

# Phase transitions in quantum chromodynamics

Hildegard Meyer-Ortmanns\*

*Institut für Theoretische Physik, Universität Heidelberg, D-69120 Heidelberg, Germany*

The current understanding of finite-temperature phase transitions in QCD is reviewed. A critical discussion of refined phase-transition criteria in numerical lattice simulations and of analytical tools going beyond the mean-field level in effective continuum models for QCD is presented. Theoretical predictions about the order of the transitions are compared with possible experimental manifestations in heavy-ion collisions. Various places in phenomenological descriptions are pointed out where more reliable data for QCD's equation of state would help in selecting the most realistic scenario among those proposed. Unanswered questions are raised about the relevance of calculations that assume thermodynamic equilibrium. Promising new approaches for implementing nonequilibrium aspects of the thermodynamics of heavy-ion collisions are described. [S0034-6861(96)00302-9]

## CONTENTS

I. Introduction	474	a. How to measure an interface tension	509
II. General Background	477	b. Multicanonical algorithms	510
A. Phase transitions in statistical systems	477	4. Phenomenological implications for the early universe	511
1. First- and second-order transitions in the infinite-volume limit	477	C. Including dynamical fermions	513
a. The Landau free energy	478	1. Pitfalls on the lattice	513
b. The renormalization-group approach in the infinite-volume limit	480	2. Finite-size scaling analysis	514
2. Finite-size scaling analysis	481	3. Finite-mass scaling analysis	515
a. Second-order transitions	482	4. Bulk transitions	518
b. First-order transitions	483	a. Identification of the lattice artifact	519
c. The two-peak Gaussian model	484	b. Basic observables	520
d. Binder's cumulant	485	c. Determination of the critical coupling	520
e. $q$ -state Potts models	485	d. Indications of chiral symmetry restoration	521
B. Phase transitions in QCD	488	5. Results for two and three flavors	522
1. Renormalization-group analysis in the chiral limit	488	6. The equation of state for two-flavor QCD	524
2. The limit of a pure $SU(N_c)$ gauge theory	490	7. Simulations with Wilson fermions	525
a. Effective vs physical temperatures	491	a. Difficulties with Wilson fermions	526
b. The phase structure of $Z(N)$ spin models	491	b. 2+1 flavors and critical quark masses	527
c. Inclusion of dynamical quarks	492	c. Improved actions	527
d. Finite quark masses and external fields	492	D. QCD at finite baryon density	528
III. The Lattice Approach to the QCD Transition	493	IV. Effective Models in the Continuum	529
A. A primer for lattice gauge theory	493	A. Models for quark degrees of freedom	529
1. The lattice action	494	1. QCD and dysprosium	530
2. The Wilson action and hopping parameters	495	a. Response to a stimulus in $T$	532
3. Staggered fermions and flavor symmetries	495	b. Response to a stimulus in $m$	532
4. Sources of error	496	c. The critical equation of state	532
5. Translation from lattice units to physical units	496	d. The role of the strange quark mass	533
6. The critical temperature $T_c$	497	2. The chiral transition in chiral perturbation theory	534
7. A test of asymptotic scaling	497	a. The free-field case	536
8. Translation to physical units	498	b. Including interactions up to three loops	536
9. Numerical simulations	498	c. Results in the chiral limit without heavier mesons	536
a. Monte Carlo on the lattice	498	d. Nonvanishing quark-masses	537
b. The hybrid Monte Carlo algorithm	499	e. Influence of heavier mesons	537
B. Pure gauge theory	500	3. Clausius-Clapeyron equation for QCD	538
1. The order of the $SU(3)$ deconfinement transition	500	4. Mass sensitivity of the chiral transition	539
a. Correlation lengths, mass gaps, and tunneling events	501	a. Mass sensitivity	539
b. Adaptation to the finite volume	503	b. Choice of the model	540
2. Thermodynamics on the lattice	505	c. Tree-level parametrization at zero temperature	540
3. Interface tensions in QCD	508	d. Critical meson masses in a mean-field calculation	541
		e. The large- $N_f$ approach	542
		f. Results for critical quark/meson masses	542
		g. An upper bound on the latent heat	543
		B. Models for gluonic degrees of freedom	544
		1. A network of strings	544
		a. The $SU(2)$ case	545
		b. The $SU(3)$ case	546
		c. Inclusion of matter fields	547
		2. Dual Ginzburg-Landau models	548
		3. Some further approaches	549
		V. Relativistic Heavy Ion Collisions	551

\*Electronic address:

[h.meyer-ortmanns@thphys.uni-heidelberg.de](mailto:h.meyer-ortmanns@thphys.uni-heidelberg.de)

A. Scales and observables	551
B. The hydrodynamic framework	554
1. Bjorken-Shuryak expansion scenario	555
2. The bag model equation of state	556
C. Signatures sensitive to the nature of the phase transition	558
1. Thermodynamic observables	558
a. Entropy measurements	558
b. Energy measurements	559
c. Transverse momentum distributions	559
d. Enthalpy and pressure	559
2. Dileptons and real photons	561
a. Dileptons	561
b. Integration over the spacetime history	562
c. Real photons	565
3. Strangeness production	566
4. Pion interferometry	571
5. Multiplicity fluctuations	576
6. Intermittency analysis	578
D. Theoretical concepts for off-equilibrium situations	581
1. Nucleation rate of hadronic bubbles	581
2. Large domain coarsening	584
3. Transport coefficients	586
4. Dynamical universality and disoriented chiral condensates	588
VI. Summary and Conclusions	590
Acknowledgments	592
References	592

## I. INTRODUCTION

Phase transitions are found over a wide range of temperatures. They start near absolute zero with the Mott transitions (Mott, 1968) from superconductor to insulator and are common between  $10^2$  K and  $10^3$  K. Nine orders of magnitude higher one finds the phase transitions discussed here. At a temperature typical of the QCD scale ( $\Lambda_{\overline{\text{MS}}}\sim 200$  MeV) nuclear matter melts. At this temperature quarks and gluons cease to be confined inside hadrons and begin moving freely. Three orders of magnitude higher is the next phase transition at which the electromagnetic and weak interactions are unified.

These two transitions may seem exotic, but according to the big bang theory they have occurred at least once in the universe. The electroweak transition is predicted to have taken place at  $10^{-10}$  s, and the deconfinement transition at  $10^{-6}$  s after the big bang. The deconfinement transition occurred when the temperature of the universe dropped to the order of tera degrees, i.e., to a scale  $T_c\sim(2.32\pm 0.6)\times 10^{12}$  K or  $200\pm 50$  MeV. [The conversion from natural energy units (MeV) to degrees Kelvin (K) is determined by the Boltzmann constant  $k_B\sim 8.6\times 10^{-5}$  (eV/K) or  $(100\text{ MeV}\sim 1.16\times 10^{12}\text{ K})$ .] These temperatures with time scales of  $10^{-23}$  s and typical distances of 1 fm may now be achievable in the laboratory. At CERN one reaches energy densities around  $1.5\text{--}2.5\text{ GeV/fm}^3$ , and higher densities are expected when Relativistic Heavy-Ion Collider (RHIC) experiments come on-line at Brookhaven (Stachel, 1994).

We have glibly mentioned “the QCD transition,” as if it were unique. However, there are really two, and they

take place if either the temperature or the density crosses some critical point.

Each transition is well defined in some limit of QCD. If the quark masses are infinite, one has the *deconfinement* transition. If the quark masses are zero, one has the *chiral* (symmetry-restoring) transition. It is not yet clear whether these transitions persist for physical quark masses or whether they will occur together. If they take place at different temperatures, one expects the chiral transition to occur after the deconfinement transition (Banks and Casher, 1980; Campbell *et al.*, 1990). Lattice results indicate, however, that both transitions coincide at the same critical point. One of them can then be regarded as driving the other, and the nature of the transition will be determined by the driving symmetry. Here it is assumed that both happen simultaneously, unless stated otherwise.

As mentioned above, a transition is expected for both increasing density and increasing temperature. This review concentrates on the effects of temperature, as this is the area in which the lattice is strong and many partial or preliminary results are available.

One usually has many length, or equivalently energy, scales playing a role at a phase transition. This makes it impossible, in most cases, to find a suitable expansion parameter for a perturbative treatment. In addition, QCD is asymptotically free in the UV limit, but has a coupling  $g(T)\approx 1$  near the phase transition. Thus a non-perturbative approach must be used; lattice regularization may be the only means of calculating the properties of the phase transition from first principles.

One expects, naively, a phase transition in QCD if the density rises to the point at which the hadrons begin to overlap. This will be around  $0.5\text{ GeV/fm}^3$ , the energy density in a proton.

Similarly, for temperatures above the QCD scale parameter,  $T>\Lambda_{\overline{\text{MS}}}=200$  MeV, one expects a phase transition. At extremely high temperatures,  $T\gg\Lambda_{\overline{\text{MS}}}$ , with typical momenta  $Q\approx T$ , the coupling  $Q^2/\Lambda_{\overline{\text{MS}}}^2$  becomes weak, and resummed finite-temperature perturbation theory (Braaten and Pisarski, 1990a, 1990b, 1990c, 1990d, 1992a, 1992b) may be applied.

There are three natural length scales in the QCD plasma,  $l_p=1/T<l_e=1/gT<l_m=1/g^2T$  (DeGrand and DeTar, 1986). The first gives the range over which perturbation theory may be applied, the second corresponds to the scale set by the chromoelectric mass, and the third to the scale set by the chromomagnetic mass. The latter two scales reflect the increasingly nonperturbative nature of long-range interactions in the plasma and the lack of screening in the magnetic sector.

Although these long-range interactions are strong and may lead to a type of dynamic confinement in the plasma, the bulk properties, the equation of state, etc. are still dominated by short-range interactions, so perturbation theory may well be valid at high temperatures for these quantities.

More evidence that the chromomagnetic, long-range sector of QCD may have a complicated, nonperturbative structure comes from the phenomenon of dimen-

sional reduction. In  $(3+1)$ -dimensional Euclidean field theory, the imaginary time dimension is proportional to the inverse temperature,  $l_\tau \propto 1/T$ . Thus the theory reduces, in the high-temperature limit, to three-dimensional QCD (Reisz and Petersson, 1991; Kärkäinen *et al.*, 1992, 1993, 1994; Reisz, 1992).

Deriving the properties of the QCD phase transition is clearly difficult. After all, it is analogous to deriving the properties of boiling water directly from the quantum Hamiltonian of water molecules. So the thermodynamics of the phases and the order of the transitions between them have only recently been settled for pure gauge theory, despite nearly two decades of computer simulations.

The *order of a phase transition* is one of the basic thermodynamic classifications. A phase transition is said to be of *first order* if there is at least one finite gap in the first derivatives of a suitable thermodynamic potential in the thermodynamic limit. A finite latent heat often goes along with a gap in the order parameter. A transition is said to be of *second order* if there is a powerlike singularity in at least one of the second derivatives of the potential. If the thermodynamic potential is analytic over the whole temperature range (for temperature-driven transitions), the phase conversion is called a *crossover phenomenon*.

In general the order of the QCD transitions depends on the number of colors  $N_c$ , the number of flavors  $N_f$ , the current quark masses, and more subtle effects related to the IR and UV cutoffs. The deconfinement transition is now believed to be of first order for three colors (Fukugita *et al.*, 1989; Gavai *et al.*, 1989), and the equation of state is known in the limit of infinite quark masses (Engels, Fingberg, *et al.*, 1990; Laermann *et al.*, 1995). Similarly the chiral transition is believed to be of second order for two massless flavors (Pisarski and Wilczek, 1984; Karsch, 1994) and of first order for three or more massless flavors (Pisarski and Wilczek, 1984; Gavai and Karsch, 1985; Gavai *et al.*, 1987; Brown *et al.*, 1990a; 1990b; Iwasaki, 1995; Iwasaki, Kanaya, Kaya, *et al.*, 1995). These results do not contradict each other. They just indicate the sensitivity of the order to the involved approximations and should be taken as a warning not to jump to conclusions.

Ultimately one is interested in the physically relevant case of two light quarks (up and down), and one heavier quark (the strange quark). This case is still under discussion. It is an open question whether the chiral transition and the deconfinement transition persist for realistic quark masses.

The order of a phase transition has far-reaching phenomenological consequences. Phenomenological implications of QCD transitions are supposed to be visible in relativistic heavy-ion collisions and astrophysics. For a first-order transition one expects metastabilities with latent heat, interfaces, supercooling and overheating, etc. The experimental consequences of a first-order transition make it relatively easy to see, especially if the plasma “explodes” into the hadronic phase (Cleymans *et al.*, 1986). A second-order transition, lacking a jump in

the energy density, may be less easy to see experimentally. However, it has divergent correlation lengths in the thermodynamic limit. These may also lead to observable consequences (Bialas and Peschanski, 1988; Wilczek, 1992; Bjorken *et al.*, 1993; Rajagopal and Wilczek, 1993a, 1993b; Gavin *et al.*, 1994a, 1994b), in analogy to the well-known phenomenon of critical opalescence.

Perhaps most likely is a smooth crossover, as in the transition to an electric plasma from a molecular or atomic gas. One will have a smooth transition from a pion gas at low temperatures to a quark gluon plasma at high temperatures, with a highly nontrivial mixture of excitations in the neighborhood of the crossover. If one has a sharp crossover phenomenon with a rapid change in thermodynamic quantities over a small temperature interval (say of the order of 10 MeV), there is still some chance for measurable effects in experiments. For example, double  $\Phi$  peaks in the dilepton invariant mass spectrum are still predicted as a signature for the phase conversion, as long as the crossover phenomenon is rapid enough (Ko and Asakawa, 1994; Ko, 1995).

The situation is further complicated by the nature of the experiments. Ultrarelativistic heavy-ion collisions have a finite volume expanding in time. The finite volume smooths out the nonanalyticities in the free energy. The high-temperature matter created in the collisions may not reach equilibrium either, and does have different processes dominating in different regions of phase space. It becomes a nontrivial task to find “thermometers” to tune and measure the temperature of the transient hot plasma. Thus no clear signal of the phase transition is known, although a number of signals has been proposed, and the quantities measured so far can usually be modeled by both a hot hadron gas and a quark-gluon plasma.

In the early universe we have different competing scales. The expansion of the universe is rather slow in units of QCD. If a typical time scale of QCD is taken as  $1/T_c$  ( $\sim 1 \text{ fm}/c \sim 10^{-23} \text{ s}$ ), the Hubble time is of the order of  $10^{19}/T_c$ . Thus there is enough time for equilibration before and after the transition. The spatial volume  $V$  of the universe,  $10^{-6} \text{ s}$  after the big bang, appears as almost infinite in units of QCD. In units of  $T_c$  it is given as  $VT_c^3 = 7.1 \times 10^{55} (200 \text{ MeV}/T_c)^3$ , if the physical correlation length  $\xi$  is taken as  $\xi(T_c)T_c = 1.38 \pm 0.24$  in the deconfinement phase. This leads to a rather small upper bound on the amount of supercooling during a (hypothetical) first-order deconfinement transition (Banerjee and Gavai, 1992). The effect of supercooling is an important ingredient in the argument for why one should see remnants of the early QCD transition even today. A first-order deconfinement transition could have led to inhomogeneities in the baryon number density in the early universe. If these inhomogeneities survived until the epoch of primordial nucleosynthesis, they could have influenced the light-element abundances and led to deviations from values obtained in the standard scenario (without a phase transition) (Applegate *et al.*, 1987; Fuller *et al.*, 1988; Schramm *et al.*, 1992). It is these deviations in the light-element abundances which have

been predicted as a visible remnant today. Meanwhile this prediction has become rather questionable. Initial inhomogeneities that are compatible with the rather small supercooling are most likely insufficient to induce well separated proton-rich and neutron-rich regions. Pronounced inhomogeneities in the proton and neutron distributions are a precondition for a considerable change in the initial conditions of primordial nucleosynthesis.

Lattice Monte Carlo simulations become rather time consuming when one tries to simulate a realistic situation with small but nonzero quark masses. The CPU time of a typical simulation with dynamical quarks is of the order of months or even years. From the viewpoint of one less familiar with the lattice approach, it is difficult to understand why, for example, extensive Monte Carlo calculations have been performed to simulate a strongly interacting system with eight nearly massless flavors, when only two flavors are light in nature. Without a deeper understanding of the lattice methods one is left with the impression that the field of QCD computer simulations has decoupled from the field of heavy-ion collisions. The connection between both areas of research seems sometimes to be lost.

It would be overly ambitious to attempt, in this article, to close the gaps between the various approaches to understanding the finite-temperature transitions of QCD. This review will focus on the progress made in determining the nature of the phase transitions, and the development of sophisticated tools for this purpose. The complexity of the task and the time needed to turn the hints of today into firm answers will be explained.

The organization of the paper is as follows. Section II.A gives some background from statistical physics. We summarize the main steps in a renormalization-group approach, for both infinite and finite volume, and recall some concepts for distinguishing first- and second-order transitions, independently of the specific QCD dynamics. In Sec. II.B we deal with a renormalization-group analysis applied to QCD. Such an analysis provides a guide for a more detailed investigation of the phase diagram.

In view of the anticipated nonperturbative nature of the transition and the need for simulating QCD rather than a simplified model, it is natural to start with the lattice approach (Sec. III). From numerous papers we select a few to explain the kind of progress that has been made. Section III.A serves mainly to give the basic ideas, to fix the notations, and to introduce some calculational tools of lattice gauge theory.

Section III.B deals with the pure gauge theory without matter fields. We describe the controversy over the order of the deconfinement transition in SU(3) gauge theory (Sec. II.B.1). Thermodynamics on the lattice is the topic of Sec. III.B.2. For a long time it has been taken for granted that the effect of matter on the transition is only to weaken the strength of its order. Hence various interface or surface tensions have been calculated as further characteristics of the transition dynamics. Interface tensions refer to free energies associated

with interfaces between different realizations of the plasma phase, between the phases of broken and restored chiral symmetry, or between the phases of deconfinement (plasma phase) and confinement (hadronic phase). A reliable estimate of their values at the transition point could play a selective role in the phase-transition scenarios. A measurement of the interface tension in SU(3) gauge theory is presented in Sec. III.B.3. Intimately related to the topic of interface tensions are the phenomenological implications of the QCD transition in the early universe. In Sec. III.B.4, we summarize the argument for finding relics of the early QCD transition.

Section III.C is devoted to the inclusion of dynamical fermions. Here we restrict the discussion mostly to the staggered fermion formulation. In Sec. III.C.1 we list the pitfalls of the lattice, including infrared, ultraviolet, and finite-mass artifacts. We describe a finite-size scaling analysis in the presence of fermions in Sec. III.C.2. Finite-mass artifacts (analyzed in a finite-mass scaling analysis) refer to the necessity of extrapolating results to the chiral limit from finite-mass calculations (Sec. III.C.3). Examples of the subtleties of UV artifacts are given in Secs. III.C.3 and III.C.4. In Sec. III.C.5 we present results for the cases of two and three light-fermion flavors, which come closest to the physical current quark masses. Recent results on the equation of state for two-flavor QCD are explained in Sec. III.C.6. Wilson fermions are the topic of Sec. III.C.7.

A brief review of lattice simulations at finite baryon density is presented in Sec. III.D.

The lattice approach plays an important role for QCD. It has the advantage of starting from first principles. But even if lattice calculations finally succeed in providing reliable predictions, there is some need for further alternatives. In a typical Monte Carlo calculation, QCD's full partition function is simulated at once (if we disregard the involved approximations for a moment). From such a simulation one does not get an intuitive insight into why a result comes out in a particular way. Effective models try to fill this lack. They are usually not derived from QCD in a strict sense, but share some important symmetries with QCD. Accordingly we consider models for quark and gluonic degrees of freedom separately (Secs. IV.A and IV.B).

In Sec. IV.A we concentrate on theoretical tools that go beyond the mean-field level: the renormalization-group approach (Sec. IV.A.1), chiral perturbation theory (Secs. IV.A.2 and IV.A.3), and a  $1/N$  expansion (Sec. IV.A.4). The renormalization group leads to specific predictions for the critical indices and the temperature dependence of meson masses close to  $T_c$ . These predictions may serve as working hypotheses for lattice simulations (Wilczek, 1992; Rajagopal and Wilczek, 1993a, 1993b).

Chiral perturbation theory is well established for describing QCD at low energies and small temperatures. We outline the derivation of chiral condensates as a function of temperature in Sec. IV.A.2. Of particular interest are the influence of finite current quark masses

and an estimate of the contribution of heavier mesons to thermodynamic quantities. Heavier mesons turn out to be non-negligible in the phase-transition region, contrary to one's naive expectation (Gerber and Leutwyler, 1989). Although chiral perturbation theory fails in the vicinity of the transition region, it leads to an upper bound for the latent heat, when Clausius-Clapeyron relations are applied to QCD in a hypothetical first-order transition (Leutwyler, 1992; Sec. IV.A.3). The bound excludes strong supercooling scenarios and thus may be of particular interest to phenomenologists.

In Sec. IV.A.4 we apply a  $1/N$  expansion to the linear  $SU(3) \times SU(3)$  sigma model. Here large  $N$  refers to the number of flavors. The topics are the mass sensitivity of the chiral transition and thermodynamics for physical values of (pseudo)scalar meson masses. An attempt is made to locate the phase boundary between first-order chiral transitions and crossover phenomena in meson mass space (Meyer-Ortmanns and Schaefer, 1996).

As an example of a simplified physical picture of the interaction dynamics in a pure gauge theory, we consider Patel's color flux-tube models (Patel 1984a; 1984b) in Sec. IV.B.1. In the flux-tube models the  $SU(3)$  deconfinement transition occurs when the network of color flux tubes becomes infinitely connected. The models abandon symmetry breaking as a driving mechanism for the deconfinement transition. Section IV.B.2 deals with dual Ginzburg-Landau models, explaining the confinement property of QCD in terms of a dual Meissner effect.

If quark and gluonic degrees of freedom are included in an effective Lagrangian for QCD, their interplay in the transition dynamics may be investigated. We consider the outlook for such attempts in Sec. IV.B.3.

Section V is devoted to heavy-ion collisions. Although the application of a thermodynamic concept is not fully established, throughout the review we adopt the viewpoint that a thermodynamic description of nucleus-nucleus collisions is meaningful. After a recapitulation of basic concepts in relativistic hydrodynamics (Sec. V.B), we point out possibilities for measuring thermodynamic observables (Sec. V.C.1). Substructures in dilepton spectra and enhanced strangeness production, which could be sensitive to the underlying transition dynamics, are the topic of Secs. V.C.2 and V.C.3. Pion interferometry, measurements of multiplicity fluctuations, and intermittency are powerful experimental tools in heavy-ion collisions. They may also be utilized for inferring the order of the QCD transition (Secs. V.C.4, V.C.5, and V.C.6). The common difficulty for all of these experimental devices is to find a *unique* signature for a certain type of transition. Thus a flattening in the  $\langle p_T \rangle$  distribution of charged pions as a function of multiplicity distributions in rapidity space is compatible with a first-order transition, but lack of available phase space serves as an alternative and much simpler explanation (van Hove, 1985).

We shall not cover the issue of preequilibrium production, although it is likely to be relevant in some ways. Section V.D is concerned with off-equilibrium theory.

Estimates of the nucleation rate and transport coefficients are two examples of attempts to incorporate the knowledge of equilibrium QCD in situations slightly off equilibrium. An alternative off-equilibrium process is large domain coarsening. Under certain conditions it is more likely to be the relevant mechanism for phase separation at the transition point (Sec. V.D.3). Plasma evolution far out of equilibrium is the topic of the final section, V.D.4. Methods of nonequilibrium physics and the powerful concept of dynamic universality classes become applicable if the hot plasma is quenched via almost instantaneous cooling (Rajagopal and Wilczek, 1993a, 1993b). Large correlation volumes may be created in spite of a small equilibrium correlation length, which is hampered from growing due to the finite pion mass. In realistic heavy-ion experiments the hypothetical plasma presumably cools neither adiabatically nor instantaneously, and the physical masses are likely not to be light enough for inducing large correlation volumes with the aid of nonequilibrium amplification. Nevertheless, we conclude Sec. V with some ideas about dynamic universality, as it may be challenging and stimulating to consider a quench of a hot quark-gluon plasma as some kind of gedanken-experiment.

In Sec. VI we summarize the most important points of the preceding sections. We list the main results, the open questions, and interesting perspectives for further studies.

Unavoidably, a selection of papers has been cited from among numerous others related to the same topic, in order to keep the article readable. If I do not refer to a paper in spite of its relevance, this may be a result of ignorance.

## II. GENERAL BACKGROUND

### A. Phase transitions in statistical systems

#### 1. First- and second-order transitions in the infinite-volume limit

The order of a phase transition is one of the basic thermodynamic classifications. It concerns the thermodynamic potential and its derivatives at the transition. The thermodynamic potential  $\Omega$  is the free energy  $F$  for a ferromagnet or the Gibbs free energy  $G = F + pV$  for a fluid. In Eqs. (2.1) we recall the basic thermodynamic formulas for a magnet:

$$Z(T, H, V) = \int \mathcal{D}U e^{-\beta \mathcal{H}(U)}, \quad (2.1a)$$

$$F = -T \ln Z, \quad (2.1b)$$

$$E = -\frac{\partial \ln Z}{\partial \beta}, \quad S = -\left(\frac{\partial F}{\partial T}\right)_H, \quad M = -\left(\frac{\partial F}{\partial H}\right)_T, \quad (2.1c)$$

$$c_H = \left(\frac{\partial E}{\partial T}\right)_H, \quad c_{H,M} = T \left(\frac{\partial S}{\partial T}\right)_{H,M}, \quad (2.1d)$$

$$\chi_T = \frac{1}{V} \left( \frac{\partial M}{\partial H} \right)_T.$$

Here  $\mathcal{H}$  denotes the spin Hamiltonian,  $\beta$  is the inverse temperature, the Boltzmann constant  $k_B$  has been set equal to 1 everywhere,  $H$  is an external magnetic field,  $V$  is the volume, and  $\int \mathcal{D}U$  stands for the sum over all spin configurations  $\{U\}$ , weighted with the Boltzmann factor  $\exp\{-\beta\mathcal{H}\}$ . Equation (2.1b) defines the free energy  $F$ , depending on  $T$ ,  $H$ , and  $V$ , of a system in a finite volume  $V$ . In the *large-volume* limit the free-energy density  $f$  is given as  $f = \lim_{V \rightarrow \infty} F(V)/V$ , with  $F(V) \equiv \int_V d^d x f(x)$  in  $d$  space dimensions. The free energy is assumed to depend on the scaling fields  $T$  and  $H$ . In view of a finite-size scaling analysis, we retain the dependence on  $V$ . The parameters  $T$  and  $H$  are on an equal footing—both may drive a phase transition. One speaks of a *temperature-driven* or a *field-driven transition*. In QCD we focus our discussion on temperature-driven transitions (although density- or “mass-driven” transitions may be considered as well). First derivatives of the free energy with respect to  $T$  or  $H$  lead to the internal energy  $E$ , the entropy  $S$ , or the magnetization  $M$  according to Eq. (2.1c). The magnetization is the conjugate variable to the external field and plays the role of an order parameter in the case of a magnet. More generally, it plays the role of an order parameter if the external field *explicitly* breaks the symmetry, which may be spontaneously broken at a phase transition.

