QCD ON COARSE LATTICES

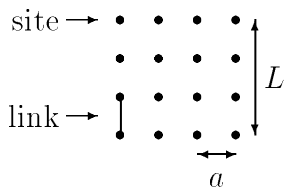
G. Peter Lepage*

Newman Lab. of Nuclear Studies, Cornell University
Ithaca, New York 14853, USA

INTRODUCTION

Lattice quantum chromodynamics is *the* fundamental theory of low-energy strong interactions. In principle, lattice QCD should tell us everything we want to know about hadronic spectra and structure, including such phenomenologically useful things as weak-interaction form factors, decay rates and deep-inelastic structure functions. In these lectures I discuss a revolutionary development that makes the techniques of lattice QCD much more widely accessible and greatly extends the range of problems that are tractable.

The basic approximation in lattice QCD is the replacement of continuous space and time by a discrete grid. The nodes or "sites" of the grid are separated by lattice spacing a, and the length of a side of the grid is L:



The quark and gluon fields from which the theory is built are specified only on the sites of the grid, or on the "links" joining adjacent sites; interpolation is used to find the fields between the sites. In this lattice approximation, the path integral, from which all quantum mechanical properties of the theory can be extracted, becomes an ordinary multidimensional integral where the integration variables are the values of the fields at each of the grid sites:

$$\int \mathcal{D}A_\mu \, \ldots \, \mathrm{e}^{-\int L \mathrm{d}t} \longrightarrow \int \prod_{x_j \in \mathrm{grid}} \mathrm{d}A_\mu(x_j) \, \ldots \, \mathrm{e}^{-a \sum L_j}. \tag{1}$$

Thus the problem of nonperturbative relativistic quantum field theory is reduced to one of numerical integration. The integral is over a large number of variables and so

*email: gpl@lns.mail.cornell.edu

Monte Carlo methods are generally used in its evaluation. Note that the path integral uses euclidean time rather than ordinary minkowski time, where $t_{\rm eucl} = \mathrm{i}\, t_{\rm mink}$; this removes a factor of i from the exponent, getting rid of high-frequency oscillations in the integrand that are hard to integrate.

Early enthusiasm for this approach to nonperturbative QCD gradually gave way to the sobering realization that very large computers would be needed for the numerical integration of the path integral — computers much larger than those that existed in the mid 1970's, when lattice QCD was invented. Much of a lattice theorist's effort in the first twenty years of lattice QCD was spent in accumulating computing time on the world's largest supercomputers, or in designing and building computers far larger than the largest commercially available computers. By the early 1990's, it was widely felt that teraflop computers costing tens of million of dollars would be essential to the final numerical solution of full QCD.

The revolutionary development I discuss here was the introduction of new techniques that allow one to do realistic numerical simulations of QCD on ordinary desktop workstations or even personal computers. To understand this development one must understand the factors that govern the cost of a full QCD simulation. This cost is governed by a formula like

$$\text{cost} \propto \left(\frac{L}{a}\right)^4 \left(\frac{1}{a}\right) \left(\frac{1}{m_\pi^2\, a}\right) \tag{2}$$

where the first factor is just the number of lattice sites in the grid, and the remaining factors account for the "critical slowing down" of the algorithms used in the numerical integration. This formula shows that the single most important determinant of cost is the lattice spacing. The cost varies as the sixth power of $1/a$, suggesting that one ought to keep the lattice spacing as large as possible. Until very recently it was thought that lattice spacings as small as .05–.1 fm would be essential for reliable simulations of QCD. As I describe in these lectures, new simulation results based on new techniques indicate that spacings as large as .4 fm give accurate results. Assuming that the cost is proportional to $(1/a)^6$, these new simulations using coarse lattices should cost 10^3–10^6 times less than traditional simulations on fine lattices.

The computational advantages of coarse lattices are enormous and will certainly redefine numerical QCD: the simplest calculations can be done on a personal computer, while problems of unprecedented difficulty and precision can be tackled with large supercomputers. In these lectures I explain why we now think coarse lattices can be made to work well. I review the techniques from quantum field theory needed to redesign lattice QCD for coarse lattices. These techniques are based on ideas from renormalization theory and effective field theory. I begin by discussing the field-theoretic implications of discretizing space and time. Then I discuss in detail how to discretize the dynamics of gluons and light quarks. These sections present a mixture of very old and very new results. I end the lectures with a simple example of an effective field theory that illustrates the key field theoretic ideas underlying lattice QCD.

FIELD THEORY ON A LATTICE

In this section I discuss the factors that limit the maximum size of the lattice spacing. Replacing space-time by a discrete lattice is an approximation. If we make the lattice spacing too large, our answers will not be sufficiently accurate. For the purpose of these lectures, I define a "sufficiently accurate" simulation to be one that reproduces the low-energy properties of hadrons to within a few percent, which is very precise for low-energy strong-interaction physics. The issue then is: How large can we make a while keeping errors of order a few percent or less?

A nonzero lattice spacing results in two types of error: the error that arises when we replace derivatives in the field equations by finite-difference approximations, and the error due to the ultraviolet cutoff imposed by the lattice. I now discuss each of these in turn, and then focus on the key role played by perturbation theory in correcting for finite-a errors.

Approximate Derivatives

In the lattice approximation, field values are known only at the sites on the lattice. Consequently we approximate derivatives in the field equations by finite-differences that use only field values at the sites. This is completely conventional, and very familiar, numerical analysis. For example, the derivative of a field ψ evaluated at lattice site x_j is approximated by

$$\frac{\partial\psi(x_j)}{\partial x} \approx \Delta_x\psi(x_j) \tag{3}$$

where

$$\Delta_x\psi(x) \equiv \frac{\psi(x+a)-\psi(x-a)}{2a}. \tag{4}$$

It is easy to analyze the error associated with this approximation. Taylor's Theorem implies that

$$\begin{aligned} 2a\,\Delta_x\psi(x) &\equiv \psi(x+a)-\psi(x-a) \\ &= \left(\mathrm{e}^{a\partial_x}-\mathrm{e}^{-a\partial_x}\right)\psi(x) \end{aligned} \tag{5}$$

and therefore

$$\Delta_x\psi = \left(\partial_x + \frac{a^2}{6}\partial_x^3 + \mathcal{O}(a^4)\right)\psi. \tag{6}$$

Thus the relative error in $\Delta_x\psi$ is of order $(a/\lambda)^2$ where λ is the typical length scale in $\psi(x)$.

On coarse lattices we generally need more accurate discretizations than this one. These are easily constructed. For example from Eq. (6) it is obvious that

$$\frac{\partial\psi}{\partial x} = \Delta_x\psi - \frac{a^2}{6}\Delta_x^3\psi + \mathcal{O}(a^4), \tag{7}$$

which is a more accurate discretization. When one wishes to reduce the finite-a errors in a simulation, it is usually far more efficient to improve the discretization of the derivatives than to reduce the lattice spacing. For example, with just the first term in the approximation to $\partial_x\psi$, cutting the lattice spacing in half would reduce a 20% error to 5%; but the cost would increase by a factor of $2^6=64$ in a simulation where cost goes like $1/a^6$. On the other hand, including the a^2 correction to the derivative, while working at the larger lattice spacing, achieves the same reduction in error but with a cost increase of only a factor of 2.

Eq. (7) shows the first two terms of a systematic expansion of the continuum derivative in powers of a^2. In principle, higher-order terms can be included to obtain greater accuracy, but in practice the first couple of terms are sufficiently accurate for most purposes. Simple numerical experiments with a^3-accurate discretizations like this one show that only three or four lattice sites per bump in ψ are needed to achieve accuracies of a few percent or less. Since ordinary hadrons are approximately 1.8 fm in diameter, these experiments suggest that a lattice spacing of .4 fm would suffice for simulating these hadrons. However QCD is a quantum theory, and, as I discuss in the next section, quantum effects can change everything.

Ultraviolet Cutoff

The shortest wavelength oscillation that can be modeled on a lattice is one with wavelength $\lambda_{\min} = 2a$; for example, the function $\psi(x) = +1, -1, +1\ldots$ for $x = 0, a, 2a\ldots$ oscillates with this wavelength. Thus gluons and quarks with momenta $k = 2\pi/\lambda$ larger than π/a are excluded from the lattice theory by the lattice; that is, the lattice functions as an ultraviolet cutoff. In simple classical field theories this is often irrelevant: short-wavelength ultraviolet modes are either unexcited or decouple from the long-wavelength infrared modes of interest. However, in a noisy nonlinear theory, like an interacting quantum field theory, ultraviolet modes strongly affect infrared modes by renormalizing masses and interactions. Thus we cannot simply discard all particles with momenta larger than π/a; we must somehow mimic their effects on infrared states. This we can do by modifying our discretized lagrangian.[†]

To see how we might mimic the effects of $k > \pi/a$ states on low momentum states, consider the scattering amplitude T for quark-quark scattering in one-loop perturbation theory (Fig. 1a). The difference between the correct amplitude T in the continuum theory and the cut-off amplitude $T^{(a)}$ in our lattice theory involves internal gluons with momentum k of order π/a or larger. This is because the classical theories agree at low momenta, and therefore all propagators and vertices agree there as well. Thus the loop contributions from low momenta will be the same in T and $T^{(a)}$, and cancel in the difference. Given that the external quarks have momenta p_i that are small relative to π/a, we can expand the difference $T - T^{(a)}$ in a Taylor series in $a\,p_i$ to obtain

$$\begin{aligned} T - T^{(a)} &= a^2\, c(a)\, \overline{u}(p_2)\gamma_\mu u(p_1)\, \overline{u}(p_4)\gamma^\mu u(p_3) \\ &+ a^2\, c_A(a)\, \overline{u}\gamma_\mu\gamma_5 u\, \overline{u}\gamma^\mu\gamma_5 u \\ &+ a^4\, d(a)\, (p_1 - p_2)^2\, \overline{u}\gamma_\mu u\, \overline{u}\gamma^\mu u \\ &+ \cdots . \end{aligned} \tag{8}$$

where the couplings c, $c_A\ldots$ are dimensionless functions of the cutoff. This difference is what is missing from the lattice theory; it is the contribution of the $k > \pi/a$ states excluded by the lattice.[‡] The key observation is that we can reintroduce this high-k contribution into the lattice theory by adding new interactions to the lattice lagrangian:

$$\begin{aligned} \delta\mathcal{L}_{4\mathrm{q}}^{(a)} &= \tfrac{1}{2}\, a^2\, c(a)\, \overline{\psi}\gamma_\mu\psi\, \overline{\psi}\gamma^\mu\psi \\ &+ \tfrac{1}{2}\, a^2\, c_A(a)\, \overline{\psi}\gamma_\mu\gamma_5\psi\, \overline{\psi}\gamma^\mu\gamma_5\psi \\ &+ a^4\, d(a)\, \overline{\psi}D^2\gamma_\mu\psi\, \overline{\psi}\gamma^\mu\psi \\ &+ \cdots . \end{aligned} \tag{9}$$

These new local interactions give the same contribution to $T(qq \to qq)$ in the lattice theory as the $k > \pi/a$ states do in the continuum theory. Although these correction terms are nonrenormalizable, they do not cause problems in our lattice theory because it

[†]The idea of modifying the lagrangian to compensate for a finite UV cutoff is central to chiral field theories for pions, nonrelativistic QED/QCD and all other effective field theories. The application to lattice field theories was pioneered in the form discussed here by Symanzik and is refered to as "Symanzik improvement" of lattice operators[1].

[‡]More precisely, this contribution is due to states above the cutoff, and to corrections needed to fix up the states below but sufficiently near the cutoff that they suffer severe lattice distortion.

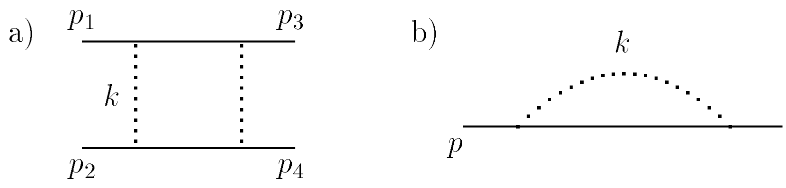


Figure 1. One-loop amplitudes contributing to: a) $qq \to qq$, and b) the quark self energy.

is cut off at π/a. On the contrary, they bring our lattice theory closer to the continuum without reducing the lattice spacing.

This simple analysis illustrates a general result of renormalization theory: one can mimic the effects of states excluded by a cutoff with extra local interactions in the cut-off lagrangian. The correction terms in $\delta\mathcal{L}^{(a)}$ are all local — that is, they are polynomials in the fields and in derivatives of the fields — since contributions omitted as a consequence of the cutoff can be Taylor-expanded in powers of ap_i. This is intuitively reasonable since these terms correct for contributions from intermediate states that are highly virtual and thus quite local in extent; the locality is a consequence of the uncertainty principle. The relative importance of the different interactions is determined by the number of powers of a in their coefficients, and that is determined by the dimensions of the operators involved: each term in $\delta\mathcal{L}^{(a)}$ must have total (energy) dimension four, and so if the operator in a particular term has dimension $n+4$ then the coefficient must contain a factor a^n. In principle there are infinitely many correction terms, but we need only include interactions with operators of dimension $n+4$ or less to achieve accuracy through order $(ap_i)^n$. In practice we only need precision to a given order n in ap_i, and there are only a finite number of local operators with dimension less than or equal to any given $n+4$.

Initially the $k > \pi/a$ contributions appear to be bad news. As we argued in the previous section, a^2 corrections are necessary for precision when a is large. However the quantum nature of our theory implies that there are a^2 terms, due to $k > \pi/a$ states, which depend in detail on the nature of the theory. Thus, for example, when we discretize the derivative in the QCD quark action we replace

$$\overline{\psi}\partial \cdot \gamma\psi \to \overline{\psi}\Delta \cdot \gamma\psi + a^2\, d(a)\, \overline{\psi}\Delta^{(3)} \cdot \gamma\psi + \cdots \tag{10}$$

where now $d(a)$ has a part, $-1/6$, from numerical analysis (equation Eq. (7)) plus a contribution that mimics the $k > \pi/a$ part of the quark self energy (Fig. 1b). The problem is that the latter contribution is, by its nature, completely theory specific; we cannot look it up in a book on numerical analysis. We must somehow solve the $k > \pi/a$ part of the field theory in order to calculate it; otherwise we are unable to correct our lagrangian and unable to use large lattice spacings.

The good news, for lattice QCD, is that $k > \pi/a$ QCD is perturbative provided a is small enough (because of asymptotic freedom). Then correction coefficients like $d(a)$ can be computed perturbatively, using Feynman diagrams, in an expansion in powers of $\alpha_s(\pi/a)$. Thus, for example, our corrected lattice action for quarks becomes

$$\begin{aligned}\mathcal{L}^{(a)} &= \overline{\psi}\,(\Delta \cdot \gamma + m(a))\,\psi \\ &+ d(a)\, a^2\, \overline{\psi}\Delta^{(3)} \cdot \gamma\psi \\ &+ c(a)\, a^2\, \overline{\psi}\gamma_\mu\psi\, \overline{\psi}\gamma^\mu\psi + \ldots\end{aligned} \tag{11}$$

where

$$d(a) = -\frac{1}{6} + d_1\, \alpha_s(\pi/a) + \cdots \tag{12}$$

$$c(a) = c_2\,\alpha_s^2(\pi/a) + \cdots \tag{13}$$
$$\cdots$$

are computed to whatever order in perturbation theory is necessary. In this way we use perturbation theory to, in effect, fill in the gaps between lattice points, allowing us to obtain continuum results without taking the lattice spacing to zero.

