PHYSICS OF THE EARLY UNIVERSE

APPLICATIONS OF THEORIES OF ELEMENTARY PARTICLES TO COSMOLOGY AT VERY HIGH TEMPERATURES

F.R. KLINKHAMER

Kast dissertaties

PHYSICS OF THE EARLY UNIVERSE

Goya: Comico Descumbrimento (Funny Discovery)
by permission of the Fitzwilliam Museum, Cambridge, England

PHYSICS OF THE EARLY UNIVERSE

APPLICATIONS OF THEORIES OF ELEMENTARY PARTICLES TO COSMOLOGY AT VERY HIGH TEMPERATURES

PROEFSCHRIFT

TER VERKRIJGING VAN DE GRAAD VAN DOCTOR IN
DE WISKUNDE EN NATUURWETENSCHAPPEN AAN DE
RIJKSUNIVERSITEIT TE LEIDEN, OP GEZAG VAN DE
RECTOR MAGNIFICUS DR. A.A.H. KASSENAAR,
HOOGLERAAR IN DE FACULTEIT DER GENEESKUNDE,
VOLGENS BESLUIT VAN HET COLLEGE VAN DEKANEN TE
VERDEDIGEN OP WOENSDAG 21 SEPTEMBER 1983
TE KLOKKE 16.15 UUR

door

FRANS RICHARD KLINKHAMER

geboren te Utrecht in 1956

promotiecommissie:

promotoren	Prof.Dr.F.A. Berends Prof.Dr.H. van der Laan
referent	Dr.F.A. Bais
overige leden	Prof.Dr.W.B. Burton Prof.Dr.J.A.M. Cox Prof.Dr.H.J. Habing Prof.Dr.H.C. van de Hulst Prof.Dr.J.H. van der Waals

aan de beste ouders

CONTENTS

Part II TEMPERATURES AT THE SCALE OF HADRONS

1. INTRODUCTION

The Universe is continuously expanding and in its earliest phase the matter content was very dense and hot. In order to describe this brief but important epoch one needs to know how matter, or better the elementary particles, behave under these conditions. This is the subject of high-energy physics, which studies reactions of particles coming from an accelerator. In this thesis I study the interface of early cosmology and elementary particle physics. The reader must have quite some background, namely the standard cosmological model based on the General Theory of Relativity and the physics of elementary particles and their interactions, but excellent textbooks on these subjects are available, e.g. [1, 2].

In this introduction I briefly outline our present knowledge of high-energy physics (section 1.1) and cosmology (section 1.2), which lies at the basis of the subjects discussed further on. In section 1.3 I give an overview of the contents of this thesis and show how the chapters interrelate. Finally some recent developments will be reviewed in section 1.4 and some concluding remarks are presented in section 1.5. The reader is advised to reread these last two sections after having digested the main chapters of this dissertation.

Natural units will be used with $h/2\pi=c=k=1$. Energy (= mass = temperature) is measured in units of GeV $\equiv 10^9$ eV and the mass scale associated with Newton's constant G of gravity is $M_{Planck} \equiv G^{-1/2} = 1.22\ 10^{19}$ GeV. To get the familiar cgs units we have the following relations: 1cm = $5.068\ 10^{13}$ GeV^{-1}, 1g = $5.610\ 10^{23}$ GeV, 1s = $1.519\ 10^{24}$ GeV^{-1}, 1K = $8.617\ 10^{-14}$ GeV. The symbol - indicates an estimated value and O(..) gives the order of magnitude.

1.1 Elementary particles and interactions

There are four types of interactions between elementary particles:

(1) the electromagnetic force;

(2) the weak force, which causes radioactive decay of a neutron into a proton, electron and antineutrino;

(3') the strong force, which binds the protons and neutrons together in the atomic nucleus;

(4) the gravitational force, which is so weak that an astronomical amount of particles is required to do something interesting, e.g. the solar system compared to an atom bound by the electric force.

The observed elementary particles are divided in two classes: hadrons and leptons. The crucial difference is that hadrons participate in the strong interaction, whereas leptons do not. This terminology also suggests that the hadrons are heavier than the (corresponding) leptons.

By now there is clear evidence that all hadrons are composed of more fundamental components: quarks. The hadrons are either baryons, which each contain three quarks, or mesons, which contain a quark and an anti-quark. The quarks have a new type of charge, called "colour", on which operates a new interaction:

(3) the colour force, which should give as a sort of spill over the strong interactions (3') between hadrons.

This colour-force has a most peculiar property: at very small distances ($\ll 1\ GeV^{-1}$) it is weak, but grows stronger over larger separations of the sources. This indicates the origin of the phenomenon of <u>confinement</u>: it is impossible to separate a hadron into isolated quarks.

The tool used to describe the elementary particles is Relativistic Quantum Field theory [2], which includes the effects both of Quantum

Mechanics and of Special Relativity. Over the space-time manifold fields are defined, which describe the creation and annihilation of (anti) particles. These fields, defined at each space-time point, may carry sets of indices, on which certain symmetry transformations operate. This means that although the fields are changed accordingly, the physical content remains the same. The force between two particles may be thought to arise from the exchange of certain (other) particles, e.g. the photon is the carrier of the electromagnetic interaction between electrons and the gluons carry the colour force between quarks or themselves. More specifically gauge theories are used to describe the fundamental interactions (1-4) and the particles on which they operate. The quintessence of a gauge theory is the following: the theory is invariant when the fields, of the electrons, say, undergo a local (or gauge) transformation in some internal symmetry space; "local" means that the parameters of the transformation can be chosen arbitrarily in each space-time point; in order to have invariance there are spin-1 gauge fields, for example the photon field; these gauge bosons then are the carriers of the interaction. Recall that particles of (half-) integer spin have (Fermi-Dirac) Bose-Einstein statistics and are called (fermions) bosons.

For the gauge theories of the electromagnetic, weak and colour forces quantum effects can be calculated. These theories are characterized by the specific form of the gauge symmetry, which has the structure of a Lie group. U(N) is the group of the NxN unitary matrices and for SU(N) these matrices have unit determinant. The standard model is

(1+2) electro-weak theory (Glashow-Weinberg-Salam model, GWS):

symmetry group = $U(1) \otimes SU(2)$

gauge bosons : W^+, W^-, $Z°$ (intermediate vector bosons)

γ (photon)

(3) Quantum Chromodynamics (QCD):

symmetry group = $SU(3)$

gauge bosons : eight gluons g.

To specify the theory completely one must also give the representations of the matter fields under these symmetry groups: for the GWS model see table 1 of Chapter 2, for QCD the quarks are in the vector ($\underline{3}$) representation of SU(3).

The gauge bosons of electromagnetism and colour are massless. But the intermediate vector bosons must be very heavy, of the order of 80 GeV. In the GWS model this is achieved by the Higgs mechanism of spontaneous breaking of a local symmetry, which can be sketched as follows:

1) introduce scalar fields ϕ into the theory;
2) the selfinteractions of these fields, described by a potential $V(\phi)$, is very special: $V(\phi)$ is invariant under the symmetry, but its minimum $\phi=v$ is not;
3) for gauge theories this gives some gauge bosons a mass of order v; also fermions ψ may get a mass $\sim \lambda v$ from a Yukawa term $\lambda\bar{\psi}\phi\psi$ in $V(\phi,\psi)$, where λ is a small coupling constant;
4) in the GWS model ϕ is a doublet under SU(2) and $v \sim 300$ GeV.

So far it has not been possible to quantize gravity, which is described classically by Einstein's General Theory of Relativity. For laboratory experiments with particle accelerators the gravitational force can be neglected. In the early Universe as described in this thesis we need only classical gravity to set the stage for the processes of the other interactions. In the following we only consider the forces (1, 2, 3).

All experiments (center of mass energy $\lesssim 50$ GeV) fit in the standard model, i.e. GWS and QCD. But is the standard model all there is? If not, what happens at interaction energies $\gtrsim v$? There appear to be two alternative roads towards a more powerful theory.

1) Grand Unification Theory (GUT). The catchwords are simplicity and straightforward extrapolation. One assumes that the Higgs mechanism of the GWS model is correct in its simple form and that the scalars ϕ are elementary. At an energy of ~ 100 GeV we know the effective interaction strengths $\alpha_i \equiv g_i^2/4\pi$, i=1,2,3, for the three gauge groups U(1), SU(2) and SU(3), respectively. From the renormalization group theory we know how the $\alpha_i(E)$ evolve towards higher energy E and find that the three α_i become equal at $M_U \sim 10^{15}$ GeV! It is assumed that the "desert" between 10^2 and 10^{15} GeV is empty of new phenomena. This suggests that there is a GUT for $E > M_U$ with one simple gauge group G ($\supset$ SU(5)), which is broken by a Higgs mechanism with $\phi_{min} = V \sim M_U$ to the smaller group U(1) x SU(2) x SU(3). For $E < M_U$ the α_i evolve away from their common value at $E \sim M_U$. In short, the three observed forces are the low energy remnant of one unified underlying force. GUTs may answer some open questions of the standard model, because the unified theory (G) sets constraints on its U(1) x SU(2) x SU(3) part. Also they make some spectacular predictions: proton decay and the physics of the very early Universe at temperatures $T \sim M_U$. For more details see Chapter 2 and references therein.

2) Compositeness. The catchword is "naturalness" [3]. The scenario 1) above required some miraculous fine-tunings in its symmetry breaking mechanism.

It may be that the motivation for this will be found in the future. But following the idea of "naturalness" one should allow only small parameters if they are guaranteed to vanish under some symmetry, e.g. small fermion mass by chiral symmetry. In this way one might be led to compositeness models, where the Higgs scalar field ϕ is a collective excitation and also the "observed" quarks and leptons are composed of more fundamental preons. The dynamical mass scale of the new preon interactions M_{preon} must be above $\sim 10^3$ GeV and one needs an unbroken (chiral) symmetry to keep the composite quarks and leptons nearly massless. Alas there does not exist (yet) a viable candidate model; for a review see [4]. Anyway it is clear that there would certainly be no "desert", but rather a jungle of strong interaction phenomena and new particles at energies $\gtrsim M_{preon}$, where the preons bind together.

Experiment must decide between these two alternatives, or others. If the Z° were found with a mass precisely as predicted in the simple GWS model, the naive treatment of the symmetry breaking and the unification scheme would become much more plausible. Of course, direct evidence [5] for proton decay and/or a superheavy magnetic monopole associated with the breaking of G to some $G' \times U(1)$ at M_U, would confirm option 1). Similarly, new phenomena at energies of order 100 GeV would point towards option 2).

1.2 Cosmology

1.2.1 Standard model: Hot Big Bang

Cosmology studies the global structure and history of the Universe. The only force operating over these large distances is gravity. [Note that electric charges can be screened, but not so for gravity, where all the "signs" are equal or, in other words, the forces always attractive]. Hence we need to have some ideas on the gravitational force. Newton already tried to build an infinite, homogeneous and static Universe, but ran into the problem how to avoid gravitational collapse [6]. We now know that his gravitational potential was undefined, cf. section 15.1 of [1]. Only with Einstein's General Theory of Relativity (GTR), which incorporates the Newton theory, can we tackle the Universe. The Einstein field equations are, cf. [1],

$$R_{\mu\nu} - \tfrac{1}{2}\, g_{\mu\nu}\, R^{\lambda}_{\lambda} = -8\pi G T_{\mu\nu}. \tag{1}$$

$R_{\mu\nu}$ is the Ricci tensor build from the metric $g_{\mu\nu}$, where μ and ν run from 1 to 4, and describes the space-time curvature. $T_{\mu\nu}$ is the energy-momentum tensor of the matter. Equations (1) relate the structure of the space-time manifold (lhs) to the energy content of the matter (rhs). The constant of proportionality is Newton's constant G, which measures the strength of the gravitational interaction. Of course, (1) looks deceptively simple: these are 10 (not 16 by $\mu\nu$-symmetry) non-linear differential equations.

While the Newton force law

$$\vec{F}_i = - \frac{G\, m_1 m_2}{\left|\vec{r}_1 - \vec{r}_2\right|^3} (\vec{r}_1 - \vec{r}_2) \tag{2}$$

only gives the interaction between two sources at spatial positions $\vec{r}_i$ with masses m_i (i=1,2), the Einstein equations (1) are global and thus more appropriate to attack the structure of the Universe. [This is a trifle demagogic, instead of (2) I could have written down the Poisson equation]. Indeed in the year following the discovery of the field equations (November 1915) Einstein was the first to study modern cosmology. It is well known that he introduced a term $-\Lambda g_{\mu\nu}$ in the lhs of (1) in order to have a static solution. Later he dropped this cosmological constant as his "greatest error", but see below for its renaissance.

Without this Λ term A. Friedmann (1922) found that the Universe expands with time t. For a homogeneous and isotropic space-time with scale factor R(t) filled with an ideal gas (pressure p, density ρ) he derived from (1)

$$\left(\frac{dR}{dt}\right)^2 + k = \frac{8\pi G}{3}\rho R^2 \tag{3a}$$

$$\frac{dp}{dt} R^3 = \frac{d}{dt}\left(R^3(\rho+p)\right). \tag{3b}$$

k = -1, 0, +1 is the curvature constant for an open, Euclidean (flat) or closed Universe. (3a) describes the expansion, whereas (3b) is similar to the thermodynamic equation of energy conservation. We also need the equation of state

$$p = p(\rho). \tag{3c}$$

For details see [1].

Modern cosmology is based on (3) together with the following two observations:

1) E. Hubble (1929) found a (linear) relation between redshift and distance of far galaxies, which shows that the Universe indeed expands;

2) A. Penzias and R. Wilson (1965) discovered a Cosmic Background Radiation (CBR), which we now know to have a blackbody spectrum with T_o = 3K. For relativistic matter ($p \propto \rho$) and non-relativistic matter (p=0) we have from (3) $\rho_{rel} \propto R^{-4}$ and $\rho_{non-rel} \propto R^{-3}$. Presently normal non-relativistic matter is dominant, but in the early phase of the Universe (R smaller)

relativistic matter such as the photons of the CBR must have been dominant.

In this way one arrives at the Hot Big Bang model, where "hot" refers to the early relativistic phase. In this relativistic phase (3a) gives us $R \propto \sqrt{t - t_0}$, which means that our presently expanding Universe started at a certain moment t_0 when matter was infinitely hot and dense (R=0): the Big Bang. See the next paragraph for a remark on this "singularity" at t_0. When the temperature of the Universe cooled to ~ 1 MeV nuclear reactions synthesized a large abundance of ^{4}He. [Historically, G. Gamov and collaborators predicted in 1948 from their nucleosynthesis calculation a low temperature CBR]. Also very specific abundances of ^{2}H, ^{3}He and ^{7}Li arise in the model, which appear to agree with the observations [7]. Hence the Hot Big Bang model incorporates naturally the observed homogeneity and expansion of the Universe together with the radiation background and explains how the primordial abundances have originated.

1.2.2 Early Universe

Hopefully it is clear why we feel confident to explore the further consequences of the Hot Big Bang model, especially in the epoch far before nucleosynthesis. In 1970 Weinberg [1] could not really do this, because it was not clear how matter would behave at $T \gtrsim 1$ GeV, where nucleons would touch or even overlap, cf. our Part II. But in section 1.1 I discussed the recent advances in the field of elementary particle physics and the standard SU(3) x SU(2) x U(1) model allows to trace the Universe back to temperatures $\lesssim 100$ GeV. Even earlier epochs can be investigated if the standard model is a remnant of a unified interaction. If on the other hand compositeness holds true then the precise history for $T \gtrsim M_{preon}$ is unknown, compositeness theories being still in a rudimentary state.

Before I give the highlights of the early Universe I must introduce the reader to the important physical processes that may occur. Notice that one can study the relativistic phase of the Universe, because there is local thermodynamic equilibrium nearly all the time, which allows a simple description. But no observable quantities can arise from equilibrium, so that the brief moments of transition are very important to us. For example at $T \sim 1$ MeV the weak interactions no longer keep up with the expansion rate of the Universe and the neutrinos decouple from the rest of the matter. Because of this transition and later reheating of the photons the present neutrino temperature should be lower than that of the photon background by a factor $(4/11)^{1/3}$, which in principle is an observable difference. Now the important processes:

1) Quark deconfinement. At T=0 the dynamical property of quark confinement is absolute, it has to do with the non-perturbative structure of the QCD vacuum. But at $T > T_c \sim 0.2$ GeV the confinement property vanishes: quarks and gluons move freely, although Debye screened. For $T \gtrsim 5\ T_c$ we have a nearly ideal gas of quarks, leptons and gauge bosons. This simplicity makes it possible to discuss the phenomena 2) and 3) at much higher temperatures.

2) Phase transitions or symmetry restoration at high temperatures. Consider the potential $V(\phi)$ of the elementary scalars ϕ. At T=0 it must have a non-symmetric minimum, so that through the Higgs mechanism some gauge bosons acquire mass (sect. 1.1). But at very high temperatures the effective potential is different. In fig. 1 symmetry restoration is illustrated for the discrete symmetry $\phi \rightarrow -\phi$ and where the "ball" indicates what the groundstate is.

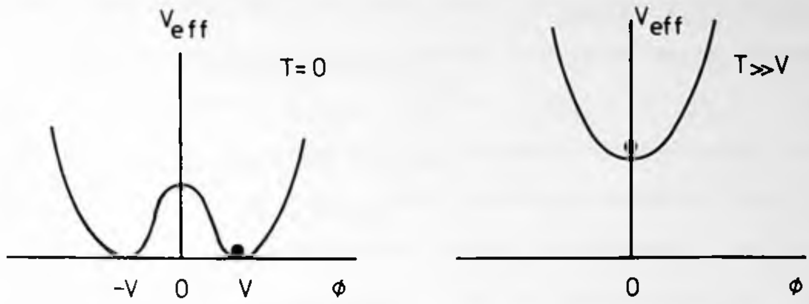

Fig. 1

This symmetry restoration should occur for the Glashow-Weinberg-Salam and Grand Unification models, i.e. at temperatures of the order of 10^2 and 10^{15} GeV. The detailed dynamics in the early Universe is very interesting and during the GUT transition superheavy magnetic monopoles may have been created (see our Part I).

3) Baryon number creation. GUTs generally lead to reactions which violate the conservation of baryon number B. A (anti)quark has $B = (-1/3)1/3$. At low energies these interactions are very weak, suppressed by powers of $1/M_U$, and predict a large, but finite, lifetime of the proton. In the early Universe when $T \sim M_U$ these reactions have their "normal" strength. In fact they may provide the answer why the present Universe has a dominance of baryons over anti-baryons (net baryon number).

Fig. 2a gives the scenario of the first second of the Universe. Note that for $T \gtrsim M_{Planck}$ one does not know what "happens", because the quantum effects of gravity are not understood. In this "epoch" the Einstein equations (1) are not valid. Even the idea of a global cosmic time t [1]

(a)

ρ[g cm^{-3}]	t[sec]	T[GeV]	
10^{95}	-10^{-43}	10^{19}	QUANTUM GRAVITY?
10^{79}	10^{-37}	10^{15}	GUT Phase Transition *
			Baryon number creation
10^{27}	10^{-11}	10^{2}	ELECTROWEAK (GWS) Phase Transition
10^{15}	10^{-5}	10^{-1}	$\frac{\text{Quarks}}{\text{Hadrons}}$ **
10^{5}	1	10^{-3}	Nucleosynthesis
~	~	~	
10^{-20}	10^{6} yr	4000 K	Recombination to atoms
10^{-29}	10^{10} yr	3 K	Present!

(b)

T[GeV]		
??		
$M_{Preon}=10^{3}$?	Preon confinement transition	$\frac{\text{preons}}{\text{quarks + leptons}}$
10^{2}?	Electroweak transition	
10^{-1}	Quark confinement transition	$\frac{\text{quarks + leptons}}{\text{hadrons leptons}}$ **

Fig. 2. History of the Early Universe if (a) Grand Unification Theories or (b) some compositeness model holds. Numerical values are only indicative. Parts I and II of this thesis discuss the epochs indicated by * and ** ,respectively.

may not work, hence the quotation marks above. We set t=0 at some temperature just below M_{Planck}, the precise definition is not so important because the error is $\Delta t \sim M_{Planck}^{-1} \sim 10^{-43}$s.

If the physics for energies $\gtrsim M_{GWS}$ involves new interactions (between preons, say), the earliest phase of the Universe is uncertain (fig. 2b). If all present particles are made out of strongly bound preons, they probably are liberated at $T > T_{c,preon} \sim M_{preon}$, just at quark deconfinement at 0.2 GeV. Also the behaviour at $T \sim M_{GWS}$ is expected to be different than as described in 1). The physics should then be understood better, before one can talk in detail about the very early Universe.

1.3 Overview

This thesis consists of two parts that address very different physics and energy regimes. Both parts start with a general chapter, where the references have been updated with respect to the original publications.

In Part 1 the very early Universe is explored when interactions from Grand Unification may operate. Especially I consider the transition from the unified gauge group G to the SU(3) x SU(2) x U(1) phase as the Universe cools below temperatures $T \sim M_U$ and Chapter 2 contains a general discussion of what may happen. There are two types of transition possible: second or first order, where the value of ϕ_{min} of the effective potential $V(\phi,T)$ goes to a non-zero value continuously or with a jump, respectively. The first order transition may happen at a very low temperature (supercooling) compared to the transition temperature if it would be second order. Also for a first order transition the precise dynamics and its feed back on the global expansion of the Universe may be spectacular. During the supercooling epoch the Universe expands exponentially, the scale factor grows as $R \propto \exp(t/\tau)$ instead of the normal $R \propto \sqrt{t}$. [This is just as in a de-Sitter Universe, which results from adding a non-zero cosmological constant to (1) and (3a).] Guth [9] has argued that, if the exponential expansion phase was long enough and the temperature after reheating $T_R \sim M_U$, the smoothness and flatness of our observed Universe could be explained. This strong expansion phase could perhaps also reduce the number density of magnetic monopoles to sufficiently low levels. However the main problem that P. Hut and I pointed out was how to end this supercooling, while keeping the density of the Universe homogeneous enough. Later this problem was discussed in more detail in [8]. The recent "new inflationary Universe scenario" (section 1.4) might resolve this problem.

Chapter 3 presents some figures with the results of numerical calculations on how the matter-antimatter asymmetry is created after the phase transition. Normally one neglects this complication of the "initial data" and indeed we find that the final asymmetry is quite independent of them. These figures illustrate the brief discussion in sections 4 and 6.2 of Chapter 2. For details the interested reader is referred to the original publication.

In Chapter 4 we argue that supercooling below $T^* \sim T_c^2 M_{Planck}^{-1}$ does not occur. The reason is that supercooling is a delicately balanced proces, which may be tipped over by expected(?) gravitational effects. If the supercooling ends at T^* the inflation is far less than required for Guth's original scenario [9]. Also the new scenario may be invalidated by this process, see section 1.4.

The major violation of "naturalness" in GUTs is the fine tuning of the Higgs potential in order to have $M_{GWS}/M_U \sim 10^{-13}$, this is called the hierarchy problem. In the last two years supersymmetry (a symmetry between particles of different spin) was invoked to alleviate this problem. If one sets the parameters of the tree level potential $V(\phi)$ so that $M_{GWS}/M_U=10^{-13}$, the special fermion-boson cancelations make that this ratio remains unchanged by higher order quantum corrections. In the standard GUTs one must tune many orders so that the final ratio is 10^{-13}. Even better, Witten suggested a mechanism that might also explain the smallness of this ratio. But in Chapter 5 I show that this model cannot generate the baryon asymmetry, because of the peculiar dynamics of the phase transition of Grand Unification, contrary to the "standard" case of Chapter 3. This problem seems quite general if supersymmetry is broken at a mass scale far below M_U, see paragraph 1.4.4 below.

Another implication of GUTs for cosmology is that they allow for a small, but non-zero, restmass of the neutrino (ν). I found that a mass in the range 10 eV $\lesssim m_\nu \lesssim$ 100 eV would be very important for the formation of structure in the Universe [10]. These articles are not included in the present thesis for the following two reasons: 1) experimentally and theoretically it is not clear that m_ν is so large, i.e. a mass significantly above $\sim M_{GWS}^2 M_U^{-1}$; and 2) the actual formation of structure occurs only when $t \gtrsim 10^8$s. The rôle of superheavy ($> 10^5$ GeV) right-handed neutrinos, the partners of the light ones mentioned above, or other superheavy fermions, in the early Universe is not very interesting [11].

In Part II I consider the phenomenon of quark liberation at high temperatures, $T > T_c \sim 0.2$ GeV. Here the starting point (QCD, see sect. 1.1) and the problems addressed are better defined than those of Part I. Although the energy scale is modest ($\sim 10^0$ GeV) the theoretical sophistication to solve this problem must be great. Ultimately the methods and knowledge obtained in solving the non-perturbative QCD should help us to understand what the Higgs mechanism, as applied in Part I, really means. High temperature quark deconfinement has the following direct applications:
1) Heavy ion collisions in accelerator experiments: for a head-on collision with center of mass energy per nucleon $\gtrsim$ 1 GeV a "fireball" is expected to be created with $T > T_c$. For details and references see [12].
2) The dynamics of the quark-hadron transition in the early Universe; see section 1.5 below for a further remark.

Theoretical expertise gained from the study of quark (de)confinement could be very useful if compositeness occurs, which should on the one hand be similar to and on the other hand differ significantly from QCD, or if 't Hooft's idea [13] of "tumbling" to larger and larger SU(N) theories for increasing energy contains some truth. In both cases the early Universe would differ from that sketched in fig. 2a and Part I. Even more speculative is the possibility that classical gravity is an induced quantum effect, with Newton's constant G arising in some way analogous to the mass scale of QCD [14]; in the very early Universe with $T \gtrsim M_{Planck}$ there would then be drastic changes in $G_{induced}$ and $\Lambda_{induced}$, somehow analogous to those of deconfinement in QCD.

But first we must get hold of the bear's skin, namely the physical mechanism that liberates the quarks in high temperature QCD. Chapter 6 reviews some background and theoretical ideas on this matter. As said above, the deconfinement transition is a change in the nature of the QCD vacuum. For zero temperature we have the successful bagmodel of hadrons: a finite region of perturbative vacuum, where quarks move freely, is immersed in the sea of impenetrable true vacuum. The non-perturbative vacuum may be thought of as having a zero "dielectric constant". Of course, this picture cannot tell us why the deconfining transition occurs, because the two vacua structure is put in by hand. One needs to understand how this phenomenologic model arises from the fundamental theory, i.e. QCD. Only then can one hope to understand what changes at high enough temperature. In Chapter 7 I show how deconfinement takes place, if one follows a certain ansatz as to how the dielectric model arises.

Whereas Chapter 7 starts the attack of the deconfinement problem from the continuum theory, one may also start already from a lattice theory. Here the fields do not live in the continuum space-time but on a discrete subspace, the lattice. Quark liberation may then be studied in the strong coupling expansion, see also paragraph 1.4.5 below. But in order to establish the physical temperature T_c, we need to take the continuum limit, which lies in the weak coupling regime. G. Mack has presented an idea on how the zero temperature confinement might persist in the weak coupling regime. In Chapter 8 I show that it is easy to understand why this fails at high enough temperatures.

It will be clear from the above and from Chapters 6-8 that a precise understanding of quark liberation will come if and only if one knows how the zero temperature theory gives confinement. In that case we have to establish explicitly how the continuum theory and the weak coupling lattice theory are linked. These remarks may seem trivial, but are nevertheless true. Also the continuum limit of the lattice gauge theory is hard to study numerically. Analytical methods should be used preferentially. One of the most successful methods may be the 1/N expansion. The SU(N) gauge theory without matter fields has only one free parameter, namely the number of "colours" N. It has been known for some time that we may get an understanding of the structure of the relevant SU(3) theory by considering the limit case SU(∞). [For some caution on the problem of confinement in the $N \to \infty$ limit see paragraph 1.4.5 below]. Even better would be to consider the limits N, $N_F \to \infty$, N/N_F fixed, if we introduce N_F flavours of "quarks" in the fundamental representation $\underline{N}$. This is because the relevant theory both has N=3 colours and 3 light quarks, named up, down and strange. Hence it is extremely interesting to consider the pure SU(N), $N \to \infty$, gauge theory on the lattice. Recently there has been significant progress for this case: nearly all space-time dependence can be neglected for $N \to \infty$. In Chapter 9 and 10 I consider this reduced theory. First I discuss what happens physically and then I address the non-trivial problem how finite temperature can be incorporated. Finally in Chapter 11 P. van Baal and I construct an elegant model, which looks very promising to study large-N (de)confinement both numerically and analytically.

1.4 Recent developments*

The subject matter of this thesis evolves rapidly. In this section I mention some relevant new results ($\leq$ January 1983). For more details the reader is referred to the original literature. First I consider four topics relevant to Part I.

1.4.1 New Inflationary Universe

The structure of Guth's article [9] is: if so much inflation occurs, then these problems are solved. But the interesting point is that first order phase transitions from GUTs may provide the "if". Alas, it turned out that there were severe problems with the reheating, cf. Chapter 2, 3 and [8], and that even apart from that the inflation would not be large enough (Chapter 4). Recently Linde and Albrecht and Steinhardt [15] realized that for potentials $V(\phi)$, which are sufficiently flat around ϕ=0, there is an extra source of inflation: after supercooling to $T \sim 10^8$ GeV (fig. 3a) a fluctuation region (FR) of size $\sim T$ makes a jump from $\phi = 0$ to $\phi \sim T$, which evolves towards the true vacuum at $\sim 10^{14}$ GeV (fig. 3b); this roll-over is quite slow and meanwhile the FR still expands exponentially. The present Universe would only be a tiny part of this vastly inflated FR and will thus be smooth and nearly Euclidean. Also the magnetic monopoles would lie at the boundary of our FR and we will not be bothered by them for a long time. In the FR the ϕ upon arriving at the true vacuum will oscillate rapidly, damp by decaying in particles, and thus reheat the FR smoothly to $T_R \sim 10^{14}$ GeV, which assures the generation of the baryon asymmetry [16]. In one stroke many problems of our Chapter 2 are evaded, but there remain three major ones:

*to be (re)read after the main chapters of the present thesis.

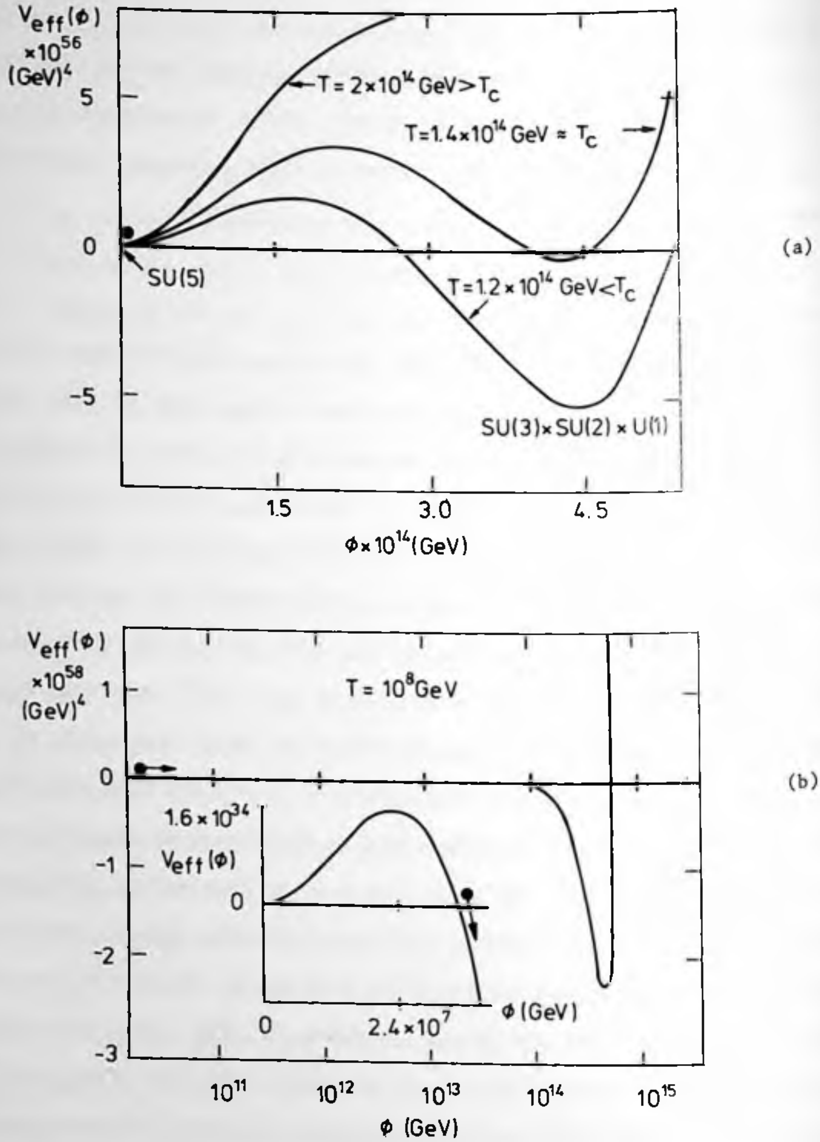

Fig. 3. The potential for the new inflationary Universe scenario as used by Albrecht and Steinhardt [15]. Notice the linear (a) and logarithmic (b) scales of the ϕ-axis.

1) The scalar potential $V(\phi)$ must be very special, named after Coleman and E. Weinberg, with a vanishing $(\text{mass})^2\phi^2$ term. This is very unnatural, but perhaps there is a yet unknown reason for it. But if the theory is coupled to gravity, the total $((\text{mass})^2 + \xi(\text{curvature}))\phi^2$ term must also vanish. This looks very strange: the particle physics has a fine tuning, which involves gravity through (curvature) = G x vacuum energy density. Note that the socalled minimal coupling $\xi=0$ is not at all minimal through its quantum effects [19]. This problem might be linked with the familiar one of the vanishing present cosmological constant Λ.

2) During the roll-over, fluctuations in ϕ develop which lead to density perturbations [17]. The good news is that these perturbations are nearly constant over all scales (< size of FR, of course), which is what one would like to find, cf. discussion in ref. [10]. The bad news is that the density amplitude δ appears to be much too large. To estimate δ some approximations were made and it is possible, but not probable, that a correct analysis gives a much smaller δ.

3) It is not clear that the Universe <u>can</u> cool to $T \sim 10^8$ GeV, or in other words, that the ϕ makes such a small jump in the FR. If the relevant scale is more like $T_H \sim M_U^2 M_{\text{Planck}}^{-1}$ there might be not enough inflation during roll-over, see the next paragraph.

1.4.2 <u>Gravitational effects on the supercooling phase</u>

In Chapter 4 we argued that quantum gravity effects terminate the supercooling at temperatures $\sim T_H$, which is above those required for the inflation scenario, old or new version [18]. Our claim was substantiated in [19], where also the vacuum contribution in a curved background to $\langle\phi^2\rangle$ was calculated, which may destabilize the false vacuum at T^*. The value of T^* depends on many details, but lies in the range T_H to $(T_C T_H)^{\frac{1}{2}}$, with the

Hawking temperature $T_H \sim 10^{10}$ GeV and $T_C \sim 10^{14}$ GeV typically. In [20] it is claimed that for certain parameters the (closed) Universe makes a homogeneous transition out of the false vacuum, so that there is enough inflation. But it is not clear what happens physically, e.g. they use the tunneling amplitude to arrive at the top of the barrier (cf. fig. 3b).

1.4.3 Fermion dynamics around a magnetic monopole

In [21] it is argued that a magnetic monopole significantly modifies the fermion vacuum around it up to distances of the order of 1 Fermi. This could lead to reactions that violate the conservation of fermion number, and possibly baryon number, with cross-sections determined by the strong interactions! If this is correct (there are still some problems to derive the proton decay rate) magnetic monopoles would strongly catalyze the decay of nearby protons.

In turn it is then quite trivial to see how many monopoles are allowed to hang around, notably using the heating of old neutron stars [22]. This conflicts with the flux implied by Cabrera's candidate event [5] of one Dirac magnetic charge g. Note that such a charge would be indirect experimental evidence that the SU(3) of colour is not broken at very low energy [23]. As said the inflation scenario predicts no monopoles in our part of the Universe.

1.4.4 Supersymmetric GUTs

The problems of Chapter 5 for a supersymmetric GUT in the early Universe, namely the dynamics of the phase transition and the baryon number creation, hold quite generally. Apart from the insufficient reheating ($T_R \sim$ scale of supersymmetry breaking $\ll M_U$) there are problems for the theory itself to generate enough baryon asymmetry [24]. Another insurmountable problem seems to be how to end up in the correct vacuum [25]. Perhaps the transition is helped somehow by strong coupling effects for $\alpha_{GUT} \sim 1$ at $T \sim 10^{10}$ GeV [25, 26], after which the baryon number can be created if there happen to be scalars with mass $m_H \sim 10^{10}$ GeV [26]. Also Witten's upside down model [ref. in Chapter 5] with supersymmetry breaking scale $M \sim 10^{10}$ GeV and $M \lesssim M_H \lesssim M_U$ has the problem that either (if M_H high) the created baryon number is not enough or (if M_H low) the proton decays too fast [27].

Apart from all the cosmological problems it is difficult to get the correct low energy theory, certainly if one wants to keep the model simple.

We can be somewhat briefer on the new results relevant to our Part II.

1.4.5 SU(N) lattice gauge theories (LGTs)

In Chapter 6 I mentioned that Weiss had calculated the effective potential for the $L(\underline{x})$ variable in the weak coupling regime. Polonyi and Szlachanyi [28] have done the same for pure SU(2) in the strong coupling expansion, and find that $\langle L \rangle_T$ shifts away from 0 for high enough temperature, apparently a second order phase transition. More generally Borgs and Seiler [29] derived for SU(N) a bound on where the phase transition line must lie in the g^2-T plane. But their bound, which is independent of the "magnetic" terms in the action, gets worse towards the continuum limit ($g^2 \to 0$).

The most interesting development is the theoretical [30] and numerical [31] indication that the deconfinement phase transition for the pure SU(3) theory is first order. Svetitsky and Yaffe [30] argue as follows: The effective $L(\underline{x})$ theory is both for high and low temperature short ranged. Assume that this holds for all temperatures then the N=2 and 3 cases are very different, namely N=2 will be in the universality class of the three dimensional Ising model, which has a second order phase transition, whereas N=3 is in the same class as the three state Potts model, which seems to have a first order phase transition. This appears to be confirmed by the numerical work of [31], where N=2 and 3 behave very differently. Notice that for N=3 the jump $\langle L\rangle_T$ at $g^2_{crit}(N_t)$ shrinks rapidly when one goes to a finer lattice (larger N_t, cf. Chapter 8). In [31] they also calculated the chiral symmetry restoration temperature T_{ch} with Kogut-Susskind fermions and neglecting fermion loops. The order parameter $\langle\bar{\psi}\psi\rangle_T$ vanishes at T_{ch} (cf. Addendum in Chapter 6):

$$SU(2):\ 1.00 < T_{ch}/T_c < 1.30 \text{ (smooth)}$$

$$SU(3):\ 1.00 < T_{ch}/T_c < 1.05 \text{ (discontinuous)}.$$

Now some new results on the zero temperature phase diagram, or in other words, the permanence of confinement (cf. Chapter 8). Numerical simulations [32] indicate that in the continuum limit (lattice spacing $\to$ 0) Lorentz invariance indeed is restored, but perhaps there are some complications as to when already asymptotic scaling sets in [33]. On the theoretical front Tomboulis [34] seems to have proven that the SU(2) LGT is in the confinement phase for <u>all</u> couplings. He uses a bound on the electric flux in a periodic box obtained by thinning out variables à la Migdal and Kadanov ("block spin"). If larger and larger blocks are made one arrives at the strong coupling LGT where confinement is trivial. For high enough temperature T one evidently does not get the bound, because of the 1/T size limitation. All this is very much in the spirit of our Chapter 8.

Finally let us discuss the large-N theories. There is the as yet unanswered question how (dis)similar the confinement of the SU(3) theory is compared to that of the $N \to \infty$ theory. In the one case the center, which may be essential for confinement [35], is discrete Z(3), whereas it tends to the continuous U(1) for $N \to \infty$. Also the string tensions of adjoint (A) and fundamental (F) representations of the gauge fields are related for $N \to \infty$ by $\sigma_A = 2\sigma_F$ [36]. This would be difficult to explain from the point of view of [35], because the adjoint representation of SU(N) is blind to the center Z(N). Perhaps this problem may be solved if the elusive masterfield [37] can indeed be constructed with the reduced model [38].

1.5 Concluding remarks

Already in section 1.1 I said that there are (at least) two crucial experiments for Grand Unified Theories (GUTs):

1) the Z^0 mass as predicted by the Glashow-Weinberg-Salam model $M_Z/M_W = 1/\cos\theta_W$ and $M_W \simeq 37$ GeV/$\sin\theta_W$, where θ_W is the Weinberg mixing angle [39];

2) a proton lifetime of the order of 10^{31} yr.

If confirmed this would point towards the simple picture of GUTs and one can then discuss seriously the processes of the very early Universe (Part I). In my opinion supersymmetric GUTs are too complicated to believe, certainly if they must overcome the problems of paragraph 1.4.4 and produce the measured value of $\theta_W = 28°$, which can be explained so beautifully by the desert of the standard GUT picture. Of course, it may be that (local) supersymmetry is an ingredient of quantum gravity.

How then would the history of the very early Universe run? The main problem is to avoid the creation of too many magnetic monopoles. My guess is the following: 1) at $\sim 10^{15}$ GeV there is a dull phase transition (second, or weakly first order); 2) the number of monopoles is reduced by thermal fluctuations if the Higgs masses are quite large typically [40];

3) the baryon number is created just after the phase transition, this requires an extension of the minimal SU(5) GUT (cf. Chapter 2); 4) the problems of smoothness, flatness and the origin of the small density perturbations for galaxy formation, all are relegated to the epoch of quantum gravity. [Of course, we can do this only if there is no later inflation, which washes out these perturbations.]

I do agree that the new inflationary Universe [15] is very interesting, but "complete solutions to the age-old problems always go down better if well salted" [41]. Some grains of salt were added in paragraph 1.4.1.

We need a better understanding of confinement and the Higgs mechanism, or in general of non-perturbative phenomena in the standard SU(3) x SU(2) x U(1) model. If the two experiments mentioned at the beginning should turn out very differently one should also consider the possibility of compositeness [4]. This is why I turned to the, seemingly less spectacular, problem of quark deconfinement in high temperature QCD (Part II). In the investigation of this problem we may learn a great deal; for some other (in)direct applications see section 1.3 above. As to cosmology I think the Universe just went through the quark-hadron transition without producing notable results. [The alleged mechanism that should produce the large perturbations of [42] is totally unclear to me]. For completeness I remark that the Glashow-Weinberg-Salam transition at $T \sim 10^2$ GeV could not be very spectacular either, because we do not want to wash away the baryon number created earlier.

