it once you have it. And in this case, kind of the main phenomenological thing that you're after is v c b and what I've described to you, but you could also think about b decays, where you're looking for new physics. And again, if you can construct how the decay rate is looking, then you can have a hope of finding new physics in the coefficient of those decay rates, effectively in the Wilson coefficients of decay rates.

OK? So we're going to turn to something else, but let me pause and see if there's any questions.

AUDIENCE: Is this the best way to [? get ?] [? v c b? ?]

IAIN STEWART: The two are actually pretty competitive, the exclusive and inclusive. Basically because the lattice is doing a good job of the matrix elements that you need in the exclusive. And again, you have a lot of kinematic information you can use from the experiment on the shape. So they're pretty competitive, actually.

AUDIENCE: Do they agree?

IAIN STEWART: They agree-- I think the disagreement is at, like, kind of the 1.8 sigma level, so they agree. There's an interesting story in v u b. So you could do the same kind of thing, not for a charm quark, as we did here, but for a light quark. And then actually, if you do b to pi l nu and you look at the inclusive version of that, and you can do the same type of OPE-- a little bit different, actually, but you can do it in OPE there, as well. Then you go through and you get a v u b in the two different ways, and the disagreement's at the 2 and 1/2 sigma level, maybe even-- yeah. It fluctuates with time, but on average it's 2 and 1/2 sigma. [CHUCKLES]

So that's interesting. There's not an understanding of what's going on there. That's more interesting than-- this one basically agrees. OK, so rather than go further into this, I want to turn now to my promised topic to you, to tell you about renormalons. Actually, renormalons have something to do with power correction, so it's not disconnected from what we've been talking about. So what are the ideas here?

So we've already seen in our discussion of renormalization that there is a freedom in defining the perturbative series. And we kind of focused on the m s bar scheme as being the simplest thing. But I told you about some subtleties with m s bar, and now we're going to address one of them that's related to this thing called renormalons.

So we have some freedom, if you like, in adjusting the cutoffs. And we had this cutoff, mu, and it was dividing up what was perturbative and nonperturbative in m s bar. But we could have done something else. We could have used a different type of cutoff. We could have used a Wilsonian cutoff, and that would have divided up things a little bit differently.

And you should ask the question, are any possible way of dividing things up equivalent? We've already saw when we were doing calculations that, for the logarithms, they were equivalent. But it turns out that, if you don't-- that the powers, sort of the power separation of power divergences can actually have an impact. And that's related to what renormalons are.

And m s bar and dim reg basically avoids thinking about power corrections. You just set them to zero. And so what can happen is, in m s bar, your matrix elements are having a little bit too much UV physics, and your Wilson coefficients are actually sensitive to the IR. They're not sensitive-- you don't really see it when you look at the coefficient. You just see a number, and it looks pretty good, but there's actually an asymptotic structure to those Wilson coefficients that comes from higher orders of perturbation theory. There's things hiding there. That's what this discussion is going to be about.

So if you do actually not make a good choice, then it can be the case that when you go to higher orders in the perturbative expansion, you get lousy convergence. And that goes hand-in-hand, actually, with kind of another thing, which is harder to visualize, but I'll explain it. And that is that you have trouble extracting the nonperturbative parameters.

And the reason that you're having trouble-- so you could think about doing some calculation like that OPE over there. And you extract a value for lambda 1 from all your fits to all these moments I was describing to you. And then you go to one higher order in alpha s, and you extract another value, and all of a sudden, this guy changes by a factor of 2 and you wonder what's going on. And you wonder, well, what's nature telling you about the kinetic energy if I change the order of my perturbation theory and all of a sudden I'm extracting a factor of 2 different value for that matrix element.

And that's related actually to this poor convergence. These things go hand in hand. If there's pure convergence in the series, then when you extract matrix elements you can get different values. Phenomenologically, you can imagine how that would be related, but it's physically related, too. It has to do with the fact that there's a poor convergence here because you haven't divided up the IR physics and the UV physics fully correctly, and that's reflected in this inability to extract these guys in a convergent fashion.