Perturbation Theory and Tadpole Improvement[2]

Improved discretizations and large lattice spacings are old ideas, pioneered by Wilson, Symanzik and others[3]. However, perturbation theory is essential; the lattice spacing a must be small enough so that $k \approx \pi/a$ QCD is perturbative. This was the requirement that drove lattice QCD towards very costly simulations with tiny lattice spacings. Traditional perturbation theory for lattice QCD begins to fail at distances of order 1/20 to 1/10 fm, and therefore lattice spacings must be at least this small before improved actions are useful. This seems very odd since phenomenological applications of continuum perturbative QCD suggest that perturbation theory works well down to energies of order 1 GeV, which corresponds to a lattice spacing of 0.6 fm. The breakthrough, in the early 1990's, was the discovery of a trivial modification of lattice QCD, called "tadpole improvement," that allows perturbation theory to work even at distances as large as[2, 4] 1/2 fm.

One can readily derive Feynman diagram rules for lattice QCD using the same techniques as in the continuum, but applied to the lattice lagrangian[5]. The particle propagators and interaction vertices are usually complicated functions of the momenta that become identical to their continuum analogues in the low-momentum limit. All loop momenta are cut off at $k_\mu = \pm\pi/a$.

Testing perturbation theory is also straightforward. One designs short-distance quantities that can be computed easily in a simulation (i.e., in a Monte Carlo evaluation of the lattice path integral). The Monte Carlo gives the exact value which can then be compared with the perturbative expansion for the same quantity. An example of such a quantity is the expectation value of the Wilson loop operator,

$$W(\mathcal{C}) \equiv \langle 0|\tfrac{1}{3}\,\mathrm{Re\ Tr\ P}\,\mathrm{e}^{-ig\oint_{\mathcal{C}} A\cdot dx}|0\rangle, \tag{14}$$

where A is the QCD vector potential, P denotes path ordering, and $\mathcal{C}$ is any small, closed path or loop on the lattice. $W(\mathcal{C})$ is perturbative for sufficiently small loops $\mathcal{C}$. We can test the utility of perturbation theory over any range of distances by varying the loop size while comparing numerical Monte Carlo results for $W(\mathcal{C})$ with perturbation theory.

Fig. 2 illustrates the highly unsatisfactory state of traditional lattice-QCD perturbation theory. There I show the "Creutz ratio" of $2a \times 2a$, $2a \times a$ and $a \times a$ Wilson loops,

$$\chi_{2,2} \equiv -\ln\left(\frac{W(2a\times 2a)\,W(a\times a)}{W^2(2a\times a)}\right), \tag{15}$$

plotted versus the size $2a$ of the largest loop. Traditional perturbation theory (dotted lines) underestimates the exact result by factors of three or four for loops of order 1/2 fm; only when the loops are smaller than 1/20 fm does perturbation theory begin to give accurate results.

The problem with traditional lattice-QCD perturbation theory is that the coupling it uses is much too small. The standard practice was to express perturbative expansions of short-distance lattice quantities in terms of the bare coupling α_{lat} used in the lattice

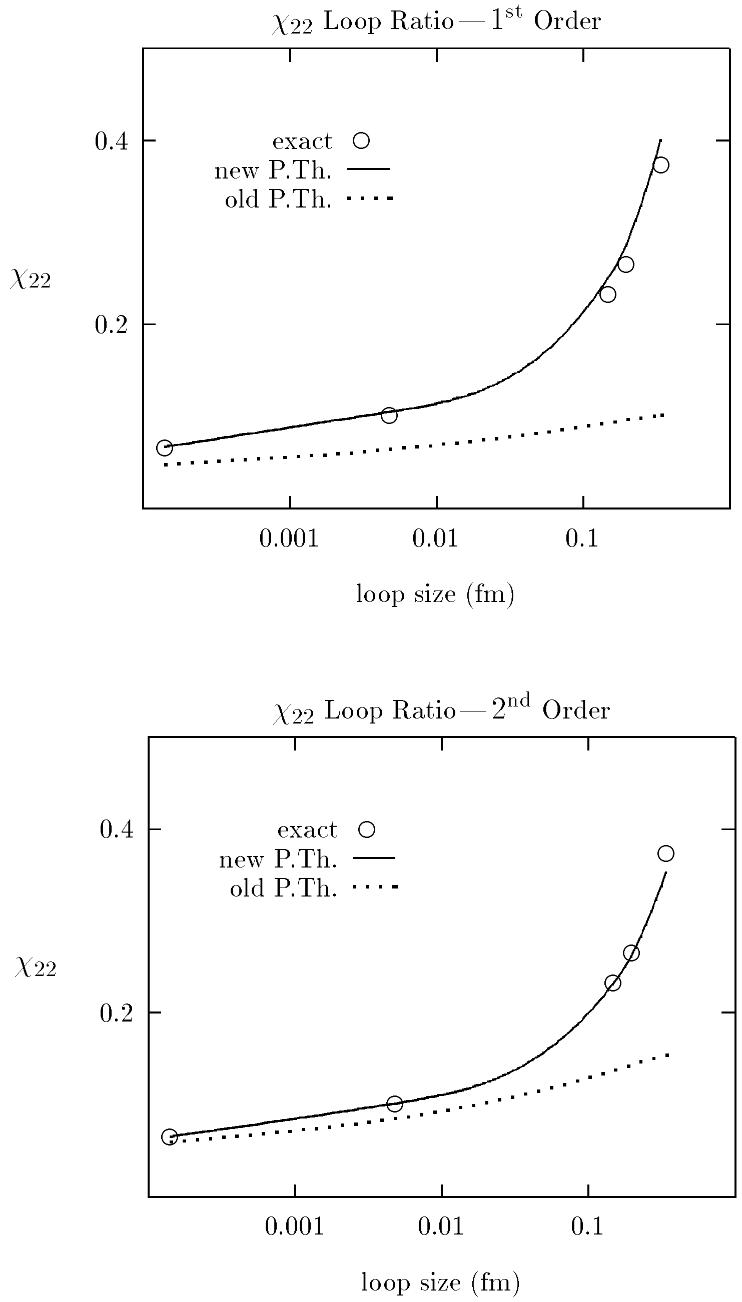


Figure 2. The χ_{22} Creutz ratio of Wilson loops versus loop size. Results from Monte Carlo simulations (exact), and from tadpole-improved (new) and traditional (old) lattice perturbation theory are shown.

lagrangian. This practice followed from the notion that the bare coupling in a cutoff theory is approximately equal to the running coupling evaluated at the cutoff scale, here $\alpha_s(\pi/a)$, and therefore that it is the appropriate coupling for quantities dominated by momenta near the cutoff. In fact the bare coupling in traditional lattice QCD is much smaller than true effective coupling at large lattice spacings: for example,

$$\alpha_{\rm lat} = \alpha_V(\pi/a) - 4.7\,\alpha_V^2 + \cdots \tag{16}$$

$$\leq \tfrac{1}{2}\alpha_V(\pi/a) \qquad \text{for } a > .1\,\text{fm} \tag{17}$$

where $\alpha_V(q)$ is a continuum coupling defined by the static-quark potential,

$$V_{\rm Q\overline{Q}}(q) \equiv -4\,\pi\,C_{\rm F}\frac{\alpha_V(q)}{q^2}. \tag{18}$$

Consequently $\alpha_{\rm lat}$ expansions, though formally correct, badly underestimate perturbative effects, and converge poorly.

The anomalously small bare coupling in the traditional lattice theory is a symptom of the "tadpole problem". As we discuss later, all gluonic operators in lattice QCD are built from the link operator

$$U_\mu(x) \equiv \mathrm{P}\,\mathrm{e}^{-\mathrm{i}\int_x^{x+a\hat\mu} gA\cdot \mathrm{d}x} \approx \mathrm{e}^{-\mathrm{i}agA_\mu} \tag{19}$$

rather than from the vector potential A_μ. Thus, for example, the leading term in the lagrangian that couples quarks and gluons is $\overline{\psi}U_\mu\gamma_\mu\psi/a$. Such a term contains the usual $\overline{\psi}gA\cdot\gamma\psi$ vertex, but, in addition, it contains vertices with any number of additional powers of agA_μ. These extra vertices are irrelevant for classical fields since they are suppressed by powers of the lattice spacing. For quantum fields, however, the situation is quite different since pairs of A_μ's, if contracted with each other, generate ultraviolet divergent factors of $1/a^2$ that precisely cancel the extra a's. Consequently the contributions generated by the extra vertices are suppressed by powers of g^2 (not a), and turn out to be uncomfortably large. These are the tadpole contributions.

The tadpoles result in large renormalizations — often as large as a factor of two or three — that spoil naive perturbation theory, and with it our intuition about the connection between lattice operators and the continuum. However tadpole contributions are generically process independent and so it is possible to measure their contribution in one quantity and then correct for them in all other quantities.

The simplest way to do this is to cancel them out. The mean value,

$$u_0 \equiv \langle 0|\tfrac{1}{3}\,\mathrm{Re}\ \mathrm{Tr}\,U_\mu|0\rangle_{\rm LG} \tag{20}$$

consists of only tadpoles and so we can largely cancel the tadpole contributions by dividing every link operator by u_0. That is, in every lattice operator we replace

$$U_\mu(x) \rightarrow \frac{U_\mu(x)}{u_0} \tag{21}$$

where u_0 is computed numerically in a simulation.

The mean link is gauge dependent and can be made arbitrarily small by changing the gauge. Thus we choose the gauge that maximizes $\langle\,\mathrm{Tr}\,U_\mu\rangle$, which happens to be Landau gauge (hence the subscript LG above). This gauge minimizes the tadpoles in u_0, pushing u_0 as close to 1 as possible; any tadpole contribution that remains cannot be a gauge artifact.

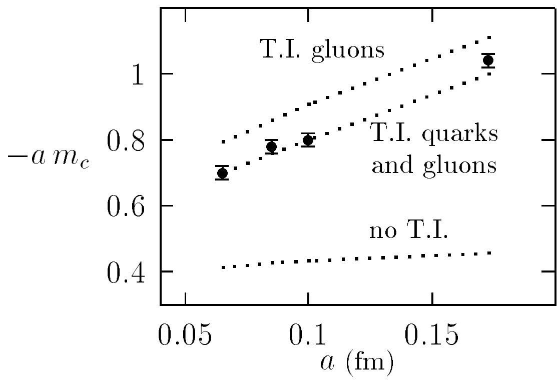


Figure 3. The critical bare quark mass for Wilson's lattice quark action versus lattice spacing. Monte Carlo data points are compared with perturbation theories in a theory with no tadpole improvement (T.I), tadpole-improved gluon dynamics, and tadpole-improved quark and gluon dynamics.

The u_0's cancel tadpole contributions, making lattice operators and perturbation theory far more continuum-like in their behavior. Thus, for example, the only change in the standard gluon action when it is tadpole-improved is that the new bare coupling $\alpha_{\rm TI}$ is enhanced by a factor of $1/u_0^4$ relative to the coupling $\alpha_{\rm lat}$ in the unimproved theory:

$$\alpha_{\rm TI} = \frac{\alpha_{\rm lat}}{u_0^4}. \tag{22}$$

Since $u_0^4 < .6$ when $a > .1\,$fm, the tadpole-improved coupling is typically more than twice as large for coarse lattices. Expressing $\alpha_{\rm TI}$ in terms of the continuum coupling α_V, we find that now our intuition is satisfied:

$$\alpha_{\rm TI} = \alpha_V(\pi/a) - .5\,\alpha_V^2 + \cdots \tag{23}$$
$$\approx \alpha_V(\pi/a). \tag{24}$$

Perturbation theory for the Creutz ratio Eq. (15) converges rapidly to the correct answer when it is reexpressed in terms of $\alpha_{\rm TI}$. An even better result is obtained if the expansion is reexpressed as a series in a coupling constant defined in terms of a physical quantity, like the static-quark potential, where that coupling constant is measured in a simulation. By measuring the coupling we automatically include any large renormalizations of the coupling due to tadpoles. It is important that the scale q^* at which the running coupling constant is evaluated be chosen appropriately for the quantity being studied[2, 4]. When these refinements are added, perturbation theory is dramatically improved, and, as illustrated in Fig. 2, is still quite accurate for loops as large as $1/2\,$fm.

This same conclusion follows from Fig. 3 which shows the value of the bare quark mass needed to obtain zero-mass pions using Wilson's lattice action for quarks. This quantity diverges linearly as the lattice spacing vanishes, and so should be quite perturbative. Here we see dramatic improvements as the tadpoles are removed first from the gluon action, through use of an improved coupling, and then also from the quark action.

The Creutz ratio and the critical quark mass are both very similar to the couplings we need to compute for improved lagrangians. Tadpole improvement has been very successful in a wide range of applications.

Summary

Asymptotic freedom implies that short-distance QCD is simple (perturbative) while long-distance QCD is difficult (nonperturbative). The lattice separates short from long distances, allowing us to exploit this dichotomy to create highly efficient algorithms for solving the entire theory: $k > \pi/a$ QCD is included via corrections $\delta\mathcal{L}$ to the lattice lagrangian that are computed using perturbation theory; $k < \pi/a$ QCD is handled nonperturbatively using Monte Carlo integration. Thus, while we wish to make the lattice spacing a as large as possible, we are constrained by two requirements. First a must be sufficiently small that our finite-difference approximations for derivatives in the lagrangian and field equations are sufficiently accurate. Second a must be sufficiently small that π/a is a perturbative momentum. Numerical experiments indicate that both constraints can be satisfied when $a \approx 1/2\,\mathrm{fm}$ or smaller, provided all lattice operators are tadpole improved.

It is important to remember that almost any perturbative analysis in QCD is contaminated by nonperturbative effects at some level. This will certainly be the case for the couplings in $\delta\mathcal{L}$. However we now have extensive experience showing that such nonperturbative effects are rarely significant for physical quantities at the distances relevant to our discussion. Should a situation arise where this is not the case (and one surely will, some day) we may have to supplement the perturbative approach outlined in this section with nonperturbative techniques. Such situations will not pose a problem if they are relatively rare, as now seems likely. Furthermore, as we discuss later, simple nonperturbative techniques already exist, should they be needed, for tuning the leading correction terms in both the gluon and quark actions.

GLUON DYNAMICS ON COARSE LATTICES

In this section I discuss first the construction of accurate discretizations of the classical theory of gluon dynamics. I then discuss the changes needed to make a quantum theory, and illustrate the discussion with Monte Carlo results. Finally I discuss actions for anisotropic lattices.

Classical Gluons

The continuum action for QCD is

$$S = \int d^4x\, \tfrac{1}{2} \sum_{\mu,\nu} \mathrm{Tr}\, F_{\mu\nu}^2(x) \tag{25}$$

where

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu + ig[A_\mu, A_\nu] \tag{26}$$

is the field tensor, a traceless 3×3 hermitian matrix. The defining characteristic of the theory is its invariance with respect to gauge transformations where

$$F_{\mu\nu} \to \Omega(x)\, F_{\mu\nu}\, \Omega(x)^\dagger \tag{27}$$

and $\Omega(x)$ is an arbitrary x-dependent SU_3 matrix.

