AN INTRODUCTION TO LATTICE QCD

R. D. Kenway
Department of Physics and Astronomy, The University of Edinburgh, The King’s Buildings, Edinburgh EH9 3JZ, Scotland

Abstract
Numerical simulation of QCD is enabling us to calculate strong-interaction effects reliably and thereby to test the Standard Model outside the perturbative regime. I introduce the key features of these simulations, describe the basic first step in which the quark masses are determined and review the status of results for the light hadron spectrum.

1 WHY DO WE NEED NUMERICAL SIMULATIONS OF QCD?
While perturbation theory is reliable for QCD at high energies thanks to asymptotic freedom, the low-energy properties of the theory are intrinsically non-perturbative. It seems unlikely that the complexities of the hadron spectrum and matrix elements will ever be obtained analytically, although we continue to hope for insight into the mechanism by which QCD confines its elementary fields. Fortunately, direct numerical simulation of lattice QCD can proceed from first principles, without any ad hoc assumptions, to compute with arbitrary precision much of the low-energy physics. At present there is a high price attached to doing this, due to the need for very high performance computers.

There are several objectives to this work. In the first instance, we hope to achieve a reliable and systematically-improvable means of calculating non-perturbative aspects of QCD. Perhaps this level of control will lead to an understanding of the mechanism of confinement. Secondly, despite its successes, the Standard Model rests on perturbation theory. Very little of it has been tested non-perturbatively. Many of the Standard Model parameters are obscured experimentally by the need to know hadronic matrix elements. Numerical simulation can provide estimates of these and, perhaps, thereby expose inconsistencies in the Standard Model. Finally, numerical simulation can predict hadronic states, such as glueballs and hybrids, which have yet to be identified experimentally, as well as new phases of QCD, such as the quark-gluon plasma and colour superconductivity, which may exist at non-zero temperature and/or baryon density, and could have important implications for astrophysics.

Thus, there is a lot of physics which we are missing, because it is intrinsically non-perturbative. Numerical simulation is becoming a powerful tool for studying non-perturbative features of quantum field theories and, specifically for QCD, it is likely to play an important role in the search for physics beyond the Standard Model.

In this lecture I will introduce lattice QCD and describe some recent results for the light hadron spectrum (further details can be found in [1]). Christine Davies’ lecture [2] will tell you much more about the phenomenological applications, particularly to the physics of heavy quarks. Several excellent reviews of lattice QCD have appeared recently [3, 4, 5] and, if you want to know the current state of the art, the proceedings of the annual lattice conference [6] gives a complete account.

1.1 Confinement from first principles
It is quite straightforward to demonstrate numerically one of the expected features of confinement, namely the linearly-rising potential between a quark and an antiquark. An example is
Fig. 1: The potential energy, \( V(r) \), between a static quark and antiquark as a function of their separation \( r \) (scaled relative to a fixed physical length \( r_0 \)), computed by the UKQCD Collaboration [7].

shown in Fig 1, where the data are from simulations with different quark masses (parametrised by \( \kappa^{\text{sea}} \) in the figure).

Our intuitive picture of confinement is that a flux tube forms between the quark and the antiquark, giving rise to a constant attractive force. The energy in this confining string therefore builds up linearly with separation, until it becomes energetically favourable for the string to break, with creation from the vacuum of a quark-antiquark pair. The energy at which this is expected to happen is roughly the mass of the two mesons which are formed, and this is indicated in Fig 1 by horizontal dashed lines. Once the string breaks, the potential should flatten out. So far, QCD simulations have only just reached large enough distances to be able to observe this, and a clear signal has not yet been obtained. However, string breaking was demonstrated this year in a gauge-Higgs model [8], so this next step in understanding quark confinement is probably close at hand.

1.2 Parameters of the Standard Model

The basic input parameters for QCD are the quark masses. Since quarks are not asymptotic states, due to confinement, the relationship between the quark masses and any physical observable of QCD is complicated. In a quantum field theory, the parameters in the Lagrangian are fixed by a set of renormalisation conditions. The most direct condition to use for quark masses is to require, for each quark flavour, that the mass of a hadron which contains that flavour is the same as observed experimentally.

The resulting quark mass depends on the regularisation scheme used to calculate the hadron mass, i.e., the lattice, and on the momentum scale, i.e., the lattice momentum cut-off. It is possible to relate one regularisation scheme to another and to use the renormalisation group to change the momentum scale. By convention, quark masses are quoted in the \( \overline{\text{MS}} \) scheme at a scale of 2 GeV.

Other parameters of the Standard Model also require QCD calculations in order to de-
Table 1: The 19 Standard Model parameters (excluding neutrino masses).