Anyway, whatever the future theories of high-energy physics will turn out to be, their link with the early phase of the Universe will be most interesting: "Combien d'évenements se pressent dans l'espace d'une seconde, et que de choses dans un coup de dé." [43]

References

[1] S. Weinberg, Gravitation and Cosmology (Wiley, 1972).

[2] T.D. Lee, Particle Physics and Introduction to Field Theory (Harwood Ac. Publ., 1981);
C. Itzykson, J.B. Zuber, Quantum Field Theory (McGraw-Hill, 1980).

[3] G. 't Hooft, in Recent Developments in Gauge Theories, eds. G. 't Hooft et al. (Plenum 1980).

[4] M. Peskin, in Lepton and Photon Interactions at High Energies, ed. W. Pfeil (1981).

[5] experiments on proton decay: D.H. Perkins, In Proceedings XXI Int. Conf. High Energy Physics (Paris, July 1982);
a detected magn. monopole?: B. Cabrera, Phys. Rev. Lett. 48, 1378 (1982), but not confirmed: D.E. Groom, E.C. Loh, H.N. Nelson, D.M. Ritson, Phys. Rev. Lett. 50, (1983), 573.

[6] M.A. Hoskin, J. Hist. Astron. 8, 77 (1977).

[7] J. Audouze, in Astrophysical Cosmology, eds. H.A. Brück et al. (Specola Vaticana, 1982).

[8] E. Weinberg, A. Guth, Nucl. Phys. B. B212, 321 (1983); S.W. Hawking, I.G. Moss, J.M. Steward, Phys. Rev. D26, 2681 (1982).

[9] A.H. Guth, Phys. Rev. D23, 347 (1982).

[10] F.R. Klinkhamer, C.A. Norman, Astrophys. J. 243, L1
E245, L97 (1981); F.R. Klinkhamer, Astron. Astrophys. 107, 235 (1982).

[11] F.R. Klinkhamer, G. Branco, J.P. Derendinger, P. Hut, A. Masiero, Astron. Astrophys. 94, L19 (1982); J.A. Harvey, E.W. Kolb, D.B. Reiss, S. Wolfram, Nucl. Phys. B177, 456 (1981).

[12] R. Anishetty, P. Koehler, L. Mclerran, Phys. Rev. D22, 2793 (1980); eds. M. Jacob, H. Satz, Quark Matter Formation and Heavy Ion Collisions (World Scientific 1982).

[13] G. 't Hooft, Phys. Lett. 109B, 474 (1982).

[14] S. Adler, Rev. Mod. Phys. 54, 729 (1982).

[15] A.D. Linde, Phys. Lett. 108B, 389; 114B, 431 (1982;
A. Albrecht, P.J. Steinhardt, Phys. Rev. Lett. 48, 1220 (1982).

[16] A. Albrecht, P.J. Steinhardt, M.S. Turner, F. Wilczek, Phys. Rev. Lett. 48, 1437 (1982); A.D. Dolgov, A.D. Linde, Phys. Lett. 116B, 329 (1982); L.F. Abbott, E. Farh681 (1982).

[17] A.H. Guth, S.Y. Pi, Phys. Rev. Lett. 49, 1110 (1982); see also S.W. Hawking, Phys. Lett. 115B, 295 (1982) and A.D. Linde, Phys. Lett. 116B, 335 (1982).

[18] R. Brandenberger, R. Kahn, Phys. Lett. 119B, 75 (1982).

[19] A. Vilenkin, Phys. Lett. 115B, 91 (1982); A. Vilenkin, L.H. Ford, Phys. Rev. D26, 1231 (1982).

[20] S.W. Hawking, I.G. Moss, Phys. Lett. 110B, 35 (1982).

[21] V.A. Rubakov, Nucl. Phys. B203, 311 (1982); C.G. Callan, Phys. Rev. D25, 2141, D26, 2058 (1982), Nucl. Phys. B212, 391 (1983).

[22] E.W. Kolb, S.A. Colgate, J.A. Harvey, Phys. Rev. Lett. 49, 1373 (1982); for a review of other implications F.A. Bais, J. Ellis, D.V. Nanopoulos, K.A. Olive, preprint CERN-TH 3383 (1982).

[23] G. Lazarides, Q. Shafi, P.W. Trower, Phys. Rev. Lett. 49, 1756 (1982).

[24] H. Haber, Phys. Rev. D26, 1317 (1982).

[25] M. Srednicki, Nucl. Phys. B202, 327; B206, 132 (1982).

[26] D.V. Nanopoulos, K. Tamvakis, Phys. Lett. 110B, 449; N + T + K. Olive, Phys. Lett. 115B, 15 (1982).

[27] T. Banks, V. Kaplunovsky, Nucl. Phys. B206, 45 (1982).

[28] J. Polonyi, K. Szlachanyi, Phys. Lett. 110B, 395 (1982).

[29] C. Borgs, E. Seiler, Nucl. Phys. B215 (FS7), 136 (1983).

[30] B. Svetitsky, L.G. Yaffe, Nucl .Phys. B210, (FS6) 423 (1982).

[31] J. Kogut et al., Phys. Rev. Lett. 50, 393 (1983).

[32] N. Kimura, A. Ukawa, Nucl. Phys. B205 (FS5), 637 (1982); C.B. Lang, C. Rebbi, Phys. Lett. 115B, 99 (1982).

[33] A. Gonzalez Arroyo, C.P. Korthals Altes, J. Peiro, M. Perrotet, Phys. Lett. 116B, 414 (1982); R.C. Brower, D.A. Kessler, H. Levine, Phys. Rev. D26, 959 (1982).

[34] E. Tomboulis, Phys. Rev. Lett. 50, 885 (1983).

[35] G. Mack, Phys. Lett. 78B, 263 (1978); ref. [8, 20] of our Ch. 8; J. Ambjørn, P. Olesen, Nucl. Phys. B170, (FS1), 265 (1980).

[36] J. Greensite, M.B. Halpern, preprint LBL-14912 (1982).

[37] E. Witten, in ref. [3].

[38] J. Greensite, M.B. Halpern, Nucl. Phys. B211, 343 (1983).

[39] J.C. Taylor, Gauge Theories of Weak Interactions (Cambridge UP, 1978), Chapter 8.

[40] F.A. Bais, S. Rudaz, Nucl. Phys. B170 (FS1), 507 (1980).

[41] E.T. Bell, Men of Mathematics (Simon & Schuster 1937).

[42] M. Crawford, D.H. Schramm, Nature 298, 538 (1982).

[43] H. de Balzac, La peau de chagrin (La Pleiade, Gallimard, 1979).

NOTE ADDED

As of 1 June 1983 the preliminary results of the two experiments mentioned in sect. 1.5 are: 1) from the CERN $p\bar{p}$-collider there are four candidate events for the Z° with approximately the GWS mass; 2) from the Irvine-Michigan-Brookhaven experiment over 130 days the 90% c.l. limit on the proton life time is $\tau_p/BR > 1\ 10^{32}$ yr, where BR is the branching ratio of the $\pi^\circ e^+$ channel ($\sim$ 40%?).

CHAPTER 2

BREAKING OF GAUGE SYMMETRIES IN THE COOLING UNIVERSE

Je m'accommode assez, pour moi, des petits corps;
Mais le vuide à souffrir me semble difficile,
Et je goûte bien mieux la matière subtile.

Molière, Les Femmes Savantes.

1. Introduction

The expanding, homogeneous and isotropic Universe. which is presently filled with low energy photons (T= 3K), is thought to have originated in a rapidly expanding and very hot phase: the Big Bang. The knowledge needed to describe the earliest epoch of the Universe has been augmented significantly by recent advances in our understanding of the interactions of elementary particles. The essential idea is that of gauge theories, to which all four fundamental interactions belong, namely gravity, electromagnetism and the weak and strong forces. The latter three forces have been proven to be renormalizable, i.e. quantum corrections are finite and calculable, whereas a renormalizable quantum theory of gravity remains elusive.

Contrary to the case of the photon the mediating particles of the weak interactions are massive, O(100 GeV), which is precisely the reason why the interaction is so weak. This mass results from the mechanism of spontaneous symmetry breaking (SSB), in which a non-symmetric ground state occurs for a theory with symmetric interactions. More specifically the vacuum state <u>is not</u> invariant under the transformations while the Lagrangian <u>is</u> symmetric.

It is probable that these three observed forces are the low energy remnants of a unified gauge theory, broken at several energy scales. Because quarks and leptons are grouped together in the representations, Grand Unified Theories (GUTs) have baryon number (B) changing reactions which, not being included in the electroweak and colour interactions, are very weak at observable energies and thus predict a large, but finite, decay time for the proton of $\sim 10^{31\pm 2}(M_X/6\ 10^{14}\mathrm{GeV})^4$ years, where M_X is the mass of the relevant boson mediating the B violating force. But in the very early Universe at temperatures larger than or of the order of M_X the B violating forces are quite effective and in fact they might well create a small asymmetry between baryons and anti-baryons, which after the annihilation of the pairs at much lower temperatures O(1 GeV) gives the observed dominance of matter over antimatter.

After some introductionary work we will review the transitions to lower gauge symmetry as the Universe cools down (sections 5 and 6). These socalled phase transitions (PTs) might modify the expansion of the Universe significantly, but we hope, of course, that they do not invalidate the "standard" results of Heliumsynthesis and baryon number creation (sections 2 and 4). Hopes that PTs might generate the required density perturbations for galaxy formation are tempered in §6.2.
Also we will briefly discuss in §6.3 the expected creation of monopoles, i.e. localized,finite energy solutions of the equations of motion with magnetic charge, which might be expected when the spatial distribution of the symmetry breaking "directions" is non-trivial.

Let us remark that the coverage of the references is only indicative. We put $\hbar = K_{\mathrm{Boltzmann}} = c = 1$ and express nearly everything in powers of GeV (roughly the proton mass), but in order to distinguish gravity, which sets the stage only, we keep $M_{Pl} = G^{-1/2} = 1.22\ 10^{19}$ GeV. Indices μ run over 0,1,2,3; x^μ and ∂_μ denote a space-time point and derivative; and all indices occurring twice are summed over.

2. The Big Bang

There are three basic cosmological observations (Weinberg, 1972).

1. On large enough scales ($\gtrsim O$ (100 Mpc)) the Universe is homogeneous and isotropic. Hubble found in 1929 a linear expansion: redshift $Z \equiv (\lambda_{observed} - \lambda_{emitted})/\lambda_{emitted} = H_o \times$ distance, with H_o a constant.
2. Penzias and Wilson discovered in 1965 the isotropic electromagnetic background radiation which has a Planck spectrum and temperature $\sim$ 3K.
3. There is a universal Helium abundance of $\sim$ 25% in mass and a very low primordial abundance of heavier elements later to be enhanced by stellar evolution processes.

The theoretical model to understand these facts is very simple (Weinberg, 1972). From the Einstein equations of classical gravity and the restrictions on the metric tensor from homogeneity and isotropy (Robertson and Walker) we have an expansion equation

$$H^2 \equiv \left(\frac{\dot{R}}{R}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{R^2} , \qquad (1)$$

where R(t) is the scale factor and k = -1,0,1 a curvature parameter. Also we have an energy conservation equation (adiabatic expansion)

$$\frac{d}{dR}(\rho R^3) = -3pR^2. \qquad (2)$$

For relativistic particles we have the energy density $\rho_{rel} = \frac{\pi^2}{30} N T^4$, with N the total effective number of helicity scates (2 for the photon), and from entropy conservation ($s \propto T^3$; volume $\propto R^3$) $T \propto R^{-1}$. Thus $\rho \propto R^{-4}$ and the curvature term in Eq. (1) can be neglected for $R \to o$. Eq. (1) then gives the Big Bang evolution

$$t = 2.4\ 10^{-6}\ N^{-1/2}\ (\frac{T}{GeV})^{-2}\ s. \tag{3}$$

Particles with mass M_Z will be roughly as abundant as the photons γ for $T > M_Z$, but for lower temperatures they will annihilate and reheat the other interacting particles somewhat. Note that they will be able to reach their equilibrium density $n \sim (M_Z T)^{3/2} \exp\left[- M_Z/T\right] \ll n_{rel} \sim T^3$ only if their interaction rate $\Gamma_Z > H$ at $T \sim M_Z$. This point will be of importance in section 4.

We note that because of the finite age of the Universe the maximal distance travelled by light, i.e. the causally connected region, is limited. This socalled particle horizon is given by $d_H(t) = R(t) \int_o^t R(\bar{t})^{-1}\, d\bar{t} = 2t$, as follows from the metric $ds^2 = dt^2 - R(t)^2 \left[d_3 s^2\right]$ and $R(t)$ given by Eq.(3).

Roughly the Big Bang scenario runs as follows. At very early times the Universe is dominated by relativistic particles; precise calculations show that at $t \sim 1$ minute some 25% He is synthesed and practically no heavier elements; still later at $T \sim 4000$ K the protons and electrons recombine and the photons expand freely henceforth, no longer being Thomson scattered ($T \propto R^{-1}$; now $T_o \sim 3K$); at roughly the same epoch the energy density starts to be dominated no longer by radiation but by non-relativistic matter, i.e. hydrogen and Helium nuclei. In this last phase, the largest in time (cf. Eq. 1), we can neglect the pressure and the precise expansion solution depends on two constants only, which we take as the present values of expansion rate and density ratio: $H_o \equiv (\dot{R}/R)_o$ = h 100 km s^{-1} Mpc^{-1} and $\Omega_o \equiv (\rho/\rho_{cr})_o$, with $\rho_{cr} \equiv 3H_o^2/8\pi\rho = h^2\ 2\ 10^{29}$ gcm^{-3}. According to present knowledge h $\sim$ 1/2 to 1 and $\Omega_o \sim 0.1$ to 1. The age of the Universe is $t_o = f\ (\Omega_o)\ H_o^{-1} \sim H_o^{-1} \sim\ h^{-1}$ 20 billion years, but we will be discussing the very earliest relativistic phase only, when $t\ (T = M_X) \sim 10^{-35}$ s.

3. Particle interactions: Unification

3.1 Immediately after the invention of quantum mechanics the relativistic theory for electromagnetic interactions was sought and the Dirac equation dates of 1928 already. But only in the late 40's was the theory proven to be renormalizable: the infinite quantum corrections can be absorbed in a finite number of constants (charge and mass of the electron for example) and the theory (Quantum Electrodynamics, QED) can be formulated using the observed finite masses and charges only and finite quantum corrections. That this can be done is highly non-trivial; infinities at each order in the perturbation expansion might each require different counterterms in the Lagrangian, whereas for QED all counterterms, say $(Z_1 e + Z_2 e^2 + --)\ \bar{\psi}\psi$, are of a limited number of forms, say mass terms $m_o\bar{\psi}\psi$, and thus only a few constants can be redefined, say mass m_{renorm}. QED has a very special form; it is a gauge theory, i.e. invariant under transformations with local parameters (as Λ (x) in Eq. 4). The global invariance (Λ = constant) leads to charge conservation.

What about the other two particle interactions? First let us consider the weak interactions. Fermi gave in 1934 the theory for β decay ($n \to p + e^- + \bar{\nu}$), in a form analogous to QED with a four fermion interaction and dimensional coupling constant $G_F \sim 10^{-5}\ \mathrm{GeV}^{-2}$. Although very succesfull the theory could not be the final one: it is not renormalizable and at high energies ($\gtrsim$ 300 GeV) the cross-sections violate the limits from unitarity (in simple terms: more out than in; cf. Taylor, 1976). In the 50's and 60's two ingredients towards a solution were put forward, but each had their problems. One was the idea of spontaneous symmetry breaking (SSB) which may give masses to the mediating bosons (see example below), but according to the Goldstone theorem massless scalars are predicted to occur which are not observed. The other idea is to extend the U(1) symmetry of QED to a larger group, these are called Yang-Mills (YM) theories. We then have more than one "photon", but they obviously are massless, whereas the weak interactions are weak because the mediating boson W <u>is</u> very heavy: $G_F = g^2/2\,M_W^2$, with g the dimensionless coupling constant (cf. e of QED). YM theories looked promising for renormalization. Higgs showed that surprisingly a spontaneously broken gauge theory only had the good properties: massive gauge bosons but no massless scalars. Later 't Hooft showed that the YM theories after the SSB remained renormalizable.

It will prove useful, for later work, to illustrate all this for the simplest model, called the Abelian Higgs model (cf. O'Raifeartaigh, 1979). The gauge symmetry group U(1) has one parameter $\Lambda(x)$ and works on the gauge field A_μ and the complex scalar field $\phi = \phi_1 + i\,\phi_2$ as follows, in infinitisimal form,

$$A_\mu(x) \to A_\mu(x) + \frac{1}{e}\,\partial_\mu\,\Lambda(x)$$
$$\phi(x) \to [1 - i\Lambda(x)]\;\;\phi(x)\;. \qquad (4)$$

The invariant Langrangian density is

$$L = -\tfrac{1}{4}\,F_{\mu\nu}\,F^{\mu\nu} - \tfrac{1}{2}|D_\mu\,\phi|^2 - V(\phi)\;, \qquad (5)$$

with field strength $F_{\mu\nu} = \partial_\mu\,A_\nu - \partial_\nu A_\mu$, gauge covariant derivative $D_\mu\phi = \partial_\mu\phi + ieA_\mu\phi$, and the scalar potential

$$V(\phi) = -\mu^2\;|\phi|^2 + \lambda|\phi|^4 \qquad (\lambda,\ \mu^2 > 0). \qquad (6)$$

The vacuum will be at the asymmetric minimum of V, which we choose $\sigma \equiv <0\;|\phi|\;0> = <0\;|\phi_1|\;0> = (\mu^2/2\lambda)^{\frac{1}{2}}$.

Develloping in fields around this minimum $\theta = \phi - \sigma$ we rewrite (5) after a gauge rotation to eliminate θ_2:

$$L = -\tfrac{1}{4}\,F_{\mu\nu}\,F^{\mu\nu} - \tfrac{1}{2}\,M_A^2\,A_\mu^2 - \tfrac{1}{2}\;\;|D_\mu\theta_1|^2 -$$
$$2\mu^2\theta_1^2\;\; + L_{interaction}\,(A_\mu,\,\theta_1)\;, \qquad (7)$$

where the gauge boson has become massive $M_A^2 = e^2\;\mu^2/2\lambda$. There is only one massive scalar field θ_1 and $L_{interaction}$ retains the gauge invariance. Realistic models require larger gauge groups and the introduction of fermions, for completeness we give the general form. Lie group G has generators t_r ($r = 1$ --- N, $N = n^2 - 1$ for $G = SU(n)$), hence for $g \in G$: $g = \exp\,[\Lambda_r\;t_r]$. Define a Lie valued potential $A_\mu = A_\mu^r\;t_r$. The finite gauge transformations are

$$A_\mu \to g\;A_\mu g^{-1} + \frac{1}{e}\;g\;\partial_\mu\;g^{-1}$$
$$\phi \to R\;(g)\;\phi\;,$$

where $g = g(x)$, R a representation for the ϕ, and e a coupling constant.

The fields and covariant derivatives are

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + e\,[A_\mu, A_\nu] \tag{8}$$
$$D_\mu \phi = \partial_\mu \phi + e\,R(A_\mu)\,\phi \;.$$

With the Cartan-Killing metric $\{,\}$, note $\{A_\mu, A^\mu\} = \sum_r A_\mu^r A^{\mu r}$, we generalise Lagrangian (5) to

$$-L = \tfrac{1}{4}\{F_{\mu\nu}, F^{\mu\nu}\} + \tfrac{1}{2}\{D_\mu\phi, D^\mu\phi\} + \bar{\psi}\gamma^\mu D_\mu \psi + \bar{\psi} m \psi + \lambda \bar{\psi}\phi\psi + V(\phi). \tag{9}$$

The 3^{rd} and 4^{th} terms of the RHS are the generalised Dirac terms for the fermions ψ in some representation R_ψ (replace R in D_μ definition (8)), while the 5^{th} term is a Yukawa interaction. Some gauge fields A_μ^s acquire masses by SSB if $V(\phi)$ has non-trivial minima, and if H is the remaining symmetry there are dim (H) fields A_μ^t which remain massless.

With these ideas Glashow, Weinberg and Salam constructed the electromagnetic-weak theory (table I; Taylor, 1976). The quarks u, d and the leptons e^-, ν_e are put in left-handed doublets and right-handed singlets (no ν_R), each generation is treated similarly. The ratio $g_1/g_2 \equiv \tan\theta_W$ is determined experimentally : $\sin\theta_w^2 = 0.21$. The SSB gives the weak bosons a mass and by the existence of the Z^o predicted neutral currents ($e^- \nu_e \rightarrow e^- \nu_e$), as confirmed later by CERN experiments. Also the prediction by Glashow, Iliopoulos and Maiani of the charm (c) quantum number (in order to cancel $s\bar{d} \rightarrow W^+ W^- \rightarrow s\bar{d}$) has been verified, still later the beauty (b) quark was found as $\Upsilon = b\bar{b}$ at $\sim$ 10 GeV (truth (t) remains the current experimenters quest). Baryons are composed of 3 quarks and mesons of 2, with the interquark forces described by a YM theory (called Chromodynamics, with "colours"

Table I: Low energy interactions

	Electroweak	Colour
gauge group G	$SU(2) \times U(1)$	$SU(3)$
coupling constants g	g_2 $\quad g_1$	g_3
strengths $\alpha = g^2/4\pi$ as observed at $O(10^2 GeV)$	$\alpha_2 \sim 0.03$ $\alpha_1 \sim 0.01$	$\alpha_3 \sim 0.2$
fermion representations	$\begin{bmatrix} u \\ d \end{bmatrix}_L$ $\quad \begin{bmatrix} \nu_e \\ e^- \end{bmatrix}_L$ $u_R, d_R, e_R.$ *	$\begin{bmatrix} u_r & u_y & u_b \\ d_r & d_y & d_b \\ c_r & c_y & c_b \\ & \vdots & \end{bmatrix}$
gauge fields	W^+, W^-, Z^o mass $\sim g\sigma$ γ photon	8 gluons
Higgs scalars	$\begin{bmatrix} \phi^+ \\ \phi_o \end{bmatrix}$, with $\langle 0 \vert \phi \vert 0 \rangle = \sigma \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ $\sigma \sim 300$ GeV	none

*similarly for the other two generations: read for (u, d, e, ν_e) respectively (c, s, μ, ν_μ) or (t, b, τ, ν_τ), where we neglect Cabbibo mixing of the down quarks.

L, R denotes left- and right-handed components.

Table II: Unified Interactions (Georgi and Glashow, 1974)

G_U : SU(5)

$\alpha_U(10^{15}\,\mathrm{GeV})$: 1/40

fermions*:

$$\underline{\underline{\bar{5}}} = \begin{bmatrix} \bar{d}_r \\ \bar{d}_y \\ \bar{d}_b \\ e^- \\ \nu_e \end{bmatrix}_L \qquad \underline{\underline{10}} = \begin{bmatrix} 0 & \bar{u}_b & -\bar{u}_y & -u_r & -d_r \\ & 0 & \bar{u}_r & -u_y & -d_y \\ & & 0 & -u_b & -d_b \\ -\text{ etc} & & & 0 & -e^+ \\ & & & & 0 \end{bmatrix}_L \text{ antisymm.}$$

scalars : ϕ: $\underline{\underline{24}}$, with $\langle 0|\phi|0\rangle \sim 10^{15}$ GeV diag $(1,1,1,-3/2,-3/2)$

H: $\underline{\underline{5}}$, with $\langle 0|H|0\rangle \sim 10^{2}$ GeV $(0,0,0,0,1)^T$

*other generations similarly

red, yellow, blue as charges, Table I). For high interaction energies q^2 the colour forces get weaker (cf. Eq. 10), probably this explains the famous lepton scattering results on protons, which apparently scattered on freely moving point particles (quarks). For larger distances g_3 gets larger and this is related to the confinement of quarks, i.e. no separate quarks exist, but rigorous calculations are still lacking. Also the hadron production relative to $\mu^-\mu^+$ in $e^- - e^+$ collisions is beautifully explained by 5 quark flavours, each in 3 colour variants.

3.2

The three elementary particle interactions can thus be understood as the gauge theories summarized in Table I. But, as always, deeper questions remain unsolved:

1) there are still (too) many arbitrary parameters, e.g. 7 for the electroweak theory.
2) why is electric charge quantised and $|Q_{electron}| = |Q_{proton}|$?
3) why do we have three so different forces?
4) what is the relation between quarks and leptons, which on the one hand are so different (quarks have strong interactions, leptons not; quarks heavier than the leptons, per generation at least), but on the other hand have some inter relations (notably the cancelling of the triangle anomalies)?
5) why is there a triplication of generations?
6) are the Higgs scalars elementary particles or composites, and what regulates the SSB precisely?
7) what is the quantum theory of gravity?

The last three problems are still unsolved (perhaps 5 has something to do with the connection GUT-supergravity, see below), but the first four might be tackled in the unification scheme (reviews: Langacker, 1980; Nanopoulos, 1980; Ellis, 1980). The idea is to have one simple gauge group G_U, hence one coupling constant g_u, and quarks and leptons together in G_U representations. At a high energy scale M_U the symmetry breaks spontaneously to the "observed" SU(3) × SU(2) × U(1). But what about one coupling strength, whereas we observe large differences (Table I)? The answer lies in the renormalization group equations (Georgi et al., 1974), which give the effective coupling as a function of the interaction energies (q^2), generally for G = SU(n)

$$\alpha_n(q^2) = \frac{12\pi}{(11n - 2F)\ \ln\ (q^2/\Lambda^2)} \qquad (10)$$

with F the number of quark flavours (6?) and Λ an energy scale. With the observed α_3 and α_1 (with n = 1 in (10)) at $q^2 = (100\ \text{GeV})^2$ we calculate equality at $M_U \sim 10^{15}$ GeV with $\alpha_U \sim 1/40$. But now we may calculate g_2(100 GeV) and find: $\sin^2\theta_W \sim 0.20$! Thus the three observed forces are low energy remnants of a unified interaction broken at energies $\sim M_U$. Also questions 2) and 4) may be dealt with (e.g. Nanopoulos, 1980). The smallest possible (rank 4) G_U is SU(5) (Georgi and Glashow, 1974). There are stringent restrictions on G which basically allow only 2 larger G, which will have more fermions (cf. Barbieri, 1980). Table II gives the fermion representations and the required Higgs scalars for the SU(5) unification, where again each generation is treated similarly.

The breaking by <0 |H| 0> in the Yukawa terms gives the down quark and charged lepton equal masses. Thus for the third generation $m_b \sim m_\tau$, which gets renormalized for the low observable energies (using F = 6: a strong indication on the number of generations) to $m_b \sim 3m_\tau$, which fits the experimental data (1.5 and $\sim$ 4.5 GeV). For the first generation the observed ratio is quite different from 3, see discussion in Ellis (1980).

But there is also something completely new in the unification hat. By the breaking through the ϕ scalars 24-(8+3+1)=12 bosons X get a high mass (the ones of the SU(3) × SU(2) × U(1) remaining massless). Some of these bosons mediate reactions between quarks and leptons ($uu \xrightarrow{X} e^{+}\bar{d}$) which at our low energies leads to a very weak proton decay ($p = uud \rightarrow e^{+} + (\pi^{o} = d\bar{d})$)with a large lifetime due to the X propagator, but perhaps measurable. Another significant effect of the baryon number changing reactions might lie in the early Universe at temperatures $T \sim M_X$.

4. Creation of the Matter-Antimatter Asymmetry

All observations indicate the total absence of antimatter in the Universe (Steigman, 1976). The baryon number density n_B thus is non-zero and observations give the numerical value

$$\frac{n_B}{n_\gamma} = \frac{n_{baryon} - n_{antibaryon}}{n_\gamma} \sim \frac{n_{baryon}}{n_\gamma} \sim 10^{-10} \left(\frac{\Omega_{bo}}{0.01}\right), \qquad (11)$$

with Ω_{bo} the present baryon density relative to the critical density (section 2;c.f. Olive et al. 1981). If baryon number would be totally conserved one might think the small number of Eq. (11) just to be given by initial conditions. But it would be more satisfactory if we could find a physical mechanism to generate this small baryon asymmetry.

The basic idea is quite simple: the density ratios of baryons:antibaryons: photons at very high temperatures is something like $10^{10}:10^{10}:10^{10}$, later the unknown mechanism produces a small asymmetry $10^{10}+1:10^{10}-1:10^{10}$, which after annihilation[1] gives the final ratios 2:0: $\sim 10^{10}$, as observed. The key problem, of course, is the second step. Four ingredients are required to generate a net baryon number density (e.g. Weinberg, 1979):

1. Baryon number (B) non-conservation,
2. Charge conjugation (C) asymmetry,
3. C and parity (CP) asymmetry,
4. non-thermal equilibrium distribution functions of some of the species involved.

Conditions 2 and 3 are to avoid a canceling of production by particles and antiparticles (operation of C changes the sign of the baryon number, because B (particle) = -B (antiparticle); P leaves B unchanged). Also condition 4 is obvious; if everything first stays in equilibrium no asymmetry can arise. As we have seen in the previous chapter unified interactions will give B violating reactions as well as C and CP asymmetries. Condition 4 is implemented by the fast expansion of the Universe as given by Eq. (1) (see also Appendix A of Kolb and Wolfram, 1980a). In order to illustrate condition 4 Weinberg (1979) has proposed the delayed decay scenario, which runs as follows. Assume complete equilibrium at T_{Pl} ($\gg M_X$); we have for the decay rate of the heavy X boson (including a time dilatation factor)

$$\Gamma_X \sim \alpha_X M_X^2 N \left[T^2 + M_X^2\right]^{-1/2} \tag{12}$$

and the expansion rate of the Universe (Eq. (1))

$$H \sim N^{1/2} T^2/M_{Pl} . \tag{13}$$

The X and their antiparticles $\overline{X}$ will decay around a temperature T_d when $\Gamma_X(T_d) \sim H(T_d)$. If the X and $\overline{X}$ decay at temperature

$$T_d < M_X \tag{14}$$

they will not be replen ished by inverse decays (Boltzmann factor), and the net baryon number to entropy ratio produced in the delayed decay is simply proportional to their equilibrium density

$$\frac{n_B}{s} \sim \frac{45}{4\pi^4} \zeta(3) \left(\frac{N_X}{N}\right) \Delta B \quad , \tag{15}$$

with s the specific entropy and ΔB the average net baryon number from the decay of a X-$\overline{X}$ pair. Before discussing the estimates of ΔB we must see whether or not condition (14) applies. Using Eqs. (12)(13) we easily see that (14) implies

$$M_X > N^{1/2} \alpha_X M_{Pl} \ . \tag{16}$$

For gauge bosons ($\alpha \sim 1/40$) with a calculated mass of $M_X \sim 6\ 10^{14}$ GeV (Ellis et al., 1980a) we see that the delayed decay condition (14) does not hold. However Higgs bosons are believed to have very small Yukawa couplings (as in the Weinberg-Salam model, cf. Taylor, 1976) and the condition (14) may apply.

In order to find ΔB one calculates the different branching ratios of the X and $\overline{X}$ (cf. simple model below; Nanopoulos and Weinberg, 1979). For SU(5) with two $\underline{5}$'s of Higgs and some honest estimates Yildiz and Cox (1980) calculated (15) to be of order of 10^{-10} (for further discussion of ΔB and the required CP breaking e.g. Ellis, 1980). Recently a direct connection was noticed (Ellis et al., 1981a) between the ΔB interference graphs and those leading to a finite renormalization ($\delta\theta_{GUT}$) of the θ parameter of the QCD (SU(3)) vacuum. [This θ defines the ground state when topologically distinct Yang-Mills vacua exist, analogous to the Bloch functions for periodic potentials; gauge transformations in topological equivalence class n implemented on the vacuum $|\theta\rangle$ give $G_n\ |\theta\rangle = e^{in\theta}\ |\theta\rangle$ (for an introduction e.g. Crewter, 1978; Jackiw, 1980).]

Assuming $\theta = 0$ at a high energy, say M_{Planck}, the present observational limits on the electric dipole moment of the neutron d_n, which gets a dominant contribution of the CP violating term with parameter θ (1 GeV) $\gtrsim \delta\theta_{GUT}$, allow for practically no entropy generation after the n_B generation at unification energies (Ellis et al., 1981a,b). The reasoning goes as follows: $d_n \gtrsim 4.10^{-16}\ \delta\theta_{GUT}$ e-cm would violate the experimental upper limit (2.10^{-24} e-cm) if the generated $\{n_B/s\}_{GUT}$ has to be significantly larger than 10^{-10} to allow for later entropy generation. We remark that in comparing $\delta\theta_{GUT}$ and ΔB graphs moduli of typical unitary matrix elements U_{dm} connecting the different contributing Higgses, namely in decay and mass terms, were naturally assumed to be O(1). If these are substantially smaller they could alleviate the nearly conflicting theoretical ($\delta\theta >$ ---$|U_{dm}|^2\ n_B/s$) and observational ($d_n <$ ---) limits (see Eq. (24) of Ellis et al., 1981a).*

If condition (14) is not strongly satisfied the final result will be less than estimated in Eq. (15) and the rate equations of the B changing reactions have to be solved numerically.

Kolb and Wolfram (1980a) have introduced a simple model which incorporates the major physical ingredients. More detailed models are not quite relevant for the moment because of the lack of knowledge on the precise unification Lagrangian and the details of CP violation.

The model consists of two types of particles: nearly massless particles b and $\bar{b}$ carrying baryon numbers B = ½ and B = −½ respectively, and massive bosons X and $\bar{X}$ mediating baryon-number violating reactions. The decay amplitudes M of these massive bosons are parametrized as

$$|M(X \to bb)|^2 = (1 + \eta)\tfrac{1}{2}\ |M_o|^2,$$
$$|M(X \to \bar{b}\bar{b})|^2 = (1 - \eta)\tfrac{1}{2}\ |M_o|^2,$$
$$|M(\bar{X} \to \bar{b}\bar{b})|^2 = (1 + \bar{\eta})\tfrac{1}{2}\ |M_o|^2,$$
$$|M(\bar{X} \to bb)|^2 = (1 - \bar{\eta})\tfrac{1}{2}\ |M_o|^2,$$

with $|M_o|^2$ of the order of a small coupling constant α. Because of unitarity

*Note that the above arguments do not hold if a global U(1) axial symmetry forces θ to be zero (Peccei and Quinn, 1977; Dine et al. 1981).

and CPT invariance only two free parameters η, $\bar{\eta}$ are left, where $\varepsilon = \bar{\eta}-\eta = O(\alpha)$ measures the amount of CP breaking. Thus a state initially containing an equal number of X and $\bar{X}$ ($n_X^o = n_{\bar{X}}^o$) will decay, in the absence of back reactions, to a system with a net baryon number $n_B = (\eta - \bar{\eta})\frac{1}{2}/(n_X^o + n_{\bar{X}}^o)$. For simplicity all particles are given only one spin degree of freedom, and obey Maxwell-Boltzmann distributions. Because in the expanding Universe all densities drop quickly, a convenient type of variable is

$$Y_A \equiv n_A/n_Y,$$

the relative number density of particle A(= b, $\bar{b}$, X or $\bar{X}$) with respect to photons. The rate equations can be written as (Hut and Klinkhamer, 1981a)

$$\frac{dY_\Delta}{dx} = \tfrac{1}{2}\, x^2\, K_1(x) - \frac{\alpha}{4x_p}\left\{x \frac{K_1(x)}{K_2(x)}\, Y_\Delta(x) + \tfrac{1}{4}\,\varepsilon\, x^3\, K_1(x)\, Y_B(x)\right\} \tag{17a}$$

$$\frac{dY_B}{dx} = \frac{\alpha}{4x_p}\left\{\varepsilon\, x \frac{K_1(x)}{K_2(x)}\, Y_\Delta(x) - x^3\, K_1(x)\, Y_B(x) - \frac{288\alpha}{\pi}\, x^{-4}\, Y_B(x)\right\}, \tag{17b}$$

with "time" parameter $x = M_X/T$, an effective expansion parameter $x_p = M_X/(7.5\ 10^{18}\ N^{-1/2}\ \mathrm{GeV})$, $K_{1,2}$ modified Bessel functions, $Y_B = Y_b - Y_{\bar{b}}$, and the deviation from equilibrium parametrised as $Y_\Delta = Y_+ - Y_+^{eq}$, $Y_+ = \frac{1}{2}(Y_X + Y_{\bar{X}})$. In deriving Eq. (17) we put $\eta + \bar{\eta} = 0$, the final Y_B being quite insensitive to the exact value as long as $|\eta + \bar{\eta}| < 1$.*

*see fig. 5 in Chapter 3 of this thesis.

In equation (17a) it is clear how the first RHS term determines the production of Y_Δ, independent of already existing Y_Δ and Y_B, as a function of temperature only ($T = M_X/x$). The second term destroys the deviation from equilibrium of the X's, and is therefore proportional to α/x_p or $\alpha G^{-\frac{1}{2}}$: the magnitude of the deviation is governed by a competition between particle reaction rates and Universe expansion. The third term is typically ten orders of magnitude smaller than the second one in our calculations ($\varepsilon = 10^{-6}$; $Y_B \lesssim \varepsilon$). Therefore the rate equation for Y_Δ is nearly completely Y_B-independent, and $Y_\Delta(x;\, x_o)$ is fixed by specifying the initial condition $Y_\Delta(x_o;\, x_o)$ independent of Y_B.

The rate equation for Y_B shows a production term $\propto Y_\Delta$, with the same reaction vs. expansion factor α/x_p, but also the small CP violation parameter ε. The next two terms $\propto Y_B$ determine the damping of Y_B, by means of all baryon number changing processes, independent of ε, which already indicates that $Y_{B\,max} \leq \varepsilon$. The first of these two terms describes inverse decays of X, $\overline{X}$ ($bb \rightarrow X$, etc.), and drops off quickly for high x ($K_1(x) \propto x^{-\frac{1}{2}} e^{-x}$ for $x \gg 1$). The last term is the Fermi approximation for $2 \rightarrow 2$ scattering processes ($bb \rightarrow \bar{b}\bar{b}$, etc.), which is the dominating, but still rather unimportant, term for $x \gtrsim 20$. For high energies this term is unimportant, since the large contribution to $2 \rightarrow 2$ scattering by the exchange of on-shell intermediate X's is already included in the previous term (see Kolb and Wolfram, 1980a, sect. 2.3.2). Therefore we just replaced x^{-4} by 1 for $x < 1$, since a detailed treatment of the $2 \rightarrow 2$ processes would be unnecessary in this regime.

Let us choose parameters $\alpha = 1/40$, $M_X = 10^{15}$ GeV, $N = 100$ and $\varepsilon = 10^{-6}$ (note that condition (16) is strongly violated). Starting from thermal equilibrium at $x_o = 0$ (or $x_o = M_X/M_{Pl} \sim 10^{-4}$) we find $Y_B(x)$ peaking around $x \sim 1$ and then, mostly because of inverse decays, falling a factor 4 to the final value $Y_B(\infty, x_o \doteq 0) = 1.43\ 10^{-8}$ (this differs from Kolb and Wolfram (1980a) who find a somewhat larger drop, this difference could only be due to our treatment of the 2-2 scattering, but we overestimated their contribution somewhat at $x \sim 1$). For other parameters Kolb and Wolfram (1980a) displayed similar damping of the Eq. (15) estimate in their figs. 3 and 4.

Finally we remark that because of the B violating processes operative at $T = O(M_X)$ any (not too large) primordial baryon number will be destroyed (second term RHS of Eq. (17b)), after which the final Y_B will be generated. The presently observed matter dominance thus sets via the required ΔB important constraints on the unification theory and the CP violation.

5. Phase Transitions

Up till now we have not included finite temperature effects in the particle theory. We are warned however by such drastic effects as super conductivity (another example of SSB, cf. section 3) vanishing for high enough temperatures: symmetry is restored by the thermal fluctuations. In the Abelian Higgs model we used the minimum σ of the classical (tree) potential: $\left[\frac{dV^o}{d\phi}\right]_\sigma = 0$. The vacuum state is determined by the effective potential (cf. Coleman and Weinberg, 1973), where quantum corrections are included (e.g. in the loop, i.e. $\hbar$, expansion). If one still requires a non-trivial minimum of V_{eff} one finds that for 1-loop in A_μ the quartic coupling must not be too small $\lambda > 3e^2/32\pi^2$ (cf. Linde, 1979). For finite temperature the effective potential is ideally suited: usual field theory applies, but with the time components of the momenta discrete and the relevant integralis replaced by sums (de Fetter and Walecka, 1971). The expectation value of operator O is $\langle O\rangle = \mathrm{Tr}\left[O \exp(-H/T)\right]/\mathrm{Tr}\left[\exp(-H/T)\right]$, with H the Hamiltonian and T the equilibrium temperature. For the $T = 0$ potential $V^o = -\frac{1}{2}\mu^2\phi^2 + \frac{\lambda}{4}\phi^4$ one finds the equation for the minimum σ modified: $\sigma(\lambda\sigma^2 - \mu^2 + \frac{\lambda}{4}T^2) = 0$. The minimum $\sigma(T)$ shifts continuously from $\sigma(T = 0) = (\mu^2/\lambda)^{1/2}$ to $\sigma(T_c) = 0$ at a critical temperature $T_c = 2\mu/\lambda^{1/2}$; this is called a <u>second order phase transition</u> (2PT). For $T > T_c$ the theory is symmetric. Strictly speaking quantum corrections will always introduce a weak discontinuity in $\sigma(T)$. For the Abelian Higgs model one finds analogously: $\sigma(\lambda\sigma^2 - \mu^2 + \frac{1}{12}(4\lambda + 3e^2)T^2) = 0$. And for $\lambda \leq e^4$ we see that although there is a non-trivial minimum σ' at $T_c \sim (15\lambda/2\pi^2)^{1/4}\mu$ with $V(\sigma = 0) = V(\sigma')$, there is a barrier in between. The transition from $\sigma = 0$ to $\sigma' \neq 0$ will be discontinuous and occurs at a temperature different from T_c: <u>a first order phase transition</u> (1PT). For completeness we remark that for a temperature range $|T - T_c| \lesssim e^2 T_c$ perturbation theory breaks down, the $m_{Higgs} \to 0$ and infra-red divergencies occur (Weinberg, 1974). For more details we refer to the reviews of Kirznits and Linde (1976) and Linde (1979). Also for more complicated models (GUTs) it holds true the 1PTs occur if the Higgs particles H are light compared to the gauge bosons G (cf. Eq. 7). Alas our understanding of the Higgs sector in GUTs is little. Generally one might say that if the scalars H are fermion composites $H = F\overline{F}$ one expects $M_H \sim M_F \sim M_G$; but if the breaking is through radiative corrections (Coleman and Weinberg, 1973) we have typically $M_H \sim \alpha^{1/2} M_G$. Perhaps there are two indications that the breaking is radiative (at least at the unification scale[2]). The unexplained hierarchy of energy scales $M_{WS} \sim 10^2$ GeV $<<< M_U \sim 10^{15}$ GeV $<< M_{Pl} \sim 10^{19}$ GeV could originate as follows

(Ellis et al. 1980b; Weinberg, 1979b): if the quartic couplings λ of the 2 Higgs systems (with bare masses zero) needed for the breakings are of the same strength $\sim g^2$ at M_{Pl} , the λ of the $\underline{\underline{24}}$ will diminish much more rapidly at lower energies than those of the $\underline{\underline{5}}$, and the first will give a breaking at $\sim 10^{-4}$ M_{Pl} when $\lambda \sim 0$, while the second will be ~ 0 at very much lower energies. Secondly Ellis et al.(1980c) have looked for the GUT which can be incorporated in the extended supergravity theory with composite states, and find only G = SU(5), fermion representations $3(\underline{\underline{10}} + \bar{\underline{\underline{5}}})$ (exactly the 3 generations as observed) and probably massless $\underline{\underline{24}}$ and $\underline{\underline{5}}$ scalars. Alas the argument involves a lot of speculations, but some have a field theoretical backing in 2 dimensional models.