And we can actually quantify that using something called renormalon techniques. So one way of saying it is that, if you make a poor choice, you're plagued by something called renormalons, so they're actually something bad and you don't want them.

So what goes wrong is that the short distance of coefficients have hidden power law sensitivity to the IR. You wanted these to be UV things, but they are sensitive to the IR. And it comes in dimensional regularization in the powers. We've been very careful about separating out logarithms, but we weren't as careful about powers, and it comes back to haunt us if we look carefully at the theory.

And there's a corresponding sensitivity to the UV in the matrix elements. So let me give you an example that's not working through formalism, but just numerics. So let's look at b to u e bar nu at lowest order. The up quark is massless. And we'll think about this-- like we were talking about for the OPE for charm, we'll think about it inclusively so that we can actually just look at this and it makes physical sense. The same thing that I told you about the charm case applies for the up quark case. The lowest-order prediction is just calculate this guy and include loop corrections. The leading power prediction is the same story as for the charm.

So that means that physically it's relevant to think about this decay rate. So what does it look like? There's some G Fermi. Integrating out the W boson. There's some factors of pi. The mass dimensions of this guy are 1. G Fermi squared is minus 4, so you need 5 powers, and the thing that's the setting the mass dimensions is m b, so you get 5 powers of m b.

And if you look at the perturbative series and you said set mu to equal m b, this is what it looks like. So epsilon-- I'm just introducing something which is a counting parameter, and it's 1. I could also make alpha s the counting parameter, but I actually want to stick a number in for alpha s, so alpha s is going to disappear in the next line, and I'll just keep the epsilon, which is just telling me the contribution is coming from this order.

So you have some choice here for how you define the b quark mass, OK? You could use the pole mass, pole in the b quark propagator. Or you could use the m s bar mass. Or you could use some other definition. And it's raised to the 5th power. So whatever definition you pick, it can be pretty sensitive to that.

So let me tell you what the results look like in three different schemes for that mass. So let's first do the pole scheme. This is the same. We have m b pole to the 5th power, and then we'd write out what the series looks like in that scheme.

You look like you're doing pretty good when you're at one loop. You went down by a factor of 5. But then you go into two loops, and you find that your correction at two loops is pretty much the same size as your correction at one loop. So you say, well, let's use m s bar. That changes the perturbative series.

And maybe you think you're doing a little bit better because at least, well-- (CHUCKLING) but this guy got bigger, and this guy is bigger than that guy, so it's also not really working. Well, if the phenomenology is as accurate as I told you, you must imagine that there is something that does work better than that. Indeed there is.

So there's other mass games, and I'll talk about one of them, continuing this table, where we switch to something called the 1S mass. Now [INAUDIBLE] looks like that. We're very happy. Or at least we're much happier than we were over here. So what is the 1S mass? The 1S mass is basically that you take half the perturbative mass for the epsilon system.

So the 1S mass is m b b bar system calculated perturbatively, divided by 2. OK, we'll talk more about why this is working in a minute. Let me also write down-- these are conversion formulas because you calculate the decay rate once and for all, so you calculate the first line. Then you know how to convert between schemes, and that's what's causing the perturbative serious to differ.

So m b pole is equal to m b m s bar, and there's a series that relates them that looks like that. And I could put in numbers here. And so that's what's changing the numbers when I go from this line to this line. Basically, kind of the 5 hours of 0.09 that you get takes you from here to here.

We could also think about switching from the pole scheme to this m b 1S scheme. We have a different series. It still doesn't look very convergent. The numbers look small, but-- this is supposed to be a 6-- but this number here is bigger than that number.

So that when I stick this series and combine it together with this series for the pole mass, then I get that series at the top of the board. And the problem is not in the m b 1S. The problem is in the m b pole. And the problem in the m b pole is being reflected in this series that relates them and its poor convergence.

AUDIENCE: Can you explain [? what ?] [? m ?] [? b ?] [? bar ?] [? being ?] [? perturbative ?] [? means? ?]

IAIN STEWART: Yeah, so you calculate-- so you can think of calculating the Coulomb potential between b and b bar, and that will give you alpha s corrections to just the mass. So you have 2 m b plus Coulomb potential plus, you know, radiative corrections to that. And you keep dressing it up and just perturbatively calculate as if it was a Coulomb problem, a QED problem, the mass of the b b bar state, and then you divide by 2.