<table>
<thead>
<tr>
<th>11 masses</th>
<th>$m_u$</th>
<th>$m_c$</th>
<th>$m_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$m_d$</td>
<td>$m_s$</td>
<td>$m_b$</td>
</tr>
<tr>
<td></td>
<td>$m_e$</td>
<td>$m_\mu$</td>
<td>$m_\tau$</td>
</tr>
<tr>
<td></td>
<td>$m_Z$</td>
<td>$m_H$</td>
<td></td>
</tr>
<tr>
<td>3 couplings</td>
<td>$\alpha$</td>
<td>$G_F$</td>
<td>$\alpha_s$</td>
</tr>
<tr>
<td>3 mixing angles</td>
<td>$V_{ud}$</td>
<td>$V_{us}$</td>
<td>$V_{ub}$</td>
</tr>
<tr>
<td>+</td>
<td>$V_{cd}$</td>
<td>$V_{cs}$</td>
<td>$V_{cb}$</td>
</tr>
<tr>
<td>1 phase</td>
<td>$V_{td}$</td>
<td>$V_{ts}$</td>
<td>$V_{tb}$</td>
</tr>
<tr>
<td>1 vacuum angle</td>
<td>$\theta$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

termine them from experiment. The Standard Model has passed an impressive number of experimental tests, but, even so, we are convinced that it is not a complete theory. Obviously, it excludes the gravitational force. We expect that all four forces have a common origin and that some more complete theory will demonstrate this. Secondly, even excluding the neutrino masses, the Standard Model has 19 free parameters, given in Table 1, which seems too many for them all to be fundamental constants. Surprisingly, only the lepton masses ($m_e$, $m_\mu$ and $m_\tau$), the fine structure constant ($\alpha$), the Fermi constant ($G_F$) and two of the CKM matrix elements ($V_{ud}$ and $V_{us}$) are known to better than 1%. Along with the quark masses, the other elements of the CKM matrix are obscured by hadronic uncertainties.

1.3 The search for new physics

Our interest in knowing better the Standard Model parameters is that the present uncertainties may obscure fundamental inconsistencies. If these are present, they may provide clues to physics beyond the Standard Model. There are several other ways to discover new physics, such as the observation of a supersymmetric particle, or the search for the Higgs boson which must reveal something of the origin of mass, but these may have to wait for the LHC. The most immediate place to look is probably the CKM matrix and the question of why there are three generations.

We know there are exactly three light neutrinos from the observed decay $Z^0 \rightarrow \nu\bar{\nu}$, which provides a measurement of their number:

$$N_\nu = 2.99(2).$$

The consequence of three generations is that the CKM matrix,

$$
\begin{pmatrix}
V_{ud} & V_{us} & V_{ub} \\
V_{cd} & V_{cs} & V_{cb} \\
V_{td} & V_{ts} & V_{tb}
\end{pmatrix}
$$

is unitary,

$$\text{unitary},$$

and so there are non-trivial relationships between the CKM matrix elements, embodied in the so-called unitarity triangles, such as

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

The three complex numbers in Eq (3) form a triangle in the complex plane, whose area is proportional to the amount of $CP$ violation in the Standard Model. If we knew the CKM matrix elements more accurately, we could test the unitarity triangles and, if they close, obtain an indirect measure of $CP$ violation. Of course, we hope to see something going wrong with this picture.
The determination of the CKM matrix elements involves measuring quark decays. The imminent B Factories will provide a wealth of new data on the decays of the $b$ quark, which correspond to the most poorly-known CKM matrix elements. However, $b$ quarks cannot be isolated in the laboratory and always decay inside hadrons. Hence, we have to be able to disentangle the strong-interaction effects in these decays and this usually means being able to compute hadronic matrix elements reliably. Numerical simulation of lattice QCD is essential for getting the most out of B Factories.

2 HOW IS THE COMPUTATION DONE?
Over the 25 years since Wilson invented lattice QCD [9], there have been very considerable technical advances, both in the formulation of QCD on a spacetime lattice and in the numerical algorithms used to simulate it. Progress towards realistic simulations during that time has been as much due to these theoretical developments as to the advances in computer technology. Here, I will try to convey an impression of what, on the one hand, makes QCD a difficult system to simulate and, on the other, gives us confidence that realistic simulations are within reach.

2.1 Lattice discretisation
The starting point is the functional integral for QCD in Euclidean spacetime (ie imaginary time),

$$\langle O \rangle = \frac{1}{Z} \int DAD\bar{q}Dq \, O[A, \bar{q}, q]e^{-S[A, \bar{q}, q]}.$$  (4)

This defines the expectation value of some product, $O$, of the gluon, quark and antiquark fields, $A$, $q$ and $\bar{q}$, as the average of $O$ over all possible field configurations, weighted with the exponential of (minus) the Euclidean action, $S$. Euclidean spacetime is used so that Eq (4) is formally identical to the canonical distribution in statistical mechanics. The resulting probabilistic interpretation of the functional integral then leads naturally to approximating the Euclidean spacetime by a finite $L^3 \times T$ hypercubic lattice, of spacing $a$, and evaluating the average by Monte Carlo.

