



SIMULATING QUARKS

By Michael Creutz

OVER THE PAST 25 YEARS, THE THEORETICAL-PARTICLE PHYSICISTS WHO PRACTICE LATTICE-GAUGE THEORY HAVE BECOME SOME OF THE MOST FREQUENT USERS OF supercomputing cycles—despite much of the community historically regarding the use of computation as somehow intellectually inferior to pure thought.

Although we've long known that atoms consist of electrons surrounding a nucleus made of nucleons (protons and neutrons), we've recently learned that at a deeper level, the nucleons themselves are composites. We can best explain the strong forces between them by assuming they are composed of three quarks interacting via fields called *gluons*. The need for three constituents helps explain much of the zoo of similar states seen in particle physics experiments. Understanding these forces is one of the main goals of particle physics. Certain intractable aspects of the interactions between quarks and gluons have driven us to the computer. Indeed, large-scale simulations have helped us make major inroads into issues highly resistant to traditional approaches. We can't predict quantities such as the transition to a quark-gluon plasma at 100 million Kelvins in a controllable way by any other means.

Lattice-gauge theory provides a controlled scheme for studying strong interactions at low energies. In this article, we'll see that the main tools are powerful but demanding algorithms (such as conjugate-gradient sparse-matrix inversions) familiar in other fields. Still-unsolved issues involve the "sign" problem and the basic formulation of parity violation on the lattice. Although we certainly need additional computing capability, we also need new ideas.

Seeing Is Believing

Why do we believe in quarks in the first place? We've never seen them isolated; they don't fit our naive concept of an elementary particle being something that can propagate over long distances through space. But several observed facts have forced us to believe that quarks do have a fundamental mean-

ing. The first arises in the basic spectrum: observed particles appear in families corresponding to combining different types, called *flavors*, of quarks. History has given rather frivolous names to quark types—up, down, and strange for the three lightest ones. There are 10 ways to combine these three quark types into a bound state. The corresponding states form a *decouplet*, all members of which have been found in experiment. The proton containing two up quarks and one down, and the neutron containing two down quarks and one up are members of a different multiplet of eight (rather than 10 because some combinations are removed for symmetry reasons).

A second piece of evidence for quarks appears in the excited states of elementary particles. These excitations correspond to bound quarks rotating around each other with increasing angular momentum. Indeed, from the observed spectrum, the quarks seem to be held together by something like a string with a 14-ton tension. In the quark-gluon theory, this string's origin is the gluonic field, which I'll discuss shortly.

A third reason to believe in quarks comes from the high-energy scattering of electrons (or muons or neutrinos) on protons. In this observation, the electron appears to be a point-like object, with a radius of less than 10^{-16} cm, but the proton has a measured radius on the order of 10^{-13} cm. Very high-energy electrons often scatter at rather large angles, like a bullet bouncing off a paper bag—obviously, something hard must be inside the bag. In the proton case, the contents are the quarks themselves.

Finally, evidence has appeared over the years for several rather exotic heavier quarks, including the—again, frivolously named—charm, bottom, and top quarks. Their masses range widely, from about 1.5 times the proton mass for the charm quark up to 185 proton masses for the top quark. These heavier quarks move relatively slowly in their bound states, making accurate predictions possible for a large spectrum of excited states. The stunning agreement of these predictions with experimental observation has made the concept of quarks incontrovertible.

Gluons

If quarks are real, how is it that we've never seen them iso-

lated? The answer might lie in the rather special properties of their interactions via the exchange of gluons. The quark-confining dynamics of QCD (which stands for quantum chromodynamics, a confusing name for something that has nothing to do with the 5,000-Angstrom scale of visible colors) proceeds in close parallel to electrodynamics—but with some elegant embellishments.

First, unlike the unique photon of electrodynamics, the strong interactions involve eight distinct gluons: free particles such as the proton contain three quarks—a concept intricately tied mathematically with the eight generators of the group $SU(3)$, the set of three by three unitary matrices of unit determinant. This group's structure is crucial to enabling bound states of three quarks, such as in the proton. Second, unlike the neutral photon of electrodynamics, these eight gluons are charged with respect to each other. In the basic picture of confinement, the quarks are the sources for gluonic fields, but the gluon self-charges to prevent these fields from spreading in electricity's familiar $1/r^2$ manner. Instead, the lines of electric flux form themselves into *flux tubes* (see Figure 1). These flux tubes are real objects with an energy-per-unit length representing the 14-ton tension holding the quarks together. We call this formation the *confinement phenomenon*, which early lattice-gauge simulations rather convincingly demonstrate.

