BOTTOM QUARK PHYSICS AND THE HEAVY QUARK EXPANSION

MICHAEL LUKE

Department of Physics
University of Toronto
60 St. George Street
Toronto, Ontario
Canada M5S1A7
E-mail: luke@physics.utoronto.ca

These lectures were given at TASI-02 on the subject of b quark physics.

1. Introduction

In these lectures I hope to give you an introduction to bottom quark physics, and a feel for the theoretical issues which are involved. Because of time constraints, I cannot cover all aspects of the field, so I have chosen to focus on a few key ideas. The most important is the idea of "factorization". This is a word which is used to mean many different things in particle physics, but here I use it in the very general sense of separating long and short-distance physics. This is important for B decays, because much of the theoretical progress in the subject relies on being able to separate the short-distance physics at scales $1/m_W$ and $1/m_b$ which determines the quark-level process (and for which perturbation theory is a useful tool) from the complex long-distance physics of hadronization. Effective field theory will be introduced as a tool which makes this factorization automatic.

The outline of these lectures is as follows: First, I will discuss the use of precision measurements of b quark transitions to probe physics at or above the weak scale. I will then discuss the extent to which b-flavoured hadrons may be used to probe the interactions of b quarks, and then look at some important applications. Finally, I will briefly discuss some current directions of research. Throughout the lectures I will use the decay $b \to s\gamma$ to illustrate the concepts, since a complete understanding of this decay requires almost all of the theoretical machinery I will discuss.

2. Precision Measurements and Flavour Physics

As you have heard in many lectures at this school, there are many good reasons to believe that the standard model is not the correct description of nature at scales above a few hundred GeV or so. A strongly-interacting symmetry-breaking sector and supersymmetry in one of its many forms are perennial favourites for new physics, as are the more recent ideas of large extra dimensions and Little Higgs models. All of these theories conjecture new particles at the 100 GeV-few TeV scale. Unfortunately, to date no experimental evidence for such particles has been found.

Historically, there have been two ways to discover new particles experimentally. The first is to build ever-larger colliders which can directly create more and more massive particles. The second is to perform precision low-energy measurements, which corresponds to looking for new physics via virtual effects. While the first is usually preferable, because it's easiest to study particles if you can actually make them, important complementary information can be found by following the second path. This is the course being pursued at the B factories at SLAC and KEK, as well as by many other precision low energy experiments, and has has a number of important successes in the past. Here are a few examples:

(1) Virtual c quark loops contribute to $K^0-\bar{K}^0$ mixing via the diagrams in Fig. 1. From the measured size of the mixing, Gaillard and Lee[1] were able to predict that m_c had to be less than a few GeV.

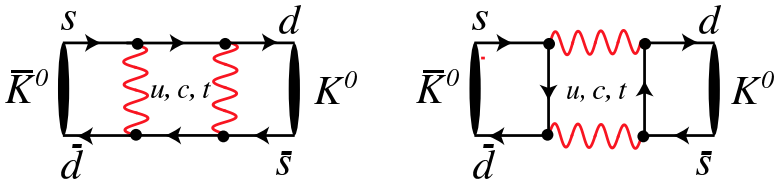

Figure 1. Quark-level diagrams contributing to $K^0 - \bar{K}^0$ mixing.

(2) Similarly, virtual t quark loops contribute to $B^0 - \bar{B}^0$ mixing. In 1986, the ARGUS and UA1 experiments measured a rate for this process that was much larger than expected[2], implying a top quark mass greater than $50 - 100$ GeV. This was the first evidence for a heavy top quark.

(3) Precision electroweak measurements (mostly at LEP) determined $m_t = 169^{+16}_{-18}\ ^{+17}_{-2}$ GeV[3], making it fractionally one of the best known

quark masses, before it was even discovered.

Indeed, the constraints from precision measurements puts strong constraints on a variety of new particles at the few TeV scale, and has even been dubbed the "little hierarchy" problem - the fact that there appears to be a sizeable hierarchy between the Higgs mass in the Standard Model and the scale of new four-fermi operators.

It is straightforward to calculate the Standard Model contribution to flavour transitions at the quark level. Since there are many more possible transitions than free parameters, one could therefore in principle compare with experiment and either discover discrepancies or place constraints on models of new physics. However, there is a problem: when comparing with experiment the theoretical predictions often have large theoretical errors arising from the difficulties in relating the underlying quark process to the observed hadronic process (denoted by the black blobs in Fig. 1). Even purely leptonic interactions can be plagued by hadronic uncertainties through loop effects, as the next example illustrates. Dealing with this issue will occupy much of these lectures.

Example: $g-2$ of the muon

An instructive example of the importance of understanding low-energy QCD is provided by the recent measurement of the anomalous magnetic moment of the muon by the E821 experiment at Brookhaven[4]. As is well known, there was initially great excitement in the community due to a reported $2.6\,\sigma$ discrepancy between theory and experiment:

$$(a_\mu^{\rm th} - a_\mu^{exp}) \times 10^{11} = 426 \pm 150^{\rm exp} \pm 67^{\rm th}. \qquad (1)$$

where $a_\mu \equiv \frac{1}{2}(g-2)$. This was seen as a possible signal of new physics, such as the supersymmetric contributions such as those in Fig. 2. Unfortunately, the discrepancy was reduced to a little over one standard deviation when a sign error was found in the theoretical calculation[5]. Nevertheless, it is worth considering what the theoretical limitations on this prediction are, to determine what level of discrepancy would be required to be a convincing signal of new physics.

The bulk of the contribution to $g-2$ arises from QED diagrams, which have been calculated to a quoted precision of $\pm 3 \times 10^{-11}$. Electroweak corrections give a much smaller contribution, but with a comparable error. These are both reliable because they are purely perturbative calculations. Unfortunately, even with muons, low-energy hadronic physics creeps in, via

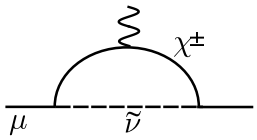

Figure 2. A supersymmetric contribution to $g-2$ of the muon.

the vacuum polarization graph in Fig. 3(a). Fortunately, while this graph is nonperturbative, the hadronic contribution to the photon vacuum polarization may be measured experimentally via e^+e^- annihilation to hadrons, as well as hadronic τ decay, via a dispersion relation[6]. (This is discussed in more detail in a different context in Section 3.4.)

Unfortunately, there is also a "light-by-light" scattering contribution, shown in Fig. 3(b), which cannot be extracted from data. This contribution is therefore typically estimated using models of hadrons, not from QCD[7]. A recent estimate[5] of this contribution to a_μ is $(80 \pm 40) \times 10^{-11}$, but the size of this error is controversial. The only completely model independent analysis is based on chiral perturbation theory, where there is a contribution from the unknown coefficient of an operator which could easily be several times larger than this uncertainty:

$$a_\mu^{LL}(\text{had}) = \left(13^{+50}_{-60} + 31\tilde{\text{C}}\right) \times 10^{-11} \tag{2}$$

where $\tilde{C}$ is a nonperturbative parameter which is expected, on dimensional grounds, to be of order 1. However, a value of 2 or 5 is equally natural; even the original discrepancy could be fit with $\tilde{C} \sim 10$. Fitting the experimental result gives

$$\tilde{C} = 7 \pm 6 \tag{3}$$

(where we have added the uncertainties in Ref. [8] in quadrature). Most

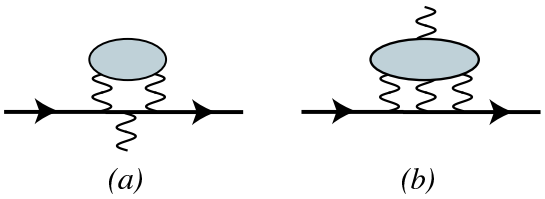

Figure 3. Contributions of (a) vacuum polarization and (b) light-by-light scattering to the muon anomalous magnetic moment. The blobs contains low-energy quarks and gluons and so are strongly interacting.

importantly, the contribution of $\tilde{C}$ is roughly the size of the proposed new physics contributions.

What this means is that it is extremely hard to convincingly extract signs of new physics from this measurement. Since we are currently unable to calculate $\tilde{C}$ from first principles, we are forced to rely on models of low-energy hadronic physics. Certainly, some models are demonstrably better than others[5], but the agreement between theory and experiment is good enough that without a model-independent treatment of the hadronic physics it is very difficult to convincingly argue that a discrepancy of this order could be a sign of new physics, rather than a failure in our understanding of low-energy QCD. Furthermore, it is difficult to see how this theoretical error could be reduced in the future without a fully nonperturbative calculation.

The moral of this story is that, when searching for new physics in precision experiments, we need precise and model-independent predictions with quantifiable error estimates. If hadrons are involved anywhere, this necessarily involves being able to systematically deal with low-energy QCD.

2.1. *b quarks as probes of New Physics*

Before looking for deviations from the Standard Model, we need to understand it. As indicated in Fig. 4, we observe a complicated pattern of flavour transitions, which follow a well-defined hierarchy. In the Standard Model, this pattern arises because of the hierarchy of the CKM matrix,

$$\mathfrak{L}_{cc} = -\frac{g}{\sqrt{2}} W_\mu \left(\bar{u}\ \bar{c}\ \bar{t}\right)_L \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L + \text{h.c.} \tag{4}$$

where

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \sim \begin{pmatrix} 1 & \lambda & \lambda^3 \\ \lambda & 1 & \lambda^2 \\ \lambda^3 & \lambda^2 & 1 \end{pmatrix} \tag{5}$$

and $\lambda = \sin\theta_C = 0.2205 \pm 0.0018$ is the Cabibbo angle. Since λ is small, it is useful to treat is as an expansion parameter. This allows us to write a particularly useful parameterization of the CKM matrix due to Wolfenstein[9]:

$$V_{\text{CKM}} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \tag{6}$$

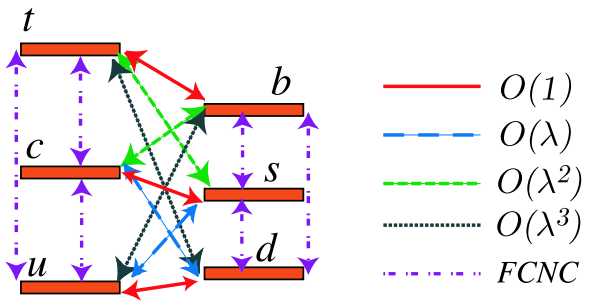

Figure 4. The relative magnitudes of quark transition amplitudes in the Standard Model.

λ is known at the $\sim 1\%$ level, $A = 0.83 \pm 0.03$ is known (due to measurements I will later discuss) at the $\sim 5\%$ level, and ρ and η have larger uncertainties.

A huge number of quark transitions is therefore determined by these four parameters (in addition to the quark masses, gauge couplings, and gauge boson masses), a few of which are shown in Fig. 5. The standard model is therefore extremely predictive, and the goal of the B factories and other precision low-energy experiments is to test these predictions as redundantly as possible.

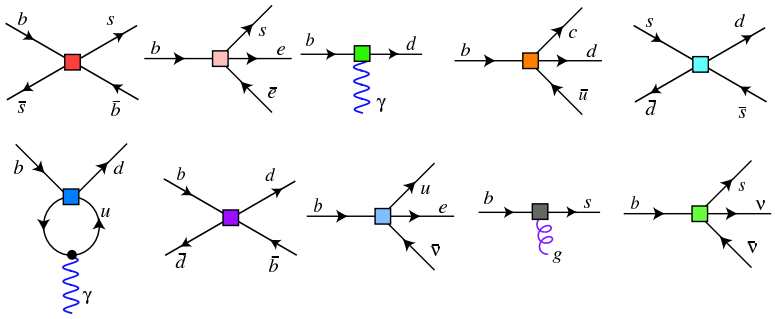

Figure 5. Some flavour-changing transitions.

Transitions between quarks of different charge, such as $b \to c\ell\bar{\nu}_\ell$ and $b \to cu\bar{d}$, are dominated by tree-level graphs in the SM and many of its extensions, and therefore simply reflect the CKM hierarchy. At one loop, however, things get more complicated. The decays denoted "FCNC" are flavour-changing-neutral current decays, in which the transition is between

quarks of the same charge. These decays are forbidden at tree level in the SM, but can arise through loop diagrams such as those in Fig. 6, and so are related in a complicated way to the various SM parameters. (In addition, the GIM mechanism[10] ensures that they are also forbidden in the limit of degenerate intermediate quarks.) Hence, the amplitudes for FCNC processes are suppressed in the SM, and so new physics at the few hundred GeV scale may contribute at a comparable level to these decays, as illustrated schematically in Fig. 7. Thus, rare b decays provide a sensitive

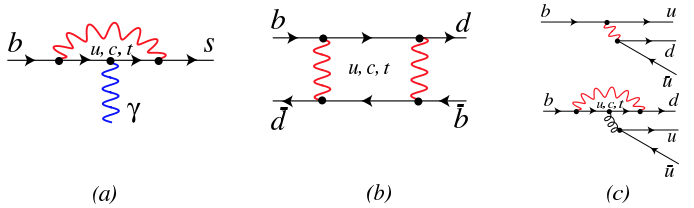

Figure 6. FCNC processes in the standard model. Note that some four-fermion transitions, such as $b \to u\bar{u}d$ (which in turn determines the hadronic process $B \to \pi\pi$) receive comparable contributions from tree-level and one-loop "penguin" amplitudes, because the tree level amplitude is suppressed by the small CKM element V_{ub}. This "penguin pollution" is a serious impediment to the theoretical prediction for such decays, as will be discussed in Sec. 4.2.2.

probe of the structure of flavour physics at high energy scales. Since there are many such transitions determined by a small number of parameters in the standard model, one can place strong constraints on (or discover!) new physics by studying as many transitions as possible.

The preceding arguments hold for transitions involving any of the quarks; however, b quarks are particularly good probes of short-distance physics. There are a number of reasons for this, both theoretical and experimental, but one of the most important is that there are a number of observables that can be cleanly predicted in the SM, despite the complications of hadronization. As we will discuss in these lectures, many of these arise because the b quark is heavy: $m_b \gg \Lambda_{\rm QCD}$, so b decays are short-distance compared with the QCD scale. However, clean observables are not restricted to the b system; for example, some rare kaon decays (such as $K \to \pi\nu\bar{\nu}$) are also very clean theoretically, and are also being intensely studied[11].

Let me make a couple of additional comments on the CKM matrix:

(1) **CP Violation:** $V_{\rm CKM}$ has a single nontrivial phase (i.e. a phase

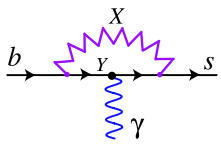

Figure 7. A contribution to $b \to s\gamma$ from new physics.

that cannot be absorbed in a field redefinition), which in the Wolfenstein parameterization is put into $V_{ub} = |V_{ub}|e^{-i\gamma}$ and $V_{td} = |V_{td}|e^{-i\beta}$ (these two parameters are not independent in the SM). This phase violates charge-parity (CP) invariance of the theory, and in the SM is responsible for the observed CP violation in the K and B systems. In the presence of CP violation, the rate for a decay and its CP conjugate may differ: $\Gamma(B \to f) \neq \Gamma(\bar{B} \to \bar{f})$, and so measurement of CP conjugate rates may be used to determine the phase.