It was Wilson [9] who realised that a spacetime lattice could be introduced while preserving exact local gauge invariance at the lattice sites. The non-zero lattice spacing regulates QCD in a very natural way, which emphasises its geometrical aspects and does not require any gauge fixing.

However, the lattice obviously breaks the Euclidean spacetime symmetries, and these must be recovered by tuning the lattice spacing to be much smaller than any other relevant length scale. An engineer’s intuition then tells you that the results should become independent of the lattice spacing. Rigorously, this amounts to tuning the gauge coupling to a critical point of the statistical mechanical system, corresponding to asymptotic freedom, ie making contact with the perturbative regime of the theory.

Less obviously, the lattice explicitly breaks chiral symmetry. This symmetry is responsible for ensuring that if the input quark mass is zero, then the quark mass remains zero at all orders in perturbation theory, and the pion is an exact Goldstone boson of the spontaneously-broken chiral symmetry. The fact that the lattice explicitly breaks chiral symmetry means that quark masses are additively renormalised and we have to adjust the up and down quark masses used in the simulations to find a value at which the pion is massless. This is only a minor practical inconvenience, although it complicates the picture of how spontaneous chiral symmetry breaking comes about.

A limitation of today’s numerical algorithms is that they only reproduce Eq (4) exactly for degenerate pairs of quark flavours. Approximate algorithms exist for non-degenerate flavours.
As we shall see, this is of little practical relevance today, because the computational cost is such that most simulations have used the quenched approximation corresponding to zero flavours, or, at best, two flavours in an attempt to simulate up and down quarks.

The simulation of lattice QCD proceeds by taking the gauge coupling and the masses of the \( N_f \) quark flavours as inputs. The Monte Carlo method is used to evaluate expectation values of products of fields from which, as I will indicate in the next section, it is possible to extract hadron masses and matrix elements. Then the experimental values of \( N_f + 1 \) hadron masses are used to fix the quark masses and the overall scale, ie the unit of measurement. In practice, this means that the quark masses are adjusted until \( N_f \) hadron masses, in units of the \((N_f + 1)\)th hadron mass, match experiment. The gauge coupling is tuned towards zero to approach the continuum limit. The lattice spacing is judged to be small enough when dimensionless ratios of hadron masses become independent of the gauge coupling, which is called ‘scaling’.

### 2.2 Quantities which can be computed easily

The starting point is usually to calculate hadron energies. These may be obtained from two-point correlation functions, which are expectation values of products of fields localised at two spacetime points, separated by a distance \( \tau \) in Euclidean time, eg

\[
\langle \mathcal{O}^\dagger(\tau)\mathcal{O}(0) \rangle = \langle 0|\mathcal{T}[\mathcal{O}^\dagger(\tau)\mathcal{O}(0)]|0 \rangle \quad (5)
\]

\[
= \langle 0|\mathcal{O}^\dagger e^{-\mathcal{H}\tau}\mathcal{O}|0 \rangle \quad (6)
\]

\[
= \frac{\sum_n |\langle n|\mathcal{O}|0 \rangle|^2 e^{-E_n\tau}}{2E_n} \quad (7)
\]

Here, in Eq (5), the basic result which relates path integrals to vacuum expectation values of time-ordered products of Heisenberg fields has been used. Next, in Eq (6) the fields have been transformed to the Schrödinger representation, by introducing the Hamiltonian operator in imaginary time and, finally, in Eq (7), a complete set of states, \( \{ |n \rangle \} \), with energies \( E_n \), has been inserted. It follows that the two-point function falls exponentially at large Euclidean time, \( \tau \), with a rate given by the energy of the lightest state which has a non-zero overlap with \( \mathcal{O}|0 \rangle \).

Choosing \( \mathcal{O} \) with particular quantum numbers enables us to extract the energy of the lightest hadron with those quantum numbers. At zero spatial momentum, this is just the hadron mass.

Numerically, the masses are obtained by fitting Eq (7) to a sum of exponentials. As a by-product we also obtain the matrix elements in Eq (7). If \( \mathcal{O}|0 \rangle \) has a non-zero overlap with a pseudoscalar meson, eg \( \mathcal{O} = \bar{q}_1 \gamma_5 q_2 \), then the leading matrix element is proportional to the pseudoscalar decay constant, \( f_{PS} \),

\[
im_{PS}f_{PS} = \langle 0|\bar{q}_1 \gamma_5 q_2|PS \rangle \quad (8)
\]

where \( m_{PS} \) is the pseudoscalar meson mass. This illustrates the basic technique for extracting masses and matrix elements.