AUDIENCE: [INAUDIBLE]

IAIN STEWART: Yeah

AUDIENCE: But you just pretend?

IAIN STEWART: There's nonperturbative corrections to it, but what we're extracting from it is the series. OK? I'll explain more why it's going to work, why this is, like, a reasonable choice, and we'll come to that. OK, so the lesson here is simply that the choice of mass scheme has a big impact on the perturbative series. And we haven't yet figured out why some things work and why some things don't, but we will.

It's absolutely crucial to use the right one. Otherwise, the predictions will not be good. Well, physically, we can actually argue right away why m b pole is not so good. And that's because, physically, there is no pole. There's no pole in the quark propagator. And that is because of confinement.

So we use this notion of pole and a quark propagator when we're doing perturbation theory. But nonperturbatively, it's not a well-defined notion, or at least there doesn't really exist poles in quark propagators. So it's only perturbatively a good notion.

So I'll say it's only perturbatively meaningful. And in reality, it's ambiguous, it's an ambiguous notion. And you can think that the amount by which it's ambiguous is related to hadronization, which is set by the scale lambda QCD. So there's an ambiguity in what you mean by a pole mass physically due to nonperturbative effects. Now, when we set up HQET, we actually used m pole. So we're using a questionable physical parameter.

So when we did these phase redefinitions, we were actually using the pole mass scheme because we were expanding about mass shell. So if you were to do other things, you can think about getting to other choices and implementing them in the HQET. No problem with doing that.

But there does exist another operator that we sort of dropped without thinking too hard about it. So if we switch to using a different mass scheme, there's some delta m, which you can think of as a series in alpha s, which we wrote on the board a moment ago. And where that delta m would show up is that you would have kind of an L delta m operator, which would just be exactly delta m, and then Q v bar Q v.

That would be left over when we cancel the masses in the Lagrangian, if you like. That's one way of thinking about it. So we have an extra term in the Lagrangian in HQET. And now you can ask, well, if we have an extra term, we'd better worry about power counting. And it turns out that the way we can understand why m s bar wasn't working is related to power counting.

So in m s bar, if you look at what delta m is, delta m is m bar itself times alpha. But by power counting, you don't want something that's growing with m b in the Lagrangian. m b is suppressed by alpha, but it's alpha at m b that is providing [? some ?] [? suppression, ?] but that's only 0.2, and that's actually just not enough suppression to get rid of the big m b here. You don't get-- you go from 5 GV-- you know, there's some number in front. Maybe you get down to 1 and 1/2, 2 GV, but you're not getting down to a small enough value.

So parametrically, this is just not good, both parametrically, because it grows with m b, and numerically it's too big. It's not order lambda QCD, which would be OK. So parametrically and numerically, delta m is just too big for HQET power counting in this m s bar scheme.

You'd include an operator that effectively is on average larger and more important than your kinetic term because the v dot D operator-- v dot D is counting like lambda QCD, and it had v dot D with two Q's. Now you have something delta m with two Q's, and if it's on average larger, then you're messed up. So that's why actually physically the m b bar is not a good choice. You can think about it from the effective theory from that way.

m b bar is actually a mass that's-- you're supposed to think of it as a mass that you use for high-energy physics, for physics really high energy, above the b quark mass scale. Then it's a good parameter. We're now doing physics below the b quark mass scale, and this is just one way of seeing why it's not such a good parameter there, because the perturbative corrections are just too big. So what about this 1S mass?

So here, it turns out that if you calculate delta m, you get-- in the same way that you think about corrections to hydrogen being order alpha squared, you get m b 1s alpha squared. And numerically, that's small enough when you put the coefficient in. So it still doesn't make us feel very good because it grows with m b, but if I just care about numerics and I'm doing b quark physics, it works because the alpha squared is enough suppression that this is a good mass scheme, and that's why the perturbation theory treating as order lambda QCD numerically, if not parametically. Is OK.