On the second level of derivatives [Eqs. (2.1d)] we have the specific heat  $c$  (at constant  $H$  and/or  $M$ , respectively) and the isothermal susceptibility  $\chi_T$ . For a fluid, the isothermal susceptibility would be replaced by the compressibility  $\kappa$ .

In the infinite-volume limit, a phase transition is signalled by a singularity (in the sense of nonanalyticity) in the thermodynamical potential  $\Omega$ . If there is a finite discontinuity in at least one of the first derivatives of  $\Omega$ , the transition is called *first order*. In the case of a ferromagnet, there is a jump in the magnetization if one passes through the transition temperature from the phase of broken symmetry to the symmetric phase. This gives  $M$  the name of an *order parameter*, as  $M$  indicates the order of spins. In this way it tells us the phase in which the system is encountered at a given temperature.

The remaining first derivatives of  $F$  with respect to  $T$ , the internal energy and entropy, usually also show a discontinuity at the transition point. A gap in the entropy is associated with a finite latent heat  $\Delta Q = T_c \cdot \Delta S$ , but there need not be such a gap. (Consider a transition in a ferromagnet between states of magnetization opposite in sign but equal in magnitude. The latent heat would vanish in this case, while the magnetization would jump between values of opposite sign.) Conversely, there may be a finite latent heat without a gap in the order parameter at the transition point.

The second derivatives of the thermodynamic potential at a first-order transition are typically  $\delta$ -function singularities (corresponding to the discontinuities in the first derivatives) or finite.

According to the original Ehrenfest classification of phase transitions, *nth-order transitions* are defined by the occurrence of discontinuities (rather than divergences) in the *nth-order* derivative of the appropriate thermodynamical potential. In M. E. Fisher’s terminology, one distinguishes between first-order and *continuous* (or higher-order) transitions. In continuous transitions the first derivatives of  $\Omega$  are continuous, whereas second derivatives are either discontinuous or *divergent*. In a *second-order transition* at least one of the second derivatives of  $\Omega$  is divergent. (If there are at most finite discontinuities in the second derivatives, the transition is of higher than second order.) Hence the order parameter  $M$  will vanish continuously at the transition point.

The susceptibility  $\chi_T$  and the specific heat  $c$  typically both diverge in a second-order transition. (Again it is not necessary that both of them diverge.) Here the divergences are power-law singularities. They are characterized by *critical indices*. For example, in the infinite-volume limit the susceptibility scales close to  $T_c$  according to  $\chi_T \equiv -(\partial^2 f / \partial H^2)_T \propto |1 - T/T_c|^{-\gamma}$  with an index  $\gamma$ ,  $f$  as defined above, and the specific heat according to  $c_H \equiv -T(\partial^2 F / \partial T^2)_H \propto |1 - T/T_c|^{-\alpha}$  with critical index  $\alpha$ .

#### a. The Landau free energy

One criterion for the order of the phase transition is given by Landau’s theory (Landau and Lifschitz, 1958; Aizu, 1970; Michel, 1980; Toledano, 1981). It consists in an expansion of the free energy in powers of the order parameter. The allowed terms in this expansion are further selected by symmetry arguments. Phase transitions can be classified according to the transformation behavior of their order parameters under a symmetry transformation. In this introductory section we discuss only an order parameter described by a scalar field  $\phi$ . In QCD applications the scalar  $\phi$  will be replaced by an  $O(N)$ -vector with  $N$  components or an  $SU(3)$  matrix parametrized by two independent directions of possible “ordering” (see Sec. IV).

The ansatz of a free-energy functional for a scalar order parameter  $\phi$  in  $d$  space dimensions is given as

$$F\{\phi(x)\} = \int d^d x \left\{ a(\nabla \phi(x))^2 + \frac{r}{2} \phi^2(x) + \frac{\lambda}{4} \phi^4(x) - h \phi(x) \right\}. \quad (2.2)$$

For vanishing  $h$  this is the simplest form that admits spontaneous symmetry breaking. Although  $F$  is reflection invariant (if  $h=0$ ), the ground state need not be so.  $F$  may take its minimum for nonvanishing values  $\pm \phi_0 \neq 0$  (later denoted as  $\langle \phi \rangle$ ), depending on the values of  $a$ ,  $r$ , and  $\lambda$ . The “couplings”  $a$ ,  $r$ ,  $\lambda$ , and  $h$  should be considered as parameters, where  $a$ ,  $\lambda > 0$ . In the example of a magnet, the condition  $\lambda > 0$  corresponds to the physical condition that the magnetization is bounded. Later  $r$  has the meaning of a mass squared,  $\lambda$  of a coupling strength of the interaction, and  $h$  of an

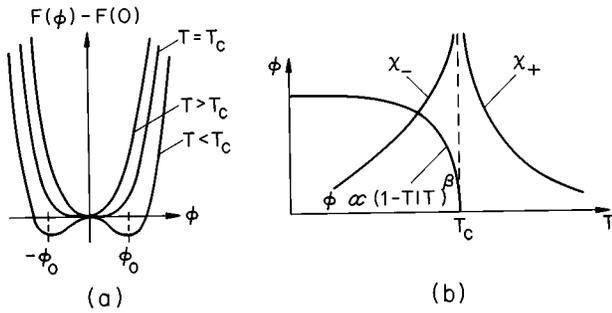


FIG. 1. Signature for a second-order phase transition: (a) Landau's free energy as function of the scalar order-parameter field  $\phi$ ; (b) order parameter  $\phi$  and associated susceptibility above ( $\chi_+$ ) and below ( $\chi_-$ )  $T_c$  as function of the temperature  $T$ .

external field. In the vicinity of a second-order transition the order parameter is small (more generally, if fluctuations in the field are allowed, its average expectation value is small), hence one drops higher powers of  $\phi$ . A further assumption is that  $\phi$  is slowly varying in space [thus there are no higher derivative terms than  $(\nabla\phi)^2$ ]. A  $\phi^3$  term is missing if a symmetry under sign inversion  $\phi \rightarrow -\phi$  is required for vanishing  $h$ . For  $h=0$  it is easily verified that Eq. (2.2) predicts a second-order phase transition. For  $r < 0$  two stable states are predicted with magnetizations  $\pm\phi_0$ . The condition  $r=0$  defines the critical temperature. Thus one may write for  $r$

$$r = \tilde{r}(T - T_c). \tag{2.3}$$

Figures 1(a) and 1(b) display the typical signatures of a second-order transition. The nontrivial minima at  $\pm\phi_0$  move continuously inwards as the temperature is increased towards  $T_c$ , where the  $Z(2)$  symmetry is restored [Fig. 1(a)]. Figure 1(b) illustrates the vanishing of the order parameter as a function of  $T$  and the power-law divergence in the susceptibility.

For fixed temperature  $T < T_c$ , a field-driven transition can be considered as a function of  $h$ . In this case the transition is of first order, and  $\phi$  jumps from  $\sqrt{(-r/\lambda)}$  to  $-\sqrt{(-r/\lambda)}$  as  $h$  changes sign.

Although the ansatz (2.2) was originally proposed for second-order transitions, temperature-driven first-order transitions can be described as well. We mention two possibilities.

(i)  $\lambda < 0$  in Eq. (2.2). The coupling  $\lambda$  can play the role of a renormalized coupling. In later applications of QCD,  $\lambda$  occurs as a renormalized coupling in an effective description of QCD. It varies as a function of the value that is chosen for the strange-quark mass  $m_s$ . For certain values of  $m_s$ ,  $\lambda$  becomes negative. If  $\lambda < 0$ , one has to include a term  $\propto \phi^6$  with positive coefficient to stabilize the free-energy functional (see Sec. IV.A.1). When we add a term  $(\kappa/6)\phi^6(x)$  to Eq. (2.2) ( $\kappa > 0$ ),  $F$  has two local minima over a certain temperature interval  $T_0 < T < T_1$ , where  $T_0$  and  $T_1$  have the meaning of stability limits of the disordered phase in the ordered, and the ordered phase in the disordered, respectively. At  $T_c$  the minima are equally deep, and the order pa-

rameter jumps from  $\phi_0 = \pm(3\lambda/4\kappa)^{1/2}$  to zero. In general  $\lambda$  may change its sign as function of an external parameter  $P$  ( $P \equiv m_s$  in our example of Sec. IV). When  $\lambda$  changes its sign at some value  $P^*$ , a line of second-order transitions ends at a so-called *tricritical point*  $T_t = T_c(P^*)$  and continues as a line of first-order transitions. Tricritical behavior is predicated to occur for QCD in certain limiting cases (see Sec. IV).

(ii) a cubic term in Eq. (2.2). If such a term is not suppressed by a symmetry argument, it admits a first-order temperature-driven transition. For  $\lambda > 0$ ,  $F$  has two minima at  $\phi_0 = 0$  and at  $\phi_0 \neq 0$  over a temperature interval  $T_0 < T < T_1$ . If the cubic term is written as  $(\alpha/3)\phi^3(x)$ , the nontrivial minimum at  $T_c$  occurs for  $\phi_0 = -9r/\alpha$ . The very existence of a cubic term in an effective potential for the electroweak phase transition has been under much debate in the last few years. In the simplest models for the electroweak transition the order parameter is an  $O(N)$  vector field, where  $r$ ,  $\lambda$ , and  $\alpha$  are renormalized parameters. Here we add a warning. A cubic term in the classical potential of Eq. (2.2) does not guarantee that the transition will be first order. If the transition is weakly first order in the sense that the maximum between zero and the nontrivial minimum is not high, the transition may be washed out by fluctuations of the order-parameter field. Such fluctuations will be discussed below.

Landau's concept of the free energy and the criterion for the order of the transition is widely used in applications of particle physics. Two caveats should be mentioned at least. The first one concerns convexity properties, the second the validity range of the mean-field approximation.

(i) *Convexity properties.* It is known from general thermodynamic principles that thermodynamic potentials in thermal equilibrium are convex functions of their variables. The nonconvex shape of Fig. 1 and the coexisting minima for a first-order transition are obviously at odds with the general expectation. Landau's free energy is a macroscopic concept. Its nonconvex and physical realization may be understood as a *coarse-grained free energy* arising in an intermediate step from a microscopic to a macroscopic scale.

To be specific, let us start with a microscopic spin Hamiltonian  $\mathcal{H}(\{s_i\})$  depending on spin variables  $s_i$  associated with sites  $i$  of a hypercubic lattice. As a first step short-wavelength fluctuations are eliminated by dividing the lattice into cells of linear dimensions  $L$  and introducing new variables  $\phi(x)$  in a block spin transformation according to

$$\phi(x) = \frac{1}{L^d} \sum_{i \in L^d(x)} s_i. \tag{2.4}$$

Here  $x$  is a site on the block lattice, and the block spin variable  $\phi$  is identified with an order-parameter field as it enters the ansatz (2.2). In this way we make the connection to Landau's "mean-field" free energy  $F = F_{mf}$ . The dynamics of a coarse-grained block lattice with field variables  $\phi(x)$  is determined by the so-called coarse-grained Hamiltonian  $\mathcal{H}_{CG}$ . It is obtained as

$$e^{-\beta\mathcal{H}_{CG}[\phi(x)]} = \int \mathcal{D}s_i P([\phi(x)], [s_i]) e^{-\beta\mathcal{H}[s_i]}, \quad (2.5)$$

where the path integral extends over all spin configurations  $[s_i]$  under the constraint that a particular configuration  $[\phi(x)]$  on the block lattice [determined according to Eq. (2.4)] is kept fixed. The constraint is denoted by  $P(\dots)$ . Such a block spin transformation may be iterated. The general folklore is then (although it is extremely hard to prove rigorously) that, after a sufficient number of iterations,  $\mathcal{H}_{CG}[\phi(x)]$  has the form of the Landau free energy  $F = F_{mf}$  if one is close to a second-order phase transition. At such a transition the correlation length  $\xi$  diverges, and the condition  $L \ll \xi$  on the linear cell size  $L$  is easily satisfied.

The bulk free energy  $F$  of a system governed by the Hamiltonian  $H_{CG}$  is then obtained by integrating over all remaining configurations  $[\phi]$  according to

$$F = -\frac{1}{\beta} \ln Z = -\frac{1}{\beta} \ln \int \mathcal{D}\phi \exp\{-\beta F_{mf}[\phi(x)]\}. \quad (2.6)$$

If it is justified to evaluate Eq. (2.6) in a saddle-point approximation, i.e., to drop  $\int \mathcal{D}\phi$  and take  $F_{mf}\{\}$  at its minimum  $\phi_0$ , the result is

$$F = F_{mf}. \quad (2.7)$$

Thus the free energy coincides with its mean-field value, the Landau free energy. Landau's free energy is inherently a mean-field approximation. For a spin system, the order parameter is the magnetization. The Landau free energy is an expansion in terms of mean values of spins. It does not include a summation over all spin configurations according to

$$F = -\frac{1}{\beta} \ln \int \mathcal{D}s_i \exp\{-\beta\mathcal{H}[s_i]\}. \quad (2.8)$$

(ii) *The validity of the mean-field approximation.* Next we address the question of when one is allowed to ignore fluctuations in the order-parameter field, i.e., to drop  $\int \mathcal{D}\phi$  in Eq. (2.6) and take  $F_{mf}$  for  $F$ . On a microscopic level this amounts to a replacement of spin-spin interactions by some average background represented by  $\phi \equiv M$ . The validity of the mean-field approximation is guaranteed if the fluctuations  $\delta\phi \equiv \phi(x) - \phi_0$  in the order-parameter field are small compared to the order parameter itself, i.e.,

$$\langle (\delta\phi(x))^2 \rangle \ll \phi_0^2, \quad (2.9)$$

where the average  $\langle \dots \rangle$  should be taken over all cells of the coarse-grained lattice. Upon using the fluctuation dissipation theorem, we find that Eq. (2.9) translates to (see, for example, Binder, 1987)

$$1 \ll R^d (1 - T/T_c)^{(4-d)/2}, \quad (2.10)$$

where  $R$  is the interaction range. Mean-field theory becomes exact if the dimensionality, the range of interactions, or the number of interacting neighbors becomes infinite (the large- $N$  approximation, where  $N$  is the number of order-parameter components).

Mean-field theory often gives correct qualitative predictions for phase diagrams of three-dimensional systems, which can be effective models for high-temperature QCD. As the dimensionality increases, mean-field theory improves, while numerical calculations get harder. The critical dimension depends on the form of the Landau free-energy expansion. Symmetry considerations are essential for constructing Landau's free energy. As both finite-temperature transitions of QCD are supposed to be driven by symmetry breaking, it is natural to construct effective actions for QCD following Landau's assumptions.

If condition (2.10) is violated, one should use the renormalization-group approach (Wegner, 1972; Fisher, 1974; Wilson and Kogut, 1974; Ma, 1976) to describe critical phenomena.

#### b. The renormalization-group approach in the infinite-volume limit

In this section we review the main steps of the renormalization-group approach in the *infinite-volume limit* as a tool for describing critical phenomena for second-order transitions. This approach leads to predictions of critical indices and scaling relations between them.

In the renormalization-group approach one attempts to solve the path integral (2.8) in iterated steps. (It is useful to visualize these steps as block spin transformations in a spin system.) One step refers to one application of a renormalization-group transformation  $R_b$  in coordinate space. The transformation increases the scale by a factor  $b > 1$ , hence  $b$  is called the scale factor. A series of Hamiltonians  $\mathcal{H}^{(n)}$  is generated along with the iterated application of  $R_b$ , and depends on the fields  $\phi^{(n)}$  and couplings  $g^{(n)}$  after  $n$  steps. For simplicity let us denote the set of couplings after  $n$  steps as  $g$  and after  $(n+1)$  steps as  $g^{(1)}$ . The relation between  $\mathcal{H}^{(n+1)}$  and  $\mathcal{H}^{(n)}$  implies for the coarse-grained free-energy density  $f(g)$  after one application of  $R_b$

$$f(g) = G(g) + b^{-d} f(g^{(1)}) \quad (2.11)$$

with the same functional dependence  $f$  on both sides of the equality sign. The rescaling factor  $b^{-d}$  is necessary to account for the reduction of the effective degrees of freedom after one step by  $N^{(1)} = b^{-d} N$ . To keep the free energy  $F = \int d^d x f(x)$  constant under the operation  $R_b$ , one has to compensate for the reduced volume on the block lattice by a factor  $b^{-d} < 1$ . The term  $G(g)$  denotes the contribution to  $f$  which is regular in  $g$ .

Similarly the *correlation length* on the "block lattice" is reduced by a factor  $b^{-1}$  in one step,

$$\xi(g^{(1)}) = b^{-1} \xi(g). \quad (2.12)$$

At criticality ( $T = T_c$ ),  $\mathcal{H}^{(n \rightarrow \infty)}$  is supposed to converge to a nontrivial fixed-point Hamiltonian  $\mathcal{H}(g^*)$ , which is invariant under the transformation  $R$ . The set of fixed-point couplings  $g^*$  that is invariant under  $R$  implies

$$\xi(g^*) = b^{-1} \xi(g^*). \quad (2.13)$$

Equation (2.13) is only compatible with a vanishing or a diverging correlation length, since  $b > 1$  by assumption.

Singularities of thermodynamic functions arise as the critical point is approached. Their power-law divergences are characterized by critical exponents. Critical exponents can be derived in the renormalization group approach, when the transformation  $R$  acting in the parameter space of couplings is linearized in the vicinity of the fixed point. Let  $\mathcal{H}(g)$  denote a Hamiltonian “close to” the fixed-point Hamiltonian, close in the sense that  $g \approx g^*$ , such that we can write

$$\mathcal{H}(g) = \mathcal{H}(g^*) + \sum_i q_i O_i + o \left[ \left( \sum_i q_i^2 \right)^{1/2} \right]. \quad (2.14)$$

The scaling operators  $O_i$  are taken as eigenfunctions of  $R$ , while the coefficients  $q_i$  are the *scaling fields*, measuring the distance in coupling parameter space between  $g$  and  $g^*$ . In a typical second order transition we have

$$\begin{aligned} q_1 &= k_1 t, \\ q_2 &= k_2 h, \\ t &= (T - T_c)/T_c, \end{aligned} \quad (2.15)$$

where  $k_1$  and  $k_2$  are constants,  $t$  is the reduced temperature, and  $h$  is an external field. The eigenvalues  $\lambda_i$  of the scaling operators  $O_i$  may be written as  $b^{\lambda_i}$  due to an imposed semigroup property of  $R$ . Thus we have

$$R_b \mathcal{H}(g) = \mathcal{H}(g^*) + \sum_i q_i b^{\lambda_i} O_i \quad (2.16)$$

under the action of  $R$ . Scaling fields for which  $\lambda_i > 0$  are called *relevant*, those with  $\lambda_i = 0$  *marginal*, and those with  $\lambda_i < 0$  *irrelevant*. In the linearized version of Eq. (2.16), the relevant and marginal scaling fields (or couplings) have to vanish (e.g.,  $T \rightarrow T_c$  and  $h \rightarrow 0$ ) for  $\mathcal{H}(g)$  converging to  $\mathcal{H}(g^*)$ .

The free-energy density  $f$  and the correlation length  $\xi$  can be expressed as functions of the scaling fields by using Eqs. (2.11) and (2.12). For a concrete application one has to specify the relevant scaling fields. Further fields  $q_\alpha$  refer to marginal and irrelevant scaling fields. As we are interested in a temperature-driven second-order phase transition in the presence of an external field  $h$  (for applications to QCD, see Sec. III.C.3), we choose  $q_1$  and  $q_2$  as the only relevant scaling fields according to Eq. (2.15). A frequently used notation is  $\lambda_1 \equiv y_t$  for the thermal exponent and  $\lambda_2 \equiv y_h$  for the magnetic exponent. Here we only indicate the subsequent steps. The singular parts of  $f$  and  $\xi$  are written in terms of the reduced temperature times a *scaling amplitude*. The scaling amplitudes are assumed to admit an analytic expansion in their arguments. This implies that in the vicinity of  $T_c$  ( $|t| \ll 1$ ) contributions of irrelevant scaling fields may be neglected. We explicitly mention this assumption because the condition  $|t| \ll 1$  is easily violated in numerical applications.

The asymptotic behavior of  $f$  is then given as

$$f(t, h=0) \approx G(0) + A_f^{(\pm)} |t|^{d/\lambda_1}, \quad (2.17)$$

where  $A_f^{(\pm)}$  denotes the scaling amplitude for  $f$ , and the signs  $(\pm)$  refer to the approach of  $T_c$  from above or below. This behavior implies the scaling relation

$$2 - \alpha = d/\lambda_1, \quad (2.18)$$

where  $\alpha$  is the critical exponent characterizing the scaling behavior of the specific heat close to  $T_c$ . For the correlation length one obtains

$$\xi(t, h=0) \approx A_\xi^{(\pm)} |t|^{-1/\lambda_1}, \quad |t| \ll 1. \quad (2.19)$$

From Eq. (2.19) one can see that the critical exponent  $\nu$  characterizing the divergence of the correlation length may be identified as

$$\nu = 1/\lambda_1. \quad (2.20)$$

Finally we denote the relations for the magnetization  $M$  and the susceptibility  $\chi$  in the zero-field limit ( $h=0$ ),

$$M(t, 0) = \left. \frac{\partial f}{\partial h} \right|_{h=0} \approx M_0 + A_M^{(-)} (-t)^{d/\lambda_1 - \lambda_2/\lambda_1}, \quad t < 0, \quad (2.21a)$$

$$\chi(t, 0) = \left. \frac{\partial^2 f}{\partial h^2} \right|_{h=0} \approx \chi_0 + A_\chi^{(\pm)} |t|^{d/\lambda_1 - 2\lambda_2/\lambda_1}, \quad |t| \ll 1 \quad (2.21b)$$

with splitting analogous to Eq. (2.17) in the singular parts and additive analytic contributions  $M_0$  and  $\chi_0$ , originating in derivatives of the analytic part  $G$ . The exponent  $\lambda_1$  is distinguished as long as we consider temperature-driven transitions.

Note that only the leading singular terms in the expansion of  $A_f^{(\pm)}|_{h=0}$  and  $A_\xi^{(\pm)}|_{h=0}$  in powers of  $(q_\alpha |t|^{-\lambda_\alpha/\lambda_1})$  with  $\lambda_\alpha/\lambda_1 < 0$  have been kept in Eqs. (2.21), (2.19), and (2.17). As emphasized above, this is justified as long as  $|t| \ll 1$ . Due to critical slowing down, it is inherently difficult in a Monte Carlo simulation to satisfy the condition  $|t| \ll 1$ . If one observes, in measured critical indices, deviations from theoretical expectations, one should keep in mind that contributions from irrelevant terms are one possible explanation.

## 2. Finite-size scaling analysis

Monte Carlo simulations are necessarily performed in a finite volume. Effects of the finite size are certainly important, both when the volume is small compared to physical length scales and when the physical scales are given by long-wavelength fluctuations or massless modes. Massless modes feel the finite volume however large the size of the system.

It is well known that finite-size effects render it difficult to infer the order of a phase transition. Even qualitative conclusions may be misleading. A truly second-order transition may look like first order showing a double-peak structure in the probability distribution of the internal energy or magnetization. The double-peak structure is a typical signal for the coexistence of phases at a first-order transition. In this case, however, it is a

finite-size artifact. It would eventually vanish if the volume were further increased.

Conversely, large correlation lengths in a finite volume suggest a second-order transition. The correlation length may be truly large, but finite. The finite value would be visible if the volume were further increased, through a further increase might involve a practical problem of computer time. A second possible reason for a large correlation length is a particle-like excitation in the spectrum, which gets its mass from the volume and decouples in the infinite-volume limit. In such a case the associated correlation length does not correspond to a bulk correlation length. Moreover, one has to be careful about the order of the critical temperature ( $T \rightarrow T_c$ ) and the thermodynamic ( $V \rightarrow \infty$ ) limits. If  $T \rightarrow T_c$  is taken first, the correlation length diverges in the case of a (truly) first-order transition. A misinterpretation of data for a correlation length was at the heart of the controversy over the order of the deconfinement transition in the pure SU(3) gauge theory (see Sec. III.B.1).

Fortunately there are characteristic signatures in a finite volume, which anticipate the behavior in the thermodynamic limit. A careful analysis of these signatures allows one to determine the order of the transition. The measurements should be performed as a function of a varying volume. A measurement for a single size is not conclusive in general.

In the following we discuss the finite-size scaling behavior of singularities in thermodynamic functions. We consider the case of second-order transitions first.

### a. Second-order transitions

The renormalization-group approach is an appropriate framework in which to discuss finite-size effects (see, for example, Barber, 1983). It puts the size of the system on an equal footing with other relevant scaling fields like temperature, mass, and magnetic field.

The central conjecture of a finite-size scaling analysis is to a thermodynamic function  $P$ , which admits a critical singularity in the infinite-volume limit, characterized by an index  $\mu$ ,

$$P_\infty(t) \sim A_P^\pm |t|^{-\mu}, \quad (2.22)$$

where the reduced temperature  $t$  has been defined in Eq. (2.15), and  $A^\pm$  are the amplitudes obtained in the limit  $T \rightarrow T_c$  with  $t > 0 (+)$  or  $t < 0 (-)$ . Consider a system, that is finite in at least one dimension with linear size  $L$ . The conjecture about the finite-size behavior of  $P$  is

$$P_L(t) = |t|^{-\mu} Q_P^{(\pm)}(L/\xi_\infty(t)), \quad (2.23)$$

where  $L$  is assumed to be large and  $T$  close to  $T_c$ ,  $L \gg 1$ ,  $|t| \ll 1$ , and  $\xi_\infty$  denotes the correlation length in the infinite-volume limit. The derivation of Eq. (2.23) in a renormalization-group approach is essentially the same as in the infinite-volume limit. The additional ingredients are arguments for why  $L$  can be considered as a scaling field. There may be a conflict in principle. Already in a single renormalization transformation nonlo-

cal terms are generated in the Hamiltonian. Thus the system size should be sufficiently large to account for all nonlocal couplings one wants to keep.