The standard discretization of this theory seems perverse at first sight. Rather than specifying the gauge field by the values of $A_\mu(x)$ at the sites of the lattice, the field is specified by variables on the links joining the sites. In the classical theory, the "link variable" on the link joining a site at x to one at $x+a\hat{\mu}$ is determined by the line integral of A_μ along the link:

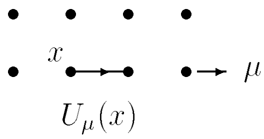
$$U_\mu(x) \equiv \mathrm{P}\, \exp\left(-i\int_x^{x+a\hat{\mu}} gA\cdot \mathrm{d}y\right) \tag{28}$$

where the P-operator path-orders the A_μ's along the integration path. We use U_μ's in place of A_μ's on the lattice, because it is impossible to formulate a lattice version of QCD directly in terms of A_μ's that has exact gauge invariance. The U_μ's, on the other hand, transform very simply under a gauge transformation:

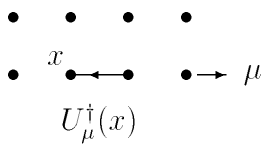
$$U_\mu(x) \to \Omega(x)\, U_\mu(x)\, \Omega(x+a\hat{\mu})^\dagger. \tag{29}$$

This makes it easy to build a discrete theory with exact gauge invariance.

A link variable $U_\mu(x)$ is represented pictorially by a directed line from x to $x+\hat{\mu}$, where this line is the integration path for the line integral in the exponent of $U_\mu(x)$:



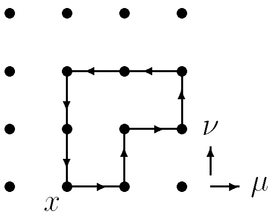
In the conjugate matrix $U_\mu^\dagger(x)$ the direction of the line integral is flipped and so we represent $U_\mu^\dagger(x)$ by a line going backwards from $x+\hat{\mu}$ to x:



A Wilson loop function,

$$W(\mathcal{C}) \equiv \tfrac{1}{3}\,\mathrm{Tr}\;\mathrm{P}\,\mathrm{e}^{-\mathrm{i}\oint_{\mathcal{C}} gA\cdot \mathrm{d}x}, \tag{30}$$

for any closed path $\mathcal{C}$ built of links on the lattice can be computed from the path-ordered product of the U_μ's and $U_\mu^\dagger$'s associated with each link. For example, if $\mathcal{C}$ is the loop



then

$$W(\mathcal{C}) = \tfrac{1}{3}\,\mathrm{Tr}\left(U_\mu(x)U_\nu(x+a\hat{\mu})\ldots U_\nu^\dagger(x)\right). \tag{31}$$

Such quantities are obviously invariant under arbitrary gauge transformations Eq. (29).

We must now build a lattice lagrangian from the link operators. We require that the lagrangian be gauge invariant, local, and symmetric with respect to axis interchanges (which is all that is left of Lorentz invariance). The most local nontrivial gauge invariant object one can build from the link operators is the "plaquette operator," which involves the product of link variables around the smallest square at site x in the $\mu\nu$ plane:

$$P_{\mu\nu}(x) \equiv \tfrac{1}{3}\,\mathrm{Re}\;\mathrm{Tr}\left(U_\mu(x)U_\nu(x+a\hat{\mu})U_\mu^\dagger(x+a\hat{\nu})U_\nu^\dagger(x)\right). \tag{32}$$

To see what this object is, consider evaluating the plaquette centered about a point x_0 for a very smooth weak classical A_μ field. In this limit,

$$P_{\mu\nu} \approx 1 \tag{33}$$

since

$$U_\mu \approx \mathrm{e}^{-\mathrm{i}gaA_\mu} \approx 1. \tag{34}$$

Given that A_μ is slowly varying, its value anywhere on the plaquette should be accurately specified by its value and derivatives at x_0. Thus the corrections to Eq. (33) should be a polynomial in a with coefficients formed from gauge-invariant combinations of $A_\mu(x_0)$ and its derivatives: that is,

$$\begin{aligned} P_{\mu\nu} \;=\; 1 \;&-\; c_1\, a^4\ \mathrm{Tr}\,(gF_{\mu\nu}(x_0))^2 \\ &-\; c_2\, a^6\ \mathrm{Tr}\left(gF_{\mu\nu}(x_0)(\mathrm{D}_\mu^2+\mathrm{D}_\nu^2)gF_{\mu\nu}(x_0)\right) \\ &+\; \mathcal{O}(a^8) \end{aligned} \tag{35}$$

where c_1 and c_2 are constants, and D_μ is the gauge-covariant derivative. The leading correction is order a^4 because $F_{\mu\nu}^2$ is the lowest-dimension gauge-invariant combination of derivatives of A_μ, and it has dimension 4. (There are no F^3 terms because $P_{\mu\nu}$ is invariant under $U_\mu \to U_\mu^\dagger$ or, equivalently, $F \to -F$.)

It is straightforward to find the coefficients c_1 and c_2. We need only examine terms in the expansion of $P_{\mu\nu}$ that are quadratic in A_μ; the cubic and quartic parts of $F_{\mu\mu}^2$ then follow automatically, by gauge invariance. Because of the trace, the path ordering is irrelevant to this order. Thus

$$\begin{aligned} P_{\mu\nu} &= \tfrac{1}{3}\,\mathrm{Re}\ \mathrm{Tr}\ \mathrm{P}\,\mathrm{e}^{-\mathrm{i}\oint_\square gA\cdot\mathrm{d}x} \\ &= \tfrac{1}{3}\,\mathrm{Re}\ \mathrm{Tr}\left[1 - i\oint_\square gA\cdot\mathrm{d}x - \tfrac{1}{2}\left(\oint_\square gA\cdot\mathrm{d}x\right)^2 + \mathcal{O}(A^3)\right] \end{aligned} \tag{36}$$

where, by Stoke's Theorem,

$$\begin{aligned} \oint_\square A\cdot\mathrm{d}x &= \int_{-a/2}^{a/2}\mathrm{d}x_\mu\mathrm{d}x_\nu\,[\partial_\mu A_\nu(x_0+x) - \partial_\nu A_\mu(x_0+x)] \\ &= \int_{-a/2}^{a/2}\mathrm{d}x_\mu\mathrm{d}x_\nu\,[F_{\mu\nu}(x_0) + (x_\mu\mathrm{D}_\mu + x_\nu\mathrm{D}_\mu)F_{\mu\nu}(x_0) + \cdots] \\ &= a^2\,F_{\mu\nu}(x_0) + \frac{a^4}{24}\,(\mathrm{D}_\mu^2+\mathrm{D}_\nu^2)F_{\mu\nu}(x_0) + \mathcal{O}(a^6, A^2). \end{aligned} \tag{37}$$

Thus $c_1 = 1/6$ and $c_2 = 1/72$ in Eq. (35).

The expansion in Eq. (35) is the classical analogue of an operator product expansion. Using this expansion, we find that the traditional "Wilson action" for gluons on a lattice,

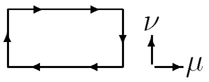
$$S_{\mathrm{Wil}} = \beta \sum_{x,\mu>\nu}(1 - P_{\mu\nu}(x)) \tag{38}$$

where $\beta = 6/g^2$, has the correct limit for small lattice spacing up to corrections of order a^2:

$$S_{\mathrm{Wil}} = \int d^4x \sum_{\mu,\nu}\left\{\tfrac{1}{2}\,\mathrm{Tr}\,F_{\mu\nu}^2 + \frac{a^2}{24}\,\mathrm{Tr}\,F_{\mu\nu}(\mathrm{D}_\mu^2+\mathrm{D}_\nu^2)F_{\mu\nu} + \cdots\right\}. \tag{39}$$

We can cancel the a^2 error in the Wilson action by adding other Wilson loops. For example, the $2a \times a$ "rectangle operator"

$$R_{\mu\nu} = \tfrac{1}{3}\,\mathrm{Re}\;\mathrm{Tr}\tag{40}$$



has expansion

$$R_{\mu\nu} = 1 - \frac{4}{6}\,a^4\,\mathrm{Tr}\,(gF_{\mu\nu})^2 - \frac{4}{72}\,a^6\,\mathrm{Tr}\,\Big(gF_{\mu\nu}(4\,\mathrm{D}_\mu^2 + \mathrm{D}_\nu^2)gF_{\mu\nu}\Big) - \cdots. \tag{41}$$

The mix of a^4 terms and a^6 terms in the rectangle is different from that in the plaquette. Therefore we can combine the two operators to obtain an improved classical lattice action that is accurate up to[6, 7] $\mathcal{O}(a^4)$:

$$S_{\text{classical}} \equiv -\beta \sum_{x,\mu>\nu} \left\{ \frac{5P_{\mu\nu}}{3} - \frac{R_{\mu\nu}+R_{\nu\mu}}{12} \right\} + \text{const} \tag{42}$$

$$= \int d^4x \sum_{\mu,\nu} \tfrac{1}{2}\,\mathrm{Tr}\,F_{\mu\nu}^2 + \mathcal{O}(a^4). \tag{43}$$

This process is the analogue of improving the derivatives in discretizations of non-gauge theories.[§]

Quantum Gluons[8]

In the previous section we derived an improved classical action for gluons that is accurate through order a^4. We now turn this into a quantum action. The most important step is to tadpole improve the action by dividing each link operator U_μ by the mean link u_0: for example, the action built of plaquette and rectangle operators becomes

$$S = -\beta \sum_{x,\mu>\nu} \left\{ \frac{5}{3}\frac{P_{\mu\nu}}{u_0^4} - \frac{R_{\mu\nu}+R_{\nu\mu}}{12\,u_0^6} \right\}. \tag{44}$$

The u_0's cancel lattice tadpole contributions that otherwise would spoil weak-coupling perturbation theory in the lattice theory and undermine our procedure for improving the lattice discretization. Note that $u_0 \approx 3/4$ when $a = .4\,\mathrm{fm}$, and therefore the relative importance of the $R_{\mu\nu}$'s is larger by a factor of $1/u_0^2 \approx 2$ than without tadpole improvement. Without tadpole improvement, we cancel only half of the a^2 error.

The mean link u_0 is computed numerically by guessing a value for use in the action, measuring the mean link in a simulation, and then readjusting the value used in the action accordingly. This tuning cycle converges rapidly to selfconsistent values, and can be done very quickly using small lattice volumes. The u_0's depend only on lattice spacing, and become equal to one as the lattice spacing vanishes. The fourth root of the plaquette is often used in place of the mean link in the definition of u_0. The two agree well at small lattice spacings, but the mean link usually gives better results (that is, more convergent perturbation theory) for coarser lattices.

Tadpole improvement is the first step in a systematic procedure for improving the action. The next step is to add in renormalizations due to contributions from $k > \pi/a$ physics not already included in the tadpole improvement. The omission of this

[§]An important step that I have not discussed is to show that the gluon action is positive for any configuration of link variables. This guarantees that the classical ground state of the lattice action corresponds to $F_{\mu\nu} = 0$. See[7] for a detailed discussion.

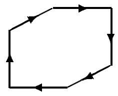
physics induces $a^2\,\alpha_s(\pi/a)$ corrections,

$$\begin{aligned}\delta\mathcal{L} &= \alpha_s\,r_1\,a^2 \sum_{\mu,\nu} \mathrm{Tr}(F_{\mu\nu}\mathrm{D}_\mu^2 F_{\mu\nu}) \\ &+ \alpha_s\,r_2\,a^2 \sum_{\mu,\nu} \mathrm{Tr}(\mathrm{D}_\mu F_{\nu\sigma}\mathrm{D}_\mu F_{\nu\sigma}) \\ &+ \alpha_s\,r_3\,a^2 \sum_{\mu,\nu} \mathrm{Tr}(\mathrm{D}_\mu F_{\mu\sigma}\mathrm{D}_\nu F_{\nu\sigma}) \\ &+ \cdots, \end{aligned} \tag{45}$$

that must be removed. The last term is harmless; its coefficient can be set to zero by a change of field variable (in the path integral) of the form

$$A_\mu \to A_\mu + a^2\,\alpha_s\,f(\alpha_s) \sum_\nu \mathrm{D}_\nu F_{\nu\mu}. \tag{46}$$

Since changing integration variables does not change the value of an integral, such field transformations must leave the physics unchanged.[¶] Operators that can be removed by a field transformation are called "redundant." The other corrections are removed by renormalizing the coefficient of the rectangle operator $R_{\mu\nu}$ in the action, and by adding an additional operator. One choice for the extra operator is

$$C_{\mu\nu\sigma} \equiv \tfrac{1}{3}\,\mathrm{Re\ Tr}$$  . (47)

Then the action, correct up to $\mathcal{O}(a^2\alpha_s^2, a^4)$, is[9]

$$S = -\beta \sum_{x,\mu>\nu} \left\{ \frac{5}{3}\frac{P_{\mu\nu}}{u_0^4} - r_{\mathrm{g}}\,\frac{R_{\mu\nu}+R_{\nu\mu}}{12\,u_0^6} \right\} + c_{\mathrm{g}}\,\beta \sum_{x,\mu>\nu>\sigma} \frac{C_{\mu\nu\sigma}}{u_0^6}, \tag{48}$$

where

$$r_{\mathrm{g}} = 1 + .5\,\alpha_s(\pi/a) \tag{49}$$
$$c_{\mathrm{g}} = .06\,\alpha_s(\pi/a). \tag{50}$$

These values are correct when u_0 is determined from the plaquette; the coefficients of α_s change slightly when the mean-link definition is used. (Of course, both definitions give the same result when the analysis is carried to all orders in α_s.)

The coefficients r_{g} and c_{g} are computed by "matching" physical quantities, like low-energy scattering amplitudes, computed using perturbation theory in the lattice theory with the analogous quantity in the continuum theory. The lattice result depends upon r_{g} and c_{g}; these parameters are tuned until the lattice amplitude agrees with the continuum amplitude to the order in a and α_s required:

$$T_{\mathrm{lat}}(r_{\mathrm{g}}, c_{\mathrm{g}}) = T_{\mathrm{contin}}. \tag{51}$$

Note that tadpole improvement has a big effect on these coefficients. Without tadpole improvement, $r_{\mathrm{g}} = 1+2\alpha_s$; that is, the coefficient of the radiative correction is four times larger. Tadpole improvement automatically supplies 75% of the one-loop contribution needed without improvement. Since $\alpha_s \approx 0.3$, the unimproved expansion for r_{g} is not particularly convergent. However, with tadpole improvement, the one-loop correction is only about 10–20% of r_{g}. Indeed, for most current applications, one-loop corrections to tadpole-improved actions are negligible.

[¶]One must, of course, include the jacobian for the transformation in the transformed path integral. This contributes only in one-loop order and higher; it has no effect on tree-level calculations.