CP violation is interesting for a number of reasons. First of all, it is interesting in its own right, in particular because the observed universe is highly CP-asymmetric: there is far more matter than antimatter. Hence, CP violation is necessary to explain the observed state of the universe[12]. For our purposes, however, it is interesting because (i) the SM is highly predictive (there is only one phase in the CKM matrix, and so all CP violation depends on the same phase) and (ii) the CP invariance of QCD may be used to get clean hadronic measurements of this phase, and hence to test the CKM picture (as will be discussed in Sec. 4.2.2). Note that the presence of a single nontrivial phase is a general feature of a 3×3 unitary matrix, so the weak interactions with three generations naturally violate CP (and hence CP violation is not a mystery, as is sometimes stated).

(2) **The Unitarity Triangle:** It is customary to express all of this graphically. The unitarity of $V_{\rm CKM}$ leads to the relation

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0. \tag{7}$$

Since each CKM element is a complex number, this relation may be drawn as a triangle the complex plane, as illustrated in Fig. 8. The length of the base of the triangle has been scaled to 1 by dividing each side by $|V_{cd}V_{cb}| = \sin\theta_C|V_{cb}|$, and the apex is at the point (ρ, η) as defined in Eq. (6). The angles β and γ are the CKM phases defined in Eq. (6), and $\alpha \equiv \pi - \beta - \gamma$. This construction is known as

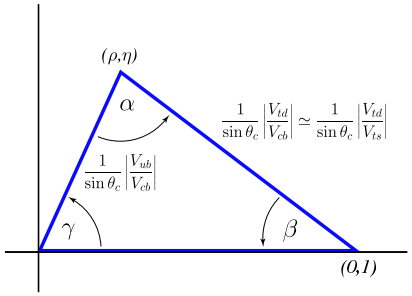

Figure 8. The unitarity relation (7) in the complex plane.

the "unitarity triangle" (UT). There are two other such triangles, arising from the two other unitarity relations, but the relation in Eq. (7) has the further property that each term is of order λ^3, so the sides of the unitarity triangle are of comparable length, while the other triangles have two sides much longer than the third and so are almost flat. This in turn means that the sides and angles of the UT should be measured with comparable precision to obtain useful constraints. Note that in the absence of CP violation, the UT would collapse to a line. Hence, by performing only CP conserving measurements of the sides of the triangle, one may determine the CP violating phase in $V_{\rm CKM}$.

The UT provides a convenient way to plot constraints on the SM from different experiments, as illustrated in Fig. 9. Note, however, that it can be misleading as a guide to new physics: two measurements of a given parameter may be redundant in the SM, but in the presence of new physics are as interesting as measurements of two different parameters. For example, the CP violation in both $B \to J/\psi K_s$ and $B \to \phi K_s$ is proportional to $\sin 2\beta$ in the SM. However, in the presence of new physics which contributes to the $b \to s\bar{s}s$ transition but not to $b \to c\bar{c}s$ this need no longer be true, and so both measurements are important. Thus, the goal is not simply to measure the sides and/or angles of the unitarity triangle, but rather to measure as many different quark-level transitions as possible, to look for deviations from the SM predictions. Currently, all determinations of sides and angles of the UT are consistent with one another within the errors (see Fig. 9).

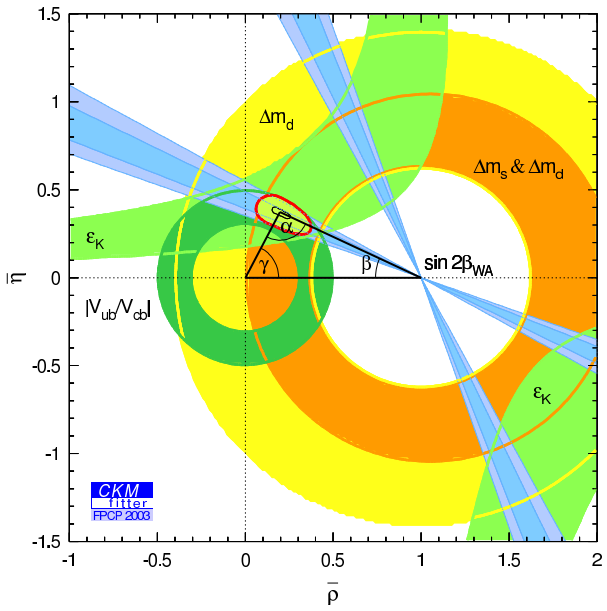

Figure 9. The experimental constraints on the unitarity triangle from different processes, and the best fit (as of summer, 2002).

3. From Quarks to Hadrons

In the last lecture we discussed how precision measurements of b quark transitions provide a sensitive probe of New Physics. Unfortunately, we can only do experiments on hadrons, not free quarks and gluons, and a simple quark transition like $b \to cu\bar{d}$ can give rise to many complicated hadronic process, such as $B \to D\pi$, $B \to D\rho$ or $B \to D\pi\pi\pi\pi$, etc., all of which are completely incalculable analytically. The problem, of course,

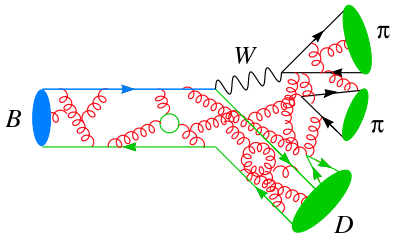

Figure 10. An incalculable mess.

is that even though the decay of a b quark is a short-distance process, occurring over a typical distance scale $r \sim 1/m_b$, hadronization is by its very nature long-distance ($r \sim 1/\Lambda_{\rm QCD}$) and nonperturbative.

Fortunately, a Feynman diagram like that in Fig. 10 is in fact a terrible representation of the physics of B decay. If you were to actually calculate this diagram, you would find loop integrals containing wildly disparate scales - m_W (the scale of the short-distance physics underlying the decay), m_b (the typical energy release of the decay) and $\Lambda_{\rm QCD}$ (the scale of hadronization). Some of the physics would be perturbative, and some nonperturbative. The short-distance physics would be independent of the complications of hadronization, whereas the long-distance physics would be insensitive to many of the details of the short-distance physics (just as fluid mechanics is insensitive to the structure of the proton). Thus, a naïve diagrammatic analysis obscures much of the important physics of the process, and the corresponding simplifications.

Example: $b \to s\gamma$

We can illustrate this with an instructive example. Suppose we were particularly sheltered theorists who had never heard of hadronization, but wanted to calculate the rare FCNC process $b \to s\gamma$ in perturbation theory. As discussed in the previous section, because this decay arises at the loop level in the SM, new physics at the weak scale can be competitive with the SM rate, and so it provides a sensitive probe of weak-scale physics. We therefore need to be able to calculate the rate in the SM, where it gets contributions from the diagrams in Fig. 11.

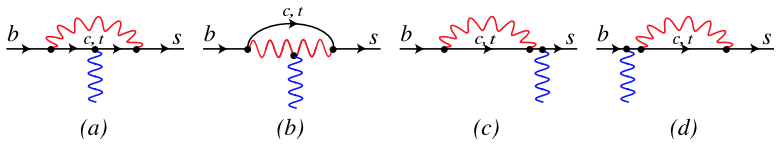

Figure 11. Feynman diagrams contributing to $b \to s\gamma$ decay in the SM at one loop. The contribution from intermediate u quarks is negligible.

Since we know about perturbation theory and infrared divergences we know we have to include graphs with both real and virtual gluons to have an infrared safe quantity. The decay is dominated by the t quark loop, and

we find at $O(\alpha_s)$[13]

$$\Gamma(b \to X_s\gamma) = \frac{\alpha G_F^2 m_b^5}{128\pi^4}|V_{ts}^* V_{tb}|^2 \left[F_2\left(\frac{m_t^2}{m_W^2}\right) + \frac{4}{3}\frac{\alpha_s}{\pi}\left(\log\frac{m_t^2}{m_b^2} + c\right)\right] \quad (8)$$

where[14]

$$F_2(x) = -\frac{2}{3}\left\{\left[-\frac{1}{4}\frac{1}{x-1} + \frac{3}{4}\frac{1}{(x-1)^2} + \frac{3}{2}\frac{1}{(x-1)^3}\right]x - \frac{3}{2}\frac{x^2}{(x-1)^4}\log x\right\}$$
$$+ \left[\frac{1}{2}\frac{1}{x-1} + \frac{9}{4}\frac{1}{(x-1)^2} + \frac{3}{2}\frac{1}{(x-1)^3}\right]x - \frac{3}{2}\frac{x^3}{(x-1)^4}\log x, \quad (9)$$

X_s denotes s or s+gluon, c is a non-logarithmic term, and I have neglected small terms suppressed by powers of m_s^2/m_b^2.

When we tried to compare with experiment, we would discover that nobody had measured the decay rate for the quark level process $b \to X_s\gamma$. However, we could argue that the final state quarks and gluons have to hadronize to *something*, so the appropriate thing to compare this to would be the sum of the decay rates to all final states containing a strange quark and a hard photon ($K^*\gamma$, $K^{**}\gamma$, $K\pi\gamma$, etc.). Such decays are called *inclusive* decays, to be distinguished from exclusive decays to a specific hadronic state. The measured inclusive branching fraction is[15]

$$\mathrm{Br}(B \to X_s\gamma) = (3.3 \pm 0.4) \times 10^{-4} \quad (10)$$

Comparing this with the theoretical prediction in Eq. (8) is, however, problematic.

At tree level, Eq. (8) gives to a branching fraction of about 1.3×10^{-4}, or almost a factor of three smaller than observation. Before we get too excited about having seen evidence for new physics, we notice that the one-loop correction to this result is huge, because it contains a term proportional to $\log m_t^2/m_b^2$. This provides an enhancement by about a factor of seven, which counteracts the α_s/π suppression from the loop. Indeed, the one-loop correction is more than a factor of two larger than the tree level result! There is also a large scale uncertainty in this result: evaluating $\alpha_s(\mu)$ at $\mu = m_b$ increases the predicted branching fraction by almost an order of magnitude to 1.2×10^{-3}, or about four times the observed rate, while choosing $\mu = m_W$ lowers the prediction to about 7×10^{-4}. Both m_W and m_b are relevant scales in the problem, so there is no obvious way to choose between them. The difference is formally a two-loop effect, but the corresponding uncertainty in the branching fraction is huge. Perturbation theory is therefore showing no signs of converging: two-loop, one-loop and tree level effects are all about the same size.

Clearly, we have a problem if we want to do precision physics. At the moment we have no idea whether the experimental result is in accord with the SM. Despite the fact that $\alpha_s(m_b)/\pi$ is less than 0.1, perturbation theory is failing badly because of these logarithmic enhancements. Nor will this problem stop at two loops - at n'th order in perturbation theory we will find terms proportional to $\alpha_s^n \log^n(m_t^2/m_b^2)$. It looks like we can only predict the decay rate to within an order of magnitude, much less the $\sim 10\%$ precision required to compete with the experimental result, and there is no way we can use this mode to search for new physics.

Physically, this makes no sense. The scale m_b is large enough for perturbation theory to be valid, and the scales m_W and m_t are completely irrelevant to a decaying b quark - a b quark can't resolve anything smaller than its Compton wavelength, so as far as the b quark is concerned, it is decaying via a point interaction. The poor behaviour of perturbation theory must just be an artifact of the fact that we are not calculating things sensibly. A sensible theory of b decay should never contain logarithms of m_t/m_b or m_W/m_b in matrix elements.

Furthermore, we have been rather glib in comparing the quark-level process with the inclusive hadronic process. The decay of a b quark to free quarks and gluons is not the same as the decay of a B meson to all possible hadrons:

(1) The initial states are different. A b quark bound in a hadron is not the same as a free b quark - it is interacting with a complicated state of strongly-interacting quarks and gluons, and so is not at rest. In addition, there are dynamical light quarks and gluons in the initial state. As a practical matter, the parton-level rate in Eq. (8) is proportional to m_b^5. Since this factor arises from phase space, we might be tempted to replace this with m_B^5 for B meson decay, which changes the prediction by about 50%. With no good reason as yet to choose one over the other, this is a large source of uncertainty.

(2) The final states are different. If we were to measure the inclusive hadronic invariant mass spectrum, we would find broad peaks at the masses of all the possible kaon resonances superimposed on a continuum from multiparticle hadronic states, while the parton level prediction has a peak at m_s followed by a perturbative tail from gluon and light-quark *bremsstrahlung*. Since the partons must hadronize to something, we may argue that if we average both dis-

tributions over a large number of resonances the two results should be equal (this idea is called "quark-hadron duality"), but how many resonances is enough? Should this argument also work for semileptonic D decay, which is dominated by only two resonances (the K and K^*)?

Both of these problems arise because of the interplay of physics at the decay scale m_b with the hadronization scale $\Lambda_{\rm QCD}$. Furthermore, in the limit $m_b \to \infty$ (or, more precisely, $\Lambda_{\rm QCD}/m_b \to 0$) both should vanish - an infinitely massive b quark is unaffected by collisions with light degrees of freedom, the difference between m_b and m_B is of order $\Lambda_{\rm QCD}$ and so a subleading effect, and in the limit $m_b \to \infty$ there is a always a large number of hadronic resonances to average over. Thus, we expect

$$\Gamma\left(B \to \sum_{\substack{X_s = K^*, K^{**}, \\ \cdots}} X_s\gamma\right) = \Gamma\left(b \to \sum_{\substack{X_s = s, sg, \\ sgg, s\bar{q}q, \ldots}} X_s\gamma\right) + O\left(\Lambda_{\rm QCD}/m_b\right). \quad (11)$$

However, on dimensional grounds these corrections should be at the 10% level, and as our phase space argument above indicates, may be much larger. If we really want to use B decays to look for new physics, we need a systematic way of calculating $\Lambda_{\rm QCD}/m_b$ corrections.

3.1. *Effective Field Theory*

The two problems raised in the previous examples ($\Lambda_{\rm QCD}/m_b$ corrections and large logarithms in perturbation theory) look different, but both have a common origin: the presence of multiple scales in the problem

$$m_W,\ m_t \gg m_b \gg \Lambda_{\rm QCD}. \quad (12)$$

To make progress, we need a tool which systematically disentangles the scales. This is the whole purpose of Effective Field Theory (EFT). Since Ira Rothstein has lectured extensively on EFT's at this school[16], I will be brief in my treatment and just stress the essential features[17].

The idea behind EFT is that the theory you use to calculate physics at a given scale μ should only contain degrees of freedom relevant at that scale. Anything happening at shorter distance scales - virtual heavy particle exchange, high momenta in loops - can't be resolved by external particles with momenta $\leq \mu$, and so is replaced in the EFT by local interactions. Short-distance physics is therefore removed from the dynamics and simply enters in the values of the coefficients of the local operators in the effective Lagrangian.

This is the same reasoning behind the multipole expansion in electrostatics: if we have a distribution of charges of size r_0 and are interested in physics at distances of order $r \sim r_0$ away, we need to know the precise charge distribution. On the other hand, at distances $r \gg r_0$, it is much more efficient to expand the theory in powers of r_0/r. At leading order, we only need to know the total charge to determine the potential. If we wish to be more accurate, we can include the dipole moment, quadrupole moment, etc, systematically expanding the true result in powers of r_0/r.