Now, if that doesn't sit pretty with you, you can do something even more fancy, and we'll talk about one example of more fancy a little later on, where you basically, instead of having m show up there, you have some parameter R, and you have something that you can just pick. And so you could pick a scheme here where R is of order lambda QCD-- by however, maybe you pick it be 1 GV or 500 MeV or whatever you like. So you can make up a scheme where this is true, and then both parametrically and numerically you're OK. And these schemes also work just as well as the 1S scheme. So we can get more fancy and also satisfy our formal requirement.

All right, so that's kind of phenomenological numbers and a bit of physics. Let's come back to some mathematics and talk about what renormalons are mathematically. Is there a mathematical way of characterizing what's going wrong? What if I didn't think about this physics? Could I do a calculation and see that something's going wrong? And the answer is yes.

So first I have to teach you a few things, if you don't know them already, about the asymptotics of perturbative series in quantum field theory. And these are actually not convergent series. So they are what are called asymptotic series. So what's the definition of an asymptotic series?

So we say a function has an asymmetric series, which I'll write as follows. Some coefficients. And we'll just call the expansion parameter alpha. You can think of it as alpha s, if you like, but this is just math, so you don't have to think about it as anything but the expansion parameter. And we say that a function has an asymptotic series if and only if the following is true. f of alpha minus the partial sum up to some level n is less than kind of the scaling that you would get from the next term in the series, which is n plus 2, for some numbers K N plus 2.

So that's actually a quite different definition than what you would have for a convergent series. For a convergence areas, you would say that you could pick any epsilon you like here. You could make the N big enough that this would get close to that. Here, I'm just saying that it's less than this with some power of alpha. So imagine that I pick alpha to be 0.1.

The thing is that this K here could still grow with N, and that actually will happen. So the truncation area in some sense is not being bounded in these asymptotic series. And perturbation theory and quantum field theory is generically of this structure, and I'll show you some examples of that. It's asymptotic rather than convergent.

So when you do a QFT calculation, a kind of typical result if you work out some asymptotics is that you would have these fn coefficients that are some power of a, and then times an n factorial. And that's kind of how they're scaling as n is getting large. You can think about the number of diagrams as just growing, but even when diagrams are not growing you can get these n factorials.

So that means if you're fixing alpha as some value of the parameter, and no matter how small you take it, at some order in perturbation theory, you're basically running out of gas, and your truncation-- you start to-- your predictions start to grow and diverge. So the corresponding values of these K's, if you had these f's that were of that form, would be that the K's are basically growing in the same way. So you'd need larger and larger numbers in order to satisfy the definition of asymptotic over there, and having these large numbers you're never able to satisfy convergent. So the series has zero radius of convergence.

And then you wonder, why have we been teaching you quantum field theory? [CHUCKLES] That's true. So what do we know? Even though these series are asymptotic, we can still make use of them. So typically what happens is that the series will decrease. For a while, it'll look like it's convergent, and then it'll start diverging. And you can characterize where it starts to diverge. And I'm not going to go through all the algebra, but I'll just tell you some facts.

And so you could imagine that you're doing some perturbation series, and it looks like it's going well. Things are converging. But then, [? some ?] [? things ?] don't start to go so well. So think of what I'm plotting here as the partial sum of your perturbation theory at N-th order. I sum up all the connections up to N, so it's like the series that I was writing over here. It's this thing.

Now, in perturbative QED, perturbative QCD, we're never getting more than a few terms. So you could say, well, OK, maybe I don't care so much about this. And in QED, that's actually the case. You don't really care so much about that. But in QCD, actually, the turnover happens already around three loops, and so as soon as you start including two loop corrections this is something you have to care about.

Now, you can say, well, the best thing I can do is stop doing perturbation theory at some N star because that's kind of where it looks the best. Like, I've gone toward something in it. And that actually is not such a bad thing. And it turns out you can characterize actually the mistake you're making by stopping there, and the mistake you're making by stopping there is of the following form.

So it's an exponential in 1 over alpha. So you would never be able to see kind of the correction that you'd need to get to the correct value, which I'm imagining is on the axis here, in perturbation theory because this doesn't have a perturbative expansion. And I'll actually show you that kind of the bad behavior here and this gap are related to power corrections. And they're exactly related to this connection that I was telling you between perturbative corrections and large-order asymptotics and power corrections.

