RAL-86-103

Science and Engineering Research Council

Rutherford Appleton Laboratory

CHILTON, DIDCOT, OXON, OX11 0QX

RAL-86-103

Topological Fluctuations and Susceptibility in SU(3) Lattice Gauge Theory

Jaap Hoek, M Teper and J Waterhouse

December 1986

Science and Engineering

Research Council

'The Science and Engineering Research Council does not accept any responsibility for loss or damage arising from the use of information contained in any of its reports or in any communication about its tests or investigations.'

TOPOLOGICAL FLUCTUATIONS AND SUSCEPTIBILITY
IN SU(3) LATTICE GAUGE THEORY

Jaap Hoek

Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, UK

M Teper^{*}

Department of Theoretical Physics, University of Oxford
Oxford OX1 3NP

J Waterhouse

Physics Department, University of Southampton
Southampton SO9 5NH

Abstract

We measure the topological charge density of the (Monte Carlo generated) SU(3) vacuum. Our algorithm involves first smoothening the generated configurations in a way that conserves the total charge in the continuum limit. We work in a range of couplings $5.6 < \beta \leq 6/g^2 < 6.0$ and on lattice sizes up to 16^4 . We find that the topological susceptibility scales like the string tension (within errors). This allows us to extract a value in physical units, $\chi_t^{1/4} = 191 \pm 16$ MeV, in good agreement with the Witten-Veneziano mass formula. We also show how χ_t is strongly suppressed as the temperature is increased through the deconfining transition, so that the $U_A(1)$ symmetry is effectively restored in the deconfined phase. Our results leave open the possibility that χ_t is an order parameter for this transition. We carefully monitor finite size effects and perform a variety of calculations - reproducibility; stability under Monte Carlo; distribution of core sizes; etc. - which give us a direct insight into the approach to continuum topology.

^{*} Research Fellow at All Souls College, University of Oxford

1. Introduction

Vacuum topological [1] fluctuations break [2] the axial U(1) symmetry of the QCD Lagrangian in the massless quark limit:

$$\partial_\mu j_\mu^5 = -2Q(x) , \quad (1.1)$$

where j_μ^5 is the axial U(1) current and $Q(x)$ is the topological charge density

$$Q(x) = - \frac{1}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} \text{tr}[F_{\mu\nu}(x) F_{\rho\sigma}(x)] . \quad (1.2)$$

The density of these fluctuations, and hence the strength of the symmetry breaking, is given by the topological susceptibility

$$\chi_t = \frac{1}{\text{vol. space-time}} \langle (\int Q(x) d^4x)^2 \rangle . \quad (1.3)$$

Experimentally, the strength of this breaking is indicated by how far the η' is from being a (near-) Goldstone boson. Approximate arguments lead to a mass formula [3]

$$\chi_t = \frac{f_\pi^2}{4N_f} (m_\eta^2 - 2m_K^2 + m_\eta^2) , \quad (1.4)$$

near the chiral limit. This relation holds in the continuum in the large number of color limit or on the lattice in the quenched approximation. Taking 3 light flavours ($N_f=3$; already implicit in our choice of η') we obtain a constraint

$$\chi_t = (180 \text{ MeV})^4 , \quad (1.5)$$

which needs to be satisfied if topology is to provide a quantitative resolution of the $U_A(1)$ problem. Since the topological susceptibility in eqns (1.4-5) is that of the pure SU(3) gauge theory, lattice Monte Carlo techniques would seem appropriate for testing this mass formula.

In the next section a number of methods will be mentioned that have been used to tackle this problem, and some issues on the vacuum topology raised. Section 3 contains a description of the cooling method and its motivation. In section 4 we present our results at zero physical temperature, followed by a discussion in section 5. Section 6 contains our results at finite temperature, and includes discussion thereon. Some more general discussion is contained in section 7. Parts of the calculation have been published before [4].

2. Vacuum topology

The first attempts [5] to test the Witten-Veneziano mass formula in QCD regularized on the lattice used operators that in the continuum limit reduce to the form (1.2). On the lattice however such operators receive non-topological contributions which, despite initial hopes, cannot be separated from the topological contribution of interest when applied directly to the Monte Carlo generated configurations at presently accessible couplings. A substantial improvement was made by Woit [6] who used a method that always leads to an integer charge. Originally these calculations were performed in $SU(2)$ but recently they have been extended to $SU(3)$ [7]. Also other methods have been developed to assign integer valued charges to Monte Carlo generated configurations [8,9]. For example, Luscher's definition [9], using an interpolating field, has been used in $SU(2)$ calculations [8]. Luscher's definition was the first geometric definition of topological charge on the lattice; a direct implementation is rather slow and has only after some explicit integrations had been carried out been used in the $SU(2)$ case. The fact however that one can assign an integer charge to a configuration does not imply that one is actually capturing a physical property of the configuration. The charge will in general depend on the operator one uses, or for some algorithms on the way one constructs auxiliary quantities. This is partly due to localized non-physical contributions to which we shall come back in other sections and below. No method can make sense for very strong coupling, where the fields are wildly fluctuating. Most methods in use allow one to investigate this dependence, eg by choosing the time direction in Woit's method or the way one "completes" a hypercubical lattice in Phillips and Stone's method [10]. Also with the method we will use one has a handle on this problem, namely independence of the way one goes through the lattice in relaxing links. We will come back to this in section 5.

A topological fluctuation must have a core(s), of non-zero topological charge density, and can be characterised by the size, ρ , of this core. One expects that the important contribution to χ_t come from $\rho \approx 0(\Lambda_{\text{mom}}^{-1})$: the typical dynamical length scale. For such ρ , $g^2(\rho)$ is not small and one cannot rely on perturbative or semi-classical methods: hence the necessity of a numerical calculation. For $\rho \ll \Lambda_{\text{mom}}^{-1}$, on the other hand, $g^2(\rho)$ is small and semi-classical methods [11] apply: the basic topological fluctuation is the well-known instanton with action (to 2-loops)

$$S_I(\rho) \underset{\rho \ll 0(\Lambda_{\text{mom}}^{-1})}{=} \frac{8\pi^2}{g^2(\rho)} \underset{\rho \rightarrow 0}{\rightarrow} \infty. \quad (2.1)$$

The density of such instantons in the vacuum will thus be

$$D(\rho) \frac{d\rho}{\rho} = \frac{d\rho}{\rho^5} e^{-S_I(\rho)}, \quad (2.2)$$

so for $g^2(\rho)$ small, where these calculations can be done analytically, the density is negligible. On the lattice, one expects

$$S_I^{\text{lattice}}(\rho) \underset{\rho \gg a}{\approx} S_I(\rho). \quad (2.3)$$

However as ρ decreases down to $\rho \approx a$ the instanton core begins to slip between the lattice links and its action goes to zero:

$$S_I^{\text{lattice}}(\rho) \underset{\rho \approx 0(a)}{<} S_I(\rho). \quad (2.4)$$

This suggest a strong relative enhancement of instantons on the size scale of the ultraviolet cut-off. Such lattice artefacts should decouple from the long-distance physics and so should not be included in a calculation of χ_t . The exclusion of all such lattice artefacts, is the principal difficulty facing attempts to calculate topology on the lattice.

Since one needs to distinguish between topological fluctuations on the basis of their size, it would be very useful to calculate the total topological charge Q from the local topological charge density $Q(x)$ rather than from

the winding of the gauge potential outside the cores. (Although this latter approach [9] has been extensively pursued [6-8] in the SU(2) case, and very recently in the SU(3) case too [12], the lattice artefact contribution has never been properly quantified and what evidence there is suggests that it is large [13,14]).