The main theoretical difficulty concerning quark confinement is that we can't see it in an expansion of the quark-gluonic charge. When the gauge coupling is turned off, we have a bunch of free quarks running around; when the coupling is turned on, the theory is qualitatively different, with only quarks bound into the physical hadrons (those particles that feel the strong nuclear force) allowed to move long distances. The traditional Feynman diagram methods are restricted to phenomena at short distances, so they aren't applicable here.

Lattice-Gauge Theory

The lattice approach to quantum field theory is much deeper than a mere calculation tool: it provides a fundamental regularization scheme for defining a general renormalizable field theory. As with all known field theories in four dimensions, the basic interactions of quarks and gluons immediately show the famous ultraviolet divergences that must be renormalized. Traditional renormalization schemes involve diagrammatic expansion: we calculate Feynman diagrams until we find an infinite factor, at which point we must control the infinity with some regularization scheme. However, Feynman diagrams are inherently

Recommended Reading

You can find a simple introduction to the basics of lattice-gauge theory in my book, *Quarks, Gluons, and Lattices* (Cambridge, 1983); you'll find a more modern and thorough discussion of the subject in I. Montvay and G. Münster's *Quantum Fields on a Lattice*, (Cambridge, 1997).

For a retrospective on the early developments in the field, see my article, "The Early Days of Lattice Gauge Theory," *The Monte Carlo Method in the Physical Sciences*, J. Gubernatis, ed., Am. Inst. of Physics, 2003; <http://arxiv.org/abs/hep-lat/0306024>.

The proceedings of annual lattice-gauge conferences contain up-to-date reviews of the field's status: two excellent examples are *Nuclear Physics B Proc. Supplement*, vol. 119, 2003, and vol. 106, 2002; <http://arxiv.org/html/hep-lat/0203004/>.

A good reference on the Z_2 experiments appears in M. Creutz, L. Jacobs, and C. Rebbi, "Experiments with a Gauge Invariant Ising System," *Physical Rev. Letters*, vol. 42, no. 21, 1979, pp. 1390–1393.

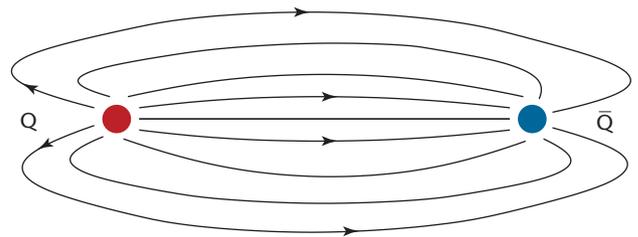


Figure 1. A flux tube. The gluonic fields connecting a quark to an antiquark arrange themselves in a flux tube that gives rise to an energy that linearly increases with the separation of the quarks.

perturbative, meaning they're based on a power expansion for small forces. To study nonperturbative phenomena such as confinement, we need a nonperturbative cutoff. Lattice-gauge theory provides just what we need.

A lattice has a minimum wavelength, given by the lattice spacing a . This is the lattice regulator's cutoff parameter, which we should extrapolate to zero at the end of any calculation. A lattice is just a mathematical trick that defines things further; this idea contrasts with solid-state physics, in which nature provides a physical underlying lattice at the atomic scale.

Gluons are gauge fields, a rather elegant concept with many definitions. The lattice approach is based on a gauge theory as a theory of phases. As a quark propagates through space-time, its wave function picks up a factor

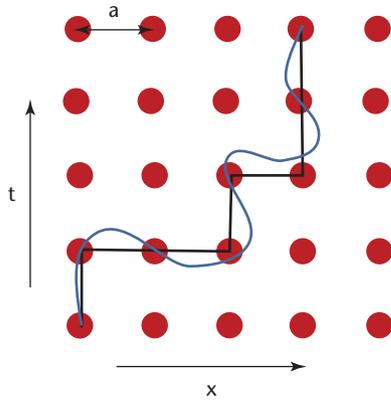


Figure 2. A quark’s motion through space-time as approximated by a sequence of discrete hops. On each hop, the quark wave function picks up a “phase” described by the gauge fields. For the strong interactions, this phase is a unitary matrix in the group $SU(3)$.