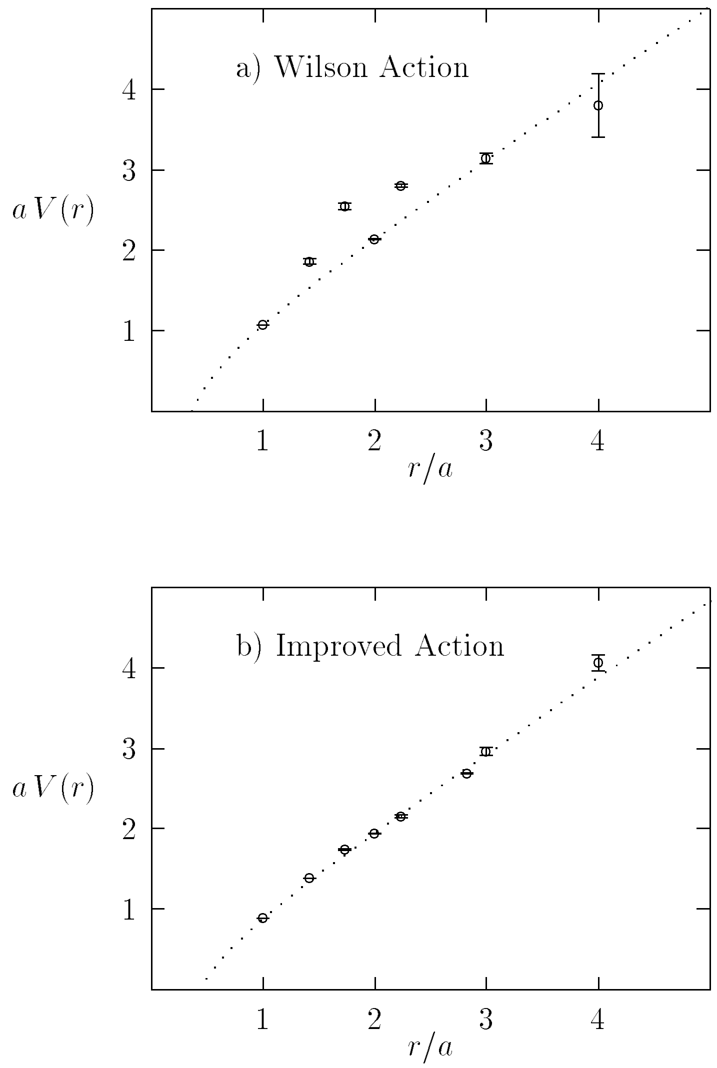


Figure 4. Static-quark potential computed on 6^4 lattices with $a \approx 0.4\,\mathrm{fm}$ using the Wilson action and the improved action. The dotted line is the standard infrared parameterization for the continuum potential, $V(r) = Kr - \pi/12r + c$, adjusted to fit the on-axis values of the potential.

Tests of the Gluon Action

Having a procedure for systematically improving the gluon action, we need to establish whether the improvements really do improve simulations.

The effects of the $\mathcal{O}(a^2)$ improvements in the gluon action are immediately apparent if one compares the static-quark potentials computed with the Wilson action Eq. (38) and with the improved action[8] Eq. (48). Monte Carlo simulation results for these potentials are plotted in Fig. 4, together with the continuum potential. The Wilson action has errors as large as 40%. These are reduced to only 4–5% when the improved action is used instead.

The dominant errors in the uncorrected simulation of $V(r)$ reflect a failure of rotational invariance, which is expected since the a^2 error in the Wilson action Eq. (38) is neither Lorentz nor rotationally invariant. The points at $r = a, 2a, 3a \ldots$ are for separations that are parallel to an axis of the lattice, while the points at $r = \sqrt{2}a, \sqrt{3}a \ldots$ are for diagonal separations between the static quark and antiquark. The $\mathcal{O}(\alpha_s)$ corrections to the action have little effect on these results, suggesting that $\mathcal{O}(a^2\alpha_s)$ corrections are comparable to those of $\mathcal{O}(a^4)$ and therefore unimportant for most current applications. Note also that the rectangle term in the gluon action is the only $\mathcal{O}(a^2)$ term that affects rotation invariance. Thus one can tune its coefficient numerically, without resorting to perturbation theory, until rotation invariance is restored at large r (for example, until $V(2a,2a,a)=V(3a,0,0)$). Obviously, from our example, such nonperturbative tuning gives couplings that are consistent with perturbation theory.

Anisotropic Lattices

It is widely thought that Monte Carlo simulations on coarse lattices are noisier than those on fine lattices. Physics is extracted from simulations by computing Monte Carlo estimates of the correlations (Greens functions) between operators at widely separated times; for example, we compute masses by looking for the large-t plateau in the logarithmic derivative of a correlation function $G(t)$:

$$\frac{1}{G(t)}\frac{dG(t)}{dt} \to m \qquad \text{as } t\to\infty. \tag{52}$$

It is true that such measurements are noisier on conventional coarse lattices, especially for correlation functions where signal/noise vanishes exponentially quickly with increasing t. This is correct despite the fact that the ratio of signal to noise is approximately independent of lattice spacing when t is large; noise overcomes signal at some fixed physical distance, independent of the lattice spacing. The problem on coarse lattices is that this fixed distance might be only one or two lattice spacings long, since the lattice spacing is large. Consequently, without large Monte Carlo data sets, the statistically usable part of a mass plateau, for example, may be too short to allow identification or verification of the plateau. A lattice with small lattice spacing allows us to identify the plateau at smaller distances, with less noise, but at increased cost from the small lattice spacing. Since signal/noise decreases exponentially with increasing t, while simulation cost grows only as a (large) power of lattice spacing, one might decide that smaller lattice spacings are ultimately more efficient. The noise problem on coarse lattices can be greatly alleviated by the careful design of sources and sinks for correlation functions, but a far more efficient approach is to reduce the lattice spacing a_t in the time direction while retaining a large spacing a_s in spatial directions. Then more values of a correlation function can be measured at small t's, where relatively few Monte Carlo measurements are needed to suppress the noise. The small increase in simulation cost that results from a smaller a_t is more than offset by the exponential improvement in signal-to-noise at the smaller t's.

Anisotropic actions are easy to design. The standard improved gluon action becomes[10, 11]

$$\begin{aligned} S_n = {} & -\beta \sum_{x,\,s>s'} \frac{a_t}{a_s}\,\eta \left\{ \frac{5}{3}\frac{P_{s,s'}}{u_s^4} - \frac{1}{12}\frac{R_{ss,s'}}{u_s^6} - \frac{1}{12}\frac{R_{s's',s}}{u_s^6} \right\} \\ & -\beta \sum_{x,\,s} \frac{a_s}{a_t}\,\eta^{-1} \left\{ \frac{4}{3}\frac{P_{s,t}}{u_s^2 u_t^2} - \frac{1}{12}\frac{R_{ss,t}}{u_s^4 u_t^2} \right\}, \end{aligned} \tag{53}$$

where u_s and u_t are the (Landau-gauge) spatial and temporal mean links, respectively, and η accounts for renormalization of the anisotropy by quantum effects. Since $a_t \ll a_s$, we have not bothered to correct for order-a_t^2 errors in this action; thus the action extends

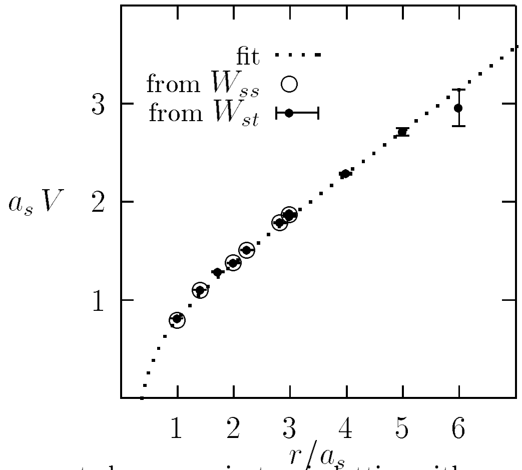


Figure 5. Static potential computed on an anisotropic lattice with $\eta a_t/a_s = 0.5$ and $\beta = 1.8$. Potentials obtained from spatial-temporal loops (W_{st}) and from purely spatial loops (W_{ss}) are shown.

at most one step in the time direction and has no gluon ghosts. An a^2-improved version of Landau gauge is obtained by maximizing

$$\sum_{x\,\mu} \frac{1}{u_\mu a_\mu^2} \,\mathrm{ReTr}\left\{ U_\mu(x) - \frac{1}{16 u_\mu}\, U_\mu(x)\, U_\mu(x+\hat\mu) \right\}. \tag{54}$$

The anisotropic gluon action seems to have six parameters, but these appear in only three independent combinations,

$$\frac{\beta}{u_s^3 u_t}, \qquad \left(\frac{\eta a_t}{a_s}\right)\frac{u_t}{u_s}, \qquad u_s, \tag{55}$$

which control the spatial lattice spacing, the anisotropy, and the size of the order a_s^2 corrections, respectively. These can be tuned numerically by choosing β and $\eta a_t/a_s$, then determining u_t and u_s self-consistently, and finally measuring η. Alternatively one might compute the last three parameters using perturbation theory[12].

One can readily test an anisotropic simulation by showing that the (euclidean) time direction, with its smaller lattice spacing, is equivalent to any other direction. For example, in the same simulation, one can compute the same quantity twice, once taking the time direction to be along the time axis and once with, for example, the t and z axes interchanged. The two calculation should agree. Fig. 5 compares the static-quark potential computed on an anisotropic lattice in these two ways. Shifted to account for renormalization, the two potentials agree well for all r.

Anisotropic lattices with $a_t \ll a_s$ are useful for any simulation involving high-energy states, including glueballs, excited hadronic states, and high-momentum final states in form factors. Very impressive results have been obtained for the glueball spectrum using these techniques[13]. A variety of quark and gluon actions designed for anisotropic lattices are currently under development.

Summary

In this section I have outlined how one designs actions for simulating gluon dynamics on coarse lattices. The simulation results show that accurate simulations are possible even on lattices as coarse as $a = .4\,\mathrm{fm}$ using very simple actions. Tadpole improvement is essential, but one-loop radiative corrections to the action are not too important. Rotational invariance of the potential and, for anisotropic lattices, space-time interchange symmetry provide sensitive tests of improved gluon actions, and can be used to tune the leading correction terms.

The coarse lattices make simulations very fast. The simple simulation results shown here, as well as the glueball simulations in[13], were all generated using desktop workstations.

LIGHT QUARKS ON COARSE LATTICES[14]

In this section I review techniques for designing very accurate discretizations of the Dirac equation for quarks in a gluonic field.

The euclidean Dirac lagrangian in the continuum is

$$\mathcal{L} = \overline{\psi}_c \,(\mathrm{D}\cdot\gamma + m_c)\,\psi_c \tag{56}$$

where

$$\{\gamma_\mu, \gamma_\nu\} = 2\,\delta_{\mu\nu} \tag{57}$$

and I take

$$\gamma_t = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \gamma_s = \begin{pmatrix} 0 & \sigma_s \\ \sigma_s & 0 \end{pmatrix} \qquad \gamma_5 \equiv \gamma_t\gamma_x\gamma_y\gamma_z = i\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \tag{58}$$

where, again, s and t as subscripts signify spatial and temporal indices respectively. The obvious discretization is

$$\mathcal{L}_{\mathrm{lat}} = \overline{\psi}_c\left(\Delta\cdot\gamma + m_c + \mathcal{O}(a^2)\right)\psi_c \tag{59}$$

where now Δ_μ is a (tadpole-improved) gauge-covariant derivative:

$$\Delta_\mu\psi(x) \equiv \frac{U_\mu(x)\,\psi(x+a\hat{\mu}) - U^\dagger_\mu(x-a\hat{\mu})\,\psi(x-a\hat{\mu})}{2\,a\,u_0}. \tag{60}$$

Unfortunately this simple lattice Dirac action has a very serious pathology; it suffers from "fermion doubling." This problem can be avoided by introducing a field transformation before discretizing the action. Starting with the continuum action, which we write as

$$\mathcal{L} = \overline{\psi}_c M_c \psi_c, \qquad M_c \equiv D\cdot\gamma + m_c, \tag{61}$$

we introduce a field transformation,

$$\psi = \Omega\psi_c \tag{62}$$

$$\Omega^2 = 1 - \tfrac{1}{2}ra(D\cdot\gamma - m_c) \tag{63}$$

to obtain a new lagrangian $\overline{\psi}M\psi$ where

$$\begin{aligned} M &= \Omega M_c \Omega \\ &= D\cdot\gamma + m_c\left(1 + \tfrac{1}{2}ram_c\right) - \tfrac{1}{2}ra\left(D^2 + \tfrac{1}{2}g\sigma\cdot F\right). \end{aligned} \tag{64}$$

The third term in M, introduced by the field transformation, breaks the doubling symmetry when we discretize the action. Allowing errors only of order a^4 and higher, we obtain the "D234c" discretization[14] of M:

$$\begin{aligned} M_{\mathrm{D}234} \equiv{}& \Delta\cdot\gamma - c_3\,\frac{a^2}{6}\Delta^{(3)}\cdot\gamma + m \\ &- \frac{ra}{2}\left(\Delta^{(2)} + c_F\frac{g\sigma\cdot F}{2} - c_4\frac{a^2}{12}\Delta^{(4)}\right) \end{aligned} \tag{65}$$

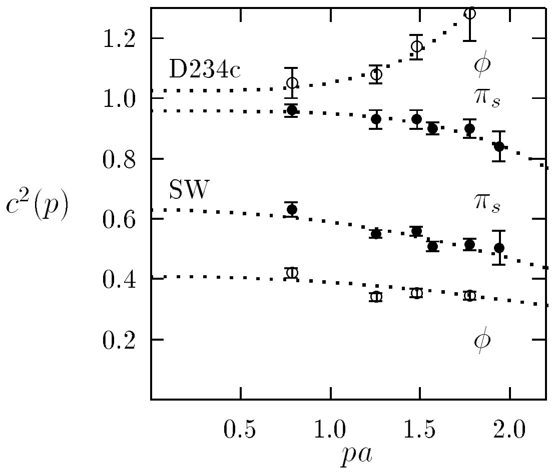


Figure 6. Speed of light squared from the quark dispersion relation using the D234c and SW actions, for pseudoscalar ("π_s") and vector ("ϕ") mesons made of strange quarks. Simulations are on isotropic lattices with $a \approx .4\,\mathrm{fm}$ and an improved gluon action. Fits shown are to $c_0 + c_4(pa)^4$ for D234c and $c_0 + c_2(pa)^2$ for SW. In the continuum, $c^2 = 1$.

where $m \equiv m_c(1 + \frac{1}{2} ram_c)$, $\Delta_\mu^{(n)}$ is the lattice discretization of the n'th gauge covariant derivative D_μ^n, and $F_{\mu\nu}$ is an a^2-improved discretization of the gluon field strength. All operators are tadpole improved; at tree level the couplings $c_F = c_3 = c_4 = 1$. This action with only the $\mathcal{O}(a)$ corrections is know as the Sheikholeslami-Wohlert (SW) action; with no correction terms, it is the Wilson (W) action. These last two actions have been extensively studied. The D234c action is quite new.

Tests of Quark Actions

The D234c action has been extensively tested on coarse lattices with[14] $a = 0.25\,\mathrm{fm}$ and $0.4\,\mathrm{fm}$. This study focused on quark masses near the strange quark mass. Hadrons made of quarks with lighter masses are generally larger and lighter; consequently discretization errors are probably larger for the strange quark than for the other light quarks.