6. Phase transitions from unification in the Universe

6.1

After all the preparations of the previous chapters we now get to business.[3] At very high temperatures ($T > T_c = 0(10^{15}$ GeV)) the vacuum is in the symmetric state and the gauge bosons and fermions (the observed ones get their masses only at the 10^2 GeV breaking) are massless and forces long range. If the vacuum gravitates normally (cf. weak equivalence principle, Weinberg 1972) the symmetric minimum at $T > T_c$ has a large energy density $\rho_v = 0(\sigma^4)$, because the present cosmological constant is zero or small $\lesssim(10^{-2}$ eV$)^4$ (cf. Kolb and Wolfram 1980b). This just amounts to setting the zero energy level of the potential: $V(\sigma, T=0) = 0$. For $T > T_c$ this $\rho_v \sim \sigma^4 \sim T_c^{\ 4}$ will be negligible compared to the particles $\rho \sim NT^4$. For a 2PT the vacuum shifts at T_c to the broken state so that never in the history of the Universe was ρ_v dominant. But for a 1PT the transition is blocked for $T \leq T_c$ and the constant energy density ρ_v will soon dominate over that of the particles which is redshifted by the expansion. From Eq. (1) and $\rho_v \sim \sigma^4$ we then have a very rapid expansion approaching the de-Sitter solution

$$R(t) \sim R(t_c) \exp \left(\frac{t - t_c}{\tau}\right) \tag{18}$$

$$\tau \sim (8\pi/3)^{-\frac{1}{2}} M_{Pl}/\sigma^2$$

which stretches the particle horizon to

$$d_H(t) \sim \tau \exp \left(\frac{t - t_c}{\tau}\right) \quad . \tag{19}$$

The particle temperature T still goes as R^{-1}, so there results a large supercooling if the symmetric vacuum is blocked from the transition (hence called the false vacuum) for times quite larger than $t_c \sim N^{-\frac{1}{2}} T_{Pl} T_c^{-2}$ (Eq. (13)).

Because we presently are in the broken state we know that the transition must have occurred. As in the usual first order phase transitions the transition goes through nucleation. The bubbles of true vacuum will have a minimum size for which the gained energy $\propto \rho_v r^3_{min}$ compensates the surface energy $\propto - r^2_{min}$. There are two ways the bubbles may originate: by thermal excitations, or by barrier penetration. The nucleation rates are, $\Gamma_T \propto \exp\left[-E_3/T\right]$ and $\Gamma_Q \propto \exp\left[-S_4\right]$, where E_3 is the free energy of the critical solution and S_4 the action of the solution in 4-dimensional Euclidean space. Γ_T peaks at $T \sim \frac{1}{2} T_c$ (Guth and Weinberg, 1980), whereas Γ_Q is constant but small, barrier penetration being a weak effect, of course (Coleman, 1977). There are basically three alternatives to end the supercooling at T_{end}:

1. By thermally nucleated bubbles, but then T_{end} is not far below T_c, otherwise the few bubbles cannot catch up the rest of the exponentially expanding Universe (Sato, 1981).
2. By the tunneled bubbles filling the Universe.
3. The false vacuum becoming unstable.

If the barrier is large the thermally nucleated bubbles will be too sparse, so we expect strong supercooling either terminated by 2 or 3. But alternative 2 is full of problems, to name a few (for monopoles see § 6.3):

a) Strictly speaking the transition never ends, Γ_Q is constant but a bubble will cover only a finite region in co-moving coordinates (cf. Eqs. 18, 19), hence there always remains some false vacuum. Guth (1981) also noted the bubbles true vacuum probably do not percolate.

b) Neglecting point a) the problem remains to thermalise (how?) the released heat, which is put in the kinetic energy of the bubble walls, of the largest bubbles, when there is only 1 second before Helium synthesis starts.

c) If the bubbles give rise to strong inhomogeneities it is not clear whether or not we reheat to T_R of Eq. (20) below, which is important for the baryon number creation (§ 6.2).

It thus appears that an instability of the false vacuum at T_1, is the cleanest way to end the supercooling period. At T_1, the vacuum at every space point shifts to the broken state. The "latent heat" $\rho_V \sim \sigma^4$ reheats the Universe to

$$T_R = \sigma\left[\frac{30}{N_R\pi^2} + \frac{N_1}{N_R}\left(\frac{T_1}{\sigma}\right)^4\right]^{1/4} \sim 0.4\sigma \quad , \tag{20}$$

as follows from energy conservation with N_1 and N_R the number of states before and after reheating (cf. Einhorn and Sato, 1981).

After these generalities we must be more specific on what <u>does</u> occur at $T \lesssim 10^{15}$ GeV. For G = SU(5) Daniel and Vayonakis (1981) and

Guth and Weinberg (1981) calculated the phase diagrams. What transition occurs depends on the coupling constants in the Lagrangian, but a not too strong 1PT is typical. For certain parameters the false vacuum destabilizes (region d of Guth and Weinberg, 1981; for Abelian Higgs $3e^4/16\pi^2 < \lambda < e^4$). If the breaking is of the Coleman-Weinberg type (CW; see section 5) Daniel (1981) and Billoire and Tamvakis (1981) found a very large supercooling $T_{end} = O(1\ \mathrm{GeV})$ when at last the nucleation rate Γ_Q equals the expansion H. This potential for $\phi << T << \sigma$ and with the adjoint scalars Φ on the critical orbit $\Phi = \phi\, U^+ \mathrm{diag}\,(1,1,1,{}^{-3}/2,{}^{-3}/2)U$ (arbitrary U) is to 1-loop (cf. Abbott,1981)

$$
\begin{aligned}
V^1_{eff} &= \frac{5}{8} g^2 T^2 \phi^2 + B\phi^4 \left(\ln \frac{\phi^2}{\sigma^2} - \frac{1}{2}\right) + \frac{B}{2}\sigma^4 \\
B &= 8\ 10^{-4} \qquad\qquad (21)\\
\sigma &\sim 10^{15}\ \mathrm{GeV}.
\end{aligned}
$$

This potential has a barrier whose width shrinks for lower temperatures and thus there always will be a T_{end} when the nucleation rate $\Gamma_Q(T_{end})$ is large enough. This need not be the case for certain barriers originating from quantum corrections, which are more or less constant as is the expansion rate of the false Universe $H \sim \frac{1}{T}$, (Eq. 18). Recently Sher (1981) noted the importance of the running SU(5) coupling constant g(T), cf. Eq. (10), because we are dealing with exponentials. A more detailed investigation (Tamvakis and Vayonakis, 1981) shows a destabilization by a non-perturbative $<F^2_{\mu\nu}>$ term, calculated for instantons, which will become important at $T = O\ (\Lambda_{SU(5)})$, with $\Lambda_{SU(5)} \sim 2\ 10^6$ GeV, cf.Eq. (10).

Another reason for the premature ending of the supercooling might lie in the small gravitational effects (cf. the chiral effects noted by Witten (1981a) in the supercooling at the Weinberg-Salam scale, inducing a transition at O(100 MeV); but see our footnote 2). Abbott (1981; see also Fujii, 1981) considered some small coupling terms $\pm R\phi^2$ added to the potential of Eq. (21). For curvature determined by ρ_v, $R = -32\pi G(\frac{B}{2}\sigma^4)$ we have two completely different histories: -) the false vacuum destabilizes at $O(10^{10}$ GeV), or +) the barrier always is too large (cf. remark above) and the transition is never completed. Hut and Klinkhamer (1981b) noted another possible gravitational effect: for $T \leq O(10^{11}$ GeV) the wavelengths associated with the barrier are larger than the event horizon $D_H = (3/8\pi)^{1/2} M_{Pl}\, \rho_v^{-1/2}$, probably invalidating the usual flat space-time calculations and global gravitational effects in analogy with the Hawking radiation might induce a shift to the vacuum.

These three different arguments thus indicate that the supercooling for a CW SU(5) potential is ended at much higher temperatures and more smoothly than first thought.

6.2*

In this paragraph we will see whether or not the simple ideas on baryon number creation of section 4 survive the complications from phase transitions (Hut and Klinkhamer, 1980a). For this purpose we will use the simple model of Kolb and Wolfram (1980a). If the vacuum is in the disordered state ($\langle\phi\rangle = 0$) the theory is symmetric and massless, so clearly no net B can be generated, the out-of-equilibrium driving mechanism is lacking. For a second order PT this occurs at $T \gtrsim T_c$. At $T \sim T_c$ the theory is broken and the baryon number producing processes start their build up. We thus solve the equations (17) starting from thermal equilibrium at $x_o = M_X/T_c \sim g\sigma/T_c = O(1)$ and find, not surprisingly, the same n_B/s as the standard calculations with $x_o = 0$.

*see fig. 2-4 in Chapter 3 of this thesis.

Larger differences might be expected for first order PTs. The baryon number will be generated after reheating, which must be quite smooth, or in other words the thermalisation must be effective, if we want to preserve the homogeneity of the Universe. This will be the case if the vacuum becomes unstable at T_1 (10^6 - 10^{10} GeV?), because when the barrier vanishes the bubbles from the last flash of nucleation probably will have sizes and thermalisation times $O(\sigma^{-1})$. But before, in the supercooling period,we have an important bonus: the washing out of any truly primordial n_B/s will be much stronger than in the standard scenario (cf. Kolb and Wolfram, 1980a). For example the B changing collisions will have $\sigma_{2-2} \sim \alpha_U^2/T^2$ (Ellis et al., 1980d) so that the damping factor is

$$\exp\left[- M_X/x_p \int_{T_{Pl}}^{T_1} \langle v\sigma_{2-2}\rangle \, dT \right] \sim \exp\left[- \alpha_U^2 \, N^{-1/2} \, T_{Pl}/T_1 \right] , \tag{22}$$

which is $\lesssim 10^{-40}$ for $T_1 \lesssim 10^{-2}\, T_c$. This exponential damping follows from the rate equation $dY_B/dt \sim -2Y_B n\gamma \langle v\sigma_{2-2}\rangle$ (cf. the third term on the RHS of Eq. 17b, where we had inserted the Fermi approximation valid for $T << M_X$; see § 4 of Kolb and Wolfram, 1980a).

With the possibly huge damping we must now consider the n_B/s generated after smooth reheating. Again this is simply starting from equilibrium ($Y_\Delta(x_o) = Y_B(x_o) = 0$) at $x_o \sim M_X/T_R \sim g\sigma/0.4\sigma \sim O(1)$, which gives the $[n_B/s]_{final}$ equal to the standard value (this holds generally for $x_o \leq 5$). Also the numerical solutions show that the n_B/s destroying processes are still effective enough at $x \geq x_o$ to wash out a possible baryon number density ($Y_B(x_o) \neq 0$) from the thermalisation of the bubbles themselves. But this contribution to $[n_B/s]_{final}$ came from the gauge bosons, whereas lighter Higgs bosons are present. These will be the most important for the n_B/s created, since 1) their CP violating diagrams are of lower order than for gauge bosons, typically (remember from Eq. (17b) $Y_{B\ final}$ roughly $\propto$ ε), and 2) the dilution will be less (cf. Kolb and Wolfram, 1980a, fig. 4).

We thus conclude that the generated baryon number density after a second order PT or after the smooth reheating in a first order PT probably will be the same as in the usual calculations neglecting PTs. Finally some remarks (Hut and Klinkhamer, 1980a, Appendix B) on the role of IPTs for galaxy formation, where the major question is the origin of the small density perturbations, which grow under self-gravity into the observed bound systems, galaxies up to clusters. Strong IPTs provide two interesting ingredients: 1) after reheating the particle horizon has been stretched by a factor $\sim N^{1/2}\ T_c/T_1$ relative to the standard horizon $\sim 2\ t_c$, as follows easily from Eqs. (18) and (19), and 2) nucleation probably will lead to density perturbations. But point 2) includes all unsolved problems mentioned earlier and whereas the wanted density perturbations on galaxy scales perhaps could be made, this arises not at all naturally nor is it clear how to avoid unwanted aspects of the perturbation spectrum, such as unobserved strong metric perturbations (primordial black holes). Probably the density perturbations already exist before the baryon number creation epoch. This would then result in density perturbations of the adiabatic type ($\delta\rho_b/\rho_b = \delta\rho_\gamma/\rho_\gamma$, or n_B/s constant): in each region of size d_H in the huge galaxy-sized density enhancement (of very small amplitude) the same n_B/s is made at $T \sim M_X$, depending on physical parameters only (ε, M_X, α etc.), albeit at different eigen-times. The only reasonable way to have isothermal perturbations

($\delta\rho_\gamma = 0$) later, appears to be from large-scale shear at $T \sim M_X$ (Bond et al., 1981), but here we are already deviating from the charming simplicity of the standard model (section 2).

6.3

Another aspect of the breaking of gauge symmetries in the early Universe is the possible creation of monopoles, which depends on the "directions" of the breaking. From topological arguments monopoles are expected to occur if the breaking has a stage with $G \to H_j$ where $\pi_2(G/H_j) \neq \{0\}$, which always is the case if G is simply connected and H_j contains a U(1) factor, because then $\pi_2(G/H_j) = \pi_1(H_j)$ which is certainly non-trivial (remember: $\pi_n(H)$ is the n^{th} homotopy group with as elements the different equivalence classes of the mapping n-sphere S_n in H). Hence in the unification scheme monopoles are to be expected. The actual 't Hooft-Polyakov solution for G = SU(2) and a triplet Higgs ϕ^a (reviews: Actor, 1979, Prasad, 1980) illustrates this. We are looking for <u>classical</u> and <u>stationary</u> solutions with finite energy, hence ϕ at infinity must be on the orbit of minima of the potential $V = -\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4$. The exact 't Hooft-Polyakov solution, which asymptotically goes as $\phi^a \xrightarrow[\infty]{r} \hat{r}_a \left[\mu\lambda^{-1/2} + (\mathrm{const}/g\, r)\right.$ $\left.\exp(-\sqrt{2}\, m\, r)\right] \sim \hat{r}_a\,\mu\,\lambda^{-1/2}$, cannot be deformed continuously to a fixed direction in group space $\hat{n}^a$, and thus is not in equivalence class n=0 but in n=1 (remember: SU(2) $\simeq S_2$). More detailed analysis shows that this solution is a magnetic monopole with magnetic charge $g_m = n/e$ and mass $M = C\,(4\pi/g^2)\,g\,\mu\,\lambda^{-1/2}$, with a constant $C(\lambda/g^2) = O(1)$. Recently great progress has been made with multimonopole solutions (Ward, 1981; Prasad, 1981, Prasad and Rossi, 1981; Jaffe and Taubes, 1980). For larger

G the analysis can be extended.

We must thus seriously consider the <u>creation</u> of monopoles (non-trivial distribution of "directions" of ϕ^a) during phase transitions ("magnitudes") at unification temperatures. Alas a correct calculation taking care of gauge subtleties has not even been attempted, but the creation of vortices in superconduction experiments is indicative of the reality of the effect. In the following we assume the monopoles are not confined (cf. the suggestion of Linde, 1980). Naively one expects the ϕ^a directions to be uncorrelated over separations > d_H, which indicates O(1) monopole per volume element $O(d_H^3)$ (cf. Kibble, 1976; Einhorn, 1980). Because of the smallness of d_H, the large monopole mass O(100 M_X) and the slow annihilation their energy density would completely destroy the standard Helium synthesis result (Presskill, 1979). It was soon realised (Guth and Tye, 1980; Einhorn et al., 1980) that the stretching of d_H in a 1PT might be relevant. If the supercooling period ends by a shift at T_1, the directions in the many small bubbles being uncorrelated, some rough arguments on the correlation length seem to require $T_1 \leq O(10^8$ GeV) (Einhorn and Sato, 1981), not too far from the values discussed in § 6.1! If there <u>is</u> a 2 PT, or in other words $M_{Higgs} \gtrsim M_{gauge}$, perhaps thermal fluctuations reduce the monopole density enough (Bais and Rudaz, 1980).

7. Conclusion

Theories on the unification of the separate gauge theories for the three types of elementary particle interactions (electromagnetic, weak and strong) have direct relevance for the earliest phases of the Universe. The history of the Universe is described by the Hot Big Bang model, which gives the epoch of important unified interactions at times $t \sim 10^{-36}$ s and temperatures and typical energies of $\sim 10^{15}$ GeV $\sim$ $\sim 10^{28}$ K. This theoretical hubris is rewarded with an explanation of the presently observed matter-antimatter asymmetry. We have reviewed in some detail the finite temperature effects on the field theory, namely the phase transitions (PTs) at temperatures of order 10^{15} GeV (and 10^{2} GeV) between the different gauge symmetries. These PTs may have dramatic effects on the history of the expansion of the Universe, but we showed that the final "standard" results on the matter-antimatter asymmetry are hardly effected. Also we briefly discussed the expected creation of monopoles at these transitions. Naive estimates indicate much too high monopole densities, invalidating the synthesis of Helium with abundance $\sim$ 25%, as observed. Perhaps strong supercooling in first order PTs reduces the monopole density. We emphasized that in order to preserve the baryon number creation and Helium synthesis phase transitions are not allowed to create strong inhomogeneities. This might be a problem for first order transitions and the best way to end a I PT appears to be a destabilisation of the false vacuum at temperatures of order $10^{10} - 10^{6}$ GeV, for which there are indications that this indeed occurs.

This is the appropriate place to acknowledge discussions with A. Guth, D.V. Nanopoulos, K. Tamvakis, E. Weinberg and especially P. Hut, who also commented on the manuscript as did C. Norman.

Notes

[1] Ideas that the value of Eq. (11) might hold locally from imperfect (statistical) annihilation do not seem to work and also give typically $n_B/n_\gamma \lesssim 10^{-18}$ (cf. Steigman, 1976).

[2] If the Weinberg-Salam breaking is radiative, Witten (1981a) has calculated an entropy generation of $10^5 - 10^6$ from the 1 PT, which appears not to be allowed by the $\delta\theta$ arguments mentioned in chapter 4. More than one Higgs doublet, as seems required for ΔB, might reduce this supercooling (Flores and Sher, 1981).

[3] For completeness we remark that all our discussion of phase transitions might be completely changed if recent speculations on a global supersymmetry hold true. The major problem of GUTs is a natural explanation of the hierarchy $M_W/M_U \sim 10^{-13}$, or why do some scalars remain massless in the breaking at unification energies M_U? The heuristic scenario runs as follows (Witten, 1981b):
The gauge symmetry is broken to SU(3) × SU(2) × U(1) at energies M_U at the tree level, without breaking a global supersymmetry (GSS); the GSS remains unbroken up to <u>all</u> finite orders in perturbation theory: at low energies the GSS (and SU(2) × U(1)) is broken by a non-perturbative mechanism. The last step is still uncertain, though there are some analogies in lower dimensions, and, of course, how to reconcile GSS with finite temperature?

References

Actor, A. (1979) Rev. Mod. Phys. 51, 461

Abbott, L.F. (1981) Nucl. Phys. B185, 233

Bais, F.A., Rudaz, S. (1980) Nucl. Phys. B170 (FS1), 507

Barbieri, R. (1980) in International School of Physics E. Fermi, Varenna (preprint CERN-TH 2935)

Billoire, A., Tamvakis, K. (1981) Nucl. Phys. B200 (FS4), 329

Bond, J.R., Kolb, E.W., Silk, J. (1982) Astrophys. J. 255, 341

Coleman, S. (1977) Phys. Rev. D15, 2929 (also D16, 1762 and D21, 3305)

Coleman, S., Weinberg, E. (1973) Phys. Rev. D7, 1888

Crewter, R.J. (1978) Acta Physica Austriaca Suppl. XIX, 47

Daniel, M. (1981) Phys. Lett. 98B, 371

Daniel, M., Vayonakis, C.E. (1981) Nucl. Phys. B180, 301

Dine, M., Fisschler, W., Srednicki, M. (1981) Phys. Lett. 104B, 199

Einhorn, M.B. (1980) in Unification of Fundamental Particle Interactions eds., Ferrara, S., Ellis, J., Nieuwenhuizen, P. van, New York:Plenum

Einhorn, M.B., Sato, K. (1981) Nucl. Phys. B180, 385

Ellis, J. (1980) in Gauge Theories and Experiments at High Energy, eds. Bowler, K.S., Sutherland, D.G.

Ellis, J., Gaillard, M.K., Nanopoulos, D.V., Rudaz, S. (1980a) Nucl. Phys. B176, 61

Ellis, J., Gaillard, M.K., Peterman, A., Sachrajda, C.T. (1980b) Nucl. Phys. B164, 253

Ellis, J., Gaillard, M.K., Zumino, B. (1980c) Phys. Lett. 94B, 343

Ellis, J., Gaillard, M.K., Nanopoulos, D.V. (1980d), in Unification of Fundamental Particle Interactions, eds. Ferrara, S., Ellis, J., Nieuwenhuizen, P. van, New York:Plenum

Ellis, J., Gaillard, M.K., Nanopoulos, D.V., Rudaz, S. (1981a) Phys. Lett. 99B, 101

Ellis, J., Gaillard, M.K., Nanopoulos, D.V., Rudaz, S. (1981b) Nature 293, 41

Fetter, A.L., Walecka, J.D. (1971) Quantum Theory of Many Particle Systems, New York:McGraw-Hill

Flores, R.A., Sher, M. (1981) Phys. Lett. 103B, 445

Fujii, Y. (1981) Phys. Lett. 103B, 29

Georgi, H., Glashow, S.L. (1974) Phys. Rev. Lett. 32, 438

Georgi, H., Quinn, H.R., Weinberg, S. (1974) Phys. Rev. Lett. 33, 451

Guth, A.H. (1981) Phys. Rev. D23, 347

Guth, A.H., Tye, S.H. (1981) Phys. Rev. Lett. 44, 631, 963

Guth, A.H., Weinberg, E. (1981) Phys. Rev. D23, 876

Hut, P., Klinkhamer, F.R. (1981a) Astron. Astrophys. 106, 245

Hut, P., Klinkhamer, F.R. (1981b) Phys. Lett. 104B, 439

Jackiw, R. (1980) Rev. Mod. Phys. 52, 661

Jaffe, A., Taubes, C. (1980) Vortices and Monopoles, Boston:Birkhauser

Kibble, T.W.B. (1976) J. Phys. A9, 1387

Kirzhnits, D.A., Linde, A.D. (1976) Ann. Phys. 101, 195

Kolb, E.W., Wolfram, S. (1980a) Nucl. Phys. B172, 224

Kolb, E.W., Wolfram, S. (1980b) Astrophys. J. 239, 428

Langacker, P. (1981) Phys. Rep. 72, 185

Linde, A.D. (1979) Rep. Prog. Phys. 42, 389

Linde, A.D. (1980), Phys. Lett. 96B, 293

Nanopoulos, D.V. (1980) in XVième Rencontre de Moriond (preprint CERN-TH 2896)

Nanopoulos, D.V., Weinberg, S. (1979) Phys. Rev. D20, 2484

Olive, K., Schramm, D.N., Steigman, G., Turner, M.S., Yang, Y. (1982) Astrophys. J. 246, 557

O'Raifeartaigh, L. (1979) Rep. Prog. Phys. 42, 159

Pagels, H., Tomboulis, E. (1978) Nucl. Phys. B143, 485

Peccei, R.D., Quinn, H.R. (1977) Phys. Rev. D16, 1791

Prasad, M.K. (1980) Physica 1D, 167

Prasad, M.K. (1981) Comm. Math. Phys. 80, 137

Prasad, M.K., Rossi, P. (1981) Phys. Rev.24,2182 (also Phys. Rev. Lett. 46, 806)

Presskill, J.P. (1979) Phys. Rev. Lett. 43, 1365

Sato, K. (1981) Monthly Notices Roy. Astron. Soc. 195, 467

Sher, M. (1981) Phys. Rev. D24, 1699

Steigman, G. (1976) Ann. Rev. Astron. Astrophys. 14, 339

Taylor, J.C. (1976) Gauge Theories of Weak Interactions Cambridge:Cambridge UP

Tamvakis, K., Vayonakis, C.E. (1981) Phys. Lett. 109B, 283

Ward, R.S. (1981) Comm. Math. Phys. 79, 319

Weinberg, S. (1972) Gravitation and Cosmology, New York:Wiley

Weinberg, S. (1974) Phys. Rev. D9, 3357

Weinberg, S. (1979a) Phys. Rev. Lett. 42, 850

Weinberg, S. (1979b) Phys. Lett. 82B, 387

Witten, E. (1979) Nucl. Phys. B149, 285

Witten, E. (1981a) Nucl. Phys. B177, 477

Witten, E. (1981b) Nucl. Phys. B188, 513

Yildiz, A., Cox, P.H. (1980) Phys. Rev. D21, 906

CHAPTER 3

BARYON NUMBER CREATION AND PHASE TRANSITIONS IN THE EARLY UNIVERSE, figures 2-5

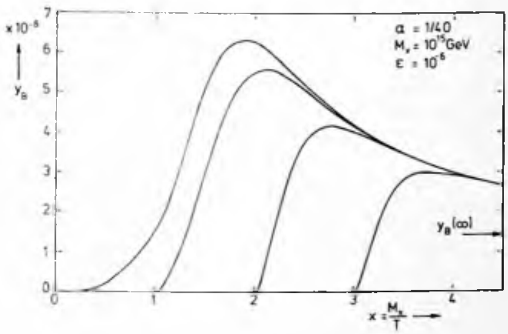

Fig. 2. The generation of the baryon number to photon ratio $Y_B(x, x_0)$, where $x = M_X/T$ and M_X the mass of the B-violating boson, is calculated for a simple model, with CP violation parameter ε. Earlier calculations (Kolb and Wolfram, 1980b) use symmetric starting conditions at $x_0 = 0$. Finite temperature effects lead to phase transitions (*PT*), *which give the boson a mass* M_X only for $T < T_c \sim M_X/g$. The $Y_B(x; x_0)$ evolution is calculated for realistic starting values after a second order *PT* ($x_0 \sim g \sim \frac{1}{2}$) or after the smooth reheating ending the period of supercooling of a first order *PT* ($x_0 \sim 1$). The final Y_B values are given in Fig. 3

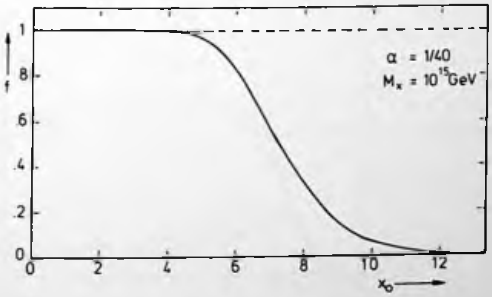

Fig. 3. Final baryon-number to photon ratio $Y_B(\infty, x_0)$ for realistic starting conditions ($x_0 \sim 1$) as compared to $x_0 = 0$: $f \equiv Y_B(\infty, x_0)/Y_B(\infty, 0)$. In this case $Y_B(\infty, 0) = 1.43\ 10^{-8}$

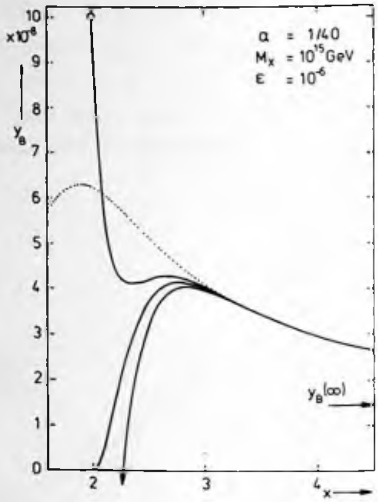

Fig. 4. The generation of the final baryon-antibaryon asymmetry for a first order phase transition, if the thermalisation of the small bubbles of true vacuum at the end of the supercooling epoch itself gives a net baryon number $Y_B^{\text{bubbles}}(x_0=2)=\pm 10^{-7}$. For comparison the curves for $Y_B(x_0=0)=0$ (dotted) and $Y_B(x_0=2)=0$ (Fig. 2) are given

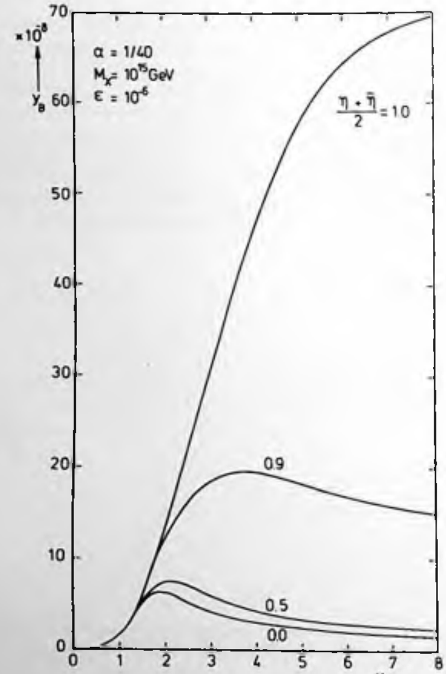

Fig. 5. $Y_B(x; x_0=0)$ for different branching ratios of X, $\bar{X}$ decays (see Appendix A). For $(\eta+\bar{\eta})/2=1$ there is a nearly total suppression of the B damping processes

GLOBAL SPACE–TIME EFFECTS ON FIRST-ORDER PHASE TRANSITIONS FROM GRAND UNIFICATION

P. HUT
Astronomical Institute, 1018 WB Amsterdam, The Netherlands

and

F.R. KLINKHAMER
Astronomical Institute, 2300 RA Leiden, The Netherlands

Received 7 May 1981

We argue that the supercooling of a first-order phase transition proceeds only to $T \sim 10^{11}$ GeV (calculated for a Coleman–Weinberg potential). Then the barrier width between real and false vacuum as calculated in flat space–time becomes comparable to the scale set by the event horizon, and mode mixing might induce the transition.

We consider phase transitions at grand unification energies, which might have taken place in the early history of the universe. First we discuss the local physics and its implications on the expansion of the universe and secondly we turn to possible effects of the global space–time structure.

Gauge theories for the unification of the strong, electromagnetic and weak interactions have at least two transitions towards a larger symmetry at high energies:

$$G \xrightarrow[M_U \sim 10^{15}\,\mathrm{GeV}]{} \mathrm{SU}(3)\times\mathrm{SU}(2)\times\mathrm{U}(1)$$

$$\xrightarrow[M_{WS} \sim 10^{2}\,\mathrm{GeV}]{} \mathrm{SU}(3)\times\mathrm{U}(1)\,,$$

where the group G describes the unified interaction with a single coupling constant g and with quarks and leptons in common representations (for reviews see refs. [1,2]). In order to apply these ideas to the early universe [3], finite-temperature field theory is used and one finds symmetry restoration for high enough temperatures, resembling phase transitions [4]. Spontaneous breaking of symmetries is due to non-zero expectation values [‡1] of the Higgs scalars φ introduced in the theory. This mechanism preserves renormalisation while providing masses to some gauge bosons and fermions [5]. In a first-order phase transition (1 PT) the shift towards non-zero $\langle\varphi\rangle$ is discontinuous, in contrast to the smooth change for a second-order phase transition. Which type of transition occurs depends on the parameters in the effective potential of the scalars $V(\varphi_c, T)$ [6]. To have breaking at very different energies some very special fine-tuning in V is required [7], which may hint [2] to symmetry breaking by radiative terms only [6]. This in turn might explain the hierarchy of hierarchies $M_{WS}/M_U \ll M_U/M_{Pl} \ll 1$, if the quartic coupling constants (λ) are of order g^2 at the Planck energy $M_{Pl} = G^{-1/2} = 1.2 \times 10^{19}$ GeV ($\hbar = c = k = 1$) [8]. Perhaps superunification [9] leads to radiatively broken $G = \mathrm{SU}(5)$. The important point here is that Coleman–Weinberg (CW) breaking leads to strongly first-order phase transitions [10,11]. In the following we will use this potential for numerical estimates.

We now consider the scenario for a 1 PT in the cooling universe. Initially the vacuum is symmetric because of a positive temperature-dependent mass^2 term in the effective potential V [12]. For temperatures $T < T_c$ [where for the two minima $V(\varphi_c = 0, T_c) = V(\varphi_c \neq 0, T_c)$] the transition to the energetically favourable broken state is blocked and the

‡1 For $T = 0$ vacuum expectation values, for $T \neq 0$ relative to a Gibbs ensemble of temperature T.

universe cools far below the critical temperature T_c to T_{end} when the transition takes place and the latent heat reheats the universe to fT_c, where $f = O(1)$ follows from entropy conservation and depends on the available particle states before and after reheating [13]. For CW breaking of SU(5) the potential is ($\varphi \ll T \ll \sigma$)

$$V = \tfrac{5}{8} g^2 T^2 \varphi_c^2 + B \varphi_c^4 (\ln \varphi_c^2/\sigma^2 - \tfrac{1}{2}) + \tfrac{1}{2} B\sigma^4 , \quad (1)$$

with φ_c a classical scalar field, $\langle\varphi\rangle_{T=0} = \sigma$ and numerical constant $B = 8 \times 10^{-4}$ [11,14]. The zero level is such that $V(\sigma, T = 0) = 0$, as required by the presently observed zero cosmological constant [15]. The part of the universe still in the symmetric state has a constant vacuum energy density $\rho_v \sim T_c^4$ which for $T < T_c = 0.3\,\sigma$ [11] leads to exponential expansion $a \propto \exp(t/\tau)$ as follows from the Friedmann equation [3]:

$$(\dot{a}/a)^2 = (8\pi/3M_{Pl}^2)(\rho_v + \tfrac{1}{30}\pi N T^4) , \quad (2)$$

where $a(t)$ is the scale factor and N the effective number of degrees of freedom of relativistic particles. The transition to the broken vacuum takes place through nucleation with a minimal bubble radius; either through thermal excitation or through tunneling [16].

How precisely the transition to the broken state for the whole universe takes place is of crucial importance and in general we can distinguish three cases:

(1) Thermal nucleation rates have a maximum just below T_c and either the bubble density gets high enough and they quickly fill the universe or else the bubbles cannot catch up with the continuously accelerated expansion of the rest of the universe [17,18].

(2) Nucleation through tunneling has a constant rate per space volume. This leads to a large supercooling, which originally was the motivation to consider 1 PT's in order to prevent high monopole densities [19] ‡2. The nucleation rate for the CW SU(5) model equals the expansion rate only at $T = O(1\,\mathrm{GeV})$ and the transition is directly to SU(3) × SU(2) × U(1), not through an intermediate SU(4) × U(1) [11]. We note that the barrier vanishes here at $T = 0$, which need not be general (e.g. if $\lambda \ll g^4$ in the abelian Higgs model [4]).

‡2 The suppression mechanism of ref. [20] requires large Higgs masses $m_H \gtrsim m_X$, which is not the case for CW breaking.

(3) Nucleation will be immediate if at T_1 the metastable symmetric vacuum becomes unstable (abelian Higgs model: $3g^4/16\pi^2 < \lambda < g^4$; in SU(5) region d of ref. [18]). The false vacuum shifts to the broken state while releasing its latent heat. If there typically is one monopole in a horizon volume their densities would be tolerable for $T_1 \lesssim 10^{12}$ GeV, although correlations starting to form after T_1 suggest the creation of more monopoles, and require $T_1 \lesssim 10^8$ GeV [13,21].

In all cases it is important to reheat locally at least to the Higgs mass so that a new baryon asymmetry can be created, since previously developed asymmetries are strongly diluted by the supercooling.

In first-order phase transitions the energy density of the metastable vacuum drives the expansion of the universe. However, we must also try to incorporate effects of the *global* structure of this accelerated expansion. In particular we will point out where the flat-space equilibrium theory used so far breaks down. Without a consistent theory of quantum gravity and one of non-equilibrium corrections to finite-temperature particle interactions, we can only attempt to give a qualitative picture, as suggested by a semiclassical approximation (see below) analogous to the Hawking effect ‡3.

The major difference between cosmology and flat space–time (laboratory) physics is the existence of horizons:

(1) A particle horizon limits the region of possible causal contact with a given observer before a given time, and thus includes all comoving particles which have intersected the observers past lightcone.

(2) An event horizon limits the region which will in future have the possibility for contact with a geodetic observer and thus is the boundary of the past lightcone of the observer for $t \to \infty$.

In the standard Big Bang model there are only particle horizons. With relativistic particles, adiabatic expansion $T \propto a^{-1}$, we have from (2) $a \propto t^{1/2}$ and hence the maximum proper distance travelled by a light signal up to time t is [3]

$$d_H = a(t) \int_0^t \mathrm{d}t'/a(t') = 2t \qquad (a \propto t^{1/2}) , \quad (3a)$$

$$\sim \tau e^{t/\tau} \quad (a \propto e^{t/\tau}), \quad (3b)$$

‡3 We thank Eardly and Press for reminding us of the Hawking radiation in de Sitter space.

The exponential expansion during a 1 PT introduces also event horizons: two observers initially separated by a distance significantly greater than the exponential timescale $\tau \sim M_{\rm Pl}\,\rho_{\rm v}^{-1/2}$ will never be able to communicate, because the intermediate region expands with a constant acceleration so that light signals never catch up. The distance to the event horizon is [22]

$$D_{\rm H} = (3/\Lambda)^{1/2} = (3/8\pi)^{1/2}\,M_{\rm Pl}\,\rho_{\rm v}^{-1/2}\,. \qquad (4)$$

Here lies the origin of all evil: the flat space–time approximation is expected to be valid for distances much smaller than that of the event horizon, for which a locally inertial reference frame is a good approximation. But for length scales comparable to or larger than the event horizon distance $D_{\rm H}$, or, equivalently, for energies $\lesssim D_{\rm H}^{-1}$ the flat space–time approximation clearly breaks down.

How does this affect our previous discussion of the 1 PT? We expect changes to be small if we discuss effects at values of the classical field $\varphi_{\rm c}$ (or $\langle\varphi\rangle$) much larger than $D_{\rm H}^{-1}$. But the description of the stability or calculations of tunneling rates of the symmetric vacuum will fail for small $\varphi_{\rm c}$. An (educated) guess of the temperature T^* below which the flat space–time approximation breaks down is made by equating the barrier width $\Delta\varphi_{\rm c}$ and $D_{\rm H}^{-1} \sim T_{\rm c}^2/T_{\rm Pl}$. For the CW potential we find roughly

$$\Delta\varphi_{\rm c}^2 \sim (g^2T^2/2B)\ln(g\sigma/T) \sim 10^{-2}g^2T^2\,,$$

where we replaced the second term in the RHS of (1) by $-2B\,\varphi^4\ln(g\sigma/T)$ [14] and hence the flat space–time treatment of the barrier will not be valid for

$$T < (T_{\rm c}/T_{\rm Pl})10^{-1}g^{-1}T_{\rm c} \sim 10^{11}\ {\rm GeV}. \qquad (5)$$

Because confinement of the Higgs expectation value in the metastable symmetric state requires local effects on scales which are globally distorted by the background metric for $T < T^*$, we expect that the universe cools to $\sim T^*$ and then shifts to the broken state, thereby ending the supercooling prematurely. Although we cannot give a rigorous proof for our assertion, we will now discuss several analogies in its favour.

Since $T^* \ll T_{\rm Pl}$ one may use a semi-classical approach where gravity is treated classically through general relativity (GR) and the particles as quantum fields. Note the contrast between the local approach in GR and the global treatment for the particle fields. In particular the definition of the Hilbert space of particle states requires a priori causality relations, whereas only the solution of GR equations provides time- or space-like separations between events (cf. ref. [22]). Unambiguous particle states can only be defined for space–time backgrounds somehow related to Minkowski space [24], to which Robertson–Walker and de Sitter spaces are linked by a conformal transformation. The problem is to treat the backreaction of the particles on the metric and as a first guess one uses a somehow regulated energy–momentum tensor of the particle fields as a classical source term in the Einstein equations.

A major result of the semi-classical approach is the prediction of thermal radiation from an isolated black hole with temperature [25]:

$$T_{\rm H} = (8\pi)^{-1}\,M_{\rm Pl}^2/M_{\rm BH}\,, \qquad (6)$$

which can be derived in several ways:

(1) A mapping of a complete set of particle states from the asymptotic past into one of the asymptotic future shows that an incoming vacuum state leads to the emergence of a thermal particle spectrum [25].

(2) Thermodynamic considerations suggest (6), up to a factor of order unity, because information on the quantum states is lost by the existence of an event horizon. The entropy of the black hole can be estimated, from which $T_{\rm H} = (\partial S/\partial M)^{-1}$ follows [26].

(3) Path integrals on a complexified Schwarzschild metric give a propagator of the form of a thermal Green's function with temperature $T_{\rm H}$ [27].

These derivations are consistent, because the asymptotic region of space–time is flat, where particle states can be defined unambiguously. All distant observers agree on the Hawking radiation (6), but an infalling observer near the horizon will hardly see any radiation [28]. This observer dependency always occurs locally in globally curved space–times for wavelengths of the order of the curvature radius (cf. Schwarzschild radius $2GM \sim T_{\rm H}^{-1}$). Even in a Minkowski vacuum a constantly accelerated observer will detect thermal radiation, because his detector measures positive frequencies with respect to his own proper time [29]. For the accelerated observer there also is an event horizon. Both the inertial and accelerated observer agree that the detector will be excited, but they differ on the interpretation, namely bremsstrahlung and absorption, respectively.

Also in our cosmological context similar phenomena occur. In a Friedmann universe the particles determine a preferred restframe. If the vacuum energy density, which is locally Lorentz invariant, dominates over that of the particles, accelerated expansion takes place (2), and the universe asymptotically approaches de Sitter space, where geodetic observers are equivalent. Gibbons and Hawking [22] showed with the same path-integral technique as for the black-hole case that every geodetic detector will see radiation with a temperature

$$T_{GH} = (12)^{-1/2}\pi^{-1}\Lambda^{1/2} \sim M_{Pl}^{-1}\rho_v^{1/2} \sim T_c^2/T_{Pl}. \quad (7)$$

Two differences with the black hole case are in order: (1) absorption of the thermal particles does not destabilize the event horizon [22], in contrast to the increasing rate of black-hole evaporation; (2) no observer independent definition of this radiation is possible [‡4]; indeed *if* this could be done, local Lorentz invariance would give an infinite total energy density from the superposition of the finite contributions (7) of all equivalent observers.