So one needs to translate $Q(x)$ in eqn (1.2) onto the lattice. The variables are SU(3) matrices $U_\mu(n)$ on the links ($\mu=1, \dots, 4$) emanating from the sites n . The product of the matrices around the elementary square emanating out of the site n in the $(\mu\nu)$ plane is denoted by $U_{\mu\nu}(n)$ (a plaquette):

$$U_{\mu\nu}(n) = U_\mu(n) U_\nu(n+\mu) U_\mu^\dagger(n+\nu) U_\nu^\dagger(n) \quad (2.5)$$

Since this operator compares neighbouring links it will contain derivatives in the continuum limit; so one expects an expansion

$$U_{\mu\nu}(n) \xrightarrow{a \rightarrow 0} 1 + ca^2 F_{\mu\nu}(n) + \dots \quad (2.6)$$

can be made. This is true and one [15] can construct a lattice topological charge density

$$Q_L^N(n) \equiv - \frac{1}{32\pi^2} \epsilon_{\mu\nu\rho\sigma} \text{Re tr}[U_{\mu\nu}(n) U_{\rho\sigma}(n)] \quad (2.7a)$$

$$\xrightarrow{a \rightarrow 0} a^4 Q(x) \quad (2.7b)$$

so that

$$\sum_n Q_L(n) \xrightarrow{a \rightarrow 0} \int d^4x Q(x) \equiv Q, \quad (2.8)$$

and

$$\chi_t^L \equiv \langle (\sum_n Q_L(n))^2 \rangle \xrightarrow[aL \rightarrow \infty]{a \rightarrow 0} \chi_t, \quad (2.9)$$

where L is the length of the lattice (in lattice units). Unfortunately the naive continuum limits displayed in eqns (2.7-9) are misleading: supposedly

higher order (dimensionless) terms, $a^n 0_n$, which are neglected in eqn (2.7) all contribute numerically at $O(a^0)$ because of the contribution of ultraviolet fluctuations, with frequencies $O(a^{-1})$, to the operator 0_n which has dimensions a^{-n} . Therefore if $\langle Q_L^2 \rangle$ is actually measured in the lattice vacuum the best one can hope for is (for small enough g^2) a linear decomposition

$$\frac{1}{L^4} \langle (\sum_n Q_L(n))^2 \rangle = \sum_m (g^2)^m + \dots + a^4 \chi_t . \quad (2.10)$$

There have been several calculations, in both SU(2) and SU(3) [5,16,17], attempting to extract χ_t using eqns (2.7-9) and assuming eqn (2.10). For technical reasons ([5] and section 3), a symmetrised version of (2.7) was used:

$$Q_L^S(n) = \frac{1}{2^4 32\pi^2} \sum_{\mu\nu\rho\sigma=\pm 1}^{\pm 4} \epsilon_{\mu\nu\rho\sigma} \text{tr} (U_{\mu\nu}(n) U_{\rho\sigma}(n)) , \quad (2.11)$$

where

$$1 = \epsilon_{1234} = - \epsilon_{2134} = - \epsilon_{-1234} \text{ etc} . \quad (2.12)$$

Unfortunately all the evidence [14,17] is that in the range of g^2 where numerical calculations are practical, Q_L is not the simple superposition of unphysical ultraviolet and physical integer-valued fluctuations that eqn (2.10) requires. This makes extremely difficult any attempt to extract the local topological charge density directly from Monte Carlo generated lattice gauge field configurations.

Given the intractability of the direct approach, we adopt a different one, the cooling method described in the next section.

3. Cooling Method

The cooling method is designed to determine the topological susceptibility (1.3) on the lattice in a way that is not plagued by contributions from localized lattice artefacts. We start with the observation that the specific problem to applying the discretised topological charge density in

eqns (2.7,2.11) is the presence of ultraviolet fluctuations, and that in terms of frequency these fluctuations are very distant from the long distance fluctuations of physical interest. One therefore expects that if one were to locally smoothen the lattice gauge field configuration, one would remove the offending ultraviolet modes while leaving the physical, extended, topological charge density undisturbed. Historically the way was paved by investigations by Berg of the $O(3)$ σ -model [18], where he found that the charge of a configuration can easily be read off from the residual action after a number of relaxation sweeps. We shall refer to the smoothening of a field configuration as cooling. Bergs calculation and a subsequent investigation [19] showed that a substantial fraction of the charge was carried by the above described lattice artefacts, which they called dislocations. Like in the 2-dimensional model investigated by Berg there is no evidence that dislocations survive a number of cooling sweeps in 4-dimensional $SU(N)$ calculations. Cooling works as described next. We use the standard simple plaquette action

$$S = -\frac{\beta}{3} \sum_{\mu\nu} \text{Re tr}(U_{\mu\nu}(n)) , \quad (3.1)$$

where $\beta=6/g^2$, throughout. The local fluctuations round a link $U_\mu(n)$ are measured by the local piece of the action that involves $U_\mu(n)$ and its neighbours:

$$\begin{aligned} S &= -\frac{\beta}{3} \sum_v \text{Re tr} [U_\mu(n) U_\nu(n+\mu) U_\mu^\dagger(n+\nu) U_\nu^\dagger(n)] + [\text{terms without } U_\mu(n)] \\ &\equiv -\frac{\beta}{3} \text{Re tr} [U_\mu(n) \Sigma] + [\text{terms without } U_\mu(n)] . \end{aligned} \quad (3.2)$$

If the link $U_\mu(n)$ is replaced by a link matrix $U'_\mu(n)$ such that the action is reduced,

$$S[U'_\mu(n)] < S[U_\mu(n)] , \quad (3.3)$$

the local variations are reduced and hence the fields are locally smoothened. In the case of $SU(2)$ [13,14], the choice

$$U'_\mu(n) = \Sigma^+ / |\Sigma| \quad (3.4)$$

will locally minimize the action. In SU(3) this does not work because $\Sigma^+ / |\Sigma|$ is not an SU(3) matrix. One route would be to modify the standard Metropolis algorithm so that one only accepts a change of $U_\mu(n)$ if it lowers the action [13]. Alternatively one can numerically find the SU(3) matrix that locally minimizes the action (3.2) [20]. Using this method one of the authors [21] found the first value for the SU(3) topological susceptibility free of lattice artefact contributions. Though not yet convincing because of lack of statistics, the measurement was, for reasons already mentioned, not plagued by large uncertainties that made the earlier measurement [5,16,17] of the SU(3) susceptibility numerically useless. In the measurements [21] the residual action after a number of cooling sweeps was used as the measure of the topological charge of the configuration. The method used in this paper consists of an approximation to the exact local minimization of the action consisting of multiplying the link-variable by a matrix from a diagonal SU(2) subgroup [4],

$$U_\mu(n) \rightarrow V U_\mu(n) , \quad (3.5)$$

where V is so chosen that within this class of transformations the action is locally minimized. In our calculation we have always used all 3 diagonal subgroups. The advantage of this type of minimization is that it lends itself well to vectorization. Moreover it turns out to be globally almost as efficient as the exact minimization. The local action minimization step consists of finding the SU(2)-subgroup matrix V that minimizes

$$- \text{Re tr} [V A] , \quad (3.6)$$

where A is the sum of the six open plaquettes $U_{\mu\nu}(n)$ that contain $U_\mu(n)$. Note that for matrices A of the form

$$H + i\lambda 1 , \quad H = H^+ , \quad \lambda \in \mathbb{R} , \quad (3.7)$$

the minimization of (3.6) implies $V = 1$. In general there are SU(3) matrices V though that lower the value (3.6). It is not easy to see explicitly that the sum of six SU(3) matrices cannot lead to a matrix of the form (3.7), though it is easy to see for infinitesimal matrices. However, the method described above is also the $\beta \rightarrow \infty$ limit of the Cabbibo-Marinari [22] Monte Carlo updating scheme which satisfies detailed balance.

Our algorithm for measuring the topological charge of a Monte Carlo generated lattice gauge field configuration will therefore proceed as follows. Focus on some link variable $U_\mu(n)$ and alter it so as to locally reduce the action. Repeating this procedure for every link on the lattice is called a "cooling sweep". A sequence of such cooling sweeps is performed and after each cooling sweep the topological charge of the smoothened configuration is measured using eqn (2.7-8,2.11). After a few cooling sweeps the ultraviolet fluctuations are removed. After 15 cooling sweeps on the 10^4 lattice the mean plaquette is as low as it would be in an ensemble of configurations generated at $\beta \approx 5000$. For configurations this smooth the naive continuum limit of (2.7) will hold to a good accuracy. As the number of cooling sweeps is increased even the longer-distance, physically relevant, quantum fluctuations will begin to be erased. However, in the continuum the topological charge is conserved under a procedure that continuously reduces the action. So, up to lattice discretisation effects, the cooling procedure leaves the physical topological charge unaltered. This means that if the cooling is continued, eventually the only remaining fluctuations will be those carrying non-zero topological charge: and since we have numerically minimized the action, they will in fact be the well-known semi-classical multi-instanton solutions. At this point the action has reached its minimum and cannot be reduced any further (up to lattice discretisation effects): it is related to the topological charge Q by

$$S = \frac{8\pi^2}{g^2} Q = \frac{4}{3} \pi^2 \beta Q . \quad (3.8)$$

If cooling is continued even further the extended instanton cores will slowly shrink because the lattice breaks scale-invariance and the action decreases as ρ decreases. When $\rho \rightarrow 0(a)$ the instanton becomes a "dislocation" and rapidly disappears from the lattice under further cooling.