The next level of sophistication is to extract single-particle matrix elements from three-point functions. These are expectation values of products of fields at three different Euclidean times, eg, using a similar argument to that in Eqs (5)–(7),

\[
\langle \pi(p_1,\tau_1)\mathcal{O}(q,\tau_2)\mathcal{K}(0) \rangle = \langle 0|\pi(p)e^{-\mathcal{H}(\tau_2-\tau_1)}\mathcal{O}(q)e^{-\mathcal{H}\tau_1}\mathcal{K}|0 \rangle \quad (9)
\]

\[
= \sum_{n,n'} \langle 0|\pi(p)|n \rangle e^{-E_n(\tau_2-\tau_1)} \frac{e^{-E_n\tau_1}}{2E_n} \langle n|\mathcal{O}(q)|n' \rangle e^{-E_{n'}\tau_1} \frac{e^{-E_{n'}\tau_1}}{2E_{n'}} \langle n'|\mathcal{K}|0 \rangle. \quad (10)
\]

The single-particle energies and the matrix elements involving the vacuum may all be obtained from two-point functions, as in Eq (7). Hence, at large time separations, \( \tau_2 \gg \tau_1 \gg 0 \), it is
possible to isolate, for example, the matrix element

$$\langle \pi(p) | s\gamma_\mu u(q) | K(p - q) \rangle$$ (11)

for the semileptonic decay $K \rightarrow \pi e\nu$, or, with other choices of operators, matrix elements for $B \rightarrow K^*\gamma$, $B^0\bar{B}^0$ mixing, etc.

Unfortunately, multiparticle final states involve complex amplitudes in Minkowski space and are much more difficult to extract from this type of analysis. So the current state of the art is limited to computing matrix elements which involve, or can be related to matrix elements which involve, only single-particle final states. Fortunately, this spans a wide range of useful phenomenology.

### 2.3 The size of the lattice

The numerical simulation is carried out in terms of dimensionless variables, with a lattice spacing of 1. So a dimensionful quantity must be input to set the scale and, thereby, determine the lattice spacing in physical units. Often the $\rho$ meson mass is used for this purpose. If the simulation is to match experiment, then the ratio of the Compton wavelength of the $\rho$ meson to the lattice spacing must be the same in the simulation and in the laboratory,

$$\frac{\text{ρ size}}{\text{lattice spacing}} = \left. \frac{m_\rho^{-1}}{1} \right|_{\text{computer}} = \left. \frac{(770 \text{ MeV})^{-1}}{a} \right|_{\text{lab}}.$$ (12)

We have seen in the previous section how to compute $m_\rho$, so this equality determines the lattice spacing, $a$, in inverse MeV.

The key question is how big a lattice is needed, because this equates directly to the cost of the simulation. The box size must be big enough to contain all the hadrons without squashing them significantly. On the other hand, the lattice spacing must be small enough on the scale of the hadrons to resolve all the relevant physics. Thus, we require

$$\text{box size} \gg (\text{masses})^{-1} \gg \text{lattice spacing.}$$ (13)

So, if we were trying to simulate mesons built from the four lightest quark flavours, this becomes

$$L = Na \gg (M_\pi, \ldots, M_{J/\psi})^{-1} \gg a$$ (14)

which would require $N \gg 20$. This type of argument pushes us towards large lattices, which are enormously costly.

Happily, there is an alternative. Since the quark masses may be varied freely in the simulations and we expect the physics to depend smoothly on them, we may simulate QCD using a compressed range of quark masses (and hence smaller lattices) and finally extrapolate the results to the light and/or heavy quark masses which correspond to the real world.

In order to understand how the computational cost varies with $N$, it is necessary to return to the functional integral in Eq (4), which, on the lattice, becomes a multiple integral over the fields at each site (and link). Computers are particularly inefficient at integrating the anticommuting Grassmann fields representing the quarks and antiquarks. Fortunately, this involves only Gaussian integrals and can be done analytically, with the following result:

$$\int dU dq dq e^{-S_G[U] + q(D + m)q} = \int dU [\det(D + m)]^{N_f} e^{-S_G[U]}.$$ (15)

This leaves only the integral over the gluon fields, represented by the link variables $U$ in the above expression, and this can be done using Monte Carlo methods. The evaluation of the
determinant is the most costly part, so that, with today’s best algorithms,

\[
\text{number of arithmetic operations} \propto \begin{cases} 
N^{10} & N_f = 2, 4, \ldots \\
N^6 & N_f = 0 \text{ (quenched)}
\end{cases}.
\]

Eq (16) The huge saving from setting \(N_f = 0\), which, in Feynman-diagram language, removes internal quark loops, motivates the so-called quenched approximation. It is equivalent to assuming that the quarks and antiquarks which can be excited out of the vacuum by gluons, are very massive. Although there is no physical justification for this assumption, it has been widely used to date, just to enable simulations on large enough lattices to make contact with the real world. As we will see, the quenched approximation turns out to be surprisingly accurate for many quantities. So quenched QCD is rather a good model of strong interactions.