The assumption is that close to the critical point one need keep only a few couplings. Once this set of couplings has been fixed,  $L$  can be chosen sufficiently large compared to the typical interaction range. In further distinction to the infinite-volume limit, the renormalization-group transformations can be iterated only a finite number of times, as the scale increases by a factor  $b > 1$  in each step. Here the allowed number of steps should be sufficiently large. It is also taken for granted that the stability of a set of fixed-point couplings  $g^* = R_b g^*$  and the semigroup property of  $R$  are preserved in a finite volume. Under these assumptions  $L^{-1}$  may be treated as an additional scaling field.

The ansatz for the free-energy density in terms of scaling fields and eigenvalues of  $R$  then takes a similar form to that of the infinite-volume analysis. It is given as

$$f(k_1 t, k_2 h, q_\alpha, L) = G(k_1 t, k_2 h, q_\alpha) + |t|^{d/\lambda_1} f^{(\pm)} \times (h|t|^{-\varphi_2}, q_\alpha |t|^{-\varphi_2}, L|t|^{1/\lambda_1}), \quad (2.24)$$

where  $b$  has been chosen large enough that  $b^{\lambda_1} |t| \sim 1$  and  $\phi_2$  stands for  $\lambda_2/\lambda_1$ . Equation (2.24) implies the finite-size scaling behavior of various derived quantities. For the zero-field susceptibility we find

$$\begin{aligned} \chi_L(t) &= \chi(t, h=0, L) = \left. \frac{\partial^2 f}{\partial h^2} \right|_{h=0} \\ &= \chi_0 + A_\chi^{(\pm)} (L|t|^{1/\lambda_1}) |t|^{d/\lambda_1 - 2\varphi_2}, \end{aligned} \quad (2.25)$$

Irrelevant terms have been dropped on the right-hand side of Eq. (2.25). The analytic part  $\chi_0$  stays finite as  $T \rightarrow T_c$ . It may be dropped for large volumes if the second term is singular as  $T \rightarrow T_c$ . If the argument of  $A_\chi^{(\pm)}$  is expressed in terms of  $\chi_\infty(t)$ , it is easily shown that the scaling behavior of  $\chi$  takes the form of Eq. (2.23) with  $P \equiv \chi$  and  $\mu \equiv \gamma$ .

Let us see what Eqs. (2.22) and (2.23) imply for the explicit  $L$  dependence of  $P$ . In order to guarantee that

$$\lim_{L \rightarrow \infty} P_L(t) = P_\infty(t), \quad (2.26)$$

the amplitude  $Q_P^{(\pm)}$  has to satisfy

$$\lim_{x \rightarrow \infty} Q_P^{(\pm)}(x) = A_P^{(\pm)}, \quad (2.27)$$

where  $x \equiv L/\xi_\infty$ . The limit  $x \rightarrow 0$  is realized if  $T \rightarrow T_c$  for fixed  $L < \infty$ . This is the limit we are interested in. In this case  $P_L(t=0)$  should be finite. Hence the singularity of  $P_\infty(t)$  has to be compensated for by the scaling function  $Q_P^{(\pm)}(x)$  according to

$$Q_P^{(\pm)}(x) \sim x^{\mu/\nu} \quad \text{for } x \rightarrow 0, \quad (2.28)$$

where we have used  $\xi_\infty(t) \sim |t|^{-\nu}$  for  $t \rightarrow 0$ . It follows from Eqs. (2.23) and (2.28) that

$$P_L(T_c) = P_L(t=0) \sim L^{\mu/\nu}. \quad (2.29)$$

The last relation (2.29) predicts for the susceptibility

$$\chi_L(T_c) \sim L^{\gamma/\nu} \quad (2.30)$$

and for the specific heat

$$c_L(T_c) \sim L^{\alpha/\nu}. \quad (2.31)$$

Finally we state the scaling behavior of the correlation length in a finite volume:

$$\xi_L(t) = \xi_\infty(t) Q_\xi^{(\pm)}(L/\xi_\infty(t)). \quad (2.32)$$

*Refined criteria* for a second-order phase transition take finite lattice sizes into account according to Eqs. (2.30), (2.31), and (2.32). Note that these formulas have been derived under the conditions  $L \gg 1$  and  $|t| \ll 1$ . One should make sure that these conditions are satisfied in the actual simulations, otherwise subleading corrections must be kept.

An important feature of finite-size scaling analysis in the renormalization-group approach is the similar treatment of the system size and other scaling fields. Treatment on an equal footing is manifest in Eq. (2.24). The inverse extension  $1/L$  appears as a relevant scaling field with eigenvalue  $\lambda_i = 1$ ,  $i = 1/L$ . To obtain the critical limit, all scaling fields that are relevant with respect to the fixed-point set of couplings have to vanish:  $t \rightarrow 0$  ( $T \rightarrow T_c$ ),  $h \rightarrow 0$ , and  $1/L \rightarrow 0$ . As long as  $\xi_\infty(t) \ll L$ , finite-size effects are negligible. If  $|t|$  is so small that  $(\xi_\infty(t)/L) \sim 0(1)$ , or if  $L$  is not sufficiently large, finite-size corrections may drastically alter the thermodynamic behavior.

Thus the formal treatment of  $1/L$  is the same as the treatment of an external field. The external field may be a mass. From physical applications it is well known that the finite volume can act like a mass. An obvious manifestation is the finite correlation length as result of a finite volume or of a nonvanishing mass. The analogy goes even further. Consider an Ising model in four dimensions in a finite volume. In the broken phase there is always a finite probability of tunneling between degenerate states with magnetization  $\pm 1$ . Below  $T_c$  the Ising model can be approximated by the effective potential of an anharmonic oscillator, where the mass in the effective action is proportional to the volume. The corresponding ‘‘particle’’ decouples in the infinite-volume limit. The kink describing tunneling between states of opposite magnetization is associated with a so-called vacuum tunneling energy. This energy goes to infinity along with the volume (spin flips of the entire system are rather unlikely in a very large volume; see Meyer-Ortmanns, 1989).

### b. First-order transitions

Finite-size scaling analysis for first-order transitions is inherently more difficult than for second-order transitions. The correlation length stays finite even in the infinite-volume limit as  $T \rightarrow T_c$  after  $L \rightarrow \infty$ , so  $\xi_\infty/L$  is no longer a sensible scaling variable as in Eq. (2.32). In the second-order case it is the thermal eigenvalue  $1/\nu$  of the renormalization-group transformation that controls

rounding and shifting of the algebraic singularities, both being of the order of  $L^{-1/\nu}$ .

The goal of finite-size scaling analysis for first-order transitions is to predict the rounding and shifting of  $\delta$ -function singularities in the second derivatives of a thermodynamic potential due to the finite volume (e.g., in the specific heat due to a latent heat, or in the susceptibility due to a jump in the order parameter).

The volume has to be finite in at least one direction. This time the rounding of the singularity, the width of the sharp crossover region, and the shift in the transition temperature are predicted to be of the order  $L^{-d}$ , where  $d$  is the dimension of the system.

Before we go into the subtleties of refined criteria, which are based on the precise form of the rounding and shifting as a function of  $L$ , we summarize some qualitative signatures of a first-order transition in a typical Monte Carlo simulation. Qualitative signatures are essentially the large or infinite-volume signatures that we call ‘‘naive’’ criteria for inferring the order of the transition:

(i) Some thermodynamic quantities (e.g., the internal energy) are almost discontinuous at the transition.

(ii) A starting configuration that is half ordered and half disordered relaxes to very different equilibrium states on both sides of  $T_c$  (rather than frequently tunneling between both sides). In a second-order transition the system relaxes to an equilibrium configuration independently of the initial condition.

(iii) At infinite volume, tunneling between both phases is completely suppressed; at small volumes tunneling may be mixed with fluctuations of statistical origin. For large volumes *tunneling events* are clearly visible in a Monte Carlo simulation. The system is in the ordered phase with a probability  $\exp\{-L^d F_0(\beta)\}$ , and in the disordered phase with a probability  $\exp\{-L^d F_d(\beta)\}$ , where  $F_0$  and  $F_d$  denote the free energies in the ordered and disordered phases, respectively. If the time history is followed over a number of Monte Carlo iterations, it shows flip-flops between states of different ‘‘magnetizations.’’ The frequency of flip-flops decreases with increasing volume.

(iv) Another manifestation of tunneling shows up in the probability distribution  $P_E$  for the internal energy  $E(\beta)$ . For large volumes it is sharply peaked at the energy values of the ordered ( $E_0$ ) and disordered ( $E_d$ ) phases. The deep valley between these peaks reflects the rare number of tunneling events. If the initial condition is an ordered start, the probability is large to find  $E_0$  for  $E$ . For a disordered start, it is large to find  $E_d$ . This is nothing but a sign of *metastability*.

(v) *Hysteresis effects* are observed even away from the transition temperature.

The qualitative signature of a pronounced double-peak structure can be made more quantitative when it is analyzed as a function of the lattice size. In this way we are led to *refined criteria*.

A precise form for finite-size scaling has been sought following different approaches. One possibility is the incorporation of first-order transitions in the

renormalization-group approach. First-order transitions are then associated with a discontinuity fixed point at zero temperature (Nienhuis and Nauenberg, 1975; van Leeuwen, 1975). In this approach one would derive the finite-size scaling as special cases of Eq. (2.24). For example, Eq. (2.30) for the susceptibility  $\chi_L \sim L^{(\gamma/\nu)} = L^{(2\lambda_2 - d)}$  gives a scaling proportional to the volume  $\chi_L \sim L^d$  with  $\lambda_2 = d$ . Similarly Eq. (2.31) for the specific heat  $c_L \sim L^{a/\nu} = L^{2\lambda_1 - d}$  gives  $c_L \sim L^d$  with  $\lambda_1 = d$ . The dimensionality  $d$  is the only eigenvalue of the discontinuity fixed point.

Scaling behavior at a first order transition can also be discussed independently of the existence of such a discontinuity fixed point. A phenomenological approach was developed by Imry (1980), Fisher and Berker (1982), Privman and Fisher (1983), Binder and Landau (1984), and Challa *et al.* (1986).

More recently exact results have been obtained by Borgs and Kotecky (1990) and Borgs *et al.* (1991) for models that can be represented by a contour expansion with small activities like the  $q$ -state Potts model for large  $q$ . The partition function  $Z(\beta, L)$  in a finite lattice of volume  $L^d$  with periodic boundary conditions is expanded according to

$$Z(\beta, L) = e^{-L^d \beta f_d(\beta)} + q e^{-L^d \beta f_o(\beta)} + \mathcal{O}(e^{-bL}) e^{-\beta f(\beta) L^d}, \quad b > 0. \quad (2.33)$$

The free energy is the minimum of  $f_d$  and  $f_o$ , where the indices  $o$  and  $d$  stand for ordered and disordered, and  $b$  is a constant larger than zero. In what follows we discuss predictions of the phenomenological two-peak Gaussian model for the energy probability distribution  $P_L(E)$ , which was been introduced by Binder and Landau (1984) and Challa *et al.* (1986). This model is related to the rigorous expansion (2.33) by an inverse Laplace transform. Therefore results of the phenomenological model may be compared to exact results and improved in a more systematic finite-size scaling analysis (Billoire, Gupta, *et al.*, 1990; Billoire, Lacaze, Morel, 1990). We present part of their results in the following.

### c. The two-peak Gaussian model

The basic quantity for temperature (energy  $E$ )-driven transitions in the first-order case is the probability distribution  $P_L(E)$  of the internal energy. [For field-driven transitions  $E$  is replaced by the magnetization  $s$  (Binder and Landau, 1984).] For large volumes and away from the transition point,  $P_L(E)$  is given by a simple Gaussian. This can be seen when the exponent is expanded around its maximum as a function of the internal energy. The double-peak structure of  $P_L(E)$  in the vicinity of a first-order transition is then easily understood when one notices that the exponent is extremal for two values,  $\tilde{E}_o(\beta)$  and  $\tilde{E}_d(\beta)$ , one for each of the coexisting phases. As a phenomenological ansatz, the energy probability density  $P_L(E)$  is now replaced by a *sum of two Gaussian distributions*. It is plausible that

their width should be determined by the distance from  $T_c$ . Expanding  $\tilde{E}(\beta)$  around  $T_c$  leads to

$$P_L(E) = A \left[ \frac{a_o}{\sqrt{c_o}} e^{-\frac{L^d \beta_c^2 (E - E_o - c_o \delta T)^2}{2c_o}} + \frac{a_d}{\sqrt{c_d}} e^{-\frac{L^d \beta_c^2 (E - E_d - c_d \delta T)^2}{2c_d}} \right]. \quad (2.34)$$

Here  $E_{o,d}$  and  $c_{o,d}$  are the infinite-volume energy and specific heat in the pure ordered and disordered phases, respectively,  $A$  is a normalization factor,  $\delta T$  is defined as  $\delta T = (\beta_c - \beta)/\beta_c^2$ , and  $a_o, a_d$  are weight factors for each phase,

$$a_o = q e^{-\Delta}, \quad a_d = e^{\Delta}. \quad (2.35)$$

The factor  $q$  corresponds to  $q$  ways of realizing the broken phase in the  $q$ -state Potts model;  $q=2$  for the ( $\pm$ ) magnetizations in the Ising model. Finally  $\Delta$  is given as

$$\Delta = \frac{L^d}{2} \beta (F_o(\beta) - F_d(\beta)) \sim \frac{L^d}{2} (\beta - \beta_c) \left( E_o - E_d + \frac{1}{2} (c_o - c_d) \delta T \right). \quad (2.36)$$

Equations (2.34)–(2.36) predict energy moments correctly up to order  $L^{1/d}$ .

A specific feature of a first-order transition is the sharpening of the double-peak structure as the volume increases. Tunneling between phases is less likely the larger the volume is. The spin configurations corresponding to mixed states lead to an energy  $E$  that lies between the peaks at  $E_o$  and  $E_d$ . [The distribution of these “intermediate” energies is not correctly represented by the superposition of Gaussians according to Eq. (2.34).] Contributions of mixed states vanish in the infinite-volume limit, as long as they are equilibrium configurations. Nothing is said about the phase conversion itself, which may proceed as an off-equilibrium process with “droplet” formation.

The double-peak structure and the criteria we present in the following are essentially based on the double-valued internal energy  $E_o$  and  $E_d$  at the transition point. The free energies are degenerate in this case.

Let us first consider rounding in the specific heat and the shifting of the transition point. The specific heat in the volume  $L^d$  is, as usual, obtained as

$$c_L = L^d \beta^2 (\langle E^2 \rangle_L - \langle E \rangle_L^2), \quad (2.37)$$

where

$$\langle E \rangle_L = \int_{-\infty}^{+\infty} P_L(E) E dE$$

with  $P_L(E)$  given by Eqs. (2.34)–(2.36). Its maximum  $c_{L \max}$  occurs for an inverse temperature

$$\beta(c_{L \max}) = \beta_c - \frac{\ln q}{E_d - E_o} \frac{1}{L^d} + \mathcal{O}(1/L^{2d}). \quad (2.38)$$

Equation (2.38) shows that the shift of the critical temperature in the finite-volume value compared to the

infinite-volume value is of the order of  $1/L^d$ . The height of the maximum in the specific heat  $c_{L \max}$  is derived to be proportional to the volume,

$$c_{L \max} = L^d \frac{\beta_c^2}{4} (E - E_d)^2 + \text{terms of order } (1). \quad (2.39)$$

Thus Eqs. (2.38) and (2.39) display the anticipated volume dependence of shifting and rounding in a first-order temperature-driven transition. Subleading corrections of  $O(1/L^{2d})$  in Eq. (2.38) and of  $O(1)$  in Eq. (2.39) have been determined by Lee and Kosterlitz (1991).

*d. Binder's cumulant*

Instead of calculations of the  $P_L(E)$  distribution itself, certain moments and cumulants of the energy probability distribution have been proven to be suitable indicators of the order of the transition in a finite volume. An important example is Binder's cumulant (Challa *et al.*, 1986), defined as

$$B = \frac{1}{3} \left( 1 - \frac{\langle E^4 \rangle}{\langle E^2 \rangle^2} \right). \quad (2.40)$$

It vanishes in the infinite-volume limit for all temperatures apart from the transition, where  $B$  gets a finite value from tunneling between coexisting phases of a first-order transition. It also vanishes in the case of a second-order transition. Calculating the moments  $\langle E^4 \rangle$  and  $\langle E^2 \rangle$  from Eqs. (2.34)–(2.36) with the use of Eq. (2.37), we find that  $B$  takes its minimum  $B_{\min}$  at (Billoire, Gupta, *et al.*, 1990; and Billoire, Lacaze, and Morrel, 1990)

$$B_{\min} = \frac{-(E_o^2 - E_d^2)^2}{12(E_o E_d)^2} + O(1/L^d) \quad (2.41)$$

at an inverse temperature

$$\beta(B_{\min}) = \beta_c - \frac{\ln[q(E_o/E_d)^2]}{E_d - E_o} \frac{1}{L^d} + O\left(\frac{1}{L^{2d}}\right). \quad (2.42)$$

The shifting of the critical temperature in Eq. (2.42) is again of order  $1/L^d$ . A nonvanishing value of  $B_{\min}$  signals a first order transition. Analogous formulas can be derived for the spin probability distribution  $P_L(s)$ , implying that the maximum of the susceptibility is proportional to the volume.

Binder's cumulant depends on the choice of an arbitrary additive constant, which may be added to the energy. This has led to consideration of the quantity  $U4$  (Billoire *et al.*, 1992)

$$U4 \equiv \frac{\langle (E - \langle E \rangle)^4 \rangle}{\langle (E - \langle E \rangle)^2 \rangle^2}. \quad (2.43)$$

Apart from a first-order transition point,  $U4$  can be shown to be larger than 1. Its minimum is given as

$$U4_{\min} = 1 + \frac{8(C_o + C_d)}{L^d \beta_c^2 (E_o - E_d)^2} + O(1/L^{2d}). \quad (2.44)$$

The power-law corrections to  $C_{L \max}$ ,  $B_{\min}$ , and  $U4_{\min}$  should be contrasted with exponential corrections to

TABLE I. Criteria for distinguishing between first- and second-order transitions in a large but finite volume.

Criterion	First order	Second order
$P_L(E)$	double peak	single peak
$c_{L \max}$	$\propto L^d$	$\propto L^{\alpha/\nu}$
$\chi_{\max}$	$\propto L^d$	$\propto L^{\gamma/\nu}$
$\beta(c_{L \max}) - \beta_c$	$\propto L^{-d}$	$\propto L^{-1/\nu}$
$B_{\min}$	$>0$	$\rightarrow 0$
$U4_{\min}$	$\rightarrow 1$	$>1$

bulk quantities like the average internal energy  $E$ , which are directly obtained as derivatives of  $Z$ . According to Eq. (2.33),  $Z$  itself has only exponential corrections from the finite volume.

In Table I we summarize criteria for distinguishing first- and second-order transitions in a finite volume. The indicated volume dependence should be understood as a leading term in a large-volume expansion. Similarly the numbers 0 or 1 indicate limiting values, which are approached as  $T \rightarrow T_c$ . Recall that  $\alpha/\nu$  and  $\gamma/\nu \leq d$  for a second-order transition. Thus a typical test in a Monte Carlo simulation could be a calculation of  $c_{L \max}/L^d$ . If this ratio as a function of  $L^d$  goes to a nonvanishing constant for large values of  $L$ , a first-order transition is signalled; if it approaches zero, the transition must be of second or higher order.

The two-phase coexistence at  $T_c$  refers to the Ising model. The analytic results for leading and subleading finite-size scaling behavior have been derived by Billoire *et al.* (1992) for the more general case of  $d$ -dimensional  $q$ -state Potts models, which include the familiar Ising model for  $q=2$ . Potts models have served as a testing ground for the above criteria.

*e. q-state Potts models*

The Hamiltonian of a  $d$ -dimensional  $q$ -state Potts model is given as

$$H = - \sum_{\langle ij \rangle} \delta_{s_i, s_j}. \quad (2.45)$$

The spin variables  $s_i$  are associated with sites on a  $d$ -dimensional hypercubic lattice. They can take  $q$  different integer values. The sum extends over nearest-neighbor pairs  $\langle ij \rangle$ . The symmetry group leaving  $H$  invariant is the permutation group of  $q$  elements. The  $q=3$  Potts model in three dimensions plays a distinguished role in QCD. It shares the global  $Z(3)$  symmetry with the pure  $SU(3)$  gauge theory. The restoration of the spontaneously broken  $Z(3)$  symmetry at finite temperature is assumed to drive the deconfinement transition in the pure gauge theory.

Beyond this application to QCD, Potts models provide a suitable testing ground for numerical methods. Exact results are available for comparison in certain special cases of  $q$  and  $d$ . The strength of the first order varies as a function of  $q$ . The  $q=5$  model has a weakly first-order transition accompanied by a tiny latent heat

of  $\Delta Q = E_d - E_o = 0.05292$  in lattice units and a large correlation length of 2000 in lattice units at the transition point (Billoire, 1991). The transition in the  $q=10$  model is strongly first order,  $\Delta Q = 0.69605$  and  $\xi \sim 6$ , both quantities measured in lattice units (Billoire, Gupta, *et al.*, 1990). Obviously the transition in the  $q=5$  Potts model is easily misinterpreted as being of second order. The linear lattice size has to exceed the large value of  $\xi \sim 2000$ .

Simulations of the  $2d-q=10$  Potts model have been performed by Billoire *et al.* (1992) to check the limiting behavior of  $(c_{L \max}/L^d)$ ,  $B_{\min}$ ,  $U_{4 \min}$ , and  $c(\beta_t)$  ( $\beta_t$  being the infinite-volume transition point). According to the above predictions these quantities should scale as  $(\text{const}1 + \text{const}2/L^d)$ . Deviations have been observed, which do not behave in a simple way as a function of  $L$ . Two possible reasons for the discrepancies are mentioned by the authors:  $q=10$  may not be in the large- $q$  limit, for which Eq. (2.33) was derived, or  $L$  may not be sufficiently large to be in the asymptotic regime.

Therefore a two-dimensional  $q=20$  Potts model was simulated in addition (Billoire *et al.*, 1992). The first-order transition there is even stronger than in the  $q=10$  case. Thus the asymptotic behavior in  $L$  should set in earlier than for  $q=10$ , "earlier" means for smaller volumes. (The ratio that counts is  $\xi/L$ .) Here the agreement with the theoretical predictions is reasonable.

Let us summarize so far. Computer simulations are always performed in a finite volume, whereas the order of a phase transition is usually formulated in the infinite-volume limit. Although this formulation is very convenient, it is not the only appropriate way to proceed. The good news comes from finite-size scaling analysis. From the information obtained in a finite volume one can predict whether the transition is going to be of first or second order in the thermodynamic limit. (Likewise a small mass can play the role of a scaling field. A finite-mass scaling analysis is suited for an extrapolation to the zero-mass limit.) From a practical point of view, a finite-size scaling analysis may be less useful, since it is derived for finite but large volumes. The original hope was to disentangle first- and second-order signatures even for moderate lattice sizes by using refined criteria. The results for the two-dimensional  $q=10$  Potts model were somewhat discouraging. One has to have such large volumes to verify the predictions of a finite-size scaling analysis that the "naive" criteria are equally applicable for inferring the order of the transition. This statement refers to weakly first-order transitions. Otherwise the asymptotic region of a large volumes would be easily realized as  $\xi/L \ll 1$  for moderate values of  $L$  when  $\xi$  was sufficiently small.

Typical volumes in Monte Carlo calculations for QCD are marginal in large finite size. For QCD transitions it seems difficult ever to get into the asymptotic regime of large volumes. Right from the beginning one is restricted to much smaller volumes when simulating full QCD rather than a spin model. The final efficiency of refined criteria is, however, a question of size, which has to be answered in the concrete model of interest. The

value of the correlation length in a first-order transition depends on details of the dynamics. For QCD the largest correlation length is likely not to be small compared to the typical lattice size.

This concludes our overview of the concepts and results of statistical mechanics. As we turn now (and for the remainder of this review) to applications in QCD, we give a dictionary of correspondences between thermodynamic quantities in statistical physics and their counterparts in QCD in (Table II).

The second and third columns refer to transitions in liquid/gas or ferromagnetic systems. Two analogies may be seen between the magnet and the fluid systems. The first is between the sets  $(T, p, V)$  of a fluid and  $(T, H, -M)$  in a magnet. The second is between  $(T, \varrho, \mu)$  in a fluid and  $(T, M, H)$  in a magnet (Stanley, 1971). The external field  $H$  is the thermodynamic variable conjugate to the order parameter  $M$ . Likewise the chemical potential is conjugate to the density  $\varrho$ , where  $\varrho$  is the order parameter for a fluid/gas transition. The replacement of  $V$  by  $(-M)$  and  $p$  by  $H$  transforms almost all equations for a fluid/gas system into the corresponding equations for a magnet.

In the last two rows various response functions are listed, describing the response of the system to a stimulus in the temperature or in an external field. They are second derivatives of the thermodynamic potential  $\Omega$ , while the order parameter is obtained as a first derivative of  $\Omega$  with respect to the conjugate field. The last column shows the associated critical exponents characterizing the singular behavior of thermodynamic functions in a second-order transition.

The QCD transitions refer to limiting cases of vanishing quark masses (chiral symmetry) or infinite quark masses [ $Z(3)$  symmetry]. A further column is devoted to the chiral transition described on an effective level (here in the linear sigma model), where the quark and gluonic substructures of mesons are disregarded. If one is interested in the evolution of a QCD-plasma "fluid,"  $(T, p, V)$  is an appropriate set of thermodynamic variables entering the equation of state.

Formally the current quark masses  $m_q$  will be shown to play the same role as an external magnetic field  $H$  in a ferromagnet. On an effective level this analogy is more manifest, as nonvanishing quark masses will be described by external fields  $\varepsilon_0, \varepsilon_8$ ; see Sec. IV.A.4. One may make use of this analogy to guess, from results in the massless or pure gauge limits, the effects of finite quark masses on the order of the QCD transitions.

The most popular order parameters in QCD are the quark condensate  $\langle \bar{q}q \rangle$  [or  $\langle \bar{q}q \rangle$  and  $\langle \bar{s}s \rangle$ ] or the mesonic condensates  $\langle \sigma_0 \rangle$ ,  $\langle \sigma_8 \rangle$  on the mesonic level for the chiral transition, and the expectation value of the Wilson line  $\langle L \rangle$  for the deconfinement transition. The ability to predict the sigma-meson and pion-meson mass from a critical equation of state (Sec. IV.A.1) sounds less surprising if one keeps this table of analogies in mind.