It is important to stress that EFT's are useful whether or not we know the "full" (short-distance) theory. Even if we had a theory of quantum gravity, we wouldn't use it to calculate projectile motion. Similarly, the fact that we have a renormalizable theory of the weak interactions doesn't mean that this is the appropriate theory to calculate low energy ($\mu \ll m_W$) processes. It may seem like we are stepping back seventy years, but Fermi's theory of the weak interactions, appropriately generalized, is actually the appropriate theory in which to calculate at low energies.

3.2. $m_W > \mu > m_b$: Four-fermi theory generalized

A familiar example of at EFT is four-fermi theory, the EFT of the weak interactions at scales much less than m_W. Suppose we were interested in nonleptonic $b \to c$ decay. At leading order this is determined by the diagram in Fig. 12(a). However, all external momenta in this graph are of order m_b, while the mass of the W boson is $m_W \gg m_b$, so according to the rules of EFT it should not be present in the theory. This is easy enough to implement: since the momentum transfer q through the W boson is $\ll m_W$, we can expand the amplitude in powers of q^2/m_W^2:

$$\begin{aligned} iA &= \frac{ig^2 V_{cb}}{8} \frac{1}{q^2 - m_W^2} \bar{c}\gamma^\mu (1-\gamma_5) b \bar{d} \gamma_\mu (1-\gamma_5) u \\ &= -\frac{4iG_F V_{cb}}{\sqrt{2}} \bar{c}_L \gamma^\mu b_L \bar{d}_L \gamma_\mu u_L \left(1 + \frac{q^2}{m_W^2} + \dots\right) \end{aligned} \tag{13}$$

where $\psi_L \equiv \frac{1}{2}(1-\gamma_5)\psi$. This amplitude may be reproduced to arbitrary order in q/m_W in an EFT with no W boson, but a series of higher-dimensional operators:

$$\begin{aligned} \mathfrak{L}_{\rm SM} \to \mathfrak{L}_{\rm eff} = \mathfrak{L}_{\rm QCD+QED} - \frac{4G_F V_{cb}}{\sqrt{2}} \Bigg[& \bar{c}_L \gamma^\mu b_L \bar{d}_L \gamma_\mu u_L \\ & - \frac{1}{m_W^2} \bar{c}_L \gamma^\mu b_L D^2 \left(\bar{d}_L \gamma_\mu u_L\right) + \dots \Bigg] \end{aligned} \tag{14}$$

This procedure of determining the coefficients of the operators in the EFT

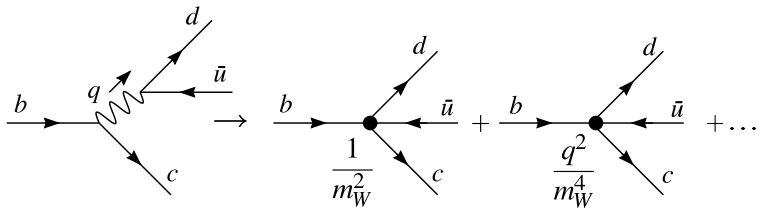

Figure 12. For momentum transfers $q \ll m_W$, the W boson may be integrated out of the effective theory and replaced by a series of nonrenormalizable operators.

by demanding that the full and effective theories give the same physics is called matching, and the W (along with the Z, higgs and t quark) is said to have been "integrated out" of the theory. In principle the EFT has an infinite number of terms, but for $q^2 \ll m_W^2$ (as is the case for b decays) it is usually sufficient just to keep the first order or two.

At tree level this is all quite trivial, but the power of this approach becomes evident when considering radiative corrections. Consider the one-loop correction to the weak matrix element in Fig. 13 (a), and the corresponding diagram in the EFT (b). The full theory graph is a mess to

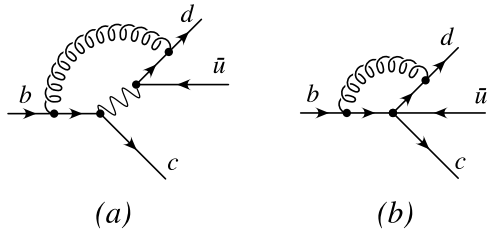

Figure 13. Radiative correction to b decay in the full and effective theories.

calculate (because it contains four propagators, two of which are massive), and is a complicated function of m_b and m_W. By power-counting, it is UV convergent. The effective theory graph is quite a bit simpler to calculate, but by power counting is UV divergent. This extra divergence looks worrisome, but it shouldn't be: the EFT is designed to get the physics right only at scales $\mu \ll m_W$. The full theory is convergent because at momenta $\mu \gg m_W$ the W propagator softens the high-momentum behaviour of the loop integral. Since the EFT isn't supposed to get the physics right at

momenta $\geq m_W$, it is not surprising that it is divergent. (Both graphs are also IR divergent.)

The EFT graph therefore requires additional renormalization. The result is then scheme-dependent, so it is not clear that it is telling us anything useful. However, the full theory graph will contain logarithms of the form

$$\sum_{ij} c_{ij} \ln \frac{m_W^2}{p_i \cdot p_j} \tag{15}$$

where the p_i's are the external momenta in the graph, in addition to infrared divergent logarithms of the gluon mass (if we choose to regulate the theory this way). These are the large logarithms which we encountered in the previous section, and which can spoil the convergence of perturbation theory. Since the EFT gets the physics of the external states right, it must reproduce these logarithms, and so must contain terms of the form

$$\sum_{ij} c'_{ij} \ln \frac{\mu^2}{p_i \cdot p_j} \tag{16}$$

where $c'_{ij} = c_{ij} + O(1/m_W^2)$, and the renormalization scale μ must appear in this way on dimensional grounds. In addition, the infrared divergences arise from low-momentum physics, so are identical (at leading order in $1/m_W^2$) in the two theories.

Thus, as advertised, at one loop the EFT correctly reproduces the low-energy physics (the logarithms) of the full theory. Of course, there are other important, non-logarithmic terms, but in the EFT these do not arise from loop graphs, but rather are contained in the coefficients of operators in the effective Lagrangian. These must be included for consistency, because before we can consistency calculate radiative corrections in the EFT, we must also calculate the radiative corrections to the matching procedure in Fig. 14 - in other words, we must adjust the coefficients of the four-fermi operators in the EFT to ensure that matrix elements are the same in the two theories up to $O(\alpha_s,\ 1/m_W^2)$. This is illustrated schematically in Fig. 14. In this case, we find that we must add an additional operator to the effective Lagrangian,

$$O_2 = \bar{c}_L T^a \gamma^\mu b_L \bar{d}_L T^a \gamma_\mu u_L. \tag{17}$$

whose coefficient is just given by the difference of the one-loop calculations in the full and effective theories. Thus, the difference in the finite pieces of the graphs goes into the coefficient $C_2(\mu)$. Furthermore, to work consistently to this order in α_s, matrix elements of O_2 in the effective theory are only evaluated at tree level.

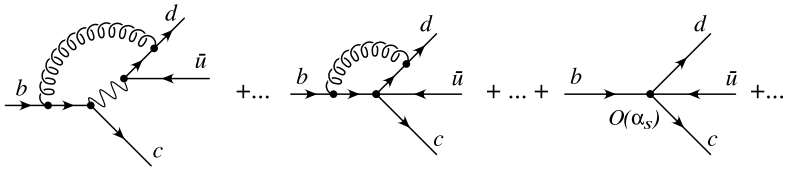

Figure 14. One-loop matching.

In general, then, the effective Lagrangian may be written

$$\mathfrak{L}_{\text{eff}} = \mathfrak{L}_{\text{QCD+QED}} + \sum_i C_i(\mu) O_i \tag{18}$$

where the O_i's denote operators renormalized at the scale μ, and the $C_i(\mu)$'s are determined by matching to a given order in α_s.

Note that in order for this whole procedure to work, it is crucial that logarithms of external momenta (as well as infrared divergences) are identical in the full and effective theories - otherwise, the difference between the graphs in the two theories could not be written as the matrix element of local operators. Tree-level matrix elements of local operators can only contain powers of external momenta (from derivatives), not logarithms.

At this point it may seem that we've gone through a long song-and-dance for no apparent reason. The EFT graphs may be simpler than the full theory graphs, but we have to calculate both to do the matching conditions. So what have we gained?

(1) We have explicitly factorized the long- and short-distance physics, with μ serving as the factorization scale. The long-distance physics (the logs of $\mu^2/p_i \cdot p_j$) is contained in the matrix elements of the operators O_i, while the short-distance logs of m_W/μ are contained in the coefficients in the effective Lagrangian. Note that the long-distance physics is process-dependent (differing for different external states) whereas the short-distance physics is incorporated into the parameters of the theory, and so is independent of the particular process being studied.

(2) Because the matching conditions are independent of the external states, we may calculate them using any external states we choose. In particular, if we are working at leading order in $1/m_W^2$, we may calculate the matrix elements in both theories between external states with zero momentum, which will greatly simplify the calculations. Both theories will give infrared divergent results, but the

infrared divergences will cancel in the matching.

(3) Finally, since the renormalization scale μ is arbitrary, we can eliminate large logarithms in matrix elements by simply choosing to renormalize the theory at $\mu \sim m_b$. Note that there appears to be a tradeoff here: choosing $\mu \sim m_b$ makes perturbation theory work well for matrix elements, but poorly for the matching conditions, since the coefficients $C_i(\mu)$ in Eq. (18) will contain the large logarithms. However, the renormalization group gives us the best of both worlds: we can calculate the matching conditions in the EFT renormalized at $\mu \sim m_W$, where perturbation works well for the C_i's, and then use the renormalization group equations (RGE's) to lower the renormalization scale to $\mu = m_b$, where we calculate matrix elements.

Let me elaborate on that last point. Matrix elements of the effective Hamiltonian are physical, and so independent of the unphysical renormalization scale μ. Thus,

$$\mu \frac{d}{d\mu} \langle C_i(\mu) O_i \rangle = 0 \tag{19}$$

and so

$$\mu \frac{d}{d\mu} C_i(\mu) = \sum_j \gamma_{ji} C_j(\mu) \tag{20}$$

where $d/d\mu$ is the total derivative, including the change in the coupling $g(\mu)$:

$$\begin{aligned} \mu \frac{d}{d\mu} &= \mu \frac{\partial}{\partial \mu} + \beta(g) \frac{\partial}{\partial g} \\ \beta(g) &= \mu \frac{dg}{d\mu} \equiv \beta_0 \frac{g^3}{16\pi^2} + \dots \end{aligned} \tag{21}$$

and $\beta_0 = -(11 - 2n_f/3)$ for n_f light quark flavours. The anomalous dimension matrix is defined by

$$\gamma_{ji} \equiv Z^{-1}_{jk} \left(\mu \frac{d}{d\mu} Z_{ki} \right) \tag{22}$$

where Z_{ij} relates the bare and renormalized operators,

$$O_i^{\text{bare}} = Z_{ij} O_j. \tag{23}$$

The RGE is solved by diagonalizing the anomalous dimension matrix; for the simple case where there is only one operator $O(\mu)$, the solution is

$$C(\mu_2) = \left(\frac{\alpha_s(\mu_1)}{\alpha_s(\mu_2)}\right)^{-\frac{\gamma_0}{2\beta_0}} C(\mu_1) \tag{24}$$

where the anomalous dimension of the operator O is $\gamma_0\alpha_s(\mu)/4\pi$. Expanding this out gives

$$C(\mu_2) = \left(1 - \frac{\alpha_s(\mu_2)}{4\pi}\gamma_0 \ln\frac{\mu_1}{\mu_2} + \frac{\alpha_s(\mu_2)^2}{32\pi^2}\gamma_0(\gamma_0+2\beta_0)\ln^2\frac{\mu_1}{\mu_2} + \dots\right) C(\mu_1). \tag{25}$$

and we see that the complete series of leading logs of μ_2/μ_1 has been summed by the RGE and put into the coefficient function $C(\mu_2)$.

The procedure is then straightforward. We calculate the matching conditions at the high scale $\mu = m_W$, where perturbation theory works well. Then, using the RGE, we evolve the coefficients from $\mu = m_W$ to $\mu \sim m_b$. The large logarithms are summed to all orders in the $C_i(\mu)$'s, and matrix elements at $\mu = m_b$ contain no logarithms of m_b^2/m_W^2.

Note that the leading logs in Eq. (25) correspond to matching at tree level, running at one loop and taking matrix elements at tree level. Including the one-loop matching corrections and one-loop matrix elements corresponds to including terms of order $\alpha_s^{n+1}\ln^n \mu_1/\mu_2$. This is the same order as the terms summed using the two-loop running, and so both must be included to work consistently at this order. Thus, as a general rule, the matching conditions and matrix elements must be evaluated at one lower order in α_s than the anomalous dimensions of the operators.

3.2.1. *Semileptonic decay*

Let me illustrate this with a few examples. In the SM, there are many flavour changing operators in the effective theory below m_W, but for a given process one only deals with the appropriate subset of operators. Semileptonic $b \to c$ (or $b \to u$) decay is the simplest: at leading order in $1/m_W$, the only operator is

$$O(\mu) = \bar{c}_L\gamma^\mu b_L \bar{\nu}_e\gamma_\mu e_L. \tag{26}$$

Since $\bar{c}\gamma^\mu(1-\gamma_5)b$ is a partially conserved current in QCD (because chiral symmetry is only softly broken by the quark masses m_c and m_b), this operator requires no additional renormalization, and so does not run below

the scale m_W. Thus,

$$C(m_b) = C(m_W) = -\frac{4G_F}{\sqrt{2}} V_{cb} \tag{27}$$

and there are no large logarithms of m_b/m_W in the semileptonic b decays to sum. The final result is

$$\Gamma(b \to X_c \ell \bar{\nu}) = \frac{G_F^2 |V_{cb}^2 m_b^5}{192\pi^3} f(m_c/m_b) + O(\alpha_s) \tag{28}$$

where

$$f(x) = 1 - 8x^2 + 8x^4 - ... \tag{29}$$

and the $O(\alpha_s)$ terms do not contain any logarithms of m_W.