To make the link with our assertion that the false vacuum indeed decays at a temperature T^* (5), we now give a general physical picture for the above results. Parker [28] notes that whenever a physical system is externally disturbed on a time-scale τ, modes with frequencies $\omega \lesssim \omega_{cr} \sim \tau^{-1}$ are excited. Both in the black-hole and the de Sitter case, the particle production results from a mixing of positive and negative frequencies of the particle fields, caused by the time dependence or curvature of the background metric. Modes are excited with energies

$$\omega \lesssim \omega_{cr}, \quad \omega_{cr} \sim (GM)^{-1} \quad \text{(Schwarzschild)}, \quad (8a)$$

$$\sim \Lambda^{1/2} \quad \text{(de Sitter)}, \quad (8b)$$

which agrees with the exponential drop in a Planck spectrum for $\omega > T$, with T given by (6) or (7). Parker also shows that relations (8) imply particle creation near enough to the event horizons so that the Heisenberg uncertainty for detection would be large enough to compensate for the negative energy of one of the particles, thus providing an energy reservoir for detection of Hawking radiation. As mentioned above, the backreaction on the event horizon is different in the black-hole and the de Sitter case.

Similarly, mode mixing will occur for energies given by (8b) during the exponential expansion in a 1 PT in the early universe. As soon as the potential barrier around the false vacuum becomes narrower than this range, decay is no longer prohibited rigorously and we expect the transition to occur.

Finally we compare our discussion with a recent article by Shore [30], who considers CW-breaking with a given strong curvature in de Sitter space [‡5]. He finds symmetry restoration for a curvature R with a Hawking temperature (7) larger than the Higgs or gauge boson mass for conformally ($\frac{1}{6}R\varphi^2$) or minimally (no $R\varphi^2$) coupled scalars, respectively. But in a 1 PT the curvature due to the false vacuum is less than this critical value by a factor $\sim T_c^2/T_{Pl}^2$. Hence this curvature will not affect the existence of an asymmetric true vacuum. In our view, the important effect of space–time curvature is not a qualitative change in the effective potential, such as the disappearance of a minimum, but the irrelevance of a stabilizing narrow barrier around the false vacuum when mode mixing occurs with respect to a flat space–time approximation.

Some implications of our suggestion of a globally induced transition to the broken vacuum at $T^* \sim 10^{11}$ GeV: (1) perhaps low enough monopole densities [13] even if there are no other suppression mechanisms [31]; (2) no unnatural ($< M_{ws} \sim 10^2$ GeV) supercooling; (3) a smooth transition and reheating, thereby saving baryon number and helium synthesis. All's well that ends well.

‡4 But perhaps two geodetic observers, passing each other, might agree, because added to the expected Doppler shift is an Unruh-type radiation from their relative acceleration [23].

‡5 Abbot [14] considers barrier penetration for an ad hoc $R\varphi^2$ term in the flat-space CW potential. But for the numerical value of the curvature R determined by ρ_V his temperature-independent barrier is of the order of our uncertainty range D_H^{-1}.

References

[1] D.V. Nanopoulos, XVième Rencontre de Moriond, CERN-TH 2896 (1980).
[2] J. Ellis, 21st Scottisch Universities Summer School in Physics, CERN-TH 2942 (1980).
[3] S. Weinberg, Gravitation and cosmology (Wiley, 1972).

[4] D.A. Kirzhnits and A.D. Linde, A.. Phys. 101 (1976) 195; and references therein.
[5] J.C. Taylor, Gauge theories of weak interactions (Cambridge U.P., 1976).
[6] S. Coleman and E. Weinberg, Phys. Rev. D7 (1973) 1888.
[7] A.J. Buras, J. Ellis, M.K. Gaillard and D.V. Nanopoulos, Nucl. Phys. B135 (1978) 66.
[8] J. Ellis, M.K. Gaillard, A. Peterman and C.T. Sachrajda, Nucl. Phys. B164 (1980) 253.
[9] J. Ellis, M.K. Gaillard and B. Zumino, Phys. Lett. 94B (1980) 343.
[10] M. Daniel, Phys. Lett. 98B (1981) 371.
[11] A. Billoire and K. Tamvakis, preprint CERN-TH 3019 (1981).
[12] S. Weinberg, Phys. Rev. D9 (1974) 3357.
[13] M. Einhorn and K. Sato, NORDITA preprint (July 1980).
[14] L.F. Abbot, preprint CERN-TH 3018 (1981).
[15] E.W. Kolb and S. Wolfram, Astrophys. J. 239 (1980) 428.
[16] S. Coleman, Phys. Rev. D15 (1977) 2929;
C.G. Callan and S. Coleman, Phys. Rev. D16 (1977) 1762.
[17] K. Sato, NORDITA preprint-80/29 (January 1980).
[18] A.H. Guth and E.J. Weinberg, Phys. Rev. D23 (1981) 876.
[19] M.B. Einhorn, D.L. Stein and D. Toussaint, Phys. Rev. D21 (1980) 3295;
A.H. Guth and S.H. Tye, Phys. Rev. Lett. 44 (1980) 631, 963.
[20] F.A. Bais and S. Rudaz, preprint CERN-TH 2885 (1980).
[21] M. Einhorn, in: Unification of fundamental particle interactions, eds. J. Ellis and P. van Nieuwenhuizen (Plenum, 1980).
[22] G.W. Gibbons and S.W. Hawking, Phys. Rev. D15 (1977) 2738.
[23] L. Smolin, preprint (September 1979).
[24] B.L. Hu, in: Recent developments of general relativity, ed. R. Ruffini (1980).
[25] S.W. Hawking, Commun. Math. Phys. 43 (1975) 199.
[26] J.D. Bekenstein, Phys. Rev. D9 (1974) 3292;
S.W. Hawking, Phys. Rev. D13 (1976) 191.
[27] J.B. Hartle and S.W. Hawking, Phys. Rev. D13 (1976) 2188.
[28] L. Parker, in: Asymptotic structure of spacetime, eds. F. Esposito and L. Witten (Plenum, 1977).
[29] W.G. Unruh, Phys. Rev. D14 (1976) 870.
[30] G.M. Shore, Ann. Phys. 128 (1979) 376.
[31] A.D. Linde, Phys. Lett. 96B (1980) 293.

SUPERSYMMETRIC UNIFICATION AND COSMOLOGY

F.R. KLINKHAMER
Leiden Observatory, 2300 RA Leiden, The Netherlands

Received 13 October 1981

We discuss the cosmological implications of a model of supersymmetric unification by Witten. Because the unification phase transition occurs with a low critical temperature, we estimate the final baryon number to be negligible.

Recently global supersymmetry has been evoked to explain the hierarchy problem of unification theories: why is the breaking scale of unification $M_U \sim 10^{15}$ GeV so much larger than that of ordinary physics, say the breaking of the Weinberg–Salam electroweak interactions at $M_{WS} \sim 300$ GeV [1]? In this letter we will consider a specific model of Witten [2] and its role in the early universe, where the expected low critical temperature might have important implications for monopoles and the creation of the presently observed matter–antimatter asymmetry [3]. In our units $\hbar = c = k_B = 1$.

Ref. [4] gives the details of having an additional $N = 1$ global supersymmetry (GSS), which survives the breaking of the unified gauge group at energy scale M_U, but is to be broken at low energies $M = 0\,(10^3$ GeV) by weak non-perturbative effects, triggering the Weinberg–Salam breaking. The important difference with respect to the breaking of internal symmetries is that GSS is broken *if and only if* the minimum of the (tree) potential is unequal to, i.e. larger than, zero. A difficulty with these models is that the SU(5) partners of the normal Higgs doublet have masses $\lesssim M$, being the supersymmetry partners of the down-like quarks, and would mediate a much too fast proton decay.

Another approach is to break the GSS explicitly at low energies M [2]. The potential for the required three complex scalar fields A, X and Y is

$$V = \lambda^2 |A^2 - M^2|^2 + g^2 |A|^2 + |2\lambda AX + gY|^2\,, \qquad (1)$$

which clearly breaks supersymmetry ($V > 0$). For $g/\lambda M \gg 1$ the minimum is at $A = 0$, but $|X|$ can be arbitrarily large (fig. 1). The GSS breaking is set by M

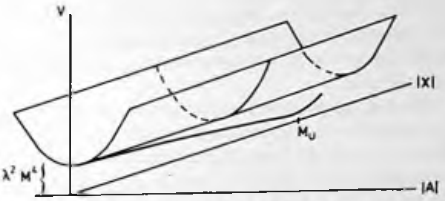

Fig. 1. Sketch of the potential V of eq. (1) with broken supersymmetry, i.e. min $V \neq 0$. The minimum of the potential is at $|A| = 0$, but the value of $|X|$ remains undetermined. One-loop corrections with gauge bosons might produce a minimum at $|X| = M_U$ (heavy line). The energy scale M_U, associated with the breaking of the unified gauge symmetry, can be much larger than the scale M of the explicit supersymmetry breaking.

($V \sim \lambda^2 M^4$; mass splittings), but an arbitrary large energy scale may arise from $\langle 0|X|0\rangle$, which up to now is undetermined. This is illustrated by the two masses $m_{1,2}$ of the A field: $m^2_{1,2} = 4\lambda^2 |\langle 0|X|0\rangle|^2 + 2g^2 \pm 2\lambda^2 M^2$.

As with the Coleman–Weinberg type of breaking [5] one-loop corrections determine the value of $\langle 0|X|0\rangle$. Including gauge fields (coupling constant e) the valley of the potential (fig. 1) is for large $|X|$

$$V_{A=0} = a\lambda^2 M^4 [1 + (b\lambda^2 - ce^2) \ln(|X|^2/\mu^2)]\,, \qquad (2)$$

with some positive constants a, b, c and μ the mass scale from renormalization. If $(b\lambda^2 - ce^2) < 0$ a minimum might be expected for large $|X|$ values [perhaps at $|X| \sim M \exp(1/e^2)$] and a hierarchy of scales has arisen. In ref. [2] a more realistic model with SU(5) broken at the large energy scale $\langle 0|X|0\rangle$ is discussed, which also has large masses for the coloured Higgs triplet saving the proton from rapid decay.

To do cosmology, temperature effects must be considered. If the usual finite-temperature field theory applies it is evident that supersymmetry is broken: the ground state is a statistical ensemble, where fermions and bosons are treated differently, and is not invariant under GSS, which results in differences of fermion and boson Green's functions [6]. Also the path integral formalism is illustrative: the Lagrange density is invariant under GSS up to a total divergence, which by the finite integration interval of imaginary time $[0, 1/T]$ gives a surface contribution. Finally, we note that for explicitly broken GSS the Goldstone fermion remains massless up to one-loop, even without protection from chirality [6]. Hence, finite temperatures will only enhance the breaking of supersymmetry, contrary to the typical case of internal symmetries, to which we will turn now.

For high temperature loops in a renormalizable theory contribute to the effective potential terms of the form T^4 and $fT^2\phi^2$, with f depending on the coupling constants. The first terms just are the contribution of the thermal fluctuations to the energy density, while the second arise from the quadratically divergent self-energy diagrams of the scalars ϕ. Normally these will lead to symmetry restoration above a critical temperature T_c [7], which can be estimated as follows. The minimum of a typical potential $V(\phi) = -m^2\phi^2 + l\phi^4$ lies at $\bar{\phi}^2 = m^2/2l$ with a value $-m^4/4l \sim -l\bar{\phi}^4$ and will disappear at temperatures of order T_c when $m^4/l \sim f\bar{\phi}^2T_c^2$, with f of order 1. Hence T_c is of the order of the vacuum expectation value $\bar{\phi}$. The same conclusion does *not* hold for the potential of eq. (2), where the vacuum energy density λ^2M^4 is much less than M_U^4:

$$T_c \sim \lambda f^{-1/2}(M/M_U)\,M\,. \tag{3}$$

Although the "derivation" of eq. (3) is not very refined, it is clear that T_c will depend on M, but perhaps the numerical factor will be less small because of some non-perturbative effect. This means that the universe cools with the unified symmetry to a very low temperature T_c, or even lower, since one might expect a first-order phase transition. More or less simultaneously the Weinberg–Salam breaking will take place.

What are the implications for the creation of monopoles at the phase transition? Although it is not clear whether one may neglect gravity, the naive monopole mass is so large ($M_{mon} \sim \alpha^{-1}M_U \gtrsim M_{Pl}$) that they might collapse immediately into black holes [8]. The standard estimate of the evaporation time is very small, $O(M_{Pl}^{-1})$, but the black holes *cannot* decay because their magnetic charge is conserved and there are no lighter monopoles. One might hope that because the transition occurs relatively late [eq. (3)] the correlation length for the Higgs fields can be large enough to have negligible monopole (black hole) density (cf. section IV of ref. [9]). As for the standard unification the monopole problem remains unsolved.

Let us neglect the usual complications of the transition and reheating to T_R [10] and consider the creation of a baryon number density at temperature $T_R \sim M$ after a smooth transition and instantaneous thermalisation of the latent heat $\rho_{vac} \sim M^4$. In the standard unification theories one worries not too much about the role of phase transitions for the creation of the observed baryon number, since $T_c \sim M_U$ is of the same order as the masses of the relevant heavy bosons X. But for Witten's model [2] the situation changes drastically [‡1]. During the symmetry period no baryon number is created, the X being massless, and any truly primordial baryon number will be washed out completely [11]. Before the transition there are roughly as many X as photons, $n_X \sim T_c^3$.

The heavy bosons acquire instantaneously, by assumption, a mass $M_X \sim M_U$ at temperatures T_R. They decay with a small asymmetry, not being replenished by inverse decays, and the net baryon number density created is proportional to $n_X(T_R)$ (cf. ref. [3]). The crucial question is if $n_X(T_R)$ still is of order T_c^3 or is reduced to $(M_XT_R)^{3/2}\exp(-M_X/T_R)$? Some estimates of the expansion time scale of the universe, and of the decay and annihilation time scales of the heavy bosons X are, at temperature T_R:

$$t \sim N^{-1/2}T_R^{-2}M_{Pl} \sim N^{-1/2}(M_{Pl}/T_R)(M_X/T_R)M_X^{-1}\,,$$
$$\tau_D \sim \alpha^{-1}N^{-1}M_X^{-1}\,,$$
$$\tau_{ann} \sim (n\langle\sigma\rangle)^{-1} \sim \alpha^{-2}(M_X/T_R)^3\,M_X^{-1}\,, \tag{4}$$

‡1 The condition to create the maximal baryon asymmetry by the delayed decay of the X [$\tau_D \sim t$ at $T_D < M_X$, cf. eq. (4)] is $M_X > N^{1/2}\,\alpha_U M_{Pl}$ [3]. Including the supersymmetric partners of the observed particles one estimates $M_U \sim 10^{18}$ GeV and $\alpha_U \sim 1/30$ [8]. Contrary to the case for the standard unification heavy gauge bosons now fulfill the condition.

with N the total effective number of states, $N = O(100)$, and $M_{\rm Pl} \equiv G^{-1/2} \sim 10^{19}$ GeV. From the hierarchy $\tau_{\rm D} \ll t \ll \tau_{\rm ann}$ one concludes that the X decay immediately and we have the baryon number to entropy ratio

$$n_{\rm B}/s \sim [0.14(N_{\rm X}/N)\,\Delta B]\,(T_{\rm c}/T_{\rm R})^3\ , \qquad (5)$$

where ΔB is the average net baryon number from the decay of a X–$\bar{\rm X}$ pair. If the estimate of eq. (3) contains some truth the created baryon number is negligible, because of the factor $(T_{\rm c}/T_{\rm R})^3 \sim (M/M_{\rm U})^3$. There are two further arguments:

(1) It might be that in the mysterious thermalisation required at the transition also the X bosons attain their equilibrium densities with the miniscule Boltzmann factor (cf. ref. [11]) [‡2].

(2) The energy densities before (b) and after (a) the transitions are roughly: $\rho_{\rm b} \sim \rho_{\rm vac} + \rho_{\rm therm} \sim M^4 + T_{\rm c}^4 \sim M^4$ and $\rho_{\rm a} \sim T_{\rm R}^4 + M_{\rm X} n_{\rm X,a}$. Energy conservation [‡3] would give for $n_{\rm X,a}$ the generous upper limit $(M/M_{\rm X})(M/T_{\rm R})^3\,T_{\rm R}^3$. Hence a suppression factor of only (!) $M/M_{\rm X}$ compared to the standard baryon number result [square brackets in eq. (5)].

To summarize, the baryon number created after a phase transition at low critical temperature appears to be negligible. This might be a severe problem for Witten's model of supersymmetric unification [2], and also the created monopole density is problematic. But one should keep in mind the proviso of our footnote 3.

I thank K. Tamvakis for introducing me to the finite-temperature breaking of GSS. This work was started in the pleasant atmosphere of the Les Houches Summerschool "Gauge Theories in High Energy Physics" and participation was made possible by a grant of the Netherlands Organization for the Advancement of Pure Research (ZWO). F.A. Bais' comments on the manuscript were appreciated.

Note added. Ginsparg [12] also has considered the phase transition of Witten's model. Our use of the high-temperature approximation is not correct and $T_{\rm c}$ is just somewhat below M [not by the factor $M/M_{\rm U}$ of eq. (3)]. Hence the reduction factor $(T_{\rm c}/T_{\rm R})^3 \sim (T_{\rm c}/M)^3$ of eq. (5) may be appreciable, but not as small as $(M/M_{\rm U})^3$. The further two arguments for reduced baryon number remain.

Ellis et al. [13] discuss the low-energy predictions from unification with GSS. Typically $M_{\rm X}$ is of order 10^{16} GeV, in contrast to the value quoted from ref. [8], and hence the relevant monopoles do not collapse gravitationally.

‡2 This is also the reason why we gave an upper limit to $n_{\rm X}(T_{\rm R})$ of $\sim T_{\rm c}^3$ and not $T_{\rm R}^3$. We expect the number of X bosons created in the thermalisation to be much less than that of massless particles.

‡3 In this letter we take the point of view that the Bose condensate gravitates normally and has its zero energy-level fixed by the observation that the present cosmological constant is vanishingly small. But we keep in mind, especially since we introduce supersymmetry and because the unification scale will not be far below $M_{\rm Pl}$, that the gravitational aspects of symmetry breaking might be radically different.

References

[1] See for a review: J. Ellis, in: Gauge theories and experiments at high energy, eds. K.S. Bowler and D.G. Sutherland (1981).
[2] E. Witten, Phys. Lett. 105B (1981) 269.
[3] S. Weinberg, Phys. Rev. Lett. 42 (1979) 850;
D.V. Nanopoulos and S. Weinberg, Phys. Rev. D20 (1979) 2484.
[4] E. Witten, Nucl. Phys. B188 (1981) 513.
[5] S. Coleman and E. Weinberg, Phys. Rev. D7 (1973) 1888.
[6] L. Girardello, M.T. Grisaru and P. Salomonson, Nucl. Phys. B178 (19810 331.
[7] D.A. Kirzhnits and A. Linde, Ann. Phys. 101 (1976) 195
A. Linde, Rep. Prog. Phys. 42 (1979) 389.
[8] S. Dimopoulos, S. Raby and F. Wilczek, Phys. Rev. D24 (1981) 1681.
[9] M.B. Einhorn and K. Sato, Nucl. Phys. B180 (1981) 385
[10] See for a review: F.R. Klinkhamer, in: Oxford Intern. Symp. Progress in cosmology (September 1981) (Reidel, Dordrecht), to be published.
[11] P. Hut and F.R. Klinkhamer, Baryon number creation and phase transitions in the early universe, submitted to Astron. Astrophys.
[12] P. Ginsparg, Phys. Lett. 112B (1982), to be published.
[13] J. Ellis, D.V. Nanopoulos and S. Rudaz, CERN preprint TH 3199.

CHAPTER 6

QUARK LIBERATION AT HIGH TEMPERATURE

Les retentissantes couleurs
Dont tu parsèmes tes toilettes
Jettent dans l'esprit des poètes
L'image d'un ballet de fleurs.

Baudelaire

1. Introduction

The standard Big Bang model provides an excellent understanding of the expanding Universe[1]. It implies that the early phase of the Universe was extremely hot and dense. Knowledge from the field of high-energy physics thus is required to describe the matter content correctly. It is widely believed that at temperatures $T \gg O(100\ \mathrm{MeV})$ there is a gas of quarks and leptons with weak interactions. For this there are some heuristic arguments[2]. 1) Hadrons may be viewed as bags of perturbative vacuum, where quarks move freely, immersed in the true non-perturbative vacuum, which has a lower energy density by some constant $B(T=0)$ and in which no free quarks propagate. One finds an energetically favoured bag size (~ 1 fermi). Clearly (?) at high temperatures the bags will touch and the quarks can propagate over large distances. For this there are two reasons. a) Thermal fluctuations will create many bags, and b) At high temperatures certain non-perturbative configurations, which at T=0 contribute to the lowering of the true vacuum energy density, are squeezed out[2], so that the bag radii $\propto B(T)^{-1/4}$ increase. 2) Quantum Chromodynamics (QCD) is the SU(3) gauge theory with quarks in the $\underline{3}$ representation. The strong interaction between nucleons is thought to be a large distance remnant of QCD. There is a very special property of QCD which contrasts with electromagnetism: at higher energies, or smaller distances, the effective coupling gets weaker (asymptotic freedom). Clearly (?) at high temperatures, when all particles have energies $\sim T$, the interactions are weak and thus the strong forces needed for confinement are absent.

But this assumed hadron-quark transition is not at all trivial, because the physics of the confinement regime is so difficult. The situation differs from the "ordinary" phase transitions resulting from the break down of the unified gauge symmetry[3], which are driven by (fundamental or composite) scalar particles. The 2-dimensional Schwinger model, where the symmetry is broken dynamically, warns us: it has *no* symmetry restoration at high temperature[4]. Confinement, also, probably is a dynamical effect of the gauge fields only. It is not at all evident *that* quark liberation occurs.

Recently new techniques for calculations in the non-perturbative regime and better theoretical understanding of quark confinement have shed their light on the question of quark liberation. This we want to discuss in the present paper.

Section 2 gives some background of quark confinement and finite temperature field theory. The clear evidence from numerical calculations of gauge theories with lattice regularization (§2.3) that quark liberation indeed occurs is presented in section 3. But this evidence is only of an experimental nature. Section 4 presents our understanding of the physical mechanisms that operate the quark liberation. The conclusions are presented in section 5. We note that in some formulae we will consider SU(N) generally instead of the SU(3) of QCD and that the numerical results presented are for SU(2). Needless to say that the covering of the references is only indicative, especially in section 2. Natural units are used so that $\hbar = c = k = 1$.

2. Background

2.1. Quark confinement

We will consider pure QCD, where the quarks act only as classical sources and not as dynamical fields, such as in vacuum polarisation. There are some arguments that confinement is only due to the gauge fields (called gluons in QCD). For example, the very fact that baryons and mesons consist of a *fixed* number of quarks and antiquarks. Also for a SU(N) , $N \to \infty$, theory the importance of quark fields relative to gauge fields goes to zero. In the vacuum polarization, for example, the contribution of a loop of fermions is negligible to that of gluons, because there are N^2-1 gauge fields and only N colours of quarks. For an introduction of why we are sure anyway that hadrons are composed of quarks see ref. (5).

The action, field strength and gauge fields are[(6)]

$$S = \int d^4x \; L(x) = -\int d^4x \; \frac{1}{2} \operatorname{tr} F_{\mu\nu} F^{\mu\nu} , \tag{1a}$$

$$F_{\mu\nu}(x) = \partial_\mu A_\nu - \partial_\nu A_\mu - ig \left[A_\mu, A_\nu\right] , \tag{1b}$$

$$A_\mu(x) = A_\mu^a(x) \; T^a , \tag{1c}$$

with T^a the group generators, $a=1..(N^2-1)$ for SU(N), and g the coupling constant. The action is invariant under *local* gauge transformations with group element $\Omega(x)$

$$A_\mu \to \Omega A_\mu \Omega^\dagger + ig^{-1} \Omega \partial_\mu \Omega^\dagger . \tag{1d}$$

For non-abelian gauge theories the non-zero commutator in (1b) gives the self interaction of the gauge fields, in contrast to the electrically neutral photon. The interaction strength $\alpha = g^2/4\pi$ at energies^2 Q^2 is determined by $d\alpha/d \ln Q^2 = \beta$. The β function can only be calculated perturbatively: $\beta = 4\pi\, b\, \alpha^2 + O(\alpha^3)$. For non-abelian SU(N) theories b is *negative* for not too many quark flavours n_q (6?), so that one finds asymptotic freedom

$$\alpha(Q^2) = \frac{12\pi}{11N-2n_q} \frac{1}{\ln(Q^2/\Lambda^2)} , \quad (\alpha << 1) \ . \tag{2}$$

Λ^2 results from the renormalization point and depends on the type of regularization used. The counter part of asymptotic freedom is called, prosaically, infrared slavery, which means that forces grow strong at large distances. This gives an intuitive idea of quark confinement, and immediately shows the necessity of non-perturbative techniques ($\alpha \gtrsim 1$). For an introduction see ref. (9), (10); but for a topological point of view of confinement see ref. (8).

We will not elaborate on the phenomenological string and bag models for hadrons[(5)], because for our problem we need to consider the basic non-perturbative QCD. For this we will discuss the Wilson criterion for confinement in pure QCD. Experimentally quark confinement is defined by the statement that the physical states are colour neutral. To pull out a coloured constituent requires infinite energy, so that the interquark potential is monotonically rising with separation. The Wilson loop is defined as[(11)]

$$W(C) \equiv \mathrm{tr}\, P \exp\left[ig \oint_C dx^\mu A_\mu(x)\right] , \tag{3}$$

with P the standard path ordering operator around curve C. The confinement criterion is

$$\langle 0|W(C)|0\rangle \propto \exp\left[-\sigma A\right] , \quad (A \to \infty) , \tag{4}$$

with A the minimal surface of C and σ a dimensional constant called the string tension. For a rectangular loop of space dimension R and time dimension T one easily sees how criterion (4) arises[(9)]. The change in the action from a heavy quark-antiquark pair created, separated to a distance R for a time T and then annihilated is just given by the logarithm of (3). For a confining potential $V \propto R$ the change in the action clearly is $\propto RT = A$, and hence we arrive at (4). One can define another loop operator, which acts as the dual of W(C), and study other phases than confinement for gauge theories in general[(12)].

2.2. Finite temperature

At thermal equilibrium the expectation value of operator O is given by

$$\langle O\rangle = Z^{-1}\ \mathrm{Tr}(O\ e^{-\beta H}) \ , \tag{5}$$

with $Z = \mathrm{Tr}\ e^{-\beta H}$ the partition function (but see below). In the path integral representation we have to integrate over (anti)periodic configurations of (fermionic) bosonic fields with period β in imaginary time[13]

$$\mathrm{Tr}\ e^{-\beta H} = Z = N(\beta) \int\limits_{\text{(anti)per}} [d\phi] \ \exp\left[-\int_0^\beta dt \int d^3x \ L_{eff}\ (\phi,\ i\dot{\phi})\right] \ , \tag{6}$$

with normalisation factor $N(\beta)$ and where ϕ denotes all sorts of fields. The first equality in (6) holds only in physical gauges, which have the correct number of degrees of freedom over which we trace. It is the trace operator that leads us to consider only (anti)periodic configurations in the path integral on the RHS of Eq. (6). L_{eff} may differ from L if the Hamiltonian density is not simply quadratic in the momenta.

The Wilson criterion is slightly modified because we now want to find the potential energy between a $q,\bar{q}$ pair averaged over a thermal ensemble. Consider for the SU(N) gauge theory

$$L(\underline{x}) \equiv N^{-1}\ \mathrm{tr}\ P\ \exp\left[i \int_0^\beta dt\ A^0\ (\underline{x},t)\right] \ , \tag{7}$$

which because of the periodicity $A^\mu(\underline{x},0) = A^\mu(\underline{x},\beta)$ is a sort of Wilson loop (but see below). The free energy relative to the vacuum $F(N_q,\ N_{\bar{q}};\ \underline{x}_1,\ ..\underline{x}_{N_q},\ \underline{y}_1,\ ..\ \underline{y}_{N_{\bar{q}}})$ of N_q quarks and $N_{\bar{q}}$ antiquarks at positions $\underline{x}_i$, $\underline{y}_i$ is[14]

$$\exp\left[-\ \beta F(N_q,N_{\bar{q}};\ ..)\right] = \langle L(\underline{x}_1)\ \dots\ L(\underline{x}_{N_q})\ L^\dagger(\underline{y}_1)\ \dots L^\dagger(\underline{y}_{N_{\bar{q}}})\rangle \ , \tag{8}$$

with <..> from path integrals as in Eq. (6). Gauge transformations which keep $A^\mu(x)$ periodic as required need not be periodic themselves

$$\Omega(\underline{x},\beta) = (c_n I)\ \Omega(\underline{x},0) \ , \tag{9}$$

with the first RHS factor in the center Z(N) of gauge group SU(N), $c_n = e^{2\pi in/N}$ for integer n. The $L(\underline{x})$ operator is not invariant under semi-periodic gauge transformations (9): $L(\underline{x}) \to c_n\ L(\underline{x})$. Hence (8) transforms by a factor $\exp[2\pi in(N_q-N_{\bar{q}})/N]$.

If the symmetry is not spontaneously broken this guarantees that $\beta F(N_q, N_{\bar{q}})$ must be infinite, unless $N_q - N_{\bar{q}} = kN$, for some integer k. For QCD (N=3) this means finite energy for baryons ($N_q=3$, $N_{\bar{q}}=0$) and mesons ($N_q=1=N_{\bar{q}}$), but not for a single quark ($N_q=1$). The RHS of (8) thus tells us at what temperatures we have confinement ($\langle L\rangle = 0$) or quark liberation ($\langle L\rangle \neq 0$). [We admit that we have been a bit cavalier about possibly divergent selfenergies in this brief discussion of the criterion on $\langle L\rangle$].

2.3. Lattice gauge theories (LGTs)

The idea is to work on a space-time lattice, where the lattice spacing a gives the momentum cut-off π/a, but to preserve gauge invariance completely. The Wilson action is[11]

$$S_W = g_0^{-2} \sum_P (\mathrm{tr}\, U_P + \mathrm{tr}\, U_P^\dagger) \tag{10}$$

$$U_P \equiv U(n)_\mu\, U(n+\hat{\mu})_\nu\, U^\dagger(n+\hat{\nu})_\mu\, U^\dagger(n)_\nu \;,$$

where the sum is over the elementary squares P, plaquettes, and $U(n)_\mu$ is a group element on the link starting at site n in direction μ ($n + \hat{\mu}$ is the site next to n in the μ direction). If we write $U(n)_\mu = \exp\left[ig_0 A(n)_\mu\right]$, we find in the limit $a \to 0$ the continuum action (1a). The action is gauge invariant under $U(n)_\mu \to \Omega(n)\, U(n)_\mu\, \Omega^\dagger(n+\hat{\mu})$. It is easy to transpose earlier continuum formulae on the lattice, e.g.

$$L(\underline{x}) = N^{-1}\, \mathrm{tr} \prod_{t=0}^{N_t} U^0(\underline{x},t) \;, \tag{11}$$

where $\beta = N_t a$.

Wilson[11] showed that for $g_0 \gg 1$ in (10) the confinement condition holds $\langle 0|W|0\rangle \propto \exp\left[-\sigma A\right]$. The major problem is to show that a small sized lattice with weak coupling leads, through many operations of thinning out of the variables, to a coarse, strong coupling theory with the Wilson action. This should happen quite independently of the details of the initial lattice. Note that the Wilson action is only one of many that give the correct continuum action. This problem can be attacked theoretically by renormalization group (RG) methods[15] or by numerical calculations, especially Monte Carlo (MC) procedures[10,16]. We will employ RG ideas in §§4.3, 4.4.

To find $\langle O\rangle$, for some operator O, MC methods avoid to calculate the many integrals over the fields, one for each degree of freedom of $U_\mu(n)$ per link. Instead they bring the system to equilibrium and then calculate $\langle O\rangle$ as the *mean* of $O(U_\mu(n))$ over a number of successive field configurations $U_\mu(n)$. The precise meaning of "equilibrium" and "successive" depends on the methods of updating the variables[16]. The results indicate that indeed the confinement at strong lattice coupling persists to weak coupling[16]. Recently spectacular results for the hadron mass spectra were obtained , where dynamical effects of the quarks were neglected[17]. With this powerful method we will now consider what happens to confinement at finite temperature.

3. Quark liberation; numerical experiments for SU(2)

In §2.2 we defined the operator $L(\underline{x})$, whose averages, as in Eq. (8), we wish to calculate numerically. These MC calculations were done in ref. (14), (18), (19) on 4-dimensional lattices with time and space sizes N_t and N_s. Confinement is destroyed for temperatures above

$$T_c(\mathrm{MC}) \sim 0.35 - 0.55\ \sigma^{\frac{1}{2}} \sim 150 - 220 \text{ MeV} , \tag{12}$$

with σ the zero temperature string tension. The true value probably is at the high end of the range, where the specific heat peaks[19]. At T_c there appears to be a second order phase transition. For this there are two arguments: the change in the order parameter $\langle L\rangle$ is smooth (fig. 1) and for temperatures near T_c large fluctuations occur during the averaging of the Monte Carlo. The physical temperature is $\beta^{-1} = (N_t a)^{-1}$. Renormalization gives the relation between the lattice spacing a and the bare coupling constant g_0: $a^2 = \sigma^{-1} f(g_0)$, where $f(g_0)$ is known for the perturbative regime and can elsewhere be calculated from MCs. It follows that the physical β is a monotonically rising function of the bare coupling g_0. This gives the physical interpretation of figs. 1 and 2. At $T \to \infty$ we *exactly* know the energy density, which is given by the Stefan-Boltzmann law up to corrections in powers of $\alpha_s(T)$. The MC results agree perfectly[19].

In principle $\langle O\rangle$ can be calculated from MCs for every operator O. Fig. 2 shows the change of the interquark potential (F(1,1) of Eq.(8)) above and below T_c. It is remarkable that the string tension below T_c rapidly approaches the $T = 0$ value $\sim(400\text{ MeV})^2$: $\sigma(T = 0.98\ T_c) = 0.4\ \sigma(T = 0)$ already[14].

For $T > T_c$ also the electric[19] and magnetic[20] screening has been calculated and found to agree with the theoretical predictions[2,21].

To conclude we may say that numerical experiments clearly show quark liberation

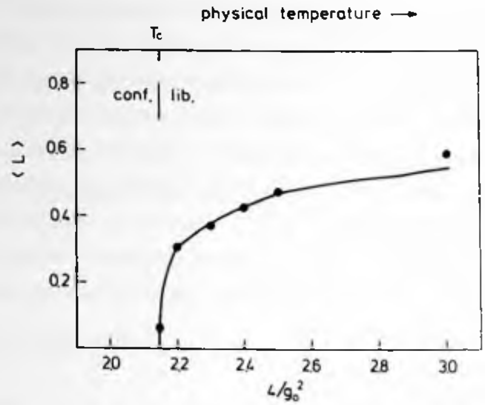

Fig. 1. Monte Carlo results from ref. (14). <L> is the expectation value of the Wilson line (11) and g_0 is the bare lattice coupling constant. The lattice had N_t=3, N_s=7 and the gauge group was SU(2). For <L> = 0 or ≠ 0 there are confined or liberated quarks. Also plotted the curve <L> = const $(4/g_0^2 - 4/g_0^{c2})^\beta$, with β = 0.21 and $4/g_0^{c2}$ = 2.15. The figure also represents <L> vs the physical temperature (sect. 3).

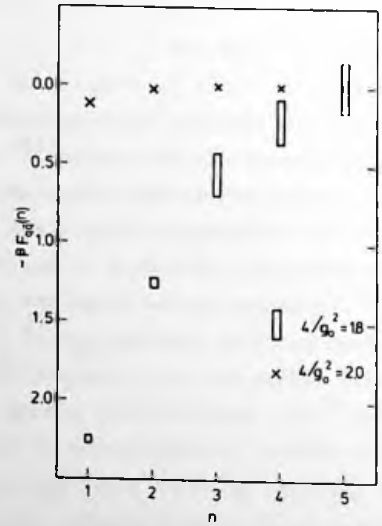

Fig. 2. Monte Carlo results for SU(2) from ref. (14) of the quark-antiquark potential vs lattice separation n. The lattice had N_t=2, N_s=10 and its critical coupling is $4/g_0^{c2}$ = 1.85. Note that the physical temperature is a monotonic function of $4/g_0^2$. Thus at a temperature below T_c we find a linear confining potential, whereas above T_c we have a screened Coulomb potential.

for temperatures above the value of Eq. (12). We now want to *understand* what happens.

4. Quark liberation: theory

4.1. Strong coupling LGT

In §2.3 we presented the LGT in a Lagrangian form. One may also use a Hamiltonian, which consists of two terms: 1) the sum over the links of the generalised momenta^2 with a factor g_0^2, and 2) the sum over a plaquette term with a factor g_0^{-2}. For strong coupling one only considers the first,electric, term. In this case Susskind[22] showed that the partition function of an Abelian LGT is similar to that of the planar Heisenberg model in the Villain approximation, but with the role of the physical temperature *inversed*. Because the phases of the Heisenberg magnet are known one finds that also the LGT has a phase transition from confinement at low temperatures to a Coulomb gas at high temperatures. For non-abelian LGTs, e.g. QCD on the lattice, the Coulomb interaction is screened by the "charged" gauge bosons, analogously to Debije screening by electrons in a plasma. For the transition temperature one finds[2,22]

$$T_c\ (\mathrm{LGT}) \lesssim \sigma a/\ln 5. \qquad (1\cdot$$

Inclusion of the other term in the Hamiltonian only reduces the confinement properties somewhat and thus the strong coupling estimate is an upper boundar The physical picture is that for $T \geq T_c$ the electric strings condensate, so tl the separation of a $q,\bar{q}$ pair to infinity, i.e. the insertion of one more string requires only a *finite* amount of energy and quarks no longer are confined.

4.2. Z(N) symmetry

In §2.2 we considered the transformation properties of the Wilson line $L(\underline{x})$ under the center Z(N) of the gauge group SU(N). Expectation values of products of $L(\underline{x}_i)$ and $L^\dagger(\underline{y}_i)$ tell us about the phase we are in. We have the following scenario[23], where one chooses a gauge so that U^0 is not 1 only on a spatial slice. At very high temperatures $\langle L\rangle$ is uniformly non-zero, but when the temperature is lowered some islands of flipped Z(N) spin appear and they condense at $T \leq T_c$, thereby restoring the symmetry $\langle L\rangle = 0$. Weiss[23] has calculated the 1-loop effective potential $V_1(L)$ for the order parameter L and indeed

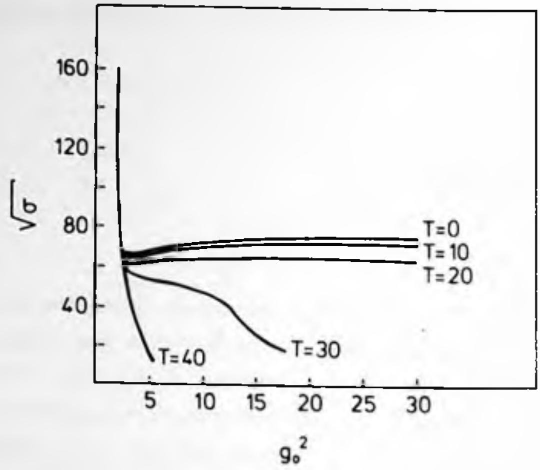

Fig. 3. Calculated string tension σ vs. the lattice coupling g_0 from the semi-classical effective theory of ref. (2). Both $\sigma^{1/2}$ and T are in units of $\Lambda_{lattice}$. For $T \gtrsim 30\ \Lambda_{lattice}$ the instanton contribution disappears and confinement no longer occurs.

finds spontaneous symmetry breaking; the minimum of V_1 lies at $L_{min} \neq 0$. This calculation will be essentially correct at high temperature, weak coupling. But at lower temperature higher orders and topological effects, notably the islands of flipped spin, will determine the behaviour. It is to be remarked that the barrier between the minima of V_1 shrinks for decreasing temperature. Hence it is plausible that the Z(N) symmetry is restored below a certain critical temperature, so that confinement occurs.

4.3. Semiclassical effective theory

The Weinberg-Salam model gives at energies << M_W = O(100 GeV) the familiar Fermi theory of β-decay, its coupling constant being determined by the parameters of the fundamental theory $G_F = \text{const}\ g^2 M_W^{-2}$. Similarly we wish to find an effective theory of QCD on the scales of interest for confinement, $\lesssim$ O(100 MeV). Here we discuss a semiclassical calculation(2): 1) integrate out the degrees of freedom on length scales $\leq$ a; 2) assume the effective action to have the Wilson form (10), we seek to determine the a dependence of the coupling constant g_0; 3) there will be contributions from ordinary perturbative fluctuations and from finite temperature instantons with sizes < a. Fig. 3 gives the resulting string tension.

The $T = 0$ result with the sharp change at $g_0^2 \sim 2$ is in good agreement with the numerical results(16). Note that in that regime the expansion parameter $g_0^2/8\pi^2$ is quite small, which justifies a posteriori the use of semiclassical methods. For temperatures above 35 Λ_{lattice} the string tension no longer is constant: the instantons are squeezed out because of the small time dimension β (§2.2). Gross, Pisarski and Yaffe find a critical temperature of

$$T_c \text{ (semi class. eff. th)} \sim 0.5\, \sigma^{1/2} , \tag{14}$$

both for SU(2) and SU(3).

4.4. Dielectric effective theory

Define a colour dielectric constant $\varepsilon(L)$ so that the effective coupling distances L is $g_L^2 = g^2/\varepsilon(L)$, with g^2 some coupling constant renormalized in the perturbative regime. Quark confinement may be described phenomenologically by having $\varepsilon(L) = 0$ for L larger than a typical hadron scale. The resulting bag model, a region of $\varepsilon \sim 1$ immersed in the $\varepsilon = 0$ vacuum, gives some impressive results(24).