For the calculations described in this paper we used the lattice version (2.7) of the FF^* operator or the symmetrized version (2.11) to measure the topological charge of Monte Carlo generated configurations during the cooling process. The main reason for the use of this operator rather than the residual action is that it converges faster. This approach - using a small number of cooling sweeps as a means to make the fields smooth enough for the local topological charge density to become directly measurable - was

first pursued by one of us [13,14] in the case of SU(2). Even after a couple of cooling sweeps the configurations are smooth enough for (2.7,2.11) to be good operators, in the sense that the sum $\sum_n Q_n$ is close to an integer.

On "hot" configurations the symmetrized version of the operator shows substantially reduced fluctuations as compared to the unsymmetrised version [5], which presumably reflects the fact that the former is the correct one to use because it has definite reflection positivity properties. On cooled configurations, however, neither form has a theoretical advantage and indeed the difference becomes rapidly much less pronounced. On most configurations where we measured the charge every sweep the non-symmetric form (2.7) was used, because it requires much less computer time. On the 16^4 configurations, where we measured only every second sweep, the symmetrized version was used because globally it still converges somewhat faster than the non-symmetric operator. Due to the fact that the cooled configurations are very smooth, we don't have problems with the large perturbative corrections that made Q_n lose its topological significance on hot configurations [5]. Using the residual action as a measure of the topological charge has the disadvantage that it may take long for all the instanton/anti-instanton pairs to annihilate. On small lattices this is not a problem but on larger ones it will be. Already on 16^4 lattices at $\beta=6.0$ one case (out of 5) has been seen where after 250 sweeps an $I\bar{I}$ pair had not yet annihilated. The advantage of the present method is that it needs only a few cooling sweeps and so we will be able to obtain usefully accurate values of χ_t for a range of couplings, which in turn allows us to test whether we are seeing continuum physics.

To come back to the motivation for using the cooling method we note that the fact that lattice artefacts exist is evident from the discrepancy between the susceptibility derived from an ensemble of configurations by direct measurement [6,8,23] and that derived using the cooling method [13,23]. The direct methods use some kind of explicit expression for the topological charge of a lattice configuration [5-8,10,12,16,17]. In the continuum limit only large scale properties of the configurations should be relevant for physical quantities and both methods should converge. At the coupling values where Monte Carlo simulations are possible however there is a clear advantage in trying to get rid of the contributions of the lattice artefacts. This is the aim of the cooling method; it is a physically

well-motivated method to improve the approach to the continuum limit.

We conclude this section with a number of examples of cooling runs. In fig 1 it is shown how the action and topological charge vary as an 8^4 configuration is systematically cooled. This configuration was originally generated at $\beta=5.80$. We observe that the action falls very rapidly to begin with - by over a factor of 100 in the first 5 cooling sweeps - and then increasingly slowly. Indeed it would seem to be plateauing to a non-zero asymptotic value of about $8\pi^2/g^2$ - the classical action of an instanton. However after about 100 cooling sweeps there is a sudden shift away from the plateau and a rapid decrease to zero. Meanwhile the total topological charge rests close to $Q=1$ from about 5 cooling sweeps till around 100 sweeps at which it also rapidly decreases to $Q=0$. We can interpret what is happening as follows. The initial very rapid decrease of the action reflects the efficiency with which the ultraviolet fluctuations are erased by our local smoothening of the fields: these ultraviolet modes contribute 0(99%) of the action of a typical lattice gauge field generated at $\beta=5.8$. There then follows the erasure of longer distance physical modes and this occurs more gradually for the same reason that the Monte Carlo took many sweeps to build up these modes in the first place. Finally, however, even these modes are erased and we reach the minimum of the action

$$\frac{\delta S}{\delta U_\mu(n)} = 0 \quad (3.9)$$

which leads to the action plateau in fig 1(a). At this point we have numerically minimised the action within the topological charge sector defined by our starting Monte Carlo generated configuration. This means that, up to lattice discretisation effects, the gauge field at this point should be simply a classical multi-instanton solution corresponding to Q - including a dilute gas of instantons and anti-instantons if the lattice is large enough. The lattice discretisation effects are ultimately important because they mean that the instanton action does depend on its size albeit by corrections that are $O((a/\rho)^2)$ and hence very small for large ρ . For the Wilson action the action decreases as ρ decreases and so the lattice instanton will slowly shrink as we continue to cool the lattice. The changes in the action are very small when ρ/a is large and will be

essentially invisible on a logarithmic scale. Once $\rho \approx 0(a)$ however, the changes become large and rapid once again as the instanton core shrinks between the lattice sites and Q drops from 1 to 0.

We emphasize the detailed consequences of cooling a gauge field configuration since ultimately the reliability of our measurements of Q depend on our interpretation of these details. It is worthwhile therefore to illustrate our above interpretation. We start with the configuration on the action plateau, after 60 cooling sweeps, where the configuration should be a solution of the equations of motion. The characteristic of these solutions is self-duality. To test this we form profile functions e.g.

$$Q(n_t) = \sum_{\vec{n}} Q(\vec{n}, n_t) \quad (3.10)$$

along all four directions with similar definitions for the action

$$\tilde{S}(n_t) = \frac{g^2}{8\pi^2} \sum_{\vec{n}} S(\vec{n}, n_t) \quad (3.11)$$

Self-duality tells us that

$$Q(n_\mu) = \tilde{S}(n_\mu) \quad (3.12)$$

We plot in fig 1(b) the profile functions of the action and topological charge along the z and t axes. Up to small corrections - due to lattice discretisation effects - we see a satisfactory confirmation of self-duality and hence of the fact that the configuration does indeed contain an instanton. Of course the instanton was expected: the cooling procedure by its nature transforms the physical topological fluctuations of the original Monte Carlo configuration into the minimum action multi-instanton solutions. Of course one cannot infer the existence of instantons in the physical vacuum from their presence on the action plateau!

By monitoring the action and Q profiles we can actually watch the instanton shrink and disappear. Equivalently we can measure the maximum value of the density since, up to lattice discretisation effects,

$$Q_{\max} = \max_{\{n\}} \{Q(n)\} \sim \rho^{-4} \quad (3.13)$$

Similarly the largest negative value of $Q(n)$

$$Q_{\min} = \min_{\{n\}} \{Q(n)\} \quad (3.14)$$

will monitor the level of background noise.

We plot these quantities in fig 1(c). By 10 cooling sweeps $|Q_{\min}| \ll Q_{\max}$ indicating that the background has become negligible. From there $Q_{\max}$ grows at an ever increasing rate indicating the shrinking of the instanton as in eqn (3.13). Just after 100 cooling sweeps the value of $Q_{\max}$ begins to plummet: the instanton core is shrinking within the boundaries of a hypercube. Now something remarkable happens. Just as $Q_{\max}$ becomes small $|Q_{\min}|$ rises to large values albeit very briefly. The location is at $n=(3,3,2,2)$ neighbouring the location of $Q_{\max}$ at $n=(3,4,2,3)$. This phenomenon - that the moment of disappearance of an instanton is accompanied by the brief appearance of an anti-instanton on a neighbouring site - is sufficiently unique as to allow us to identify the disappearance of a unit of topological charge very early in the cooling procedure, long before the background noise has been suppressed and before any well-formed plateaus in Q or S . We have an example in fig 1(c). Close examination reveals a brief peak in $Q_{\max}$ at 5 cooling sweeps. The coordinates at this peak are $n=(5,5,5,3)$ which neighbour onto the site $n=(5,5,6,2)$ where $Q_{\min}$ was located just before - and differ from the coordinates $n=(4,1,7,7)$ of $Q_{\max}$ just before the peak and the coordinates $n=(1,8,5,1)$ just after the peak. As added evidence the maximum value of the action density

$$S_{\max} = \max_{\{n\}} \{S(n)\} \quad (3.15)$$

also peaks at the third and fourth cooling sweeps. We can therefore identify that a fluctuation with $Q=-1$ vanished at about the 3rd or 4th cooling sweeps. This is also consistent that Q rises to $Q=1$ as in fig 1(a). Once can construct and measure various other quantities that provide signatures of the disappearance of a topological charge during cooling. As a result we are certainly able to measure Q from 5 cooling sweeps onwards (for $\beta \gtrsim 5.8$) and frequently earlier than that. This means that we are able

to identify all the topological fluctuations within the original Monte Carlo configurations except for those that have $\rho \approx a$ and disappear under the first couple of cools.