3.2.2. *Nonleptonic $b \to c$ decay*

Nonleptonic decays are considerably more complicated. The simplest nonleptonic decay, the $\Delta b = 1$, $\Delta c = 1$ (i.e. $b \to cq\bar{q}$) transition, is mediated at tree level by the operator

$$O_I = \bar{c}_L \gamma^\mu b_L \bar{d}_L \gamma_\mu u_L \tag{30}$$

where

$$C_I(m_W) = -\frac{4G_F}{\sqrt{2}} V_{cb} V_{ud}^*. \tag{31}$$

However, under renormalization, diagrams such as that in Fig. 13 (b) require a counterterm with a different colour structure:

$$O_{II} = \bar{c}_{L,i} \gamma^\mu b_L^j \bar{d}_{L,j} T^a \gamma_\mu u_L^i \tag{32}$$

where i and j are explicit colour indices. (O_{II} is just a convenient linear combination of O_1 and O_2 from the previous discussion.) The corresponding RGE is therefore a matrix equation

$$\mu \frac{d}{d_\mu} \begin{pmatrix} C_I(\mu) \\ C_{II}(\mu) \end{pmatrix} = \frac{\alpha_s(\mu)}{4\pi} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} C_I(\mu) \\ C_{II}(\mu) \end{pmatrix} \tag{33}$$

where, at one loop,

$$\gamma_{11} = \gamma_{22} = -2, \quad \gamma_{12} = \gamma_{21} = 6. \tag{34}$$

Since γ_{12} and γ_{21} are nonzero, O_I and O_{II} mix under renormalization. The solution to Eq. (33) is

$$C_{I,II}(\mu) = \frac{1}{2} \left[\left(\frac{\alpha_s(m_W)}{\alpha_s(\mu)} \right)^{\frac{6}{23}} \pm \left(\frac{\alpha_s(m_W)}{\alpha_s(\mu)} \right)^{-\frac{12}{23}} \right]. \tag{35}$$

Thus, even though $C_{II}(m_W) = 0$ at tree level, it is induced for $\mu < m_W$. Using $\alpha_s(m_W) = 0.12$ and $\alpha_s(m_b) = 0.22$, we find

$$C_I(m_b) = 1.11,\ C_{II}(m_b) = -0.26. \tag{36}$$

3.2.3. $b \to s\gamma$ *decay*

The $\Delta b = 1$, $\Delta s = 1$ FCNC decay $b \to s\gamma$ is more complicated still. The diagrams in Fig. 11(a, b) match onto an operator in the EFT of the form

$$O_7 \sim \bar{s}_L \sigma^{\mu\nu} b_R F_{\mu\nu} \tag{37}$$

whose matrix element determines the decay rate at tree level. However, O_7 mixes under renormalization with a whole slew of other operators. For example, the operator $\bar{b}_L \gamma^\mu c_L \bar{c}_L \gamma_\mu s_L$ requires O_7 as a counterterm from the graph in Fig. 15. It turns out that a total of eight operators are required[18]:

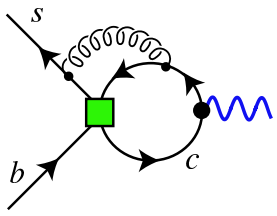

Figure 15. O_2 and O_7 mix under renormalization. This is a two-loop graph, but it contains only one strong loop and so contributes to the leading order anomalous dimension matrix.

$$H_{\text{eff}} = \frac{4G_F}{\sqrt{2}} V_{tb} V_{ts}^* \sum_{i=1}^{8} C_i(\mu) O_i \tag{38}$$

where

$$\begin{aligned}
O_1 &= (\bar{c}_{L\beta}\gamma^\mu b_{L\alpha})\,(\bar{s}_{L\alpha}\gamma_\mu c_{L\beta})\,, & O_2 &= (\bar{c}_{L\alpha}\gamma^\mu b_{L\alpha})\,(\bar{s}_{L\beta}\gamma_\mu c_{L\beta}) \\
O_3 &= \sum_{q=u..b} (\bar{s}_{L\alpha}\gamma^\mu b_{L\alpha})\,(\bar{q}_{L\beta}\gamma_\mu q_{L\beta})\,, & O_4 &= \sum_{q=u..b} (\bar{s}_{L\alpha}\gamma^\mu b_{L\beta})\,(\bar{q}_{L\beta}\gamma_\mu q_{L\alpha}) \\
O_5 &= \sum_{q=u..b} (\bar{s}_{L\alpha}\gamma^\mu b_{L\alpha})\,(\bar{q}_{R\beta}\gamma_\mu q_{R\beta})\,, & O_6 &= \sum_{q=u..b} (\bar{s}_{L\alpha}\gamma^\mu b_{L\beta})\,(\bar{q}_{R\beta}\gamma_\mu q_{R\alpha}) \\
O_7 &= \frac{e}{16\pi^2} m_b \bar{s}_{L\alpha} \sigma^{\mu\nu} b_{R\alpha} F_{\mu\nu}, & O_8 &= \frac{g}{16\pi^2} m_b \bar{s}_{L\alpha} \sigma^{\mu\nu} T^a_{\alpha\beta} b_{R\beta} G^a_{\mu\nu}
\end{aligned} \tag{39}$$

and the α's and β's are colour indices. At leading order in α_s only $C_2(m_W)$, $C_7(m_W)$ and $C_8(m_W)$ are nonzero, but the other operators mix with these

under renormalization, so one needs an 8×8 anomalous dimension matrix to obtain $C_7(m_b)$. Furthermore, the calculation is complicated considerably by the fact that the mixing of $O_2 - O_6$ with O_7 and O_8 first occurs at the two-loop level, through diagrams like that in Fig. 15. Since there is only one QCD loop in these diagrams, they contribute to the anomalous dimension matrix at leading order.

The leading order (LO) RGE's for $b \to X_s\gamma$ were first solved by Grinstein, Springer and Wise[18] in the simplified approximation in which the mixing of $O_3 - O_6$ with O_7 and O_8 was neglected. With this simplification, they found

$$C_7(m_b) = \left[\frac{\alpha_s(m_W)}{\alpha_s(m_b)}\right]^{\frac{16}{23}} \left\{ C_7(m_W) - \frac{8}{3} C_8(m_W) \left[1 - \left(\frac{\alpha_s(m_b)}{\alpha_s(m_W)}\right)^{\frac{2}{23}}\right] + \frac{232}{513}\left[1 - \left(\frac{\alpha_s(m_b)}{\alpha_s(m_W)}\right)^{\frac{19}{23}}\right]\right\}. \quad (40)$$

The complete anomalous dimension matrix was later completed by Misiak[19]; the results differ from this by about 10%. Resumming the leading logs in this way increases the tree level theoretical prediction in Eq. (8) by roughly a factor of three, bringing the prediction into agreement with the measurement, Eq. (10). However, there is still about a 25% theoretical uncertainty in this prediction, just by varying the renormalization scale by about a factor of two about $\mu = m_b$. For precision physics, this uncertainty is still disturbingly large, and to reduce it the next-to-leading-order (NLO) calculation is required.

The complete NLO calculation of $b \to X_s\gamma$ is probably the most involved QCD calculation ever performed. There have been many papers over the past ten years to get to the current state of the art[20]. The complicating factor is that, since the leading order RGE equations involve two-loop diagrams, the NLO calculation requires three-loop anomalous dimensions, as well as two-loop matching conditions and matrix elements. The current state of the art gives[20]

$$\mathrm{Br}(b \to X_s\gamma)^{\mathrm{theory}}_{E_\gamma > 1.6\ \mathrm{GeV}} = (3.57 \pm 0.30) \times 10^{-4}. \quad (41)$$

Comparing this with the experimental result, Eq. (10), we see that as long as the $1/m_b$ corrections relating this to the branching fraction for $B \to X_s\gamma$ are small, there is little room for discovering new physics in this decay. Since the NNLO calculation requires four-loop calculations, it is unlikely that the

theoretical error in this prediction will be able to be signficantly reduced anytime soon.

3.3. $\mu < m_b$: HQET

In the parton model calculation of inclusive B widths no scales below m_b appear in the calculation, so it is sufficient to work with the effective four-fermi theory discussed in the previous section, renormalized at $\mu = m_b$. This is a consequence of the fact that at leading order in $1/m_b$ the decay is entirely determined by short ($r \leq 1/m_b$) distance physics. On the other hand, decays to specific final states (exclusive decays) do get contributions from long distance physics, since that determines the details of hadronization. For example, to calculate the exclusive process $B \to K^*\gamma$ instead of the quark level process $b \to X_s\gamma$, we need to know the matrix elements of all eight operators (39) in the effective Hamiltonian (38) between a B meson and a K^*, which are all determined by long-distance physics. Thus, in keeping with the EFT philosophy, we would like to work in an EFT which factorizes the physics of the parton-level decay at $\mu = m_b$ from the physics at lower scales.

Since the long-distance contribution to exclusive decays is nonperturbative, it's not clear that switching to an EFT will be particularly useful, since EFT doesn't give us any new tools to solve strongly coupled theories. However, there are still reasons to switch to an EFT description:

(1) The EFT allows us to identify simplications (in particular, approximate symmetries) which (in some cases) still allow us to say something about the hadronic physics.
(2) The only way to calculate nonperturbative physics is to use lattice gauge theory. It is extremely inefficient to attempt to simulate a wide range of energy scales on the lattice, since the lattice spacing must be very fine (to simulate short distances correctly) and the lattice must be large (to fit the long-distance physics in). It is much more efficient to take care of the short-distance physics analytically, and only simulate the long-distance physics on the lattice. EFT does this automatically for us: an EFT renormalized at μ can be matched smoothly onto a theory with a lattice spacing $a \sim 1/\mu$.

I won't have much to say about lattice gauge theory in these lectures, so I will instead focus on the first point.

In general, factorizing physics at m_b from lower scales in B decay is

a much trickier procedure than simply integrating heavy particles out of the theory, since in a given decay there are partons carrying large ($\sim m_b$) momenta (such as the constituents of the K^* in $B \to K^*\gamma$ decay), so there may be both hard ($\sim m_b$) and soft ($\sim \Lambda_{\rm QCD}$) momentum transfers in the same process. This should be contrasted with four-fermi theory, where no external momenta are of order m_W. It is not obvious how to unravel the physics at these different scales, making it complicated to construct the appropriate EFT. I will say a little bit at the end of these lectures on the general problem, but in this section I will discuss an important case where we do not have to worry about such subtleties.

"Heavy Quark Effective Theory"[21,22,23] (HQET) is the EFT for a single heavy quark in which all hadronic momentum transfers are much smaller than m_b. This is much simpler than the general case because the scale m_b never enters the dynamics, so there is no complicated dynamical unraveling of scales required. This is the appropriate theory for semileptonic B to D decay, for example, at least in the region of phase space in which the D recoils slowly. However, it is not appropriate for $B \to K^*\gamma$ decay, since in the limit $m_b \to \infty$ the recoiling K^* carries energy of order $m_b/2$, which is not $\ll m_b$ in any limit.

HQET is conceptually a bit different from four-fermi theory. We can't simply integrate the b quark out of the theory as we did the W boson[a], because there are b quarks in the external states. Instead, since all hadronic momentum transfers in HQET are (by definition) $\ll m_b$, it is convenient to split the b quark momentum into a "large" piece (scaling like m_b) and a "residual" piece (scaling like $\Lambda_{\rm QCD}$):

$$p_b^\mu = m_b v^\mu + k^\mu \tag{42}$$

where v^μ is a fixed four-velocity (usually chosen to be the four-velocity of the meson, although any velocity differing from this by at most $O(\Lambda_{\rm QCD}/m_b)$ will do). Since all momentum transfers are soft, scattering in the EFT can only change the residual momentum k^μ. The EFT can then be obtained by expanding amplitudes in powers of k^μ/m_b.

This is easy to do: for example, expanding the b quark propagator in powers of k^μ/m_b gives

$$\frac{\not p + m_b}{p^2 - m_b^2 + i\epsilon} = \frac{m_b \not v + m_b}{2 m_b v\cdot k + i\epsilon} + O\left(\frac{k^\mu}{m_b}\right) = \left(\frac{1+\not v}{2}\right)\frac{1}{v\cdot k + i\epsilon} + \dots \tag{43}$$

[a] although we do integrate out b quark loops

which can then be inverted to give the kinetic term in the EFT:

$$\mathfrak{L}_{\rm HQET} = \bar{b}_v iD \cdot v b_v + O\left(\frac{1}{m_b}\right). \tag{44}$$

Note that the propagator includes the projector $\frac{1}{2}(1+\not{v})$, which projects out the upper 2 (particle) components on the b field, so the EFT is a theory of 2 component spinors b_v. The b_v fields have also been defined with a large phase $\exp(im_b v \cdot x)$ relative to the QCD field, so that derivatives acting on the b_v's pull down factors of k^μ instead of p_b^μ. The Feynman rules for HQET are shown in Fig. 16.

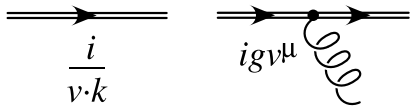

Figure 16. Feynman rules for HQET at leading order.

Let me make some comments:

(1) Working for simplicity in the rest frame of the meson, $v^\mu = (1,0,0,0)$, we see that the propagator is just the nonrelativistic energy k^0 of a static quark, and that it couples only to the electric field (via the scalar potential A^0). Such objects had actually been extensively studied before the advent of HQET - they are called Wilson lines, and are simply static colour charges. Indeed, this is exactly what a heavy quark looks like in the $m_Q \to \infty$ limit: since all external momenta are negligible compared with m_Q, interactions can't change the 4-velocity of the heavy quark, so in its rest frame it is just a static source of colour.

(2) The leading order HQET Lagrangian has no reference either to m_b or to the spin of the quark. Physically, this is because of the previous comment: as long as m_Q is much greater than any external momentum, the mass of the heavy quark is irrelevant to its dynamics - it's just a static source. Furthermore, the magnetic moment of a particle is inversely proportional to its mass, so the colour field of an infinitely massive quark only depends on its total charge, not its spin. This simple observation has profound implications, because it means that HQET has an enhanced symmetry compared with QCD: for n_h heavy flavours, long-distance QCD is symmetric under the interchange of either spin state of any of the heavy flavours

with the same four-velocity. Hence, the Lagrangian for HQET with n_h heavy flavours with the same four-velocity,

$$\mathfrak{L} = \sum_{i=1}^{n_h} \bar{h}_{i,v} iD \cdot v h_{i,v} \tag{45}$$

has an SU($2n_h$) spin-flavour symmetry. This will prove very useful shortly. Note that the symmetry exchanges quarks of the same velocity, not momentum.

(3) We can easily go to higher orders in $1/m_b$ by matching Green's functions[24,25]. For example, the three-point quark-quark-photon vertex can be expanded using the Gordon decomposition

$$\begin{aligned} \bar{u}' \not{A} u &= \bar{u}' \left[\frac{(p+p')^\mu}{2m_b} + i\frac{q_\nu \sigma^{\mu\nu}}{2m_b} \right] A_\mu \\ &= v \cdot A \bar{u}'\, u + \frac{1}{2m_q}(k+k') \cdot A \bar{u}'\, u + i\frac{q_\nu}{2m_b} \bar{u}' \sigma^{\mu\nu} A_\mu u. \end{aligned} \tag{46}$$

The first term arises from the leading order kinetic term in the Lagrangian; the second two arise from dimension five operators which can be added:

$$\begin{aligned} \mathfrak{L}_{\rm HQET} = \bar{b}_v\, iD \cdot v\, b_v + \frac{1}{2m_b} \left(\bar{b}_v (iD)^2 b_v + C_F(\mu) \frac{g}{2} \bar{b}_v \sigma \cdot G b_v \right) \\ + O(1/m_b^2). \end{aligned} \tag{47}$$

The first correction term is just the $k^2/2m$ kinetic energy term, while the second corresponds to the chromomagnetic moment of the b quark, and $C_F(m_b) = 1$. This term is the leading operator which breaks the heavy quark spin symmetry.