The indices  $b$  and  $s$  at  $T_c$  refer to the approach of  $T_c$  from the symmetric or the broken phase. For the

TABLE II. Correspondences between statistical physics and QCD.

	Liquid gas	Ferromagnet	QCD chiral symmetry	QCD Z(3) symmetry	Linear sigma model	Critical exponent
Thermodynamic set of variables	$T$ $P$ $V$	$T$ $H$ $-M$	$T$ $P$ $V$ or chirally broken to chiral symmetry	$T$ $P$ $V$ or confinement to deconfinement	$T$ $P$ $V$ or chirally broken to chiral symmetry	$\varepsilon_0, \varepsilon_8$ $\langle \sigma_0 \rangle, \langle \sigma_8 \rangle$
Phase transition from... to	liquid to gas	broken to symmetric phase	chirally broken to chiral symmetry	confinement to deconfinement	chirally broken to chiral symmetry	$\propto (T - T_c)^\beta$ $T \rightarrow T_c^b$
Order parameter	$\rho$	$M$	$\langle \bar{q}q \rangle$	$\langle L \rangle$	$\langle \sigma_0 \rangle, \langle \sigma_8 \rangle$	
Its conjugate variable	$\mu$	$H$	$m_u, m_d, m_s, \dots$		$\varepsilon_0, \varepsilon_8$	
Specific heat	$c_{V,P} = T \left( \frac{\partial S}{\partial T} \right)_{V,P}$	$c_{M,H} = T \left( \frac{\partial S}{\partial T} \right)_{M,H}$	Response to a stimulus in $T$ $c_{V,P} = -T(\partial^2 F / \partial T^2)$			$\propto (T - T_c)^{-\alpha}$ $T \rightarrow T_c$
Correlation length $\xi$	$-\ln G(r) \rightarrow  r /\xi$ $r \rightarrow \infty$	$-\ln \Gamma(\bar{r}) \rightarrow  \bar{r} /\xi$ $r \rightarrow \infty$	$-\ln \langle \bar{q}(x)q(x) \bar{q}(0)q(0) \rangle \rightarrow  x /\xi$ $ x  \rightarrow \infty$	$T = \xi$	$-\ln \langle \sigma_0(x) \sigma_0(0) \rangle \rightarrow  x /\xi$ $ x  \rightarrow \infty$	$\propto (T - T_c)^{-\nu}$ $T \rightarrow T_c^s$
Critical pair correlation $\Gamma(\bar{r})$	$\langle n(\bar{r})n(\bar{r}') \rangle$ $n(\bar{r}) \equiv \sum_{i=1}^N \delta(\bar{r} - \bar{r}_i)$	$\langle s(\bar{r})s(\bar{r}') \rangle$ $s(\bar{r}) = \text{spin at } \bar{r}$	$\Gamma(\bar{r}) \propto  \bar{r} ^{-(d-2+\eta)}$ $\langle \bar{q}(x)q(x) \bar{q}(0)q(0) \rangle \propto \exp(-\sigma x /T)$	$\langle L(0)L^+(\bar{x}) \rangle$	$\langle \sigma_0(\bar{x}) \sigma_0(0) \rangle$	$ \bar{r} ^{-(d-2+\eta)}$ $T = T_c$
Compressibility, susceptibility, "magnetization", i.e., $\rho, M, \langle \bar{q}q \rangle, \dots$ respectively	$\kappa_{T,S} = -\frac{1}{V} \left( \frac{\partial V}{\partial P} \right)_{T,S}$ $\rho$	$\chi_{T,S} = \frac{1}{V} \left( \frac{\partial M}{\partial H} \right)_{T,S}$ $M$	Response to a stimulus in the "external field" $\chi_{T,S} = \frac{\partial \langle \bar{q}q \rangle}{\partial m_q}$	$\chi = \Sigma \Gamma(\bar{x})$	$\frac{\partial \langle \sigma_0 \rangle}{\partial \varepsilon_{0,8}}$ $\langle \sigma_0 \rangle, \langle \sigma_8 \rangle$	$\propto (T - T_c)^{-\gamma}$ $T \rightarrow T_c^c$ $\propto \mu^{1/\delta}, H^{1/\delta}, m_q^{1/\delta}, \dots$ $T = T_c$

deconfinement transition, the phase of broken  $Z(3)$  symmetry is realized *above*  $T_c$ ; thus the commonly used ( $\pm$ ) signs to indicate the approach from above or below  $T_c$  would be misleading in this case. Note that the singular behavior of the ‘‘magnetization,’’ i.e., the order parameter, is characterized by a thermal exponent  $\beta$  as  $T \rightarrow T_c^b$  and by a ‘‘magnetic’’ exponent  $1/\delta$  as ‘‘H’’  $\rightarrow 0$  at  $T_c$ .

## B. Phase transitions in QCD

In a continuum notation the QCD Lagrangian is given as

$$L = L_g + L_m$$

where

$$L_g = -\frac{1}{4} F_{\mu\nu}^a(x) F^{a\mu\nu}(x)$$

$$L_m = \bar{\psi}(x)(i\gamma^\mu D_\mu - M)\psi(x)$$

and

$$F_{\mu\nu}^a \equiv \partial_\mu A_\nu^a(x) - \partial_\nu A_\mu^a(x) + gf^{abc} A_\mu^b(x) A_\nu^c(x)$$

$$D_\mu \equiv \partial_\mu - i \cdot g \frac{\lambda_a}{2} A_\mu^a(x). \quad (2.46)$$

Here  $A_\mu^a$  are the gauge fields,  $\psi_{\alpha,f,c}$ ,  $\bar{\psi}_{\alpha,f,c}$  denote the quark fields where  $\alpha$  is a Dirac index,  $f=1, \dots, N_f$  labels the flavors, and  $c=1, \dots, N_c$  labels the colors. The gauge coupling constant is denoted as  $g$ , the structure constants as  $f^{abc}$ , and  $\lambda_a$  are the generators of the fundamental representation of  $SU(N_f)$ ,  $a=1, \dots, 8$  for  $N_f=3$ , where  $\lambda_a$  are the Gell-Mann matrices. The quark mass matrix is denoted as  $M$ .

As a first step in the investigation of the phase structure one may consider certain limiting cases of Eq. (2.46). One limit is the pure gauge theory, where  $L = L_g$ . This is the limit of infinitely heavy quark masses and will be discussed in Sec. II.B.2. The topic of the following section is the limit of massless quarks, which is called the chiral limit. In the chiral limit the QCD Lagrangian is invariant under global  $U(1)_V \times SU(N_f)_L \times U(1)_A \times SU(N_f)_R$  transformations. The  $U(1)_V$  invariance corresponds to the baryon number conservation. The invariance under axial  $U(1)_A$  transformations is only classically preserved. On the quantum level it is broken via the axial anomaly. Even in the presence of an anomaly, axial  $Z_A(N_f)$  symmetry remains (Callan *et al.*, 1976). At zero temperature the invariance under  $Z_A(N_f) \times SU(N_f)_L \times SU(N_f)_R$  chiral transformations is assumed to be spontaneously broken by the QCD vacuum to the vector  $SU(N_f)_V$  symmetry. For  $N_f=2$  this isospin symmetry is realized in the hadronic spectrum to a very good approximation. For  $N_f=3$  the realization of  $SU(3)_V$  symmetry is a little more questionable. Nevertheless it is also frequently considered as an approximate symmetry of QCD.

TABLE III. Order of the chiral transition as a function of the number of flavors ( $N_f$ ) and the anomaly strength ( $g$ ), partly conjectural.

$N_f$	$g=0$ $N_c=\infty$	$g=\text{const}$ of order 1	$g=g(T)$ $g(T) \sim d_I(T)$
1	Second order $O(2)$ exponents	no transition	no transition
2	First order	Second order $O(4)$ exponents	First order $O(2) \times O(4)$ symmetry
3	First order	First order	First order $U_A(1) \times SU(3) \times SU(3)$ symmetry
$\geq 4$	First order	First order	First order

### 1. Renormalization-group analysis in the chiral limit

Rather than directly studying the QCD Lagrangian (2.46) one can analyze chiral symmetry breaking in an effective Lagrangian, which shares the chiral symmetry properties of QCD. The Lagrangian is formulated in terms of a self-interacting  $N_f \times N_f$  matrix field  $\phi$ . Chiral symmetry breaking is parametrized in terms of  $\phi_{ij}$  according to  $\phi_{ij} \sim \langle \bar{q}_i(1 + \gamma_5)q_j \rangle$ , where  $\phi$  transforms under transformations of  $G_f \equiv U_A(1) \times SU(N_f)_L \times SU(N_f)_R$  according to

$$\phi \rightarrow \phi' = \exp(i\alpha)U_+ \phi U_- . \quad (2.47)$$

Here  $U_+$  and  $U_-$  are arbitrary and independent  $SU(N_f)$  matrices, while  $\alpha$  generates a  $U_A(1)$  transformation. The most general renormalizable Lagrangian in terms of  $\phi$  that is consistent with the chiral symmetry properties of QCD is given as (Pisarski and Wilczek, 1984)

$$L = \frac{1}{2} \text{Tr}(\partial_\mu \phi^+) (\partial_\mu \phi) - \frac{m^2}{2} \text{Tr}(\phi^+ \phi) - \frac{\pi^2}{3} f_1 (\text{Tr} \phi^+ \phi)^2 - \frac{\pi^2}{3} f_2 \text{Tr}(\phi^+ \phi)^2 + g(\det \phi + \det \phi^+). \quad (2.48)$$

As a necessary condition for stability at large values of  $\phi$ ,  $f_2$  and  $(f_1 + f_2/N_f)$  have to be larger than zero. The determinant term accounts for the anomaly. It vanishes in the pure gauge case ( $N_f=0$ ) and in the limit of infinite colors  $N_c=\infty$  (Witten, 1979). At zero temperature the vacuum expectation value  $\langle \phi \rangle$  is  $SU(N_f)$  symmetric and different from zero,  $m^2 < 0$ . Spontaneous symmetry breaking is associated with an  $SU(N_f)$  massless multiplet of Goldstone bosons and a massive flavor singlet  $\eta'$ .

In Table III we summarize the results of a renormalization-group analysis by Pisarski and Wilczek (1984) of the order of chiral phase transitions. These transitions are driven by chiral symmetry restoration as the temperature is raised.

For each number of flavors three cases are distinguished: a vanishing anomaly corresponding to a vanish-

ing number of flavors or an infinite number of colors,  $g = \text{const}$  of order 1, where “const” refers to the assumed temperature independence of  $g$ , and  $g = g(T)$ . Here  $g(T)$  is taken to be approximately equal to the instanton density  $d_I$ , which is supposed to vanish at high temperatures. In the following we summarize how the conjectures of Table III arise in a renormalization group analysis.

In the previous section we mentioned the real-space renormalization-group approach as a tool for describing critical phenomena. An alternative approach to the construction of renormalization-group equations, which is utilized here, is the  $\varepsilon$  expansion performed in momentum space. Starting from a Hamiltonian in terms of an order-parameter field  $\phi$ , one integrates out large momenta (large momenta in contrast to short length scales in the real-space approach). The new Hamiltonian in terms of the new field variables is arranged to have the same form as the old one. The new couplings should be understood as renormalized couplings. They are obtained in a perturbative expansion. The coefficients of the expansion depend analytically on the dimension  $d$ . For our applications the appropriate small expansion parameter is  $\varepsilon = d - 4$ .

Normally a change in the scale of momenta induces intricate changes in the action and in derived quantities like correlation functions, with certain exceptions which we know from the previous section. Simple scaling behavior is recovered if the set of couplings reaches a value such that any further change in the scale of momenta does not affect them. This is the point in coupling parameter space where the Hamiltonian approaches a fixed-point Hamiltonian.

In the present framework fixed points occur as zeros of the  $\beta$  functions. In the case of one coupling, the  $\beta$  function gives the change in the renormalized coupling under a change in momentum scale. Equivalently the  $\beta$  function gives the change of the dimensionless renormalized coupling  $u$  under a change of the dimensionless bare coupling  $u_0$  according to

$$\beta(u) = -\varepsilon \left( \frac{\partial \ln u_0}{\partial u} \right)^{-1} = \sum_n a_n(\varepsilon) u^n, \quad (2.49)$$

where the coefficients  $a_n$  depend on  $\varepsilon$ . Given a set of bare couplings in parameter space at a certain momentum scale, the question arises whether, and under what conditions, the renormalized couplings  $v$  flow into the fixed point  $v^*$  (and hence lead to a second-order phase transition). The answer is given as a condition for the occurrence of an infrared-stable fixed point  $v^*$ . The stability criterion is that the matrix

$$\omega_{ij} \equiv \frac{\partial \beta_i}{\partial v_j} \quad (2.50)$$

have real and positive eigenvalues at  $v = v^*$ . Here  $v$  and  $v^*$  stand for sets of couplings,  $\beta_i$  are the associated  $\beta$  functions.

The  $\beta$  functions for a chiral  $SU(N_f) \times SU(N_f)$  linear sigma model in the absence of an anomaly (case  $g = 0$  in

Table III) were derived within an  $\varepsilon$  expansion (Pater-son, 1981). When the two functions  $\beta_1$  and  $\beta_2$  for  $f_1$  and  $f_2$  are inserted in Eq. (2.50), the stability criterion may be applied to search for IR-stable fixed points. The results are summarized in the first column of Table III. For  $0 \leq N_f < \sqrt{2}$ , the IR-stable fixed point has  $f_2^* = 0$  with  $O(2N_f)$  critical exponents. No infrared-stable fixed point occurs for  $N_f > \sqrt{3}$ , if  $f_1, f_2$  are of order  $\varepsilon$ .

For some time it has been taken for granted that the absence of an infrared-stable fixed point implies a first-order phase transition (Bak *et al.*, 1976). The absence or existence of an IR-stable fixed point is, however, less conclusive than was originally supposed. The existence of an IR-stable fixed point does not exclude a first order-phase transition (there may be a region in coupling parameter space that is not attracted by the fixed point). Conversely, the absence of an IR-stable fixed point does not exclude a second-order phase transition. Instead of the fixed-point criterion, precise conditions for the occurrence of a first-order phase transition have been specified by Iacobson and Amit (1981). The framework is again the renormalization-group approach, using the  $\varepsilon$  expansion. The prediction of first-order phase transitions applies to all multicomponent  $\phi^4$  theories with more than one dimensionless coupling constant. The interaction term of a multicomponent  $\phi^4$  theory can be written as

$$\propto g_{ijkl} \phi_i \phi_j \phi_k \phi_l, \quad (2.51)$$

where  $i, j, k, l \in \{1, \dots, N\}$ . The quartic terms of the linear sigma model, Eq. (2.48), can be recast in this form. Explicit expressions can be found, for example, Paterson (1981). As long as the quartic terms proportional to  $f_1$  and  $f_2$  in Eq. (2.48) are independent of each other (as is the case for  $N_f = 3$ ), the criteria of Iacobson and Amit are satisfied by the linear sigma model, and a first-order chiral transition is predicted for  $N_f \geq 2$ .

Note that these predictions are neither conjectures nor rigorous statements. They are not only based on the absence of an IR-stable fixed point, but derived within the perturbative framework with  $\varepsilon \ll 1$ . Nonperturbative features are not attainable in this approach. One may try to extrapolate the results from four- to three-dimensional models by setting  $\varepsilon = 1$  (although  $\varepsilon \ll 1$  has been assumed in deriving the  $\varepsilon$  expansion). Examples are known (Bak *et al.*, 1976) in which the results to leading order in  $\varepsilon$  remain a good guide for  $\varepsilon = 1$ .

The second column of Table III subsumes the suggestions for the order of the chiral transition in the presence of a temperature-independent strength of the anomaly, and  $g$  is assumed to be of the order of the other couplings  $f_1$  and  $f_2$ . For two flavors the det term in Eq. (2.48) acts like a mass term. Thus it may change the order of the transition depending on the magnitude of  $g$ . For three flavors, the det-term is trilinear in the matrix elements of  $\phi$ . A cubic term on the classical level is usually regarded as sufficient for inducing a first order phase transition (although we know from the discussion in Sec. II.A that a weakly first-order transition, predicted on the classical level, may be wiped out by fluc-

tuations; hence the classical cubic term makes the first order transition likely, but is no guarantee of its occurrence).

For  $N_f=4$  Pisarski and Wilczek (1984) argue again with the absence of an IR-stable fixed point in favor of a first-order transition. (The  $g \neq 0$  case is not covered by the analysis of Iacobson and Amit.) As claimed by Paterson (1981), the fixed-point structure of the  $SU(4) \times SU(4)$  linear sigma model is unchanged, when a term  $\propto g(\det \phi + \det \phi^\dagger)$  is included. For  $N > 4$  the det term is an irrelevant operator and should not change the critical behavior.

The effective symmetry of the linear sigma model, Eq. (2.48) may change as a function of temperature, if the anomaly strength is temperature dependent and determined by the density of instantons  $d_I(T)$ . Since  $d_I(T) \rightarrow 0$  as  $T \rightarrow \infty$  (Gross *et al.*, 1981),  $g$  could be small at the transition temperature compared to the  $T=0$  value. Thus the axial  $U_A(1)$  symmetry would be partially restored. Predictions accounting for this partial symmetry restoration as a result of decreasing  $g$  are listed in the third column of Table III. For further details on the consequences of an approximate  $U_A(1)$  restoration, we refer the reader to the original references.

The effective Lagrangian (2.48) can be extended to include nonzero bare (meson) masses. The simplest ansatz has the form  $(\text{tr } M \phi)$  and is linear in the mass matrix  $M$ . Formally it acts as a background magnetic field. The formal analogy is evident if we recall that the linear sigma model can be rewritten as a multicomponent  $\phi^4$  theory with an additional (multicomponent)  $\phi^3$  term (the det term) and a symmetry breaking mass term, that is linear in  $\phi$ . In this way it takes the form of the Landau free-energy functional Eq. (2.2) for a multicomponent order-parameter field  $\phi$  in an external field.

For later comparison we note that the magnetic-field term is proportional to the mass and vanishes in the chiral limit. In the other extreme case of infinitely heavy quark masses, the effective magnetic field of pure gauge theories will be shown to be proportional to  $(e^{-m})$ , thus vanishing in the infinite-mass limit ( $m \rightarrow \infty$ ). We shall come back to this point at the end of the next section.

To summarize so far, conjectures about the order of the chiral transition as function of  $N_f$  are based on a perturbative renormalization-group analysis in momentum space. The analysis has been performed for an effective  $SU(N_f) \times SU(N_f)$  linear sigma model sharing the chiral symmetry properties of QCD. None of the conjectures is rigorously proven. Some of them ( $g=0$ ) are verified within an  $\varepsilon$  expansion, others ( $g \neq 0$ ) are based on the absence of an infrared-stable fixed point when a first-order transition is predicted, a criterion that should be used with care.

From statistical physics it is known that a first-order transition remains first order when a background field is introduced and the field is sufficiently weak. Otherwise the transition may be washed out completely. The relative size of the latent heat compared to the strength of the external field, i.e. the values of the quark masses, decides whether the  $N_f=3$  chiral transition is preserved

under realistic QCD conditions or not. This question cannot be answered within a renormalization-group analysis, but only by detailed calculations. The renormalization-group approach is nevertheless a good place to start.

## 2. The limit of a pure $SU(N_c)$ gauge theory

In this section we deal with the *quenched limit* of QCD. The quenched limit is obtained as the number of flavors goes to zero or the quark masses are sent to infinity. The gluonic vacuum in a background of infinitely heavy quarks may be probed by test quarks. In perturbation theory this means that virtual quark loops are suppressed.

In the quenched limit the theory has an extra global symmetry, which results from the periodicity of the gauge fields in the temperature direction. The partition function may be represented as a functional integral over gauge fields, which are periodic in Euclidean time with period  $\beta=1/T$ . The periodicity condition arises as a consequence of the trace in the definition of the thermodynamical potential

$$Z = \int_{A_\mu(\beta, \mathbf{x}) = A_\mu(0, \mathbf{x})} \mathcal{D}A \times \exp \left\{ -\frac{1}{g^2} \int_0^\beta dt \int d^3x \frac{1}{4} \text{Tr} F_{\mu\nu}^2 \right\}. \quad (2.52)$$

(We still use a continuum notation for the path integral over all gauge fields  $A_\mu$  with Yang-Mills field strength  $F_{\mu\nu}$ .) Gauge transformations that are compatible with the periodicity condition need only be periodic up to an element  $c_N$  of the center  $Z(N)$  of the gauge group  $SU(N)$ . Thus a gauge transformation must obey

$$V(\mathbf{x}, 0) = c_N V(\mathbf{x}, \beta) \quad \text{for all } \mathbf{x} \quad (2.53)$$

with  $c_N \in Z(N)$ . The  $n$ th element  $c_N^n$  of  $Z(N)$  is given as  $\exp(2\pi i n/N)$ . It is easily checked that the generalized gauge transformation (2.53) leaves all topologically trivial Wilson loops invariant (in particular, those appearing in the action) and thus the action itself.

This means that the extra symmetry is generated by the action of local gauge transformations, which are periodic up to an arbitrary element of the center *modulo* strictly periodic local gauge transformations. The essential ingredient, which ensures the invariance of topologically trivial loops under the additional symmetry, is the property of center elements to commute with all elements of the group.

The issue of the transition from a confinement phase at low temperatures to the deconfinement phase at high temperatures can be related to the issue of whether the pure glue vacuum of QCD is  $Z(N)$  invariant like the action. As it turns out, the transition from the confinement to the deconfinement phase may be explained as a spontaneous breaking of the extra  $Z(N)$  symmetry at finite temperature.

Qualitatively the quark gluon plasma can be probed by a heavy test quark. The free energy of this test quark

should be infinite in the confinement phase, but finite in the deconfinement phase. It may be computed as the expectation value of the Wilson line also called a Polyakov loop or a thermal Wilson loop. The Wilson line  $L$  is defined as the spatially local operator

$$L(\mathbf{x}) \equiv \text{Tr} P \exp \left( \int_0^\beta dt A_0(t, \mathbf{x}) \right), \quad (2.54)$$

where  $P$  stands for the path-ordered product. The Wilson line is a topologically nontrivial loop. It is closed due to the periodic boundary conditions and transforms nontrivially under the center transformations

$$L(\mathbf{x}) \rightarrow c_N L(\mathbf{x}). \quad (2.55)$$

Thus  $L$  can have a vanishing expectation value. Conversely, a nonvanishing expectation value can be taken as the signal of a spontaneous breakdown of the global  $Z(N)$  symmetry of the action. We have

$$\langle L(\mathbf{x}) \rangle = e^{-\beta F(\mathbf{x})} = \begin{cases} 0 & \text{confinement} \\ \text{finite} & \text{deconfinement.} \end{cases} \quad (2.56)$$

Here  $F(\mathbf{x})$  denotes the free energy of an isolated test quark. This behavior qualifies the Wilson line expectation value as an order parameter for the confinement/deconfinement phases of QCD in the absence of dynamical quarks.

It is natural to look for an effective action of the  $SU(N)$  gauge theory in terms of the order parameter field, which means in terms of  $L$ . Such an effective action could simplify the investigation of the phase structure of pure QCD, if universality arguments may be used. Svetitsky and Yaffe (1982a, 1982b) have given plausible arguments that a suitable candidate for such an effective action is a  $Z(N)$  spin theory.

The result of an integration over spatial gauge fields is an effective action in terms of Wilson lines. The Wilson lines [originally given in terms of elements of the  $SU(N)$  Lie algebra], are represented by  $Z(N)$  variables. The action is argued to be short ranged. The high-temperature behavior of the  $(3+1)$ -dimensional theory is determined by the dynamics of the three-dimensional  $SU(N)$  gauge theory. Three-dimensional  $SU(N)$  gauge theories show an area law for Wilson loops (Feynman, 1981) with short-range correlations. The range of interactions is supposed to stay finite over the entire temperature range.

The effective action is invariant under  $Z(N)$  transformations. The  $Z(N)$  symmetry can be spontaneously broken at high temperatures, but is restored at low temperatures and ensures confinement.

The path-integral representation of the original  $SU(N)$  gauge theory is then replaced by

$$e^{-\beta V \Omega} = \sum_{\{s_i\}} \exp \left( - \frac{1}{g^2(T)} H[\{s_i\}] \right). \quad (2.57)$$

Here  $s_i \in Z(N)$ , the sum extends over all  $Z(N)$  spin configurations, and  $\Omega$  is the thermodynamic potential.

#### a. Effective vs physical temperatures

In the original action of Eq. (2.52)  $g$  parametrizes the interaction strength. When the time dependence of the

field is dropped, the  $\int_0^\beta dt$  can be performed to yield a prefactor of the remaining three-dimensional action  $(-\beta/g^2)$ . The  $\beta$  factor can be absorbed by a rescaling of the fields such that the action in three dimensions takes the same form as the original action in four dimensions. Now the prefactor plays the role of an inverse temperature of the  $3d$  model. Thus  $g^2$  is identified with  $T_{\text{eff}}$ , the effective temperature of a classical spin system described by the partition function of Eq. (2.57). For  $T_{\text{eff}}$  small, the “spin system” is in the ordered phase, and we have  $\langle L \rangle \neq 0$  corresponding to deconfinement. Deconfinement is realized at high physical temperatures  $T$ , for which  $g$  is small due to asymptotic freedom. A small value of  $g$  means a small value of  $T_{\text{eff}}$ , which is consistent with the initial assumption. Thus a low effective temperature  $T_{\text{eff}}$  corresponds to a high physical temperature  $T$  and vice versa. This explains why the order parameter vanishes in the low- $T$  phase and signals “order” in the high- $T$  phase.

The coupling  $g$  becomes manifestly temperature dependent when high frequency contributions are integrated out in passing from the  $(3+1)$ -dimensional to the 3-dimensional theory.

#### b. The phase structure of $Z(N)$ spin models

The phase structure of  $Z(N)$  spin models has been studied in statistical physics for various values of  $N$  and space dimension  $d-1$ . If a second-order transition is predicted in the spin model, one could attempt to locate the renormalization-group fixed point and consider the simpler spin model as a fixed point theory in the universality class of the  $SU(N)$  theory.

A table for various  $Z(N)$ ,  $SU(N)$ , and  $U(1)$  models in dimensions  $d=2, 3$ , and  $\geq 4$  can be found in Svetitsky and Yaffe, (1982b). Here we mention the cases of two and three colors.