In fig 3 a number of cooling histories of 10^4 lattices are shown. They were generated at $\beta=5.9$. In the range of couplings where our calculations were performed $5.6 \leq \beta \leq 6.0$ the separation between the lattice cut-off and physical length scales is not large, and things do not always follow the above pattern. In fig 4 we show a number of cases where the topological charge changes during the cooling run (10^4 lattices, $\beta=5.9$). These examples where physical topological fluctuations disappear early on in the cooling process will contribute to systematic errors if ignored. We will come back to this in section 5.

Generally, the smaller the core of the topological fluctuations the more quickly will the cooling erase it from the lattice by shrinking it down to $\rho < a$. Qualitatively, the tail of lattice artefacts with $\rho \approx 0(a)$ should disappear within a few cooling sweeps while physical fluctuations with $\rho \gg a$ should be essentially stable. The range of g^2 we investigate is one in which such a clear-cut distinction in size-scales is not yet operative: we work at the margin of possible continuum physics. Nonetheless on some smaller lattices, at larger β values, we have observed a clear distinction between short-lived and longer-lived components of the lattice topological charge, see section 5. This indicates a substantial pollution by lattice artefacts, in this range of β , which would render impossible the extraction of the physical topology from the winding of the gauge potential [6-8,10,12]. With the present method it appears that after about 15 cooling sweeps one is generally free of artefact contributions. Some will survive and can either be recognized by their characteristics or will contribute to the systematic error, cf section 5. In this paper we shall present results for the topological charge as measured after 25 cooling sweeps, unless otherwise stated.

4. χ_t at zero temperature

The method described in the previous section allows us to calculate the topological charge Q of a Monte Carlo generated gauge field configuration. If we do so for a sequence of such configurations labelled by $I=1, \dots, N$, on an L^4 lattice at some β we can extract an estimate of the topological

susceptibility χ_t at that β :

$$a^4 \chi_t = \frac{1}{L^4} \frac{1}{N} \sum_{I=1}^N (Q^I)^2 \pm O\left(\frac{1}{\sqrt{N}}\right) \quad (4.1)$$

To make contact with the Witten-Veneziano mass formula (1.4) we need χ_t at zero physical temperature. This essentially means that the lattices used must be large enough for finite size effects to be negligible.

Finite size effects

One might imagine that the relevant scale for finite size effects would be the pseudoscalar glueball mass, $m(0^{++})$, since

$$\langle Q(x) Q(0) \rangle \sim_{x^2 \rightarrow \infty} \exp[-m(0^{++}) |x|] \quad (4.2)$$

It turns out however that the dominant finite size effects occur when the size La of the lattice shrinks through the inverse deconfining temperature T_c at that β :

$$La(\beta) = T_c^{-1}(\beta) \quad (4.3)$$

(This is a reflection of the suppression χ_t undergoes when the temperature rises through the deconfining transition [4] - see below). In the continuum limit these two scales, $m(0^{++})$ and T_c , would be simply proportional; in our range of β this need not be the case. Accordingly we shall test for finite size effects as follows. We compare in fig 4 χ_t as measured on 4^4 , 6^4 and 8^4 lattices at $\beta=5.70$. We observe that within errors there is no suppression on the 6^4 lattice as compared to the 8^4 lattice. Now for $L=6$ eqn (4.3) is satisfied at $\beta \approx 5.86$ [24]; therefore the 6^4 lattice at $\beta=5.70$ is $\Delta\beta \approx 0.16$ from the β value where its χ is very strongly suppressed. (N

logical susceptibility

In table 1 we present our measured values of χ_t . The 8^4 lattices at $\beta=5.60$, 5.70 , 5.80 , the 10^4 lattice at $\beta=5.90$ and the 16^4 lattice at $\beta=6.00$ satisfy our above criterion and will be used to determine χ_t . The 8^4 lattice at $\beta=5.85$, on the other hand, is like the 6^4 lattice at $\beta=5.73$ and is expected to suffer substantial finite size effects (as a detailed comparison confirms).

In fig 5 we plot our measurements of $a^4\chi_t$ on lattices which we believe to suffer at most minor finite size corrections. We see that $a^4\chi_t$ decreases rapidly with increasing β in qualitative accordance with what we expect from asymptotic freedom.

To extract a value of χ_t in physical MeV units we need to obtain $\chi_t^{1/4}/M$ where M is some quantity with the dimensions of mass, whose value is already known in MeV units. Of course such a procedure only makes sense if we have evidence that the ratio $\chi_t^{1/4}/M$ has already attained its continuum limit: we need to find this ratio independent of β over a range of β such that, given the statistical errors, this independence is actually meaningful. In fig 6a we plot the value of $\chi_t^{1/4}/\sqrt{K}$ where K is the string tension as obtained in the very high statistics calculation of ref [25], and we plot values of this ratio only at those β values where K has been calculated. We observe that this ratio is consistent with being independent of β for $\beta \geq 5.70$. How significant is this? Between $\beta=5.70$ and $\beta=6.00$ both $\sqrt{K}$ and $\chi_t^{1/4}$ fall by about a factor of 2; the ratio $\chi_t^{1/4}/\sqrt{K}$ on the other hand changes by less than 15% within statistical errors. So while the observed independence of β is certainly significant it is far from being totally convincing.

Nonetheless we now possess a means and rationale for extracting a value for $\chi_t^{1/4}$ in MeV units. If we use the conventional value for the string tension $\sqrt{K}=420$ MeV, we obtain

$$\chi_t^{1/4} = \begin{array}{lll} 183 \pm 5 \text{ MeV} & \text{at } \beta = 5.7 \\ 190 \pm 7 \text{ MeV} & \beta = 5.9 \\ 191 \pm 16 \text{ MeV} & \beta = 6.0 \end{array} \quad (4.4)$$

which is in remarkably good agreement with the value expected on the basis of the Witten-Veneziano mass formula.

It is interesting to attempt to extract the ratio $\chi_t^{1/4}/\Lambda_L$. When β is large enough $a(\beta)$ can be related to Λ_L by the usual 2-loop perturbative expression:

$$a(\beta) = \frac{1}{\Lambda_L} e^{-\frac{4\pi^2}{33}\beta} \left(\frac{8\pi^2}{33}\beta\right)^{\frac{51}{121}} \quad (4.5)$$

If we assume that this expression is already accurate in our range of β , we can extract $\chi_t^{1/4}/\Lambda_L$ from our measured values of $a^4\chi_t$. In fig 6b we show the result of doing so. It is very clear that $\chi_t^{1/4}/\Lambda_L$ is not independent of β : this is no surprise because it is well known that neither is $K^{1/2}/\Lambda_L$. From β -function measurements and measurements of the deconfining temperature one does not expect asymptotic scaling, i.e. scaling according to (4.5), in the β region where we performed our measurements. This is corroborated by our result. But we do see scaling already in this β -range for the ratio of susceptibility to the square of the string tension.

Some more details

a) number of cools.

We have defined the topological charge to be that which remains after 25 cooling sweeps. It is interesting to ask how dependent are our measured values of $a^4\chi_t$ on this specific choice. To do so we measure $a^4\chi_t$ at 25 cools, at 10 cools and as early as possible on most of our data on 8^4 lattices at $\beta=5.6, 5.7, 5.8$ and on the 10^4 lattice at $\beta=5.9$. How early one can identify Q - this depends on how early one can identify the disappearance by shrinking of a unit of topological charge - depends somewhat on β and lattice size but typically will correspond to 5 cooling sweeps or fewer. In fig 7 we plot the values of $a^4\chi_t$ obtained using these three different criteria. We observe that the different definitions give values of χ_t that differ by at most about 10% - which means differences of about 2% in $\chi_t^{1/4}/\sqrt{K}$ - and, more importantly, that the variation of $a^4\chi_t$

with β is independent of the definition used, so that our ability to use the string tension to normalise $\chi_t^{1/4}$ in MeV units is not just an artefact of our choice of 25 cooling sweeps.

b) errors

Our measurements of Q were performed every 20 or 25 Monte Carlo sweeps for $\beta \leq 5.80$, every 50 sweeps at $\beta = 5.90$ and every 112 sweeps at $\beta = 6.0$, in an attempt to obtain independent measurements. Suppose we obtain N values of Q , $Q_{i=1, \dots, N}$, at a fixed β . Assuming these to be independent we can calculate

$$\langle Q^2 \rangle = \frac{1}{N} \sum_{i=1}^N Q_i^2 \quad (4.6)$$

and an error, σ , by

$$\sigma = \frac{1}{\sqrt{N-1}} \left\{ \frac{1}{N} \sum_{i=1}^N (Q_i^2 - \langle Q^2 \rangle)^2 \right\}^{1/2} \quad (4.7)$$