Weak Currents:

In addition, we can match weak currents between the full and effective theories. At leading order in α_s and $1/m_b$, this is trivial. For example, the weak $b \to c$ current matches onto the current

$$C(\mu) \bar{b}_v \gamma^\mu (1-\gamma_5) c, \qquad C(m_b) = 1 + O(\alpha_s) \tag{48}$$

in HQET. However, unlike in the full theory, this operator requires additional renormalization in the EFT. From the diagrams in Fig. 17, the current has an anomalous dimension[21,22]

$$\gamma_O = -\frac{g^2}{4\pi^2} \tag{49}$$

and so

$$C(\mu) = \left(\frac{\alpha_s(m_b)}{\alpha_s(\mu)}\right)^{-6/25} C(m_b) \tag{50}$$

But now we must be careful - this gives the correct running between $\mu = m_b$

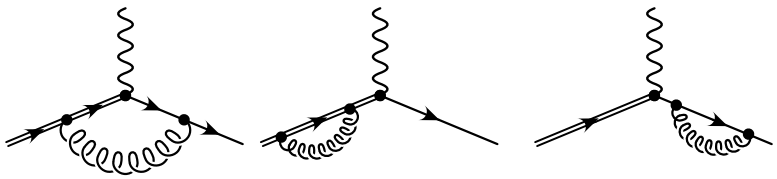

Figure 17. Diagrams contributing to the anomalous dimension of a heavy-light current in HQET.

and $\mu = m_c$. But below $\mu = m_c$, we now restrict all momentum transfers to be $< m_c$, which means we must match onto a new EFT where the c quark is treated as heavy as well,

$$\mathfrak{L} = \bar{b}_v iD \cdot v b_v + \bar{c}_{v'} iD \cdot v' c_{v'} + O(1/m_{b,c}) \tag{51}$$

and the current is

$$C(v \cdot v', \mu)\bar{b}_v \gamma_\mu (1 - \gamma_5) c_{v'} + O(1/m_{b,c}). \tag{52}$$

The anomalous dimension for this current is[26]

$$\gamma_T = \frac{g^2}{3\pi^2} (v \cdot v' r(v \cdot v') - 1) \tag{53}$$

where

$$r(w) = \frac{1}{\sqrt{w^2 - 1}} \ln\left(w + \sqrt{w^2 - 1}\right) \tag{54}$$

and so

$$C(v \cdot v', \mu) = \left(\frac{\alpha_s(m_b)}{\alpha_s(m_c)}\right)^{-6/25} \left(\frac{\alpha_s(m_c)}{\alpha_s(\mu)}\right)^{..} C(m_b), \quad \mu < m_c. \tag{55}$$

Note that at when $v = v'$ ($v \cdot v' = 1$), the anomalous dimension γ_T vanishes, and the heavy-heavy current doesn't run. This is because of the enhanced symmetry I mentioned earlier: when $v = v'$, the Lagrangian (51) has an $SU(4)$ spin-flavour symmetry which interchanges the spin components of both heavy flavours. The current $\bar{b}_v \Gamma c_v$ (for any Γ) is a generator of this symmetry. It is therefore related to a conserved charge, and so its anomalous dimension in HQET vanishes.

3.4. *Inclusive Decays and the OPE*

For general inclusive decays, we can't naïvely use HQET. Because the final state particles typically carry large ($O(m_b)$) momentum, it is not true that all the hadronic momentum transfers are soft. However, as we have argued on physical grounds, at leading order in $\Lambda_{\rm QCD}/m_b$, B meson decay is identical to free b quark decay. The argument depended on scale separation, so it is not surprising that we should be able to formulate the problem in the EFT language. The trick is that because the final state particles are highly energetic, through a bit of cleverness in the complex plane we can relate the physical decay amplitude to an amplitude in which the decay products off their mass shells by $O(m_b)$. In this region they may be integrated out of the theory just like a massive W boson, and we may treat the rest of the problem in HQET. This trick is called an operator product expansion (OPE), and its use long predates b physics.

$e^+e^- \to$ hadrons and τ decays

In a classic 1975 paper, Poggio, Quinn and Weinberg[27] showed that the ratio

$$R(s) = \frac{\sigma(e^+e^- \to \text{hadrons} + X)(s)}{\sigma(e^+e^- \to \mu^+\mu^- + X)(s)} \tag{56}$$

(where X denotes additional photons and fermi-antifermion pairs) was calculable in perturbation theory. It is instructive to go through their argument to see how it applies to the case of B decays.

The optical theorem tells us that $R(s)$ is proportional to the imaginary part of the photon self-energy (see Fig. 18)

$$R(s) = 12\pi \,\mathrm{Im}\,\Pi(s + i\epsilon) \tag{57}$$

where

$$\Pi^{\mu\nu}(q^2) = i \int d^4x\, e^{iq\cdot x} \langle 0|T j^\mu_{\rm EM}(x) j^{\nu\dagger}_{\rm EM}(0)|0\rangle. \tag{58}$$

Independent of perturbation theory, we know the analytic structure of

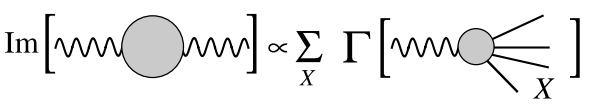

Figure 18. The optical theorem.

$\Pi(q^2)$: it is analytic everywhere except along the real axis for $q^2 > 4m_\pi^2$, where there is a branch cut due to real intermediate states, as shown in Fig. 19(a). (There are also isolated poles due to positronium bound states, but these are not important for our discussion.) Near the branch cut there are resonances which are sensitive to the details of long-distance QCD, so perturbation theory breaks down; however, in unphysical regions far from the cut we expect perturbation theory to be valid.

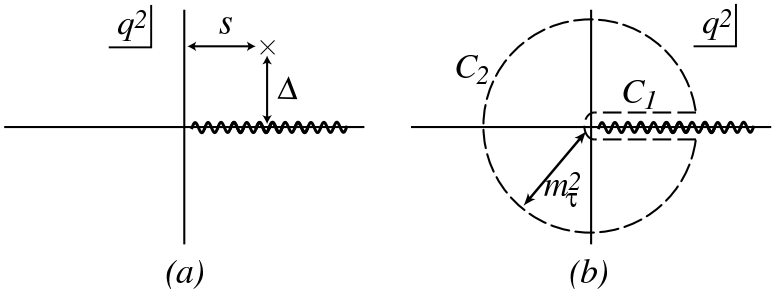

Figure 19. The analytic structure of $\Pi(q^2)$ in the q^2 plane. The wavy line denotes the branch cut due to real external states. (a) The smeared ratio for $e^+e^- \to$ hadrons is determined by $\Pi(q^2)$ in the unphysical region. (b) The phase space integral along the path C_1 for hadronic τ decay may be deformed to the path C_2, in the unphysical region.

Of course, we are interested in the physical region. However, Poggio et. al. pointed out that if we average the value of R over a region of size Δ, Cauchy's theorem allows us to relate this to $\Pi(q^2)$ in the unphysical region, where there are no hadrons:

$$\bar{R}(s,\Delta) \equiv \frac{\Delta}{\pi}\int_0^\infty \frac{R(s')}{(s'-s)^2+\Delta^2}ds' = \frac{1}{2i}\left[\Pi(s+i\Delta) - \Pi(s-i\Delta)\right]. \quad (59)$$

This shouldn't come as a surprise. For $q^2 \gg \Lambda^2_{\rm QCD}$ annihilation is a short-distance process. For a fixed physical value of q^2 the cross-section might be sensitive to long-distance physics (since q^2 could be sitting just above or below a threshold), but if we average over a number of states these effects should cancel out, leaving us with a purely short-distance process. This idea is called "global duality", and it suggests that perturbation theory should be appropriate for sufficiently smeared observables. (The definition of "sufficient" here can be problematic, but in general it means that many resonances are averaged over.)

This argument only strictly holds in the $q^2 \to \infty$ limit. To see how

corrections arise, let's study this in a bit more detail. In the physical region, $\Pi(q^2)$ is nonlocal: the intermediate hadronic state is on shell, and so can travel a long distance. On the other hand, in the unphysical region the final-state quarks, like the virtual W in four-fermi theory, are far offshell, and so can be integrated out of the EFT and replaced by a series of local operators, as illustrated in Fig. 20. This procedure is called an operator

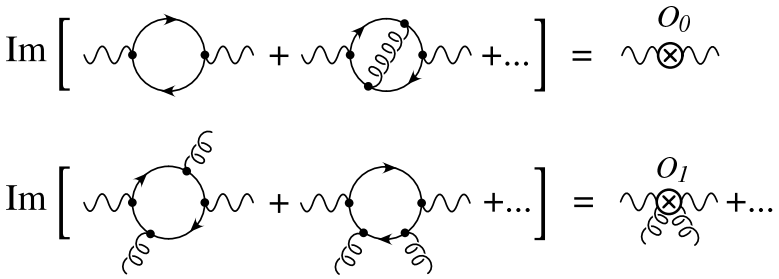

Figure 20. The operator product expansion (OPE) for $\Pi(q^2)$. The sum is over all operators with the quantum numbers of the vacuum. O_0 is the identity operator, and $O_1 = G^a_{\mu\nu}G^{\mu\nu a}$.

product expansion (OPE) - morally, it's the same as integrating heavy particles out of the EFT, but in this case the intermediate particle isn't necessarily massive, just constrained by the kinematics to be far offshell. Just as in an EFT, the matrix elements of operators of increasing dimension are suppressed by powers of $\Lambda^2_{\rm QCD}/q^2$, and so for $q^2 \ll \Lambda^2_{\rm QCD}$ we can truncate the series after the first order or two. The expression for the cross-section therefore has the form

$$R(s) \sim \sum_i C_i(\mu)\langle 0|O_i|0\rangle \tag{60}$$

where the O_i's include all operators with the quantum numbers of the vacuum, and as in an EFT, the $C_i's$ are calculated order by order in perturbation theory. The leading order operator (denoted O_0 in the figure) is the unit operator. Since its matrix element receives no radiative corrections, its coefficient is given by the imaginary part of the perturbative vacuum polarization, which is just the parton-level cross-section. Thus, as advertised, at leading order the OPE simply reproduces the parton model. The leading correction arises from the operator $O_1 = G^a_{\mu\nu}G^{\mu\nu a}$, whose coefficient is obtained by calculating the T-product of two currents in which two soft gluons are also emitted, and expanding the result in powers of the

external gluon momenta. By dimensional analysis, the matrix element of O_1 is proportional to $\Lambda^4_{\rm QCD}$, so the leading corrections to the parton model are suppressed by $\Lambda^4_{\rm QCD}/s^2$.

This technique has many applications. In addition to e^+e^- annihilation, it can also be easily extended to hadronic τ decays[28],

$$\tau \to \nu_\tau + \text{hadrons} \tag{61}$$

which has been used to extract $\alpha_s(m_\tau)$ with high precision. In this case the smearing is automatically done by the phase space integral, as illustrated in Fig. 19(b). The phase space integral of the imaginary part of $\Pi(s)$ from $s = 0$ to $s = m_\tau^2$ (weighted by the appropriate kinematic functions) is equivalent to the weighted contour integral of $\Pi(s)$ along the curve C_1. C_1 may be deformed to C_2, which lies almost entirely in the unphysical region, and so the OPE may be applied to the integrated rate.

OPE for B decays:

For B decays similar arguments hold, although in this case the initial state contains a B meson rather than the vacuum[29]. This is in fact not much of a complication, since the QCD vacuum is also nonperturbative and complicated, and in both cases the hadronic physics enters in the matrix elements of local operators. In addition, in both cases the matrix element of the leading operator is normalized, so that the leading order calculation reproduces the parton model. As with τ decays, the phase space integral over the lepton momentum in B decays performs the smearing automatically.

As in Eq. (58), the decay rate is proportional to the imaginary part of the T-product of two weak currents. We can then decompose this into different Lorentz structures:

$$\begin{aligned} T^{\mu\nu} &= -\frac{i}{2m_B}\int \langle B|T j^{\dagger\mu}_{bc}(x) j^\nu_{bc}(0)|B\rangle e^{-iq\cdot x} d^4x \\ &= -g^{\mu\nu}T_1(q^2, q\cdot v) + v^\mu v^\nu T_2(q^2, q\cdot v) - i\epsilon^{\mu\nu\alpha\beta}v_\alpha q_\beta T_3(q^2, q\cdot v) \\ &\quad + q^\mu q^\nu T_4(q^2, q\cdot v) + (v^\mu q^\nu + v^\nu q^\mu)T_5(q^2, q\cdot v). \end{aligned} \tag{62}$$

Because the heavy meson rest frame defines a vector v^μ, the T_i's are functions of q^2 and $q\cdot v$ instead of just q^2. For fixed q^2, the analytic structure of the T_i's in the $q\cdot v$ plane is shown in Fig. 21. The left-hand branch cut corresponds to physical intermediate states for the decay process, while the right-hand cut corresponds to intermediate states containing a $\bar q$ quark and two b quarks (which is related to scattering, not decay). The differential

decay rate is given by the integral over the left-hand cut,

$$\begin{aligned} d\Gamma &\sim \int_0^{m_b^2} \mathrm{Im}\, T^{\mu\nu} L_{\mu\nu}\, d(P.S.) \\ &\sim \oint_{C_1} \mathrm{Im}\, T^{\mu\nu} L_{\mu\nu}\, d(P.S.) \\ &\sim \oint_{C_2} \mathrm{Im}\, T^{\mu\nu} L_{\mu\nu}\, d(P.S.) \end{aligned} \tag{63}$$

where $(dP.S.)$ denotes the relevant phase space factors, and the curves C_1 and C_2 are shown in the figure. Because of the analytic structure of Π, we are free to deform the contour C_1 into C_2, which lies far from the resonances. We then perform an OPE for the intermediate state, as illustrated in Fig. 21. In addition to integrating out the intermediate hadronic state, in performing

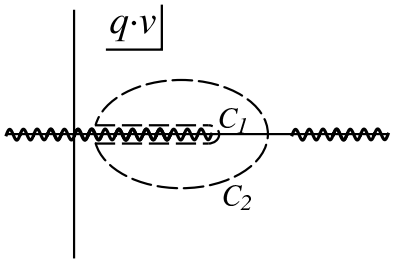

Figure 21. The analytic structure of $T^{\mu\nu}$. The phase space integral on the path C_1 may be deformed to the path C_2.

the OPE for B decay we also switch to HQET, since the matrix elements of the resulting operators are to be evaluated at a renormalization scale below m_b. Explicitly, for semileptonic $b \to u$ decay, the forward matrix element of $T^{\mu\nu}$ between free quarks with momenta $p_b = m_b v + k$ is

$$\begin{aligned} \frac{\bar{u}\gamma^\mu P_L(m_b \not{v} - \not{q} + \not{k})\gamma^\nu P_L u}{(m_b v - q + k)^2 + i\epsilon} &= (m_b v_\alpha - q_\alpha)\frac{\bar{u}\gamma^\mu\gamma^\alpha\gamma^\nu P_L u}{(m_b v - q)^2 + i\epsilon} + O(k^\mu/m_b) \\ &= P^{\mu\nu}\frac{\bar{h}_v P_L h_v}{(m_b v - q)^2 + i\epsilon} + O(k^\mu/m_b) \end{aligned} \tag{64}$$

where $P^{\mu\nu} = \frac{1}{2}(m_b v_\alpha - q_\alpha)(v^\nu g^{\alpha\mu} + v^\mu g^{\alpha\nu} - v^\alpha g^{\mu\nu} + i\epsilon^{\alpha\mu\nu\lambda} v_\lambda)$. Note that in the expansion, the external current q is taken to be of order m_b, while $k^\mu \sim \Lambda_{\rm QCD}$. Taking the imaginary part converts the denominator to a

delta function,

$$-\frac{1}{\pi}\mathrm{Im}\,\frac{1}{(m_b v-q)^2+i\epsilon}=\delta\left((m_b v-q)^2\right). \tag{65}$$

The axial piece of the matrix element of $\bar{b}_v P_L b_v$ vanishes in B-mesons, so at leading order the imaginary part of the nonlocal T-product is given by the matrix element of the local operator

$$O_0=\bar{b}_v b_v \tag{66}$$

with coefficient given by Eq. (64). This operator is a symmetry current of the theory - it just counts the number of b quarks in the state - so its matrix element is fixed by symmetry. With the conventional normalization of states,

$$\langle B|\bar{h}_v h_v|B\rangle=2m_B. \tag{67}$$

Thus, just as for e^+e^- annihilation and τ decay, the leading term in the OPE simply reproduces the free quark process, as calculated in perturbation theory.