Recently Nielsen and Patkos(25) showed how a dielectric theory may result from QCD. They consider the following average over curves C, running from x_0 to $x_0 + \varepsilon$, inside a 4-d box of dimension L_0 (the gauge group is SU(2))

$$\overline{U}(x_0 + \varepsilon, x_0; L_0) \equiv \underset{C}{\text{av}}\; P \exp\Big[i \int_{x_0}^{x_0+\varepsilon} A_\mu\, dx^\mu\Big] , \tag{15a}$$

which can be developped in the small parameter ε^μ

$$\overline{U} = K(L_0; x_0) + i\, \varepsilon^\mu B_\mu(x_0) \tag{15b}$$

$$K(L_0; x_0) \equiv \underset{C}{\text{av}}\; P \exp\Big[i \oint_C A_\mu\, dx^\mu\Big] . \tag{15c}$$

For small L_0 they derive an effective potential $V_1(s; L_0)$, where $s \equiv \frac{1}{2}$ tr K. This thinning out procedure should be continued to $L_n \geq$ a typical hadron scale. *Assume* that the final effective potential $V_*(s)$ has the same form as V_1, but with L_0 replaced by a value L_* such that $s_{min} = 0$; $V_*(s) = V_1(s; L^*)$. Because the effective fermion term in the Lagrange density probably is $i\, \overline{\psi}(s\, \partial_\mu - B_\mu)\gamma^\mu \psi$, we have a dielectric theory with no quark propagation over large distances. Klinkhamer[26] considered the finite temperature effects and showed quark liberation to occur above

$$T_c(\text{diel. eff. th}) = L_*^{-1} \sim 0.48\ \sigma^{1/2}\ . \tag{16}$$

The origin of the transition is that finite temperature squeezes the boxes to βL_n^3, which gives a strong reduction in the curve average (15c) and forces the minimum of $V_{1,\beta}(s; L_n)$ away from the confinement value 0 towards the perturbative value 1.

5. Conclusion

Monte Carlo calculations of lattice gauge theories clearly show that at high temperature quark liberation does occur. The best numerical value[19] for the transition temperature is $T_c(MC) \sim 0.56\ \sigma^{\frac{1}{2}} \sim 220$ MeV. This result is for a pure SU(2) gauge theory but we expect no basic difference for real QCD. In section 4 we gave several explanations of this phenomenon. Also the theoretical estimates of T_c of Eqs. (13), (14), (16) are in good agreement. The situation thus is quite satisfactory. The major problem is at $T = 0$, namely to show rigorously that the perturbative theory with asymptotic freedom is linked to the strong coupling LGT with quark confinement. From that investigation one could also determine the nature of the transition at T_c, which the MC results indicate to be a second order phase transition.

Let us briefly return to cosmology. For $T > 200$ MeV the matter content of the Universe indeed is a quark-lepton gas. Baryonnumber creation and changes of gauge symmetries can be calculated for this gas[3]. The quark-hadron transition probably left no direct remnants in the present Universe. The density fluctuations that might result from a second, or first, order phase transition probably are negligible for galaxy formation, which seeks an explanation for the origin of the initial density perturbations. This is because of the scales involved ($\sigma^{1/2}$ and the horizon at T_c). Still it is extremely interesting to understand the quark liberation problem, perhaps mostly because of the implications for the nature of quark confinement itself.

Acknowledgements

Learning about confinement started during a visit at the Astronomy and Astrophysics Center of the University of Chicago, with support from the grants NSF AST 78-20402 and DOE-DE-AC 02-80 ER10773. I thank O. Napoly for reading through the manuscript and the organizers for this interesting symposium.

References

(1) S. Weinberg, Gravitation and Cosmology (Wiley, 1972)

(2) D.J. Gross, R.D. Pisarski, L.G. Yaffe, Rev. Mod. Phys. 53 (1981), 43

(3) A.D. Linde. Rep. Prog. Phys. 42 (1979), 389; F.R. Klinkhamer, in A.W. Wolfendale (ed.) Progress in Cosmology (Reidel, 1982)

(4) L. Dolan, R. Jackiw, Phys. Rev. D9 (1974) 3320 [We admit that the analogy is rather weak between the Schwinger model and QCD with respect to symmetry breaking and confinement, which are, respectively, UV and IR effects. Temperature is expected to change the IR behaviour only].

(5) C. Quigg, in Gauge Theories, Les Houches 1981 (preprint Fermilab-Conf-81/78-Thy)

(6) E.S. Abers, B.W. Lee, Phys. Rep. 9C (1973), 1

(7) H.D. Politzer, Phys. Rep. 14C (1974), 129

(8) A.M. Polyakov, Nucl. Phys. B120 (1977), 429; G. 't Hooft, Nucl. Phys. B190 (FS3) (1981), 455

(9) S. Mandelstam, Phys. Rep. 67C (1980), 109

(10) P. Hasenfratz, in W. Pfeil (ed.), 1981 Int. Symp. on Lepton and Photon Interactions at High Energies

(11) K.G. Wilson, Phys. Rev. D10 (1974), 2445

(12) G. 't Hooft, Nucl. Phys. B138 (1978), 1; B153 (1979), 141

(13) C.W. Bernard, Phys. Rev. D9 (1974), 3312

(14) L.D. McLerran, B. Svetitsky, Phys. Rev. D24 (1981), 450

(15) L.P. Kadanov, Rev. Mod. Phys. 49 (1977), 267; and ref. therein

(16) M. Creutz, in XVIIIth Karpacz Winter School of Theoretical Physics, Febr. 1981; G. Parisi, in XX Int. Conf. on High Energy Physics, Madison 1980 (AI Inst. Phys., 1981)

(17) H. Hamber, G. Parisi, Phys. Rev. Let. 47 (1982), 1792; E. Marinari, G. Parisi, C. Rebbi, Phys. Rev. Let. 47 (1982), 1795; H. Hamber, E. Marinari, G. Parisi, C. Rebbi, Phys. Lett. 108B (1982), 314

(18) J. Kuti, J. Polonyi, K. Szlachanyi, Phys. Lett. 98B (1981) 199

(19) J. Engels, F. Karsch, H. Satz, I. Montvay, Phys. Lett. 101B (1981), 89; Nucl.Phys. B205 (FS5) (1982),545

(20) A.Billoire,G.Lazarides and Q. Shafi, Phys. Lett. 103B (1981), 450; T.A. de Grand, D. Toussaint, Phys. Rev. D25 (1982), 526

(21) K. Kajantie, J. Kapusta, Phys.Lett.110B (1982),299

(22) L. Susskind, Phys. Rev. D20 (1979),261 ; see also A.M. Polyakov, Phys. Lett. 72B (1978), 477

(23) N. Weiss, UBC preprint Sept. 1980; Phys. Rev. D24 (1981), 475

(24) T.D. Lee, Particle Physics and Introduction to Field Theory (Harwood Ac. Publ. 1981), ch. 17, 20

(25) H.B. Nielsen, A. Patkos, Nucl. Phys. B195 (1982), 137

(26) F.R. Klinkhamer, subm. Phys. Lett. B (preprint 9 March 1982) = Chapter 7

(27) H. Kluberg-Stern, A. Morel, O. Napoly, B. Petersson, Nucl. Phys. B190 [FS3] (1981), 504; N. Kawamoto, J. Smit, Nucl. Phys. B192 (1981), 100

(28) R.D. Pisarski, Phys. Lett. 110B (1982), 155

(29) J. Engels, F. Karsch, H. Satz, Phys.Lett. 113B (1982),398

Addendum

There has been progress recently on Monte Carlo (MC) simulations with fermions[17]. At finite temperature there are two important problems to address: 1) is the effect of dynamical quarks on (de)confinement indeed small (cf. §2.1)?, and 2) is the chiral symmetry, which at T=0 is broken with the pions as Goldstone bosons, restored at high temperature? In this addendum we will briefly discuss the second question.

In ref. (27) the spontaneous breaking of the chiral symmetry, $\langle 0|\bar{\psi}\psi|0\rangle \neq 0$, is calculated using lattice regularisation and various other approximations. The symmetry restoration at a temperature T_{ch} is discussed phenomenologically by Pisarski[28], who gives some arguments for the following relations to hold: $T_{ch} \geq T_c$ and $T_{ch} - T_c \ll T_c$, where T_c is the temperature of quark liberation. Engels, Karsch and Satz[29] have done MC simulations. The fermion part of the action used is

$$S_W^F = \sum_n \{\bar{\psi}_n\psi_n - \sum_{\mu=0}^{3} K_\mu [\bar{\psi}_n(1 - \gamma_\mu)U_{n,n+\hat{\mu}}\,\psi_{n+\hat{\mu}}$$

$$+ \bar{\psi}_{n-\hat{\mu}}\,(1 + \gamma_\mu)U^{\dagger}_{n-\hat{\mu},n}\,\psi_n]\} , \tag{17}$$

where colour and spinor indices are suppressed, n denotes a lattice site, the link variable $U_{n,m}$ runs between sites n and m, and the "hopping" parameters K_μ may differ in spacelike and timelike directions. In the partition function the ψ and $\bar{\psi}$ fields can be integrated out analytically. From the resulting determinant *only* the leading term in the hopping parameter expansion is retained. The Wilson formulation of fermions on the lattice, Eq. (17), does not lead to species doubling, but breaks the chiral invariance *explicitly*. Information about the physical chiral symmetry breaking and its restoration at T_{ch} can be obtained by considering the difference $\langle\bar{\psi}\psi\rangle_{SB} - \langle\bar{\psi}\psi\rangle$, where $\langle\bar{\psi}\psi\rangle_{SB}$ is the average for an ideal gas of massless fermions (Stefan-Boltzmann). Fig. 4 presents the energy density ε, for the pure Yang-Mills system (YM) and with two flavours of fermions, and the measure for chiral symmetry breaking, both as a function of temperature. Two conclusions may be drawn, partially answering the two questions above: 1) the inclusion of quarks does not change the liberation temperature T_c much, and 2) the temperature T_{ch} of chiral symmetry restoration is somewhat higher than T_c. The values for SU(3) and 2 quark flavours are

$$T_c(\mathrm{MC}) \sim 80\ \Lambda_{\mathrm{lattice}} \sim 0.40\ \sigma^{1/2} \sim 160\ \mathrm{MeV}$$

$$T_{ch}(\mathrm{MC}) \sim 100\ \Lambda_{\mathrm{lattice}} \sim 210\ \mathrm{MeV}\ , \tag{18}$$

where we used $\sigma(T=0) = (400\ \mathrm{MeV})^2$. These preliminary results are encouraging. It will be interesting to calculate the change in T_c when virtual quark loops are included and in T_{ch} when other actions than Eq. (17) are used. Also it should be checked that the measure for chiral symmetry used, $\langle\bar{\psi}\psi\rangle_{SB} - \langle\bar{\psi}\psi\rangle$, is correct at the low temperatures of order $100\ \Lambda_{\mathrm{lattice}}$, where g_0^2 is not very small.

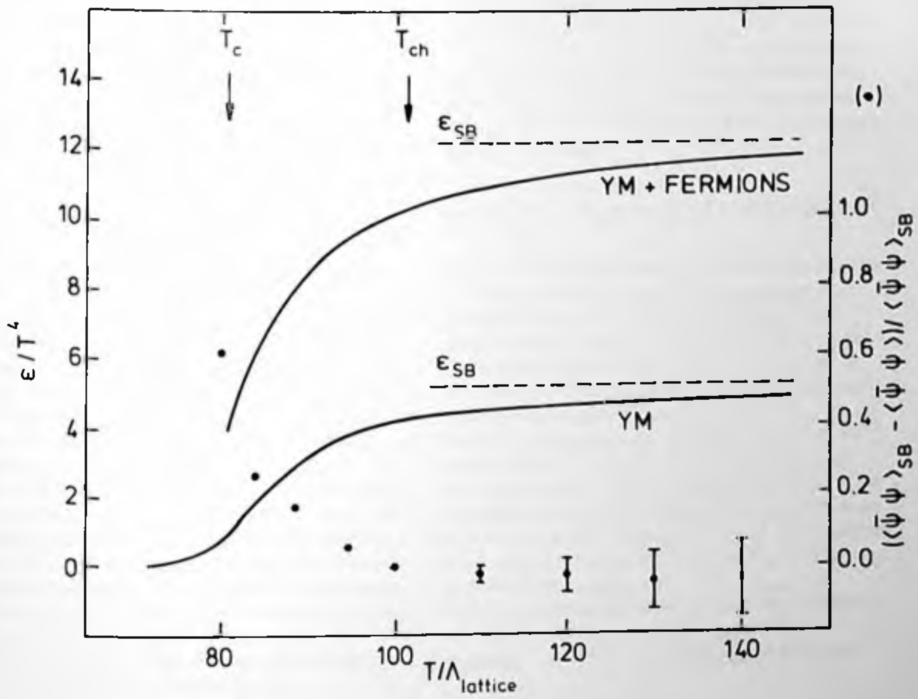

Fig. 4. Monte Carlo results for SU(3) and two flavours of quarks from ref. (29). ε is the energy density (curves) and the dots refer to the chiral symmetry breaking measure. YM denotes the pure gauge theory and SB the Stefan-Boltzmann value of ε. The curves show quark liberation above T_c and the dots chiral symmetry restoration above T_{ch}. [warning: curve values are only indicative].

CHAPTER 7

ON THE HADRON–QUARK TRANSITION AT HIGH TEMPERATURE

F.R. KLINKHAMER
Leiden Observatory, PO Box 9513, 2300 RA Leiden, The Netherlands

Received 16 March 1982
Revised manuscript received 7 July 1982

We discuss quark liberation from the point of view of the effective dielectric theory of Nielsen and Patkos and calculate the transition temperature.

Some heuristic arguments, based for example on the bag model for hadrons and the phenomenon of asymptotic freedom of QCD tell us that at temperatures much higher than typical hadron or QCD scales quarks no longer are confined [1]. Recently this has been confirmed by Monte Carlo calculations of lattice gauge theories (LGTs), which clearly show the transition from the confinement phase to the (screened) Coulomb phase. For SU(2) the transition temperature is $T_c \sim 0.35 - 0.50\,\epsilon^{1/2} \sim 150$–200 MeV, with ϵ the zero temperature string tension [2,3]. Because of the smoothness in the order parameter ($\langle L \rangle$, see below) and the large fluctuations around T_c, this probably is a second order phase transition. Along with these brute force results we would like to understand the *mechanism* of the highly non-trivial phenomenon of quark liberation. For strong coupling LGTs one indeed finds a transition [4], with as physical picture [1] the condensation of electric strings at $T \geqslant T_c$, so that it costs little to insert one more quark–string–antiquark (entropy versus energy). The second explanation focusses on the global Z(N) symmetry, in case of an SU(N) gauge theory [5]. Consider $L(\boldsymbol{x}) \equiv N^{-1}$ tr $P \times \exp[\mathrm{i} \int \mathrm{d}t\, A^0(\boldsymbol{x}, t)]$, with the integral over $[0,\beta]$ and $A^\mu(\boldsymbol{x}, t)$ Lie algebra valued. The expectation value $\langle L \rangle$ acts as an order parameter, because L is not invariant under Z(N). At high temperatures there is spontaneous symmetry breaking $\langle L \rangle \neq 0$, but when the temperature is lowered islands of flipped Z(N) spins occur and condense at $T \leqslant T_c$, restoring the symmetry $\langle L \rangle = 0$. It is easy to see that these phases correspond, respectively, to quark liberation and confinement [3].

In this letter we will take the point of view of the perfect dielectric theory [6], but as an *effective* theory resulting from QCD on large distances [7]; cf. ref. [1]. Nielsen and Patkos [7] derive for small L_0 an effective potential $V_1(\sigma)$, which has $\sigma_{\mathrm{minimum}} < 1$ {put in (3): $n = 0, f = 1, \beta = L_0$}. Here and in the following we consider a pure SU(2) gauge theory. The procedure was to consider an average over all curves, nearly closing on x_0, inside a box of volume L_0^4 of the following integral and to develop it

$$\underset{C}{\mathrm{av}}\, P \exp\left(\mathrm{i} \int_{x_0}^{x_0+\epsilon} A_\mu(x)\, \mathrm{d}x^\mu\right) = K + \mathrm{i}\epsilon^\mu B_\mu(x_0), \qquad (1)$$

with K the closed-path integrals ($\epsilon = 0$) of the lhs. This procedure could be followed for gauge groups SU(N) in general, but for $N = 2$ K as given by eq. (2) is a gauge invariant Lorentz scalar and the effective theory for its trace $\sigma \equiv \frac{1}{2}\mathrm{tr}\, K$ will be considered. B_μ then is the effective field and σ should describe the influence of the non-perturbative vacuum on scales $\sim L_0$; the effective fermion term in the lagrangian density should then be $\mathrm{i}\bar{\psi}(\sigma\, \partial_\mu - B_\mu)\gamma^\mu \psi$. Let the final potential after many steps of this block-spin procedure be $V_*(\sigma)$ and assume that (1) the minimum is at $\sigma_{\mathrm{min}} = 0$, in order to have a perfect dielectric theory of quark confinement, and (2) it has the same form as $V_1(\sigma)$, only now a special value L_* replacing L_0 so that $\sigma_{\mathrm{min}} = 0$. Nielsen and Patkos then calculate the string solution and find (for $\alpha = 4$, see below): $\Lambda_{\mathrm{MOM}} = 0.70\,\epsilon^{1/2}$, $M_{\mathrm{glueball}} = 1.15\,\epsilon^{1/2}$, $\langle \mathrm{vac}| g^2 G^a_{\mu\nu} G^a_{\mu\nu} |\mathrm{vac}\rangle = 10.38$

$\times \Lambda_{MOM}^4$. These reasonable results will become impressive once it is shown, probably using renormalization group methods, that the fixed point indeed fulfills the assumption (2). Another open problem, of course, is the inclusion of dynamical fermions. We will assume that the above picture contains some truth (hence the somewhat lengthy summary) and will consider its claims at finite temperature. Eqs. (x) of ref. [7] will be referred to as (NPx) and natural units are used $\hbar = c = k = 1$.

Field theory at equilibrium at finite temperature $T = \beta^{-1}$ is described by integrating in the path integrals of the generator of Green's functions over (anti)periodic configurations of the (fermionic) bosonic fields, with period β in imaginary time [8]. As long as $L_0 \ll \beta$ the effective theory after averaging over boxes L_0^4 will differ little from that of ref. [7] and if after many iterations an $L_* < \beta$ is reached we will still have confinement at this temperature. But if β is so low that $L_* > \beta$, confinement might not be reached, because in the successive steps of block-spinning (L_n) finite temperature effects (see below) become important before we reach the L_* scale with the assumed confinement.

In such an intermediate step ($L_n > \beta$) we consider βL_n^3 boxes instead of L_n^4 ones. Let us do a near local analysis, i.e., in K Taylor expanding both $A_\mu(x)$ on the curve in terms of $A_\mu(x_0)$ and the exponential. One finds (cf. NP2.11)

$$K(x_0) \sim 1 - (g^2/48) G_{\mu\nu}(x_0) G_{\mu\nu}(x_0) L_n^4 \alpha(\beta/L_n), \quad (2)$$

with $G_{\mu\nu}$ the Lie algebra valued field strength. The factor αL_n^4 comes from the curve average of $\oint \mathrm{d}x_2^{[\mu_2} x_2^{\rho_2]} \times \oint \mathrm{d}x_1^{[\mu_1} x_1^{\rho_1]}$ (both integrals over the same curve) which is proportional to the average of the squared area of non-intersecting minimal surfaces $\overline{S^2}$ (NP2.9). For a square box $\overline{S^2} = \alpha(1) L_n^4$, for some constant $\alpha(1)$. The numerical calculations of ref. [7] give us the normalization $\alpha(1) = 4$. But for a flattened box the constant $\alpha(\beta/L_n)$ will be a monotonically decreasing function for decreasing argument (see note 1 at the end of the paper). Using (2) as a constraint and integrating the A_μ fields out one expects an effective action for σ (cf. NP3.15, 3.17), with (see note 2 at the end of the paper)

$$V_1(\sigma, \beta) \sim L_n^{-4} \{ [22/\pi^2 \alpha(\beta/L_n)] \, \sigma \ln(L_n^2 \Lambda_{\text{box}}^2) - \tfrac{5}{2}(L_n/\beta) f \ln(1-\sigma) \} \quad (3)$$

as the high temperature effective potential over distances $\sim L_n$. $\Lambda_{\text{box}}^2 \equiv \Lambda_M^2 \exp[\psi(\tfrac{1}{2})]\, 4/\pi$ and Λ_M^2 results from minimal subtraction. With the strong decrease of $\alpha(\beta/L_n)$, see note 1, the minimum is shifted towards the value 1, i.e., away from confinement. Under the assumption (2) that the form of the effective potential remains unchanged, we have two possibilities. (1) $\beta > L_*$: For all $L_n \leqslant L_*$ we use the unmodified NP potential [put $\beta = L_n$ in (3)] and we arrive at a dielectric theory of confinement at these low temperatures. (2) $\beta < L_*$: For $L_1, \ldots, L_M = \beta$ we can use the unmodified potential, but for $L_{M+n} > \beta$ we have (3). Its minimum would be at $\sigma_{\min} = 0$ for an $\bar{L}_*$ given by

$$\bar{L}_* \Lambda_{\text{box}} = \exp(-20\pi^2 f \beta / 88 \bar{L}_*), \quad (4)$$

where we used $\alpha(x) = 4x^2$ ($x < 1$). But (4) has *no* solution for $\bar{L}_* > 0$ and $\beta \neq 0$. In other words the minimum of the effective potential can never be at the confinement value 0. We have no confinement for these high temperatures. Thus, we estimate the transition temperature

$$T_c \sim L_*^{-1} = \Lambda_{\text{box}} \exp[5\pi^2 \alpha(1)/88] = 0.48\, \epsilon^{1/2}, \quad (5)$$

where we used the zero temperature results of ref. [7]. Independent of this somewhat simplified discussion of the expected renormalization group behaviour we can derive an upper bound on T_c. Strong coupling LGT gives an estimate $T_c \sim \epsilon a / \ln 5$, with a the lattice spacing [1,4]. The effective zero temperature lagrangian [7] is an LGT with $a = L_*$, so that

$$T_c \lesssim \epsilon L_* / \ln 5 \sim 1.32\, \epsilon^{1/2}. \quad (6)$$

The inequality sign in (6) might result from the neglected magnetic terms in the hamiltonian if the bare coupling $g_0^2 < \infty$ [4], which will be relevant in our case where $g_0^2(L_*)/4\pi = 0.38$.

To summarize we may say that, if the Nielsen–Patkos approach proves to be correct, its remarkable numerical results are augmented by the hadron–quark transition temperature (5), which is in good agreement with the Monte Carlo results [2,3]. Also the reader is invited to compare our work with that of ref. [9], where the effective potential at finite temperature is calculated with a constant colour magnetic field background.

This work was started during a visit at the

Astronomy and Astrophysics Center of the University of Chicago with support from the grants NSF AST 78-20402 and DOE-DE-AC 02-80 ER 10773. Some informative discussions on LGTs with S. Wadia are acknowledged. Finally I thank the referee for urging me towards a somewhat clearer presentation of note 2 and for bringing ref. [9] to my attention.

Note 1. We will not elaborate on the precise form of $\alpha(\beta/L_n)$, but give a geometrical argument for its behaviour.

In a final analysis a scheme should be used which maximizes α, in order to have small L_*, cf. (5). Note that $\overline{S^2}$ can be viewed as an average over flat surfaces [2] of varying sizes that range through the box. A two-dimensional analogue would be:

$$\overline{l^2} = \int_0^{\pi/2} \mathrm{d}\theta\, l(\theta)^2 = \widetilde{\alpha}(L_1/L_2)L_2^2,$$

for an L_1L_2 rectangle (corners at x, y coordinates 0, 0 and L_2, L_1) and with $l(\theta)$ the length from the coordinate center to the boundary of the box at polar angle θ. We find $\overline{l^2} \propto L_1L_2 = (L_1/L_2)L_2^2$. Consider now $\overline{l^2}$ over a cube. This is $\overline{l^2}$ over a face, which fans through the cube, and the result is a factor γ times the two-dimensional result. Thus $\overline{l^2}$ over a three-dimensional box lL^2 again is proportional to $(l/L)L^2$. Similarly one finds for $\overline{S^2}$ averaged over a four-dimensional box βL^3 the dependence $(\beta/L)^2L^4$, if $\overline{S^2}$ over a rectangle βL is $\propto\beta^2L^2$, in analogy to $\overline{l^2}$.

Note 2. We indicate the changes expected for finite temperature compared to the derivation of $V(\sigma; T = 0)$ of ref. [7]. Two general remarks: firstly, space–time integrals in the euclidean actions should run over $R^3 \times [0, \beta]$ and the field configurations are periodic $A^\mu(x, 0) = A^\mu(x, \beta)$, and secondly, it again seems reasonable to assume that $\sigma(x_0)$ is only influenced by the low frequency part of the gauge fields, $|k| < \pi/L_n$ (cf. NP3.3).

The effective potential after integrating out the gauge fields under the constraint from the trace of (2) has two parts. (1) The high frequency contribution, which in one loop gives the linear term in (3), is expected to have the same form as (NP3.9). In its derivation one should use now a homogeneously magnetized box of volume βL_n^3, but this will give no difference for the linear part of V because the space–time integral in (NP3.9) also is restricted (see above). Only α will now be reduced because of the flattened box [see note (1)]. (2) For the low frequency contribution (NP3.10), which gives the $\ln(1-\sigma)$ term in (3), also a βL_n^3 box should be used. This gives an extra factor L_n/β multiplying $\ln(1-\sigma)$, cf. (NP3.14). The other change would be that, because of the periodicity condition on the gauge field configurations, in the estimate of the entropy part the dimensionality essentially is reduced by one, which should give another factor $f < 1$. Admittedly this estimate of the low frequency part of $V(\sigma, \beta)$ is rather crude. Note that the precise value of f is not relevant for our purposes: (4) has no solution for any $0 < f < 1$, so that β_c is given by the L_* from the zero temperature theory. (Of course, our approximation of the renormalization group behaviour is very simplified, if not naive.)

To summarize, the high temperature ($\beta < L_n$) effective potential is expected to have the same dependence on σ as at zero temperature, except for the three different factors as given in (3).

References

[1] A review: D.J. Gross, R.D. Pisarski and L.G. Yaffe, Rev. Mod. Phys. 53 (1981) 43.

[2] J. Kuti, J. Polonyi and K. Szlachanyi, Phys. Lett. 98B (1981) 199;
J. Engels, F. Karsch, H. Satz and I. Montvay, Phys. Lett. 101B (1981) 89.

[3] L.D. McLerran and B. Svetitsky, Phys. Rev. D24 (1981) 450.

[4] L. Susskind, Phys. Rev. D20 (1979) 2610;
A.M. Polyakov, Phys. Lett. 72B (1978) 477.

[5] N. Weiss, Phys. Rev. D24 (1981) 475; UBC preprint (September 1980).

[6] An introduction: T.D. Lee, Particle physics and introduction to field theory (Harwood Academic, 1981) Ch. 20.

[7] H.B. Nielsen and A. Patkos, Nucl. Phys. B195 (1982) 137.

[8] C. Bernard, Phys. Rev. D9 (1974) 3312.

[9] B. Müller and J. Rafelski, Phys. Lett. 101B (1981) 111;
N. Ninomiya and N. Sakai, Nucl. Phys. B190[FS3] (1981) 316.

Quark Deconfinement at High Temperature and Thick Vortices

F.R. Klinkhamer
Leiden Observatory P.O. Box 9513, NL-2300 RA Leiden, The Netherlands

Received 24 August 1982

Abstract. First we describe Mack's effective $Z(2)$ theory of quark confinement in the $SU(2)$ lattice gauge theory at zero temperature. Then we show how quarks get liberated above a critical temperature, which has a numerical value somewhat below the glueball mass (in natural units).

Hadrons are composed of quarks, which have a new type of charge, called colour (red, yellow, blue say). It is believed that the interaction between quarks, apart from the usual electric and weak interactions, is described by the non-Abelian $SU(3)$ gauge theory, called Quantum Chromodynamics (QCD). One observes that quarks are confined, which means that the physical states carry no net colour (just as white is the sum of the three basic colours). The major task of the last years has been to establish that QCD is in the confinement phase. Three introductory reviews on quarks, gauge theories and confinement are given in [1].

If we consider gauge fields only the condition for confinement is that the vacuum expectation value of the Wilson loop $W(L)$, see Sect. 1 for its definition, decreases rapidly as $\exp(-\sigma \times \text{minimum area of } L)$ for large loops L[2]. σ is a constant, the (zero temperature) string tension. Let us sketch the physics behind this condition. For a rectangular loop with size T in the time direction and $R(\ll T)$ in a spatial direction the exponent is the change in the action from a quark–antiquark pair separated at a distance R and created and annihilated over a time T. The exponent will then be $-TV(R)$, with $V(R)$ the minimum energy of a static quark–antiquark pair. For a linear potential $V(R) = \sigma R$ the Wilson condition is fulfilled. Because of the constant tension (experimentally $\sigma \sim (400\,\text{MeV})^2$) it would cost an infinite amount of energy to completely separate a quark–antiquark pair, hence no naked quarks or anti-quarks can exist. A major theoretical tool is to consider the gauge theory on a discretized 4 dimensional space-time [2]. A hypercubic lattice is used, which consists of sites x, links l, elementary squares called plaquettes p, and cubes c. The gauge fields $U(l)$, which are group elements, live on the links, whereas quarks $\psi(x)$ sit on the sites. Lorentz invariance, or better Euclidean invariance since we rotate the time axis to imaginary values so that we have a Euclidean space instead of the usual Minkowsky space, has been sacrificed momentarily, but gauge invariance is preserved exactly (see Sect. 1). Lattice gauge theories (LGTs) are, of course, ideally suited for numerical calculations.

In this paper we will consider the destruction of the quark confinement property by high temperatures. We will not discuss the applications of this phenomenon, notably for the theory of the early Universe or collisions of heavy nuclei in accelerator experiments. We "just" want to have a picture of the physical mechanism that operates the change of phase.

Computer simulations of LGTs show quark liberation above a critical temperature $T_c = 40(\pm 5)\,\Lambda_{\text{lattice}} \sim 0.44\,\sigma^{1/2} \sim 200$ MeV [3]. Here Λ_{lattice} is an energy scale, which arises in the continuum limit of the lattice spacing $a \to 0$ from a special dependence of the bare coupling $g_0(a)$ (asymptotic freedom, see (11)) so that physical masses and other quantities remain constant. We used the relation $\Lambda_{\text{lattice}} = 0.011\,\sigma^{1/2}$ as determined in [4]. This value of T_c is for the pure $SU(2)$ LGT and follows from expectation values of a sort of Wilson loop. For $SU(3)$ the result is roughly the same. Simulations including dynamical quarks are being started, but are still at a rather crude level of sophistication. The deconfinement temperature T_c remains approximately the same and for $T \geqq T_{ch} = 0(T_c)$ chiral symmetry is no longer spontaneously broken as observed at $T = 0$[5].

Several theoretical ideas on deconfinement at high temperature for the case of the pure gauge theory have been advanced, see for a review [6]. Here we wish to present a new semi-quantitative picture of quark liberation, which is not restricted to the pure gauge theory. This will be based on Mack's idea that (zero temperature) confinement arises as a hierarchy of effective $Z(N)$ theories [7]. $Z(N)$ is the discrete center group of the gauge group $SU(N)$. The physical case of interest obviously is $N=3$, but the study of the somewhat easier case of $N=2$ will already indicate the essential features of QCD. In the effective $Z(2)$ theory the role of the underlying (thick) vortices is essential. The precise definition of a vortex will be given in Sect. 1, one may consider it as the world sheet of a loop of magnetic flux. Vortices encircling the Wilson loop will have a disordering effect, so that its vacuum expectation value drops as an area law, which means that there is confinement. At strong coupling thin (= one plaquette) vortices do the job, while at weak coupling thick vortices are relevant. This is of paramount importance if the confinement of the strong coupling regime is to persist in the continuum limit.

In Sect. 1 we will discuss this zero-temperature theory of confinement in such a way that the finite temperature modification makes it clear that above T_c confinement no longer persists in the continuum limit (Sect. 2). All this is for the pure gauge theory, but in Sect 3 we present some ideas how to test for quark liberation including the fermionic quarks as dynamical fields. The conclusions are presented in Sect. 4. Throughout we will use the Euclidean Lagrangian formalism in $D=4$ dimensions. The standard Feynman–Kac path integrals reduce on a finite lattice to ordinary multiple integrals (one integration over the group per link). The theory is a special problem of statistical mechanics, this also explains some terminology we will use. Hence it is easy to incorporate finite temperature in this formalism. As the reader has noticed we work in natural units, where $h=c=k=1$.

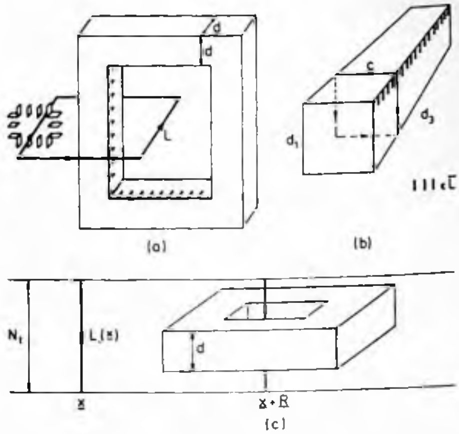

Fig. 1 a – c. Vortices for $D=3$ dimensions (the relevant case $D=4$ is analogous). The cubic lattice is not drawn, but the links in L indicate what the lattice spacing a is. In this figure we set $a=1$. **a** Vortices disorder the expectation value of the Wilson loop along L. Drawn are a thin vortex with a thickness of one plaquette and a thick vortex (container) with thickness $d^2>1$. To give a twist to the boundary conditions of the container change the link variables $U(1)$ to $\gamma U(1)$ for $1 \in \bar{L}$, which encircles the loop (indicated as heavy short lines), and with γ in the center $Z(N)$ of the gauge group $SU(N)$. **b** Long box to study the vortex free energy of (13). Twisted boundary conditions arise in the same way as in **a**. The twist changes $U(C)=\Pi U(1)$, $1 \in C$, to $\gamma U(C)$ for every C as indicated. **c** Thick vortex (container) disordering the expectation value of the thermal Wilson lines $L(\mathrm{x})L(\mathrm{x}+\mathrm{R})^\dagger$ (see Sect. 2). The lattice has a temperature $T=(N_t)^{-1}$. At high temperatures, small N_t, the thick vortices required for confinement cannot disorder $LL^\dagger$ enough and quarks are liberated.

1. Quark Confinement at $T=0$

The Abelian LGT with as gauge group $Z(2)=\{1;-1\}$ is known to have a strong coupling regime ($\beta<\beta_c$) with confinement and a weak coupling regime ($\beta>\beta_c$) where confinement is lost. From duality follows $\beta_c(Z(2))=\frac{1}{2}\ln(1+\sqrt{2})=0.44$. The partition function Z is the multiple integral over the link variables $\sigma(1)\in Z(2)$ of $\exp(-\beta \times \text{Euclidean Action})$. For the connection between "inverse temperature" β and coupling constant g_0 see the $SU(2)$ case below. In general the Wilson loop for a path L and gauge group G is defined as

$$W(L)=\chi\Big(\prod_{1\in L} U(1)\Big) \qquad U(1)\in G, \tag{1}$$

where the product is path ordered along L and χ is a character, i.e. $\chi(U)=\operatorname{trace} R(U)$ for unitary representation R of G. For $G=Z(2)$ and surface S so that its boundary $\partial S=L$ we have, because G is Abelian,

$$W(L)=\prod_{p\in S}\sigma(\partial p), \tag{2}$$

with (in general path ordering is important)

$$U(\partial p=1_1+1_2+1_3+1_4)\equiv U(1_1)U(1_2)U(1_3)U(1_4). \tag{3}$$

Take for simplicity $D=3$ dimensions and lattice spacing a. The confinement of the strong coupling regime arises from the condensation of thin (thickness $=a$) vortices. A thin vortex is a ring (closed surface for $D=4$) of plaquettes with $\sigma(\partial p)=-1$ (see Fig. 1a.) From (2) we see that numerous long vortices intersecting S may disorder $W(L)$ so that the Wilson parameter $\langle 0|W(L)|0\rangle$ drops sharply $\propto \exp(-\sigma\, A(L))$, signalling the confinement phase. These vortices are condensed for $\beta<\beta_c$ ("high temperature") because an entropy contribution compensates the large energy of long vortices. But for $\beta>\beta_c$ ("low temperature") these long vortices are rare and only short vortices encircle L so that $\langle 0|W(L)|0\rangle$ drops only as a perimeter law and quark confinement is lost (see [8]).

Now we go to the non-Abelian LGT with $G=SU(2)$, in which we are interested. The standard Wilson action, measure and partition function are for

$G = SU(N)$ in general [2]

$$A(U) = \sum_p \beta[1 - N^{-1} \operatorname{Re} \operatorname{tr}(U(\partial p))], \tag{4a}$$

$$d\mu(U) = Z^{-1} e^{-A(U)} \prod_l dU(l), \tag{4b}$$

$$Z = \int d\mu(U), \tag{4c}$$

with $dU(l)$ the normalised Haar measure over G. Note that we use the fundamental representation, so that $\chi(U) = \operatorname{tr} U$. There is invariance under *local* gauge transformations

$$\Omega(x) \in G: U(l) \to \Omega(x) U(l) \Omega(y)^{-1},$$

for l between sites x and y. The vacuum expectation value of observable $F(U)$ is

$$\langle 0|F(U)|0\rangle = \int d\mu(U) F(U). \tag{4d}$$

From the classical continuum limit one has $\beta = 2N/g_0^2$. The bare coupling $g_0(a)$ is given by the perturbative ($g_0^2 \to 0$) renormalization group equation (11) for $G = SU(N)$, which gives asymptotic freedom $g_0^2(a) \to 0$ for $a \to 0$. For $G = SU(2)$ at low enough β there is confinement just as if $G = Z(2)$. We see this as follows: 1) define plaquette variables

$$\sigma(\partial p) = \operatorname{sign} \chi(U(\partial p)), \tag{5}$$

2) note that the following inequalities hold [9, 8]

$$|\langle 0|\chi \prod_{l \in L} U(l)|0\rangle| \leqq 2\langle 0| \prod_{p \in S} \sigma(\partial p)|0\rangle$$
$$\leqq 2\langle 0| \prod_{l \in L} \sigma(l)|0\rangle_{Z(2)}, \tag{6}$$

where we compare the $SU(2)$ Wilson loop, which is bounded by the term in the middle, with the $Z(2)$ one at the same β, 3) from the discussion above we know that the $Z(2)$ theory for $\beta < \beta_c$ has the area law behaviour for the Wilson parameter and (6) implies the same for $SU(2)$ (q.e.d). Hence, thin vortices, i.e., configurations of $U(l)$ with on a closed surface $\sigma(\partial p) = -1$, give confinement for low enough β in the $SU(2)$ LGT.

The phase diagram we would like to establish for $SU(2)$ (and, of course, for the $SU(3)$ of QCD) has a confinement phase for all finite β and a second order phase transition at $\beta_c = \infty$. At the second order phase transition the correlation length in units of a diverges, so that the details of the fine scale lattice become irrelevant. Because of asymptotic freedom (running coupling constant vanishing with high energy) we want β_c at the free limit $g_0 = 0$. There are some indications that the $SU(2)$ LGT has the wanted phase structure in 4 dimensions [10] and computer simulations [4] appear to confirm this.

The question is now to understand how the $SU(2)$ LGT *avoids the deconfining phase transition*, which occurs in the $Z(2)$ LGT at finite β_c. We will give a brief discussion of Mack's answer ([7], see there for further hypotheses and arguments). The idea is that the $SU(2)$ Wilson loop is suppressed as an effective $Z(2)$ theory with lattice spacing d and coupling $\beta_{eff}(d) = \beta_c(Z(2)) - \delta$, where $0 < \delta \ll 1$. This results from the condensation of thin vortices with thickness $d = a$ for $\beta < \beta_1$ and of thick vortices with $d \geqq d_c$ for $\beta > \beta_1$. $\beta_1 \sim 2.0$ is a constant determined from low β expansions of the string tension. When going towards the continuum limit ($\beta \to \infty$) there are always thick vortices condensed, so that for any β the Wilson loop $W(L)$, with size $\gg d_c/a$, is strongly disordered by these thick vortices, just as in the confinement phase of the $Z(2)$ theory. The thick vortices may be viewed as having spread out* the magnetic flux of a thin vortex (Fig. 1a). This spreading out is possible in the Wilson action of the *continuous* $SU(2)$ group. In the thick vortex for every plaquette $\frac{1}{2}\operatorname{tr} U(\partial p)$ may be $1-\delta$, contrary to the -1 of the plaquette in the thin vortex. This shows how *thick* vortices may continue to play a role in the continuum limit ($\beta \to \infty$), while thin vortices of $\sigma(\partial p) = -1$, see (5), are virtually non-existant because for large β only configurations with $U(\partial p)$ near 1 are highly favoured, see (4abc).

The remaining problem, of course, is to *prove* that the $SU(2)$ LGT behaves on large scales as a $Z(2)$ theory with an effective coupling $< \beta_c$ so that it is in the confinement phase. One could use the idea of block spin [10], where the basic algorithm runs as follows: on a lattice with spacing $a_{(n)}$ define new variables on a larger scale $a_{(n+1)}$ and calculate their effective interactions. The gain of reducing the number of variables may be countered by the complicated nature of these effective interactions. But in many problems of statistical mechanics the phenomenon of universality occurs, meaning that after many iterations of the algorithm the solution converges, irrespective of initial details, to one simple form. In our case this should be the $Z(2)$ Wilson action, where the $\beta_{eff}(d)$ remains to be determined. Let us consider some fixed $\beta > \beta_1$, of the original $SU(2)$ LGT. After many block spin iterations we hopefully arrive at an effective $Z(2)$ LGT with lattice spacing d_c and coupling $\beta_{eff}(d_c)$. If $\beta_{eff} < \beta_c(Z(2))$, then thin vortices in the effective theory, which were thick vortices in the underlying $SU(2)$ theory, will give confinement as discussed in the beginning of this section.

Up till now this scenario of solving the theory b[y] block spinning and proving confinement has not be[en] performed, but it provides us the conceptual frame[work for] Mack's (hypothesized) effective $Z(2)$ theory of confi[ne]ment. Another line of attack is the notion that $\beta_{eff}(d_c)$ measures the free energy of a vortex with thickness d_c (see Sect. 3). In [11] a so-called vortex container was introduced, which is a rectangular sublattice of the form of a thick vortex (Fig. 1a). One considers then the change of its free energy when a unit of magnetic flux is forced through by a "twist" of the boundary conditions

* See [8] for the analogous role of thick Bloch walls in the $D = 2$ ferromagnet with continuous symmetry. Here for all β there is disorder in contrast to the Ising model. Even at large β long vortices, i.e. Bloch walls, exist. A longer vortex can counter the increase of its energy from its greater length by thickening, which reduces the gradient contribution to its total energy

(see Sect. 3). For non-overlapping containers winding around the Wilson loop a sufficient condition for confinement is that the free energy change should drop exponentially with the thickness of the container [11]. In Sect. 3 we will discuss somewhat further this method to determine the physics of thick vortices from the response of (sub) lattices to twisted boundary conditions. Clearly one determines in this way the contribution to the free energy from the internal structure of the thick vortex, whereas the configurational entropy may be derived on a coarse lattice where they are just thin vortices. In this way we also arrive at an effective $Z(2)$ theory. This ends our intermezzo on *how* Mack's effective $Z(2)$ theory might arise.