A better alternative - if there are correlations between neighbouring measurements of Q - is to bin the sequence of N values of Q into subsequences of M measurements each. From each subsequence we obtain an estimate of Q^2

$$Q_j^2(M) = \frac{1}{M} \sum_{i=1+(j-1)M}^{jM} Q_i^2 \quad (4.8)$$

and if M is large enough these estimates will be independent, so that a good error estimate will be

$$\sigma = \frac{1}{\sqrt{N/M-1}} \left\{ \frac{1}{N/M} \sum_{i=1}^{N/M} (Q_i^2(M) - \langle Q^2 \rangle)^2 \right\}^{1/2} \quad (4.9)$$

In practice M is limited by the need for N/M to be large enough. The errors quoted in table 1 are batched errors of this kind. In the table we compare the two kinds of error estimates. We see that the differences are

not very large suggesting that our choice of how frequently to measure Q was sensible, although the small number of measurements at $\beta=6.0$ makes the error estimate rather unreliable there. We shall return to a more precise estimate of correlations later on.

c) $\langle Q^2 \rangle$ in the dilute gas approximation

When one measures Q^2 as in eqn (4.6) one is obviously most affected by the largest values of Q_1^2 which are also the rarest. Indeed the error estimate in eqn (4.7) will frequently be dominated by the single largest Q_1^2 measurement: even in our highest statistic data this value contributes significantly to the error estimate. This problem can be overcome if one knows the functional form of the Q distribution. In this case the parameters will be largely determined by the accurately measured small Q values. Of course we do not know the functional form: but we could try the dilute gas approximation and this would at the same time serve as a way of checking the validity of that approximation.

In the dilute gas [11] we assume that the probability distributions for the numbers of instantons and anti-instantons are independently Poisson

$$P_i = e^{-m} \frac{m^i}{i!} \quad (4.10)$$

so that the probability to obtain charge Q is given by a simple convolution

$$P_Q = e^{-2m} m^{|Q|} \sum_{i=0}^{\infty} \frac{m^{2i}}{i!(i+|Q|)!} \quad (4.11)$$

We fit our measured distributions of Q ; we assume that all the measurements are independent so that if the value Q appears $n(Q)$ times the percentage error is $1/\sqrt{n(Q)}$. The best fit determines m for each β and lattice size and hence $\langle Q^2 \rangle$:

$$\langle Q^2 \rangle = 2m \quad (4.12)$$

In table 2 we compare these "dilute gas" values of $\langle Q^2 \rangle$ with the usual "naive" values and show the minimum $\chi^2/\text{d.f.}$ of the fit. We observe that the

dilute gas values are consistent with the naive values. If on the other hand, we are interested in the validity of the dilute gas approximation, the relevant quantity is the goodness of fit, $\chi^2_{\min}/df.$, and especially so for the highest statistics data. Here we find that the fits are not good but acceptable; interestingly, for the 8^4 data at $\beta=5.85$ and the 6^4 at $\beta=5.73$ which we know are strongly afflicted by finite size effects the fits are very good. This is interesting because the topological charge carriers on lattices which are becoming too small, tend to be relatively pointlike themselves - perhaps enhancing the validity of the dilute gas approximation.

5. Topology: lattice or continuum?

The method we have used to measure the topological charge Q is designed so that it will correctly measure physical topological fluctuations deep in the continuum limit where the scales of these fluctuations are far from the lattice regularisation scale a . Our actual calculations, on the other hand, have been performed for values of the bare coupling where the dynamical length scale is not much more than 2 or 4 times the lattice spacing. This leads to problems. For example, with $a(\beta)$ large, the field fluctuations between neighbouring links may be so large that one loses the notion of there being any unique continuum configuration into which the lattice configuration can be naturally embedded and then the topology we measure may be more a product of the way the first few cooling steps deform these highly fluctuating fields rather than being a meaningful property of the original field configuration. To test whether this is so or not we shall take a Monte Carlo generated field configuration and perform several different cooling sequences on it: if the measured Q is the same then indeed we see that it must be a property of the original configuration and not of the cooling procedure.

In this section we shall be performing various calculation designed to reveal to what extent the topology we are measuring by the method described in section 3 is indeed "continuum" topology. Firstly we shall perform some

reproducibility tests of the kind described in the previous paragraph. Then we shall measure Q on every member of a sequence of Monte Carlo configurations to test how stable Q is under Monte Carlo updating: this stability will grow with increasing β . Next we shall determine the distribution of ρ after 25 cooling sweeps and shall see whether, in our range of β , there is beginning to occur a clear separation between ultraviolet and physical length scales. At this point it will be appropriate to compare our results with recent low statistics measurements of Q by geometric methods. Finally we shall calculate the correlation between the Q measurements described in section 4 and see how rapidly the correlation length grows with β .

reproducibility

To what extent is the topology we measure a product of the cooling procedure rather than being intrinsic to the Monte Carlo generated field configuration? To answer this question we shall perform several independent cooling sequences from the same Monte Carlo configuration and inspect whether the topology obtained differs for the different sequences or not i.e. we check for reproducibility (The sequences will differ by the order in which one chooses the sites during a cooling sweep: the choice will be random). Our expectation is that in the strong coupling limit there should be no reproducibility so that any onset of reproducibility as β increases is a signal for the onset of at least one aspect of continuum topology.

Starting with strong coupling, we thermalise a 6^4 lattice at $\beta=1.0$ and perform 2 independent cooling sequences from the same configuration. We show in fig 8 how Q varies with the number of cooling sweeps for each sequence: there is a clear lack of reproducibility.

We now turn to $\beta=5.80$, which is in the middle of the β range we used to calculate χ_t and where we would hope to begin to see signs of reproducibility. We focus on five Monte Carlo generated 8^4 field configurations, labelled A,B,...,E, and perform 3 independent cooling sequences on each. At the end of each sequence of 25 cools we measure the total topological charge Q and in addition identify the coordinates of the site where each topological fluctuation has its maximum topological charge density (provided this maximum is greater than $1/16\pi^2$). The results are set

out in table 3. What is immediately striking is the extent of the observed reproducibility: in most cases not only Q but also the detailed locations of the "instanton" cores are the same for all three cooling sequences. The only exception is configuration C where an extra "instanton" survives the first cooling sweep: upon inspection this is a very narrow fluctuation, close to shrinking out of the lattice.

At $\beta=6.0$ a test is provided by the fact that some configurations have been treated by both the exact minimization method of [20,21] and the method presented in this paper. The results of the exact minimization are after 250 cooling sweeps, moving through the lattice in a regular way. The results with the method of this paper were obtained by going through the lattice in a checkerboard fashion, and after 25 sweeps as usual. In table 4 we summarize the results of the two methods on five configurations. Here the ease of the present method to treat $\bar{I}\bar{I}$ pairs correctly shows up in ED6029, where the correct total charge vanishes, whereas with the residual action criterion tentatively a total charge one was assigned (though the fact that the value .78 was relatively far from an integer made this assignment a bit suspect). In case of ED6037 one charge was lost unnoticed, with the residual action criterion, but the two sets of measurements surely are consistent.

The conclusion is that in our β range we have largely regained reproducibility so that the Q we measure is indeed an intrinsic property of the original field configurations and is not just an artefact of the cooling algorithm.

Stability under Monte Carlo updating

As we go deeper into the continuum limit we expect that the physical topological charge should become increasingly stable against the local changes of which the Monte Carlo updating is composed. Indeed, for large enough β a sequence of Monte Carlo configurations should remain trapped in a sector of fixed Q . It is therefore interesting to see how close or far we are from this ideal continuum limit.

We measure Q on each of a sequence of 90 configurations at $\beta=5.7$, 5.8 and 5.85. The results are presented in fig 9. The presence of correlations is certainly evident - especially at $\beta=5.8$ and 5.85. There appears to be one component of Q which is stable over about 20 Monte Carlo sweeps and to this is added a rapidly fluctuating component. Looking at the configurations in

some detail confirms the idea that the rapid fluctuations are due to relatively pointlike topological charges.

correlations in Q

Suppose we measure the correlation of Q between configurations that are separated by j Monte Carlo sweeps. We expect that for large j

$$C(j) \equiv \frac{\langle Q_i Q_{i+j} \rangle}{\langle Q_i^2 \rangle} \propto \exp(-j/\xi) \quad (5.1)$$

where the correlation length ξ will depend on the details of the updating algorithm and on β . We will attempt to extract ξ from our measurements of Q at $\beta=5.8, 5.9$ and 6.0 .