At subleading order, the only dimension four operator is $\bar{h}_v D^\mu h_v$. By Lorentz invariance, its matrix element must be proportional to v^μ, so

$$\langle B|\bar{b}_v D^\mu b_v|B\rangle=v^\mu\langle B|\bar{b}_v v\cdot D b_v|B\rangle. \tag{68}$$

However, from Eq. (44), the leading order equation of motion in HQET[b] is $v\cdot D h_v=0$, so the matrix element vanishes at this order. Thus, there are no corrections to the parton model at $O(1/m_b)$! We will discuss higher order corrections in the next section.

Note that one immediate implication of the OPE is that inclusive decay rates for b-flavoured mesons are determined by parton-level kinematics, not hadron kinematics. In particular, the rate for $B\to X_s\gamma$ decay is therefore proportional to m_b^5, not m_B^5, since the latter would introduce a $1/m_b$ correction to the parton rate, which we have seen is absent.

4. Applications

In the last few sections I introduced the main theoretical tools that allow us to factorize long and short-distance physics. Now we proceed to some applications.

[b] An important, but sometimes unappreciated, feature of the equations of motion is that they may always be applied at the operator level, even though the quarks in hadronic matrix elements are not on shell. See Ref. [30].

4.1. *Inclusive Decays: Corrections to the parton model*

We saw in the previous section that up to $O(1/m_b)$, inclusive B decay rates are the same as that of free b quarks to quarks and gluons. To obtain higher accuracy, we must include in the OPE operators of dimension five, of which there are only two. Their matrix elements are conventionally denoted as

$$\begin{aligned}\lambda_1 &= \frac{1}{2m_B}\langle B|\bar{b}_v(iD)^2 b_v|B\rangle \\ \lambda_2 &= \frac{1}{6m_B}\langle B|g\bar{b}_v\sigma_{\mu\nu}G^{\mu\nu}b_v|B\rangle.\end{aligned} \tag{69}$$

(In another convention, these matrix elements are denoted $\mu_\pi^2 = -\lambda_1 + \ldots$ and $\mu_G^2 = 3\lambda_2 + \ldots$, where the ellipses denote terms suppressed by powers of $1/m_b$.) These are the same operators which appear in the HQET Lagrangian Eq. (47), and correspond to the kinetic energy of the heavy quark in the meson, and the chromomagnetic energy. Since the latter is the leading operator which violates spin symmetry, it may be determined from the experimentally measured $B - B^*$ mass splitting,

$$\lambda_2 = \frac{1}{4}(m_{B^*}^2 - m_B^2) + O(\Lambda_{\rm QCD}/m_b) \simeq 0.12\,{\rm GeV}^2 \tag{70}$$

while λ_1 cannot be simply related to meson masses. Thus, up to $O(\Lambda_{\rm QCD}^2/m_b^2)$, inclusive quantities may be expressed in terms of only two unknown parameters, m_b and λ_1. From the HQET Lagrangian (47), the meson and quark masses are related via

$$m_B = m_b + \bar{\Lambda} - \frac{\lambda_1 + 3\lambda_2}{2m_b^2} + \ldots \tag{71}$$

so m_b may be exchanged for $\bar{\Lambda}$, the energy of the light degrees of freedom in the meson.

At the price of introducing even more parameters, one may continue to expand to higher orders in $1/m_b$. At $O(1/m_b^3)$ there are two more operators, whose matrix elements are defined as

$$\begin{aligned}\rho_1 &\equiv \frac{1}{2m_B}\langle B|\bar{b}_v(iD_\mu)(iv\cdot D)(iD^\mu)b_v|B\rangle \\ \rho_2 &\equiv -\frac{i}{6m_B}\epsilon^{\alpha\mu\beta\delta}\langle B|\bar{b}_v(iD_\alpha)(iD_\mu)(iD_\beta)\gamma_\delta\gamma_5 b_v|B\rangle.\end{aligned} \tag{72}$$

In addition, there are the T-products of $1/m_b$ corrections to the HQET Lagrangian with the operators (69), for a total of six unknown quantities. For most quantities, expressions are available up to $O(1/m_b^3)$, although uncertainties at this order mean that the $1/m_b^3$ terms are typically used to estimate theoretical uncertainties.

4.1.1. *Total decay widths:*

For example, for $b \to s\gamma$ we have, at tree level[31],

$$\Gamma(B \to X_s\gamma) = \frac{\alpha G_F^2 m_b^5}{32\pi^4}|V_{tb}V_{ts}|^2|C_7(m_b)|^2 \left[1 + \frac{\lambda_1}{2m_b^2} - \frac{9\lambda_2}{2m_b^2} + \dots\right] . \tag{73}$$

The λ_2 term, which is known, reduces the rate by about 2%, while the λ_1 term is less than a 1% effect for reasonable values of λ_1. The $1/m_b$ corrections are therefore very small for this decay, and do not appreciably shift the prediction (41).

The semileptonic $B \to X_c\ell\nu$ decay width is proportional to the CKM matrix element V_{cb}. The result is[29]

$$\Gamma(B \to X_c e\bar{\nu}) = \Gamma_0 \left[f_0(\hat{m}_q) + \frac{1}{2m_b^2} f_1(\hat{m}_q, \lambda_1, \lambda_2) + A_0(\hat{m}_q)\,\frac{\alpha_s}{\pi}\right] , \tag{74}$$

where

$$\begin{aligned}
\Gamma_0 &= \frac{G_F^2 m_b^5 |V_{cb}|^2}{192\pi^3} \\
f_0(\hat{m}_q) &= 1 - 8\hat{m}_q^2 + 8\hat{m}_q^6 - \hat{m}_q^8 - 24\hat{m}_q^4 \ln \hat{m}_q \,, \\
f_1(\hat{m}_q, \lambda_1, \lambda_2) &= \lambda_1 \left(1 - 8\hat{m}_q^2 + 8\hat{m}_q^6 - \hat{m}_q^8 - 24\hat{m}_q^4 \ln \hat{m}_q\right) \\
&\quad + \lambda_2 \left(-9 + 24\hat{m}_q^2 - 72\hat{m}_q^4 + 72\hat{m}_q^6 - 15\hat{m}_q^8 - 72\hat{m}_q^4 \ln \hat{m}_q\right)
\end{aligned} \tag{75}$$

and A_0 contains the one-loop radiative corrections[32]. Hence, with a precise determination of $\bar{\Lambda}$ and λ_1, a precise determination of V_{cb} is possible. Finally, setting $m_c = 0$ in this expression, we obtain the rate for $B \to X_u\ell\nu$,

$$\Gamma(B \to X_u e\bar{\nu}) = \frac{G_F^2 m_b^5 |V_{ub}|^2}{192\pi^3} \left[1 + \frac{\lambda_1 - 9\lambda_2}{2m_b^2} + \left(\frac{25}{6} - \frac{2}{3}\pi^2\right)\frac{\alpha_s}{\pi} + \dots\right] . \tag{76}$$

4.1.2. *Differential widths and spectra:*

We can construct an OPE not only for total decay rates, but also for differential rates and therefore spectral moments. There has been much recent interest in such quantities because they allow the nonperturbative parameters $\bar{\Lambda}$, λ_1 (and, in principle, higher order terms) to be determined experimentally. This not only allows the rates in the previous section to be predicted with high accuracy, but consistency between different observables provides a stringent test of the whole OPE picture for inclusive decays. Some quantities of particular recent interest include the first moment of the photon energy spectrum in $B \to X_s\gamma$ decay[31,33], the first moment of the hadronic

energy spectrum in $B \to X_c \ell \bar{\nu}$ decay[34], and various weighted spectral moments of the charged lepton spectrum in $B \to X_c \ell \bar{\nu}$ decay[35]. Working to $O(1/m_b^2)$, each of these observables constrains a linear combination of $\bar{\Lambda}$ and λ_1. For example, the results of several recent CLEO analyses are shown in Fig. 22. This analysis obtained[36]

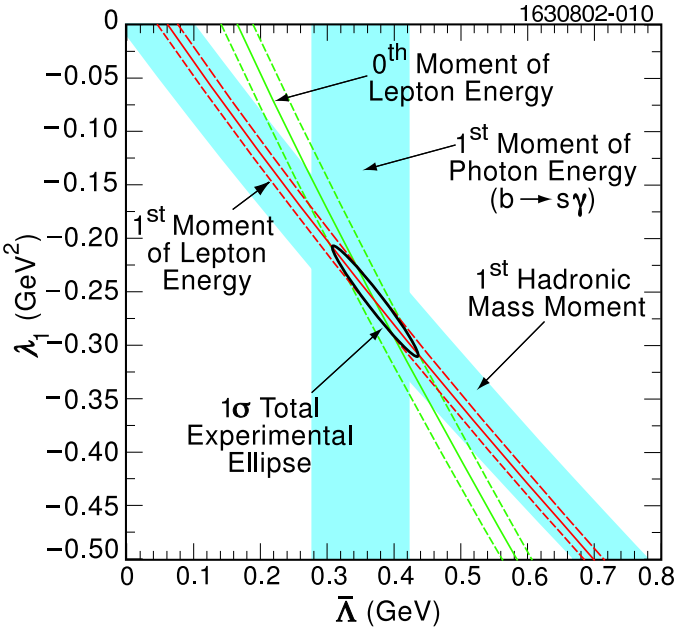

Figure 22. Experimental constraints on $\bar{\Lambda}$ and λ_1, from [36]. Note the consistency between the different measurements.

$$m_b^{(1S)} = 4.82 \pm 0.07_E \pm 0.11_T \,\mathrm{GeV},$$
$$\lambda_1 = -0.25 \pm 0.02_{ST} \pm 0.05_{SY} \pm 0.14_T \,\mathrm{GeV}^2 \qquad (77)$$

where the "1S" mass $m_b^{(1S)}$ is an example of a short-distance b quark mass, which is preferred for practical purposes to $\bar{\Lambda}$ because of the poor behaviour of perturbation theory when the quark pole mass is used in calculations.[c] Global fits to a variety of spectral moments have also been performed by various groups, using expressions up to $O(1/m_b^3)$, which give comparable results[38].

[c]This is a fascinating side issue which I have no time to discuss here. See Ref. [37] for further discussion.

Combining these values with the experimental measurement of the semileptonic $B \to X_c$ width gives[36]

$$|V_{cb}| = (40.8 \pm 0.5 \pm 0.4 \pm 0.9) \times 10^{-3} \tag{78}$$

where the first error is from the uncertainty in the total semileptonic width, the second from m_b and λ_1, and the third an estimate of higher order corrections.

4.1.3. *V_{ub} and the shape function*

In principle, one could use Eq. (76) and the values of m_b and λ_1 from the moment fits to determine V_{ub} from the total inclusive semileptonic $b \to u$ width, with a theoretical uncertainty at the $\sim 5\%$ level. However, because of the $\sim 100\times$ background from $b \to c$ decays, the inclusive $b \to u$ rate can only be measured in restricted regions of phase space where the charm background is absent. In these restricted regions, the usual OPE breaks down.

This is easiest to see by instead considering the analogous situation in inclusive $b \to X_s\gamma$ decay. Suppose we were interested not in the total rate but instead in the differential rate $d\Gamma/dE_\gamma$ This is in fact an important experimental quantity, since one has to impose a hard cut on the photon energy $E_\gamma \gtrsim 2.2$ GeV to eliminate background from $b \to c$ decay, and it is important to know how much of the $b \to s\gamma$ this cuts out. Unfortunately, the OPE gives us little information. At tree level the decay is purely two-body, so the photon energy spectrum is just a delta function at $E_\gamma = m_b/2$. At higher orders in the OPE it doesn't look much better - higher order terms are proportional to derivatives of delta functions at the same point, so this spectrum makes no sense unless we integrate it over some region of E_γ. Radiative corrections populate the region $E < m_b/2$ because of gluon bremmstrahlung, but the resulting theoretical spectrum does not have peaks due to the various hadronic resonances. This is an example of quark-hadron duality failing if the rate is not sufficiently inclusive. But how large a region must we integrate over?

The OPE tells us. The denominator T-produce of the currents for $b \to X_s\gamma$ is

$$D = \frac{m_b^2}{(m_b v + k - q)^2 + i\epsilon} = \frac{1}{(m_b v - q)^2 + 2k \cdot (m_b v - q) + k^2 + i\epsilon} \tag{79}$$

where k^μ is the residual momentum of the b meson, defined in Eq. (42). Writing $q^\mu = x m_b \bar{n}^\mu/2$, where $\bar{n}^\mu$ is a light-like vector, and $x = 1$ corre-

sponds to the maximum photon energy at the parton level, this becomes

$$D = \frac{1}{1 - x + \hat{k} \cdot n + (1 - x)\hat{k} \cdot \bar{n} + \hat{k}^2 + i\epsilon} \tag{80}$$

(where $\hat{k} \equiv k/m_b$). There are three possible regions:

(1) $1 - x \gg \Lambda_{\rm QCD}/m_b$: This is the usual assumption for the OPE. In this region, the denominator in (80) may be expanded in powers of k^μ/m_b, and the result matched onto local operators.
(2) $1 - x \lesssim \Lambda_{\rm QCD}^2/m_b^2$: In this region, there is no sensible $1/m_b$ expansion of the denominator, and the OPE breaks down. This is physically reasonable: in this region, the invariant mass of the final hadronic state is constrained to be of order $\Lambda_{\rm QCD}$, and so the decay is dominated by a small number of resonances, and one would expect an inclusive description to fail. Indeed, in this region the leading term of the OPE is the same size as all the other terms in the denominator, and so the expansion in powers of $1/m_b$ doesn't converge.
(3) $1 - x \sim \Lambda_{\rm QCD}/m_b$: In this region, the maximum invariant mass of the final state is of order $(\Lambda_{\rm QCD} m_b)^{1/2} \gg \Lambda_{\rm QCD}$, and one would expect an inclusive description to be valid. However, in this region $\hat{k}\cdot n$ and $1-x$ are both of order $\Lambda_{\rm QCD}$, and one cannot treat the $\hat{k}\cdot n$ term as a perturbation. This is often known as the shape function region.[39]

In the shape function region, the usual OPE fails, but we may still expand the denominator in powers of $\Lambda_{\rm QCD}/m_b$:

$$D = \frac{1}{1 - x + \hat{k} \cdot n + i\epsilon} + \cdots \Rightarrow \mathrm{Im}\, D = -\pi\delta(1 - x + \hat{k} \cdot n) + \ldots \tag{81}$$

A delta function containing the residual momentum is not related to the matrix element of a local operator; instead, it is given by the function

$$f(\omega) \equiv \langle B|\bar{b}\,\delta(\omega + i\hat{D} \cdot n)b|B\rangle. \tag{82}$$

This is a strange-looking object, but it looks less peculiar if we perform a Fourier transform in ω and look at it in position space,

$$\tilde{f}(t) \equiv \bar{B}|\bar{b}(0)P \exp\left[\int_0^t n \cdot A(t')dt'\right] b(t)|B\rangle \tag{83}$$

where the P denotes path-ordering, and t measures the distance along the light-like direction n^μ: $b(t) \equiv b(tn/m_b)$. This is a nonlocal operator where

the b and $\bar{b}$ fields are separated by a light-like displacement, and the path-ordered exponential is there to restore gauge invariance. (This is technically known as a light-like Wilson line.) In $n\cdot A = 0$ gauge, the function $f(\omega)$ gives the amplitude for the light-cone component of the residual momentum k^μ to be ω. $f(\omega)$ is therefore the light-cone distribution function of the heavy quark residual momentum; it is often referred to as the "shape function". (The shape function was actually originally defined as an integral of $f(\omega)$, but this nomenclature seems to have stuck.)