The spin model, which is associated with a  $(3+1)$ -dimensional  $SU(2)$  gauge theory at high temperatures, is the Ising model in three dimensions. The 3d Ising model is known to have a second-order phase transition.

The case of  $N_c=3$  in  $d=3$  is particular. In the space of three-dimensional  $Z(3)$ -symmetric theories no infrared-stable renormalization-group fixed point is known. A specific realization of a  $Z(3)$ -symmetric spin theory is the three-state Potts model (see Sec. II.A), which is known to have a first-order transition. The famous cubic term on the classical level, driving the transition to first order, is allowed by the  $Z(3)$  symmetry. The potential of a  $Z(3)$ -symmetric theory of a single complex scalar field  $L(\mathbf{x})$  may be written as the sum of a  $U(1)$ -symmetric term depending on  $|L|^2$  and a term depending on  $\text{Re } L^3$ . In  $d=3$ , a  $U(1)$  gauge theory has a second-order transition (see for example, Pfeuty and Toulouse, 1977). A term  $\text{Re } [L(\mathbf{x})^3]$  explicitly breaks  $U(1)$  symmetry down to  $Z(3)$ . In the renormalization-group sense it is relevant enough to affect the critical behavior of the  $U(1)$  theory.

Thus the conjecture of Svetitsky and Yaffe is that in the absence of an IR-stable fixed point the original (3+1)-dimensional SU(3) gauge theory has a first-order finite-temperature phase transition as well. The prediction was confirmed in early Monte Carlo calculations (e.g., Celik *et al.*, 1983a, 1983b), then cast into doubt (Bacilieri *et al.*, 1988, see also the introduction and Sec. III.B below), and has been reestablished in more recent calculations by Fukugita *et al.* (1989). Before we report on this controversy in Sec. III.B.1, let us consider the influence of dynamical quarks on the deconfinement transition.

### c. Inclusion of dynamical quarks

From spin systems we know that a first-order transition stays first order if the magnetic field is sufficiently small. The gaps in thermodynamic quantities are only continuously deformed for a perturbatively small external field. The magnetization, however, ceases to be a good order parameter, as it is always finite due to the presence of an external background field.

Similarly the Wilson line may be expected to fail as an order parameter when dynamical quarks are included. The physical reason is easily understood. The free energy of an isolated test quark no longer diverges ( $\langle L \rangle \neq 0$  for all temperatures). When the flux tube between two test quarks is sufficiently stretched, a  $q\bar{q}$  pair is popped out of the vacuum. The test quarks may always form finite-energy bound states with dynamical quarks.

Since the deconfinement transition was supposed to be driven by the spontaneous symmetry breaking of global  $Z(3)$  symmetry, we consider the invariance of the QCD action under  $Z(3)$  transformations in the presence of dynamical quarks. The generalized gauge transformations, Eq. (2.53), are periodic only up to an element of the center of the gauge group. While the periodic boundary conditions on the gauge field are preserved under the generalized transformations, the anti-periodic boundary conditions on the quark fields are not. Imposing

$$\psi(\mathbf{x}, \beta) = -\psi(\mathbf{x}, 0) \quad (2.58)$$

on the quark fields as before, we find that Eq. (2.58) transforms under a generalized gauge transformation with  $U(\mathbf{x}, \beta) = c_N^i U(\mathbf{x}, 0)$ , where  $i$  labels the center elements, according to

$$\begin{aligned} \psi(\mathbf{x}, \beta) \rightarrow U(\mathbf{x}, \beta) \psi(\mathbf{x}, \beta) &= c_N^i U(\mathbf{x}, 0) \psi(\mathbf{x}, \beta) \\ &= -c_N^i U(\mathbf{x}, 0) \psi(\mathbf{x}, 0). \end{aligned} \quad (2.59)$$

Equation (2.59) shows that the gauge transformed fields  $U \cdot \psi$  (under the generalized gauge transformation) no longer satisfy antiperiodic boundary conditions. If we compare the path integral over all gauge-field configurations in the absence and presence of dynamical quarks satisfying Eq. (2.58), configurations differing by generalized gauge transformations get the same Boltzmann weight in the absence of quarks, but a different weight in their presence. Hence the effect of dynamical quarks is

an explicit symmetry breaking of the  $Z(N)$ -symmetry. Its strength is determined by the values of the quark masses.

### d. Finite quark masses and external fields

The analogy between an external field in a spin system and dynamical quarks in QCD becomes manifest when the quarks are integrated out. Integration over quark degrees of freedom induces an external field on the effective level of spin models (Banks and Ukawa, 1983; DeGrand and DeTar, 1983). We refer the reader to the work of DeGrand and DeTar (1983) and anticipate results from lattice gauge theory.

Consider a four-dimensional SU(3) lattice gauge theory with fermions at high temperatures. The original action consists of a gauge field part  $S_g$  and a fermionic part  $S_F$ . The fermionic part may be written as  $S_F(U) = \sum \bar{\psi} M(U) \psi$  [dropping all indices, where  $U$  is the SU(3) gauge field and  $M$  the fermion matrix, explicit expressions will be given later]. When the Grassmann variables are integrated out, the effective action takes the form

$$S_{\text{eff}}[U] = S_g[U] + \text{Tr} \ln M[U]. \quad (2.60)$$

In the high-temperature and strong-coupling limit of the SU(3) gauge theory it can be shown that the gauge part  $S_g$  transforms into the action of a 3d three-state Potts model. (For the Potts model, see Sec. II.A.) The fermionic part of Eq. (2.60) simplifies, when it is treated in a hopping-parameter expansion. The hopping parameter  $\kappa$  may be related to the bare quark mass  $m$  of the original action according to

$$\kappa \sim \frac{1}{2} e^{-m\tau}, \quad (2.61)$$

where  $m \cdot \tau$  is the bare quark mass in lattice units and  $\tau$  is the lattice spacing in the time direction. [Equation (2.61) holds only at strong coupling and for small values of  $\kappa$ .] The result for the fermionic term together with the simplified gauge part leads to the following effective action (for zero chemical potential):

$$S_{\text{eff}} = - \left[ \frac{\beta a}{\tau} \sum_{n, \mu} \text{Re} z_n^* z_{n+\hat{\mu}} + h(\kappa) \sum_n \text{Re} z_n \right]. \quad (2.62)$$

Here  $a$  is the lattice spacing in spatial directions. The sums go over all sites  $n$  and nearest-neighbor pairs  $(n, n+\hat{\mu})$ ,  $\hat{\mu}$  is the unit vector in the  $\mu$ th direction, and  $z_n$  are elements of  $Z(3)$ . The coupling  $h$  is the external field, which is a remnant of the dynamical fermions. The  $\kappa$  dependence is explicitly known and approximately given as

$$h(\kappa) \sim 24\kappa. \quad (2.63)$$

With the help of Eq. (2.61) we see that  $h$  vanishes with  $\kappa$  for infinitely heavy quark masses, in the limit of a pure gauge theory. The strength of the external field grows with decreasing quark mass. Hence sufficiently light quark masses will completely wash out the first order

transition in the  $3d$  three-state Potts model. DeGrand and DeTar (1983) obtain for the critical strength of the magnetic field  $h_{cr} = 2/3 (\ln 2 - 2/3)$ . This result was obtained in a mean-field analysis. It holds for a cubic lattice in three dimensions. At the critical field (mass), the line of first order transitions terminates at a second-order critical point and disappears for larger fields (smaller masses).

A particular virtue of the lattice “derivation” of the effective action Eq. (2.62) is that it allows a calculation of the effective coupling in terms of parameters of the original action [here  $h = h(\kappa)$ ]. In this respect the effective Potts model differs from the effective Lagrangian of chiral perturbation theory (see Sec. IV.A.2), where the Lagrangian parameters have to be fixed from an experimental input. The price one has to pay for such a derivation on the lattice are the involved approximations (strong coupling and high-temperature expansions). Strong-coupling results on the lattice must be extrapolated to continuum results over a long distance in coupling parameter space. It remains to be shown that the qualitative predictions of the effective  $Z(N)$  Potts models survive the continuum limit.

In summary of Sec. II.B, the renormalization-group analyses of Pisarski and Wilczek, and of Svetitsky and Yaffe refer to the idealized limits of vanishing or infinite quark masses, respectively. They were performed in the  $SU(N_f) \times SU(N_f)$  linear sigma model and in  $Z(N)$  spin models rather than directly in QCD. Studies in these limiting cases reveal a dependence of the order of QCD transitions on the number of flavors and the number of colors. The case of physical interest is included for three colors and two or three light flavors. The outcome of the renormalization-group analysis is the message that one has to do hard work in the following sense. Model calculations in terms of a scalar field theory with an  $N$ -component order-parameter field would suffice to model QCD, if the transition were also conjectured to be of second order in the case of three light flavors, or, more realistically, two light and one heavier flavor. One would be free to choose as simple a model as possible within the conjectured universality class. Since the deconfinement and the chiral transitions are expected to be of first order (for three colors and three massless flavors; see, for example, Karsch, 1990), the transition depends on details of the dynamics, in this case the dynamics of full QCD.

Compared to the scale of  $T_c$ , two quarks (up, down) are light, three are heavy (charm, bottom, top), but the strange quark happens to be just of the order of  $T_c$ . In the thermodynamics of QCD, the quark masses play the same role as external magnetic fields in temperature-driven transitions of ferromagnets. From this analogy one must expect that the effects of finite masses depend on their actual values. Results in the idealized limit may even change qualitatively under their influence. Thus a challenge for further investigations of the phase structure of QCD is to find out which of the alternatives displayed in Figs. 2(a) and 2(b) is realized. Partly conjectural diagrams are shown in the  $(m, T)$  plane, where

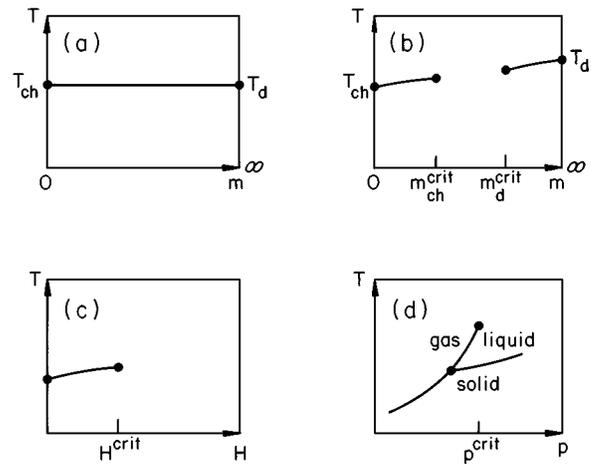


FIG. 2. Hypothetical phase diagrams of QCD in the  $(m, T)$  plane, where  $T$  is the temperature, and  $m$  stands for generic current quark masses: (a) the transitions survive the external fields and coincide, (b) both transitions are washed out for intermediate (realistic?) mass values; (c) an  $(H, T)$  diagram for a ferromagnet; and (d) a  $(p, T)$  diagram for a liquid/gas system. For further explanations see the text.

$m$  stands for a generic current quark mass, and  $N_f = 3 = N_c$  is assumed. For comparison, an  $(H, T)$  diagram for a ferromagnet and a  $(p, T)$  diagram for a liquid/gas system are shown in Figs. 2(c) and 2(d). For  $m = 0$ , a first-order chiral transition is predicted at some temperature  $T_{ch}$ . The chiral transition continues to be of first order, but in a weakened form, as long as the mass is smaller than a critical value (corresponding to  $h_{ch} < h_{cr}$ ). For  $m = \infty$  or  $\kappa = 0$  the first-order deconfinement transition occurs at some temperature  $T_d$ . It will persist as long as  $h(\kappa) \equiv h_d$  is smaller than some  $\tilde{h}_{cr}$  or  $m > m_{cr}$ . Both transitions may “meet” and coincide for intermediate mass values [Fig. 2(a)] or they may not meet [Fig. 2(b)]. In the latter case the discontinuities disappear completely. The dynamical quark masses are then too large for the chiral transition and too small for the deconfinement transition to persist. The chiral symmetric deconfining high-temperature region is then smoothly connected with the chiral symmetry-broken, confined, low-temperature world of daily life. To date lattice results suggest that Fig. 2(b) gives a more realistic description. Lattice calculations are the topic of the next section.

### III. THE LATTICE APPROACH TO THE QCD TRANSITION

#### A. A primer for lattice gauge theory

This section serves mainly to establish the notation of lattice gauge theory. We list the basic definitions, summarize some tools that are relevant to the phase transitions, and point out sources of error for misleading results. For pedagogical introductions we refer the reader to the literature (see, for example, Kogut, 1983; Creutz *et al.*, 1983; or more recently, Rothe, 1992; Montvay and Münster, 1994).

The partition function of an  $SU(N)$  gauge theory interacting with matter fields can be written in the Euclidean path-integral formulation as

$$Z = \int \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} \exp[-S(A_\mu, \bar{\psi}, \psi; g, m_i)], \quad (3.1)$$

where the action depends on the gauge fields  $A_\mu$ , quark fields  $\bar{\psi}$ ,  $\psi$ , the gauge coupling  $g$ , and the quark masses  $m_i$ . The action is given by

$$S = \int_0^\beta dt \int d^3x \mathcal{L}(A_\mu, \bar{\psi}, \psi; g, m_i), \quad (3.2)$$

where  $\mathcal{L}$  is the QCD Lagrangian [Eq. (2.46)] for  $N_f$  flavors and  $N_c$  colors. The chemical potential is set to zero. At finite temperature the gauge and matter fields have to satisfy the following boundary conditions:

$$\begin{aligned} A_\mu(\mathbf{x}, 0) &= A_\mu(\mathbf{x}, \beta) \\ \psi(\mathbf{x}, 0) &= -\psi(\mathbf{x}, \beta) \quad \text{for all } \mathbf{x}, \mu \\ \bar{\psi}(\mathbf{x}, 0) &= -\bar{\psi}(\mathbf{x}, \beta). \end{aligned} \quad (3.3)$$

One usually has periodic boundary conditions in the spatial directions in computer simulations. Note that the finite temperature enters only via the boundary conditions;  $\beta$  denotes the inverse temperature.

In finite-temperature physics, a typical task is to evaluate the thermal expectation value of a physical observable  $O$ . The thermal expectation value, defined as

$$\langle O \rangle = \frac{1}{Z} \text{Tr} e^{-\beta H} O, \quad (3.4)$$

takes the following form in the path-integral formulation:

$$\langle O \rangle = \frac{\int_{pbc} \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} O e^{-S}}{\int_{pbc} \mathcal{D}A_\mu \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{-S}}, \quad (3.5)$$

where  $pbc$  stands for the periodic (antiperiodic) boundary conditions as specified in Eq. (3.3). The functional integral (3.5) is by itself not well defined. Many schemes may be used to perform the functional integration in an approximate way.

One possibility, which is particularly suited to the nonperturbative nature of the phase transition, is the lattice formulation. The (3+1)-dimensional spacetime continuum is discretized commonly on a hypercubic lattice. The lattice provides a gauge-invariant regularization scheme. Lattice results may be extrapolated to the continuum limit in the very end. Lattice artifacts can be controlled via renormalization-group equations and should vanish in the continuum limit. In practice it may be quite subtle and intricate to disentangle lattice artifacts from continuum physics.

A gauge-invariant lattice action is conveniently defined in terms of link variables  $U(x, \mu) \equiv U_x^\mu$ , associated with a link  $(x, \mu)$  leaving site  $x$  in direction  $\mu$ , and matter variables  $\psi(x) \equiv \psi_x$ ,  $\bar{\psi}(x) \equiv \bar{\psi}_x$ , associated with sites on the lattice. We keep the continuum notation for the discrete sites  $(\mathbf{x}, x_0)$ . The link variables are elements of

the gauge group  $SU(N)$  and replace the continuum gauge fields  $A_\mu$ . In the continuum limit these are related as

$$U_x^\mu = \exp\left(iga \sum_{i=1}^8 \lambda^i A_\mu^i(x)\right). \quad (3.6)$$

If the matter fields are fermions, as in the QCD action,  $\psi$  and  $\bar{\psi}$  are Grassmann variables satisfying the usual anticommutation relations

$$\{\psi(x), \psi(y)\} = 0, \quad \{\bar{\psi}(x), \bar{\psi}(y)\} = 0. \quad (3.7)$$

The matter fields carry spin, color, and flavor indices, as indicated in Eq. (2.46). The lattice discretized version of the partition function then takes the form

$$\begin{aligned} Z &= \int_{pbc} \prod_{x, \mu} dU_x^\mu \prod_x d\psi(x) d\bar{\psi}(x) \\ &\quad \times \exp[-S(U_x^\mu, \bar{\psi}_x, \psi_x; g^2, m_i, a_\sigma/a_\tau)]. \end{aligned} \quad (3.8)$$

In contrast to the symbolic notation in Eq. (3.5), the measure  $\mathcal{D}U\mathcal{D}\psi\mathcal{D}\bar{\psi} \equiv \prod dU \prod d\psi d\bar{\psi}$  has a well-defined meaning in Eq. (3.8). The products extend over all link and site variables of the lattice,  $dU_x^\mu$  refers to the (gauge-invariant) Haar measure on the  $SU(N)$  gauge group, and  $d\psi d\bar{\psi}$  is the usual measure over Grassmann variables for each site  $x$ . The action depends on the gauge coupling  $g^2$  and the quark masses  $m_i$  ( $i=1, \dots, N_f$ ) as before, but has an additional dependence on the lattice spacings in spatial ( $a_\sigma$ ) and temporal ( $a_\tau$ ) directions. The lattice spacing, or equivalently the bare coupling, may be chosen independently in the spatial and temporal directions. This is sometimes utilized in finite-temperature calculations of thermodynamical quantities.

### 1. The lattice action

Two requirements have to be satisfied by a lattice action. The first is the reproduction of the classical continuum limit. In the limit of  $a_\sigma = a_\tau \rightarrow 0$ , the action must approach the continuum form of Eq. (3.2). The second is local gauge invariance, i.e., an invariance under local gauge transformations  $V(x)$ , acting according to

$$\begin{aligned} U_x^\mu &\rightarrow (U_x^\mu)' = V(x) U_x^\mu V(x + \hat{\mu}), \\ \psi(x) &\rightarrow \psi'(x) = V(x) \psi(x). \end{aligned} \quad (3.9)$$

(The second requirement is commonly used, but actually not necessary. Actions with non-compact gauge groups recover local gauge invariance only in the continuum limit.) These requirements are not very restrictive. They leave us with a variety of possible lattice formulations. Here we state only the most popular choices, which are sufficient for what follows.

For the gauge part of the action the Wilson form (Wilson, 1974), is the most widely adopted, while for the fermionic part, two choices are *Wilson fermions* (Wilson, 1974) and *staggered or Kogut-Susskind fermions* (Kogut and Susskind, 1975; Susskind, 1977). We distin-

guish explicitly between spacelike and timelike quantities, to make the appearance of temperature explicit.

The Wilson form for the gauge part is defined as

$$S_g = \frac{2N \cdot a_\tau}{g_\sigma^2 \cdot a_\sigma} \sum_{\mu < \nu < 4} P_x^{\mu\nu} + \frac{2N \cdot a_\sigma}{g_\tau^2 \cdot a_\tau} \sum_{\mu < 4} P_x^{\mu 4},$$

where

$$P_x^{\mu\nu} = 1 - \frac{1}{N} \text{Re Tr} U_x^\mu U_{x+\hat{\mu}}^\nu U_{x+\hat{\mu}}^{\mu+} U_x^{\nu+} \quad (3.10)$$

contains a product of gauge-field variables along the boundary of an elementary plaquette of the lattice,  $N=3$  for SU(3), and Tr denotes the trace in color space.

The first term of the action is a sum over plaquettes, that contain only spacelike links, whereas the second term involves timelike links as well. It can be easily checked that Wilson's formulation (3.10) satisfies both requirements for a lattice action.

To find an appropriate form for the fermionic part of the action turns out to be more difficult. A naive translation of the continuum Dirac action leads to a lattice action that actually describes  $2^d (=16)$  species of fermions in the naive continuum limit rather than the intended one species of the original continuum action. "No-Go" theorems by Nielsen and Ninomiya (1981a, 1981b) explain the failure of naive transcription and tell us that no lattice formulation of the fermionic action exists that is satisfactory in all aspects. If one insists on a local action, either continuous chiral symmetry is completely lost on the lattice or one ends up with too many flavors.

We consider two popular choices, *Wilson fermions* and *staggered fermions*. The former choice gets rid of the species doubling at the expense of breaking all continuous chiral symmetries explicitly in the limit of  $M \rightarrow 0$ . The latter choice keeps a  $U(1) \times U(1)$  chiral symmetry for all lattice couplings, a welcome feature for an investigation of chiral symmetry restoration. The price is too many flavors in the continuum limit (although the number is reduced with respect to the naive formulation) and a broken flavor symmetry on the lattice. For derivations and details of lattice fermions we refer the reader to the textbooks (e.g., Rothe, 1992; Montvay and Münster, 1994) or reviews (Creutz *et al.*, 1983; Kogut, 1983). Here we only summarize the results.

## 2. The Wilson action and hopping parameters

In the Wilson action the bare masses  $m_i$  are hidden in the hopping parameters in spacelike ( $\kappa_\sigma^{(i)}$ ) and timelike ( $\kappa_\tau^{(i)}$ ) directions, one for each flavor  $i$ , where  $i$  labels the  $N_f$  species of flavors in the continuum limit. For the free gauge theory, ( $U_x^\mu = 1$ ) and  $a_\sigma = a_\tau$ ,  $\kappa_\sigma^{(i)} = \kappa_\tau^{(i)} = \kappa^{(i)}$ , the relation to the bare quark masses is

$$\frac{1}{2} (1/\kappa^{(i)} - 1/\kappa_c) = \exp(m_i a) - 1 \quad (3.11)$$

with  $\kappa_c = 1/8$ . For massless quarks  $\kappa^{(i)} = \kappa_c$ . In the interacting theory, the dependence of  $\kappa_c$  on  $g^2$  is not known. This is an undesirable feature in Monte Carlo calculations when hopping parameter values should be translated to quark masses to check the relevance of the results for realistic mass values.

While the flavor symmetry is well defined and conserved for all lattice spacings by the Wilson action, all chiral symmetries are explicitly broken even in the massless case ( $m_i = 0$  or  $\kappa^{(i)} = \kappa_c$ ). "All" chiral symmetries should be contrasted with remnants of the full invariance under certain subgroups. Recall our original intention to study the phase structure within the lattice approach. A chiral symmetry restoration at finite temperature in the massless limit should be signalled by a melting of the condensate  $\langle \bar{\psi} \psi \rangle$ . Such a "melting" is prejudiced right from the beginning, if an explicit chiral symmetry breaking is involved, as it is in the Wilson formulation.

## 3. Staggered fermions and flavor symmetries

Staggered fermions are frequently used in studies of the chiral phase transition. A careful inspection of the origin of the species doubling in the naive formulation suggests the possibility of eliminating the unwanted fermions by doubling the effective lattice spacing. This amounts to a distribution of the fermionic degrees of freedom over the original lattice in such a way that the effective lattice spacing for each type of Grassmann variable is twice the fundamental lattice spacing. It turns out that in  $d$  dimensions  $2^{d/2}$  fermion fields are necessary to place a different fermionic degree of freedom at each site of an elementary hypercube on the lattice. This indicates the origin of the integer multiple of four flavor degrees in the continuum for four dimensions, which is attainable by a description with staggered fermions.

A realization of these ideas is rather involved. The sites of the hypercube will be occupied by single-component spinors  $\chi_f, \bar{\chi}_f$ , which may be multicomponent in flavor space,  $f=1, \dots, n$ . (This number  $n$  should be distinguished from the desired number  $N_f$  of continuum flavors.) The  $\chi$ 's and  $\bar{\chi}$ 's are certain linear combinations of the original fields  $\psi, \bar{\psi}$ . In the end of a simulation in terms of  $(\chi, \bar{\chi})$  fields, the results for the  $(\psi, \bar{\psi})$  fields have to be reconstructed.

The action for staggered fermions in terms of  $(\chi, \bar{\chi})$  fields is defined by

$$S_F \equiv \sum_{x, x'} \sum_{f=1}^n \bar{\chi}_f(x) Q_{x, x'}^f \chi_f(x')$$

with

$$Q_{x, x'}^f = \frac{1}{a_\sigma} \sum_{i=1}^3 D_{x, x'}^{(i)} + \frac{1}{a_\beta} D_{x, x'}^{(4)} + m_f \delta_{x, x'}$$

and

$$D_{x,x'}^{(\mu)} = \frac{1}{2} \Gamma_{\mu}(x) [U_x^{\mu} \delta_{x,x'-\hat{\mu}} - U_{x'}^{\mu+} \delta_{x,x'+\hat{\mu}}] \Gamma_{\mu}(x) \\ = (-1)^{x_1+x_2+\dots+x_{\mu-1}}. \quad (3.12)$$

The only remnant of the Dirac structure is hidden in the phases  $\Gamma_{\mu}$ . The staggered fermion action depends on the quark masses  $m_f$ . The index  $f$  labels the species of “staggered” flavors. Their number  $n$  is often set to 1, since  $n$  staggered species on the lattice correspond to  $N_f=4 \cdot n$  flavors in the continuum. The case  $n=1$  comes closest to  $N_f=3$ , for which QCD is approximately chiral invariant. More precisely, for  $n=1$  and  $d=4$  one has 16  $\chi$ 's in each elementary hypercube. It can be shown that these 16  $\chi$ 's can be combined to define a quark field with four flavors in the continuum, more generally with  $N_f=4 \cdot n$  flavors. A nonzero lattice spacing breaks the flavor symmetry between the  $4n$  flavors, and it breaks the chiral  $U(4n) \times U(4n)$  symmetry of the continuum limit of the staggered fermion action for  $m_f=0$  down to  $U(n) \times U(n)$ .

It is the  $U(1) \times U(1)$  remnant of the full chiral symmetry (which is left for  $n=1$  and arbitrary lattice spacing) that is the desired feature for investigating chiral symmetry restoration.