In [7] the minimum thickness d_c for vortex condensation is derived to be ($\beta_1 \sim 2.0$)

$$a^{-1}d_c = 1 \qquad \beta < \beta_1 \tag{7a}$$

$$a^{-1}d_c = \exp\{(3\pi^3/11)(\beta - \beta_1)\} \qquad \beta > \beta_1. \tag{7b}$$

This completely determines σa^2 as a function of β and agrees with numerical results [4]. Also the predictions for the vortex free energy (see Sect. 3) are confirmed by direct numerical simulations with $\beta_1 = 2.06$ [12, 13].

In the standard $SU(2)$ theory magnetic strings (lines of plaquettes with $\sigma(\partial p) = -1$ from (5)) are observable, since they contribute to the action (4a) differently than if $\sigma(\partial p) = +1$. Note that if the Lagrangian density would be proportional to $|\mathrm{tr}\, U(\partial p)|^2$ this would not be the case. [This variant theory basically has a gauge group $G = SO(3)$, which fulfills the continuum condition $\pi_1(G) \neq \varnothing$ for the existence of monopoles with invisible strings, whereas $\pi_1(SU(2)) = \varnothing$[13]] Vortices are *closed* surfaces (in $D = 3$ a closed line, Fig. 1a) of these negative plaquettes and are the world sheet of a closed magnetic string. But the magnetic string may also end in a monopole (for $SU(2)$ there is no distinct antimonopole, $Z(2)$ having only one element $\neq 1$). A monopole is identified in a cube c if the configuration $U(l) \in SU(2)$ is such that with (5)

$$\prod_{p \in \partial c} \sigma(\partial p) = -1. \tag{8}$$

These monopoles are important for $\beta \lesssim \beta_1$. They determine the cross over region [14] and contribute via the free energy of the thin vortices to the string tension [13]. It is not unexpected that these small (size $\leqq a$) monopoles contribute to quantities defined over a plaquette size, respectively specific heat and thin vortices. For large operators, e.g. the Wilson loop, they will not be important. Even apart from this size argument one knows that small monopoles are very rare in the continuum limit $\beta \to \infty$ where for nearly all plaquettes $\sigma(\partial p) = 1$ (see above).

In principle one can define a larger monopole on the scale of a large cube C, for example, with in (8) the product of plaquettes running over ∂C. We only consider thick vortices for confinement in the weak coupling regime and not corresponding large monopoles and their thick strings*.

With this picture, in the zero temperature LGT, of confinement persisting towards the continuum limit by the presence of thick vortices, we will address what happens at finite temperatures.

2. Quark Liberation at $T \geqq T_c$

Field theory at equilibrium at a temperature $T > 0$ is easy to incorporate [17] in the path integral formalism used in the previous paragraph. Four dimensional space integrals run over $R^3 \times [0, T^{-1}]$, i.e. the "time" axis is shortened, and path integrals use only periodic field configurations with $A_\mu(\mathbf{x}, 0) = A_\mu(\mathbf{x}, T^{-1})$, with $A_\mu(x)$ the Lie algebra valued field. On a lattice its size also is flattened $N_t \times N_s^3$ ($N_t \ll N_s$). For simplicity we take the lattice spacings and coupling constants equal in timelike and spacelike directions (for a more precise treatment see [18]). In numerical simulations, specifically those of the Monte Carlo (MC) type, a typical lattice with $N_t = 2$–5, $N_s = 10$ gives a good indication for the behaviour near T_c [3].

The free energy $F(R)$ of a static quark–antiquark ($q\bar{q}$) pair at a spatial distance $|\mathbf{R}|$ is given by [19]

$$\exp[-T^{-1}F(\mathbf{R})] = \langle L(0)L(\mathbf{R})^\dagger \rangle_T, \tag{9}$$

with $\langle\,\rangle_T$ from the normalized multiple integral over periodic configurations. The Wilson line $L(\mathbf{x})$ is defined as $\prod U(l)$, where the product runs over the links of the straight line from $(x, y, z, 0)$ to (x, y, z, T^{-1}).

We will now turn to the phenomenon of quark liberation at $T \gg M$, with $M \sim (2 \text{ to } 3)\sigma^{1/2}$ the glueball mass [23]. Our starting point will be the ideas discussed in Sect. 1, notably the role of thick vortices.

Consider a lattice of fixed size $N_t \times N_s^3$, with N_t small (but see below) and N_s large enough. Because $T^{-1} = N_t a < M^{-1}$ we are in the weak coupling regime ($\beta > \beta_1$; $a < M^{-1}$), where thick vortices can be important a priori. But the thick vortices, which in case of confinement ($F \propto R$) must disorder between the L, $L^\dagger$ operators so that $\langle L(0)L(\mathbf{R})^\dagger \rangle_T \propto \exp(-\sigma T^{-1} R)$ are now restricted in their thickness (Fig. 1c). If we approximate by using vortex containers [11] their maximum thickness is $d_{\mathrm{crit}} \sim N_t a$, regardless of the limit $R \to \infty$. From (7b) we have the minimum thickness d_c/a required for confinement at β. Because the hierarchy of ever thicker vortices for increasing β only applies if $d_c/a \leqq N_t$ we find a critical coupling above which confinement fails

$$\beta_{\mathrm{crit}}(N_t) = \beta_1 + (11/3\pi^2)\ln(N_t). \tag{10}$$

The perturbative ($\beta \equiv 2N/g_0^2$ large) renormalization

* For a variant action analogous to that of [15] over larger sizes than a plaquette the phase diagram is similar [16]. The relation between thick monopoles and vortices with respect to the vortex internal structure should be compared to that of the thin ones, where small monopoles contribute significantly to σa^2 in the strong coupling region [13]

group equation gives the relation between g_0^2 and a for $SU(N)$ in general

$$a = \Lambda_{\text{lattice}}^{-1} \exp\left[-\frac{24\pi^2}{11N} g_0^{-2} - \frac{51}{121} \ln \frac{11N}{48\pi^2} g_0^2 \right], \tag{11}$$

where Λ_{lattice} a specimen of the notorious Λ parameters, which arise from regularization, and different from those of the continuum theory by a calculable factor. From $T = (N_t a)^{-1}$ and (11) one finds the physical temperature $T_c = 0(100\Lambda_{\text{lattice}}) = 0(\sigma^{1/2})$ corresponding to the lattice results $\beta_{\text{crit}}(N_t)$ of (10).

"But, objects the alert reader, doesn't the very same argument hold against confinement at *low* temperatures $T \ll M$?" The negative answer is twofold:

1. In standard MC studies of confinement at $T = 0$ [4], the Wilson loop L is taken to be (much) smaller than the lattice size N_s^4, in order to *avoid* finite size and boundary effects. Also β must be smaller than $\sim \beta_{\text{crit}}(N_t = N_s)$[19]. Then the results of $\langle 0|W(L)|0\rangle$ tell us about the potential energy of a $\bar{q}q$ pair at $T \simeq 0$.

2. In principle one can also study the $q\bar{q}$ interaction averaged over a thermal ensemble of (anti) quarks, be it at low temperatures $0 < T \ll M$. For a $q\bar{q}$ separation R the free energy $F(R)$ again can be determined through (9) from the value of $\langle L(0)L(\mathbf{R})^\dagger \rangle_T$. For our lattice we have the physical condition $N_t a \ll M^{-1}$ and in order to mimic the infinitely extended continuum space $N_s \gg N_t$. For the small N_t we study we are now in the strong coupling regime (β small, $a \gtrsim M^{-1}$). Here confinement can be directly calculated to occur [2]. The physical picture is the condensation of *thin* ($\lesssim a$) vortices, just as in the $Z(2)$ gauge theory. For N_t a few these relevant vortices indeed fit in between the L, $L^\dagger$ operators, so as to give the strong disorder leading to confinement at these low temperatures $T \ll M$.

The second part of the above answer seems to require a small N_t. This cannot be, intuition tells us we should find the same, or even better, physical results on a fine scale lattice with $1 \ll N_t \ll N_s$. We will show that this is the case. Crucial is the observation that the minimum vortex thickness d_c is a *physical* quantity, only when measured in units a will it increase for larger β. The exponent of (7*b*) should include an extra factor, which comes from the two loop calculation, just as the one in (11). The end result for d_c then is independent of β, as long as $\beta > \beta_1$. This is obvious because from [7] we have $d_c^2 = 0.54\,\sigma^{-1}$, which relates d_c directly to the physical string tension σ. In the derivation one uses $\sigma_{Z(2)}(\beta = \beta_c - \delta) = 0.54 a^{-2}$ for the $Z(2)$ string tension, which follows from strong coupling expansions. Similarly $\beta_1 \sim 2.0$ is determined from $\sigma_{SU(2)}(\beta = \beta_1) = 0.54 a^{-2}$. We now reassure the alert reader from above completely: for low temperatures there is confinement, because thick vortices can always disorder between the $L(\mathbf{0})$ and $L(\mathbf{R})^\dagger$ operators as long as $T^{-1} \gg d_c$, for all $N_t = (Ta)^{-1}$ if $\beta(a) > \beta_1$. Mutatis mutandis we have quark liberation for $T^{-1} \ll d_c$.

In Mack's effective theory it is assumed implicitly that only vortices with thickness $d = d_c$ make the disorder. However it may very well be that a somewhat thicker vortices contribute. Generally ve much thicker vortices will be less important becau their configurational entropy is smaller than a thinn vortex of the same length. We may take an avera thickness $d = Fd_c, F \geqq 1$. We then estimate, usir $\sigma d_c^2 = 0.54$,

$$T_c = (1.4/F)\sigma^{\frac{1}{2}}. \tag{12}$$

The predictive power of our reasoning for the numeri cal value of T_c is not great, but it states correctly that T is at least a factor 2 below the glueball mass [23], which is the only mass scale in pure LGTs. The Monte Carlo results [3] seem to imply a moderate $F \simeq 3$.

To summarize, the semi quantitative theory discussed in Sect. 1 explains (de) confinement at temperatures (above) below some T_c by the (absence) presence of vortices of a thickness a few times d_c. Up to now we only discussed the pure gauge theory with quarks acting as classical sources. But the mechanism presented for (de) confinement at (high) low temperatures also remains valid if quarks act as truly dynamical fields. In the following paragraph we will present some ideas how high temperature quark liberation in a LGT with fermion fields might be established.

3. Vortex Free Energies

Another order parameter to determine the phases of LGT is the vortex free energy μ [20, 21]. On a finit lattice Λ with a non-simply connected boundary $\partial\Lambda$ one considers the partition function $Z(U(\partial\Lambda))$ as a function of its boundary conditions (*bc*).

Hence in (4*bc*) we do not arbitrarily integrate over the $U(l \in \partial\Lambda)$'s. Specifically one looks at the relative change from a twist:

$$\mu_\gamma(\Lambda, bc) = -\tfrac{1}{2} \ln [Z(\Lambda, \text{twisted } bc)/Z(\Lambda, bc)]. \tag{13}$$

where (Fig. 1*b*)

1. Λ is a parallelopiped in $D = 4$ dimensions with sides $d_1 \times d_2 \times d_3 \times d_4$ and $d_1 = d_2 \ll d_3 = d_4$
2. boundary conditions are periodic for the gauge fields (free for the fermions)
3. a twist is introduced by some singular gauge transformation so that $U(C) \to \gamma U(C)$, γ is a non-trivial element of the group center ($= -1$ for $Z(2)$) and C on $\partial\Lambda$ winds around the short dimensions of the box.
4. this twist may be obtained from $U(l) \to \gamma U(l)$ for $l \in \tilde{L}$ in Fig. 1b.

The physics behind this scheme is that the twist forces a unit of magnetic flux through the box in the long direction. Compare with the vortex container of Fig. 1a where also a twist is introduced by $U(l) \to \gamma U(l)$, $l \in \tilde{L}$ [11]. An analogous expression to (13) with for Λ a container with thickness d determines the effective $Z(2)$ coupling $\beta_{\text{eff}}(d_1)$ of Sect. 1 [13]. Hence it is interesting to study (numerically) (13), which is a sort of straightened out vortex container. Most importantly the possible different behaviours of (13) with respect to the

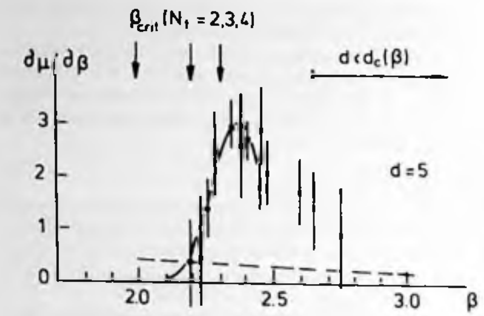

Fig. 2. Numerical results of [12] on the vortex free energy μ of (13) for a box d^4. The theoretical curve based on Mack's theory of confinement agrees well with the data points. For a fixed box size, here $d = 5$, at high enough β the minimum thickness of the vortices d_c is larger than the box considered. Here the results should follow a perturbative curve (dashed line?, see [12, 24]). The high temperature phase transition from confinement to liberation for β rising above $\beta_{crit}(N_t)$ is expected to show up for $N_t \times d^3$ box sizes. The numerical results should then follow the perturbative curve down to $\beta_{crit}(N_t)$ and then for lower β along Mack's curve. This method to calculate numerically $\beta_{crit}(N_t)$, and hence T_c, from the difference by flattening for the vortex free energy μ, perhaps can be done also with dynamical quark fields present.

box dimensions also distinguish the phases of the gauge theory when dynamical quark fields are present [21].

The theory of confinement at $T = 0$ as described in Sect. 1 makes a prediction for the behaviour of μ as a function of β. This has been confirmed by numerical simulations [12, 13] for d^4 boxes, with $d = 3 - 6$. These simulations calculate $\partial\mu/\partial\beta$ from the difference of the averaged $\mathrm{tr}\, U(\partial p)$ with or without twisted boundary conditions on the box, see (4) and (13). Following the ideas of Sect. 2 on the mechanism of quark liberation it would be interesting to compute μ for a $N_t \times d^3$ block with $N_t \ll d$. We know that for $\beta > \beta_{crit}(N_t)$ there are no thick vortices to give confinement and we expect some drastic changes in the numerical results in the range $\beta_{crit}(N_t)$ up to $\beta(da = d_c)$. Even for the hypercubic box d^4 above $\beta(da = d_c)$ there are no more thick vortices operating in it. Below $\beta_{crit}(N_t)$ we are still in the confinement phase and the results for the hypercubic and flattened box should agree approximately. Figure 2 sketches the expected results. As said this method to determine $\beta_{crit}(N_t)$ for the pure gauge theory probably can be extended to include dynamic fermion fields [21], because their contribution to the partition function can be calculated exactly in MC simulations over the gauge fields [22]. The twist operates on the gauge fields only, while the fermion fields contribute to the internal structure. In a MC on a $N_t \times d^3$ lattice one should try to establish a $\beta_{crit}(N_t)$, which separates different behaviour of μ, or a quantity determined by it, as a function of β. Many details are unresolved or undecided for the moment, but this appears to be an alternative* way to determine T_c for QCD on the lattice.

4. Conclusions

Quark confinement at low temperatures and liberation at $T > T_c$ has been, or will be shortly, established by numerical simulations of the pure gauge theory on the lattice. The inclusion of dynamical quark fields in these calculations is possible in principle, cf. [22]. But these are brute force results, and we would like to isolate the components in the theory which are responsible for the "observed" behaviour. For the persistance of confinement to the continuum theory at $T = 0$ they are the thick vortices of Sect. 1 [7, 8, 11]. Focussing on them we showed (Sect. 2) how to understand the phenomenon of quark liberation above a critical temperature T_c. We estimated T_c in (12), which is not too far from the Monte Carlo value. In Sect. 3 we turned to the change of free energy in boxes when a magnetic flux line is introduced by a twist in the boundary conditions of the box. This is very much like the vortex containers that approximate the physics of the thick vortices [11]. Here we showed how numerical simulations could be used to determine T_c. This was for the pure gauge theory, but the vortex free energy remains a relevant order parameter when dynamical quarks are present [21], contrary to the Wilson loop parameter [2]. Perhaps T_c can be calculated analogously in this case. To summarize we may say, that the thick vortices of the lattice gauge theory, which probably guarantee the persistance of confinement in the continuum limit, simultaneously explain by their absence at high temperatures the liberation of the quarks.

Acknowledgements. I thank Prof. M. Rees for the hospitality of the Institute of Astronomy, Cambridge, where this work was started, and Dr. J. Smit for valuable comments on an earlier version of the present article.

References

1. C. Quigg: Les Houches Summerschool 1981, preprint Fermilab-Conf-81/78-Thy; E.S. Abers, B.W. Lee: Phys. Rep 9C, 1 (1973); S. Mandelstam: Phys. Rep. 67C, 109 (1980)
2. K.G. Wilson: Phys. Rev. D10, 2445 (1974)
3. J. Engels, F. Karsch, H. Satz, I. Montvay: Nucl. Phys. B205, 545 (1982), and references herein
4. M. Creutz: Phys. Rev. D21, 2308 (1980) and Phys. Rev. Lett. 45, 313 (1980); G. Bhanot, C. Rebbi,: Nucl. Phys. B180, 469 (1981)
5. J. Engels, F. Karsch, H. Satz: preprint BI-TP 82/8; J. Kogut et al.: preprint ILL-TH-82-5
6. F.R. Klinkhamer: In Proceedings XVIIth Rencontre de Moriond 1982, Gif-sur-Yvette: Frontieres, (in press)
7. G. Mack: Phys. Rev. Lett. 45, 1378 (1980)
8. G. Mack: In Recent developments in gauge theories. eds. G. 't Hooft et al., New York: Plenum Press 1980

* The Bielefeld collaboration of [5] calculated the energy density $\varepsilon(\beta)$ with the leading term of the hopping parameter expansion for the Wilson action of the fermions, $\varepsilon(\beta)$ changes rapidly near $\beta_{crit}(N_t)$. Similarly $\partial\mu/\partial\beta$ might be calculated perhaps

9. G. Mack, V.B. Petkova: Ann. Phys. **123**, 442 (1979)
10. L.P. Kadanov: Rev. Mod. Phys. **49**, 267 (1977); K. Wilson: In G. 't Hooft et al. (eds): Recent developments in gauge theories. New York: Plenum Press 1980
11. G. Mack, V.B. Petcova: Ann. Phys. **125**, 117 (1980)
12. G. Mack, E. Pietarinen: Phys. Lett. **94B**, 397 (1980)
13. G. Mack, E. Pietarinen: Nucl. Phys. **B205**, 141 (1982)
14. R.C. Brower, D.A. Kessler, H. Levine: Nucl. Phys. **B205**, 77 (1982)
15. L. Caneshi, I.G. Halliday, A. Schwimmer: Nucl. Phys. **B200**, 77 (1982)
16. I.G. Halliday: private communication
17. C.W. Bernard: Phys. Rev. **D9**, 3312 (1974)
18. F. Karsch: Nucl. Phys. **B205**, 285 (1982)
19. L.D. McLerran, B. Svetitsky: Phys. Rev. **D24**, 450 (1981)
20. G. 't Hooft: Nucl. Phys. **B153**, 141 (1979)
21. G. Mack, H. Meyer: Nucl. Phys. **B200**, 249 (1982)
22. A. Duncan, M. Furman: Nucl. Phys. **B190**, 767 (1981); A. Duncan, R. Roskies, H. Vaidya: Phys. Lett. **114B**, 439 (1982)
23. B. Berg, A. Billoire, C. Rebbi: Ann. Phys. **142**, 185 (1982)
24. A. Gonzalez-Arroyo, J. Jurkiewicz, C.P. Korthals-Altes: preprint CPT-81/P. 1336

HEATING THE QUENCHED EGUCHI-KAWAI MODEL

F.R. KLINKHAMER

Leiden Observatory, Postbus 9513, 2300 RA Leiden, The Netherlands

Received 12 November 1982
(Revised 3 January 1983)

We consider the Eguchi–Kawai reduction, in the momentum-quenched prescription, of the SU(N) lattice gauge theory for $N \to \infty$ and address the problem of how finite temperature might be incorporated. This is of interest in order to establish quark deconfinement at high temperatures. We also show that different quenching procedures may be inequivalent.

1. Introduction

An approach towards solving quantum chromodynamics (QCD) is to consider the large-N limit of the SU(N) gauge theory [1]. Some experimentally observed characteristics of QCD, which has only $N = 3$ colours, may be explained in this way*. Another line of attack is the non-perturbative regularization from a space-time lattice, a method initiated by Wegner and Wilson [3]. Both strong coupling expansions and numerical, Monte Carlo, simulations of lattice gauge theories (LGTs) with $N = 2$ and 3 provide some tentative results, notably on the persistence of confinement and the hadron spectrum [4]. These simulations also show that in QCD, deconfinement occurs at a critical temperature $T_c \sim 200$ MeV and that at $T \gtrsim 5T_c$ there is an asymptotically free gluon gas [5]. As an important theoretical task it remains to determine the mechanism and the dynamical variables that operate this quark liberation; for some general papers see [6, 7]. Here we will try to open a new way using the large-N LGT, for which recently a reduced model has been proposed [9, 10, 12, 13, 26, 27].

In sect. 2 we discuss the reduced model [12], trying to be as explicit as possible in order to clarify some ambiguities. How to incorporate finite temperature will be discussed in sect. 3, where we show that the finite temperature expectation value of a *local* operator may be calculated in the reduced model from (9). How one might then look for high-temperature deconfinement will be the topic of sect. 4. In our way of introducing temperature into the reduced model this should be done by calculating intensive thermodynamic quantities such as the energy density (14). We also mention in this section that the claimed equivalence of the models of [10] and [12] is not precise, or equivalently that the method of quenching is somewhat

* See for example ref. [2].

arbitrary. In sect. 5 we present the conclusions and discuss briefly the hamiltonian point of view.

Before we proceed with the reduced large-N LGT we remark that Thorn [8] has argued that the $N\to\infty$ continuum theory indeed has a deconfining *phase transition* at T_c. The argument is that for $T>T_c$ the free energy density $f=F/V$ is of order N^2, because of asymptotic freedom, while at $T<T_c$ it is $O(1)$, because glueball interactions are suppressed $O(N^{-1})$ [2]. This shows a different analytic behaviour above/below T_c, which is the definition of a phase transition. Whether for large, but finite, N this phase transition persists or that deconfinement is more like the gradual dissociation of molecules, say, is not clear yet.

Finally note that the difference between $U(N)$, which we will consider in the following, and $SU(N)$ is trivial, $-f_{U(N)}=-f_{SU(N)}+\pi^2T^4/45$, because the singlet field $\mathrm{Tr}\,(A_\mu(x))$ in $U(N)$ is free [8]. The space-time dimension is $D=d+1$ and if not specified we take the relevant case of $d=3$ space dimensions. Natural units with $h/2\pi=c=k=1$ are used.

2. Reduced model

Recently Eguchi and Kawai [9] showed that the standard [3] $U(N)$ LGT for $N\to\infty$ can be reduced significantly: from an infinitely extended lattice to a periodic lattice with one site only. But numerical simulations [10, 11] proved this reduction of the number of variables not to be correct in the weak coupling regime. The reason is that for couplings $g^2\leqslant g^2_{EK}\sim 6.7/N$ [11], a $U(1)^D$ global symmetry of the EK model [9] is broken spontaneously, which invalidates the equivalence with the standard model. This symmetry breaking arises from the strong effective attractions [10] at weak coupling in the EK model between the eigenphases p^i_μ ($i=1,\ldots,N$; see below) of the link variables, which have this $U(1)^D$ symmetry. Consequently Bhanot et al. [10] proposed to suppress these interactions, in the hope that their quenched EK model (QEK) also equals the standard model at weak coupling. The QEK model was vindicated (but see sect. 4) by Gross and Kitazawa [12], who showed that the planar diagrams of QEK are equivalent to the integrands of the standard perturbation expansion [1]. Here the (fixed) p^i_μ variables of QEK play the role of the loop momenta (cf. [13, 27]) of the standard theory and thus should be integrated over at the end.

Now let us be more specific. The Wilson action is [3]

$$S_W=g^{-2}\sum_x\sum_{\mu<\nu}\mathrm{Tr}\,[U_\mu(x)U_\nu(x+\hat\mu)U^+_\mu(x+\hat\nu)U^+_\nu(x)+\text{h.c.}]\,, \tag{1}$$

with $U_\mu(x)\in U(N)$ residing on the link (length a) from site x in the μ direction, or is in other words the bond between site x and its neighbour $x+\hat\mu$. There is invariance under gauge transformations $U_\mu(x)\to S(x)U_\mu(x)S(x+\hat\mu)^{-1}$ for any mapping $S(x)$: $Z^D \to U(N)$. The reduced action is [12]

$$S_R = g^{-2} a^{D-4} \sum_{\mu<\nu} \mathrm{Tr}\,[U_\mu D_\mu U_\nu D_\nu (U_\mu D_\mu)^+ (U_\nu D_\nu)^+ + \text{h.c.}]\,, \tag{2}$$

with $D_\mu = \exp(iP_\mu a)$, $(P_\mu)_{ij} = p^i_\mu \delta_{ij}$, $|p^i_\mu| \leq \pi/a$. The integration measure for this QEK model is [12]

$$\int \prod_\mu \mathrm{d}U_\mu\, C(U_\mu, D_\mu)\,, \tag{3}$$

$$C(U_\mu, D_\mu) = \int \prod_\mu \mathrm{d}V_\mu\, \Delta(D_\mu)\, \delta(U_\mu D_\mu - V_\mu D_\mu V_\mu)\,,$$

$$\Delta(D_\mu) = \prod_{i<j} \sin^2((p^i_\mu - p^j_\mu)\tfrac{1}{2}a)\,,$$

where the standard Haar measure $\prod_\mu \mathrm{d}U_\mu$ is constrained by C so that the variables D_μ cannot be eliminated from the action (2) by a change of variables, which would give us back the EK action. Note that by integrating out the U_μ variables, one arrives precisely at the QEK partition function of [10]. But it appears that there is some difference or arbitrariness in how to treat observables (see sect. 4). For other ways to counter the spontaneous breaking of $U(1)^D$, see the recent papers [26, 27], which focus on reproducing the standard Makeenko–Migdal loop equations. At large, but finite, g^{-2} and N the rôle of the different [10, 12, 26, 27] reduced lattice actions, which all seem to reproduce the planar graphs of perturbation theory and the loop equations with boundary conditions, is not clear. Note that the same problem, although less severe, arises for the standard LGT (Wilson, heat-kernel, . . . actions). The reduced model is invariant under the following $U(N)$ transformation (no sum over μ)

$$U_\mu D_\mu \to S U_\mu D_\mu S^{-1}\,, \qquad S \in U(N)\,, \tag{4}$$

which may be viewed as a gauge transformation $U_\mu \to S(x) U_\mu S(x+\hat{\mu})^{-1}$, cf. (6b). There is not a separate transformation for the D_μ, because these are not dynamical variables, but momenta only. The vacuum expectation value of operator $O(U_\mu, D_\mu)$ is given by

$$\langle 0|O|0\rangle = \int_{-\pi/a}^{\pi/a} \prod_{\substack{\mu \\ i}} (\mathrm{d}p^i_\mu a/2\pi) Z_R^{-1} \int \prod_\mu \mathrm{d}U_\mu\, C\, O \exp(S_R)\,, \tag{5}$$

where the normalized momentum integrations have a UV cut-off from the space-time lattice with spacing a and where there is an obvious normalisation by the quenched partition function $Z_R = Z_R(D_\mu)$, see (11b).

The equivalence of the QEK and Wilson models probably applies not only for strong [9, 10] and weak [12] coupling, but also at intermediate values. Numerical simulations [14] of QEK show the same first-order phase transition at $(g^2N)^{-1} \sim 0.3$

as that of the standard large-N SU(N) LGT with Wilson action [15]. From the work on mixed actions [16] we expect this phase transition not to be deconfining, but "only" to indicate different analytic behaviour of observables, for example the internal energy $E(g^2)$ for g^2 greater or smaller than $g^2_{\text{critical}} \sim (0.3N)^{-1}$. Note that also the heat-kernel action for $D=4$ dimensions has a phase transition [17], so that this analytical barrier for large-N theories does not seem to be an artifact of the lattice action used. Perhaps the reduced model has an even less smooth transition from strong to weak couplings; at least this has been shown to occur for $D=2$ [29].

Finally, we make the obvious remark that the reduced model can also be investigated numerically for confinement by the usual Wilson criterion [3]. For a rectangular contour C (sides T and L, $T \gg L$) one considers $W(T,L) \equiv \langle 0|N^{-1} \operatorname{Tr} \prod_C U_\mu(x)|0\rangle$ with the action (1) and standard measure. If $W(T,L)$ contains a term which drops exponentially with the area $A=LT$ for $A \to \infty$, there is quark confinement. Numerically [18] the string tension σ follows from

$$-a^{-2}\left[\ln \frac{W(T,L)W(T-1,L-1)}{W(T-1,L)W(T,L-1)}\right], \qquad T, L \gg \xi a^{-1},$$

where ξ is the physical correlation length. In this way, one eliminates a perimeter $P=2(L+T)$ contribution to $W(T,L) \propto \exp(-\sigma a^2 A - maP)$. If $a^2\sigma(g^2)$ for $g^2 \to 0$ scales according to the renormalisation group we may determine the physical tension σ in units of Λ_{lattice} and then in MeV. With the reduction $U_\mu(x) \to U_\mu$ [9]

$$\prod_C U_\mu(x) \to (U_\mu)^L (U_\nu)^T (U_\mu^+)^L (U_\nu^+)^T, \tag{6a}$$

for a $T \times L$ rectangle in the μ, ν plane, one could follow the same procedure by calculating $W(T,L)$ in the EK model. For the quenched model one uses the translation rule [12]

$$U_\mu(x) \to \exp(iP\cdot x)U_\mu \exp(-iPx) \tag{6b}$$

to calculate $W(T,L)$ with (2), (3) and (5), which should also be correct at weak couplings. The very same rule, of course, gives the reduced action (2) from (1).

3. Finite temperature

For a quantum field theory in equilibrium at temperature T [19], the expectation value of operator O is $\langle O\rangle_T = Z^{-1} \operatorname{Tr}(O\, \mathrm{e}^{-H/T})$, $\langle 1\rangle_T = 1$, with H the hamiltonian and the trace running over physical states. The standard perturbation expansion applies with the following changes in the Feynman rules

$$p^0 \to i\pi T(2n), \qquad [(2n+1) \text{ for fermions}], \tag{7}$$

$$\int \mathrm{d}^4 p \to 2\pi i T \int \mathrm{d}^3 p \sum_{n=-\infty}^{\infty},$$

$$\delta^4(p-p') \to (2\pi i T)^{-1}\, \delta_{nn'}\, \delta^3(p-p'),$$

where the factor i comes from the use of a euclidean metric. These rules also follow from the path-integral formalism, where for (fermionic) bosonic fields only (anti)periodic configurations contribute with period T^{-1} in euclidean time. For the standard LGT one considers thus a $N_t \times N_s^d$ lattice, with (time-) space-like lattice spacing $(a_t)a_s$ and $N_t a_t \ll N_s a_s$. The *physical* condition is that $N_t a_t = T^{-1}$. One may hope that this LGT describes the continuum QCD correctly if $\max(a_t, a_s) \ll 1$ fm. Numerical calculations for SU(N), $N = 2, 3$, have been performed successfully for lattices with $a_t = a_s$ and N_s and N_t of order 10 and 2–5, respectively [5]. For simplicity we will set $a_t = a_s = a$ at the end of the manipulations that are to follow. Consider now the reduced model, which was shown [12] to give the standard perturbation theory, *provided* the p_μ^i variables were treated as momenta. Clearly to put temperature into QEK the momentum integrations of (5) should be changed according to rule (7).

The prescription for the reduced action to be used at finite temperature is somewhat more subtle. The $T = 0$ QEK model is reduced from a symmetric lattice $N_s \times N_s^d$, $N_s \to \infty$. As mentioned, the standard $T > 0$ LGT is formulated on a $N_t \times N_s^d$ lattice, where the periodicity in euclidean time results in the Fourier sum of (7) in momentum space. Define "local" operators to have a support of links, which is not closed by the periodic boundary conditions of the finite lattice used. For a local operator O_{loc} and large N_t, i.e. for any fixed temperature a or $g^2(a)$ very small, we expect that in order to calculate $\langle O_{\text{loc}} \rangle_T$ only a trivial transcription [5] of the reduced action (2) is needed,

$$S_R^T = g_s^{-2}(a_t/a_s) \sum_{\substack{\mu<\nu \\ =1,\dots,d}} \mathrm{Tr}\,[\mathrm{UD}(\mu\nu)] + g_t^{-2}(a_s/a_t) \sum_{\substack{\mu= \\ 1,\dots,d}} \mathrm{Tr}\,[\mathrm{UD}(\mu t)]$$

$$+\text{h.c.}\,, \tag{8}$$

with $\mathrm{UD}(\mu\nu)$ an obvious abbreviation. At the end of the manipulations needed to extract thermodynamic quantities (see below), we will set $a_t = a_s$ in (8), because, while in the standard LGT $a_t \neq a_s$ is allowed by the *extra* N_t periodicity, which links a_t to the physical temperature $a_t = (N_t T)^{-1}$, this is not possible in the reduced model. The constraint of integration measure (3) remains unchanged, because it only guarantees that the p_μ^i variables cannot be transformed away, whereas the effect of temperature concerns the values the p_0^i may take. The finite temperature version of (5) is thus

$$\langle O_{\text{loc}} \rangle_T = \int_{-\pi/a_s}^{\pi/a_s} \prod_{\substack{i \\ \mu=1,\dots,d}} (\mathrm{d}p_\mu^i a_s/2\pi) \prod_i \Big(\sum_{n=-N_t/2}^{N_t/2} N_t^{-1} \Big) Z_R^{T^{-1}} \int \prod_\mu dU_\mu (\mathrm{e}^{S_R^T} O_{\text{loc}})\,,$$

$$p_0^i = 2\pi n T\,, \qquad N_t \equiv (T a_t)^{-1}\,, \qquad \{a_t = a_s\}_{\text{end}}\,, \tag{9}$$

where we indicated symbolically that we set $a_t = a_s$ at the end of the manipulations. This must give the correct perturbation expansion, cf. chapter IVa of [12], where again the tadpole diagrams from the constraint C vanish upon integration in (9) because of the reflection symmetry $p^i_\mu \to -p^i_\mu$ and $n \to -n$.

Finally we comment on the limits to be taken on the number of colours N, the lattice volume $N_t N_s^d$ and the lattice cut-off a. Recall that in our lattice approximation of the continuum we have the symmetrical picture $a \equiv a_t = a_s$ for simplicity. The problem is that these limits probably do not commute, and hence we must be specific about what we mean. We take the practical (optimistic) point of view of having a large fixed $N = \bar{N}$. The $N \to \infty$ theory is used to obtain an *approximation* to the $\bar{N}$ theory, which also hopefully contains the quintessence of QCD. If we are looking for the physical deconfinement transition (sect. 4) we must try to avoid the large-$\bar{N}$ phase transition [15, 17] mentioned in sect. 2. For a fixed temperature this is always possible with a large enough N_t ($a \equiv a_t = N_t^{-1} T^{-1}$), and of course we must simultaneously have $N_s \gg N_t$, which results, say, from setting $N_s = 100 N_t$. So to arrive at a correct, i.e. not an approximate, theory we should take $\lim_{\bar{N}\to\infty} \lim_{N_t \to \infty}$ (reduced model), where the N_t limit implies the limits of N_s and a (by the temperature condition). That these limits exist has not yet been proved, cf. the discussion in [24].

4. How to look for deconfinement

Finite temperature (de)confinement in the standard LGTs is established by calculating the connected Green functions of the operator [5, 20]

$$L(\underline{x}) = L(x_1, \dots, x_d) = N^{-1} \operatorname{Tr} \prod_{n=1,\dots,N_t} U_0(\underline{x}, n)\,,$$

which is gauge invariant because the line is closed by the periodic boundary condition in the time direction (period T^{-1}, see sect. 3). But (Q)EK is derived from an infinite lattice, or at least a large one where boundary effects are negligible; also our finite temperature version (9) is only to be trusted for "local" operators.

Suppose we naïvely use (9) to calculate expectation values of products of $L(\underline{x})$'s translated from the standard to the reduced model by recipe (6b). Alas our boldness will not be rewarded, consider namely the following two operators:

(i) $\langle L(\underline{x})\rangle_T$ serves as an order parameter in numerical simulations; (un)equal to zero it signals (de)confinement [20]. But its naive translation into QEK according to (6b) $\langle (U_0 D_0)^{N_t} \exp(-iP_0 a N_t)\rangle_T$ vanishes under the p_0^i integrations of (9). Also in EK with rule (6a) it is zero simply by the U(1) symmetry $U_0 \to e^{i\alpha} U_0$, if unbroken.

(ii) $\langle L(\underline{0}) L(\underline{R})\rangle_T - \langle L(\underline{0})\rangle_T \langle L(\underline{R})\rangle_T$ determines the finite temperature quark-antiquark potential for separation $|\underline{R}|$ and thus signals the (de)confinement phase too [20]. Translated to QEK it would vanish $O(N^{-2})$, because of large-N factorization, cf. [12], which on the other hand is a necessary [9] ingredient in order to arrive at the reduced model. We remark that the major weakness of the reduced model is precisely how to calculate corrections subdominant in N, cf. [12, 27, 28].

A more promising approach is to use a physical quantity, which in the standard model is local. An obvious candidate to consider is the temperature dependence of the energy density $\varepsilon(T)$. For $T \gg T_c$, the energy density should approach the Stefan–Boltzmann expression $\varepsilon_{SB} = N^2(\pi^2/15)T^4$, whereas in the confinement regime ($T < T_c$) it will be much diminished because of the reduction in degrees of freedom, i.e. heavy bound states versus N^2 free gluons. For SU(N), $N = 2$ and 3, the Monte Carlo simulations for $\varepsilon(T)$ indeed show a strong reduction at $T_c \sim 40\Lambda_{\text{lattice}} \sim 0.4\sqrt{\sigma}$ (lattice values quoted for $N = 2$) and the existence of a strongly interacting gas only in a small temperature interval $1-2\times T_c$ [5].

The well-known definitions of free energy F, energy density ε and pressure P

$$F = -T\ln Z\,, \qquad \varepsilon = -V^{-1}\left(\frac{\partial \ln Z}{\partial T^{-1}}\right)_V, \qquad P = -\left(\frac{\partial F}{\partial V}\right)_T, \tag{10}$$

are easily transcribed for the lattice theory, keeping N_t fixed $\partial/\partial T^{-1} = N_t^{-1}\,\partial/\partial a_t$, up to finite-size corrections [5].

We will give an equation for the free energy F of the standard theory in terms of the reduced model. But note that this finite temperature F is not directly related to the vacuum free "energy" F_{vac} mentioned in [9, 10, 12], which we first consider. From [12] one may derive

$$\frac{F_{\text{vac}}}{N_s^D} = \int [\mathrm{d}p]_{\text{norm}} \ln\left\{\int [\mathrm{d}U]C\exp\left(-S_R(U_\mu D_\mu)\right)\right\}$$

$$= \int [dp]_{\text{norm}} \ln\left\{\int [\mathrm{d}V]\sin^2(\cdots)\exp\left(-S_R(V_\mu D_\mu V_\mu^+)\right)\right\}, \tag{11a}$$

in a symbolic notation, cf. (2, 3, 5), and where in the last equation we have integrated out the U_μ variables. This is different from the expression of [10], which is in the same notation

$$\frac{F_{\text{vac}}}{N_s^D} = \int [\mathrm{d}p\,\sin^2(\cdots)]_{\text{norm}} \ln\left\{\int [\mathrm{d}V]\exp\left(-S_R(V_\mu D_\mu V_\mu^+)\right)\right\}, \tag{12}$$

where we used that the θ_μ^i variables of [10] are equal to $p_\mu^i a$. Gross and Kitazawa have shown, again integrating over the U_μ, that their quenched partition function is

$$Z_R = \int [\mathrm{d}p]_{\text{norm}} \int [\mathrm{d}U]C\exp\left(-S_R(U_\mu D_\mu)\right)$$

$$= \int [\mathrm{d}p]_{\text{norm}} \sin^2(\cdots)\int [\mathrm{d}V]\exp\left(-S_R(V_\mu D_\mu V_\mu^+)\right), \tag{11b}$$

where the last expression is identical to that of [10].

But the quenching procedure of [10] and the momentum approach of [12] appear to lead to the different F_{vac} above and also to different expressions of $\langle 0|O(L)|0\rangle$, for an operator $O(L)$ over a set of closed loops L on the standard lattice translated according to (6b). Exactly how one takes the quenching is somewhat arbitrary and we follow the prescription of [12], which has a clear physical interpretation.

At finite temperature in the continuum limit $a \to 0$ or N_t, $N_s \to \infty$ for a *fixed* temperature (see sect. 3) it seems reasonable to have a relation analogous to (11a) for $F(T, V)$

$$\frac{T^{-1}F}{N_t N_s^d} = -\int \sum [\mathrm{d}p] \ln (Z_\mathrm{R}^T) , \tag{11c}$$

with $\int \sum [\mathrm{d}p]$ an abbreviation of the integrals and sums in (9). From (11c) we have with $V(d) = (N_s a)^d$ the space volume

$$f(T) \equiv \frac{F}{V(d)} = -a^{-d-1} \int \sum [\mathrm{d}p] \ln (Z_\mathrm{R}^T) . \tag{13}$$

With the first and second thermodynamic relations of (10) we have from (11c) for the energy density $\varepsilon(T)$

$$\varepsilon(T) = -a^{-d} \left[\int \sum [\mathrm{d}p] \frac{\partial \ln Z_\mathrm{R}^T}{\partial a_t} \right]_{a_t = a_s} - \varepsilon_0 , \tag{14}$$

where we used, with N_t fixed, $\partial/\partial T^{-1} = N_t^{-1}\, \partial/\partial a_t$ and the fact that $\int \sum [\mathrm{d}p]$ of (9) is independent of a_t. We subtracted a constant from the RHS of (14) because in the partition function $Z = N(T) Z_\mathrm{R}^T$ we had omitted the temperature-dependent normalization $N(T)$, cf. [5, 19]. The first RHS term of (14) is the expectation value of the time-like plaquettes. When $T \to 0$ the sum in (9) becomes an integral and the ε_0 term subtracts a vacuum contribution so that $\varepsilon(T=0) = 0$. Similarly a complicated expression for the pressure P could be written down.