Thus, at leading order in $\Lambda_{\rm QCD}/m_b$ the nonperturbative contribution to the photon spectrum in $b \to X_s\gamma$ is given, in the region $1 - x \sim \Lambda_{\rm QCD}/m_b$, is determined by the light-cone distribution of the heavy quark residual momentum.[39]

At this stage we don't seem to have gained anything. However, the function $f(\omega)$ is universal. Consider, for example, the shape of the hadronic invariant mass spectrum in $b \to X_u\ell\bar{\nu}$ decay. In this case,

$$D = \frac{m_b^2}{\hat{s}_0 + 2\hat{k}\cdot(v - \hat{q}) + \hat{k}^2 + i\epsilon} \tag{84}$$

where q^μ is the momentum transfer to the leptons and $\hat{s}_0 \equiv (v - \hat{q})^2$ is the parton-level hadronic invariant mass.[d] If $\hat{s}_0 \ll 1$, then $m_b v - q$ is almost a lightlike vector, so we write $m_b v - q = m_b/2\bar{n}^\mu + \Delta^\mu$. Thus, when $\Delta^\mu \sim \Lambda_{\rm QCD}$, we have $s_0 \sim \bar{n}\cdot\Delta \sim O(\Lambda_{\rm QCD} m_b)$, and

$$D \sim \frac{1}{\hat{s}_0 + \hat{k}\cdot n + i\epsilon} + \cdots \Rightarrow \mathrm{Im}\, D = -\pi\delta(\hat{s}_0 + \hat{k}\cdot n) + \ldots \tag{85}$$

Comparing this with Eq. (81), we see that the hadronic invariant mass spectrum is determined in the region $s_0 \lesssim \Lambda_{\rm QCD} m_b$ by the same function that determines the shape of the photon spectrum near the endpoint of the $b \to X_s\gamma$. A similar argument shows that the rate for $B \to X_u\ell\bar{\nu}$ in the region $E_\ell > m_b - \Lambda_{\rm QCD}$ is given by the same distribution function.

More precisely, at leading order in $1/m_b$ the various spectra are deter-

[d] This differs from the physical hadronic invariant mass $(m_B v - q)^2$; however, we won't worry about that complication here.

mined by convoluting $f(\omega)$ with the appropriate kinematic functions:

$$\begin{aligned}
\frac{1}{\Gamma}\frac{d\Gamma}{d\hat{E}_\gamma}(B \to X_s\gamma) &= 2f(1-2\hat{E}_\gamma) + \dots \\
\frac{1}{\Gamma}\frac{d\Gamma}{d\hat{E}_\ell}(B \to X_u\ell\bar{\nu}) &= 2\int \theta(1-2\hat{E}_\ell-\omega)f(\omega)\,d\omega + \dots \\
\frac{1}{\Gamma}\frac{d\Gamma}{d\hat{s}_H}(B \to X_u\ell\bar{\nu}) &= \int \frac{2\hat{s}_H^2(3\omega-2\hat{s}_H)}{\omega^4}\theta(\omega-\hat{s}_H)f(\omega-\hat{\Lambda})\,d\omega + \dots
\end{aligned} \tag{86}$$

This is useful phenomenologically for determining the CKM mixing angle $|V_{ub}|$. As mentioned in Sec. 4.2.1, the semileptonic width for $b \to u$ decay is only easily experimentally accessible in restricted regions of phase space, such as $E_\ell > (m_B^2 - m_D^2)/2m_B$, or $s_H < m_D^2$[40], where background from $b \to c$ decay is absent. Because $m_c^2/m_b^2 \simeq \Lambda_{\rm QCD}/m_b$, this corresponds in both cases to the shape function region. The shape function $f(\omega)$ may then be measured in $B \to X_s\gamma$ decay and then used to predict either the charged lepton or hadronic invariant mass spectrum in semileptonic $B \to X_u$ decay to obtain a theoretically clean measurement of $|V_{ub}|$.

At subleading order in $1/m_b$ the universality is broken, and new subleading shape functions are required, introducing theoretical uncertainty into the extraction of $|V_{ub}|$. The expansion has been carried out systematically to $O(1/m_b)$, while large $O(1/m_b^2)$ effects due to annihilation graphs has also been considered[41].

4.2. *Exclusive Decays and Symmetries*

4.2.1. $B \to D^*\ell\bar{\nu}$ *and* V_{cb}

In Sec. 3.3 I discussed the fact that HQET has a spin-flavour symmetry which is not manifest in QCD. We made us of this in Sec. 3.4, where we used the fact that the operator $\bar{b}_v b_v$ is a generator of that symmetry, and therefore its matrix element is fixed (matrix elements of symmetry generators are related to conserved charges). However, heavy quark symmetry gives us much more information. Any operator of the form

$$\bar{h}_{i,v}\Gamma h_{j,v} \tag{87}$$

for any Dirac structure Γ and heavy flavours i and j is a generator of the symmetry, and so has an absolutely normalized matrix element. Physically, this just corresponds to the fact that exchanging the heavy quark in a meson with another heavy quark of arbitrary spin has no effect on the light degrees of freedom in the meson. This is true because, as we have discussed, in the

$m_Q \to \infty$ limit a heavy quark simply behaves as a static source of colour charge, and the so the light degrees of freedom are insensitive to the mass and spin of the quark.

The most important phenomenological application of this symmetry is to semileptonic $B \to D/D^* \ell \bar{\nu}$ decays. Isgur and Wise[22] showed that heavy quark symmetry implies that all matrix elements of the form $\langle H_i^{(*)}(v')|\bar{h}_{i,v'} \Gamma h_{j,v}|H_j(*)(v)\rangle$ (where $H_i(*)$ denotes either the pseudoscalar (D or B) or vector (D^* or B^*) meson with a heavy quark of flavour $i = c,\ b$) could be related to a single universal function:

$$\langle H_i^{(*)}(v')|\bar{h}_{i,v'} \Gamma h_{j,v}|H_j^{(*)}(v)\rangle \sim \xi(v \cdot v'). \tag{88}$$

$\xi(v \cdot v')$ has been dubbed the "Isgur-Wise function". It has the additional property that, because of the absolute normalization discussed in the previous paragraph,

$$\xi(1) = 1. \tag{89}$$

This means that all form factors for semileptonic $B \to D$ and $B \to D^{(*)}$ decay are known at the kinematic point where the c quark moves with the same velocity as the c quark (i.e., the q^2 transfer to the leptons is maximal.)

We can use the EFT machinery to calculate the corrections to this prediction. Quite generally, we can write the differential decay rate

$$\frac{d\Gamma}{dw}(\bar{B} \to D^{(*)} e \bar{\nu}_e) = \frac{G_F^2}{48\pi^3}|V_{bc}|^2 (m_B - m_{D^*})^2 m_{D^*}^3 (w+1)^3 \sqrt{w^2-1} \times \left[1 + \frac{4w}{w+1}\frac{m_B^2 - 2w m_B m_{D^*} + m_{D^*}^2}{(m_B - m_{D^*})^2}\right] |F(w)|^2 \tag{90}$$

where

$$F(w) = \xi(w) + O(\alpha_s) + O(1/m_{c,b}) \tag{91}$$

and $w \equiv v \cdot v'$. The complicated w dependence just comes from the phase space integrals.

Radiative Corrections: These arise from perturbative corrections to matching conditions, as well as running of the $\bar{b}\Gamma c$ current below $\mu = m_b$, and so are completely calculable. The one-loop running calculation has already been discussed in Sec. 3.3. From our previous results, then, we get the leading perturbative correction to Eq. (90),

$$\bar{c}\gamma^\mu(1-\gamma_5)b \to \left(\frac{\alpha_s(m_b)}{\alpha_s(m_c)}\right)^{6/25} \bar{c}_v \gamma^\mu (1-\gamma_5) b_v \tag{92}$$

and so

$$F(1) = \left(\frac{\alpha_s(m_b)}{\alpha_s(m_c)}\right)^{6/25} + O(\alpha_s^{n+1}\log^n m_c/m_b) + O(1/m_{b,c}). \qquad (93)$$

At next order, the current must be matched to $O(\alpha_s)$ accuracy, and the running performed to two loops (resumming all terms of order $\alpha_s^{n+1}\log^n m_c/m_b$). This calculation has been done[42]. For practical purposes, however, $\log m_c/m_b$ is not very large, and so it is not necessary to sum terms of order $\alpha_s^n \log m_c/m_b$ to all orders. Instead, one typically uses a fixed order calculation. This therefore includes terms of order α_s^2 which are not enhanced by powers of $\log m_c/m_b$, but neglects terms of $O(\alpha_s^3)$ and higher.

Power Corrections: Just as the Lagrangian has corrections corresponding to higher dimension operators, so the weak current matches in the EFT onto additional operators. For general v and v', the only operator at $O(1/m_c)$ is

$$\bar{c}_{v'} i\overleftarrow{\not{D}}\gamma^\mu(1-\gamma_5)b_v \qquad (94)$$

and so at $O(1/m_c)$ we need the matrix element

$$\langle D^{(*)}(v')|\bar{c}_{v'} i\overleftarrow{\not{D}}\gamma^\mu(1-\gamma_5)b_v|B(v)\rangle \qquad (95)$$

There are three possible form factors required to describe the matrix elements of this operator in HQET. Furthermore, additional $1/m_Q$ corrections arise due to T-products of $1/m_Q$ operators in the effective Lagrangian and the leading order current,

$$\langle D^{(*)}(v')|T(\bar{c}_{v'}\gamma^\mu b_v, \mathfrak{L}_1)|B(v)\rangle. \qquad (96)$$

However, at zero recoil ($v = v'$) things simplify dramatically: both of these corrections *vanish*, although each for a different reason.

- The T-product in (96) vanishes by the Ademollo-Gatto theorem when $v = v'$. The theorem states that the symmetry breaking corrections to matrix elements of an approximate symmetry current are second order in the breaking terms in the Lagrangian.
- The matrix element (95) can be related, when $v = v'$, to the matrix element of the operators $\bar{c}_v \overleftarrow{D}\cdot v b_v$, which vanishes by the equations of motion.

Thus, we have the nice result that the leading nonperturbative effects to the absolute normalization of the Isgur-Wise function at zero recoil are of order $\Lambda_{\rm QCD}/m_c^2$, not $\Lambda_{\rm QCD}/m_c$.[25] This immediately raises this prediction to a reasonable level of precision, since corrections of order $\Lambda_{\rm QCD}^2/m_c^2$ are expected to be of order 5%.

Putting everything together, we find

$$\begin{aligned} F(1) &= 1 + \eta_A^{(1)} \frac{\alpha_s(m_b)}{\pi} + \eta_A^{(2)} \left(\frac{\alpha_s(m_b)}{\pi} \right)^2 + \delta_{1/m^2} + O\left(\alpha_s^3, 1/m_{b,c}^3\right) \\ &= 0.960 \pm 0.007 + \delta_{1/m^2} + O\left(\alpha_s^3, 1/m_{b,c}^3\right) \end{aligned} \tag{97}$$

where $\eta_A^{(1)}$ and $\eta_A^{(2)}$ are perturbatively calculable, and δ_{1/m^2} refers to the incalculable $O(1/m_{c,b}^2)$ corrections. There have been a number of attempts to estimate the size of these corrections from quark models, sum rules and other methods. Combining the results in the literature gives the rough estimate[43]

$$\delta_{1/m^2} = -5.5 \pm 3\% \tag{98}$$

which gives the result

$$F(1) = 0.91 \pm 0.04. \tag{99}$$

Combining this with the world averaged measured value of $d\Gamma/dw$ at the endpoint gives the value[44]

$$|V_{cb}| = (41.9 \pm 1.1_{\rm expt} \pm 1.8_{\rm theory}) \times 10^{-3}. \tag{100}$$

Note that this is not as precise as the result from inclusive decays, Eq. (78). However, it is based on a completely different approach, both theoretically and experimentally, and so the agreement between the two results within errors is a nontrivial check on the validity of the $1/m_b$ expansion.

4.2.2. $B \to J/\psi K_s$ *and* $\sin 2\beta$

I have been focusing in these lectures on the heavy quark expansion and its application to decay rates, etc. in CP conserving quantities, but I must say something about CP violation. The B factories were designed, after all, to study the SM picture of CP violation and while I cannot give this subject its due, I would like to at least mention it[45]. It also provides a nice illustration of another situation where we can use symmetry considerations to make model-independent predictions for hadronic processes. In this case, the relevant symmetry of QCD is CP.

Consider charged B decay to some final state f. The amplitude for this decay may be written as

$$A_f = \langle F|H|B^+\rangle = \sum_k A_k e^{i\delta_k} e^{i\varphi_k}. \tag{101}$$

The sum over k corresponds to the sum over different amplitudes (or operators) which contribute to the decay. In general, such amplitudes are complex, since there may be real intermediate states, as shown in Fig. ... . The phase φ_k is the weak phase which is present in the weak Hamiltonian itself.

Similarly, the amplitude for the CP-conjugate process $B^- \to \bar{f}$ may be written

$$\bar{A}_{\bar{f}} = \langle \bar{f}|H|B^-\rangle = \sum_k A_k e^{i\delta_k} e^{-i\varphi_k}. \tag{102}$$

The important thing is that, because QCD is CP-conserving, the strong amplitude is unchanged in the CP-conjugate process. The only change in the amplitude arises because the CP-violating weak phase flips sign.

This immediately can give rise to CP-violation in the rates, since the different amplitudes can interfere, and so we can have $|A_f|^2 \neq |\bar{A}^2_{\bar{f}}$. The problem is that in general, this interference depends on the values of the strong amplitudes $A_k e^{i\delta k}$, which in turn depend on long-distance QCD. For most decays, then, it is difficult to translate measured CP violation into a determination of the weak phases φ_k.