Wilson fermions and staggered fermions appear quadratically in the action. In general, fermion fields can only occur quadratically in a renormalizable four-dimensional field theory. Thus fermions are usually integrated out. This seems to be the only tractable way in numerical and analytical calculations, since the fermion fields are Grassmann variables. The integrals are performed using the well-known formulas of Berezin integration (Berezin, 1966). The result is

$$Z = \int_{pbc} \prod_{x,\mu} dU_x^{\mu} e^{-S_g \det^l(Q)}, \quad (3.13)$$

where  $Q$  is given by the action for Wilson fermions and by Eq. (3.12) for staggered fermions. The power  $l$  equals  $N_f$ , the number of continuum flavors, for Wilson fermions, and  $l$  equals  $n$  for staggered fermions (recall that  $n=N_f/4$ ). For a positive-definite  $Q$  one may use  $\det Q = \exp(\text{Tr} \ln Q)$ , which leads to

$$Z = \int_{pbc} \prod_{x,\mu} dU_x^{\mu} e^{-S_{\text{eff}}}$$

with

$$S_{\text{eff}}(\{U_x^{\mu}\}) = S_g(\{U_x^{\mu}\}) - l \text{Tr} \ln[Q(\{U_x^{\mu}\})], \quad (3.14)$$

where the trace extends over spin, color, and flavor indices. Thus a simulation of Eq. (3.14) with  $l=1$  describes  $N_f=4$  in the continuum limit, if staggered fermions are used for  $Q$ . For Wilson fermions  $\det Q(U)$  is real and positive for values of the hopping parameter less than  $1/8$  (Seiler, 1982).

#### 4. Sources of error

For staggered fermions (and vanishing chemical potential)  $\det Q$  is positive, but its eigenvalues are not al-

ways positive. This is an annoying feature of a representation in terms of pseudofermions. Therefore in actual simulations  $[-l \text{Tr} Q_x^{(\mu)}]$  in the effective action of Eq. (3.14) is replaced by  $(-l \text{Tr} \ln[\hat{Q} \hat{Q}^+(U)])$ . Consider first  $\hat{Q} = Q$ . The term  $\text{Tr} \ln Q Q^+$  is proportional to  $(\det Q)^2$ , thus the replacement induces a further doubling of flavor degrees of freedom from  $n$  to  $2n$ . To compensate for this doubling, one reduces the number of degrees of freedom in  $Q$  by a factor of 2 by a further doubling of the effective lattice spacing. Let us call the corresponding fermion operator  $\hat{Q}$ . Due to the even-odd symmetry of the determinant, one has  $\det Q = \det \hat{Q}_{\text{even}} \det \hat{Q}_{\text{odd}}$ . Thus  $\text{Tr} \ln Q$  is actually simulated as  $\text{Tr} \ln \hat{Q} \hat{Q}^+$ . This induces an error, which is not well under control. Another uncontrolled error entering the formulation for staggered fermions is due to flavor-exchanging currents. The flavor symmetry is violated for a finite lattice spacing. By claiming that their contribution vanishes in the continuum limit, one may induce an error similar to that induced when one says that the irrelevant additive terms in the Wilson action vanish in the continuum limit. It is well known that the very irrelevant terms in the Wilson action are essential for reproducing the right axial anomaly in the continuum limit.

#### 5. Translation from lattice units to physical units

A basic step in understanding the lattice approach is the translation of lattice results into physical units. For illustration let us consider the measurement of a mass. Such a measurement on the lattice typically yields a dimensionless number of the order of 1. The dimensionless lattice mass  $m_{\text{latt}}$  is related to the physical mass  $m$  via the lattice spacing  $a$ . From simple dimensional arguments we have

$$m = m_{\text{latt}} a^{-1} \quad (3.15)$$

in units where  $\hbar c=1$ . (In these units  $100 \text{ MeV} \sim 0.5 \text{ fm}^{-1}$ .) In other words, the lattice mass  $m_{\text{latt}}$  is measured in units of  $a^{-1}$ . Other variables are obtained similarly, energy densities in units of  $a^{-4}$ , etc.

When a lattice mass is interpreted as  $m \cdot a$  and  $m$  (MeV) is known from experiment, Eq. (3.15) gives “the” lattice spacing in physical units. More precisely it gives  $a(g)$  ( $\text{MeV}^{-1}$ ) at coupling  $g$ , if  $g$  stands for the bare input parameter(s) of the lattice Lagrangian, which have been used in the measurement of  $m_{\text{latt}}$ . In a pure gauge theory  $g$  is the bare gauge coupling. The lattice spacing obtained in this way is unique, i.e., independent of the choice of physical input  $m$  [MeV], only if one is in (or close to) the continuum limit.

Let us assume we have determined  $a$  ( $\text{MeV}^{-1}$ ) or  $a$  (fm) from a first mass measurement as indicated above and measure a second mass  $\tilde{m}_{\text{latt}}$  in lattice units. The physical value  $\tilde{m}$  [MeV] is then predicted from the lattice simulation and can be compared with the experimental value. Furthermore, once  $a(g)$  is known as a function of  $g$  in units of  $\text{MeV}^{-1}$  or fm, it makes sense to associate strong couplings ( $g \gg 1$ ) with coarse-grained

lattices (say  $a > 1$  fm), and weak couplings with a fine grain size close to the continuum description. The very existence of such a universal mapping between bare couplings and lattice spacings in physical units is based on the renormalization-group equation. In case of an  $SU(N_c)$  gauge theory with  $N_c$  colors and  $N_f$  massless flavors, the renormalization-group equation relates  $g^2$  and  $a$  in the continuum limit (for perturbatively small couplings  $g^2 \rightarrow 0$ ) according to

$$\begin{aligned} a\Lambda_L &= (b_0 g^2)^{-b_1/2b_0^2} e^{-1/2b_0 g^2}, \\ b_0 &= \frac{1}{16\pi^2} \left[ 11 \frac{N_c}{3} - \frac{2}{3} N_f \right], \\ b_1 &= \left( \frac{1}{16\pi^2} \right)^2 \left[ \frac{34}{3} N_c^2 - \left( \frac{10}{3} N_c + \frac{N_c^2 - 1}{N_c} \right) N_f \right]. \end{aligned} \quad (3.16)$$

If lattice calculations are performed at sufficiently small values of  $g$ , an observable in lattice units should scale as a function of  $g$  in a way that is determined by  $a(g)$  according to Eq. (3.16). For a mass in lattice units this implies

$$(ma)(g) = m[a(g)]. \quad (3.17)$$

If such a scaling behavior is observed, the asymptotic scaling regime has been reached. The lattice spacing can then be replaced by the scale parameter  $\Lambda_{\text{latt}}$ , which represents the only parameter in QCD with  $N_f$  massless flavors and has to be fixed from experiment.

The physical volume  $V$  and temperature  $T$  are given in terms of the lattice spacing as

$$V = N_\sigma^3 a_\sigma^3, \quad \frac{1}{\beta} \equiv T = \frac{1}{N_\tau a_\tau}. \quad (3.18)$$

For  $a_\tau = a_\sigma = a$ , the volume and temperature may be converted in to units of fm if  $a(g)$  is taken from Eq. (3.16).

For  $a_\tau \neq a_\sigma$ , the relations  $a_\tau(g_\sigma, g_\tau)$  and  $a_\sigma(g_\sigma, g_\tau)$  can be determined in the case of a pure gauge theory [Eq. (3.10)] as follows. We use the notations

$$\frac{2N}{g_\sigma} \equiv \beta_\sigma, \quad \frac{2N}{g_\tau} \equiv \beta_\tau, \quad \xi \equiv \frac{a_\sigma}{a_\tau} \equiv \frac{a}{a_\tau}, \quad \beta_g \equiv \frac{2N}{g^2}, \quad (3.19)$$

where  $g_\sigma$ ,  $g_\tau$  have been introduced in Eq. (3.10) and  $\beta_g$  denotes the bare coupling on an isotropic Euclidean lattice. Here  $\beta_g$  is large when  $g$  is small, as for the physical temperature (cf. the remarks in Sec. II.B.2). Thus one may also think of  $\beta_g$  as some kind of *effective temperature* (rather than an effective *inverse* temperature as the notation suggests). The two relations that replace  $g(a)$  for the isotropic case are written as

$$\begin{aligned} \beta_\sigma &= \xi^{-1} \beta [1 + c_\sigma(\xi) g^2 + O(g^4)], \\ \beta_\tau &= \xi \beta [1 + c_\tau(\xi) g^2 + O(g^4)]. \end{aligned} \quad (3.20)$$

The coefficients  $c_\sigma$  and  $c_\tau$  have been perturbatively determined by Karsch (1982). With  $\xi = a/a_\tau$  and  $g = g(a)$  from Eq. (3.16), Eq. (3.20) can be solved for  $a = a_\sigma$  and  $a_\tau$  in terms of  $\beta_\sigma$  and  $\beta_\tau$ , the couplings of the pure

gauge action. The relations  $a_\sigma(\beta_\sigma, \beta_\tau)$  and  $a_\tau(\beta_\sigma, \beta_\tau)$  replace  $a(g)$  on an anisotropic lattice in thermodynamic calculations.

## 6. The critical temperature $T_c$

In QCD the critical temperature  $T_c$  depends on the number of colors ( $N_c$ ), the number of flavors ( $N_f$ ), the current quark masses  $m_q$ , and the volume  $V$ . The finite volume induces a shift of  $T_c$  of the order of  $1/V$ , if  $V$  denotes the  $d$ -dimensional volume of the system. The finite light-quark masses give an effect of a few percent compared to the value of  $T_c$  in the chiral limit. The strongest dependence comes from the number of flavors.  $T_c$  varies about 100 MeV between  $T_c \sim 150$  MeV for two light flavors and  $T_c(N_f=0) \sim 260$  MeV for a pure gauge theory (DeTar, 1995). This is easily understood in a percolation picture. The transition is assumed to occur when a critical hadron density is reached at which the hadrons start overlapping, a condition we mentioned briefly in the introduction. A much lower temperature is then needed for creating a critical hadron density out of light pions with a mass of the order of  $T_c$ , than out of heavy glueball states. The lightest states in a pure gauge theory are glueballs with a mass of the order of  $5T_c$ .

For a smooth crossover phenomenon  $T_c$  is no longer defined. However, if there is a narrow temperature interval with rapid changes in thermodynamic quantities, it makes some sense to associate a “*pseudocritical*” or “*crossover*” temperature  $T_c$  with the rapid crossover region. More precisely, the crossover temperature is defined as the temperature of maximum change in an observable (e.g., the chiral condensate). Alternatively it is the temperature of the peak in some susceptibility. Both pseudocritical values for  $T_c$  must agree, only if the susceptibility is the derivative of the particular observable whose maximum change defines  $T_c$ .

## 7. A test of asymptotic scaling

Usually one measures the critical coupling  $g_c$  rather than the critical temperature  $T_c$ . Order parameters are plotted as functions of  $g$  [or  $\beta = 1/(6g^2)$ ] to show their behavior as a function of  $T$ . For simplicity let us consider an isotropic lattice  $a_\sigma = a_\tau = a$  and the limit of a pure gauge theory with bare coupling  $g$ . For a given number  $N_\tau$  of lattice sites in the timelike direction, the temperature may be varied by tuning  $a$  via  $g$  according to Eq. (3.18). From a physical point of view it is not surprising that the temperature may be implicitly varied via the coupling, as both are related through asymptotic freedom. Starting in the strong-coupling region  $g \gg 1$  and lowering  $g$ , we encounter the transition from confinement to the deconfinement phase at a certain coupling  $g_c$ . Alternatively,  $g$  can be kept fixed, but  $N_\tau$  varied from smaller to larger values (say from  $N_\tau = 2$  to  $N_\tau = 8$ ). In this way the transition region is passed from deconfinement to the confinement phase, if  $g$  and  $N_\tau$  are in an appropriate range.

One would like to simulate the lattice system for large values of  $N_\tau$  to keep the finite-size effects small. For

large  $N_\tau$  a rather small coupling is necessary to reach the high-temperature phase. The larger  $N_\tau$ , the smaller is the critical coupling  $g_c$ , which corresponds to the transition temperature  $T_c$ . In practice, the computer time rises rapidly if  $g$  is small. This explains why typical extensions in the time direction are limited to  $N_\tau = 6, 8, \text{ or } 12$ , when fermions are included.

The effect of fermions is a further reduction of the effective lattice spacing  $a_\tau$  at the same coupling  $g$  compared to the pure gauge case. Even larger values of  $N_\tau$  are needed to reach the transition region. Small values of  $N_\tau$  require stronger couplings  $g$  in the presence of fermions. One should keep in mind that one is far outside the asymptotic scaling region, if dynamical fermions are simulated at small  $N_\tau$ .

A *test of asymptotic scaling* is the very first check of whether the measured critical coupling  $g_c$  has some relevance for a critical temperature in the continuum limit. When the temporal size  $N_\tau$  is increased, the transition should occur at a smaller lattice spacing  $a(g_c)$  so that  $T_c = 1/[N_\tau a(g_c)]$  remains constant. For small couplings  $g$ , the relation  $g_c(N_\tau)$  should scale as

$$T_c / \Lambda_L = N_\tau^{-1} (b_0 g_c^2)^{b_1/2} b_0^2 e^{1/2 b_0 g_c^2}. \quad (3.21)$$

Equation (3.21) follows from Eqs. (3.16) and (3.18). Calculations for an SU(2) and SU(3) gauge theory have been performed up to  $N_\tau = 16$ . A measurement of  $g_c$  shows strong violations of asymptotic scaling according to Eq. (3.21) (see, for example, Karsch, 1993).

Asymptotic scaling violations have also been observed in other observables. They seem to be universal. Thus it should be possible to absorb the scaling violations in a “renormalization” of the bare coupling to an effective coupling which accounts for rapid fluctuations in the action in an intermediate coupling range. This idea turned out to be successful (Altmeyer *et al.*, 1993). As a function of an effective coupling  $T_c / \Lambda_{\overline{\text{MS}}}$  has a much weaker dependence on  $N_\tau$ . An extrapolation to the continuum limit seems to be justified, leading to (Karsch, 1993)

$$\frac{T_c}{\Lambda_{\overline{\text{MS}}}} = \begin{cases} 1.23 \pm 0.11, & \text{SU(2) gauge theory} \\ 1.03 \pm 0.19, & \text{SU(3) gauge theory.} \end{cases} \quad (3.22)$$

## 8. Translation to physical units

A second check as to whether the critical temperature on the lattice has some relevance for the continuum limit is a translation to physical units. The outcome should be independent of the choice of the experimental input. In the pure SU(3) gauge theory,  $T_c$  is approximately independent of the physical input. It is 239(13) MeV from the string tension  $\sqrt{\sigma} = 420 \pm 20$  MeV, 239(23) MeV from the  $\rho$  mass and 225(30) MeV from the nucleon mass (Karsch, 1993).

For QCD with two light flavors in the staggered fermion formulation the *crossover temperature* is estimated as 140–160 MeV (DeTar, 1995). The experimental input comes from the  $\rho$  mass with 770 MeV. The critical temperature is then determined as

$$T_c \text{ (MeV)} = \frac{770 \text{ (MeV)}}{(m_\rho a)(6/g_c^2, m_q a) N_\tau}. \quad (3.23)$$

The  $\rho$  mass in lattice units  $m_\rho a$  (depending on the bare gauge coupling  $g_c$  and the bare quark masses  $m_q$ ) is calculated as a function of  $g_c$  and  $m_q a$  from a fit to several zero-temperature simulations (see DeTar, 1995, and references therein). Note that in Eq. (3.23) the  $\rho$  mass [MeV] is kept fixed at its physical value of 770 MeV, while the bare quark masses  $m_q a$  are allowed to vary to unphysical values leading to unphysical  $\rho$  masses as well. As emphasized by DeTar (1995), one should further keep in mind that the nucleon to-rho mass ratio is approximately 20% above its physical value over the range of lattice parameters, where  $m_\rho a$  is measured. Also the  $m_\pi/m_\rho$  ratio comes out twice to three times the experimental value, because the bare quark masses are still too large. Thus the  $\rho$  mass is not (yet) a perfect candidate for conversion from lattice units into physical units.

In earlier lattice simulations thermodynamic quantities were plotted as a function of  $6/g^2$ . Nowadays one often finds plots vs  $T$  (MeV). The translation should be made with care, as the mapping between  $6/g^2$  and  $T$  is not (yet) unique.

## 9. Numerical simulations

So far we have specified the measure, boundary conditions, and lattice action for the path integral of Eq. (3.8). We have sketched how numbers from the lattice can be translated into physical units, and what a criterion for the relevance of lattice results for continuum physics looks like. Numerical and analytical methods can be utilized to attack the functional integrations of Eq. (3.8). Analytical methods are usually applicable in limits where small expansion parameters are available. The *hopping parameter* expansion is an expansion in small values of  $\kappa$ , around the limit of infinite quark masses. The *strong coupling* expansion applies for large values of the bare gauge coupling  $g$ . As mentioned above, it is intrinsically difficult to find a small expansion parameter in the transition region, where  $g$  is neither weak nor strong. Thus it is not surprising that many results on the phase structure in the vicinity of  $T_c$  are based on *numerical* simulations. The most important approach is the Monte Carlo method.

### a. Monte Carlo on the lattice

In a Monte Carlo procedure a set of field configurations  $\{U_x^{\mu(1)} \dots U_x^{\mu(N)}\}$  is generated such that the Boltzmann factor is absorbed in the selection of configurations. For a positive-definite and real action  $S[\{U_x^\mu\}]$ , the expectation value of an observable  $O$  defined as

$$\langle O \rangle = \frac{\int \prod_{x,\mu} dU_x^\mu e^{-\beta S[\{U_x^\mu\}]} O(\{U_x^\mu\})}{\int \prod_{x,\mu} dU_x^\mu e^{-\beta S[\{U_x^\mu\}]}}, \quad (3.24)$$

[a lattice version of Eq. (3.5) for a pure gauge theory] equals

$$\langle O \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N O(\{U_x^{\mu i}\}) \quad (3.25)$$

and is approximated by the arithmetic average over a finite number of configurations,

$$\langle O \rangle_N \approx \frac{1}{N} \sum_{i=1}^N O(\{U_x^{\mu i}\}). \quad (3.26)$$

The observables in Eq. (3.26) are evaluated on an ensemble of configurations which are representative of the coupling (temperature)  $\beta^{-1}$ . The configurations are selected according to the probability distribution

$$P(U)\mathcal{D}(U) \equiv \frac{e^{-\beta S[U]}}{Z} DU, \quad (3.27)$$

where  $DU$  and  $Z$  are shorthand notations of the expressions in Eq. (3.24). The fundamental idea of the Monte Carlo method is stated by the law of large numbers, which tells us that under very general conditions on the probability measure  $\lim_{N \rightarrow \infty} \langle O \rangle_N = \langle O \rangle$ .

The sequence of configurations is generated with a Markov process. An easy way of constructing a Markov process, that has a particular distribution  $P$  as its fixed point is to choose the transition probabilities  $Q: (\{U_x^{\mu i}\} \rightarrow \{U_x^{\mu f}\})$  from an initial configuration  $\{U_x^{\mu i}\}$  to a final configuration  $\{U_x^{\mu f}\}$  to satisfy the detailed balance condition

$$P[U^{(i)}]Q(U^{(i)} \rightarrow U^{(f)})\mathcal{D}U = P[U^{(f)}]Q(U^{(f)} \rightarrow U^{(i)})\mathcal{D}U. \quad (3.28)$$

The fixed-point distribution is in our application the Boltzmann equilibrium distribution. A particular way of realizing the detailed balance condition is the Metropolis algorithm (Metropolis *et al.*, 1953).

The Metropolis algorithm is an example of a local updating procedure. One starts from an initial configuration, which may be chosen ordered [ $U_x^{\mu} = 1$  for all links (cold start)] or disordered [ $U_x^{\mu} \in SU(N)$  random for all links (hot start)], or mixed. A single new link variable  $U_x^{\mu'}$  is then chosen randomly and always accepted as a replacement for the old variable if it lowers the action, i.e.,  $\Delta S = S(\{U'\}) - S(\{U\}) < 0$ . Otherwise the change is accepted with a conditional probability. For  $\Delta S > 0$  a random number  $r$  with  $0 \leq r \leq 1$  is selected. If  $r < e^{-\Delta S}$ , the new variable  $U'$  is still accepted for replacing the old one; otherwise one goes back to  $\{U_x^{\mu}\}$  and repeats the steps as indicated above.

In this way all links of the lattice are changed, either randomly or successively. Such a sweep through the entire lattice is counted as one Monte Carlo iteration. Usually hundreds or thousands of such iterations are necessary before the equilibrium distribution is reached. The expectation value of the observable is then obtained according to Eq. (3.26), where the sum extends only over the last  $N$  equilibrated (or thermalized) configurations.

When fermions are included, one encounters a technical problem due to the nonlocality of the effective action Eq. (3.14). The nonlocality refers to the  $\det$  term or equivalently to the  $\text{Tr} \ln Q$  term. For staggered fermions

the matrix  $Q$  has  $(3N_{\sigma}^3 \cdot N_{\beta})^2$  complex elements. In the local Metropolis algorithm this matrix should be calculated for every link in every iteration, which renders it impracticable.

#### b. The hybrid Monte Carlo algorithm

In the last decade much effort has been invested in improving algorithms for QCD with dynamical quarks (for a review, see the lectures by Toussaint 1988 or Herrmann and Karsch, 1991). Integration over the fermions led to the factor  $\det(Q)$  in Eq. (3.13). While  $Q$  was a local operator, coupling only nearest neighbors on the lattice,  $\det Q$  is nonlocal. A calculation of  $\det Q$  with  $\exp\{-S_f\}$  as Boltzmann factor should be avoided. The idea is to bosonize  $\det Q$  and to “lift”  $Q$  into the exponent. The bosonization is performed with the *pseudofermion method* of Petcher and Weingarten (1981).

The *pseudofermion method* (Petcher and Weingarten, 1981) is based on the following formula for the determinant:

$$\det Q = \frac{1}{\det Q^{-1}} = \int \mathcal{D}\lambda^* \mathcal{D}\lambda e^{-\sum_{x,y} \lambda_x^* Q_{xy}^{-1} \lambda_y}. \quad (3.29)$$

Here  $\lambda, \lambda^*$  are complex bosonic (pseudofermionic) variables. Note that  $Q$  has to be a positive-definite matrix for the Gaussian integral in Eq. (3.29) to converge. If we identify  $Q$  with the fermionic matrix of Eq. (3.13),  $Q$  has negative eigenvalues. This is the reason why  $Q$  is replaced by  $Q^+ Q$  at the cost of doubling the number of fermionic flavors or (for compensating this doubling) by  $\hat{Q}^+ \hat{Q}$ , where  $\chi$  fields live on even and  $\bar{\chi}$  fields on odd sites. (Another reason for the replacement is that it is convenient for easily generating the pseudofermion fields from a heat bath.) Via the doubling ( $Q \rightarrow Q Q^+$ ) configurations leading to negative eigenvalues of  $Q$  are mapped on the same contribution to the effective action as those with positive eigenvalues, although  $\ln \det Q$  in the original effective action is not given as  $\text{Tr} \ln Q$  for negative eigenvalues. Thus the error entering this way is difficult to control.

Like  $\det Q$ , the action in Eq. (3.29) is nonlocal. Hence we see, when the bosonic part ( $-\sum_{x,y} \lambda_x^* Q_{xy}^{-1} \lambda_y$ ) is added to the gauge part of the QCD action to account for fermions, we need a nonlocal updating procedure. A practical method, which is free of uncontrolled errors and often used nowadays in QCD applications, is the *hybrid Monte Carlo algorithm* (Duane *et al.*, 1987). It has to be combined with the pseudofermion representation to yield a feasible simulation of QCD. The hybrid Monte Carlo algorithm is a stochastic hybrid of a Langevin and a microcanonical algorithm, for which an additional Monte Carlo acceptance test has been built in to remove the systematic errors of the stochastic hybrid. The stochastic hybrid algorithm combines the virtues of the Langevin and the microcanonical (or molecular dynamics) algorithms. Both algorithms use equations of motion to select new variables.

In the *Langevin algorithm* (Parisi and Wu, 1981; Fukugita and Ukawa, 1985; Batrouni *et al.*, 1985), the

equation of motion is a partial differential equation, which is of first order in time. The finite step size in time entering the discretized version has to be extrapolated to zero by varying its size. This is time consuming and an unwanted feature.

The *microcanonical algorithm* (Callaway and Rahman, 1982, 1983) is based on the observation that the Euclidean path integral of a quantum field theory in four dimensions can be written as a partition function for a system of classical statistical mechanics in four spatial dimensions with a canonical Hamiltonian that governs the dynamics in a fifth new “time” variable. This time  $\tau$  may be identified with the simulation time. The algorithm determines the new variables  $\Phi_x(\tau)$  and their canonically conjugate momenta  $\pi_x(\tau)$  in a fully deterministic way. Here one moves faster through phase space  $(\phi, \pi) \rightarrow (\phi', \pi') \rightarrow \dots$ , but at the price of losing ergodicity.

The *stochastic hybrid algorithm* (Duane, 1985) interrupts the integration of Hamilton’s equation of motion along a single trajectory with fixed energy. The interruption is made once in a while for a refreshment of momenta. The new parameter, which enters and should be optimized, is the frequency of momenta refreshment. If it is low, the algorithm is as slow as the Langevin algorithm, if it is too high, ergodicity will be violated. Yet there is a systematic error in the hybrid algorithm introduced by the finite time step in the integration procedure, which leads to a violation of energy conservation.

The idea of the *hybrid Monte Carlo algorithm* is to absorb this energy violation in a “superimposed” Monte Carlo procedure. The phase-space configurations at the end of every molecular dynamics chain are subjected a Metropolis acceptance test. If the time step is too large, the energy violation too strong, the  $(\Phi, \pi)$  configuration is likely to be rejected. When the energy is conserved, the configuration is always accepted. In this way the systematic error due to finite integration steps is eliminated.

Above we have sketched a number of intermediate steps in developing the hybrid Monte Carlo algorithm, in order to give an idea of how complex the problem of finding efficient algorithms is. Although the hybrid Monte Carlo algorithm moves the system quickly through configuration space, it is slow due to a matrix inversion, which enters the equations of motion. The matrix inversion is usually performed with the conjugate gradient method. The number of conjugate gradient steps is proportional to  $(m_f^{-1})$  and takes more than 90% of the CPU time. The hybrid Monte Carlo acceptance rate is proportional to  $(m_f^{-3/4})$  and the autocorrelation time  $\propto m_f^{-1}$  (Karsch, 1992). For example, one lattice update on a  $16^3 \times 8$  lattice with  $m_f = 0.01$  takes about 20 min on a 200 MFlop machine. Here we see why it is so difficult to work with small bare-quark masses. Quark masses should be as small as possible to avoid an uncontrolled extrapolation to the chiral limit, but some extrapolation is unavoidable.

More generally we can see why Monte Carlo simulations of lattice QCD are so time consuming when dy-

namical fermions are included. It is extremely hard to satisfy the following requirements simultaneously:

(i) The masses are small enough to guarantee a safe extrapolation to the chiral limit.

(ii) The volume is large enough to avoid finite-size effects.