Our measurements of Q are made every δ Monte Carlo sweeps where δ was originally chosen with the intention that it should be large enough for the correlations to be small. In table 5 we present the values we find for the normalised correlation function (5.1). We find a strong correlation at all β values indicating that there is a component of Q that is relatively stable under Monte Carlo update and, because δ increases with β , this stability also increases with β .

To quantify these observations one should at this point extract the correlation length from

$$\xi/\delta = - \ln \left\{ \frac{C(2\delta)}{C(\delta)} \right\} \quad (5.2)$$

Unfortunately our value of δ is already so large that we do not obtain a statistically significant signal for $C(2\delta)$. Accordingly we can do no better than to extract an effective correlation length ξ_{eff} using

$$\xi_{\text{eff}}/\delta = - \ln C(\delta) \quad (5.3)$$

which probably understates the true correlation length. We show the variation of ξ_{eff} with β in fig 10. We see that ξ increases rapidly with β and is already large at $\beta=6.0$: further evidence for the continuum-like behaviour of a significant component of Q.

Fig 10 also has some general implications for attempts to calculate Q . Since we would space our measurements no closer than ξ sweeps apart, we can see that the measurement time - 25 cooling sweeps in our case - is already much less than the update time - 112 sweeps - at $\beta=6.0$. So at $\beta=6.0$ we are already being dominated by this aspect of critical slowing down: our sequences of Monte Carlo configurations are being trapped in sectors of a fixed physical topological charge. It is clear that if we use local Monte Carlo update procedures we shall not be able to measure χ_t at β values that are much larger than $\beta=6.0$ [26].

Q at various length scales

We emphasized at the beginning of this paper that one of the attractive features of our method is that it allows us to discriminate between topological fluctuations on different length scales. So far we have not made any real use of this possibility except insofar as our choice of 25 cooling sweeps was motivated by our experience that the narrowest "instantons" would have typically disappeared during the first few cooling sweeps. Now we shall return to our measurements of χ_t as presented in section 4 and see what happens if we place increasing lower bounds on the sizes of the topological fluctuations which we include in the calculation.

In practice we shall gauge the size of the core of an "instanton" by the maximum value that the topological charge density $Q(n)$ takes within that core. The relationship is clearly

$$Q_{\max} = \max_{n \in \text{core}} |Q(n)| \sim \rho^{-4} \quad (5.4)$$

where ρ is the core size (and we ignore lattice discretisation effects). We shall calculate Q for various cut-offs Q_c , i.e. we only include in Q those topological charges whose cores satisfy $|Q_{\max}| < |Q_c|$, and we examine how the resulting χ_t depends on Q_c . As β increases we expect the density of core sizes to separate into a possible peak around $\rho \approx a$ due to lattice artefacts and a second peak around $\rho \approx 0(1 \text{ fermi})$ due to the physical continuum topology. So if Q_c is chosen so that lattice artefacts are excluded, we would expect that, as β increases, χ_t should become increasingly independent of the precise choice of Q_c as long as Q_c is within a range

$$Q_c \in [\alpha \tilde{Q}, \infty] \quad (5.5)$$

where $\tilde{Q}$ is the typical $Q_{\max}$ for an "instanton" that is about to shrink out of the lattice and α is some fixed factor, $\alpha < 1$.

A typical value of $\tilde{Q}$ is

$$32\pi^2 \tilde{Q} \approx 10 \text{ to } 20 \quad (5.6)$$

so we shall choose cut-offs Q_c with values $32\pi^2 Q_c = 2, 3, 5$ and ∞ . We present in fig 11 the variation of χ_t with Q_c , after 25 cooling sweeps, for $5.7 \leq \beta \leq 5.9$. (Not all the data of section 4 are included, for operational reasons. Errors are not shown: the relevant errors are in the differences of the χ_t for various Q_c and these are small compared to those differences).

We see that while at $\beta=5.7$ χ_t varies quite rapidly with Q_c - indicating that there is no real separation between physical and regularisation length scales - this variation with Q_c decreases sharply as β increases. This provides direct evidence that even as early as $\beta=6.0$ χ_t , as measured this way, will be dominated by topological fluctuations well separated from the lattice regularisation scale.

We should also remark that fig 11 exposes an ambiguity in determining the value of χ_t in physical units. While at $\beta=5.9$ the value of $(a^4 \chi_t)^{1/4}$ varies by only $\approx 8\%$ over the whole range of cut-offs, the variation with β of χ_t does depend strongly on Q_c : so, for example, if we were to use $Q_c = 2/32\pi^2$ we would find that $\chi_t^{1/4}/\sqrt{K}$ would certainly not be independent of β , thus undermining - in this case - our motivation for using $\sqrt{K}=420$ MeV to set the scale of χ_t in physical units. These ambiguities are of course inevitable in this range of couplings: it is a strength of our method that it provides us with the means to explicitly monitor the contributing length scales and so to make visible such ambiguities.

Comparison to geometric methods

Geometric methods which measure the net winding on the lattice (this net winding being absorbed into gauge singularities) will measure not only the physical topology which we measure with our method but also all the $\rho \approx a$ lattice artefacts which we have largely and deliberately excluded by our 25 cooling sweeps. Since one expects the contribution of these lattice artefacts to disappear in the continuum limit the results of the various methods should converge.

In fig 12 we compare our results to some very recent geometric calculations [7,12]. The latter are obviously, at the moment, very imprecise and, in the case of ref [12], on lattice sizes too small to be taken very seriously. Nonetheless there is evidence that all the differences are getting smaller as β increases in marked contrast to the SU(2) case where the differences are much larger [13,14,23] and, apparently, β independent. This again provides us with evidence of the general onset of continuum topology, although it seems clear that the geometric calculations will only be rid of their lattice artefact contributions at rather large β .

6. χ_t at non-zero temperature

The calculations described in sections 4 and 5 show topological fluctuations at zero temperature gradually discarding their knowledge of the lattice as β is made larger. The non-zero behaviour of χ_t is also of interest: we know from semi-classical calculations that χ_t must vanish as $T \rightarrow \infty$, but of greater interest [27] is what does, or does not, happen as T increases through the deconfining temperature T_c . We have in fact shown elsewhere [4] that χ_t suffers a rapid suppression as T increases through T_c , just as in the case of SU(2) [13], but this suppression did not appear abrupt enough for χ_t to qualify as an actual order parameter for the transition. Here we shall examine things more carefully and shall be led to a more open conclusion.

The physical temperature of a periodic Euclidean $L^3 L_t$ lattice is

$$T = \frac{1}{aL_t} \tag{6.1}$$

(where we require L to be large enough for thermodynamics to be relevant). We perform our calculations on a $8^3 4$ lattice; by increasing β the temperature increases as $T=1/4a(\beta)$. In fig. 13 we plot χ_t as measured on the $8^3 4$ lattice and compare it to the " $T=0$ " results previously displayed in fig 5. We see that χ_t is strongly suppressed as the system's temperature is increased through the deconfining transition, which occurs for the $8^3 4$ lattice at $\beta \approx 5.68$ [24]. This suppression seems to occur in a β interval of about $\Delta\beta=0.1$ i.e. in a range of temperatures within about 10 to 15% either side of T_c . So the suppression is rapid, but not rapid enough to suggest that χ_t is really an order parameter smeared out by finite spatial volume effects. Nonetheless the fact that the vacuum loses its topological fluctuations precisely across the deconfining transition - effectively restoring the axial $U(1)$ symmetry - is remarkable and unexpected.

In the above calculation Q was, as usual, measured as the topological charge remaining after 25 cooling sweeps. It is interesting to see how χ_t depends on the number of cooling sweeps and also what is the distribution of core sizes remaining after 25 cooling sweeps. To determine the latter we perform the same analysis as in the previous section: we identify the maximum value of the topological charge density $Q_{\max}$ in each "instanton" core and count only those "instantons" whose $Q_{\max}$ is less than the same cut-off Q_c . Since $Q_{\max}$ is proportional to ρ^{-4} where ρ is the instanton core size (up to lattice effects), when we decrease Q_c our measured Q will include only the broader topological charge carriers.