However, if we are clever, we can find processes for which the strong phase drops out. A famous example is the decay $B \to J/\psi K_s$, which proceeds dominantly via the $b \to c\bar{c}s$ process. (There is also a "penguin" diagram which contributes to the decay, but it is proportional to the same weak phase, so the argument still holds.) Since it is dominated by one amplitude, it does not appear that we will get any CP violation, since CP violation requires the interference of at least two amplitudes. However, there is another amplitude hidden in the process: the B^0 can oscillate into a $\bar{B}^0$, which can then decay, and this mixing carries a relative phase (due to V_{td} in the Wolfenstein parametrization) which can interfere with the decay. Thus, we really do have two interfering amplitudes, and the relative phase between them is β.

More precisely it is easy to show that

$$\frac{\Gamma(\bar{B}^0(t) \to \psi K_S) - \Gamma(B^0(t) \to \psi K_s)}{\Gamma(\bar{B}^0(t) \to \psi K_S) + \Gamma(B^0(t) \to \psi K_s)} = -\sin 2\beta \sin \Delta m t + O(\lambda^2) \tag{103}$$

and thus this measurement provides a clean measurement of $\sin 2\beta$. The current world average, from Belle and Babar, is[15]

$$\sin 2\beta = 0.79 \pm 0.14. \tag{104}$$

This is shown in Fig. 8 along with the other experimental constraints, and is clearly in good agreement with the SM expectation.

A similar analysis holds for the rarer decay $B \to \phi K_s$, which proceeds primarily via the penguin decay $b \to s\bar{s}s$. Since in the SM in the Wolfenstein parameterization, neither $b \to c\bar{c}s$ nor $b \to s\bar{s}s$ contains a weak phase, both measurements measure the phase β in mixing. However, since the second process occurs at the loop level in the SM, it is easy to cook up models of new physics which contribute a large phase to the $b \to s\bar{s}s$ amplitude, so consistency of the measured β from the two processes provides an important test of the SM. Indeed, there has been recent excitement over a possible discrepancy between these measurements[46]:

$$\begin{aligned} \sin 2\beta(B \to \phi K_s) &= -0.96^{\pm} 0.50^{+0.09}_{-0.11} \text{ (BELLE)} \\ \sin 2\beta(B \to \phi K_s) &= -0.19^{+0.52}_{-0.50} \pm 0.09 \text{ (BABAR)} \end{aligned} \tag{105}$$

However, the experimental errors are high enough that the discrepancy with Eq. (104) cannot be considered more than a tantalizing hint at present.

Unfortunately, it is not so easy to come up with decays which are dominated by a single amplitude like this. For example, we could try a similar trick with $b \to u\bar{u}s$ to try to predict CP violation in $B \to \pi\pi$. In this case, the relative phase between the amplitudes is $\gamma + \beta = \pi - \alpha$, so one could hope to use this decay to determine $\sin 2\alpha$. However, as shown in Fig. 6(c), this process receives contributions from both the tree-level $b \to u\bar{u}s$ amplitude and a penguin amplitude. The latter is a one-loop process, but since it is a strong loop there is no reason for it to be suppressed, and it contributes with different weak and strong phases. Thus, without knowing the relative amplitudes between the tree and penguin processes, we can no longer make a model-independent determination of α in this manner, and more complicated and/or less theoretically rigorous approaches must be used.

5. Conclusions

In these lectures I have discussed some of the most important applications of the heavy quark expansion to B decays. This is a large and rapidly evolving field. The applications I have discussed are mostly "classic", in that the theory was developed a number of years ago and is well understood.

Recently, there has been much work attempting to further exploit (generalized) factorization in B decays, and to make use of the resulting simplifications, in particular to better understand nonleptonic decays. As an important example, several years ago, Beneke, Buchalla, Neubert and Sachrajda[47] proposed a generalized form of factorization applicable to two-body nonleptonic decays. Within their framework, they showed that so-called naïve factorization, in which matrix elements of two currents are written as the product of matrix elements of the individual currents, arises model-independently from QCD leading order in $1/m_b$ and α_s. This claim has generated much interest in the field, both phenomenological and theoretical, as the validity of the arguments has been heatedly debated. More recently, an EFT description which encompasses the kinematics of these decays, dubbed Soft-Collinear Effective Theory (SCET) has been developed[48]. Since the kinematics are much more complicated than in HQET - with hard, soft and collinear momenta possible - the EFT is more complicated, and (I believe) not completely understood. However, much progress has been made in understanding the interplay of the various scales, and this is an active and rapidly-evolving field.

References

This list of references is by no means exhaustive; I have only referenced the original literature for specific points made in the text, along with some useful reviews on the subject.

1. M. K. Gaillard and B. W. Lee, Phys. Rev. D **10**, 897 (1974).
2. H. Albrecht *et al.* [ARGUS COLLABORATION Collaboration], Phys. Lett. B **192**, 245 (1987); C. Albajar *et al.* [UA1 Collaboration], Phys. Lett. B **186**, 247 (1987) [Erratum-ibid. **197B**, 565 (1987)].
3. L. Montanet *et al.* [Particle Data Group Collaboration], Phys. Rev. D **50**, 1173 (1994).
4. H. N. Brown *et al.* [Muon g-2 Collaboration], Phys. Rev. Lett. **86**, 2227 (2001).
5. M. Knecht and A. Nyffeler, Phys. Rev. D **65**, 073034 (2002); M. Knecht, A. Nyffeler, M. Perrottet and E. De Rafael, Phys. Rev. Lett. **88**, 071802 (2002).
6. S. Narison, Phys. Lett. B **513**, 53 (2001) [Erratum-ibid. B **526**, 414 (2002)] [arXiv:hep-ph/0103199].
7. K. Melnikov, Int. J. Mod. Phys. A **16**, 4591 (2001) [arXiv:hep-ph/0105267].
8. M. Ramsey-Musolf and M. B. Wise, Phys. Rev. Lett. **89**, 041601 (2002).
9. L. Wolfenstein, Phys. Rev. Lett. **51**, 1945 (1983).
10. S. L. Glashow, J. Iliopoulos and L. Maiani, Phys. Rev. D **2**, 1285 (1970).
11. see, for example, G. Buchalla and A. J. Buras, Phys. Rev. D **54**, 6782 (1996).

12. A. D. Sakharov, Pisma Zh. Eksp. Teor. Fiz. **5**, 32 (1967) [JETP Lett. **5**, 24 (1967 SOPUA,34,392-393.1991 UFNAA,161,61-64.1991)].
13. M. A. Shifman, A. I. Vainshtein and V. I. Zakharov, Phys. Rev. D **18**, 2583 (1978) [Erratum-ibid. D **19**, 2815 (1979)]; B. A. Campbell and P. J. O'Donnell, Phys. Rev. D **25**, 1989 (1982); S. Bertolini, F. Borzumati and A. Masiero, Phys. Rev. Lett. **59**, 180 (1987); N. G. Deshpande, P. Lo, J. Trampetic, G. Eilam and P. Singer, Phys. Rev. Lett. **59**, 183 (1987).
14. T. Inami and C. S. Lim, Prog. Theor. Phys. **65**, 297 (1981) [Erratum-ibid. **65**, 1772 (1981)].
15. K. Hagiwara *et al.* [Particle Data Group Collaboration], Phys. Rev. D **66**, 010001 (2002).
16. I. Z. Rothstein, "TASI lectures on effective field theories," arXiv:hep-ph/0308266 (in these proceedings.)
17. For other reviews of effective field theory, see D. B. Kaplan, arXiv:nucl-th/9506035; A. V. Manohar, arXiv:hep-ph/9606222.
18. B. Grinstein, R. P. Springer and M. B. Wise, Phys. Lett. B **202**, 138 (1988); Nucl. Phys. B **339**, 269 (1990).
19. M. Misiak, Phys. Lett. B **269**, 161 (1991); Nucl. Phys. B **393**, 23 (1993) [Erratum-ibid. B **439**, 461 (1995)].
20. see A. J. Buras, A. Czarnecki, M. Misiak and J. Urban, Nucl. Phys. B **631**, 219 (2002) and references therein.
21. M. A. Shifman and M. B. Voloshin, Sov. J. Nucl. Phys. **45**, 292 (1987) [Yad. Fiz. **45**, 463 (1987)]; Sov. J. Nucl. Phys. **47**, 511 (1988) [Yad. Fiz. **47**, 801 (1988)]. H. D. Politzer and M. B. Wise, Phys. Lett. B **208**, 504 (1988); "Leading Logarithms Of Heavy Quark Masses In Processes With Light And Heavy Phys. Lett. B **206**, 681 (1988); E. Eichten and B. Hill, Phys. Lett. B **234**, 511 (1990); H. Georgi, Phys. Lett. B **240**, 447 (1990).
22. N. Isgur and M. B. Wise, Phys. Lett. B **232**, 113 (1989); Phys. Lett. B **237**, 527 (1990).
23. For reviews, see A. V. Manohar and M. B. Wise, Cambridge Monogr. Part. Phys. Nucl. Phys. Cosmol. **10**, 1 (2000); M. Neubert, Phys. Rept. **245**, 259 (1994); M. A. Shifman, arXiv:hep-ph/9510377; A. F. Falk, arXiv:hep-ph/0007339.
24. E. Eichten and B. Hill, Phys. Lett. B **243**, 427 (1990); A. F. Falk, B. Grinstein and M. E. Luke, Nucl. Phys. B **357**, 185 (1991).
25. M. E. Luke, Phys. Lett. B **252**, 447 (1990).
26. A. F. Falk, H. Georgi, B. Grinstein and M. B. Wise, Nucl. Phys. B **343**, 1 (1990).
27. E. C. Poggio, H. R. Quinn and S. Weinberg, Phys. Rev. D **13**, 1958 (1976).
28. E. Braaten, S. Narison and A. Pich, Nucl. Phys. B **373**, 581 (1992).
29. M. A. Shifman and M. B. Voloshin, Sov. J. Nucl. Phys. **41**, 120 (1985) [Yad. Fiz. **41**, 187 (1985)]; J. Chay, H. Georgi and B. Grinstein, Phys. Lett. B **247**, 399 (1990); I. I. Y. Bigi, N. G. Uraltsev and A. I. Vainshtein, Phys. Lett. B **293**, 430 (1992) [Erratum-ibid. B **297**, 477 (1993)]; I. I. Y. Bigi, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, Phys. Rev. Lett. **71**, 496 (1993); A. V. Manohar and M. B. Wise, Phys. Rev. D **49**, 1310 (1994);

T. Mannel, Nucl. Phys. B **413**, 396 (1994).
30. H. D. Politzer, Nucl. Phys. B **172**, 349 (1980); H. Georgi, Nucl. Phys. B **361**, 339 (1991).
31. I. I. Y. Bigi, B. Blok, M. A. Shifman, N. G. Uraltsev and A. I. Vainshtein, arXiv:hep-ph/9212227; A. F. Falk, M. E. Luke and M. J. Savage, Phys. Rev. D **49**, 3367 (1994).
32. Y. Nir, Phys. Lett. **B221**, 184 (1989).
33. A. Kapustin and Z. Ligeti, Phys. Lett. B **355**, 318 (1995); C. Bauer, Phys. Rev. D **57**, 5611 (1998) [Erratum-ibid. D **60**, 099907 (1998)]; A. L. Kagan and M. Neubert, Eur. Phys. J. C **7**, 5 (1999); Z. Ligeti, M. E. Luke, A. V. Manohar and M. B. Wise, Phys. Rev. D **60**, 034019 (1999).
34. A. F. Falk, M. E. Luke and M. J. Savage, Phys. Rev. D **53**, 2491 (1996); Phys. Rev. D **53**, 6316 (1996); M. Gremm and A. Kapustin, Phys. Rev. D **55**, 6924 (1997); A. F. Falk and M. E. Luke, Phys. Rev. D **57**, 424 (1998).
35. M. B. Voloshin, Phys. Rev. D **51**, 4934 (1995); M. Gremm, A. Kapustin, Z. Ligeti and M. B. Wise, Phys. Rev. Lett. **77**, 20 (1996); M. Gremm and I. Stewart, Phys. Rev. D **55**, 1226 (1997);
36. R. A. Briere *et al.* [CLEO Collaboration], arXiv:hep-ex/0209024.
37. M. Beneke, Nucl. Phys. B **405**, 424 (1993); M. E. Luke, A. V. Manohar and M. J. Savage, Phys. Rev. D **51**, 4924 (1995); M. Neubert and C. T. Sachrajda, Nucl. Phys. B **438**, 235 (1995).
38. C. W. Bauer, Z. Ligeti, M. Luke and A. V. Manohar, Phys. Rev. D **67**, 054012 (2003); M. Battaglia *et al.*, eConf **C0304052**, WG102 (2003) [Phys. Lett. B **556**, 41 (2003)].
39. M. Neubert, Phys. Rev. D49 (1994) 3392; D49 (1994) 4623; I.I. Bigi *et al.*, Int. J. Mod. Phys. A9 (1994) 2467; T. Mannel and M. Neubert, Phys. Rev. D50 (1994) 2037.
40. V. Barger, C. S. Kim and R. J. Phillips, Phys. Lett. B251, (1990) 629; A.F. Falk, Z. Ligeti, and M.B. Wise, Phys. Lett. B406 (1997) 225; I. Bigi, R.D. Dikeman, and N. Uraltsev, Eur. Phys. J. C4 (1998) 453; R. D. Dikeman and N. Uraltsev, Nucl. Phys. B509 (1998) 378.
41. C. W. Bauer, M. E. Luke and T. Mannel, arXiv:hep-ph/0102089; Phys. Lett. B543 (2002) 261; M. Neubert, Phys. Lett. B **543** (2002) 269; A. K. Leibovich, Z. Ligeti and M. B. Wise, Phys. Lett. B539 (2002) 242.
42. D. J. Broadhurst and A. G. Grozin, Phys. Lett. B **267**, 105 (1991).
43. P. F. . Harrison and H. R. . Quinn [BABAR Collaboration], "The BaBar physics book: Physics at an asymmetric B factory," SLAC-R-0504 .
44. http://www.slac.stanford.edu/xorg/hfag/semi/summer02/summer02.shtml
45. I. I. Y. Bigi and A. I. Sanda, Cambridge Monogr. Part. Phys. Nucl. Phys. Cosmol. **9**, 1 (2000).
46. K. Abe *et al.* [Belle Collaboration], arXiv:hep-ex/0308035; B. Aubert *et al.* [BABAR Collaboration], Phys. Rev. Lett. **89**, 201802 (2002).
47. . Beneke, G. Buchalla, M. Neubert and C. T. Sachrajda, Phys. Rev. Lett. **83**, 1914 (1999); Nucl. Phys. B **591**, 313 (2000); arXiv:hep-ph/0007256.
48. C. W. Bauer, S. Fleming and M. E. Luke, Phys. Rev. D **63**, 014006 (2001); C. W. Bauer, S. Fleming, D. Pirjol and I. W. Stewart, Phys. Rev. D **63**,

114020 (2001); C. W. Bauer and I. W. Stewart, Phys. Lett. B **516**, 134 (2001); C. W. Bauer, D. Pirjol and I. W. Stewart, Phys. Rev. Lett. **87**, 201806 (2001); C. W. Bauer, D. Pirjol and I. W. Stewart, Phys. Rev. D **65**, 054022 (2002); M. Beneke, A. P. Chapovsky, M. Diehl and T. Feldmann, Nucl. Phys. B **643**, 431 (2002);