The leading correction in M_{D234c} beyond those in the more conventional SW formalism is

$$- c_3 \frac{a^2}{6} \sum_\mu \Delta_\mu^{(3)} \gamma_\mu. \tag{66}$$

This violates Lorentz invariance; it cancels similar violations in the leading-order terms. Thus we can test this correction by computing

$$c^2(\mathbf{p}) \equiv \frac{E_h^2(\mathbf{p}) - E_h^2(0)}{\mathbf{p}^2} \tag{67}$$

where $E_h(\mathbf{p})$ is the energy of hadron h at three-momentum $\mathbf{p}$. In a Lorentz invariant theory, $c(\mathbf{p})$ should equal 1, the speed of light, for all $\mathbf{p}$. Results at $0.4\,\mathrm{fm}$ are shown in Fig. 6 for both the D234c and SW actions. D234c is dramatically superior, deviating from 1 by only 3–4% at $\mathbf{p} = 0$. Also the dispersion relation is accurate to within 10% out to three-momenta of order $1.5/a$.

We could tune the coupling c_3 in the D234c action nonperturbatively until $c^2(0) = 1$ for the pseudoscalar, thereby restoring Lorentz symmetry through order a^2. Such tuning

Table 1. Percentage change in hadron masses when the redundant coefficient r is changed from 1 to 2/3. The hadrons are all made out of strange quarks; π_s and N_s are analogues of the pion and nucleon. In the continuum, masses are r-independent.

Action:	Wilson	SW	D234c	D234c
a (fm):	0.4	0.4	0.4	0.25
π_s	16.0 (7) %	8 (2) %	2.0 (7) %	1.0 (1) %
ϕ	15.2 (11)	8 (2)	2.0 (5)	1.3 (2)
N_s	9.0 (14)	5 (2)	1.1 (6)	0.7 (2)
Ω	9.0 (20)	8 (2)	1.5 (6)	0.9 (2)

at .4 fm indicates that $c_3 \approx 1 + \alpha_s/2$, which is sufficiently close to $c_3 = 1$ that there is little need to bother. There is no value of c_3 for which both the pseudoscalar and vector have exactly $c^2 = 1$; uncorrected a^4 errors are larger than the radiative corrections to c_3 at this lattice spacing and quark mass.

There are two different definitions of a hadron's mass in lattice simulations. One is the "static mass," $E_h(0)$, and the other the "kinetic mass," which equals $E_h(0)/c^2(0)$. In D234c these two definitions agree to within 3-4%. In SW they differ by 40–60%. The kinetic and static masses in either formalism must be equal for zero mass mesons, because of the axis-interchange symmetry of the actions. The deviations seen here are because the strange quark is relatively massive at 0.4 fm; the ϕ mass, for example, is $2.1/a$ with D234c.

Another test of the D234c action involves varying the parameter r. The terms multiplied by r in the D234c action are approximately redundant at low momenta, since they originate from a field transformation. Thus, insofar as our discretization is accurate, hadron masses and other quantities are independent of r. Results from simulations both at the standard value $r = 1$ and at $r = 2/3$ are shown in Table 1. With D234c the hadron masses changed by no more than 1–2% at .4 fm, and, as expected, less at .25 fm. This suggests that finite-a errors due to the r dependent terms in the action are of order 6% or less at .4 fm. With SW, where the r terms are less highly corrected and therefore less redundant, shifts are of order 5–8% at .4 fm, while they are 10–15% for the Wilson action, which is the least corrected.

The coefficients c_F, c_3, and c_4 are all renormalized by residual quantum effects, even after tadpole improvement. These can be varied in simulations to assess the importance of the potential corrections. Simulations show that only c_F has a strong effect on the static hadron masses. D234c with $c_F = 1$ gives errors in the ϕ mass of order 2% at 0.25 fm and only 7% at 0.4 fm; but a radiative correction of order $\pm\alpha_V/6$ in c_F would either remove these errors or double them. Thus to achieve precisions of order a couple percent at such large lattice spacings, we require a calculation of the order α_V correction to c_F. Work has begun on the perturbative evaluation of this correction; nonperturbative tuning is also possible[14]. Preliminary results suggest that radiative corrections are small.

These tests show that improved actions work very well for light quarks. Work continues on applications to phenomenology and on variations for anisotropic lattices. Simulations at .4 fm, with errors of order 5% or less, seem likely.

EFFECTIVE FIELD THEORY: AN EXERCISE[15]

Lattice QCD is a highly nontrivial example of an effective field theory. An effective field theory is a low-energy approximation to some more general theory. Here low-energy refers to wavelengths longer than the lattice spacing, while the more general theory is full continuum QCD, which is valid at wavelengths far smaller than any practical lattice spacing. A thorough understanding of effective field theories is essential for work in lattice QCD, and therefore it is useful to have simple examples of effective field theories with which to explore one's understanding. In this section I present a very simple example of an effective theory, one that uses nothing more sophisticated than the Schrödinger equation. This example, nevertheless, illustrates nonperturbative renormalization, effective field theories, operator product expansions, and many other crucial aspects of renormalization theory.

In this section, we design an effective Schrödinger theory that reproduces a given collection of low-energy data. We begin by discussing the data we will use. We then examine an obvious, well-known procedure for modelling such data, and its limitations. Next we construct a somewhat less obvious effective theory that overcomes these limitations. This modern approach allows us to model the data with arbitrary precision at low energies. It requires that we nonperturbatively renormalize the Schrödinger equation. The approach is numerical throughout, because this is the simplest approach to the particular problems under study.

Synthetic Data

To begin we need a collection of low-energy data that describe a particular physical system. While we could use real experimental data from a physically interesting system, it is better here to avoid the complexities associated with experimental error by generating "synthetic data."

I generated synthetic data for use in these lectures by inventing a simple physical system, and then solving the Schrödinger equation that describes it. I obtained binding energies, low-energy phase shifts, and matrix elements. For the system, I chose the familiar one-particle Coulombic atom, but with a short-range potential $V_s(\mathbf{r})$ in addition to the Coulomb potential:

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) \tag{68}$$

where

$$V(r) = -\frac{\alpha}{r} + V_s(\mathbf{r}). \tag{69}$$

I arbitrarily set $m = 1$ and $\alpha = 1$.

For our purposes, the short-range potential may be anything; I made one up. The whole point of effective theories is that we can systematically design them directly from low-energy data, with no knowledge of the short-distance dynamics. Thus the form of $V_s(\mathbf{r})$ is irrelevant to our analysis. To underscore this point I will not reveal the functional form of the $V_s(\mathbf{r})$ I used. In what follows we need only know that it has finite range.

Such short-range interactions are common in real Coulombic atoms. The effect of the proton's finite size on the hydrogen atom's spectrum is an example. Another is the weak interaction between the electron and proton, which generates very short-range potentials.

Having chosen a particular $V_s(\mathbf{r})$, I wrote a simple computer code to numerically solve for the radial wavefunctions of energy eigenstates. I used this to compute a

Table 2. Binding energies for various S-wave eigenstates of the synthetic hamiltonian.

level	binding energy	level	binding energy
$1S$	1.28711542	$6S$	0.0155492598
$2S$	0.183325753	$\ldots$	
$3S$	0.0703755485	$10S$	0.00534541931
$4S$	0.0371495726	$20S$	0.00129205010
$5S$	0.0229268241		

Table 3. S-wave phase shifts for the synthetic hamiltonian. Phase shifts are computed for $r = 50$.

energy	phase shift	energy	phase shift
10^{-10}	-0.000421343353	.03	1.232867297
10^{-5}	-0.133227246	.07	-0.579619620
.001	-1.319383451	.1	-1.156444634
.003	0.900186195	.3	-0.106466466
.007	-0.146570028	.7	-1.457426179
.01	-0.654835316	1	1.160634967

variety of binding energies, phase shifts and matrix elements. Some of the S-state binding energies are listed in Table 2. Note that the energies would have been given by $-\alpha^2 m/2n^2 = -1/2n^2$, with $n = 1, 2\ldots$, had there been no V_s. Thus the $1S$ energy is more than doubled by the short-range potential; V_s is not a small perturbation. Also the short-range nature of V_s is evident in this data since the energies approach the Coulombic energies for the very low-energy, large-n states.

Sample S-wave phase shifts are given in Table 3. These phase shifts depend upon the radius at which they are measured because of the long Coulomb tail in the potential $V(\mathbf{r})$; I arbitrarily chose $r = 50$ for my phase-shift measurements, this being much larger than the Bohr radius, $1/\alpha\, m$, of my atom, and also much larger than the range of $V_s(\mathbf{r})$. (Alternatively, one can compute the phase shift with and without $V_s(\mathbf{r})$, and take the difference; the Coulomb divergence cancels in the difference.) I also computed $\langle \mathbf{p}^4 \rangle$ for several S-states, as well as the wavefunctions at the origin; these are tabulated in later sections.

The numerical analysis required to generate such data is minimal. The computation can easily be done using standard numerical analysis packages or symbolic manipulation programs on a personal computer.

Our challenge is to design a simple theory that reproduces our low-energy data with arbitrarily high precision. We must do this using only the data and knowledge of the long-range structure of the theory (that is, we are given m and α).

A Naive Approximation

A standard textbook approach to modelling our data is to approximate the unknown short-range potential by a delta function, whose effect is computed using first-

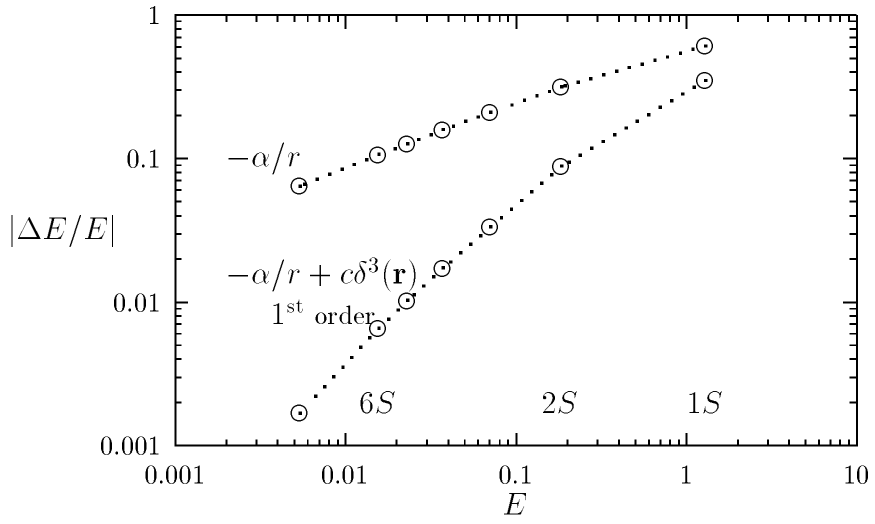


Figure 7. Relative errors in the S-wave binding energies are plotted versus binding energy for the Coulomb theory, and for the Coulomb theory augmented with a delta function in first-order perturbation theory.

order perturbation theory. The approximate hamiltonian is

$$H_{\rm app} = \frac{\mathbf{p}^2}{2m} - \frac{\alpha}{r} + c\,\delta^3(\mathbf{r}), \tag{70}$$

where c is a parameter. Using first-order perturbation theory, the energy levels in our approximate theory are

$$\begin{aligned} E_n^{\rm app} &= E_n^{\rm coul} + c\left|\psi_n^{\rm coul}(0)\right|^2 \\ &= -\frac{1}{2n^2} + c\,\frac{\delta_{l,0}}{\sqrt{\pi}\,n^3}. \end{aligned} \tag{71}$$

Our approximation has only a single parameter, c. This we must determine from the data. We do this by fitting formula (71) to our lowest-energy data, the $20S$ binding energy; this implies $c = -.5963$. We use the lowest-energy state because that is the state for which the replacement $V_s(\mathbf{r}) \to c\,\delta^3(\mathbf{r})$ is most accurate.

In Figure 7 I show the relative errors in several S-wave binding energies, plotted versus binding energy, obtained using just a Coulomb potential ($c = 0$) and using our approximate formula, Eq. (71), with $c = -.5963$. Both approximations become more accurate as the binding energy decreases, but adding the delta function in first order gives substantially better results. Our approximation to V_s is quite successful.

The limitations of this approximation become evident if we seek greater accuracy. We have made two approximations. First we used only first-order perturbation theory. There will be large contributions from second and higher orders in perturbation theory if the short-range potential is strong. So we might wish to compute the second-order contribution to $E_n^{\rm app}$:

$$\sum_{m \neq n} \frac{\langle n|c\,\delta^3(\mathbf{r})|m\rangle\;\langle m|c\,\delta^3(\mathbf{r})|n\rangle}{E_n - E_m}. \tag{72}$$

Unfortunately this expression gives an infinite shift; the sum over scattering eigenstates diverges as scattering momentum $\mathbf{p} \to \infty$. The delta function is too singular to be meaningful beyond first-order perturbation theory.

The second approximation is replacing V_s by a delta function. The nature of this approximation can be appreciated by Fourier transforming V_s. Since V_s has a very short range, its transform $v_s(q^2)$ depends only weakly on the momentum transfer q. Thus we might try approximating v_s by the first few terms in its Taylor expansion when calculating low-energy matrix elements:

$$v_s(q^2) = v_s(0) + q^2\, v_s'(0) + \cdots. \tag{73}$$

This is an economical parameterization of the short-distance dynamics since it replaces a function, $v_s(q^2)$, by a small set of numbers: $v_s(0)$, $v_s'(0)\ldots$. Keeping just the first term is equivalent to approximating $V_s(r)$ by a delta function. The additional terms correspond to derivatives of a delta function, and correct for the fact that the range of V_s is not infinitely small — the expansion is in powers of q times the range. This suggests that we might improve our approximation by adding a second parameter, again to be tuned to reproduce data:

$$V_s(\mathbf{r}) \to c\,\delta^3(\mathbf{r}) + d\,\nabla^2\delta^3(\mathbf{r}). \tag{74}$$

Unfortunately the additional first-order shift in $E_n^{\rm app}$, $\langle n|d\,\nabla^2\delta^3(\mathbf{r})|n\rangle$, is infinite. Again the delta function is too singular at short distances for such matrix elements to make sense.

Our first attempt to refine the approximate model has failed. We replaced the true, but unknown short-distance behavior of the potential by behavior that is pathologically singular as $r \to 0$. Conventional wisdom, often to be seen in books and even recent conference proceedings, is that our approximations must be abandoned once infinities appear, and that only knowledge of the true potential can get us beyond such difficulties. This wisdom is incorrect, as we now see.

Effective Theory

The infinities discussed in the previous section are exactly analogous to those found in relativistic quantum field theories. Among other things, they indicate that even low-energy processes are sensitive to physics at short distances. Modern renormalization theory, however, tells us that the low-energy (infrared) behavior of a theory is independent of the *details* of the short-distance (ultraviolet) dynamics. Insensitivity to the short-distance details means that there are infinitely many theories that have the same low-energy behavior; all are identical at large distances but each is quite different from the others at short distances. Thus we can generally replace the short-distance dynamics of a theory by something different, and perhaps simpler, without changing the low-energy behavior.

The freedom to redesign at short-distances allows us to create effective theories that model arbitrary low-energy data sets with arbitrary precision. There are three steps:

1. Incorporate the correct long-range behavior: The long-range behavior of the underlying theory must be known, and it must be built into the effective theory.

2. Introduce an ultraviolet cutoff to exclude high-momentum states, or, equivalently, to soften the short-distance behavior: The cutoff has two effects. First it excludes high-momentum states, which are sensitive to the unknown short-distance dynamics; only states that we understand are retained. Second, it makes all interactions regular at $r=0$, thereby avoiding the infinities that plague the naive approach of the previous section.