The significance of the computational complexity in Eq (16) is dramatically demonstrated when you consider a simple practical test to see if a particular simulation is close enough to the continuum limit. The obvious thing to do is to halve the lattice spacing and see if the results change significantly. For quenched QCD, the calculation on the lattice with twice the number of sites in each direction costs 64 times more computer time. For full QCD, the factor is over 1000. So, even if the original simulation has been done on a PC, to check it requires a supercomputer!

### 2.4 Computers

Since the early days, lattice QCD simulations have pushed against the limits of computer speed and, consequently, have played a part in the development of high-performance computing, particularly parallel computing.

Due to the translational invariance and locality of quantum field theories, their lattice versions are intrinsically parallel. Thus, the lattice may be divided up equally amongst an array of microprocessors, each responsible for the simulation on its part of the lattice and communicating with its neighbours to update fields on the boundary between them. The amount of computation which each microprocessor has, grows with the volume of its portion of the lattice, while the amount of communication it has to do, grows only with the surface area. Thus, the communications overhead is relatively small and may even be completely overlapped by simultaneous computation. The bottom line is that the simulation may be implemented so that the speed is directly proportional to the number of microprocessors.

This means that the major limit on QCD simulations is the size of the parallel computer used, which is dictated by the financial budget available. There is no hard technological barrier in view, given the relentless increase in microprocessor speed, which has been doubling every 19 months since the 1950’s [10]. Currently, the most highly-parallel machine being used for QCD is the 12,000-processor QCDSF at the RIKEN-Brookhaven Center, which has a peak speed of 600 Gflops and is amongst the fastest computers in the world [11].

### 2.5 Improved lattice formulations

The \(N^{10}\) complexity for full QCD is strong motivation for doing better than just waiting for more powerful computers. Very considerable progress has been made in recent years in improving the lattice formulation to cancel systematically the leading discretisation errors in physical quantities [4]. The obvious consequence of this is that we should see scaling at larger lattice spacings, so that \(N\) can be kept small while maintaining a large enough box.

This standard numerical technique is called Symanzik improvement when applied to quantum field theories [12]. For QCD, the \(O(a)\) corrections are cancelled by including one additional
A comparison of the approach to the continuum limit of the vector meson mass computed using different lattice formulations, including two $O(a)$-improved actions (denoted by circles and diamonds), for which the leading discretisation errors are proportional to $a^2$ \cite{13}. Points to the left of the dashed line are the continuum extrapolations of the corresponding data points.

The (dimension 5) term in the action,
\[
c_{SW} \frac{a^5}{4} \sum_x \bar{q}(x)i\sigma_{\mu\nu}F_{\mu\nu}(x)q(x),
\]
together with explicit linear quark-mass dependence and mixing of the composite fields,
\[
O^R = Z_{O}(1 + b_O am_q)(O + \sum_n c_n aO_n),
\]
used in correlation functions, such as Eq (4), (5) and (9). The improvement coefficients ($c_{SW}$, $b_O$, $c_n$, etc) may be determined non-perturbatively, eg by imposing symmetries broken by the lattice.

An example of how effective improved actions can be is shown in Fig 2. This plots the vector meson mass, $m_V$, at fixed quark mass (defined by holding the ratio of the pseudoscalar and vector meson masses constant at 0.7), in units of the string tension, $\sqrt{\sigma}$, as a function of the square of the lattice spacing, in the same units, so that the data from improved simulations should fall on straight lines. The data labelled ‘Wilson Quarks’ is unimproved, showing curvature due to the $O(a)$ errors, and deviating significantly from the continuum limit, $a = 0$, as $a$ increases. The two sets of improved data, labelled ‘Improved SW Quarks’, which employ different choices of lattice gluon action, are close to the continuum limit even at large lattice spacings. The small remaining dependence on $a$ can be confidently extrapolated away. (The other data in this plot refer to different choices of lattice action with various degrees of improvement.) The extra computational cost of using an improved formulation is usually around 20%, so the efficiency gain is very significant.
2.6 The lattice theorist’s toolkit

In summary, today’s lattice theorist employs the following range of tools.

A supercomputer is needed to generate gauge field configurations and to calculate correlation functions, particularly if the objective is phenomenology, as this often requires large lattices to achieve adequate momentum resolution in the calculation of form factors.

Chiral perturbation theory is needed to guide the extrapolation of simulation results to light quark masses. The high cost of simulating light quarks directly seems to make this extrapolation inevitable for some years yet and it is a major source of uncertainty.

Keeping the lattice spacing as large as possible, means the momentum cutoff is low and some form of extrapolation to large quark masses is needed to study $b$ physics. Heavy-quark effective theory can guide this extrapolation. Alternatively, the heavy quarks can be simulated directly using non-relativistic QCD. For more about this, see [2].