In the standard LGT one derives from (10) similar expressions, which reduce in the continuum limit $a \to 0$ to [5]

$$\varepsilon a^4 = 6Ng^{-2}(\bar{P}_t - \bar{P}_s) , \qquad (\varepsilon - 3P)a^4 = 6Na \frac{\partial g^{-2}}{\partial a} (\bar{P}_t + \bar{P}_s - 2\bar{P}_\mathrm{sym}) ,$$

with $\bar{P}_r$, r = s or t, the expectation value of

$$\sum_{\{P_r\}} [1 - N^{-1} \operatorname{Re} \operatorname{Tr} U(P_r)] / \sum_{\{P_r\}} ,$$

and $\bar{P}_\mathrm{sym}$ the same for all plaquettes over a large symmetric lattice, which comes from the term $-\varepsilon_0$ of (14). It is evident, cf. [21], that any $\bar{P}_r = \mathrm{O}(g^2 N^2)$, with r = s, t, sym, so that we see analytically that for $T \to \infty$ the ideal gluon gas arises, which is confirmed by the numerical simulations [5]. For the reduced model the situation is less transparent, but for $T \to \infty$ (weak couplings) we must have $Z_\mathrm{R}^T \to Z$, which then gives the SB values of ε and P [6]. Monte Carlo studies must be carried out for lower temperatures of ε and $\partial f / \partial g^{-2}$ after which the latter is to be integrated over numerically to give f, see sect. 5 for some remarks.

5. Discussion

In [12, 13] it was shown that the quenched Eguchi–Kawai model produces identically the planar graphs of the standard perturbation theory, *if* the diagonal variables P_μ are treated as momenta. This model was reduced from a $D = d+1$ dimensional symmetric lattice of infinite volume. We have proposed in sect. 3 that finite temperature might be incorporated by the standard change of the momentum integrations (7) and temporarily of the action (8) in order to extract thermodynamic observables (sect. 4), after which we set the space- and time-like lattice parameters equal. At a fixed temperature we expect this procedure to be correct, for "local" observables at least, in the continuum limit, where in the standard LGT on a $N_t N_s^d$ lattice $N_t = (Ta)^{-1} \to \infty$. Numerical calculations of these thermodynamic observables can be done in principle. The search for high-temperature deconfinement will be somewhat different than for the standard LGT, where a $g_{\mathrm{crit}}(a, N_t)$ is found, from which T_c follows by scaling. For the reduced model we have as parameters besides a the physical T directly, and we always want to remain quite deep in the weak coupling regime. Note that if one uses random spatial p_μ^i instead of integrating over them [12], $N \gg (2\pi)^d$ is required. This follows if one wants good coverage of momentum space: the size $2\pi/(aN^{1/d})$ of the subcubes, which contain one p_μ^i each, must be $\ll M$, which gives the quoted condition on N even on a coarse lattice of $a \sim M^{-1}$. M is the dynamical mass scale. It seems better to try to do the integrations directly, cf. [14]. Another problem is the implementation in Monte Carlo simulations of the constraint in (3).

Up to now we have used the lagrangian formalism, but let us briefly consider the hamiltonian approach. For the EK action [9] derived from (1) by use of (6a) one immediately writes down the hamiltonian analogous to the standard one of [22]

$$H_{\mathrm{EK}} = \tfrac{1}{2} g^2 \sum_{\mu, b} E_\mu^b E_\mu^b - \frac{2}{g^2} \sum_{\mu<\nu} \mathrm{Re}\,\mathrm{Tr}\,(U_\mu U_\nu U_\mu^+ U_\nu^+)\,, \tag{15}$$

where μ, ν are, of course, space-like and where the E_μ^b, $b = 1, \ldots, N^2$, are the momenta. Also there is a Gauss law on the electric flux to constrain to the set of physical states. For the standard hamiltonian at strong coupling, Susskind [23] showed quark liberation for $T > T_c$ by inserting two widely separated test charges (source terms in Gauss law). It is not clear to me how this method might be implemented for (15). Of course, one can follow the same road as in sect. 4 with, for the partition function, $Z = \mathrm{Tr}\exp(-H_{\mathrm{EK}}/T)$, or better with (15) replaced by the conjectured quenched hamiltonian of [25]. The hypothesized rôle of "instantons" corresponding to permutations of the $P_\mu a$ variables [25], which should lead to the non-perturbative confining vacuum, is perhaps significantly reduced at high temperatures analogously to the situation in [6], leading to deconfinement.

To summarize, the reduction method of Eguchi and Kawai of the large-N lattice gauge theory may also be interesting in order to understand finite-temperature behaviour, but its non-perturbative aspects are not yet completely clear.

I thank Maria Vonk of NIKHEF for providing me with some preprints and Dr. A. Gonzalez-Arroyo for his comments on the manuscript.

Note added

Having finished this article I was informed of the following.

(i) An early hint to the symmetry breaking at weak coupling is contained in [29]. Also using the idea of twist, a quenched reduced model has been derived [30], which leads to the planar graphs of perturbation theory and which is more efficiently simulated on a computer than QEK [10]. Up to large [5] finite size effects a reduced model with "symmetric twist" [30] for $N = L^2$ might give the behaviour of the standard SU(N) LGT with $N_t = L$.

(ii) Okawa [31] also has noted the different expressions of [10] and [12] for expectation values. His numerical results of the internal energy using (5) show the absence of a large-N phase transition, contrary to the case when the prescription of [10] is used, cf. [14]. For the search for high-T quark liberation as advocated in sects. 4 and 5 this may be beneficial.

(iii) Neuberger [33] looked for high-T deconfinement using a space-like reduced model along a time-like torus of length N_t. From the $g^2_{\mathrm{crit}}(N_t)$, where $\langle L \rangle$ becomes non-zero, he appears to find a quite large T_c in terms of Λ_{MOM}, say, compared to that for $N = 2$ and 3 [5]. It is clear from the present article that in my opinion this approach is not guaranteed to mimic the finite-temperature LGT. The point is that he collapses not only the space-like but also the time-like plaquettes spatially, while keeping the time direction unreduced. At strong coupling this may be correct, and indeed gives H_{EK} of (15). But at weak coupling it is not clear that *all* loop equations are identical to the original ones, indeed see [33].

References

[1] G. 't Hooft, Nucl. Phys. B72, 461; B75 (1974) 461
[2] S. Coleman, in Pointlike structures inside and outside hadrons, ed. A. Zichichi (Plenum, NY, 1982)
[3] K.J. Wilson, Phys. Rev. D10 (1974) 2445;
F. Wegner, J. Math. Phys. 12 (1971) 2259
[4] P. Hasenfratz, in Lepton and photon interactions at high energies, ed. W. Pfeil (1981); in proc. John Hopkins Workshop, Florence June 1982, in press
[5] J. Engels, F. Karsch, H. Satz and I. Montvay, Nucl. Phys. B205 [FS5] (1982) 545;
J. Engels, F. Karsch and H. Satz, Phys. Lett. 113B (1982), 398
[6] D.J. Gross, R.D. Pisarski and L.G. Yaffe, Rev. Mod. Phys. 53 (1981) 43
[7] F.R. Klinkhamer, in The Birth of the universe, Moriond Astrophysics meeting 1982, eds. J. Audouze and J. Tran ThanhVan (Frontières, Gif-sur-Yvette, 1982);
F.R. Klinkhamer, Z. Phys. C, C16 (1982), 163;
B. Svetitsky and L.G. Yaffe, Nucl. Phys. B210 [FS6] (1982) 423
[8] C.B. Thorn, Phys. Lett. 99B (1981) 458
[9] T. Eguchi and H. Kawai, Phys. Rev. Lett. 48 (1982) 1063
[10] G. Bhanot, U.M. Heller and H. Neuberger, Phys. Lett. 113B (1982) 47;
U.M. Heller and H. Neuberger, Nucl. Phys. B207 (1982) 399
[11] M. Okawa, Phys. Rev. Lett. 49 (1982) 353

[12] D.J. Gross and Y. Kitazawa, Nucl. Phys. B206 (1982) 440
[13] S.R. Das and S.R. Wadia, Phys. Lett. 117B (1982) 228;
G. Parisi and Y.C. Zhang, Phys. Lett. 114B (1982) 319
[14] G. Bhanot, U.M. Heller and H. Neuberger, Phys. Lett. 115B (1982) 237
[15] M. Creutz and K.J.M. Moriarty, Phys. Rev. D25 (1982) 1724;
F. Green and S. Samuel, Nucl. Phys. B194 (1981) 107
[16] L. Caneshi, I.G. Halliday and A. Schwimmer, Nucl. Phys. B200 (1982) 77
[17] S. Samuel, Nucl. Phys. B205 [FS5] (1982) 221
[18] M. Creutz, Phys. Rev. D21 (1980) 2308; Phys. Rev. Lett. 45 (1980) 313;
M. Creutz and K.J.M. Moriarty, Phys. Rev. D26 (1982) 2166
[19] S. Weinberg, Phys. Rev. D9 (1974) 3357;
C.W. Bernard, Phys. Rev. D9 (1974) 3312;
L. Dolan and R. Jackiw, ibid 3320
[20] L.D. McLerran and B. Svetitsky, Phys. Rev. D24 (1981) 450
[21] V.F. Muller and W. Rühl, Ann. of Phys. 133 (1981) 240
[22] J. Kogut and L. Susskind, Phys. Rev. D11 (1975) 395
[23] L. Susskind, Phys. Rev. D20 (1979) 2610;
A.M. Polyakov, Phys. Lett. 72B (1978) 477
[24] E. Seiler, Gauge theories as a problem of constructive quantum field theory and statistical mechanics (Springer, Berlin, 1982)
[25] H. Neuberger, Phys. Lett. 119B (1982) 179
[26] T. Chen, C. Tan and X. Zheng, Phys. Lett. 116B (1982) 419
[27] A.A. Migdal, Phys. Lett. 116B (1982) 425
[28] G. Maiella and P. Rossi, preprint CERN-TH 3416
[29] A. Gonzalez-Arroyo, J. Jurkiewicz and C.P. Korthals-Altes, in Proc. 1981 Freiburg NATO Summer Institute (Plenum, New York, 1982)
[30] A. Gonzalez-Arroyo and M. Okawa, Phys. Lett. 120B (1983), 174; BNL preprint (November 1982)
[31] M. Okawa, Phys. Rev. Lett. 49 (1982) 705
[32] H. Neuberger, private communication; IAS preprint (December 1982)
[33] F.R. Klinkhamer, preprint (February 1983)

CHAPTER 10

MORE ON THE EGUCHI-KAWAI REDUCTION AT FINITE TEMPERATURE

F.R. Klinkhamer

Leiden Observatory

Postbus 9513

2300 RA Leiden

The Netherlands

Abstract:

I consider the (quenched) Eguchi-Kawai reduction in d dimensions only of the SU($N \to \infty$) gauge theory on a d+1 dimensional lattice. Then I discuss reduced models at finite temperature both with or without reduction of the link variables of the remaining dimension.

February 1983, revised May 1983.

1. Introduction

Eguchi and Kawai [1] have shown that the SU($N \to \infty$) lattice gauge theory (LGT) with the Wilson action [2] on an infinite lattice may be reduced to a similar model *without* space-time dependence. Later several authors [e.g. 3-8] gave the necessary modifications of this reduced model so that is valid at weak couplings too. The problem how finite temperature might be incorporated in the single point model I discussed in [9]. At the end of that article I expressed some doubts on another possible finite temperature reduced model, which was studied by Neuberger [10]. In this article I elaborate these doubts, dispel them and learn how to improve the single point model [9]. Also this may be of importance if one tries to construct a reduced Hamiltonian [11,12] by taking the continuum limit on one direction of the Euclidean theory [13].

The content of this paper is as follows. In sect. 2 I have to repeat the proof [1] of the Eguchi-Kawai reduction in some detail, so that it is easy to see in sect. 3 how the Schwinger-Dyson loop equations are reproduced if one would like to reduce the d+1 dimensional LGT in d dimensions only. If this unreduced ("time") dimension has length $N_t a_t$, where a_t is its lattice spacing, this would be precisely the model used in [10] to study the large-N behaviour at temperature $T = (N_t a_t)^{-1}$. In sect. 4 I discuss the implications of a (solved) problem of sect. 3 for the study of a truly single point model and propose three possible solutions. Finally, section 5 contains some further discussion.

2. Eguchi-Kawai reduction

Consider a D-dimensional infinite hypercubic lattice with spacing $a = 1$. On it define the gauge theory with group $G = SU(N)$ by the partition function

$$Z = \int \prod_{x,\mu} dU_\mu(x) \quad \exp\left[\beta\, S_W\right] \tag{1a}$$

with the Wilson action [2]

$$S_W = \sum_{x,\mu\neq\nu} \mathrm{tr}\,(U_\mu(x)\, U_\nu(x+\mu)\, U_\mu(x+\nu)^\dagger\, U_\nu(x)^\dagger). \tag{1b}$$

The notation is standard with $U_\mu(x)\ \epsilon G$ the variable on the link in the μ-direction between the sites x and x+μ, and with dg, gϵG, the normalized Haar measure over the group G. The large β expansion gives the relation $\beta = g^{-2}$ with the bare coupling constant g(a).

A single point model may be defined by

$$Z_{EK} = \int \prod_\mu dU_\mu \ \exp\left[\beta \sum_{\mu\neq\nu} \mathrm{tr}(U_\mu U_\nu U_\mu^\dagger U_\nu^\dagger)\right]. \tag{2}$$

Eguchi and Kawai [1] proved in the limit $N \to \infty$ that the reduced model (2) is equivalent to the standard lattice gauge theory (LGT) of (1). The translation rule

$$T:\ U_\mu(x) \to U_\mu \tag{3}$$

must be used to find in the EK model the counter part of the expectation values $W(C) = \langle \mathrm{tr} \prod_C U_\mu(x)\rangle$, where C is some closed path. Their proof is that the observables W(C) in the standard model and the translated $W_{EK}(C)$ in the reduced model satisfy the same set of Schwinger-Dyson (S-D) equations. Two properties are used in this proof:

(i) large-N factorization $\langle \mathrm{tr}(\prod_{C'} U)\ \mathrm{tr}(\prod_{C''} U)\rangle =$

$\langle \mathrm{tr}(\prod_{C'} U)\rangle\langle \mathrm{tr}(\prod_{C''} U)\rangle + O(1)$, and

(ii) the $U(1)^D$ symmetry of (2) $U_\mu \to e^{i\phi_\mu} U_\mu$ (no sum over μ).

At weak coupling this symmetry is spontaneously broken [3] (there are only D links, but the limit $N \to \infty$ gives an infinite number of degrees of freedom) and the EK model (2) must be modified in order to keep equivalence with the standard theory [2]. For the moment forget about these modifications.

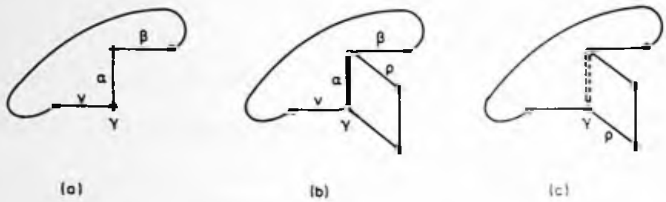

Fig. 1. Curves for the Schwinger-Dyson equation (4). (a) Shows the original curve C and (b) and (c) the curves C'_ρ and C''_ρ, respectively, where a $\alpha\rho$-plaquette has been added. In (b) the order of the vertices is $(..,y-\nu,y,y+\alpha,y+\alpha+\rho,y+\rho,y,y+\alpha,y+\alpha+\beta,..)$. In (c) the dashed lines indicate a cancellation $U_\alpha(y)^\dagger\ U_\alpha(y)=1$ and the order of the vertices is $(..,y,y+\rho,y+\rho+\alpha,y+\alpha,y+\alpha+\beta,..)$. Only near y has the curve been drawn specifically, the rest is arbitrary.

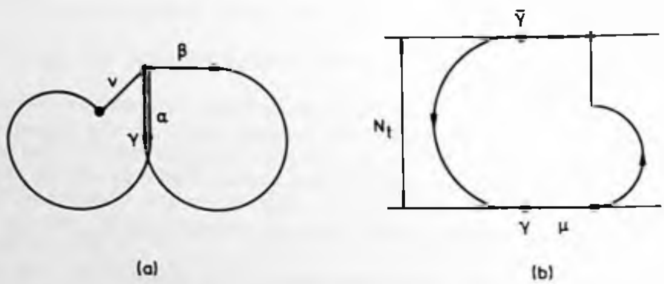

Fig. 2. Example of curves which lead to source terms in the Schwinger-Dyson equations. (a) This curve gives a source term in (4), where the two paths are $(y,y+\alpha,y+\alpha+\beta,..,y)$ and $(y,y+\alpha,y+\alpha+\nu,..,y)$. (b) If the time direction is finite (N_t) and has periodic boundary conditions, these make this curve intersect itself (in opposite direction in this example) and a source term arises in (8). These source terms probably are important for the phenomenon of quark deconfinement at a critical temperature T_c.

The S-D equations in the two models are derived as follows; see the original paper [1] for the missing details, the example paths used and illustrated (figs. 1,2) here have the same notation. Consider for a closed path $C = (x, x+\mu, \ldots, y-\nu, y, y+\alpha, y+\alpha+\beta, \ldots, x-\sigma, x)$ the expectation value $\overline{W}(C, T_j, y, \alpha)$, where a group generator T_j is inserted before $U_\alpha(y)$. For the change of variables $U_\alpha(y) \to (1+i\varepsilon\, T_j)U_\alpha(y)$ it follows from invariance (order ε terms vanish) and summation over j that these S-D equations hold

$$N\, W(C) + \frac{N}{\lambda} \sum_{\rho\neq\alpha} W(C'_\rho) - \frac{N}{\lambda} \sum_{\rho\neq\alpha} W(C''_\rho) = S(C,y,\alpha)\; W(C'_{yy})W(C''_{\bar{y}\bar{y}}). \tag{4}$$

As usual $\lambda = N/\beta$ is to be held fixed in the limit $N \to \infty$. The paths of the lhs of (4) are shown in figs. 1a-c. The curves C_ρ follow when the change of variables is made on the $U_\alpha(y)$ contained in the plaquettes of the action (1b). The rhs of (4) is a possible source term, where large N factorization has been used already.

$S = 0$ or $-1/+1$ if the $(y,y+\alpha)$ link is traversed only once or twice in the same/ opposite direction, respectively. In the latter case there arise from the j summation the traces over two closed curves C'_{yy}, $C''_{\bar{y}\bar{y}}$, where $\bar{y}=y$ ($\bar{y} = y+\alpha$) for the same (opposite) sense of direction. The possible y and $\bar{y}$ difference comes from a cancellation $U_\alpha(y)$ and $U_\alpha(y)^\dagger$ in the trace and is trivial here but not entirely in the next section. Fig. 2a illustrates the case $\bar{y} = y$.

With the translation (3) $W(C) \to \langle \mathrm{tr}\,(U_\mu \cdots U_\nu\, U_\alpha\, U_\beta \cdots)\rangle_{EK}$ one derives the same equation (4) for the reduced model, but with extra source terms. These additional terms arise if in C a direction is traversed twice, say $\mu = \alpha$ (similarly if in the opposite direction). In the standard model $U_\mu(x)$ and $U_\mu(y)$ are different, but not after the translation (3) where they become the same U_μ and would thus give a source term

$$\langle \mathrm{tr}(\prod_{L'_1} U_\mu)\; \mathrm{tr}(\prod_{L'_2} U_\nu)\rangle_{EK}\,. \tag{5}$$

In (5) the products of the U_μ are the translations (3) of the $U_\mu(x)$ on the open lines $(x, x+\mu, \ldots, y-\nu, y)$ and $(y, y+\mu, y+\mu+\beta, \ldots, x-\sigma, x)$. In the standard model gauge invariance would give $\langle \mathrm{tr} \prod_{L_1'} U_\mu(x) \rangle = 0$. In the reduced model (5) also vanishes, but now by the two properties mentioned above: Factorization splits (5) and the symmetry $U_{\mu_1} \to e^{i\psi} U_{\mu_1}$ and $U_{\mu_2} \to e^{i\psi} U_{\mu_2}$ make each part vanish. Here μ_1, is a direction where the number of U_{μ_1} and $U_{\mu_1}^\dagger$ of the product L_1' are unequal, analogously for μ_2. The reader may look for an example at fig. 3. The links $(y, y+\nu)$ and $(z, z+\nu)$ in the path shown in (a) give in the loop equation of the reduced model a source term (5) with "supports" as given in (b) and (c). But the U(1) symmetry of the reduced model for the encircled link U_t makes this source term vanish.

This means that the W(C) of the standard theory (1) and W(C) of the EK reduced model (2,3) have the same S-D equations. As said at weak coupling the EK model must be changed so that extra source terms as (5) still vanish.

3. Partial reduction

Consider LGT(1) on a $N_t \times (\infty)^d$ lattice, where the "time" direction, index t, is finite and has periodic boundary conditions. The lattice may have different spacings $a_t(a_s)$ and couplings $\beta_t(\beta_s)$ in the action (1b) for time and space-like directions [14]. A reduction in d = 3 space-like dimensions only would give the following model,

$$Z_{pEK} = \int \prod_{n=1}^{N_t} \left[dU_t(n) \prod_{\mu=1}^{d} dU_\mu(n)\right] \exp\left["\beta" \; S_{pEK}\right] \tag{6a}$$

$$"\beta" \; S_{pEK} = \sum_{n=1}^{N_t} \Big[\beta_s (a_t/a_s) \sum_{\mu\neq\nu=1}^{d} \mathrm{tr}(U_\mu(n) U_\nu(n) U_\mu^\dagger(n) U_\nu^\dagger(n)) + \beta_t (a_s/a_t) \times$$

$$\sum_\mu \mathrm{tr}(U_\mu(n) U_t(n) U_\mu^\dagger(n+1) U_t^\dagger(n) + \mathrm{H.C.})\Big] \;. \tag{6b}$$

It is understood that at large β_s the necessary modifications have been applied on the spatial parts of (6). Specifically $U_\mu(n)$ in (6b) may be replaced by $U_\mu(n)D_\mu$, with D_μ independent of n, and ΠdU_μ by the constrained integration, see [4].

The S-D equations for the standard model are the same, up to the trivial replacements "β", as in (4) but with extra source terms $s(c,y,\alpha,N_t)$ <....> from the time periodicity. These were also found by Gocksch and Neri [15]. Consider the curve of fig. 2b, which is intersecting by the periodicity condition in the time direction, and gives a source term $\langle tr(\prod_{L_1} U_\mu(x)\; tr(\prod_{L_2} U_\mu(x))\rangle$, with L_1 the line $(y,y+\mu,\ldots.\bar{y})$ via the right and L_2 the line $(\bar{y},..y)$ via the left. I will discuss these terms in sect. 4.

The S-D equations for the model (6) are the same as for the unreduced model. Spurious source terms (5) from the identification of the space-like links in each time layer n vanish as in the preceding section. Now consider even a further simplification (6') with $U_\mu(n)$, $\mu \neq t$, the same for all n. Then a special subset S of curves may cause some trouble, namely those where the separate parts of the source terms (5) with supports $L_{1,2}$ can be forced to zero *only* by the phase symmetry in time like links. This means that the subset S contains links $(y,y+\nu)$ and $(z,z+\nu)$ with $z = y+k\,e_t$, where k is a non-zero integer inequal to a multiple of N_t and e_t is unit vector in the time direction. Fig. 3 gives an example, whereas the curve $(y,y+\nu,y+\nu+t,\ y+t,\ y+t+\mu = z',\ y+t+\mu+\nu,\ y+\mu+\nu,\ y+\mu,y)$ is not included in S, because the source term (5) can be made to vanish by $U_\mu \rightarrow e^{i\phi}U_\mu$ or by the implied quenching modification (see below (6)).

Thus at large enough β_t, when the $U_t \rightarrow e^{i\psi}U_t$ symmetry is broken, the S-D equations derived from the change of variables $U_\nu(y) \rightarrow (1+i\varepsilon\, T_j)U_\nu(y)$ might be different for all curves in S. Note that (6) has a *local* symmetry $U_\mu(n) \rightarrow \Omega(n)\; U_\mu(n)\Omega(n)^{-1}$ and $U_t(n) \rightarrow \Omega(n)\; U_t(n)\Omega(n+1)^{-1}$, which would make terms as (5) for curves in S vanish $(\Omega(z) \neq \Omega(y))^{F1}$. Hence only the pEK model (6) has the correct S-D equations including the source terms from the N_t periodicity. In

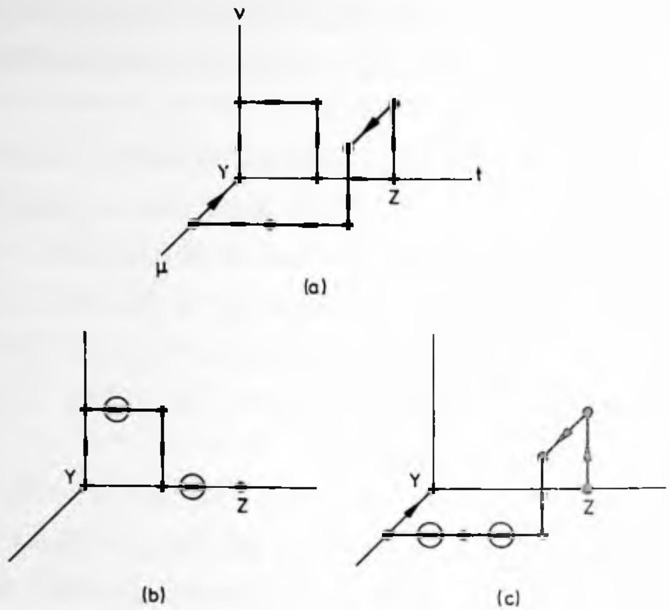

Fig. 3. Example of a curve C (a) that may lead to an extra source term (5) in the Schwinger-Dyson equations for the partially reduced model (6'), where t is the unreduced dimension. This source term (5) has as support the two open lines $L'_{1,2}$ shown in (b) and (c). Only if there is a phase symmetry on U_t can the uncircled links make this source term vanish.

the next section I consider single point models where the spurious source terms from S do cause problems.

4. Further reduction at finite temperature

It is well known [14] that the $N_t \times \infty^d$ theory of sect. 3 describes the SU(N) gauge theory with a lattice regularization at finite temperature $T = (N_t a_t)^{-1}$. The reduced model (6) has been used to study numerically [10] this system at large-N. Also in that paper it was shown analytically that the ground-state spectrum of the eigenvalues $\exp[i\theta_j]$, $j = 1, .., N$, of $\prod_n U_t(n)$ is

$$\beta_t \to 0: \theta_j = \frac{2\pi}{N}\,(j - \frac{N+1}{2}),$$

$$\beta_t \to \infty: \theta_j = \frac{2\pi}{N}\,k,$$

with k an arbitrary integer and $\Sigma\theta = 0$ (mod 2π) because SU(N) is the gauge group. Note that similar calculations of $\langle \mathrm{tr} \prod_n U_t(n)\rangle$ have been done in [16]. The Z(N) symmetry $\theta_j \to \theta_j + 2\pi k/N$ is thus (un)broken at (small) large β_t. Recall that the order of eigenvalues is $e^{i\theta_j}$ is arbitrary. This would be the same whether or not the space-like dimensions are reduced.

In the standard LGT the $\beta^{crit}(N_t)$ of the Z(N) phase transition is related [16] to the deconfinement transition at temperature T_c, which is extracted from the renormalization group scaling of $\beta^{crit}(N_t)$. From his Monte Carlo simulations, which gave $\beta^{crit}_{pEK}\,(N_t)$, Neuberger [10] claimed the physical deconfinement temperature T_c at large-N to be $12 - 50\ \Lambda_{mom}$. What would happen if we reduce the torus model (6) further by setting $U_\alpha(n)$, $\alpha = t, 1, .., d$, equal for all n, which I will call the replica model (rEK)?

First let me specific on what rEK is. Basically it is a single point model with d+1 matrix variables: the integration measure of (6) is changed to

$$\prod_n \prod_{\alpha=t,1,..,d} dU_\alpha(n) \prod_{n\neq m} \delta(U_\alpha(n)-U_\alpha(m)), \tag{7a}$$

and the translation rule (3) is modified

$$T': U_\alpha(n,x_1,..,x_d) \to U_\alpha(n), \tag{7b}$$

where n becomes the label of the replica. In this way the distinction is more transparent between the wanted and unwanted source terms below. But it is clear that now the order parameter $\langle \mathrm{tr} \prod_{n=1}^{N_t} U_t(n)\rangle_{rEK} = \langle \mathrm{tr}\, U_t^{N_t}\rangle_{EK}$ gets a non-zero value for any $\beta > \beta_{EK} \sim 0.15\, N$ [21], where the $U(1)^{d+1}$ symmetry of the single-point EK model gets broken. [Notice that the Z(N) symmetry as discussed above is part of this U(1) symmetry on the time link U_t.] This will be independent of N_t and we can never obtain from rEK a non-trivial behaviour of $\beta^{crit}(N_t)$, which is to be linked to the physical high-temperature deconfinement transition. All this is not surprising, because as said this replica model is equal to single point EK model. Somehow we must improve the completely reduced model (7) so that it becomes equivalent to the standard theory also for $\beta > \beta_{EK}$, including the non-trivial dynamics around $\beta^{crit}(N_t)$.

The standard S-D equations of the unreduced theory are

$$\mathrm{lhs}(4) = s(C,y,\alpha)\langle\ldots\rangle + s(C,y,\alpha,N_t)\langle\ldots\rangle\ , \tag{8}$$

see (4) and sect. 3, where the second source term from the time periodicity (e.g. fig. 2b) was found. The second term on the rhs of (8) may be non-zero in the deconfined phase [14,16]. The reduced model (7) has the same equations (8) except for a spurious term on the rhs

$$s(C,y,\alpha)\langle\ldots\rangle_{rEK} \text{ if } C\varepsilon S \tag{9}$$

as found in the preceding section. The sought-for cure of rEK should kill terms like (9) but not those like the second of the rhs of (8). There are (at least) two methods along the ideas of [4,5] and [6], which appear to achieve this, so that their loop equations are equal to those (8) of the LGT.

1) Modify somehow the ground state of model (7) at large β_t; $\{\theta_t{}^j(n)\} = \{\theta_t{}^j(n')\}$ for $n \neq n'$, where $e^{i\theta_t{}^j(n)}$ are the eigenvalues of $U_t(n)$. The θ_t^j, j=1, N, must have the following spectrum: cover the interval $[0,2\pi]$ uniformly, but bunched together at values $2\pi k/N_t$ for integers k. For this to be possible one must have $N = mN_t$, with m a positive integer (see the remark below). With these "momenta" θ_t we do the momentum quenching prescription of [4] in the t direction. Specifically replace in (6,7a)

(i) $U_t(n)$ by $U_t\ D_t$, $D_t^{ij} = \delta^{ij} \exp[i\theta_t^j]$

(ii) $\int \prod_n dU_t(n)$ by $\int dU_t \times C(U_t, D_t)$,

where

$$D_t{}^{ij} = \delta^{ij}\ e^{i\theta_t^j}$$

and C is the constraint on the eigenvalues of U_tD_t of [4].

But (iii) now we do *not* integrate over the momenta θ_t^j at the end of the calculation. There is a $Z(N_t)$ symmetry $U_t \to e^{i2\pi k/N_t}\ U_t$, which kills (9) indeed, while leaving $\mathrm{tr}U_t^{N_t}$ invariant. With the condition $N = mN_t$ the model will be correct (good momentum coverage) in the limit $N_t \to \infty$. Note that the interval $2\pi/N_t$ of [9] has been recovered. Now the remark on the condition $N = mN_t$. This condition differs for the quenching prescriptions of [3] and [4], where the structure of the integrations over the "momenta" θ_μ in the expectation value of operator O is respectively [cf. 9]

$$\int \prod_{\mu,j} d\theta_\mu^j \prod_\mu \prod_{i>j} \sin^2((\theta_\mu^i - \theta_\mu^j)/2) \quad \langle O\rangle_\theta \tag{a}$$

$$\int \prod_{\mu,j} d\theta_\mu^j \quad \langle O\rangle_\theta . \tag{b}$$

With our random momenta the measure of (a) would give zero for coinciding angles ($N>N_t$), so that one should take $N=N_t$. The weak coupling limit $N_t \to \infty$ then drags the N along, a somewhat strange situation. For the prescription of [4] one may have $N > N_t$, but one should then correct for the extra weight given to zero loop momenta in the Feynman diagrams. Perhaps in both cases it is easiest to have $N=N_t$. Note that I took only the time-like momenta θ_t random. If all momenta are taken at random the numbers work out differently, for example if the spacing of θ_μ is $2\pi/N_s$ one has $N = m\, N_t N_s^d$.

2) In order to disfavor [6] the broken symmetry state vs. the symmetric one the following action could be used in (6a, 7a)

$$"\beta"\, S_{\text{mod rEK}} = "\beta"\, S_{\text{rEK}} - \beta_t^2\, N^{-1} \sum_{\substack{L_i = \text{part } C \\ C \varepsilon S}} c_L \left| \mathrm{tr}\, (\prod_{L_i} U_\mu) \right|^2, \tag{10}$$

where $c_L > 0$ are arbitrary constants and S is the set of troublesome curves, which split in the lines L_1 and L_2 (see sect. 3 and fig. 3). This is only a formal solution of the problem. To get something more practical one could follow the consistent field (cf) approach of [6] to replace the $\sum_L$ term of (10). This replacement is exact in the limit $\beta \to \infty$ and it is hoped that the same cf addition to the action S_{rEK} leads to the desired symmetric ground state also at intermediate β. The cf approach is to replace the link variables by their diagonalized form D_μ, calculate the trace of the product of D_μ's over L, and finally sum over the open lines L. If we do this also for the space-like dimensions [6] we get for the second term of rhs of (10)

$$-\beta^2 N \sum_{i \neq j} \prod_{\mu \neq 1} \delta(\theta_\mu^i - \theta_\mu^j)\, F(\theta_t^j, N_t) \tag{11}$$

where F is a messy expression, because in the set of troublesome curves of S we excluded those like fig. 2b, whose source terms we do not want to kill. Also for simplicity we went to a symmetric lattice $\beta_t = \beta_s = \beta$.

Finally let me mention another possibility. The twisted (TEK) model [7], with the trace in the action of (2) multiplied by a fixed twist $Z_{\mu\nu} \varepsilon Z(N)$, looks very promising, especially for numerical simulations. For $N=L^2$ a particular twist can be choosen so that perturbation theory gives the planar Feynman graphs and that the expectation values of open loops from inside a L^4 box vanish. Hence the 1/N corrections are of order of the finite size correction [14] of the standard LGT on a L^4 lattice. Perhaps it is possible to construct another twist so that this box becomes of size $N_t \times N_s^d$, $N_t < N_s$, which would approximate the standard large-N theory at temperature $T = (N_t a_t)^{-1}$ and also be useful in the construction of a TEK Hamiltonian. But that this can be done generally is not certain, because the L^4 arises as a generalization of the dimensionality (2^4) of the Clifford algebra of Dirac matrices.

The problem is to find a solution to $\Omega_\mu \Omega_\nu = Z_{\nu\mu} \Omega_\nu \Omega_\mu$, $Z_{\nu\mu} \equiv \exp(2\pi i n_{\nu\mu}/N)$, with such a twist that the Ω_μ saturate the action and such that "open loops" vanish only inside a box $N_t N_s^d$, $N_t < N_s$ and $d = 3$. A class of solutions with the first two properties is given by the ansatz [20]

$$\Omega_\mu = P^{a_\mu} Q^{b_\mu} R^{c_\mu} S^{d_\mu} \tag{12}$$

with arbitrary integers a,b (c,d) modulo p(q).

The SU(N) matrices S P Q R are given in [20], whose notation I follow. Because $PQ = a_p QP$ and $RS = a_q SR$, $a_n \equiv \exp(2\pi i/n)$ and $N = pq$, while the other commutators vanish, it is straightforward to calculate the twist $n_{\nu\mu}$ for the ansatz above. Note that for any integer k $\mathrm{tr}\Omega^k$ = (phase factor) x $\mathrm{tr}(P^{ka_\mu} Q^{kb_\mu} R^{kc_\mu} S^{kd_\mu})$. A necessary condition to have tr AB=0 is AB=a BA with a≠1. [For A traceless and B the unit matrix we see that the opposite is not true.] For the ansatz (12) to give vanishing traces of open loops one can write necessary conditions, c which I consider first those generated by the P and Q's. For a part with N_μ μ-links and N_ν ν-links this condition is (no sum over u and ν)

$$(N_\mu a_\mu + N_\nu a_\nu)(N_\mu b_\mu + N_\nu b_\nu) \neq 0 \pmod{p} \,. \tag{13a}$$

For the three other classes of paths one has similar equations, up to

$$\Big(\sum_{\mu=t\ldots d} N_\mu a_\mu\Big)\Big(\sum_\mu N_\mu b_\mu\Big) \neq 0 \pmod{p} \,. \tag{13b}$$

Other equations hold with a_μ and b_μ replaced by c_μ and d_μ. All this is quite complicated, but it is easy to construct an example for $N_t < N_s$ by choosing p > 1 prime and by use of the following lemma:
if a positive integer n has prime factors < p then n≠0 (mod p).

Now the left-hand sides of conditions (13) are less than $(1+d)^2 N_s^2$ AB $\equiv$ M, with $A \equiv \max_\mu (a_\mu)$ and B similarly, and suppose all are $\geq 1^{F2}$. For $M < p$ the lemma shows that the conditions (13) are fulfilled; one could do better, of course. But we want slightly more, namely tr $\Omega_t^{N_t} \neq 0$. An example that solves this problem goes as follows, Choose $q = N_t$, hence $N = pN_t$, and for (12) $\Omega_t = S$. This achieves the required non-vanishing of the trace, but because $a_o = b_o = c_o = 0$ the *necessary* conditions (13) may be violated and one must check that the traces of open loops inside the $N_t N_s^3$ box *still* vanish. For tr Ω_t^k, $k < N_t$, this happens automatically. Some numerical values will show that this construction is not very practical and one should find explicit solutions to the whole set of equations of which (13) is a part: p = 1601, then N_s is at least 10 (if A and B = 1) so that for N_t = 3,5 one needs N = 4803, 8005. But in this example N_s probably is of the order of $(N^2/N_t)^{1/3}$, so that a more practical twist might be constructed for $N \sim 100$ with $N_t < N_s$ and $N_s \sim 20/N_t^{1/3}$. After the present article had been finished this program was realized [22].

5. Discussion

I have argued that unlike model (6) the completely reduced one (7) is not to be trusted at large β_t to mimic the large-N theory at finite temperature. This is because the Schwinger-Dyson loop equations of this model get spurious source terms if the loop belongs to the class S, of which an example is given in fig. 3. Also I gave two or three possible remedies, where the problem was to get rid of these spurious source terms but *not* of those very similar source terms arising from the time periodicity (fig. 2b), which are crucial [17] to give the phenomenon of quark deconfinement at a high temperature T_c. Note that these terms vanish for $T<T_c$, which implies that in leading order of 1/N the physical content is temperature independent, see also [15].

It is possible to construct [13] from the Euclidean theory by use of the transfer matrix and the limits $\beta_t \to \infty$ and $a_t \to 0$ the Hamiltonian of the theory. Indeed the authors of [11, 12] have done that starting from the partially reduced model (6), which now has been proved to be correct (sect. 3). Only along these lines could one obtain, say, the glueball spectrum [18]. This is because in the Euclidean theory the connected part of the Green's function vanishes in leading order of $1/N$ by factorization [4]. A transparant way to understand why in the quenched EK models these subleading terms are inherently unobtainable follows if one considers the fluctuations around the master field [19].

Just as for the original reduction by Eguchi and Kawai [1] our arguments for the case of finite temperature were based on the loop equations. But notice that many questions such as (which S-D loop equations (together with boundary conditions) determine the physical content of "the" large N theory?" are yet unanswered. Indeed the loop equations derived [4] for the momentum-quenched reduced model are slightly different than (4), but perhaps they have a smoother continuum limit. In [9] I gave the most naive finite temperature reduced model, which equals the d+1 dimensional one of the momentum-quenching prescription [4,5], but with the time-like momentum integral replaced by the standard Fourier sum with interval $2\pi/a_t N_t$. Because of the quenching in the timelike direction the loop equations for this model do not have the spurious source terms, but still those from curves as in fig. 2b[F3]. Thus, while this model is expected to describe the perturbative regime (or high temperatures) reasonably well, its validity in the non-perturbative regime (e.g. the deconfinement transition) may be correct also. The same should hold for the simpler model of sect. 4. A problem that remains is to select between the different modifications [3-7] of the EK model, which hopefully are equivalent at $N = \infty$, for the best one at large but finite N, where they are inequivalent [cf. 9,22].

Acknowledgements

This work resulted from a visit to the USA, which was generously supported by the Netherlands Organization for the Advancement of Pure Research (ZWO) and the Lorentzfonds. I thank Dr. P. Hut and Prof. J. Bahcall for the hospitality at the Institute for Advanced Study in Princeton. There Dr. H. Neuberger kindly showed me his numerical results [10].

Note added

I have heard that just as in sect. 4 D. Gross and L. Yaffe tried to do something with the time-like momenta.

Footnotes

F1. I thank H. Neuberger for pointing this out to me.

F2. Admittedly I am a bit cavalier about zeros in the coefficients $a_\mu, \ldots, d_\mu$. But recall that (13) and others are only necessary conditions and a subset of them might be enough. I hope to consider the problem in more detail in a future publication, see [22].

F3. One has $N^{-1}\mathrm{tr}\, D_t^k = N^{-1}(e^{2\pi i\, kn_1/N_t} + \ldots + e^{2\pi i\, kn_N/N_t})$, which after summation $N_t^{-1} \sum_{n_i=1}^{N_t}$ for all n_i gives 1(0) for $k \neq 0$ (= 0) mod N_t or, in other words, a (non-)zero expectation value for these (closed) open loops. [This corrects a remark in sect. 4 of ref. 9]. Note that in our model of sect. 4 for $N = mN_t$ we fix $n_i = i$, which achieves the same.