3. Add local correction terms to the effective hamiltonian: These mimic the effects of the high-momentum states excluded by the cutoff in step 2. Each correction term consists of a theory-specific coupling constant, a number, multiplied by a theory-independent local operator. The correction terms systematically remove dependence on the cutoff. Their locality implies that only a finite number of corrections is needed to achieve any given level of precision.

We now apply this algorithm to design an effective hamiltonian that describes our data.

To begin, our effective theory is specified by a hamiltonian,

$$H_{\text{eff}} = \frac{\mathbf{p}^2}{2m} + V_{\text{eff}}(\mathbf{r}), \tag{75}$$

where the effective potential, V_{eff}, must become Coulombic at large r: $V_{\text{eff}}(r) \to -\alpha/r$, with $\alpha = 1$, for large r. We also need an ultraviolet cutoff. I chose to introduce a cutoff into the Coulomb potential through its Fourier transform:

$$\frac{1}{r} \quad \xrightarrow{\text{F.T.}} \quad \frac{4\pi}{q^2} \quad \xrightarrow{\text{cutoff}} \quad \frac{4\pi}{q^2}\,\mathrm{e}^{-q^2a^2/2} \quad \xrightarrow{\text{F.T.}} \quad \frac{\operatorname{erf}(r/\sqrt{2}a)}{r}, \tag{76}$$

where

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2}\,dt \tag{77}$$

is the standard error function. The new, regulated potential is finite at $r=0$, but goes to $1/r$ for $r \gg a$. It inhibits momentum transfers of order $\Lambda \equiv 1/a$ or larger. The exact form of the cutoff is irrelevant; there are infinitely many choices all of which give similar results.

Short-distance dynamics is explicitly excluded from the effective theory by the cutoff. We mimic the effects of the true short-distance structure by adding local correction terms. As discussed in the previous section, the low-momentum behavior of any short-range potential is efficiently described in terms of the Taylor expansion in momentum space. Transforming back to coordinate space gives a series that is a polynomial in the momentum operator $\mathbf{p} \equiv \boldsymbol{\nabla}/\mathrm{i}$ multiplied by a delta function. We need an ultraviolet cutoff to avoid infinities, and therefore we smear the delta function over a volume whose radius is approximately the cutoff distance a. I chose a smeared delta function defined by

$$\delta_a^3(\mathbf{r}) \equiv \frac{\mathrm{e}^{-r^2/2a^2}}{(2\pi)^{3/2}\,a^3}, \tag{78}$$

but, again, the detailed structure of this function is irrelevant; other choices work just as well.

Remarkably, the structure of the correction terms is now completely determined, even though we have yet to examine the data. The effective potential must have the form

$$\begin{aligned} V_{\text{eff}}(\mathbf{r}) = {} & -\frac{\alpha}{r}\operatorname{erf}(r/\sqrt{2}a) + c\,a^2\,\delta_a^3(\mathbf{r}) \\ & + d_1\,a^4\,\nabla^2\delta_a^3(\mathbf{r}) + d_2\,a^4\,\boldsymbol{\nabla}\cdot\delta_a^3(\mathbf{r})\boldsymbol{\nabla} \\ & + \cdots + g\,a^{n+2}\,\nabla^n\delta_a^3(\mathbf{r}) + \cdots, \end{aligned} \tag{79}$$

where coupling constants c, d_1, d_2... are dimensionless. It consists of a long-range part together with a series of local "contact" potentials. The contact terms are indistinguishable, to a low-momentum particle, from the true short-distance potential, provided the coupling constants are properly tuned. The potential is nonrenormalizable in the traditional sense, but that is not a problem here since the cutoff prevents infinities.

Generally the effective potential reflects the symmetries of the true theory. Here our effective potential is rotationally invariant because our data are. When the data are not rotationally invariant, additional terms like $a^3\,\mathbf{g}\cdot\boldsymbol{\nabla}\,\delta_a^3(\mathbf{r})$ must be included in V_{eff}, where now the coupling constants include vectors, tensors… that characterize the rotational asymmetries of the true theory.

One might worry that our Taylor expansion of the short-distance dynamics would fail when we computed physical quantities like the scattering amplitude. This is because high-momentum states affect even low-energy processes through quantum fluctuations. For example, in the true theory the scattering amplitude is

$$\langle f|T(E)|i\rangle = \langle f|V|i\rangle + \sum_n \frac{\langle f|V|n\rangle\,\langle n|V|i\rangle}{E-E_n} + \cdots. \tag{80}$$

The sums over intermediate states in second-order and beyond include states with arbitrarily large momentum. A momentum-space Taylor expansion of the potential would not converge in matrix elements involving such states. Furthermore when we replace the exact potential V by our effective potential V_{eff}, our ultraviolet cutoff in effect limits the sums to states with momenta less than the cutoff $1/a$; contributions from high-momentum intermediate states apparently are completely absent from our effective theory. Our effective theory is saved by the fact that high-momentum intermediate states are necessarily highly virtual if, as we assume, the initial and final states have a low energy. By the uncertainty principle, such states cannot propagate for long times or over large distances. Thus any contribution from these states is very local and can be incorporated into the effective theory using the same set of (smeared) delta-function potentials already included in V_{eff}. In this way, the high-momentum states that are explicitly excluded, or badly distorted, by the cutoff are included implicitly through the coupling constants. Note that this means the coupling constants in our effective theory depend nonlinearly on the true potential $V(r)$.

While the true theory is obviously independent of the value of the cutoff a, results computed in the effective theory are only approximately a independent. The residual a dependence in such a result is typically a power series in qa, where q is the characteristic momentum associated with the process under study (for example, the initial momentum or the momentum transfer in a scattering amplitude). The contact terms remove these a-dependent errors order-by-order in a. Thus, for example, the $\mathcal{O}(a^2)$ term in V_{eff} removes errors of order $(qa)^2$ and is the most important correction. Errors of order a^4 are removed by the $\mathcal{O}(a^4)$ terms, of which there may be only two since there are only two independent ways of combining two momentum operators $\boldsymbol{\nabla}/\mathrm{i}$ with the smeared delta function. Obviously only a finite number of contact terms is needed to remove errors through any finite order in a.

The coupling constants vary with a since they must account for quantum fluctuations excluded by the cutoff. More or less is excluded as a is increased or decreased, and therefore the coupling constants must be adjusted to compensate. They are said to be "running coupling constants." When the short-distance potential is weak, the coupling to high-momentum intermediate states is weak and the coupling constants change only slowly with a.[‖] On the other hand, the coupling constants are often very

[‖]This behavior is modified if the cutoff distance a is very large. When a is larger than the range of the *long-range* potential, or when there is no long-range potential, the coefficients of the contact operators tend to go to a constant. Thus for example, c in V_{eff} will decrease like $1/a^2$ as a increases so that the coefficient of the delta function operator becomes constant.

cutoff dependent when the short-distance interactions are strong. The dependence of the effective theory on its coupling constants becomes highly nonlinear in this case. In fact this behavior can be so strong that the relative importance of different contact terms can change: an a^4 operator can act more like an a^2 operator, and vice versa. The physical systems I discuss here do not suffer from this complication, although it is easy to study the phenomenon by modifying the synthetic data appropriately.

Highly nonlinear behavior also results when the cutoff distance a is made too small. From the previous section we know that taking $a \to 0$ is a bad idea. In general it makes little sense to reduce a below the range r_s of the true potential. By assumption, our data involves energies that are too low — wavelengths that are too long — to probe the true structure of the theory at distances as small as r_s. When $a < r_s$, high-momentum states are included that are sensitive to structure at distances smaller than r_s. But the structure they see there is almost certainly wrong. Thus taking a smaller than r_s cannot improve results obtained from the effective theory. In fact, as the nonlinearities develop for small a's, results often degrade, or, in more extreme cases, the theory may become unstable or untunable.

Finally a comment on the physical significance of the cutoff: A common practice in applications of potential models, for example to nuclear or atomic physics, is to use a cutoff like our's to account for some physical effect, for example finite nuclear size. Then one must worry about whether the functional form of the cutoff is physically correct. We are *not* doing this here. Our cutoff is only a cutoff; physical effects like finite nuclear size are incorporated systematically through the contact terms. We need never worry, for example, about what the true charge distribution is inside a nucleus. Indeed this is the great strength of the effective theory. The effective theory gives us a simple, *universal* parameterization for the effects of short-distance structure. The form of the contact terms is theory independent. Only the numerical values of the coupling constants c, $d_1 \ldots$ are theory specific; for our (low-energy) purposes, everything that we need to know about the short-distance dynamics of the true theory is contained in these coupling constants. As discussed earlier, this situation is conceptually similar to a multipole analysis of a classical field. The couplings here are the analogues of the multipole moments, while the delta-function potentials are analogous to the multipoles that generate the field.

The exact form of the cutoff, and the exact definition of the smeared delta function are irrelevant. Different definitions give similar results for binding energies and phase shifts, but with different values for the coupling constants, and possibly quite different.

Tuning the Effective Theory; Results

We now tune the parameters of our effective theory so that it reproduces our low-energy data through order a^4. For simplicity, we examine only S-wave properties. Thus we need only the c and d_1 terms in V_{eff}, Eq. (79); we can drop the d_2 term in $V_{\text{eff}}(\mathbf{r})$ since it is important only for P-wave states.[**]

To generate the results discussed here I first chose a reasonable value for the cutoff distance: to begin with, $a = 1$, the Bohr radius for the Coulombic part of the interaction. Then I varied the coupling constants c and d_1 until the S-wave phase shifts from the effective hamiltonian agreed with the data at energies $E = 10^{-5}$ and 10^{-10}.

[**]It is obvious that the d_2 term couples only to P-waves in the limit $a \to 0$. When a is nonzero, however, this term has a small residual coupling to S-waves. This results in an interaction of order a^6, which may be ignored to the order we are working here. One can easily design contact terms that contribute only to a single channel in orbital angular momentum. For example, our smeared delta function could be replaced by a local potential that is separable.

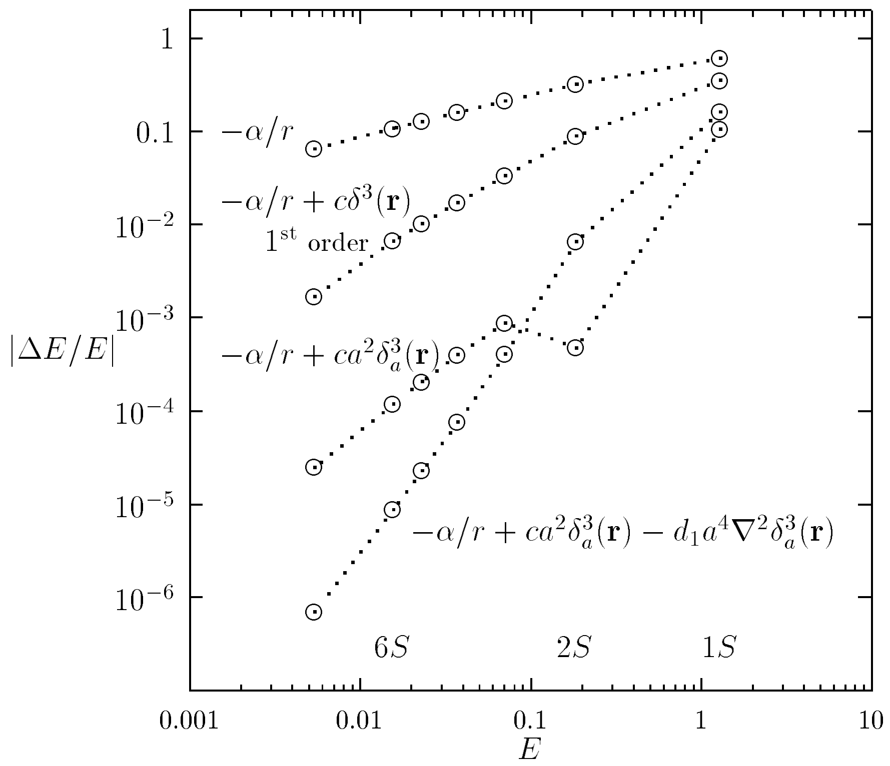


Figure 8. Relative errors in the S-wave binding energies are plotted versus binding energy for the Coulomb theory, the Coulomb theory augmented with a delta function in first-order perturbation theory, the nonperturbative effective theory through a^2, and the effective theory through a^4.

I found $c/4\pi = -3.18$ and $d/4\pi = -0.199$. Having found the coupling constants, I then generated binding energies, phase shifts, and matrix elements using the effective theory. I also generated results with only the a^2 correction ($d_1 = 0$); the effective theory requires a different value of c in this case, since the two contact terms affect each other.

I tuned the couplings using phase shifts at very low energies. This was to minimize the effect of the a^6 errors, which arise because we truncated our effective potential at order a^4. These "truncation" errors are smallest for the lowest-energy data, since they are generally proportional to qa raised to some positive power. In general one should use the most infrared data available when tuning. Alternatively one might attempt a global fit to all of the data; but then it is crucial to give greater weight in the fit to low-energy data than to high-energy data: some estimate of the $(qa)^n$ error due to truncation must be added to the experimental errors used to weight the data in the fit[16].

So long as $a > 0$, $V_{\text{eff}}(\mathbf{r})$ is a simple, nonsingular function of r. Consequently the effective theory is simple to solve numerically. To compute the results here, I reused the code that I wrote to generate the synthetic data, but with the true potential $V(\mathbf{r})$ replaced by the effective potential $V_{\text{eff}}(\mathbf{r})$. The effective theory is no harder to solve than the other; in particular, because of the cutoff, there are no infinities.

In Figure 8 the errors in the binding energies obtained from the tuned effective theory are compared with those obtained above from first-order perturbation theory. Even with only the a^2 correction there is a sizable reduction in errors, due to contributions from second and higher orders in the correction — our numerical solution of the effective theory is nonperturbative, and so includes all orders. The accuracy at low energies is further enhanced by the a^4 correction, which begins to account for the

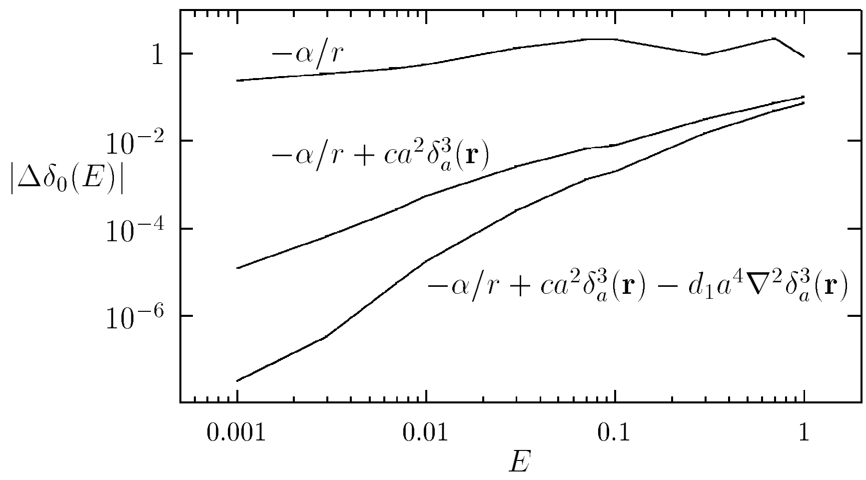


Figure 9. Errors in S-wave phase shifts (in radians) computed with the effective theory corrected through order a^2 and a^4. Values are for $a = 1$ and are plotted versus energy. Results are also shown for the theory with no contact terms.

finite range of the short-range interaction. The crossing of the curves for the a^2 and a^4 theories around $E = 0.1$ has no significance; the error in the a^2 theory changes sign there, passing through zero.