Although I have not discussed this in detail, quantities computed in lattice QCD often need to be translated into a more conventional perturbative regularisation scheme, in order for the results to be useful in phenomenology. This matching involves either a lattice perturbation theory calculation (if the momentum scale is high enough for perturbation theory to be valid), or some non-perturbative renormalisation.

Finally, the use of improved actions to reduce discretisation errors has become almost mandatory, and certainly so for simulations involving dynamical quarks.

3 THE DETERMINATION OF QUARK MASSES

The starting point for any simulation of QCD is to fix the bare quark masses which enter the Lagrangian. A further matching calculation enables these lattice quark masses to be related to quark masses defined in a perturbative regularisation scheme, so that they can be used in phenomenology.

3.1 Fixing the bare quark mass parameters

Conventionally, the bare quark mass enters the lattice QCD action via a so-called hopping parameter, $\kappa$, which is essentially its inverse. Simulations are performed for a range of hopping-parameter values in order to find the value, $\kappa = \kappa_{\text{crit}}$, which corresponds to zero quark mass (recall that chiral symmetry is broken explicitly by the lattice and so quark mass is additively renormalised). This is defined as the hopping parameter at which the pseudoscalar meson mass vanishes,

$$m_{\text{PS}}(\kappa = \kappa_{\text{crit}}) = 0,$$

(19)

corresponding to it being the Goldstone boson of spontaneously-broken chiral symmetry (although the way this is actually realised in the lattice theory is quite subtle). Chiral perturbation theory tells us how $m_{\text{PS}}$ depends on quark mass, when this is small, and so can be used to extrapolate the data to $\kappa_{\text{crit}}$. Then the bare quark mass for flavour $q$ is

$$m_q a = \frac{1}{2\kappa_q} - \frac{1}{2\kappa_{\text{crit}}},$$

(20)

where $\kappa_q$ is fixed by matching, for instance, the ratio of the pseudoscalar and vector meson masses to the experimental ratio of the masses of suitably flavoured mesons,

$$\frac{m_{\text{PS}}}{m_{\text{V}}} = \frac{M_\pi}{M_\rho}, \frac{M_K}{M_\rho}, \ldots$$

(21)
3.2 Chiral behaviour of hadron masses

Simulations close to $\kappa_{\text{crit}}$ are prohibitively costly. Instead, the simulations have to be performed at relatively large quark masses and, consequently, the chiral extrapolation becomes a significant source of error in the resulting estimates for the $u$ and $d$ quark masses, since $\kappa_{ud} \sim \kappa_{\text{crit}}$.

For full QCD, chiral perturbation theory predicts for the Goldstone pion,

$$m_{\text{PS}}^2 = Am_q + O(m_q^2) \quad (22)$$

and for all other hadrons,

$$m_{\text{hadron}} = B + Dm_q + O(m_q^2), \quad (23)$$

where $A$, $B$ and $D$ are constants.

However, the behaviour of hadron masses is different in quenched QCD. This is because the absence of quark loops eliminates the infinite series of loop diagrams which gives the $\eta'$ meson its large mass. Instead, the $\eta'$ meson remains light and, in fact, the singlet two-point function has a double pole, which introduces logarithmic terms in the quark mass [14], namely

$$m_{\text{PS}}^2 = Am_q \{1 - \delta \log m_q\} + O(m_q^2) \quad (24)$$

$$m_{\text{hadron}} = B + C\delta m_{\text{PS}} + Dm_{\text{PS}}^2 + O(m_{\text{PS}}^3), \quad (25)$$

where $\delta$ parametrises the strength of these quenched chiral logarithms. The simulation data from the CP-PACS Collaboration [15], which represents the state of the art for the quenched hadron spectrum, are consistent with the presence of quenched chiral logarithms, although, as can be seen in Fig 3, the differences from ordinary chiral perturbation theory only set in at very small quark masses.

3.3 Light quark masses

Since electromagnetic effects are not included in the QCD simulations, the up and down quarks are degenerate in mass. Recent quenched results for their mass, $m_{ud}$, and the mass of the strange
Fig. 4: Quenched QCD data for the light quark masses, obtained by the CP-PACS Collaboration [15] at a range of lattice spacings, and using two different definitions for the renormalised mass: the Vector Ward Identity (VWI) and the Axial Vector Ward Identity (AWI). In all but the top set of data, quenched chiral perturbation theory (qχPT) was used for the chiral extrapolation.