In fig 14 we show how $a^4\chi_t$, as measured on our $8^3 4$ lattices after 25 cooling sweeps, depends on the cut-off Q_c . For the purposes of comparison we also include the same calculation on our $T \approx 0$ 8^4 and 10^4 lattices (taken from fig 11). We see that χ_t on the $8^3 4$ lattices is very sensitive to the cut-off Q_c and that this sensitivity increases as β increases - in marked contrast to the behaviour of χ_t on the low temperature lattices. This is telling us that, as β increases, a rapidly decreasing proportion of the instantons remaining at 25 cools are broad - most are relatively pointlike. The implication of this is that χ_t should also be much more sensitive to the number of cooling sweeps at high T than at low T . In fig 15 we show that this expectation is fulfilled: at high T , χ_t decreases rapidly with the

number of cooling sweeps indicating that the distribution of core sizes is strongly peaked towards $\rho \approx a$, so that most of them shrink out of the lattice after only a few cooling sweeps.

To make all the above observations more concrete we show in fig. 16 the profile of a typical "instanton" remaining on the $8^3 4$ lattice at $\beta=5.70$ after 25 cools. The profile $Q(t)$ is defined

$$Q(t) = \sum_{\vec{x}} Q(\vec{x}, t) \quad (6.2)$$

with corresponding definitions for $Q(x)$ etc. We see in fig. 16 that the "instanton" at high temperatures does not possess a core that is roughly spherical, but rather one that is like a tube closed through the periodic timelike boundaries. In the spatial directions it is narrow which leads to a large $Q_{\max}$.

What we see then is that if we include only the "broad" topological fluctuations remaining after 25 cools, the suppression of χ_t as T rises through T_c is much sharper than if we include fluctuations of all sizes. This raises the exciting possibility that the part of χ_t due to physical length fluctuations may actually be an order parameter for the deconfining transition in the continuum limit. Of course at $\beta=5.7$ there is no clear separation between the narrow, $\rho \approx 0(a)$, fluctuations and those on typical dynamical length scales, $\rho \approx 0(\Lambda_{\text{mom}}^{-1})$. So what one needs to do is to repeat these calculations on L^{36} , L^{38} etc. lattices around $\beta=5.88$, 6.00 etc., so as to determine whether the functional form of the suitably cut-off χ_t is tending towards being discontinuous at T_c , with $\chi_t=0$ for $T > T_c$, as β is increased, or not.

7. Discussion

The fact that in the presently accessible range of couplings the lattice cut-off scale is not well separated from physical length scales, emphasizes the value of methods which are able to monitor directly the separation of these length scales as β increases. The method we have used in this paper has this desirable feature, by virtue of the fact that it can

monitor the sizes ρ of the cores of the topological fluctuations during the cooling procedure. We have measured the sensitivity of the susceptibility to changes in the range $\rho \gg \rho_0$ of core sizes included and have seen explicitly how this sensitivity decreases with increasing β , thus providing a direct demonstration of the desired separation of scale sizes.

While our method certainly reveals the ambiguities inherent in measuring the topological charge for $\beta \leq 6.0$, it would be obtuse to ignore the possibility of extracting physics from our measurements where there is evidence that it makes sense to do so. In particular, the fact that we observe the ratio $\chi_t^{1/4}/\sqrt{K}$ to be independent of β for $5.7 \leq \beta \leq 6.0$, with errors small enough for this observed β independence to be statistically significant, encourages us to extract a value for χ_t in physical units. Using the conventional value for the string tension, $\sqrt{K}=420$ MeV, and taking for χ_t the value measured on 16^4 lattices at $\beta=6.0$ - where the statistical errors are the largest but the systematic errors are the smallest - we obtain

$$\chi_t = (191 \pm 16 \text{ MeV})^4 \quad (7.1)$$

which is in very good agreement with the expectations of the Witten-Veneziano mass formula. This provides strong support for the idea that the large η' mass is indeed driven by the anomaly and associated topological vacuum fluctuations.

A second piece of interesting physics we have observed is the rapid suppression of χ_t as the temperature is raised through the deconfining temperature [4]. The advantage of examining this suppression in SU(3) rather than in SU(2) - where it was first observed [13] - is that the SU(3) transition is 1st rather than 2nd order. This makes it practical to ask whether χ_t is in fact an order parameter for the deconfining phase transition. We had some indications that this might be the case, in that the transition became sharper if we excluded the more pointlike charges from Q , but a definitive answer will require calculations over a range of β .

Since we can measure the location and distribution of "instanton" cores, we can in principle perform non-trivial tests of the dilute gas approximation.

In practice the only such test we have performed is to see whether the observed distributions of Q can be fitted by a convolution of Poisson distributions for positive and negative charges. We found the fit to be acceptable on the low temperature data, but the statistics are insufficient to draw any strong conclusions.

What should one do next? Certainly one should increase statistics at $\beta=6.0$, and, less urgently, at $\beta=5.9$: in our experience we find that about 500 more-or-less independent measurements are needed at any β to give errors that are as small as usefully needed. Also measurements at larger β are needed. However our measurements of effective correlation lengths indicate that critical slowing down will prevent one from measuring Q much deeper than $\beta \approx 6.2$ with present Monte Carlo methods. This has been shown by one of us in the β -range 6.3-6.9 [26]. Multigrid methods [28] will probably solve this problem once they have been implemented for SU(3). Nonetheless, accurate measurements of χ_t and K for $\beta \leq 6.2$ might well be enough to pin down whether the ratio $\chi_t^{1/4}/\sqrt{K}$ has attained its continuum limit. A second question that urgently requires further work is whether χ_t is indeed an order parameter for the deconfining transition. Calculations on L^{36} and L^{38} lattices near $\beta=5.87$ and 6.00 respectively, with various cuts on core sizes, could provide very useful information on this question.

Acknowledgement

One of us (M.T.) would like to thank RAL and especially the University of Southampton for their hospitality during the course of this work. Two of us (M.T. and J.H.) would like to thank CERN for its hospitality, and donation of computer time. We acknowledge SERC grant no NG11849 under which the 16^4 configurations were produced by the Edinburgh collaboration. The numerical calculations reported on were performed on the DAP at QMC, London, on machines at CERN and the Rutherford Appleton Laboratory, and on the Cyber 205 in Manchester. One of us (J.W.) thanks the SERC for financial support.

References

1. A.A. Belavin, A.M. Polyakov, A.S. Schwarz, Yu S. Tyupkin: Phys. Lett. 59B (1975) 85.
2. G. 't Hooft: Phys. Rev. Lett. 37 (1976) 8; Phys. Rev. D14 (1976) 3432.
3. E. Witten: Nucl. Phys. B156 (1979) 269.
G. Veneziano: Nucl. Phys. B159 (1979) 213; Phys. Lett. 95B (1980) 90.
J. Smit: talk at the 1986 Coseners House meeting on Lattice Gauge Theory.
4. J. Hoek, M. Teper and J. Waterhouse, Phys. Lett. 180B (1986) 112.
5. P. Di Vecchia, K. Fabricius, G.C. Rossi and G. Veneziano, Nucl. Phys. B192 (1981) 392; Phys. Lett. 108B (1982) 323.
6. P. Woit, Phys. Rev. Lett. 51 (1983) 638; Nucl. Phys. B262 (1985) 284.
Y. Arian and P. Woit, Phys. Lett. 169B (1986) 402; Nucl. Phys. B268 (1986) 521.
7. Y. Arian and P. Woit, Topological Susceptibility in SU(3) Lattice Gauge Theory, SUNY preprint, 1986.
8. I.A. Fox, J.P. Gilchrist, M.L. Laursen and G. Schierholz, Phys. Rev. Lett. 54 (1985) 749.
9. M. Luscher: Comm. Math. Phys. 85 (1982) 39.
10. A. Phillips, D. Stone: Lattice Gauge Fields, Principal Bundles and the Calculation of Topological Charge, SUNY preprint, 1985.
G. Lasher, A. Phillips, D. Stone: LBL Meeting on "Quark Confinement and Liberation", May 1985 (Proceedings: World Scientific, Eds. Klinkhammer and Halpern).
11. S. Coleman: Uses of Instantons (Erice Lectures 1977).
12. M. Gockeler, A.S. Kronfeld, M.L. Laursen, G. Schierholz, U-J. Wiese: Topology in SU(3) Lattice Gauge Theory: First calculation of the Topological Susceptibility, DESY 86-107.
13. M. Teper: Phys. Lett. 162B (1985) 357; 171B (1986) 81; 171B (1986) 86.
14. M. Teper: Instantons, topology and the $U_A(1)$ anomaly: a lattice Monte Carlo investigation of the quantised SU(2) vacuum, Oxford preprint (1986).
15. M. Peskin: 1978 thesis.