The phase shifts tell a similar story; see Figure 9. At low energies, the errors decrease steadily as the correction terms are added, order-by-order in a^2. The slope of the error curve changes as each new correction is added, getting steeper by one power of $E \propto p^2$ each time the order of the error is increased by a^2, as expected. Similar behavior is apparent in the previous figure, which shows the errors in the binding energies.

Both of these figures show that the correction terms have little effect at energies $E > 1$. At these energies, the particle's wavelength is sufficiently short that the particle can probe the detailed structure of $V_s(\mathbf{r})$. Effective theories are useless in this limit. To go much beyond $E = 1$ in this example, one must somehow uncover the true short-distance structure of the theory.

I redid the analysis for a variety of different a's to explore cutoff dependence. Figure 10 shows how the error in the phase shift depends on a. As a is reduced from $10 \to 3 \to 1$, the errors decrease at all but the highest energies. For this range, a is larger than the range r_s of the true short-distance interaction, and therefore the errors are proportional to a power of qa, which decreases with decreasing a. The errors stop decreasing for $a < 1$, becoming almost insensitive to a. From our discussion in the previous section, this suggests that $r_s \approx 1$. The coupling constants also vary with a in the manner expected if $r_s = 1$. For example, with $d_1 = 0$, coupling constant $c/4\pi$ is roughly -3 and only weakly dependent on a when $a \geq 1$. It grows rapidly as a is reduced below one; for example, $c = -74.0$ when $a = 0.01$.

I also looked for multiple solutions for the coupling constants[17]. Setting $d_1 = 0$ and $a = 1$, for example, I found that the phase shift at $E = 10^{-10}$ is correct when $c/4\pi$ equals -12.4 or 1.6, as well as when it equals -3.18, the value used above. These extra solutions give reasonably good results for the phase shifts and binding energies, but not as good as our original value. The reason is that with these new c's the theory's short-distance structure is significantly modified. In particular, with $c = 24.7$, there is an extra bound state with a very high binding energy (≈ 6). We do not expect our effective theory to work well at high energies, so there is nothing wrong with having the

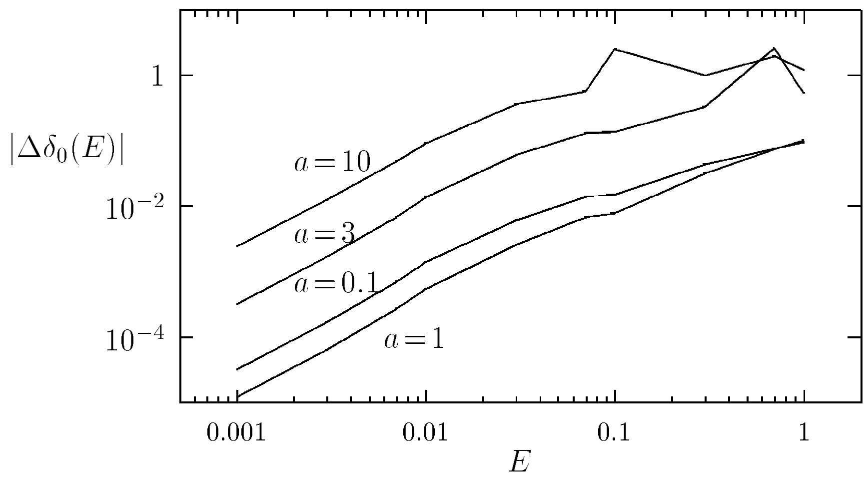


Figure 10. The dependence on cutoff of the$S$-wave phase shift errors plotted against energy is shown. Errors decrease as a is reduced until $a \approx 1$, beyond which point there is no improvement. The a^2 theory was used here ($d_1 = 0$).

extra state; the usual $1S$, $2S\ldots$ states are all still present, each with an extra node in its wavefunction at short distances. When $c/4\pi = 1.6$, the $1S$ state disappears, while the other states are missing a node at short distances.

It is remarkable that a simple two-parameter effective theory, tuned using only two very low energy phase shifts, can reproduce the full range of our data so successfully. The binding energy of the $2S$ state, for example, is accurately predicted by the effective theory to better than 1%, despite the fact that the cutoff scale $a = 1$ is almost a quarter the radius of the atom. Our most infrared data is reproduced to better than six digits.

Operators and the Operator Product Expansion

So far we have used our effective theory to compute binding energies and phase shifts. We now examine quantities that depend in detail on the wavefunctions. Consider, for example, the matrix element $\langle n|\mathbf{p}^4|n\rangle$, which might be important if we wished to include relativistic corrections in our potential model. In Table 4 values of this matrix element are listed for several S-states both for the true theory, and for our corrected theory (with $a = 1$). The values disagree by more than a factor of two, even for very low-energy states, despite the fact that the two theories agree on the corresponding binding energies to several digits.

The problem is that the operator in the effective theory that corresponds to $\mathbf{p}^4$ in the true theory is not $\mathbf{p}^4$. As is true of the hamiltonian, there are local corrections to $\mathbf{p}^4$ in the effective theory. Thus, for any S state, we expect

$$\langle \mathbf{p}^4\rangle_{\rm true} = Z\,\langle \mathbf{p}^4\rangle_{\rm eff} + \frac{\gamma}{a}\,\langle \delta_a^3(\mathbf{r})\rangle_{\rm eff} + \eta\, a\,\langle \nabla^2\delta_a^3(\mathbf{r})\rangle_{\rm eff} + \mathcal{O}(a^3) \tag{81}$$

where Z, γ and η are dimensionless constants that are independent of the state. I tuned these constants so that Eq. (81) gave correct values for the $10S$, $15S$ and $20S$ bound states, and found $Z = 1$, $\gamma = -96.2$ and $\eta = -140.6$ when $a = 1$. This formula gives excellent results, as shown in Table 4. Just as for the binding energies, the relative error decreases steadily with decreasing energy.

This example illustrates the subtlety of the equivalence between the effective theory and the true theory. Although the infrared spectra and phase shifts are quite

Table 4. Expectation values of $\mathbf{p}^4$ for S-states in the true theory, and in the effective theory ($a\!=\!1$). These are compared with matrix elements of the renormalized operator in the effective theory that corresponds to $\mathbf{p}^4$ in the true theory.

level	$\langle \mathbf{p}^4 \rangle$	$\langle \mathbf{p}^4 \rangle_{\rm eff}$	$\langle Z\mathbf{p}^4 + \gamma\delta_a^3/a + \cdots \rangle_{\rm eff}$
$1S$	69	5.9	28
$2S$	5.50	1.7	5.34
$3S$	1.309	0.4	1.306
$4S$	0.5070	0.17	0.5068
$5S$	0.24740	0.08	0.24738
$6S$	0.138784	0.05	0.138780

Table 5. Wavefunctions evaluated at $r\!=\!0$ for S-states in the true theory, and in the effective theory ($a\!=\!1$). These are compared with the renormalized expression in the effective theory that corresponds to the wavefunction at the origin in the true theory.

level	$\psi(0)$	$\psi_{\rm eff}(0)$	$\overline{\gamma}\int \psi_{\rm eff}\,\delta_a^3 + \cdots$
$1S$	1.50	0.53	-3.4
$2S$	0.383	0.19	0.369
$3S$	0.1837	0.09	0.1830
$4S$	0.11353	0.06	0.11344
$5S$	0.079005	0.04	0.078986
$6S$	0.059031	0.03	0.059025

similar, the short-distance structure of corresponding wavefunctions in the two theories may be completely different. This means that matrix elements of operators sensitive to short distances, like $\mathbf{p}^4$, are not equal, even for very low-energy states. Renormalization theory, however, tells us how to define correction terms for the operators in the effective theory that correct for the differences in short-distance behavior between the two theories. Just as in the hamiltonian, the corrections are local operators with state-independent coupling constants, and, again, only a finite number of terms is needed to any given order in a.

Another example of the same idea is provided by the wavefunction at the origin, which is often needed in phenomenological work. In Table 5 values of the wave function at the origin are listed for several S-wave bound states. Again the effective theory gives results that are completely different from the true theory. Renormalization theory, however, implies that

$$\begin{aligned}\psi_{\rm true}(r\!=\!0) &= \overline{\gamma}\int d^3r\,\psi_{\rm eff}\,\delta_a^3(\mathbf{r})\\ &\quad+\overline{\eta}\,a^2\int d^3r\,\psi_{\rm eff}\,\nabla^2\delta_a^3(\mathbf{r})\\ &\quad+\mathcal{O}(a^4)\end{aligned} \tag{82}$$

where $\overline{\gamma}$ and $\overline{\eta}$ are dimensionless, state-independent constants. For $a = 1$, I found $\overline{\gamma} = -28.53$ and $\overline{\eta} = -3.615$ by tuning this formula to give correct results for the $20S$ and $15S$ bound states. With these values, the results from the effective theory for

$\psi_{\text{true}}(0)$ are given in the last column of Table 5. Again the agreement is excellent for all but the most ultraviolet of the states, and it gets better as the states are more infrared — just as expected.

Both of these applications are closely related to the operator product expansion in quantum field theory and to factorization theorems of the sort used in perturbative QCD. Each quantity is a sum of terms, each of which factors into a long-distance state-dependent matrix element times a short-distance universal coefficient. This is exactly the structure exploited in lattice QCD to compute matrix elements of currents or four-quark operators, especially for applications in heavy-quark physics.

An Application: Nucleon-Nucleon Interactions

An example of an effective theory is the pion-nucleon theory of low-energy nuclear interactions. This problem is accurately described by a Schrödinger formalism, rather than a field theory, because pion retardation is negligible — that is, pion exchange is effectively instantaneous. The potential in our nucleon-nucleon hamiltonian can be determined by "matching" scattering amplitudes computed in the potential model with the same amplitudes computed using the chiral field theory, as above. There are both long-range parts, involving the exchange of one or more pions, and short-range or "contact" potentials like the delta function potentials above. The effective theory is an expansion in powers of the momentum divided by Λ_c, the momentum scale at which QCD color becomes apparent (500–1000 MeV).

The leading-order long-range potential is due to the exchange of a single pion; it gives

$$V_{1\pi} = \alpha_\pi\, \boldsymbol{\tau}_1\!\cdot\!\boldsymbol{\tau}_2\, \frac{\boldsymbol{\sigma}_1\!\cdot\!\boldsymbol{\nabla}\, \boldsymbol{\sigma}_2\!\cdot\!\boldsymbol{\nabla}}{m_\pi^2}\, v_\Lambda(r) \tag{83}$$

where $v_\Lambda(r)$ is a Yukawa potential at large r,

$$v_\Lambda(r) \overset{r\,\text{large}}{\longrightarrow} \frac{\mathrm{e}^{-m_\pi r}}{r}, \tag{84}$$

and the $\boldsymbol{\sigma}$'s and $\boldsymbol{\tau}$'s are Pauli matrix operators for the nucleon spin and isospin respectively. Chiral symmetry implies that the dimensionless coupling α_π is related to the pion decay constant, $f_\pi \approx 93\,\text{MeV}$, the nucleon's axial vector coupling, $g_A \approx 1.25$, and the pion mass m_π:

$$\alpha_\pi \approx \frac{g_A^2\, m_\pi^2}{16\pi\, f_\pi^2} \approx 0.06\text{–}0.08. \tag{85}$$

The coupling is small because it is proportional to m_π^2/Λ_c^2, Λ_c being of order $4\sqrt{\pi} f_\pi$. Projecting the one-pion potential onto different spin-orbit eigenstates, we obtain

$$V_{1\pi} \to -\alpha_\pi \left[b\, v_\Lambda(r) + b_T\, v_T(r) \right] \tag{86}$$

where b and b_T are constants that depend upon the channel, and

$$v_T(r) \equiv \frac{1}{m_\pi^2}\left(v_\Lambda'' - \frac{v_\Lambda'}{r} \right) \tag{87}$$

$$\overset{r\,\text{large}}{\longrightarrow} \frac{\mathrm{e}^{-m_\pi r}}{r}\left[1 + \frac{3}{m_\pi r} + \frac{3}{m_\pi^2 r^2} \right]. \tag{88}$$

Here we drop a part of the potential that is indistinguishable from the contact terms we add below. We introduce an ultraviolet cutoff in the Fourier transform, as above, to obtain

$$v_\Lambda(r) \equiv \frac{1}{2r}\left[\mathrm{e}^{-m_\pi r}\mathrm{erfc}\Big(-(r\Lambda - m_\pi/\Lambda)/\sqrt{2} \Big) - (r \to -r) \right], \tag{89}$$

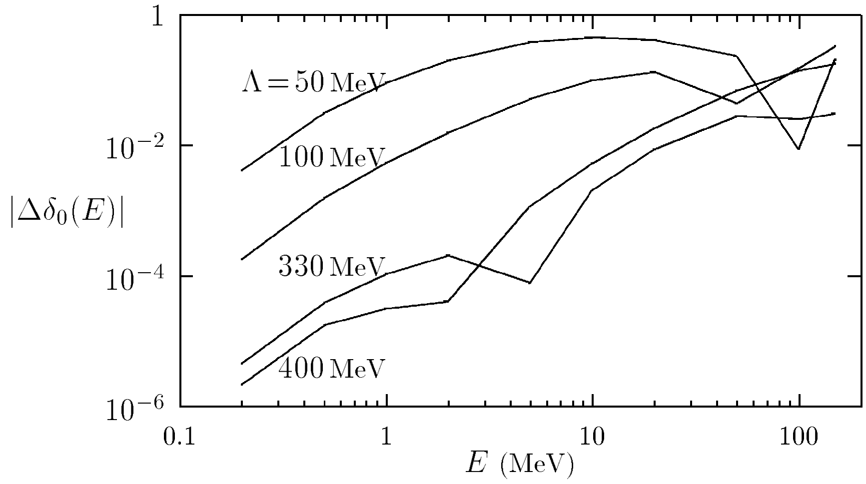


Figure 11. Errors in the $^1S_0$ phase shifts (in radians) versus energy for the full effective theory with different values of the cutoff Λ.

where $\mathrm{erfc}(x) = 1 - \mathrm{erf}(x)$, and $\Lambda = 1/a$ is the momentum cutoff.

Higher-order (in p/Λ) contributions to the long-range potential come from two-pion exchange. We will not include these here; they are described in the literature[18].

The short-distance terms in the potential are smeared delta functions, with $a = 1/\Lambda$, multiplied by powers of the momentum operator $\mathbf{p} = \boldsymbol{\nabla}/\mathrm{i}$, and the spin and isopin operators. It is convenient for our analysis to classify these by the spin-orbit channels that they affect.

The $^1S_0$ and $^3S_1$ channels have Λ^{-2} terms with no $\boldsymbol{\nabla}$'s; there is one for each channel, each with its own coupling constant. In order Λ^{-4} there are two contact potentials, one for each channel, proportional to $\mathbf{p}^2\delta^3_{1/\Lambda}(\mathbf{r}) + \delta^3_{1/\Lambda}(\mathbf{r})\mathbf{p}^2$. The form of this potential can be simplified by adding $-2\,\mathbf{p}\cdot\delta^3_{1/\Lambda}(\mathbf{r})\mathbf{p}$, which doesn't couple to S-states in this order; the resulting potential is then proportional to $(\nabla^2\delta^3_{1/\Lambda})$, just as before.