This type of exponential is something that we can relate to power corrections, and we'll see how that pans out, although we might-- probably won't get there until next class. OK, so this is kind of a prelude, and now we're going to go into more detail. And it turns out, in order to go into more detail here, that, much as you make Fourier transforms to explore another space, we're going to do a transform to something called Borel space.

So it turns out that, when you have a divergent series or an asymptotic series, there's still degrees of divergence. And we can classify how divergent it is by using something called the Borel transform. Yeah?

AUDIENCE: [INAUDIBLE] will the series still be asymptotic?

IAIN STEWART: So no, it's not. Yeah, so 1S mass-- let's see. Yeah, it is still asymptotic, but it's-- in a way that I'll describe once we understand what this Borel is, it's much less divergent than the other series, than the pole series. If you think about the pole series as being degree-0 divergent, then this is like a few orders down less divergent. So once we define less divergent of a divergent series, then I'll be able to make that more precise.

It's a question of when power corrections come in. And if you look at, like-- so physically, the 1S mass is a physical thing for an epsilon state. If you look at how power corrections come in, then they're suppressed because there's an ambiguity between power corrections in perturbation theory. If the power corrections are suppressed, you have less ambiguity in the perturbation theory. That's another way of thinking about it. But the answer to your question.

AUDIENCE: [INAUDIBLE]

IAIN STEWART: It basically affects this a. All right, lots of good questions. Let's see how we get there. So when I transform, I'm going to call the transform function capital F, and I'm going to call its argument B for Borel. And so what's the definition of the Borel transform? We're going to define it as follows.

There's a first-order term, and I'm just going to put a delta function there. The real thing that matters is the series of terms that come next. And instead of having F to the n alpha to the n, I'm going to divide by an extra n factorial, and that's the definition of, given a set of F's, how I construct what's called the Borel transform. And because of the n factorial, I get improved convergence, so the n factorial is making it converge better.

So if there's a transform, there should be an inverse transform. So here's the inverse transform. It's an integral from 0 to infinity, b over alpha F of b. And that would get me back to the F. So if you have a convergent series and you think of this transform, then you just can go back and forth, and you get back to the original-- the function that you started with.

So if you have some series that's like that, you get back-- you could calculate that series, and you get some F of alpha just before Borel transforming. If you do this Borel transform, then you calculate this integral, you get back the same F of alpha from the inverse transform.

Let me give you an example that shows you that, if you have a divergent series, it's not too crazy to define the sum of the divergent series by using these transforms. So I claim that, for a divergent series, where F of b and the inverse transform exist-- so if we start with a divergent series, but these things exist, then we could just use this transform as a way of defining F of alpha.

So let me give you a simple example of that to convince you that it's not such a bad thing to do. So let's consider the following series. Just an alternating series, but alpha is a parameter greater than 1. That series doesn't converge. Of course, for alpha less than 1, it would converge.

And if I asked you what it should be, you would say, well, I calculate for alpha less than 1, I analytically continue, and it's 1 plus alpha. Well, that's one way of getting there. Another way of getting there is using this Borel transform. So if you calculate it here, F of b, you sum up 0 to infinity minus b to the n-- once you put the minus 1 in there and there's an extra n factorial.

So that's e to the minus b, perfectly well-behaved function. And then you have to do an interval, 0 to infinity, db. You could have e to the minus b over alpha times the F of b, which is e to the minus b, and that just gives you alpha over 1 plus alpha. And this is a perfectly well-defined interval for large alpha. Large alpha is even making it better. So the integral is perfectly well-defined, and we get an answer that we like.

OK, so this Borel transform is a useful way of dealing with divergent series. The real question is, what happens if the inverse transform doesn't exist? And that's what's going to be the thing that we're most interested in.

So if the F of the integral, integral over the F of b doesn't exist, then the integrand can tell us about the severity of the singularities, I.e., the severity of the divergence. And that's the real power of this Borel method. So let's do another example, where actually things will diverge. So let me take F n to be a to the minus n, n plus k factorial.