16. K. Fabricius, G.C. Rossi: Phys. Lett. 127B (1983) 229.
N.V. Machaldiani, M. Muller-Preußker: Calculation of the topological susceptibility in SU(3) lattice gauge theory, DESY L-Trans-280 (1983).
17. K. Ishikawa, G. Schierholz, H. Schneider, M. Teper: Phys. Lett. 128 (1983) 309.
18. B. Berg, Phys. Lett. 104B (1981) 475.
19. M. Luscher: Nucl. Phys. B200 [FS4] (1982) 61.
20. J. Hoek: Computer Physics Comm. 39 (1986) 21.
21. J. Hoek: Phys. Lett. 166B (1986) 199.
22. N. Cabbibo, E. Marinari: Phys. Lett. 119B (1982) 387.
23. E.-M. Ilgenfritz, M.L. Laursen, M. Muller-Preußker, G. Schierholz, H. Schiller, Nucl. Phys. B268 (1986) 693.
24. A.D. Kennedy, J. Kuti, S. Meyer, B.J. Pendleton: Phys. Rev. Lett. 54 (1985) 87.
25. Ph. de Forcrand, G. Schierholz, H. Schneider, M. Teper: Phys. Lett. 160B (1985) 137.
26. J. Hoek, Long time-scale correlations in topological quantities, Minnesota Supercomputer Institute preprint, 1986.
27. R.D. Pisarski, F. Wilczek: Phys. Rev. D29 (1984) 338.
28. J. Goodman, A.D. Sokal: Phys. Rev. Lett. 56 (1986) 1015.
A. Brandt, D. Ron and D.J. Amit, Multi-level approach to Discrete-State and Stochastic Problems, preprint 1986.
Ph. de Forcrand, talk at Conference on Perspectives in Computational Physics, Oct 29-31, 1986, Trieste.
J. Smit, private communication.

Table 1

β	<u>lattice</u>	<u>number</u> <u>measurements</u>	<u>$\langle Q^2 \rangle$</u>	<u>errors</u> <u>ind./batched</u>
5.60	8^4	69	4.46	0.78/0.87
5.70	8^4	474	2.65	0.16/0.14
	6^4	360	0.73	0.07/0.05
	4^4	150	0.000	0.007
5.73	8^4	159	1.69	0.15/0.20
	6^4	237	0.29	0.04/0.05
5.80	8^4	480	1.35	0.10/0.10
5.85	8^4	302	0.69	0.08/0.10
5.90	10^4	220	1.71	0.17/0.23
6.00	16^4	48	4.92	1.28/1.34

Table 2

β	<u>lattice</u>	$\langle Q^2 \rangle$ (naive)	$\langle Q^2 \rangle$ (dilute gas)	$\chi^2/\text{df.}$
5.60	8^4	4.46(78)	4.33(1.25)	.56
5.70	8^4	2.65(16)	3.00(22)	1.20
	6^4	0.73(7)	0.76(6)	.21
5.73	8^4	1.69(15)	2.08(23)	6.04
	6^4	0.29(4)	0.30(4)	.35
5.80	8^4	1.35(10)	1.41(9)	.86
5.85	8^4	0.69(8)	0.66(7)	.30
5.90	10^4	1.71(17)	1.79(20)	1.65
6.00	16^4	4.92(1.28)	3.42(1.38)	.51

Table

reproducibility on 8^4 at $\beta = 5.80$

configuration	Q (coords. $Q_{\max}$) sign charge		
	cool 1	cool 2	cool 3
A	1 (3,3,3,8)+	1 (3,3,3,8)+	1 (3,3,4,8)+
B	0 (3,2,5,2)+ (6,5,2,2)-	0 (3,2,5,3)+ (6,5,2,2)-	0 (3,2,5,3)+ (6,5,2,2)-
C	2 (1,7,2,8)+ (6,6,4,7)+ (3,6,8,1)- (7,1,1,6)+	1 (1,7,2,8,)+ (6,6,4,6)+ (3,7,7,1)-	1 (1,7,2,8)+ (6,6,4,6)+ (3,7,7,1)-
D	-1 (7,1,8,5)- (6,8,6,8)- (7,3,4,6)+	-1 (7,1,8,5)- (6,8,6,7)- (7,3,4,6)+	-1 (8,1,8,5)- (6,8,6,8)- (7,3,4,6)+
E	1 (3,4,1,6)+ $\bar{\text{II}}$ below cut-off	1 (3,4,1,6)+ $\bar{\text{II}}$ below cut-off	1 (4,3,1,5)+ (5,6,1,2)- I below cut-off

Table 4

configuration	normalized residual action	interpreted charge in [21]	charge with method in this paper
ED6029	.78	1	0
ED6007	1.05	1	-1
ED6037	5.98	6	-7
ED6047	1.07	1	1
ED6055	2.01	2	2

Table 5

correlations in Q

β (lattice)	δ	$\frac{\langle Q_i Q_{i+\delta} \rangle}{\langle Q_i^2 \rangle}$
5.80 (8 ⁴)	20	0.31±.06
5.90 (10 ⁴)	50	0.24±0.07
6.00 (16 ⁴)	112	0.24±0.14

Table 6

8³⁴ lattice topology

<u>β</u>	number <u>measurements</u>	<u>$\langle Q^2 \rangle$</u>
5.60	50	1.90(0.33)
5.65	54	1.28(0.20)
5.67	45	0.84(0.27)
5.70	50	0.48(0.15)
5.75	50	0.06(0.02)
5.85	40	0.000(0.025)

Figure Captions

fig 1 Cooling an 8^4 lattice from $\beta=5.8$:

- (a) action and topological charge as a function of the number of cooling sweeps;
- (b) testing self-duality on the action plateau;
- (c) maximum positive and negative values of $Q(n)$ on the lattice

fig 2 Examples of cooling runs on a 10^4 lattice at $\beta=5.9$.

fig 3 Examples of cooling runs on a 10^4 lattice at $\beta=5.9$ where the charge changes because of the disappearance of (anti-) instantons.

fig 4 Topological susceptibility at 25 cooling sweeps as a function of lattice volume, L^4 .

fig 5 Topological susceptibility as a function of β .

fig 6 The measured topological susceptibility in units of the string tension and Λ_L : search for continuum physics.

fig 7 Topological susceptibility from Q as measured as early as possible or after 10 cooling sweeps or after 25 cooling sweeps.

fig 8 (a) Two different cooling sequences from a single 6^4 configuration generated at $\beta=1.0$.
(b) Three different cooling sequences from a single 8^4 configuration generated at $\beta=5.8$.

fig 9 Q measured from each of a sequence of 90 8^4 Monte Carlo configurations generated at (a) $\beta=5.7$ (b) $\beta=5.8$ (c) $\beta=5.85$.

fig 10 Effective correlation length for Q in units of Monte Carlo sweeps for various β .

fig 11 Topological susceptibility at 25 cooling sweeps for various cuts on the minimum sizes of the topological fluctuation to be included.

fig 12 Comparison of χ_t obtained with the cooling method (from fig 5) and with geometric methods [7,12].

fig 13 Susceptibility on an 8^{34} lattice as its temperature increases (with increasing β) through the deconfining transition and a comparison with "zero" temperature values.

fig 14 As for fig 11 but for the high temperature 8^{34} lattice as well.

fig 15 As for fig 7 but for the high temperature 8^{34} lattice as well.

fig 16 Topological charge profile of an instanton on an 8^{34} lattice at high temperature: the "instanton" is a narrow tube closed on itself through the short periodic time boundary.

FIG.1.(a)

FIG.1(b)

FIG.1.(c)

FIG. 3

FIG. 2

FIG. 4.

χ_t AT ZERO TEMPERATURE

FIG. 5.

FIG. 6.

FIG. 7.

FIG. 8.

(b) 8^4 LATTICE AT $\beta = 5.80$

FIG.8.

SEQUENCES OF 8^4 CONFIGURATIONS

FIG. 9

FIG. 9

FIG. 10

FIG.11.

FIG.12

HIGH TEMPERATURE

FIG. 13

NARROW vs BROAD - HIGH T

FIG 14

FIG.15.

INSTANTON ON $8^3 4$ AT HIGH T

FIG. 16