References

[1] T. Eguchi and H. Kawai, Phys. Rev. Lett. 48, 1063 (1982)

[2] K.J. Wilson, Phys. Rev. D10, 2445 (1974)

[3] G. Bhanot, U.M. Heller and H. Neuberger, Phys. Lett. 113B, 47 (1982);
U.M. Heller and H. Neuberger, Nucl. Phys. B207, 399 (1982)

[4] D.J. Gross and Y. Kitazawa, Nucl. Phys. B206, 440 (1982)

[5] G. Parisi and Y.C. Zhang, Phys. Lett. 114B, 319 (1982);
S.R. Das and S.R. Wadia, Phys. Lett. 117B, 228 (1982)

[6] T. Chen, C. Tan and X. Zehng, Phys. Lett. 116B, 419 (1982)

[7] A. Gonzalez-Arroyo and M. Okawa, Phys. Lett. 120B, 174 (1983);
preprint BNL 32393 (November 1982) to be publ. Phys. Rev. D

[8] J. Greensite and M.B. Halpern, Nucl. Phys. B211, 343 (1983);
J. Alfaro and B. Sakita, Phys. Lett. 121B, 339 (1983)

[9] F.R. Klinkhamer, Nucl. Phys. B218, 32 (1983)

[10] H. Neuberger, preprint IAS (December 1982)

[11] Y. Kitazawa and S.R. Wadia, Phys. Lett. 120B, 377 (1983)

[12] H. Neuberger, Phys. Lett. 119B, 179 (1982)

[13] M. Creutz, Phys. Rev. D15, 1128 (1977)

[14] J. Engels, F. Karsch, H. Satz and I.Montvay, Nucl. Phys. B205 (FS5), 545 (1982)

[15] A. Gocksch and F. Neri, Phys. Rev. Lett. 50, 1099 (1983)

[16] N. Weiss, Phys. Rev. D25, 2667 (1982);
J. Polonyi and K.Szlachanyi, Phys. Lett. 110B, 395 (1982)

[17] B. Svetitsky and L.G. Yaffe, Nucl. Phys. B210 (FS6), 423 (1982)

[18] H. Levine and H. Neuberger, Phys. Rev. Lett. 49, 1603 (1982)

[19] G.C. Rossi and M. Testa, preprint CERN-TH-3480 (December 1982)

[20] Y. Brihaye, Phys. Lett. 122B, 154 (1983)

[21] M. Okawa, Phys. Rev. Lett. 49, 353 (1982)

[22] F.R. Klinkhamer and P. van Baal, preprint (May 1983)

HOT TWISTS FOR THE SINGLE-POINT MODEL OF LARGE-N QCD

Frans R. Klinkhamer

Leiden Observatory
P.O. Box 9513
2300 RA Leiden
The Netherlands

Pierre van Baal

Institute for Theoretical Physics
Princetonplein 5, P.O. Box 80.006
3508 TA Utrecht, The Netherlands

Abstract

We construct two types of twists for the $SU(N \to \infty)$ Twisted-Eguchi-Kawai model, which mimic a periodic boundary condition in the temporal direction only and over an arbitrary extent N_o. In this way we introduce finite temperature ($T = N_o^{-1}$ in lattice units) in the single-point model. In weak coupling one gets the correct planar expansion.

May 1983

submitted to Nucl.Phys. B

1. Introduction

Recently much attention has been given to the reduction [1,2] of the Wilson SU($N \to \infty$) lattice gauge theory [3] to a single-point model. Especially a model with twisted boundary conditions looks very promising for numerical [5] and analytic [6] calculations. Also the problem how to incorporate finite temperature in (partially) reduced models has been considered [7,8,9]. In this paper we present general twists, which achieve this for a single-point model and which we will call Hot Twists. These can be used in Monte Carlo simulations to study the large-N deconfinement transition at the critical temperature T_c. Preliminary results indicate that $T_c \sim 3\sqrt{\sigma}$, where σ is the string tension. We extracted from [5] $\Lambda_L \sim 3.10^{-3}\sqrt{\sigma}$ and from [8] $T_c \sim 10^3\Lambda_L$ (*). This is not far from the value for SU(2) and SU(3) [10]: $T_c \sim 0.5\sqrt{\sigma}$. So one might suspect that the mechanism of confinement at $N \to \infty$ is *not* very different from that at $N = 2,3$, contrary to the claim of e.g. ref. 11.

The program of this paper is as follows. In sect. 2 we briefly discuss the reduction procedure at finite temperature. In sect. 3 we show that non-perturbatively, i.e. at the level of the loop equations, there is agreement for $N \to \infty$ between the one point model with the Hot Twists and the Wilson theory at finite temperature. In sect. 4 we construct the appropriate Hot Twists. Sect. 5 deals with the weak coupling limit, in which one retrieves the planar expansion at finite temperature. Finally in sect. 6 we give some further discussion of our Hot Twists, notably their potential value for Monte Carlo simulations.

(*) We *assumed* equal Λ_L for the twisted [5] and the quenched [8] Eguchi-Kawai model.

2. Reduction

We start from the pure SU(N) gauge theory on a hypercubic lattice (spacing a = 1) in euclidean space-time with dimension 4, where indices are denoted by greek symbols ($\mu,\nu,\ldots$). If necessary we distinguish the time-(space-) direction by $\mu = o$ ($i,j,k\ldots = 1,2,3$). On a lattice of size $\Lambda = N_o \times \Lambda_s$ with periodic boundary conditions in the time direction the partition function Z_W and the action S_W [3] used are(*):

$$Z_W = \int \prod_{\mu,x\in\Lambda} dU_\mu(x)\exp(-\beta S_W) \tag{2.1}$$

$$S_W = \sum_{\mu\neq\nu,x\in\Lambda} \mathrm{Tr}\left(1-U_\mu(x)U_\nu(x+\hat{\mu})U_\mu(x+\hat{\nu})^\dagger U_\nu(x)^\dagger\right) \tag{2.2}$$

This gives the gauge theory at equilibrium temperature $T = N_o^{-1}$, as long as the extent of the spatial directions is much larger than N_o. Eguchi and Kawai [1] showed that for $N \to \infty$ the theory (2.1) is equivalent to the reduced single-point model under the reduction:

$$R:\ U_\mu(x) \to U_\mu \quad . \tag{2.3}$$

For weak-coupling ($\beta \to \infty$), modification of (2.3) proved to be necessary, first done with a quenching procedure, see [2] and ref. therein. Later Gonzalez-Arroyo and Okawa [4,5] proposed the more elegant approach of introducing appropriate twist $Z_{\mu\nu} \in Z_N$ (the discrete centre of SU(N)).

This Twisted-Eguchi-Kawai (T.E.K.) model is defined by:

$$Z_{TEK} = \int \prod_\mu dU_\mu \exp\left(-\beta\ S_{TEK}\right) \tag{2.4}$$

$$S_{TEK} = \sum_{\mu\neq\nu} \mathrm{Tr}\left(1-Z_{\mu\nu}U_\mu U_\nu U_\mu^\dagger U_\nu^\dagger\right) \quad . \tag{2.5}$$

(*) Here and in the following β is the inverse coupling constant squared and *not* the inverse temperature!

The twist $Z_{\mu\nu}$ will be labelled by the twist tensor $n_{\mu\nu} = -n_{\nu\mu}$ through:

$$Z_{\mu\nu} = \exp\left(\frac{2\pi i\ n_{\nu\mu}}{N}\right) . \tag{2.6}$$

To guarantee in weak coupling agreement of the internal energy, we demand the twist to be orthogonal:

$$\kappa(n) = \frac{1}{8} n_{\mu\nu} n_{\alpha\beta} \varepsilon_{\mu\nu\alpha\beta} = \sigma . N \tag{2.7}$$

with σ integer. Furthermore the twist must be nonsimple, which usually means $\sigma \neq o$, and the corresponding minimum action solution must be unique up to a gauge transformation, which is equivalent to the statement $i(n) = 1$, with:

$$i(n) = \text{g.c.d.} \left(n_{\mu\nu}, N, \kappa(n)/N\right) \tag{2.8}$$

(g.c.d. = greatest common divisor, see [6] for details).

In order to establish further correspondence we follow ref. 5 in identifying Wilson loop operators. One makes the change of variables:

$$U_\mu(x) \rightarrow Z(x,\mu) U_\mu(x) \ , \tag{2.9}$$

with $Z(x,\mu) \in Z_N$ such that

$$Z_{\mu\nu} = Z(x,\mu) Z(x+\hat{\mu},\nu) Z(x+\hat{\nu},\mu)^{-1} Z(x,\nu)^{-1} \tag{2.10}$$

is independent of x. After the reduction R one retrieves the action of the TEK model. In order to see that this change of variables is possible we have to show that we can choose the $Z(x,\mu)$ such that we still respect the boundary condition:

$$Z(x,\mu) = Z(x+N_o\hat{o},\mu) \tag{2.11}$$

for all x and μ. We first construct $Z(x,\mu)$ satisfying (2.10) on the same lattice but neglecting the condition (2.11), which is easily done. We shall show that with a suitable Z_N-gauge transformation (2.11) will be satisfied. Each time layer represents a configuration on an infinite 3-dimensional Z_N-lattice. They have the same plaquette values Z_{ij} and so by an appropriate Z_N-gauge transformation we can choose:

$$Z(x,i) = Z(x+\hat{o},i) \tag{2.12}$$

for all x and i. Now consider the ratio of the Z_N-factors picked-up by two neighbouring straight timelike lines: $\prod_{k=o}^{N_0-1} Z(x+k\hat{o},o) \prod_{\ell=o}^{N_0-1} Z(x+\hat{\imath}+\ell\hat{o},o)^{-1}$. Using (2.12) it represents the Z_N-Wilson factor for a closed loop (see fig. 1a) and with the definition (2.10) this ratio becomes $Z_{oi}^{N_o}$. It will be shown later that we must, and can, choose the twist such that N_o is the smallest positive integer with $Z_{oi}^{N_o} = 1$ for all i. Therefore the Z_N-factor:

$$Z_t \equiv \prod_{k=o}^{N_0-1} Z(x+k\hat{o},o) \tag{2.13}$$

is independent of x_i. This is necessary to have a consistent reduction of temporal loops. Unlike for $Z(x,i)$ we have $Z(x,o)$ dependent on x_o, but we can consistently choose $Z(x,o)$ to be periodic:

$$Z(x,o) = Z(x+N_o\hat{o},o) \ , \tag{2.14}$$

because with (2.12) equation (2.10) remains valid. We can even choose a spatially independent Z_N-gauge transformation which makes $Z_t \equiv 1$, which will be assumed below.

As in the TEK-model one has the following correspondence for the Wilson loop C:

$$\frac{1}{N}\,\mathrm{Tr}\Big(\prod_{x\in C} U_\mu(x)\Big) \leftrightarrow \frac{1}{N} \prod_{x\in S} Z_{\mu\nu}\ \mathrm{Tr}\Big(R \prod_{x\in C} U_\mu(x)\Big)\ , \tag{2.15}$$

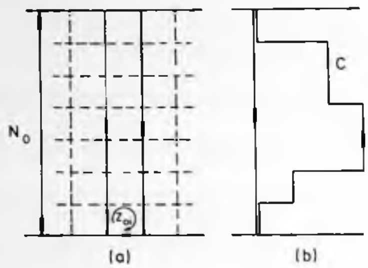

Fig. 1: a) The time-like lines considered in the construction of the change of variables $U_\mu(x) \to Z(x,\mu)U_\mu(x)$.

b) The "closing" of a single temporal loop in order to define the appropriate surface in (2.15).

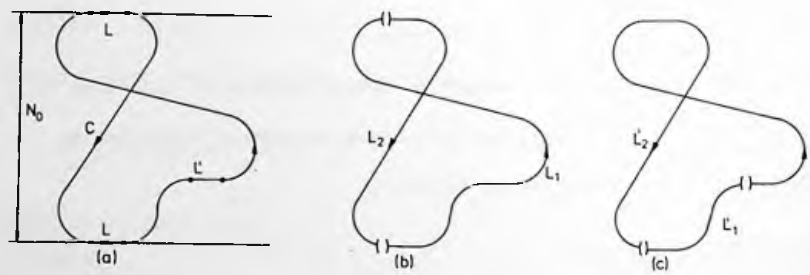

Fig. 2: The loop equation for $< N^{-1} \operatorname{Tr} \prod_{x \in C} U_\mu(x) >_W$ contains a source term $< N^{-1} \operatorname{Tr} \prod_{L_1} U_\mu(x) >_W < N^{-1} \operatorname{Tr} \prod_{L_2} U_\mu(x) >_W$ with L_1 and L_2 as in (b). These terms should survive in the reduced model, but not those (c) coming from the identification of the links L and L' in (a). The twist makes these spurious source terms vanish.

where S is any surface with boundary C ($C = \partial S$). For single temporal "loops" closing by the boundary conditions, S is defined by "closing" C with a straight time-like line (see fig. 1b).

3. Loop equations

We will briefly sketch how the loop equations in the presence of temperature can be the same for the Wilson theory and TEK model. This imposes conditions on the construction of the twists in sect. 4. We have to show that spurious source terms [1] arising from the identification of different link variables ($U_\mu(y)$ and $U_\mu(z) \to U_\mu$) vanish. These terms contain

$$< N^{-1} \, \mathrm{Tr} \, R \prod_{C_{yz}} U_\mu(x) >_{TEK} \tag{3.1}$$

with C_{yz} the part of the loop C running from y to z. For $\beta \to \infty$ the tw eating configurations [5] Ω_μ minimize S_{TEK} and the source term vanishe if $\mathrm{Tr}\left(\prod_{C_{yz}} \Omega_\mu\right) = o$. It is assumed that the twist effects are strong enough to keep (3.1) zero not only at $\beta = \infty$ but down to $\beta_{EK} \sim 0.15N$, here the strong coupling region sets in and the unbroken [12] U(1) symmetry $U_\mu \to e^{i\varphi_\mu} U_\mu$ forces (3.1) zo zero. Monte Carlo simulations appear to confirm this conjecture [5].

In the presence of temperature the loop equations in the Wilson theory have source terms due to the periodicity, which should be maintained after reduction [9]. This implies that N_o must be the smallest positive integer with

$$\mathrm{Tr}\left(\Omega_o^{N_o}\right) \neq o \, . \tag{3.2}$$

The spurious source terms vanish if:

$$\mathrm{Tr}\left(\Omega_o^{k_o}\Omega_1^{k_1}\Omega_2^{k_2}\Omega_3^{k_3}\right) = o \tag{3.3}$$

for k_μ inside a box of dimensions $N_o \times N_1 \times N_2 \times N_3$ with $N_i \to \infty$ for $N \to \infty$, but N_o *fixed*. To illustrate this consider the loop equations for the path C of fig. 2a. On the link L the change of variables [1] $U \to (1+i\epsilon T^j)U$ is performed. In the reduced model there are the genuine (temperature) source terms (3.1) with $C_{yz} = L_{1,2}$ (fig. 2b) and the spurious ones $L'_{1,2}$ from the identification of the links L and L'. The source terms from $L'_{1,2}$ are forced to zero by the twist, but not those from $L_{1,2}$ as follows from (3.2) and the cancelling of the Ω_i's. Note that if the link L' in fig. 1a is shifted upwards to the "boundary" it gives a spurious source term, which also vanishes because $\mathrm{Tr}\left(\Omega_o^{N_o}\Omega_1^{k_1}\Omega_2^{k_2}\Omega_3^{k_3}\right) = o$ for $k_i \neq o \bmod N_i$. This is necessary if we are to mimic finite temperature.

4. Construction of Hot Twists

We will first translate the constraints (3.2) and (3.3) in constraints on $n_{\mu\nu}$, which is related to Ω_μ by:

$$\Omega_\mu\Omega_\nu\Omega_\mu^\dagger\Omega_\nu^\dagger = \exp\left(\frac{2\pi i\, n_{\mu\nu}}{N}\right) . \tag{4.1}$$

As in [5] we find that condition (3.3) is satisfied if and only if k_μ is not an element of the sublattice $A \subset \mathbb{Z}^4$:

$$A = \left\{k \in \mathbb{Z}^4 \middle| \sigma k_\mu = \tilde{n}_{\mu\nu} q_\nu \ , \ q \in \mathbb{Z}^4\right\} , \tag{4.2}$$

where $\tilde{n}_{\mu\nu} = \frac{1}{2}\varepsilon_{\mu\nu\alpha\beta}n_{\alpha\beta}$ and σ is defined by (2.7). Also for all μ we have

$$\Omega_\mu \Omega_o^k \Omega_\mu^\dagger = Z^k_{o\mu}\Omega_o^k \tag{4.3}$$

and as long as $Z^k_{oi} \neq 1$ we have $\mathrm{Tr}\left(\Omega_o^k\right) = o$. So in order to have $\mathrm{Tr}\left(\Omega_o^{N_o}\right) \neq o$ we must have that N is a multiple of N_o, and n_{oi} is a multiple of N/N_o. However for at least one value of i n_{oi} should equal N/N_o. For practical purposes we take A to be a rectangular lattice, so $A = N_o\mathbb{Z} \times N_1\mathbb{Z} \times N_2\mathbb{Z} \times N_3\mathbb{Z}$. We also wish N_1, N_2 and N_3 to be of the same order for $N \to \infty$. Furthermore, anticipating that we want to reproduce the planar expansion we require the algebra generated by

$$A(k) = \Omega_o^{k_o}\Omega_1^{k_1}\Omega_2^{k_2}\Omega_3^{k_3}\ , \ k \in \mathbb{Z}^4/A\ ,\ k \neq o \tag{4.4}$$

to be $SU(N)$. Obviously $A(k)A(k') = Z_{k,k'}A(k')A(k)$ with $Z_{k,k'} \in Z_N$ and $Z_{k,k'} = 1$ iff $k = k'$ mod A. This guarantees that they are linearly independent. Namely suppose that $A(k^{(i)})$ for $i = 1$ to n are independent and $\sum_{i=1}^{n+1} \alpha_i A(k^{(i)}) = o$. Conjugation with $A(k^{(n+1)})$ yields $\sum_{i=1}^{n+1} \alpha_i Z_{k^{(i)},k^{(n+1)}} A(k^{(i)})$ and so $\sum_{i=1}^{n} \alpha_i \left(1 - Z_{k^{(i)},k^{(n+1)}}\right) A(k^{(i)}) = o$. This implies $\alpha_i = o$ for all i up to $n+1$. In conclusion the volume $\prod_\mu N_\mu$ of $\mathbb{Z}^4/A$ has to equal N^2 $(= \dim(SU(N)) + 1)$.

In ref. 6 it was shown that one can easily construct Ω_μ once $n_{\mu\nu}$ is given. So we shall concentrate on finding $n_{\mu\nu}$.

Let us first take $\sigma = 1$ in (2.7). Then $i(n) = 1$ (see (2.8)) gives no extra constraint on $n_{\mu\nu}$. The condition on A is that its 4 generators $k_\mu^{(\nu)} = \tilde{n}_{\mu\nu}$ are related by a $SL(4,\mathbb{Z})$ transformation to the four basis vectors $N_\mu e^{(\mu)}$, therefore $N_\mu^{-1}\tilde{n}_{\mu\nu}$ is an integer (in the last two expressions no summation over μ). The appropriate $SL(4,\mathbb{Z})$ transformation can be written as $Y = -(L^{-1}\tilde{n})^t$ where $L = \mathrm{diag}(N_o, N_1, N_2, N_3)$. One can check that all conditions on $n_{\mu\nu}$ are met by the choice:

$$n_{\mu\nu} = N_0 \begin{pmatrix} 0 & K(4K^2-1) & K(4K^2-1) & K(4K^2-1) \\ & 0 & K(2K-1) & 4K^2-1 \\ & * & 0 & K(2K+1) \\ & & & 0 \end{pmatrix} \tag{4.5}$$

$$N = N_0^2 K(4K^2-1) \ , \ N_1 = N_0 K(2K+1) \ , \ N_2 = N_0(4K^2-1)$$
$$N_3 = N_0 K(2K-1) \ .$$

Using the methods of ref. 6 we find:

$$\begin{aligned} \Omega_0 &= P_1 \otimes 1_2 \\ \Omega_1 &= Q_1 \otimes P_2^{(K+1)(2K-1)} \ Q_2^{1-4K^2} \\ \Omega_2 &= Q_1 \otimes P_2^{K(2K+1)} \ Q_2^{-4K^2} \\ \Omega_3 &= Q_1 \otimes P_2^{K(2K+1)} \ Q_2^{1-4K^2} \ . \end{aligned} \tag{4.6}$$

Here P_i, Q_i are elements of $SU(M_i)$ satisfying the basic commutation relations $P_i Q_i P_i^\dagger Q_i^\dagger = \exp\left(\frac{2\pi i}{M_i}\right)$. For eq. (4.6) $M_1 = N_0$ and $M_2 = N_0 K(4K^2-1)$ and $\otimes$ is the tensor product of $SU(M_1)$ and $SU(M_2)$ which lies in $SU(M_1 M_2 = N)$.

One always has that N_i is proportional to N_0. Here we constructed $n_{\mu\nu}$ such that the N_i are as close together as possible. For $N \to \infty$ we find $N_1 : N_2 : N_3 = 1:2:1$. We would rather have that for $N \to \infty$ all N_i become equal, although it certainly is not necessary. For this we had to compromise a little by allowing for $\sigma \neq 1$, supplied with a constraint on N_0. The twist we found is also more economic in Monte Carlo simulations (in sect. 6 we will elaborate on this). This other Hot Twist is given by

$$n_{\mu\nu} = N_o \begin{pmatrix} o & 2K(4K^2-1) & 4K(4K^2-1) & 2K(4K^2-1) \\ & o & 2K(2K+1) & 4K^2-1 \\ & * & o & 2K(2K-1) \\ & & & o \end{pmatrix} \tag{4.7}$$

$$N = 2N_o^2 K(4K^2-1)\ ,\ N_1 = 2N_o K(2K-1)\ ,\ N_2 = N_o(4K^2-1)\ ,$$
$$N_3 = 2N_o K(2K+1)\ ,\ N_o = odd$$

and the twist eating configuration is found to be:

$$\begin{aligned} \Omega_o &= Q_1^{-2} \otimes P_2^{2K(2K+1)(4K^2-1)}\ Q_2^{4K(1-4K^2)} \\ \Omega_1 &= P_1^{K+1} \otimes P_2^{2K(2K+1)(K+1)}\ Q_2^{-(2K+1)^2} \\ \Omega_2 &= P_1 \otimes P_2^{2K(2K+1)}\ Q_2^{-4K^2} \\ \Omega_3 &= P_1^{1-K} \otimes P_2^{(1-2K^2)(2K-1)}\ Q_2^{(2K-1)^2}. \end{aligned} \tag{4.8}$$

Here $M_1 = N_o$, $M_2 = 2N_o K(4K^2-1)$ define P_i, Q_i as above. This twist has $\sigma = 2$. One must be careful in constructing the sublattice A. Here it will be generated by $k_\mu^{(o)} = n_{\mu o}$, $k_\mu^{(1)} = \frac{1}{2} n_{\mu 1}$, $k_\mu^{(2)} = n_{\mu 2}$ and $k_\mu^{(3)} = \frac{1}{2} n_{\mu 3}$. In (4.2) q is of the form $(2\mathbf{Z},\mathbf{Z},2\mathbf{Z},\mathbf{Z})$ in order to keep k integer.

To guarantee a unique minimum action configuration (up to a gauge) we need $i(n) = 1$, which reduces here to g.c.d. $(\sigma, N_o) = 1$ or N_o = odd.

5. Planar graphs

For $N \to \infty$ we want to reproduce in weak coupling the planar expansion in the presence of finite temperature. The only influence of finite temperature is that the integrals over temporal momenta p_o get replaced by a Fourier sum $\sum_{n=o}^{N_o-1}$ with $p_o(n) = 2\pi nT$. There is no interference with planarity because this is a property of the SU(N) index structure only [13], unlike on the lattice where space-time momenta come from the color degrees of freedom. We will ignore in the continuum the infrared (zero mode) problem coming from the n = o term in the summation$^{(*)}$.

Recall that also for T = o there is only a formal correspondence between the continuum and the TEK-model at $N \to \infty$. Here also one ignores the infrared problems which still plague the theory.

We can closely follow ref. [5] in establishing the correspondence. Conjugation with Ω_μ corresponds to translation, from which one identifies the momenta q:

$$\Omega_\mu A(k)\Omega_\mu^\dagger = \exp\left(\frac{2\pi i}{N} n_{\mu\nu}k_\nu\right)A(k) \ . \tag{5.1}$$

For convenience we suppose $\sigma = 1$ (generalization is straightforward) and we label $k \in Z^4/A$ through

$$k_\nu = \sum_\mu \frac{\tilde{n}_{\mu\nu}q_\mu}{N_\mu} \ , \tag{5.2}$$

then $q_\mu = \frac{N_\mu}{N} n_{\mu\nu}k_\nu$ and $1 \leq q_\mu \leq N_\mu$.

The propagator is that of a lattice with periodic boundary conditions and size $\prod_\mu N_\mu$. For $N \to \infty$ the spatial momenta become continuous but the temporal momenta retain their correct discrete character, so that the temperature propagator in the continuum limit is reproduced (a = lattice

(*) We thank G. 't Hooft for a discussion on this point.

spacing; $aN_o = T^{-1}$):

$$\lim_{a\to o} 2a^{-2}\left(1-\cos\left(\frac{2\pi a q_o}{aN_o}\right)\right) = \left(\frac{2\pi q_o}{aN_o}\right)^2 = p_o(q_o)^2 \tag{5.3}$$

Finally a non-planar graph will acquire a phase factor $\exp\left(-\frac{2\pi i}{N}\sum_{i<j}\sum_{\mu>\nu} k_\mu^{(i)} n_{\mu\nu} k_\nu^{(j)}\right) \neq 1$, where $q_\mu^{(i)}$ are the loop momenta. Let us explicitly evaluate the phase factor for the twist (4.5) and $q^{(i)} = (o, q_1^{(i)}, o, o)$, for which $k_\mu^{(i)} = (-1, o, 2K-1, 1-2K)q_1^{(i)}$. We find: $\exp\left(\frac{2\pi i}{N}\sum_{i<j} k_2^{(i)} n_{23} k_3^{(j)}\right) = \exp\left(\frac{-2\pi i\ (2K-1)}{N_o}\sum_{i<j} q_1^{(i)} q_1^{(j)}\right)$. For $K \to \infty$ $\phi_1 = \frac{2\pi q_1}{N_o K(2K+1)}$ becomes continuous and the phase factor is:

$\exp\left(\frac{-i}{2\pi}\frac{NK(2K+1)}{N_o}\sum_{i<j}\phi_1^{(i)}\phi_1^{(j)}\right)$ which rapidly oscillates. Integration for $N \to \infty$ over $\phi_1^{(i)}$ will yield zero. So all non-planar graphs vanish for $N \to \infty$.

6. Discussion

We constructed the Hot Twists (4.5) and (4.7) for the SU(N) single-point model, which guarantees equivalence for $N \to \infty$ with the Wilson lattice gauge theory at temperature T, both perturbatively and non-perturbatively (loop equations). This is rigorous in the weak coupling limit ($\beta \to \infty$) and in the strong coupling region, but probably holds for all β. This simple model might be useful for analytical calculations. As in [6] one can easily construct surviving fluxons (with an action of $\mathcal{O}(1/N)$ such that $e^{-\beta S\text{ (fluxon)}}$ remains finite for $N \to \infty$). The spectrum is the same ($\Delta S = \frac{8\pi^2}{N}$ is the spacing in the action), but the "occupation numbers" are different. However one can argue that their entropy becomes too small for $N \to \infty$ to give a significant contribution, to f.e. the string tension.

From a practical point of view, e.g. Monte Carlo simulations, one has to keep N finite. Just as for the symmetric twist of [5] we have that for finite N the twist mimics a SU(N) lattice gauge theory with periodic boundary conditions and size $N_o \times N_1 \times N_2 \times N_3$. From [14] we know that finite size effects are small for min $N_i \geq 2N_o$, and we assume this to be true for the one point model also. For the twist (4.5) $N_\mu = N_o(1, K(2K+1), 4K^2-1, K(2K-1))$ and $N = N_o^2 K(4K^2-1)$, so that we must have $K \geq 2$ and therefore $N \geq 30N_o^2$. We now see why the twist (4.7) is more economic: $N = 2N_o^2 K(4K^2-1)$ can be as low as $6N_o^2$, because $N_\mu = N_o(1, 2K(2K-1), 4K^2-1, 2K(2K+1))$ allows for $K = 1$, which gives $N_i \geq 2N_o$.

To be honest we have to compare this number of degrees of freedom $4(N^2-1) = 144N_o^4-4$ of the one point lattice with $24N_o^4$ (or $64N_o^4$) of a $N_o \times 2N_o \times 2N_o \times 2N_o$ lattice with gauge group SU(2) (or SU(3)).
There is only something to be gained for gauge group SU(M) with $M > 4$. Of course from a numerical point of view one has to compare the one point lattice with the lattice theory for $M = N$ and *then* the reduction is enormous. However compared to the torus model of [8], where one reduces only in the spatial directions, our model is much easier to handle. The Monte Carlo calculations are just as easy as for the symmetric twist [5].

One would like to determine $\beta_c(N_o)$, above which the global Z_N symmetry is broken and deconfinement sets in, by calculating the expectation value of a single temperature loop. In order to establish the correct scaling behaviour of $\beta_c(N_o)$ and extract T_c from it, one would like to perform the calculations for quite large N_o values. For the twist (4.7) we have to work with at least SU(6), SU(54) and SU(150) for $N_o = 1$, 3 and 5 respectively. This may seem somewhat large but only in this way are we guaranteed of reasonably small boundary effects.

For given N and N_o our model is better in this respect ($N^2 = N_o \prod_i N_i$) than the single-point (torus) model of ref. 9 (8) with chosen space-time (space) quenched momenta, where $N = N_o \prod_i N_i$ ($N = \prod_i N_i$).

Acknowledgements

We thank Gerard 't Hooft for a critical reading of the manuscript.

References

[1] T. Eguchi and H. Kawai, Phys. Rev. Lett. 48 (1982) 1063.

[2] D.J. Gross and Y. Kitazawa, Nucl. Phys. B206 (1982) 440.

[3] K. Wilson, Phys. Rev. D10 (1975) 2445.

[4] A. Gonzales-Arroyo and M. Okawa, Phys. Lett. 120B (1983) 174.

[5] A. Gonzalez-Arroyo and M. Okawa, Phys. Rev. D27 (1983) 2397.

[6] P. van Baal, Utrecht preprint (February 1983)

[7] F.R. Klinkhamer, Nucl. Phys. B218 (1983) 32.

[8] H. Neuberger, IAS preprint (December 1982)

[9] F.R. Klinkhamer, to be published in Nucl. Phys. B (Leiden preprint, February 1983).

[10] J. Engels, F. Karsch, H. Satz and I. Montvay, Nucl. Phys. B205 (FS5) (1982) 545;
J. Engels, F. Karsch and H. Satz, Phys. Lett. 113B (1982) 398.

[11] J. Greensite and M.B. Halpern, LBL-14912 preprint (1982).

[12] M. Okawa, Phys. Rev. Lett. 49 (1982) 353

[13] G. 't Hooft, Nucl. Phys. B72 (1974) 461.

[14] J. Engels, F. Karsch and H. Satz, Nucl. Phys. B205 (FS5) (1982) 239.

SAMENVATTING

Kosmologie beschrijft de globale struktuur en geschiedenis van het Heelal. Het Heelal expandeert voortdurend en was dus in het begin zeer heet en dicht. Om het vroege Heelal te bestuderen moeten we weten hoe de materie zich gedraagt bij de toen heersende zeer hoge temperaturen ($> 1 GeV \sim 10^{13}$ Kelvin).* Zó belandt men bij de hoge energie fysica, die het gedrag van de elementaire deeltjes en hun wisselwerkingen bij zeer hoge interaktie energieën beschrijft. De kennis hiervan is het laatste decennium sterk toegenomen. Hiermee gewapend is het mogelijk de fysische processen, die werkzaam zijn in de eerste seconde van het Heelal, te bestuderen. Dit dan is het onderwerp van mijn proefschrift.

Waarom maken we ons eigenlijk druk om die ene seconde? Wat mij betreft zijn er drie redenen: (1) puur menselijke nieuwsgierigheid; (2) in deze zeer vroege fase kunnen bepaalde processen werkzaam zijn geweest die de kenmerken van het huidige Heelal bepaalden; (3) sommige theorieën van de elementaire wisselwerkingen opereren bij zulke hoge energieën dat in de komende eeuw(en) deeltjesversnellers op Aarde machteloos zijn, terwijl het zeer vroege Heelal wel over dergelijke energieën beschikte. Een voorbeeld van deze punten zijn de zogenaamde Grand Unification Theories (GUTs). Als we gravitatie even buiten beschouwing laten, voorspellen zij dat bij zeer hoge energieën ($> M_U \sim 10^{15}$ GeV) er slechts één geunificeerde kracht is. De drie waargenomen krachten, namelijk 1) de elektromagnetische, 2) de zwakke, die radioaktiviteit veroorzaakt, en 3) de sterke, die de protonen en neutronen bindt in de atoomkern, zouden dan slechts een "restant" van deze GUT zijn. Punt (3) is evident als we onze sterkste versnellers (~ 1000 GeV) vergelijken met de benodigde energie

*de betekenis van >, < en ~ is groter dan, kleiner dan en ongeveer gelijk aan.

M_U. M.b.t. punt (2) zouden GUTs in het vroege Heelal (leeftijd ~ 10^{-35}s) de nu waargenomen materie-antimaterie asymmetrie kunnen genereren. Ook als er inderdaad een magnetisch monopool deeltje is gevonden moet deze in een zeer vroege fase zijn gemaakt.

De laatste jaren is er een zekere symbiose ontstaan tussen de fysica van elementaire deeltjes en de kosmologie, waarbij de één voorspellingen doet die de ander toetst en vice versa. Voordat ik zeg wat dit proefschrift behandelt bespreek ik eerst de huidige kennis van deze twee vertrekpunten.

Kosmologie. Het standaard model is dat van de zogenaamde Hot Big Bang. Het is gebaseerd op Einstein's theorie van gravitatie. Het model incorporeert een aantal fundamentele waarnemingen en doet een aantal voorspellingen, die bevestigd zijn. Het Big Bang model voldoet voor een tijdsspanne van 1 seconde tot ~ 15 miljard jaar. Niets let ons dan ook om de eerste seconde van de Big Bang te exploreren, maar zoals gezegd is hiervoor kennis van het gedrag van de materie bij zeer hoge temperaturen vereist.

Hoge energie fysica. De elementaire deeltjes zijn in drie categorieën te splitsen: 1) hadronen, 2) leptonen, en 3) "ijkbosonen", die de interakties overbrengen. Hadronen, bijv. het proton, nemen aan de sterke kracht deel, maar leptonen, bijv. het electron, niet. Het is nu zeker dat hadronen opgebouwd zijn uit nog fundamentelere deeltjes: quarks. De kenmerken van een elementair deeltje hangen af van hoe het aan de interakties deelneemt. De kennis van deze krachten is sterk gegroeid: alle interaktes worden beschreven door ijktheorieën. Deze ijktheorieën zijn nogal mathematisch, voor een poging tot uitleg zie bijvoorbeeld mijn artikeltje in de Grote Winkler Prins (8e druk). Het zijn relativistische quantum-velden theorieën met een bijzondere interne symmetrie, de ijkinvariantie. De struktuur van deze ijksymmetrie bepaalt de fysische inhoud van de theorie. Voor interaktie energieën $\lesssim$ 100 GeV is er een standaard model van deze struktuur, die tot nu toe alle toetsen met glans heeft doorstaan. Maar het standaard model laat nog vele vragen open. Er lijken twee alternatieven te zijn. 1) Unificatie (GUTs), waar het standaard model een effectieve theorie bij lage energieën is van één simpelere. Een aantal open vragen van het standaard model zou kunnen worden verklaard. Tevens wordt er een spectaculaire voorspelling gedaan: het proton is instabiel, met gemiddelde leeftijd ong. 10^{31} jaar. Een aantal experimenten zal dit binnenkort toetsen. 2) Compositeness, er is een nog fijnere struktuur, die samenbindt tot het waargenomen standaard model. Helaas is er nog geen goede kandidaat theorie.

Nu zal ik de inhoud van het proefschrift bespreken, zie ook de inleiding 1.1 - 1.3. In deel I gaan we uit van unificatie theorieën en bekijken hun eventuele rol in het vroege Heelal. In deel II beschouwen we het ingewikkelde dynamische probleem van quark "confinement" in hadronen en wat er verandert bij hoge temperaturen ($T \sim 1$ GeV).

Het belangrijkste proces in dit proefschrift is de overgang tussen verschillende fasen van de wisselwerkingen als het Heelal afkoelt. Merk op hoe gecompliceerd de "gewone" fase overgangen van de statistische mechanica al zijn, bijv. het koken van water. Deze twee soorten van "fase overga[n]" hebben veel overeenkomsten.

In deel I beschouw ik de fase overgang van de geunificeer[d] symmetrie naar die van het standaard model. Dit gebeurde toen het [H] een temperatuur had van ongeveer 10^{15} GeV $\sim 10^{28}$ Kelvin en een leeftijd va[n] $\sim 10^{-35}$s. Deze phase overgang kan zeer spectaculair zijn geweest. Hoofdstuk 2 geeft een wat algemenere bespreking, terwijl de hoofdstukken 3 t/m 5 specifieker zijn. Het probleem is dat de overgang "netjes" moet plaatsvinden, met name moet het Heelal na de overgang homogeen blijven en de materie-antimaterie asymmetrie creeëren. Een voorbeeld van hoe de kosmologie een bepaalde deeltjes theorie inacceptabel kan maken levert hoofdstuk 5.

In deel II beschouw ik de overgang van vrije naar absoluut gebonden quarks als de temperatuur van het Heelal daalt onder $T_c \sim 0.2$ GeV = $2\ 10^{12}$ K. Het Heelal was tijdens deze "condensatie" ongeveer 10^{-5}s oud, maar van dit proces zijn er waarschijnlijk geen verdere "tastbare" resultaten meer over. Een andere toepassing is bij toekomstige experimenten met botsingen van versnelde atoomkernen, waar mischien een "fireball" met $T > T_c$ wordt gevormd. In dit deel probeer ik een gevoel te krijgen voor wat er zo drastisch verandert in de vacuum struktuur als de temperatuur in de buurt komt van T_c. Hoofdstuk 6 geeft een overzicht, terwijl ik in hoofdstuk 7 en 8 twee semiquantitatieve modellen van het mechanisme van de overgang voorstel. Deze twee modellen sluiten elkaar niet uit, maar benaderen het probleem eerder van twee verschillende kanten. De laatste hoofdstukken 9-11 zijn iets technischer. In de ijktheorie voor de interaktie tussen de quarks zijn er N=3 soorten lading, kleurrijk "kleur" genoemd. Eigenlijk is N de enige vrije parameter van de theorie en er treedt een aanmerkelijke vereenvoudiging op in de limiet $N \to \infty$. Als we de $N = \infty$ theorie zouden kunnen oplossen zitten we misschien al dicht ($\sim 1/N^2 \sim 10\%$) bij de oplossing van de relevante N=3 theorie. Recentelijk vonden Eguchi en Kawai dat de $SU(N=\infty)$ ijktheorie op een ruimte-tijd rooster enorm vereenvoudigd kan worden. In hoofdstuk 9 en 10 bestudeer ik deze gereduceerde modellen en kijk hoe een eindige temperatuur kan worden ingebouwd. Hoofdstuk 11 geeft een eenvoudig model waarmee voor $N \to \infty$ (de)confinement van quarks bij eindige temperatuur kan worden onderzocht.

Tenslotte verwijs ik de lezer in paragraaf 1.4 van de introductie naar recente resultaten die dit proefschrift aanvullen, en geef in 1.5 enige globale conclusies.

DANKBETUIGING

Allereerst wil ik mijn erkentelijkheid betuigen aan de Rijksuniversiteit van Leiden en haar Sterrewacht voor de volstrekte vrijheid waarmee mijn onderzoek zich heeft kunnen ontplooien. Ik dank Colin Norman voor het geschonken vertrouwen toen mijn belangstelling zich meer en meer op de fysica ging richten.

Om au courant te raken en te blijven was het belangrijk enkele congressen bij te wonen en ik dank het Kerkhoven-Bosschafonds, het Lorentzfonds en Z.W.O. voor de verleende subsidies.

Op de Sterrewacht hielpen Lena, Lenore, Marja, Petra en Wanda mij op efficiente wijze met talloze practische zaken. Tenslotte dank ik de Directeur van het Fitzwilliam Museum voor zijn toestemming om de tekening van Goya als frontispice te mogen gebruiken.

STUDIEOVERZICHT

Frans Richard Klinkhamer, geboren 9 september 1956 in Utrecht, behaalde in 1974 het einddiploma Gymnasium-β aan het Stedelijk Gymnasium te Leiden. Daarna begon hij zijn studie aan de Rijksuniversiteit van Utrecht en deed begin 1980 cum laude het doctoraal examen algemene sterrenkunde met de bijvakken wiskunde en theoretische plasmafysica. Vanaf 1 april 1980 was hij werkzaam als wetenschappelijk assistent bij de Sterrewacht van de Rijksuniversiteit Leiden. Zijn onderzoek bewoog zich op het terrein van de kosmologie en de fysica van elementaire deeltjes.

STELLINGEN

1. Proust[1] heeft opgemerkt dat Flauberts titel "L'éducation sentimentale" wel mooi compact is, maar eigenlijk grammaticaal onjuist. Een voorbeeld van wat hij bedoeld kan hebben is te vinden in ditzelfde boek (hoofdstuk 2 van deel 2): "Elle se plaignit de ses rares visites, (..)" i.p.v. "de la rareté de ses visites".

[1]*M. Proust, A propos du 'style' de Flaubert, N.R.F. janvier 1920.*

2. Een manco in de discussie van Hildesheimer[1] over de literaire antecedenten en de door Mozart gecreëerde diepgang van het type Don Giovanni is het ontbreken van enige verwijzing naar Moliere's toneelstuk "Don Juan ou le festin de pierre", terwijl Mozart daar al vroeg[2] mee in aanraking was gekomen.

[1]*W. Hildesheimer, Mozart, 1977.*

[2]*Mozart Briefe und Aufzeichnungen, II, ed. W.A. Bauer, O.E. Deutch, 1962, brief van 24 maart 1778.*

3. De reductie in de topologische limiet van QCD op een rooster naar een éénpunts model[1] is onjuist[2].

[1]*H. Levine, H. Neuberger, Phys. Lett. 119 B (1982), 183.*

[2]*F.R. Klinkhamer, preprint (April 1983).*

4. Met de twist van hoofdstuk II is het mogelijk een gereduceerde Hamiltoniaan te vinden[1].

[1]*F.R. Klinkhamer, preprint (June 1983).*

5. De radiostraling bij Antares B geeft een indirecte bepaling van het massaverlies van een B 2.5 V ster, ongeveer 10^{-10} tot 10^{-9} zonsmassa per jaar.[1]

[1]*F.R. Klinkhamer, J. Kuijpers, Astron. Astrophys. 100 (1981), 291.*

6. Het zijn onjuiste simplificaties te stellen dat het expanderend Heelal begon met een *explosie* en dat de roodverschuiving van verre sterrenstelsels het *gevolg* is van het Doppler effect.