And if we do that Borel transform of that-- if you have a series, you can do it. And there's only one piece of it that we care about, so I'm just going to write that piece. So there's this piece that has a kind of pole-like structure, and it has a pole at b equals a. And what you say is you call this pole a renormalon. This is the renormalon. You call it a b equals a renromalon.

So you're characterizing the renormalon by where that pole is. A renormalon is a flavor of a particle. This has nothing to do with a particle. Usually, you think of poles as having something to do with particles, and that's where the name comes from. It's due to 't Hooft. Blame him. [CHUCKLES]

So if a is less than 0, your integration contour is positive, and the pole is on the other side, so you don't have a problem. The inverse transform exists. These are called UV renormalons. They still can guide the perturbation theory, and they can be important for thinking about why the perturbation theory is behaving the way it is, but you can do the inverse transform. The real ones that are kind of problematic are a greater than 0 because then the pole is on the integration contour, and we can't do the transform.

So we're integrating d b from 0 to infinity, and we just have a pole sitting on the axis, so what do we do? So you could think about characterizing how poorly behaved a perturbation theory is by looking at these poles, and that's kind of what I meant by having some kind of notion of how divergent a series is.

So you look in this Borel space, and you look at the real axis, and those poles actually can be on both sides. The most severe pole is the one that's closest to the origin, and that has the smallest value of b. And you can see if you have a small value of b, then you're multiplying here with some large numbers. So if a was 0.1, you'd be multiplying by 10 to the n. So these guys are more severe.

As you go to larger values, you get additional suppression from this a to the n, so if a was-- if you went all the way out to 100, then you'd have 1 over 100 to the n. That's good, but factorial eventually wins. So it's still divergent, but we look like it was a lot better. OK? So you can characterize how severe the series is diverging by the location of these poles.

And you can characterize the ambiguity in doing an integration by-- if you just think about one pole, you can characterize the ambiguity, so let's just imagine there's one pole. You can characterize the ambiguity by going above or below. So you have two possible ways of defining the interval. You don't know which to pick. One is to go above. The other is to go below.

And then you can think about what the ambiguity is. And the ambiguity is the contour 1. Call this c1. This is c2. The ambiguity is c1 minus c2, and that is circling the pole. So you can look at the residue of the pole, and that gives you the ambiguity.

So what we'll do next time is we'll look at an example in perturbative QCD. We'll see that there's a series that has a renormalon, and we'll find it. And then we'll look at the residue, and we'll find actually that out will pop lambda QCD. So we'll see, we'll calculate, for some perturbative series, that the pole mass really has an ambiguity of order lambda QCD by using this technique of going around a residue of a pole. OK?

AUDIENCE: [INAUDIBLE] when we say that the b quark has a 4 GV mass, what's the mass? [INAUDIBLE]

IAIN STEWART: Yeah, so if you say the b quark has, like, 4.2 GV mass, that's the m s bar mass. If you say it has a 4.73 GV mass, that's like the m b 1S mass. And if you look at how the PDG measures the b quark mass, they measure it in one of these schemes that I'm telling you about, like the 1S scheme. They [INAUDIBLE] result for that. You can convert between the 1S and the m s bar mass.

The m s bar mass is a perfectly well-defined quantity it doesn't actually have a renormalon, OK? That's not its problem. The pole mass has a renormalon. The m s bar mass, its problem is related to power counting in the low-energy theory. It's a good mass for high-energy physics. It's a bad mass for low-energy physics. That's the problem with m s bar. It's not it's not a technical problem, that it has a-- [INAUDIBLE] severity of renormalon.

It can have a higher-order renormalon, but it doesn't have the same severity of renormalon as the pole mass. The pole mass basically has the most severe possible renormalon, which is, in some units, 1/2, so there's an extra 2 to the n. We'll talk about that next time. So basically, it's the u equals 1/2 renormalons that really are causing a problem in perturbative [? QCD ?] because they're the most important-- there's higher-order ones. Even alpha s actually has a normalon, alpha s in the m s bar scheme, I suspect. [CHUCKLES] But even up to five loops you won't see it, or it's very hard to see. Let's stop there.