(iii) The bare coupling is small enough to be in the asymptotic scaling regime.

## B. Pure gauge theory

### 1. The order of the SU(3) deconfinement transition

The deconfinement transition in a pure SU(3) gauge theory is believed to be of first order. For several years this point was the subject of a lively debate in the literature, due to the use of different criteria to determine the order from a numerical analysis. In Sec. II we distinguished between “naive” and “refined” criteria. In this section we shall illustrate their application to the deconfinement transition.

Early calculations of Kogut *et al.* (1983) and Celik *et al.* (1983a, 1983b) found a strong first-order deconfinement transition. The conclusions were based on hysteresis effects, signals for coexisting states, and abrupt quantitative changes in bulk thermodynamic quantities like the internal energy density. One finds in fact an explicit suggestion that hydrodynamic models of the quark gluon plasma and models of hadronization should incorporate a “hard, first-order transition” (Kogut *et al.*, 1983).

These results were questioned by Bacilieri *et al.*, 1988, who claimed to see strong indications for a second-order deconfinement transition. The claim was based on finding a correlation length that grows proportional to the lattice size, as the size was increased. The results of Bacilieri *et al.* led to a careful reinvestigation of the deconfinement transition.

From Sec. II.B we recall the renormalization-group analysis of Svetitsky and Yaffe (1982a, 1982b), in which it was argued that an effective theory for an order parameter of a (3+1)-dimensional SU( $N$ ) gauge theory shares the global symmetry with three-dimensional  $Z(N)$  spin models. In particular, knowledge about  $Z(3)$  spin models was exploited to predict a first-order phase transition for the case of an SU(3) gauge theory, based on absence of an IR-stable fixed point in  $Z(3)$  spin systems.

A possible second-order phase transition in an SU(3) gauge theory would cast some doubts on the universality arguments. While it was not difficult to “derive” the  $Z(3)$  symmetry of the effective action in terms of the order-parameter field, it was harder to argue in favor of the locality of the spin model. Thus the question of nonlocal interaction terms in the effective action was addressed by Fukugita *et al.* (1989 and 1990a) and by Gvai *et al.* (1989) and Gvai and Karsch (1989) in subsequent papers.

In general the order of the transition does depend on the type of effective couplings. These couplings are gen-

erated when the original degrees of freedom are partially integrated out. Nonlocal couplings could cause the SU(3) transition to be of second order. This suspicion was supported by earlier results of Fucito and Vulpiani (1982), who stated that already small anti-ferromagnetic next-to-nearest-neighbor (NNN) couplings in addition to the leading ferromagnetic nearest neighbor (NN) coupling induce a second-order transition in a three-dimensional Z(3) Potts model. It was suggested that likewise a small antiferromagnetic nnn coupling could induce a second-order transition in the SU(3) deconfinement transition.

Both transitions in variants of the Z(3) Potts model and in the original SU(3) gauge theory were then studied by applying refined criteria compared to earlier calculations. Here we report in greater detail on the results of Gavai *et al.* (1989) on the Z(3) Potts model and of Fukugita *et al.* (1989) on the SU(3) gauge theory. In both models the first-order nature of the transition was verified with a variety of criteria, and the results of the APE group are now thought to have been most likely a misinterpretation of the measured data.

Let us recall from Sec. II the distinction between naive and refined criteria. Naive criteria are signs of metastabilities, hysteresis effects, two-state-signals, or jumps in thermodynamic quantities. Typically all these signatures occur in a truly first-order transition. If we call them “naive,” we refer only to their incautious use when the signatures are taken for granted from measurements for a single lattice size. Early lattice simulations were “naive” in this sense. Compared to present standards they were also performed on rather small lattices.

Refined criteria refer to a finite-size scaling analysis of bulk and “special” quantities. Examples of bulk quantities are the internal energy density  $\varepsilon$ , the specific heat  $c$ , or the order-parameter susceptibility  $\chi$ . Finite-size corrections to the average internal energy are *exponentially* suppressed. The shifting and rounding effects in the peaks of  $c$  and  $\chi$  are of the order of  $1/V$  in first-order transitions (see Sec. II).

By “special quantities” we mean distribution functions of the order parameter  $P(O)$ , and derived moments or cumulants like Binder’s cumulant; see Table II. In general they have power-law finite-size correction and are practically applicable for very large volumes, where these power-law corrections are suppressed.

Similarly correlation lengths show *power-law* corrections. Even their very definition in a finite volume is rather intricate compared to bulk quantities. Their practical utility for a finite-size scaling analysis turns out to be questionable, as we show below.

We turn now to concrete applications of both kinds of criteria. Kogut *et al.* (1983) found large gaps in the internal energies of gluons (and fermions in the quenched approximation). The lattice size was  $2 \times 8^3$  and  $4 \times 8^3$ . Celik *et al.* (1983a, 1983b) recorded two-state signals from hot and cold starts, indicating two coexisting phases with different expectation values of the Wilson line at the critical coupling. On lattices with  $8^3$  and  $10^3 \times 2$  (or  $\times 3, 4$ ) sites they also found a large latent

heat measured as the difference in gluonic internal energies ( $\Delta\varepsilon/T_c^4 = 3.75 \pm 0.25$ ). Lattice expressions for the latent heat and other thermodynamic quantities will be given in Sec. III.B.2.

Brown, Christ, *et al.* (1988) performed a study concurrent with that of Bacilieri *et al.* (1988). They still found a first-order deconfining transition, but it was less strong, with signs of an increasing correlation length. The lattice volumes were enlarged to  $16^3 \times 4$  and  $16^3 \times 6$  and  $24^3 \times 4$  and  $24^3 \times 6$ . The values for the latent heat [obtained from the gap in  $(\varepsilon + p)$ ] were  $\Delta\varepsilon/T_c^4 = 2.54 \pm 0.12$  for  $N_\beta = 4$  and  $\Delta\varepsilon/T_c^4 = 2.48 \pm 0.24$  for  $N_\beta = 6$ . This is 60% and 25% smaller than the earlier values on smaller lattices, respectively.

Correlation lengths are determined from a fit of Wilson line correlation functions [see Eq. (3.40) and (3.41) below for the definition] with a string potential ansatz in the confinement phase and a screened Yukawa potential in the deconfinement phase. The results are increasing correlation lengths by a factor of 2 to 3 on both sides of the transition. The relatively large correlation length can be traced back to the  $3d$  Z(3) Potts model (see our discussion in Sec. II.A).

Bacilieri *et al.* (1988) focused their study on Wilson line correlation functions. Large correlation lengths were found, increasing proportional to the lattice size. This increase was interpreted as suggestive of a second-order transition. Only a careful finite-size scaling analysis of correlation lengths can show that this interpretation is most likely not adequate. Such an analysis is rather involved. Therefore we devote the following section to it.

### a. Correlation lengths, mass gaps, and tunneling events

We use a notation that is independent of the application in a spin or gauge model. We shall distinguish *physical* and *tunneling correlation lengths*  $\xi^{(p)}$  and  $\xi^{(t)}$  with associated mass scales  $m^{(p)}$  and  $m^{(t)}$  both in the finite and the infinite volume. A physical or bulk correlation length in the infinite-volume limit can be defined via the decay of the connected pair correlation  $C_c$  as a function of the distance  $z$  according to

$$C_c(z) = C(z) - B \xrightarrow{z \rightarrow \infty} A e^{-m^{(p)}z}, \quad (3.30a)$$

$$\xi^{(p)} \equiv 1/m^{(p)}. \quad (3.30b)$$

Here  $C$  is the unsubtracted, disconnected pair correlation and  $B$  stands for the constant disconnected part, which is different from zero in the phase of spontaneously broken symmetry. The inverse correlation length  $\xi^{(p)-1} = m^{(p)}$  is called *the physical mass gap*.

In the conventional lore it is the infinite-volume correlation length which diverges as  $\beta \rightarrow \beta_c$  for a second-order transition and stays finite in the first-order case. Here we would like to stress the order of limits,

TABLE IV. Infinite-volume behavior of the mass gap, the tunneling mass  $m^{(\text{unp})}$ , and their corresponding correlation lengths.

Phase	First-order transition	Second-order transition
Symmetric	$m^{(p)}$ stays finite as $\beta \rightarrow \beta_c^-$	$m^{(p)} \rightarrow (\beta - \beta_c)^\nu \rightarrow 0$ as $\beta \rightarrow \beta_c^-$ $\xi^{(p)} \rightarrow \infty$
Broken	$m^{(p)}$ and $\xi^{(p)}$ are finite	$m^{(p)}$ and $\xi^{(p)}$ are finite
Symmetric	$m^{(\text{unp})}$ stays finite as $\beta \rightarrow \beta_c^-$	$m^{(\text{unp})} \sim (\beta_c - \beta)^\nu \rightarrow 0$ as $\beta \rightarrow \beta_c^-$ , $\xi^{(t)} \rightarrow \infty$
Broken	$m^{(\text{unp})}=0, \xi^{(t)}=\infty$	$m^{(\text{unp})}=0, \xi^{(t)}=\infty$

$$\lim_{\beta \rightarrow \beta_c^-} \lim_{L \rightarrow \infty} \xi_L^{(p)}(\beta, L) = \lim_{\beta \rightarrow \beta_c^-} \xi^{(p)}(\beta) \begin{cases} \rightarrow \infty & \text{second order} \\ \rightarrow \text{const} < \infty & \text{first order.} \end{cases} \quad (3.31)$$

In the opposite order,

$$\lim_{L \rightarrow \infty} \lim_{\beta \rightarrow \beta_c^-} \xi_L^{(p)}(\beta, L) = \lim_{L \rightarrow \infty} \xi_L^{(p)}(\beta_c, L) \begin{cases} \rightarrow \infty & \text{second order} \\ \rightarrow \infty & \text{first order.} \end{cases} \quad (3.32)$$

The ‘‘critical’’ correlation length  $\xi_L^{(p)}(\beta_c, L)$  also diverges at  $\beta_c$  in the infinite-volume limit in the case of a first-order transition (Blöte and Nightingale, 1982; Privman and Fisher, 1983). This behavior is familiar, if we recall from Sec. II.A that singularities in second derivatives of thermodynamical potentials occur for first-order transitions as well. The only difference in the second-order case is the type of divergence:  $\delta$ -function singularities may occur in the specific heat and the magnetic susceptibility, when the infinite-volume limit is taken at criticality. Thus a diverging correlation length  $\xi_L^{(p)}(\beta_c, L)$  for  $L \rightarrow \infty$  is not conclusive evidence for a second-order transition.

In the following discussion we consider infinite-volume correlation lengths according to Eq. (3.31). In Table IV we compare the behavior of the physical and tunneling correlation lengths and their associated mass scales in the infinite-volume limit.

In the first row of Table IV we have listed the behavior of the physical mass gap and its associated bulk correlation length, when the critical coupling  $\beta_c$  is approached from the symmetric phase (i.e.,  $\beta_c$  from below), denoted as  $\beta \rightarrow \beta_c^-$ . In the first-order case  $\xi^{(p)}$  stays finite, while it diverges as  $\beta \rightarrow \beta_c^-$  for a second-order transition. In the broken phase the physical mass gap stays finite in both cases (second row).

In the lower two rows we introduce an ‘‘unphysical’’ mass gap  $m^{(\text{unp})}$  and its corresponding tunneling correlation length  $\xi^{(t)}$ . In the symmetric phase both quantities coincide with  $m^{(p)}$  and  $\xi^{(p)}$ , respectively, while  $\xi^{(t)}$  is

always infinite in the broken phase in the infinite-volume limit. Here  $m^{(\text{unp})}$  may be defined via an analogous relation to Eq. (3.30a), but this time via the decay of the disconnected correlation function,

$$C(z) \xrightarrow{z \rightarrow \infty} e^{-m^{(\text{unp})}z}. \quad (3.33)$$

The vanishing of  $m^{(\text{unp})}$  defined via Eq. (3.33) just reflects the nonvanishing constant  $B$  in the broken phase, which prevents an exponential decay of the correlation due to spontaneous symmetry breaking. The relation to the associated tunneling correlation length is not as simple as in Eq. (3.30b).

From an analysis of the transfer matrix one expects in general that the (disconnected pair) correlation function will be a coherent sum of its eigenvalues. If we consider only the two smallest eigenvalues above the ground state, the correlation  $C(z)$  may be written as

$$C(z) \sim a_1 e^{-z/\xi^{(p)}} + a_2 e^{-z/\xi^{(t)}}, \quad (3.34)$$

which should be compared with Eq. (3.33). Equation (3.34) defines the tunneling correlation length  $\xi^{(t)}$ , or more precisely,  $\xi^{(t)}$  arises in the spectrum of the transfer matrix from the level splitting associated with tunneling between different degenerate vacua. (We anticipated this effect in Sec. II.A.)

This tunneling has a finite probability as long as the volume is finite, resulting in a finite  $\xi_L^{(t)}$ , but the potential barrier between the degenerate ground states becomes infinitely high in the infinite-volume limit, leading to

$$\lim_{L \rightarrow \infty} \xi_L^{(t)} = \infty. \quad (3.35)$$

From Eq. (3.34) we recognize that as  $L \rightarrow \infty$  the diverging  $\xi^{(t)}$  gives rise to the constant disconnected term  $B$  in the correlation function Eq. (3.30a).

In a *finite volume* both length scales  $\xi^{(p)}$  and  $\xi^{(t)}$  are necessarily finite and, when considered at criticality [cf. Eq. (3.32)] both increase with increasing volume at a first-order transition. Thus there is a twofold risk of misinterpreting an increasing correlation length: the order of limits has not been properly arranged and  $\xi^{(t)}$  is intertwined with  $\xi^{(p)}$ .

*Concrete realizations* of the correlation function  $C$  are spin-spin or Wilson line pair correlations, if we deal with applications in the SU(3) deconfinement transition. We give the infinite-volume definitions, written for a large but finite volume  $V = L^3$ .

In the 3d three-state Potts model considered by Gavai *et al.* (1989), the spin-spin correlations are defined as

$$C_1(z) = \frac{1}{6V} \left\langle \sum_i s_i s_j^+ \right\rangle, \quad (3.36)$$

where  $z = |i - j|$  and the corresponding zero-momentum projection reads

$$C_0(z) = \frac{1}{6L} \left\langle \sum_{i=1}^L \bar{s}_i \bar{s}_j^+ \right\rangle. \quad (3.37)$$

Here  $s_i$  is the spin at site  $i$  and  $\bar{s}_i$  the average spin over

a plane  $i$ , while  $z$  denotes the distance between planes  $i$  and  $j$  along one of the principal axes of the lattice. The connected part

$$C_{i,c}(z) = C_i(z) - \lim_{z \rightarrow \infty} C_i(z), \quad i = 0, 1, \quad (3.38)$$

defines the physical correlation length  $\xi^{(p)}$  as

$$1/\xi^{(p)} = - \lim_{z \rightarrow \infty} \frac{1}{z} \ln C_{i,c}(z) \quad (3.39)$$

in accordance with Eq. (3.30a).

The analogous quantity  $C_0(z)$  for an SU(3) lattice gauge theory was considered by Fukugita *et al.* (1989) and is the zero-momentum correlation function of Wilson lines

$$C_0(z) = \frac{1}{3L} \sum_{i=1}^3 \sum_{z'=1}^L \langle \bar{\Omega}_i(z') \bar{\Omega}_i^+(z'+z) \rangle, \quad (3.40)$$

where  $i$  labels the planes, and  $\bar{\Omega}_i(z)$  is the average of Wilson lines  $\Omega_n$  over the  $i$ th plane. In this model  $n_i = z$ , where

$$\Omega_n = \frac{1}{3} \text{Tr} \left( \prod_{\tau=1}^{N_\tau} U_{(n,\tau)}^{(4)} \right), \quad (3.41)$$

i.e., the trace of the product over gauge-field variables  $U_{(n,\tau)}^{(4)} \in \text{SU}(3)$  associated with timelike links leaving the site  $(n, \tau)$  in direction 4;  $n$  denotes the spatial coordinates. The additive constant  $B$  of Eq. (3.30a), which can be split off in  $C_0(z)$ , is given by  $|\langle \Omega \rangle|^2$ .

Bacilieri *et al.* (1988) calculated Wilson line pair correlations essentially with the same formulas as (3.40) and (3.41). They used a source method, in which all link variables with  $z$  coordinate 1 were fixed to the identity matrix 1. The ir correlation reads

$$C_1(z) = \left\langle \frac{1}{L_x L_y} \text{Re} \sum_{x,y} \Omega_{(x,y,z)} \cdot 1 \right\rangle, \quad (3.42)$$

where  $\Omega_{(x,y,z)} \equiv \Omega_n$  from Eq. (3.41).

### b. Adaptation to the finite volume

The above definitions for the various correlation lengths are applicable in the large-volume limit, that is, if  $\lim_{L \rightarrow \infty}$  in Eq. (3.31) is approximately realized. As subleading corrections from the finite-volume limit are powerlike, the large-distance behavior of correlations ( $z \rightarrow \infty$ ) is not accessible on lattices of moderate size.

For first-order transitions it is rather hard to predict the finite-size dependence from first principles. [The approach of Borgs and Kotecky (1990) comes closest to a rigorous derivation.] In general one has to make at some place an ansatz. The double-peak Gaussian ansatz for the distribution function  $P_L(E)$  has led to definite predictions of the finite-size behavior.

Likewise it requires an ansatz for the functional form of the correlation function  $C_L(Z)$  in a finite volume. A procedure that was adopted by Gavai *et al.* (1989) and Bacilieri *et al.* (1988) is to form ratios of correlation functions,

$$R_i(z) = \frac{C_i(z)}{C_i(z+1)}. \quad (3.43)$$

A finite-volume ansatz for  $C_i(z)$ , which was inspired by the infinite-volume decay and adapted to the choice of periodic boundary conditions on a cubic lattice, is

$$C_i(z) = A_i [\exp(-z/\xi) / z^i + \exp(-[L-z]/\xi) / (L-z)^i], \quad (3.44)$$

where  $\xi$  stands for a generic correlation length and  $i=0,1$ . Equation (3.44) includes only the first image of a source in the  $z$  direction, the one that is a distance  $L$  apart. Multimirror images (accounting for images a distance  $nL$  apart) should be also included, if the correlation can become large compared to the considered volume.

The careful analysis of Gavai *et al.* for the 3d Z(3) Potts model shows that different methods of extracting the physical correlation length  $\xi^{(p)}$  from data on finite lattices do not agree in the critical region, although they do agree away from  $T_c$ . Depending on the ansatz for  $R_i$ ,  $\xi$  seems to grow with  $L$ , or to stay finite ( $\xi \sim 10$  in lattice units) in the vicinity of  $T_c$ . The former behavior is compatible with  $\xi = \xi^{(l)}$  as  $L \rightarrow \infty$ , the latter with  $\xi = \xi^{(p)}$ , i.e., a physical mass gap which stays finite at  $T_c$  and indicates the first-order nature of the transition.

To summarize, it is the very definition of a correlation length in a finite volume, the order of limits ( $L \rightarrow \infty$  and  $\beta \rightarrow \beta_c$ ) and the occurrence of the tunneling correlation length, which impede the finite-size analysis of correlation functions.

After this digression we proceed with concrete applications in the SU(3) deconfinement transition. Bacilieri *et al.* (1988) considered ratios of finite-volume correlations according to Eq. (3.44), with  $C_1$  given by Eq. (3.44), very similar to the observables that were later considered by Gavai *et al.* (1989). Rather than measuring the correlations  $C_i(z)$ , they applied a so-called smearing procedure (Albanese *et al.*, 1987) to generate operators weakly coupled to the high-energy fluctuations. We do not go into technical details, although the smearing technique is quite important for following the decay of the correlation over a large distance up to 15  $z$  slices on a lattice of cylindrical geometry  $12^2 \times 48 \times 4$ .

Alternatively “the” correlation length  $\xi$ , defined via Eqs. (3.30a) and (3.42) was measured from a three-parameter fit with  $C(z) = A \cosh[m \frac{1}{2} L_z - z]$ . The statements of the paper by Bacilieri *et al.* about the increasing correlation length refer to this length scale  $\xi = 1/m$ . When  $m$  coincides with the physical mass gap  $m^{(p)}$ , it is related to the string tension  $\sigma$  via  $m^{(p)} = \sigma N_\beta$ . This relation holds in a pure gauge theory. Thus the statement about the physical correlation length  $\xi^{(p)}$  tells us the behavior of the string tension in the critical region of the SU(3) gauge theory.

The resolution of the contradicting statements about the order of the phase transition seems to lie in this ansatz and its interpretation,

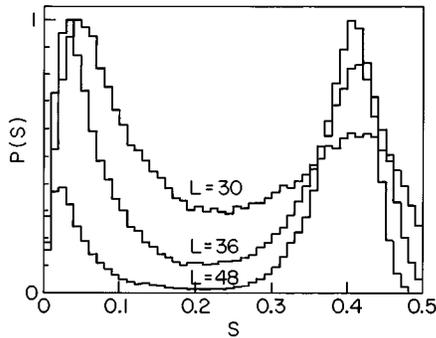


FIG. 3. Double-peak structure of the probability distribution  $P_L(S)$  of the order parameter  $S$  for  $L=30$  at  $\beta=0.36675$ ,  $L=36$  at  $\beta=0.367$ , and  $L=48$  at  $\beta=0.367025$ . From Gavai *et al.* (1989).

$$C(z) \rightarrow A e^{-mz} + B. \quad (3.45)$$

In the critical region it should probably be replaced by an ansatz like Eq. (3.34) to account for both correlation lengths  $\xi^{(p)}$  and  $\xi^{(l)}$ . If  $\xi^{(l)}$  dominates, the mass  $m$  in Eq. (3.45) reflects the increase of  $\xi^{(l)}$  rather than  $\xi^{(p)}$ .

A systematic finite-size study of the  $3d$  3-state Potts model was performed by Gavai *et al.* (1989). The Hamiltonian of the Potts model was specified in Sec. II. It is equivalent to a  $Z(3)$  spin model. The volume dependence was studied for bulk quantities like the average action  $E$  and the order parameter  $S$ . When  $S$  and  $E$  were plotted as a function of  $\beta$ , the crossover region became sharper with increasing volume. A typical manifestation of coexisting phases was verified, and the function  $\beta(L, E)$  was multivalued. One coupling (temperature) was compatible with several values of the free energy at the transition point.

Flip-flop signals in the order parameter were observed as a function of the Monte Carlo time. They are reflected in a double-peak structure of the order-parameter distribution  $P_L(S)$ . Figure 3 verifies what we expect from the general considerations of Sec. II. With increasing volume the double-peak structure becomes more pronounced. (In a truly second-order transition it fades away at larger volumes.)

In the critical region the rounding and shifting of thermodynamic quantities was governed by the thermal exponent  $\lambda_1 \equiv y_T = 1/\nu = d/(2-\alpha)$ , where  $y_T = d$  for a first-order transition (see Sec. II). In particular, the critical coupling should be shifted according to

$$\beta_{c,L} - \beta_c \propto L^{-y_T}, \quad (3.46)$$

and the width of the critical region  $\sigma_L$  should shrink as

$$\sigma_L \propto L^{-y_T}. \quad (3.47)$$

The finite-volume critical coupling  $\beta_{c,L}$  and the width  $\sigma_L$  were extracted from a Gaussian fit for the relative population density of both phases in the critical region, and, alternatively, by the midpoint and length of the coupling interval ( $0.367 \leq \beta \leq 0.36703$ ), where flip-flops

were observed on a  $48^3$  lattice. Results for both quantities were compatible with  $y_T = d$ , i.e., with a first-order finite-size scaling behavior.

The careful analysis of finite-volume correlation lengths was anticipated in the previous subsection. An estimate for the inverse physical mass gap leads to  $\xi^{(p)}(\beta_c) \sim 10$  in lattice units. Thus it stays finite at  $T_c$  as it should for a first order transition.

The inclusion of an antiferromagnetic next-to-nearest-neighbor coupling with a relative coupling strength of  $\gamma = -0.2$  does not change the qualitative result of a first-order transition (Gavai and Karsch, 1989). The discontinuity in the order parameter is of the same magnitude as for  $\gamma = 0.0$ , but the latent heat is found to be smaller by a factor of 2.

This result is interesting, as it corrects the naive expectation that a large gap in the order parameter goes along with a large latent heat. It raises the possibility that the deconfinement transition in the  $SU(3)$  gauge theory is associated with a small latent heat due to slightly nonlocal effective couplings, even if the jump in the order parameter is not small. For the actual size of the latent heat in the  $SU(3)$  gauge theory, see Sec. III.B.2.

Finally we come to the  $SU(3)$  gauge theory as considered by Fukugita *et al.* (1989). Their lattice size ranged over  $8^3 \times 4$  to  $36^3 \times 4$  sites. Two-state signals were observed for the size  $N_\sigma = 8 - 28$ . The distinction between the two-states became conspicuous with increasing volume. Likewise the double-peak structure in the order parameter distribution, where  $P(|\Omega|)$  ( $\Omega$  was the Wilson line operator of Eq. (3.41) was more pronounced for a larger lattice size.

The susceptibility of the Wilson line was given as

$$\chi = V[\langle (\text{Re}\Omega)^2 \rangle - \langle \text{Re}\Omega \rangle^2]. \quad (3.48)$$

Here  $\Omega$  was defined as in Eq. (3.41), where  $\text{Re}\Omega$  is taken as the projection of  $(\sum_n \Omega_n / V)$  onto the nearest  $Z(3)$  axis on the complex  $\Omega$  plane. The results for the maximum in the susceptibility led to a scaling of  $\chi_{\max} \propto V^{\gamma/(d \cdot \nu)}$  with  $\gamma/(d \cdot \nu) = 0.99(6)$  and a width  $\Delta\beta \propto V^{-1/(d \cdot \nu)}$  with  $1/(d \cdot \nu) = 0.95(5)$ , in reasonable agreement with the expected values of 1.

The ‘‘unphysical’’ mass gap  $m^{(\text{unp})}$  defined in Eq. (3.33) was obtained from a fit of  $C_0(z) = A \cosh[m^{(\text{unp})}(z - N_\sigma/2)]$ . The expected behavior in the infinite-volume limit is shown in Table IV. For a first order transition in a *finite* volume,  $m^{(\text{unp})}$  should decrease more sharper as  $V \rightarrow \infty$  than for a second-order transition. In fact, a crossing of the family of functions  $m^{(\text{unp})}(L, \beta)$  was observed, which is qualitatively in agreement with the expectations, although a detailed quantitative understanding is missing.