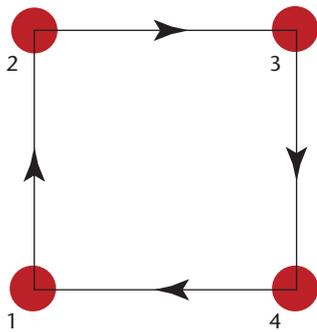


Figure 3. An elementary square, or *plaquette*. In analogy with Stoke’s law, we can find the flux through an elementary square of the lattice from the product of gauge matrices around that square. We determine the dynamics by adding the real part of this product’s trace over all elementary squares, and then inserting this “action” into a path integral. The resulting construction is formally a partition function for a system of “spins” existing in the group $SU(3)$.

from its interactions with gluons. In electrodynamics, the corresponding factor is a phase in the electron wave function. For the non-Abelian generalization with our eight gluons, the phase becomes a matrix from the group $SU(3)$; specifically, it becomes a 3×3 unitary matrix of unit determinant. The quark field has an internal degree of freedom, usually (again, misleadingly) called color, which introduces the three-valued index needed for multiplying by this matrix.

This concept leads to the basic method for formulating quark and gluon fields on a lattice. We approximate a general quark-world line by a set of jumps or *hoppings* of the

quark between nearest-neighbor sites along the lattice bonds (see Figure 2). We then introduce the gauge field as matrices on these bonds; these fields form a set of $SU(3)$ matrices, one such associated with every nearest-neighbor bond on our four-dimensional hypercubic lattice.

In terms of these matrices, gauge self-interactions take a simple, natural form. Similar to regarding electromagnetic flux as the vector potential’s generalized curl, we are led to identify the flux via an elementary square, or *plaquette* (see Figure 3). We work in four-dimensional space-time, giving rise to two types of plaquette. Those with sides in ordinary three-dimensional space are called “spatial” and represent magnetic effects. When two opposite sides are in the time direction, the plaquette is called *timelike* and represents the electric effects. This motivates the conventional “action” used for the gauge fields as a sum over all the lattice’s elementary squares. Around each square, we multiply the phases; to get a real number, we take the real part of the trace

$$S_g = \sum_p \text{Re Tr} \prod_{l \in p} U_l,$$

where the fundamental squares are denoted with p and the links with l . Because we’re dealing with noncommuting matrices, the product around the square is meant to be ordered.

To formulate a quantum theory of these variables, we would use the Feynman path integral. For this, we exponentiate the action and integrate over all dynamical variables to construct

$$Z = \int (dU) e^{-\beta S},$$

where the parameter β controls the bare coupling. Doing so converts the three-dimensional quantum field theory of gluons into a classical statistical mechanical system in four space-time dimensions. Such a many-degrees-of-freedom system cries out for Monte Carlo simulation, which now dominates the field of lattice QCD. Note the close analogy with a magnetic system; we can think of our matrices as “spins” interacting through a four-spin coupling expressed in terms of the plaquettes.

The usual lattice-gauge formulation is in Euclidian four-dimensional space, based on an underlying replacement of the time-evolution operator e^{-iH} with e^{-H} . Despite involving the same Hamiltonian, excited states are inherently suppressed, and extracting information on high-energy scattering is particularly difficult. However, low-energy states and matrix elements are the natural physical quantities to explore numer-

The Z_2 Theory

A lattice-gauge program's basic structure is quite simple, starting with nested loops over all the lattice's links. Inside the innermost loop is a probabilistic step that determines the new value for the dynamical link variable. The simple lattice-gauge model based on the group $Z_2 = \{\pm 1\}$ provides an amusing exercise for getting familiar with the concept. This model has a strong first-order phase transition, which is easy to see even on quite small lattices. The accompanying program listing presents a 150-line (with comments) Z_2 lattice-gauge program, on the basis of which we can perform a variety of experiments.

A Simple Experiment

Perhaps the simplest first experiment is a thermal cycle. For this, order the system with all links initially set to unity, and then run the update procedure starting at, say, $\beta = 1$ and gradually reducing it to $\beta = 0$. Then, we can return to the initial β and expose a strong hysteresis effect.

For this model, the transition temperature is analytically known:

$$\beta_t = \frac{1}{2} \log(1 + \sqrt{2}).$$

If we run at this beta with an initially ordered or initially disordered start, with each link randomly ± 1 , then the runs don't converge to each other in an observable time: the system remains in the corresponding phases. In particular, the average energies observed in the runs remain quite different.

To explore the different phases, the concept of a Wilson loop is useful. Multiply the links around a large closed loop:

the average value of this product decreases with the loop size. For the small β phase, this falloff is exponential with the loop area, whereas in the large β phase, the decrease is exponential only in the loop parameter.

To verify the transition temperature, we can start with a mixed state—for example, we might have the first third of the lattice random and the remainder ordered. Just above or below the transition temperature, we can watch the appropriate phase come to dominate.