The P states do not couple to contact terms in $\mathcal{O}(\Lambda^{-2})$. Thus our effective theory predicts that P states are more accurately described by just the long-range potential than S-states. The leading P-wave contact terms are $\mathcal{O}(\Lambda^{-4})$, and, after projecting out the orbital and spin angular momenta, are all proportional to $\mathbf{p}\cdot\delta^3_{1/\Lambda}(\mathbf{r})\mathbf{p}$. This operator can be simplified by adding $-1/2(\mathbf{p}^2\delta^3_{1/\Lambda}(\mathbf{r}) + \delta^3_{1/\Lambda}(\mathbf{r})\mathbf{p}^2)$, which doesn't couple to P states in this order, to obtain again a potential proportional to $(\nabla^2\delta^3_{1/\Lambda})$. There is one such potential for each of the four spin-orbit P-wave channels, and each has its own coupling constant.

There are no contact terms in $\mathcal{O}(\Lambda^{-2}, \Lambda^{-4})$ that couple just to D states; thus D states should be even more accurately described by the long-range potential than P states. The only contact term in order Λ^{-4} that affects a D states is one that couples the $^3D_1$ and $^3S_1$ states.

There are nine contact terms in all through order Λ^{-4}: two Λ^{-2} terms and two Λ^{-4} terms for the S-states, four Λ^{-4} terms for the P states, and one Λ^{-4} term coupling S and D states. Each of these has its own coupling constant. This seems like a lot of coupling constants, but it is only one or two per channel. And so we analyze the np scattering data channel by channel.

We now compare this model with experimental data. For simplicity, we restrict our study to np scattering, thereby avoiding the problem of Coulomb distortion in pp scattering. We fit phase shifts for different channels, using phase-shift data from the

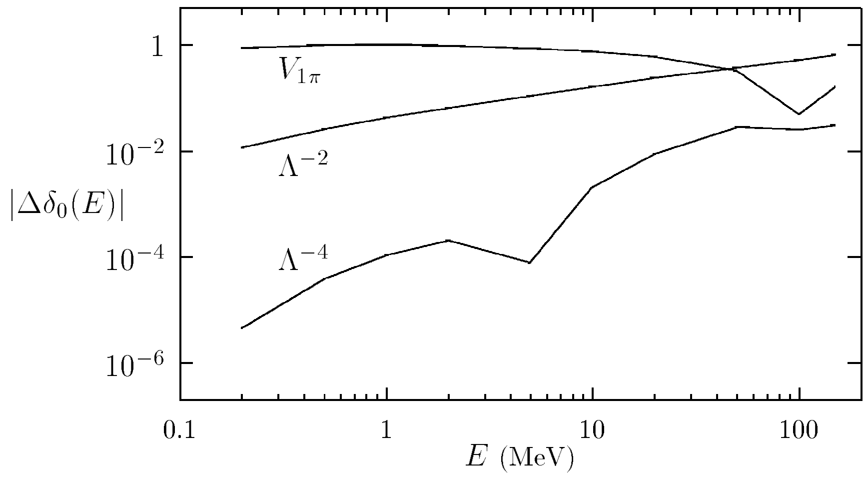


Figure 12. Errors in the $^1S_0$ phase shifts (in radians) versus energy for the effective theory through orders Λ^{-2} and Λ^{-4}. Results from the theory with just pion exchange ($V_{1\pi}$) are also shown. The cutoff was $\Lambda = 330\,\text{MeV}$ in each case.

multi-energy fits by the Nijmegen group[19].

The potential for $^1S_0$ phase shifts, including one-pion exchange and contact terms through order Λ^{-4}, is

$$V(r) = -\alpha_\pi v_\Lambda(r) + c\,\frac{\delta^3_{1/\Lambda}(\mathbf{r})}{\Lambda^2} - d\,\frac{\nabla^2\delta^3_{1/\Lambda}(\mathbf{r})}{\Lambda^4}, \tag{90}$$

where $\alpha_\pi \approx 0.075$ and $m_\pi \approx 140\,\text{MeV}$. Coupling constants c and d must be tuned to the data. When I tried this, I used the Nijmegen phase shifts at $E = 0.05\,\text{MeV}$ and $0.1\,\text{MeV}$ to tune c and d for Λ's ranging from 50 MeV to 1000 MeV. Here and elsewhere in this section E refers to the (nonrelativistic) center-of-mass energy.

In Figure 11 the errors obtained in the phase shift are plotted as a function of center-of-mass energy E for a variety of different values of the cutoff Λ. As in our synthetic example, the errors start out large when Λ is small, and decrease steadily until $\Lambda \approx 300\,\text{MeV}$. No further improvement is seen for larger Λ's.

In Figure 12 the errors obtained from the Λ^{-2} theory ($d = 0$) are compared with those from the Λ^{-4} theory. At low energies the errors in the Λ^{-2} theory grow approximately like $E^{0.5}$ with energy. The errors in the Λ^{-4} theory are much smaller, but grow like $E^{1.5}$ — that is faster by two powers of p, exactly as expected. So here, as in our synthetic data, we see clear evidence that the theory is improved by adding higher-order corrections, and that it is improved in just the manner predicted. This convinces me that the effective theory is working.

The curves in the logarithmic plots of $\Delta\delta_0(E)$ versus E dip at various different points. These dips result when the error changes sign. They should be ignored.

The last figure also shows errors for the effective theory with one-pion exchange but no contact terms ($V_{1\pi}$). Evidently one-pion exchange contributes little to this phase shift at most energies.

In Figure 13 the theoretical $^1S_0$ phase shifts are shown together with the data. The results from the full effective theory are almost independent of the cutoff. The variation at high energies gives an indication of the size of the Λ^{-6} corrections yet to be included.

Results for the $^1P_1$ and $^3D_2$ phase shifts are also shown in Figure 13. The effective

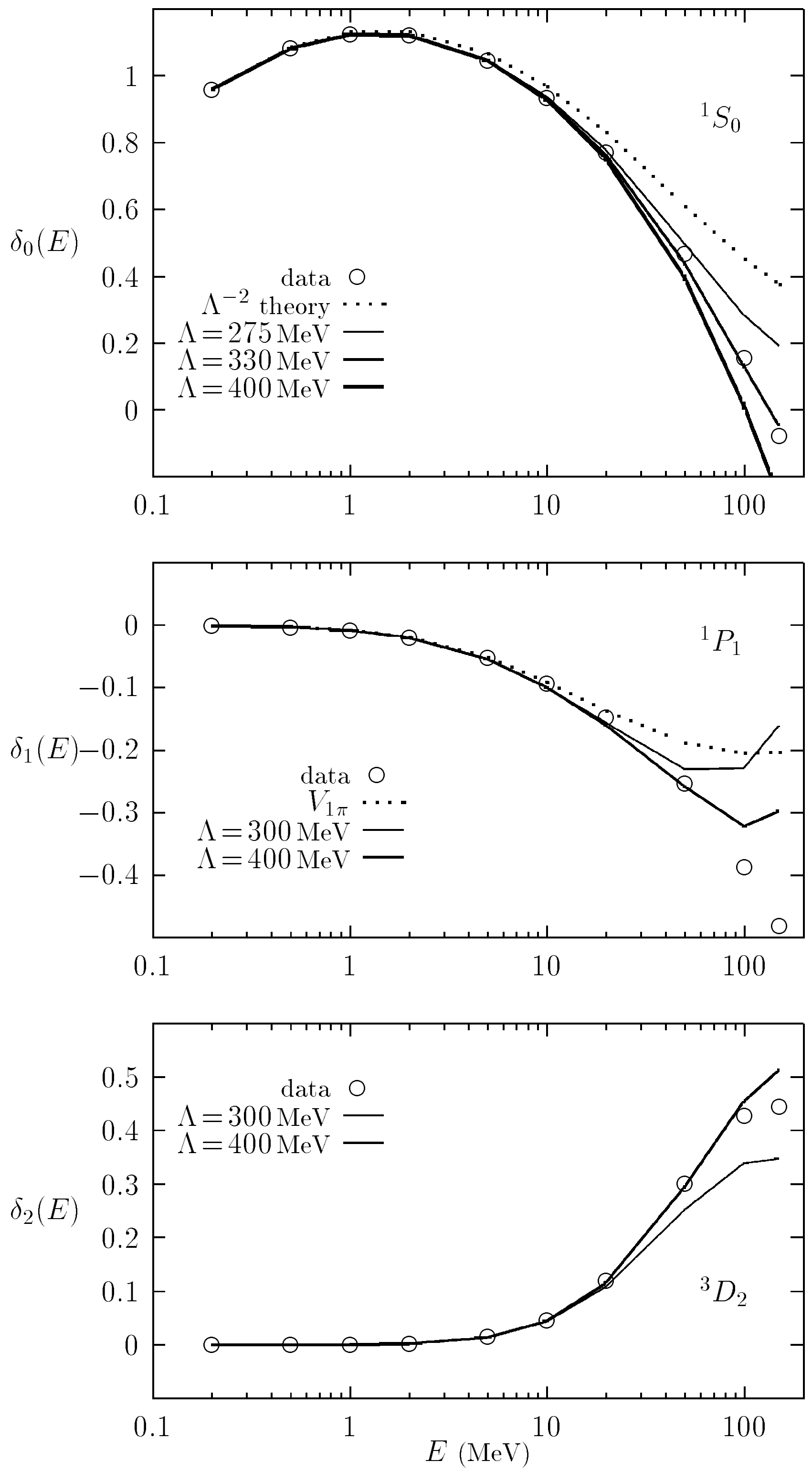


Figure 13. Phase shifts (in radians) for different channels from the effective theory plotted versus energy. Results are shown for various values of the cutoff Λ, and are compared with experiment. S-wave results are for the Λ^{-2} version of the theory ($d=0$, $\Lambda=200$ MeV) and for the full effective theory. P-wave results are for the uncorrected ($V_{1\pi}$) theory, and for the full effective theory. D-wave results are for the uncorrected one-pion theory with different values of Λ; there are no Λ^{-2} or Λ^{-4} corrections for this D-wave channel.

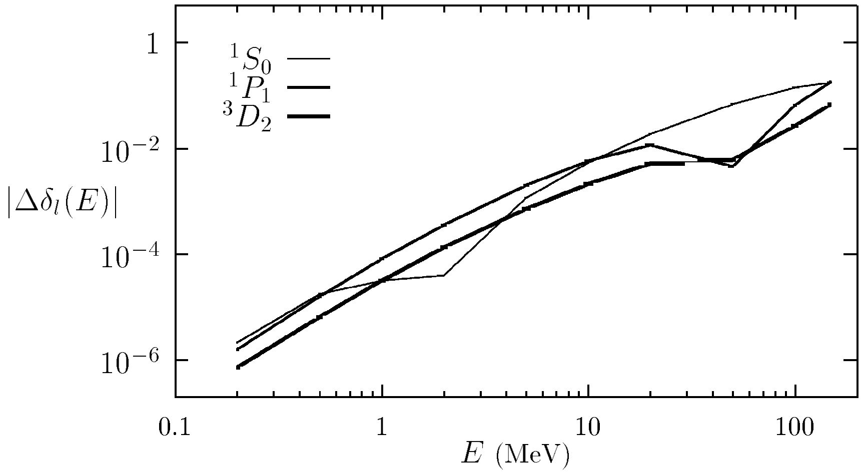


Figure 14. Errors in the phase shifts (in radians) for different channels from the effective theory corrected through order Λ^{-4}.

potentials for these two cases are, through order Λ^{-4},

$$V(^1P_1) = 3\,\alpha_\pi\, v_\Lambda(r) + c(^1P_1)\,\frac{\nabla^2\delta^3_{1/\Lambda}(\mathbf{r})}{\Lambda^4}, \tag{91}$$

and

$$V(^3D_2) = -\alpha_\pi\,[\,v_\Lambda(r) + 2\,v_T(r)]\,. \tag{92}$$

The data confirm that the long-range potential plays a much more important role for P-waves, which have no Λ^{-2} contact term, than for S-waves. And the D-wave analysis has no contact terms at all (until order Λ^{-6}). In both these cases the phase shifts are fairly linear in the couplings, unlike the low-energy S-wave phase shifts.

The relative errors are much smaller for our S-wave phase shifts than for the others. The absolute magnitude of the errors in each channel, however, is roughly the same when the effective theories for each channel are corrected to the same order in $1/\Lambda$. Also the energy dependence of these errors is the same. This is illustrated in Figure 14. It is exactly what is expected based upon the effective theory.

Effective theory does an excellent job of explaining nucleon-nucleon interactions — as it must.

Acknowledgements

These lectures are abridged and updated versions of lectures given at the Schladming Winter School[11] and at the Jorge André Swieca Summer School[15]. My work was supported by the National Science Foundation (USA) and by a fellowship from the Guggenheim Foundation.

REFERENCES

1. K. Symanzik, *Nucl. Phys.* **B226** (1983) 187.
2. G.P. Lepage and P.B. Mackenzie, *Phys. Rev.* **D48** (1993) 2250.
3. For an overview see K.G. Wilson, *Rev. Mod. Phys.* **55** (1983) 583.
4. G.P. Lepage, *Lattice QCD for Small Computers* in *The Building Blocks of Creation,* edited by S. Raby and T. Walker (World Scientific Press, Singapore, 1994).

5. See, for example, H. Kawai, R. Nakayama and K. Seo, *Nucl. Phys.* **B189** (1981) 40.
6. G. Curci, P. Menotti, and G. Paffuti, *Phys. Lett.* **130B** (1983) 205; Erratum: *ibid.* **135B** (1984) 516.
7. M. Lüscher and P. Weisz, *Comm. Math. Phys.* **97** (1985) 59.
8. M. Alford, W. Dimm, G.P. Lepage, G. Hockney and P. Mackenzie, *Phys. Lett.* **B361** (1995) 87.
9. M. Lüscher and P. Weisz, *Phys. Lett.* **158B**, (1985) 250, and references therein.
10. M. Alford, T. Klassen, G.P. Lepage, C. Morningstar and M. Peardon, in preparation.
11. For a review see G.P. Lepage, e-print hep-lat/9607076, in the proceedings of the 1996 Schladming School on *Perturbative and Nonperturbative Aspects of Quantum Field Theory.*
12. G.P. Lepage, e-print hep-lat/9707026, to be published in the proceedings of the International Workshop on Lattice QCD on Parallel Computers (Tsukuba, Japan, Mar 1997).
13. C. Morningstar and M. Peardon, *Phys. Rev.* **D56** (1997) 4043.
14. M. Alford, T. Klassen, and G.P. Lepage, e-print hep-lat/9709126, to be published in the proceedings of Lattice 97 (Edinburgh, Scotland, Jul 1997); *Nucl. Phys.* **B496** (1997) 377; and references therein.
15. G.P. Lepage, e-print nucl-th/9706029, to be published in the proceedings of the 8th Swieca Summer School: Nuclear Physics (Sao Paulo, Brazil, 1997).
16. This point is emphasized in E. Angelos, Cornell University Ph.D. Thesis, 1996.
17. I thank Martin Savage explaining the multiple solutions to me.
18. A particularly useful reference is C. Ordonez, L. Ray and U. van Kolck, *Phys. Rev.* **D53** (1995) 2086. They reference much of the older material as well.
19. Partial wave analysis of the Nijmegen University group, obtained from the World Wide Web page http://nn-online.sci.kun.nl/.