Quark, $m_s$, in the $\overline{\text{MS}}$ scheme at a scale of 2 GeV, are shown as functions of the lattice spacing in Fig 4. There are two different definitions of the renormalised quark mass, depending on which of two Ward Identities is used. Although these give different results at non-zero lattice spacing, both definitions have consistent continuum limits. The result for the quenched QCD up and down quark mass is [15]

$$m_{ud} = 4.6(2) \text{ MeV}.$$  

(26)

The results for the strange quark mass show the first evidence of a breakdown of the quenched approximation. The two sets of data in Fig 4 have been obtained by using the experimental $K$ and $\phi$ meson masses to fix the bare strange quark mass, and it is clear that they do not give the same continuum limit. The reason is simply that the quenched strange meson spectrum does not agree with experiment for any choice of the strange quark mass. The continuum results are [15]

$$m_s = \begin{cases} 143(6) \text{ MeV} & (\phi \text{ input}) \\ 115(2) \text{ MeV} & (K \text{ input}) \end{cases}. $$  

(27)
The next step is to simulate QCD with two degenerate flavours of quark. Since the gluon configurations ‘feel’ the quarks through the determinant term in Eq (15), separate Monte Carlo runs have to be performed for each value of the quark mass used as input. I’ll call this the sea-quark mass. On each set of configurations, we can compute hadron correlation functions built from valence quark fields with a different mass from the sea-quark mass, if we so choose. The two degenerate sea-quark flavours have to be identified with the up and down quarks, at least in the limit of small mass, but the valence quarks can be used to approximate other quenched flavours, such as strange or charm, propagating in gluon fields which are only sensitive to the up and down quarks. The chiral limit may be approached in two ways: by maintaining $\kappa_{\text{valence}} = \kappa_{\text{sea}}$, or by fixing $\kappa_{\text{sea}} = \kappa_{udd}$ and varying $\kappa_{\text{valence}}$, called ‘partial quenching’. Different values of $\kappa_{\text{crit}}$ result from these two approaches, at non-zero lattice spacing, and this mostly affects the estimate of the up and down quark mass, but the differences appear to go away in the continuum limit.

Some of the best data obtained so far is shown in Fig 5. Even so, the data are not of sufficient quality to justify a continuum extrapolation. The trends are, however, encouraging. Although significant differences between the various definitions are found at non-zero lattice spacing, all these differences decrease as the continuum is approached. Secondly, the strange quark mass estimates obtained from the $K$ and $\phi$ mesons appear to be converging in the con-
tinuum limit, unlike in quenched QCD. Finally, and most interesting of all, the quark mass estimates are significantly lower than those obtained in the quenched approximation. At a time when it is proving difficult to establish clear effects from including dynamical quarks in our simulations, this is a welcome result!

While the strange quark mass estimates have dropped from around 130 MeV to around 80 MeV, by including dynamical up and down quarks, it should be noted that the latter result still corresponds to a quenched strange quark and so may be an upper bound on the real value.

4 THE QCD SPECTRUM

Attempts to compute the spectrum of QCD began with Wilson’s invention of lattice QCD, and represent an essential step in validating QCD. The whole exercise has proved to be much more difficult than originally hoped. This is because it requires demonstrating complete control of all aspects of the simulation. In fact, the quenched QCD light hadron spectrum, which was tackled first for the reasons given above, has turned out to be so close to the real world that, just four years ago, the GF11 Collaboration claimed it agreed with experiment to within 6% [16]. But, we know quenched QCD must be wrong at some level!

4.1 The quenched spectrum and decay constants

It has taken the last four years, and an enormous computational effort, to show that the quenched QCD spectrum does deviate significantly from experiment. This was achieved by the CP-PACS Collaboration [15] and I have already described the inconsistencies they found in the quenched strange meson spectrum. Although it awaits independent confirmation, the CP-PACS tour de force has effectively concluded a 20-year effort to calculate the quenched QCD spectrum.

CP-PACS have obtained high precision results with control of all systematic errors (except quenching itself) on lattices up to $64^3 \times 112$. They perform a continuum extrapolation from data at four lattice spacings, spanning the range $a^{-1} \approx 2 - 4$ GeV, and in a fixed volume of linear size $L \approx 3$ fm. Although not proving the existence of quenched chiral logarithms, their data are consistent with this pathology of quenched QCD and their chiral extrapolations use the quark-mass behaviour predicted by quenched chiral perturbation theory. They are able to conclude that the spectrum is inconsistent with experiment.

The spectrum obtained by CP-PACS is shown in Fig 6, along with the GF11 results for comparison. It is evident that the major achievement has been to reduce the errors to the point where disagreement with experiment, within the uncertainty in the GF11 results, could be exposed. As I have already mentioned, the strange quark mass cannot be fixed unambiguously and, hence, CP-PACS present two spectra in Fig 6, corresponding to two possible choices. Generally, the meson hyperfine splitting is too small in the quenched approximation and it gets worse for charmonium. The octet baryon masses and the decuplet baryon mass splittings are also too small, with $M_K$ as input, although somewhat better with $M_\phi$ as input.