With slight variations on the program, we can change the physical dimension. With two space-time dimensions, the model has no transition. In three, the model can be related via duality to the Ising model, which has a second-order transition. In four or more dimensions, we have the situation of a strong first-order transition.

Being a rather simple model, this system requires no particular tricks to program. Nevertheless, you might enjoy attempting to store the links not as integers, but as single bits in computer words. With this arrangement, you can update several links in parallel by using logical operations.

A Simple Z_2 Program

The following listing shows a lattice-gauge program's simple structure as a set of nested loops over the links of a hypercubic lattice. The example is based on the group Z_2 , the group of two elements $\{+1, -1\}$.

```
/* a Z_2 lattice gauge simulation      */
/* Michael Creutz <creutz@bnl.gov>   */
/* http://thy.phy.bnl.gov/~creutz/z2 */
```

continued on p. 6

ically. Such studies are the lattice theorist's bread and butter.

Using Algorithms

Restricted to gauge fields alone, the Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller algorithm provides an excellent framework for lattice-gauge simulations (see the sidebar). With the growth in small computers' power, we don't even need supercomputing to study extensively this restricted version of the theory. Difficulties arise only with the introduction of anticommuting quark fields.

A pure gauge simulation provides a set of background gauge fields on top of which we can study quark propagation. Until relatively recently, most lattice work with quarks was in the so-called *valence*, or *quenched*, approximation. (The approximation is to ignore any feedback from the quarks on the gauge fields.) Because quarks involve large sparse matrices, the conjugate-gradient algorithm is ideally suited. Combining the resulting propagators into hadronic combinations gives predictions on physical quantities such as spectra, matrix elements, and so on. The rather random

nature of the relevant background fields has hampered the application of standard multiscale techniques; more work in this area is needed.

A dramatic increase in computational complexity occurs when we try to include the quarks' dynamical effects back in the gauge fields. Quarks are fermions, which means they should satisfy the Pauli exclusion principle (two quarks cannot simultaneously occupy the same state). To implement this principle mathematically, we treat the quark fields as anticommuting variables. Their interactions involve operators in a Grassmann space, not ordinary numbers, so the standard approach replaces the quarks with an integration over a new set of commuting variables that interact with the inverse of the quark propagator. This brings us back to a classical statistical mechanics problem, which now involves a computationally intensive nonlocal interaction. We can apply well-developed hybrids of molecular dynamics and Monte Carlo simulations to the resulting system, but as a rule of thumb, doing so requires about two orders of magnitude more computer time than the valence approximation.

```

continued from p. 5
#include <stdio.h>
#include <stdlib.h>
#include <math.h>

/* the lattice is of dimensions SIZE**4 */
#define SIZE 6
int link[SIZE][SIZE][SIZE][SIZE][4]; /* last
index gives link direction */

/* utility functions */
void moveup(int x[],int d) {
    x[d]+=1;
    if (x[d]>=SIZE) x[d]-=SIZE;
    return;
}
void movedown(int x[],int d) {
    x[d]-=1;
    if (x[d]<0) x[d]+=SIZE;
    return;
}
void coldstart(){/* set all links to unity */
    int x[4],d;
    for (x[0]=0;x[0]<SIZE;x[0]++)
        for (x[1]=0;x[1]<SIZE;x[1]++)
            for (x[2]=0;x[2]<SIZE;x[2]++)
                for (x[3]=0;x[3]<SIZE;x[3]++)
                    for (d=0;d<4;d++)
                        link[x[0]][x[1]][x[2]][x[3]][d]=1; return;
}
/* for a random start: call coldstart() and

```

```

then update once at beta=0 */
/* do a Monte Carlo sweep; return energy */
double update(double beta){
    int x[4],d,dperp,staple,staplesum;
    double bplus,bminus,action=0.0;
    for (x[0]=0; x[0]<SIZE; x[0]++)
        for (x[1]=0; x[1]<SIZE; x[1]++)
            for (x[2]=0; x[2]<SIZE; x[2]++)
                for (x[3]=0; x[3]<SIZE; x[3]++)
                    for (d=0; d<4; d++) {
                        staplesum=0;
                        for (dperp=0;dperp<4;dperp++){
                            if (dperp!=d){
                                /* move around thusly:
                                    dperp      6--5
                                    ^          | |
                                    |          1--4
                                    |          | |
                                    ---> d      2--3   */
                                /* plaquette 1234 */
                                movedown(x,dperp);
                                staple=link[x[0]][x[1]][x[2]]
[x[3]][dperp]
                                *link[x[0]][x[1]][x[2]]
[x[3]][d];
                                moveup(x,d);
                                staple*=link[x[0]][x[1]]
[x[2]][x[3]][dperp];
                                moveup(x,dperp);
                                staplesum+=staple;

```

This avenue of attack represents the current state of the art, driving the lattice-gauge community's voracious appetite for computer cycles.

The Road Ahead

Several interesting technical issues remain. Because the fermionic matrices become poorly conditioned as the quark masses are reduced, the essential conjugate-gradient steps converge slowly. To circumvent this, most simulations rely on quark mass extrapolations, assisted by chiral symmetries. Such simulations require underlying lattice actions that possess as much of these symmetries as possible. Much recent work is concentrated in this area, with researchers finding actions that give great accuracy, but at a penalty of an additional one to two orders of magnitude in computer time.

Going beyond this standard approach, several algorithmic obstacles further hamper the application of lattice methods to a variety of tantalizing problems. The basic Monte Carlo approach has a "sign problem" whenever the quark matri-

ces' determinant is not positive. Shuffling the signs into observables gives rise to theoretically valid algorithms, but requires computer time that grows exponentially with volume, thus precluding simulation on any but the smallest systems.

These issues become particularly severe in some of the more interesting physical situations. The most notorious involves the quark-gluon plasma in the presence of an excess of matter over antimatter, also called background baryon density. Recent progress has shown the feasibility of simulations for small density, allowing studies of the critical point at the end of a first-order line separating the quark-gluon plasma from low-temperature hadronic matter. For much higher densities, though, theoreticians have predicted fascinating superconducting phases at low temperature. Equivalent mathematical problems occur in attempts to simulate doped strongly correlated electron systems. Numerical algorithms to simulate such phenomena are not known, but they would be highly desirable. Frustration with the standard approaches suggests we might try to work independently from Monte Carlo—for example, by using exact di-

```

        /* plaquette 1456 */
        staple=link[x[0]][x[1]]
        [x[2]][x[3]][dperp];
        moveup(x,dperp);
        movedown(x,d);
        staple*=link[x[0]][x[1]]
        [x[2]][x[3]][d];
        movedown(x,dperp);
        staple*=link[x[0]][x[1]]
        [x[2]][x[3]][dperp];
        staplesum+=staple;
    }
}
/* calculate the Boltzmann weight */
bplus=exp(beta*staplesum);
bminus=1/bplus;
bplus=bplus/(bplus+bminus);
/* the heatbath algorithm */
if (drand48() < bplus){
    link[x[0]][x[1]]
    [x[2]][x[3]][d]=1;
    action+=staplesum;
}
else{
    link[x[0]][x[1]]
    [x[2]][x[3]][d]=-1;
    action-=staplesum;
}
}
action/=(SIZE*SIZE*SIZE*SIZE*4*6);
/* normalize the action */

return 1.-action;
}
/*****/
int main(){
    double beta, dbeta, action;
    srand48(1234L); /* initialize random number
    generator */
    /* do your experiment here; this example is
    a thermal cycle */
    dbeta=.01;
    coldstart();
    /* heat it up */
    for (beta=1; beta>0.0; beta-=dbeta){
        action=update(beta);
        printf("%g\t%g\n",beta,action);
    }
    printf("\n\n");
    /* cool it down */
    for (beta=0; beta<1.0; beta+=dbeta){
        action=update(beta);
        printf("%g\t%g\n",beta,action);
    }
    printf("\n\n");
    exit(0);
}

```

agonalization methods—but such approaches remain in their infancy.

A more fundamental issue arises for lattice theories of the standard model when parity violation is included. No known lattice formulation of left-handed neutrinos has been found that does not introduce mirrored right-handed particles. This is troubling because the lattice is the most precise way to define a quantum field theory nonperturbatively. This is not a practical problem for experimental predictions because weak interactions in most cases can be handled perturbatively. Nevertheless, from a conceptual viewpoint, the absence of a lattice formulation for the full standard model might point to deep issues that are as yet not understood.

Acknowledgments

This manuscript was authored under contract number DE-AC02-98CH10886 with the US Department of Energy. Accordingly, the US Government retains a nonexclusive, royalty-free license to publish or reproduce the published form

of this contribution, or allow others to do so, for US Government purposes.

Michael Creutz is a senior physicist at Brookhaven National Laboratory. His technical interests include theoretical particle physics and computational physics. He has a PhD in physics from Stanford. He is a member of the American Physical Society. Contact him at creutz@bnl.gov.