As we have seen, decay constants are a by-product of spectrum calculations. They are also sensitive to quenched chiral logarithms and CP-PACS’s decay-constant data also support the presence of chiral logarithms, with a similar strength, $\delta$ in Eq (24), to the spectrum data. The continuum results for $f_\pi$ and $f_K$ in quenched QCD are significantly smaller than experiment, as shown in Fig 7.

4.2 The QCD spectrum with two degenerate flavours

Present-day simulations with dynamical up and down quarks, and all other flavours quenched, have not reached anything like the same degree of control as has been achieved for fully quenched simulations. The lattices are typically smaller, with $L \sim 2$ fm, although this is not a serious
Fig. 6: CP-PACS Collaboration’s results for the quenched light hadron spectrum in the continuum limit [15]. Experimental values are denoted by horizontal lines and the results from GF11 [16] are shown for comparison.

Fig. 7: CP-PACS Collaboration’s results for the quenched $K$ meson and pion decay constants in the continuum limit [15]. Experimental values are denoted by diamonds.
problem, because the up and down quark masses actually used are large (so that $m_{PS}/m_{V} \geq 0.6$) and the hadrons are small. Consequently, a big chiral extrapolation is necessary, from data of relatively poor statistical quality. So it is not particularly surprising that the spectrum obtained agrees with experiment within large errors [15]. It will be a considerable challenge to expose any effects from quenching the heavy quarks, given the effort needed to disprove quenched QCD. The most dramatic result from dynamical-quark simulations so far has been the drop in the quark masses.

4.3 Hybrid mesons
From the point of view of phenomenology, a much more interesting aspect of the hadron spectrum calculations is the ability to predict states which are not found in the quark model and which have yet to be identified experimentally.

Hybrid mesons have gluonic excitations which permit non-quark-model quantum numbers. A while ago, the UKQCD Collaboration showed that, in quenched QCD, the lightest $s\bar{s}$ exotic meson has $J^{PC} = 1^{--}$ and a mass of 2.0(2) GeV [17].

This year, the calculation has been repeated for $N_f = 2$ sea quarks. Again, the $1^{--}$ meson is the lightest exotic and its mass is found to be 1.9(2) GeV in the chiral limit [18], consistent with the UKQCD result. Of course, hybrid mesons can mix with four-quark states, and the effect of this must be understood before the result can be considered relevant to phenomenology. Mixing in the lattice simulation is currently under investigation.

4.4 Glueballs
Glueballs are an important prediction of QCD and the low-lying quenched spectrum has been known for some time [19]. This is already a useful guide for experimental searches [20].

Although there are no results from dynamical-quark simulations yet, the mixing of the lightest glueball with scalar quarkonium states has been studied recently in quenched QCD [21]. The unmixed results are shown in Fig 8. The continuum limit of the $s\bar{s}$ state has a mass well below the lightest glueball, suggesting naively that, of the two experimental candidates, the $f_0(1710)$, rather than the $f_0(1500)$, is predominantly glue.

A crude mixing calculation gives

$$|f_0(1710)\rangle = 0.86(5)|g\rangle + 0.30(5)|s\bar{s}\rangle + 0.41(9)|n\bar{n}\rangle$$

$$|f_0(1500)\rangle = -0.13(5)|g\rangle + 0.91(4)|s\bar{s}\rangle - 0.40(11)|n\bar{n}\rangle$$

$$|f_0(1390)\rangle = -0.50(12)|g\rangle + 0.29(9)|s\bar{s}\rangle + 0.82(9)|n\bar{n}\rangle$$

where $|n\bar{n}\rangle$ is the scalar state built from up and down quarks. The opposite sign of $|s\bar{s}\rangle$ and $|n\bar{n}\rangle$ in $|f_0(1500)\rangle$ is a possible explanation of why its decay to $K\bar{K}$ is suppressed.

5 CONCLUSIONS
I hope I have given you an impression of the challenges and promises of lattice QCD. Ultimately, lattice QCD offers the prospect of a precise, model-independent tool for dealing with the hadronic uncertainties which currently hinder the search for physics beyond the Standard Model at B Factories. To date, quenched QCD has proven to be a surprisingly accurate effective theory of the strong force, but, as a result of arduous and painstaking numerical work, its failure even at the few percent level is now established. The quenched light hadron spectrum has been shown to disagree with experiment and, in particular, the strange quark mass cannot be chosen unambiguously. Simulations with dynamical quarks are underway and sea-quark effects are beginning to show up. Notably, the up, down and strange quark masses are around 40% lower.
than in quenched QCD. Continued progress is guaranteed by the relentless advance of computer technology. However, a speedier approach to fully realistic QCD simulations can be anticipated due to further developments in the theoretical formulation, and the range of non-perturbative physics which can be addressed will grow accordingly.

References