The physical mass gap  $m^{(p)}$  was read off from an ansatz as Eq. (3.30a) with  $B = |\langle \Omega \rangle|^2$ , where the exponential was replaced by a cosh to account for the periodic boundary conditions. It is small, but stays finite at  $\beta_c$  (see Table IV).

Fukugita *et al.* (1989; Fukugita, Okawa, and Ukawa, 1990) also addressed the question of nonlocal couplings

on an effective level of the SU(3) gauge theory. An answer requires explicit construction of the effective Z(3) action from the original SU(3) gauge theory. The Z(3) variables are not introduced *ad hoc* (i.e., by assumption), but expressed in terms of the original field variables. An “effective” Z(3) spin  $\omega_n$  is defined as the projection of  $\Omega_n$  to the nearest Z(3) axis (normalized to unit length). The generic form of an effective action in terms of  $\omega_n$  variables is written as

$$S_{\text{eff}} = \sum_{\alpha} \beta_{\alpha} O_{\alpha}(\{\omega_n\}). \quad (3.49)$$

Here  $O_{\alpha}$  are Z(3)-invariant monomials in  $\omega_n$ ,  $\beta_{\alpha}$  are the associated couplings, local as well as nonlocal, e.g., NNN couplings. First correlation functions of the monomials  $O_{\alpha}$  are calculated from the ensemble of  $\omega_n$ 's, generated in the original SU(3) Monte Carlo runs. The number of terms labeled by  $\alpha$  is truncated by hand. Next the couplings  $\beta_{\alpha}$  are determined with the aid of Schwinger-Dyson equations. This method is known from Monte Carlo renormalization-group calculations (Gonzalez-Arroyo and Okawa, 1987).

The results are as follows. Effective two-spin couplings decreased to values less than  $10^{-3}$  for distances larger than 3 between the two spins (on a  $24^3 \times 4$  lattice with  $\beta = 5.6925$ ). Three- and four-spin couplings were smaller than  $(1-2) \times 10^{-3}$ . The effective couplings decreased exponentially with distance, as anticipated by Svetitsky and Yaffe from the phase structure of the associated *three-dimensional* SU(3) gauge theory.

The two-spin couplings (in particular the NNN couplings), which were derived quantities in this case, turned out to be all positive. There is no sign of “antiferromagnetism.”

Like other numerical investigations this derivation is not rigorous. In particular, it involves a truncation in the effective action, which is, however, unavoidable. In this sense none of the Monte Carlo results actually *prove* the first order of the deconfinement transition, but different derivations give strong hints pointing in the same direction: the deconfinement transition in the pure SU(3) gauge theory is of first order. The controversy about this topic has been settled. The situation becomes considerably more involved when fermions are included.

## 2. Thermodynamics on the lattice

An important contribution of lattice QCD to phenomenological applications of the QCD transition is a prediction of QCD's equation of state. The behavior of the pressure  $p$ , the internal energy density  $\epsilon$ , and the entropy density  $s$  are of most interest for seeking observable effects in heavy-ion collisions. In this section we summarize the lattice expressions of thermodynamic quantities and outline an approach due to Engel, Fingberg, *et al.* (1990) for deriving the equation of state in an entirely nonperturbative framework. The method overcomes former inconsistencies observed for the pressure. Here we consider its application to the pure SU(N)

gauge theory followed by a discussion of finite-size effects in the regions of critical and high temperature.

From Sec. II.A we recall the continuum expressions for the energy density  $\epsilon$  and the pressure  $p$  as  $\epsilon = (T^2/V) \partial/\partial T \ln Z$  and  $p = T(\partial/\partial V) \ln Z$ . One way of calculating these quantities on the lattice is to transcribe the derivatives with respect to  $T$  and  $V$  on the lattice. This has led to introduction of different lattice spacings for the temporal ( $a_{\beta}$ ) and spatial ( $a_{\sigma}$ ) directions and is the reason why we have specified the lattice action directly on an asymmetric lattice. The derivatives are replaced according to

$$\begin{aligned} \frac{\partial}{\partial T} &\rightarrow \frac{1}{N_{\tau}} \frac{\partial}{\partial a_{\tau}} \\ \frac{\partial}{\partial V} &\rightarrow \frac{1}{3a_{\sigma}^2 N_{\sigma}^3} \frac{\partial}{\partial a_{\sigma}}. \end{aligned} \quad (3.50)$$

After performing the appropriate lattice derivatives of  $\ln Z$ , we set the lattice spacings equal again,  $a_{\sigma} = a_{\beta} = a$ .

Let us consider the first derivatives of the partition function. The order parameter for the pure gauge theory is the thermal Wilson line. Its lattice expression has been given in Eqs. (3.41) and (3.42). The order parameter for the chiral condensate will be given in Eqs. (3.112a) and (3.112b) in Sec. III.C.

The lattice expression for the internal energy density gets a contribution from the gluonic part and the fermionic part,

$$\epsilon = \epsilon_G + \epsilon_F. \quad (3.51)$$

For completeness we give also the formula for the fermionic part. The continuum expression

$$\epsilon = -\frac{1}{V} \frac{\partial(\ln Z)}{\partial(1/T)} = \frac{1}{V} \left\langle \frac{\partial S}{\partial(1/T)} \right\rangle \quad (3.52)$$

is transcribed to

$$\epsilon = \frac{1}{N_{\sigma}^3 a_{\sigma}^3 N_{\tau}} \left\langle \frac{\partial S}{\partial a_{\tau}} \right\rangle \quad (3.53)$$

with  $S$  given by the Wilson action or the action for staggered fermions (see Sec. III.A).

The energy density contains a contribution from the vacuum, which is a  $\beta$ -independent infinite constant (see Bernard, 1974). When the vacuum contributions from the gluons  $\epsilon_{G,\text{vac}}$  and the fermions  $\epsilon_{F,\text{vac}}$  at  $T=0$  are subtracted, the resulting lattice expressions are

$$\begin{aligned} \epsilon_G &= 2N(N_{\sigma}^3 N_{\tau} a^4) \left[ g^{-2} \left\langle \sum_{\mu < \nu < 4}^x P_x^{\mu\nu} - \sum_{\mu < 4}^x P_x^{\mu 4} \right\rangle \right. \\ &\quad \left. - c'_{\sigma} \left\langle \sum_{\mu < \nu < 4}^x P_x^{\mu\nu} \right\rangle - c'_{\tau} \left\langle \sum_{\mu < 4}^x P_x^{\mu 4} \right\rangle \right] - \epsilon_{G,\text{vac}} \end{aligned} \quad (3.54)$$

and

$$\epsilon_F = N_f (4N_{\sigma}^3 N_{\tau} a^4)^{-1} \langle \text{Tr} \{ D^{(4)} (D + m)^{-1} \} \rangle - \epsilon_{F,\text{vac}},$$

where  $m_f$  has been set to  $m$  for all flavors  $f$  in Eq. (3.54) and

$$a = a_\sigma = a_\tau, \quad D = \sum_{\mu=1}^4 D^{(\mu)},$$

$$c'_\sigma = -a \partial g_\sigma^{-2} / \partial a_\tau |_{a_\tau = a_\sigma}, \quad c'_\beta = -a \partial g_\tau^{-2} / \partial a_\tau |_{a_\tau = a_\sigma},$$

$$\epsilon_{G,\text{vac}} = N(N_\sigma^4 a^4)^{-1} [-c'_\sigma - c'_\tau] \sum_{\substack{x \\ \mu < \nu}} P_x^{\mu\nu},$$

$$\epsilon_{F,\text{vac}} = 3N_f (16a^4)^{-1}. \quad (3.55)$$

The notation has already been introduced in Sec. III.A. Note that the energy densities scale with  $(1/a^4)$ . An uncertainty in  $a(g)$  in units of [MeV] is amplified to the fourth power in the results for  $\epsilon$ . Further, we see that  $\epsilon$  is expressed in certain plaquette expectation values, which are easily accessible in a Monte Carlo simulation. In principle the coefficients  $c'_\sigma$  and  $c'_\tau$  can be determined nonperturbatively from lattice simulations (for first attempts see Blum *et al.*, 1995a), but in the past the perturbative values (Karsch, 1982; Trincherò, 1983) were used in the otherwise nonperturbative expressions. The use of perturbative values is questionable from the beginning. For the SU(3) gauge theory, large deviations from the perturbative  $\beta$  function are known to occur for  $g \geq 1$ , and the  $\beta$  function  $a dg/da$  is related to the coefficients  $c'_\sigma$  and  $c'_\tau$  via

$$a \frac{dg}{da} = g^3 \left( \frac{\partial g_\sigma^{-2}}{\partial \xi} + \frac{\partial g_\tau^{-2}}{\partial \xi} \right)_{\xi=1} \quad (3.56)$$

with  $\xi \equiv a_\sigma/a_\tau$ , (see Sec. III.A). Therefore significant nonperturbative contributions must be expected for  $c'_\sigma$  and  $c'_\tau$  as well.

The perturbative ingredients have actually led to inconsistencies in measurements of the pressure. The pressure turned out to be negative and discontinuous across the transition point. In consequence the values for the latent heat were not consistent. For a continuous pressure the latent heat  $\Delta L$  should be the same, whether it is calculated as

$$\Delta(\epsilon+p)/T_c^4 \equiv \Delta L_1 \quad (3.57)$$

or

$$\Delta(\epsilon-3p)/T_c^4 \equiv \Delta L_2.$$

Both Eqs. (3.57) are simple expressions in terms of plaquette expectation values. Fukugita, Okawa, and Ukawa (1990) found for an SU(3) gauge theory

$$\Delta L_1 = 2.54(12) \quad \text{and} \quad \Delta L_2 = 3.78(20). \quad (3.58)$$

The discrepancy between  $\Delta L_1$  and  $\Delta L_2$  is clearly outside the statistical error bars. In an independent simulation, Deng (1989) found

$$\Delta p/T_c^4 = -0.29(15) \quad (3.59)$$

for the discontinuity in the pressure.

These difficulties have been overcome by a different approach to lattice thermodynamics, which can be per-

formed entirely within a nonperturbative framework. The additional quantity now is the  $\beta$  function  $a dg/da$  from QCD, which is known in a nonperturbative region of couplings  $g$  from Monte Carlo renormalization group calculations. Otherwise only plaquette expectation values have to be calculated, as before. This approach was proposed by Engels, Fingberg, Karsch, *et al.* (1990). The steps are as follows.

(i) The free-energy density is calculated from an integration over its derivative with respect to  $\beta$ , since the logarithm of the partition function is not directly accessible within the Monte Carlo approach. With  $f = -T/V \ln Z$  one takes the derivative with respect to the gauge coupling  $\beta = 2N/g^2$

$$-\frac{\partial \ln Z}{\partial \beta} = \langle S_G \rangle = 6N_\sigma^3 N_\tau P_T, \quad (3.60)$$

where  $S_G$  is the gauge part of the action [e.g., given by Eq. (3.10)] and  $P_T$  is a short notation of the plaquette expectation value at temperature  $T$ , calculated on a lattice of size  $N_\sigma^3 N_\tau$ . If  $P_0$  denotes the corresponding zero-temperature expectation value, evaluated on a lattice of size  $N_\sigma^4$ , the difference of free-energy densities  $f$  at coupling  $\beta$  and  $\beta_1$  is

$$\frac{f}{T^4} \Big|_{\beta_1}^{\beta} = -6N_\tau^4 \int_{\beta_1}^{\beta} d\beta' [P_0 - P_T]. \quad (3.61)$$

(ii) Now an additional assumption enters, which is strictly satisfied only for an infinitely large lattice. It is the relation  $\ln Z = V \partial \ln Z / \partial V$ , which is valid for homogeneous systems in large volumes. From  $p = T(\partial/\partial V) \ln Z$  and Eq. (3.61) one obtains

$$p(\beta) = -[f(\beta) - f(\beta_1)]. \quad (3.62)$$

In Eq. (3.62)  $\beta_1$  has been chosen small enough that  $p(\beta_1) \sim 0$ .

(iii) A second quantity, which has appeared in Eq. (3.57) and is easily calculable on the lattice, is the interaction measure  $\Delta$

$$\Delta = \frac{\epsilon - 3p}{T^4} = -12NN_\tau^4 a \frac{dg^{-2}}{da} [P_0 - P_T]. \quad (3.63)$$

As the name suggests,  $\Delta$  vanishes for an ideal gas (where  $\epsilon = 3p$ ; see Sec. V).

(iv) The advantage of calculating  $\Delta$  first rather than  $\epsilon$  is that  $\Delta$  no longer depends on  $c'_\sigma$  and  $c'_\tau$ , but on the QCD  $\beta$  function  $B(g)$ ,

$$B(g) = -a \frac{dg^{-2}}{da}. \quad (3.64)$$

To evaluate  $\Delta$  in a nonperturbative coupling regime requires a knowledge of the  $\beta$  function for nonperturbative couplings as well. This function has been inferred from Monte Carlo renormalization-group studies (Hoek, 1990; Akemi *et al.*, 1993). The functional form suggested in by Hoek (1990) was used by Laermann *et al.* (1995) in fits to the more recent data from Akemi *et al.* (1993). The remaining terms in  $\Delta$  and  $f$  are

calculated in a Monte Carlo simulation. Once we have  $f$  and  $\Delta$ , we know  $\epsilon$  and the entropy density  $s$  (with  $f = \epsilon - Ts$ ). Using this approach, we find that the previously mentioned problems with the pressure have disappeared;  $p$  stays positive and behaves continuously at  $T_c$ .

A fully nonperturbative approach alone does not guarantee that the equation of state on the lattice will be relevant for continuum physics. Earlier simulations have been performed for a time extension of four lattice spacings, although  $N_\tau \sim 10$  is needed for the continuum limit.

The  $N_\tau$  dependence was been studied in the case of the SU(3) deconfinement transition by Boyd *et al.* (1995). The temporal lattice size was varied from ( $N_\tau=4$  with  $N_\sigma=16$ ) to ( $N_\tau=6$  with  $N_\sigma=32$ ) and ( $N_\tau=8$  with  $N_\sigma=32$ ). The dominant type of finite-size effect depends on the temperature. At high temperatures there is an intimate relation between finite-size (IR) artifacts and finite-cutoff (UV) artifacts, although the distinction between them may naively suggest their decoupling. The relation between IR and UV artifacts may be seen as follows. A reduction of  $a_{\tau,\sigma}(g)$  (fm) as a function of  $g$  goes along with an increase in temperature  $T$  (MeV), if  $N_\tau$  is kept fixed. In this way continuum behavior is mixed with high-temperature behavior. For a small lattice spacing, only the high-momentum modes are cut off which give the main contribution to the energy density and pressure at high temperatures. For a larger lattice spacing at lower temperatures, lower-momentum modes are also cut off, but the energy density and pressure grow approximately with  $T^4$ . Thus the finite-cutoff effects, induced by the finite number of time slices, are largest for high temperatures.

The  $N_\tau=1/(aT)$  dependence of the gluonic part of the energy density of an ideal gluon gas is estimated as (Karsch, 1995)

$$\epsilon_{\text{gluonic}}(N_\tau) = (N_c^2 - 1) \left[ \frac{\pi^2}{15} + \frac{2\pi^4}{63} \frac{1}{(N_\tau)^2} + O\left(\frac{1}{N_\tau^4}\right) \right]. \quad (3.65)$$

For  $N_\tau=4$ ,  $\epsilon_{\text{gluonic}}(4)$  is 50% larger than the corresponding Stefan-Boltzmann value in the continuum. The leading  $1/N_\tau^2$  dependence has been used by Karsch *et al.* to extrapolate  $N_\tau=6, 8$  data for  $\epsilon$  and  $p$  to the continuum limit. Figures 4(a) and 4(b) show the energy density and pressure for the various lattice sizes, including the expected increase of finite-size/finite-cutoff effects with temperature. There are strong deviations by about 20% from ideal-gas behavior (indicated by the horizontal lines) for temperatures at least up to  $3T_c$  in the continuum limit. It should further be noticed that the pressure  $p$  increases more slowly with  $T$  than the energy density  $\epsilon$ . This leads to a strong deviation of the sound velocity from the ideal-gas value. The velocity of sound can be an input from first principles to applications in heavy-ion collisions; see Sec. V.

The latent heat is indicated by the vertical bar in Fig. 4(a); the value is taken from Iwasaki *et al.* (1992). It is further reduced compared to earlier results (e.g., of Brown *et al.*, 1988). The result is

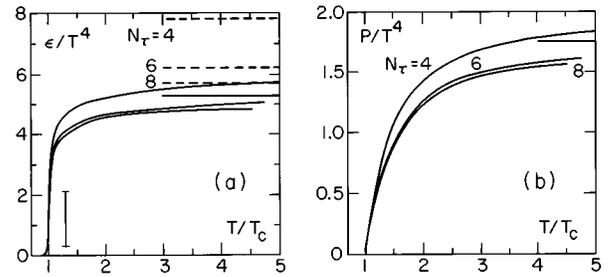


FIG. 4. Energy density (a) and pressure (b) in an SU(3) pure gauge theory on  $16^3 \times 4$ ,  $32^3 \times 6$  and  $32^3 \times 8$  lattices. Ordinate quantities are divided by  $T^4$ , rendering them dimensionless. The dashed horizontal lines indicate the corresponding results for an ideal gas on lattices of the same size. The solid horizontal line shows the Stefan-Boltzmann result in the continuum limit. From Karsch (1995).

$$\frac{\Delta \epsilon}{T_c^4} = \begin{cases} 2.44 \pm 0.24 & N_\tau = 4, \\ 1.80 \pm 0.18 & N_\tau = 6. \end{cases} \quad (3.66)$$

The latent heat is only 30% of the energy of an ideal gas at  $T_c$ .

The interaction measure  $\Delta$  is shown in Fig. 5 for three lattices. The deviation from ideal-gas behavior is largest in the transition region, going logarithmically to zero with increasing temperature. The interaction measure is obviously less sensitive to finite-size/finite cutoff effects than  $\epsilon$  and  $p$  separately. The deviations from the ideal gas behavior are assumed to be due to the nonperturbative infrared structure of QCD, which is less sensitive to UV effects. The results of Fig. 5 support this assumption and indicate that QCD's IR structure plays an important role even at high temperatures. As mentioned in the introduction (Sec. I), it is an oversimplification from asymptotic freedom that QCD behaves fully perturbatively at high temperatures.

In physical units the result for the interaction measure is given by (Karsch, 1995)

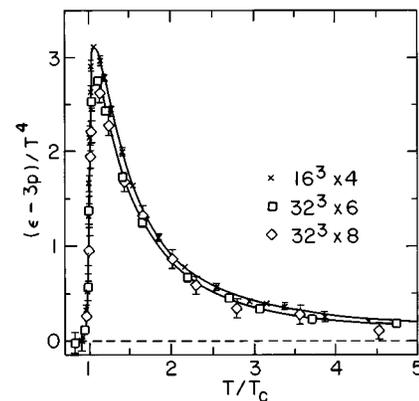


FIG. 5. Interaction measure  $\epsilon - 3p$  normalized to  $T^4$  (dimensionless units) for a pure SU(3) gauge theory for different lattice sizes. From Karsch (1995).

$$\epsilon - 3p \approx 2.4 \text{ GeV/fm}^3 \quad \text{at } T = 1.1T_c. \quad (3.67)$$

In Eq. (3.67) the string tension  $\sigma$  with  $\sqrt{\sigma} \sim 420 \text{ MeV}$  has been used as physical input. Thus the main contribution to the interaction measure comes from the zero temperature gluon condensate  $\langle G^2 \rangle_0 = 2.0 \text{ GeV/fm}^3$ , noting that the trace anomaly  $(\epsilon - 3p)$  is related to the gluon condensates at zero and finite temperature according to

$$\epsilon - 3p = \langle G^2 \rangle_0 - \langle G^2 \rangle_T. \quad (3.68)$$

As argued above, UV artifacts induced by finite temporal extension play a dominant role in the high-temperature region but a minor role in the transition region. Their main contributions to finite-size effects in the transition region come from the low-momentum modes, in particular if a correlation length becomes large. In case of the pure SU(2) gauge theory, the standard finite-size scaling analysis for second order phase transitions can be applied to the nonanalytic behavior of the free-energy density. The finite-size effects are controlled by the ratio  $N_\sigma/N_\tau = TV^{1/3}$  rather than by  $1/N_\tau$  as in the high-temperature region. In particular, the normalized critical energy density  $\epsilon^{\text{crit}}/T_c^4$  should scale as a function of  $TV^{1/3}$  with critical indices of the three-dimensional Ising model (cf. Sec. II.B.2). This scaling behavior has been verified by Engels *et al.* (1995). An extrapolation to the infinite-volume limit leads to

$$\epsilon^{\text{crit}}/T_c^4 = 0.256(23) \quad (3.69)$$

for a pure SU(2) gauge theory. This result may be considered as free of lattice artifacts. It is a continuum result, although not continuum physics, because of two colors and the absence of dynamical fermions.

To summarize, reliable data for QCD's equation of state exist for the limiting cases of a pure SU(2) and SU(3) gauge theory. Earlier discrepancies in results on the pressure have been removed by using fully nonperturbative ingredients in the derivation. To test the relevance of the lattice results for continuum physics, a careful finite-size scaling analysis has been performed. The finite-size effects are under control, as the predicted dependence on  $N_\tau$  and  $N_\sigma$  has been confirmed in the simulations. Therefore an extrapolation to the continuum limit [ $N_\tau \rightarrow \infty$ ,  $a \cdot T \rightarrow 0$ ,  $T$  (MeV) fixed] is safe. Such tests should become standard in future lattice simulations.

The fully perturbative equation of state can also be applied when dynamical fermions are included. We report on first attempts in this direction in Sec. III.C.6. However, a finite-size scaling analysis gets much more involved in the presence of dynamical fermions. For sufficiently large quark masses, the singularities of the first-order SU(3) or the second-order SU(2) deconfinement transitions will be rounded anyway. One then has to disentangle which part of the rounding comes from finite volume and which from finite masses (see Secs. III.C.2 and III.C.3).

### 3. Interface tensions in QCD

In recent years the interface (or surface) tension has been frequently calculated for models of QCD. Like the

latent heat, the surface tension is an important measure of the strength of a first-order transition. The interface occurs between phases, which coexist at the critical temperature. A large surface tension leads to a strong supercooling effect. The onset of the phase conversion is delayed as the critical temperature is reached. One possible scenario for the phase conversion is droplet formation. Once the system has sufficiently supercooled below the critical temperature, the gain in energy from the conversion to the new phase can compensate for the energy loss in interface free energy, and the phase conversion sets in.

In applications to QCD, different kinds of interfaces must be distinguished. In a first-order chiral transition we have interfaces between phases with broken and restored chiral symmetry. Interfaces between the deconfinement and the confinement phases may occur in a first-order deconfinement transition. They will be considered below for calculations of interface tensions. Furthermore different realizations of the plasma phase are separated by interfaces. Such interfaces occur not only in the region around  $T_c$ , but throughout the phase of broken  $Z(3)$  symmetry above  $T_c$ . One specific phase realization corresponds to a spatial domain characterized by a certain expectation value of the Wilson line operator. The associated interface free-energy density is called an *ordered/ordered interface tension*  $\sigma_{oo}$ , whereas an interface between the plasma and the hadronic phase leads to an *ordered/disordered interface tension*  $\sigma_{od}$ . We write “plasma phase” as a common name for the deconfinement and/or chiral symmetric phase and “hadronic phase” for the confinement and/or chiral symmetry-broken phase. Via the relative magnitude of both quantities at  $T_c$  one may gain some insight about the degree of wetting.

*Wetting* is an alternative or competing mechanism to droplet formation. One phase may spread along a two-dimensional front into the other. As stated in a liquid/gas picture, a droplet of liquid may wet an interface between a gas and a solid when these phases coexist. Whether the wetting is complete or not depends on the relative size of  $\sigma_{oo}$  and  $\sigma_{od}$ .

In QCD the issue of wetting was raised by Frei and Patkós (1989). It was studied in systems that share the global  $Z(3)$  symmetry, in the three-state Potts model, and in an effective model of Polyakov loops. Later the question of wetting at the deconfinement transition was investigated in an effective three-dimensional  $Z(3)$ -symmetric  $\Phi^4$  model for the Polyakov loop (Trapenberg and Wiese, 1992). In this model the confinement phase completely wets the different realizations of the deconfinement phase at  $T_c$ .

So far these results are not applicable to QCD transitions under realistic conditions, since fermions have been excluded. Recall that the inclusion of fermion masses may completely wash out the first-order nature of the chiral and the deconfinement transition. A smooth crossover phenomenon prevents the occurrence of different *coexisting* phases and their associated interfaces. If the hints of today are confirmed in the future,

that both transitions fade away for physical fermion masses, the various interface tensions have no impact on applications in the early universe or relativistic heavy-ion collisions.

#### a. How to measure an interface tension

Consider a system in which a domain with volume  $V_H$  in the (hadronic) phase  $H$  is embedded in a second phase  $Q$  (the quark-gluon plasma phase) with volume  $V_Q$ . If we denote by  $f_h$  and  $f_Q$  the free-energy densities of both phases, the free energy  $F$  of the total system is given as

$$F = F_s + (V_H f_H + V_Q f_Q), \quad (3.70)$$

where  $F$  differs from the sum of the domain free energies just by an amount  $F_s$  and  $F_s$  equals the free energy associated with the interface separating both domains.

As long as  $V_H$  and  $V_Q$  are finite, and a homogeneous external field  $\beta$  is applied,  $F_s$  will actually be zero. Thermal equilibrium forbids the coexistence of phases as stable configurations in a finite volume. The system will tunnel from one phase into the other. A transient coexistence can only result from metastability effects. If one wants to measure the surface tension directly as the excess free energy due to an interface in a finite system, one has to stabilize the interface by an external field gradient  $\Delta\beta$ . Again the order of limits is essential,

$$\lim_{\Delta\beta \rightarrow 0} \lim_{V_H, V_Q \rightarrow \infty} F_s/A = \sigma, \quad (3.71)$$

where  $\sigma$  is the surface free energy  $F_s$  per unit interfacial area  $A$  and  $\Delta\beta$  stands for a generic field gradient.

For temperature-driven transitions in QCD an appropriate “field” gradient is provided by the temperature. To have a preferred direction, we consider a spacetime lattice with cylindrical spatial geometry, i.e.,  $L_x \times L_y \times L_z \times N_\tau$  with  $L_x = L_y \ll L_z$ . Choosing half of the lattice links ( $z=1, 2, \dots, L_z/2$ ) at a temperature larger than  $T_c$  (corresponding to a coupling  $\beta \equiv 6/g^2 > \beta_c$ ) and the other half ( $z=L_z/2+1, \dots, L_z$ ) below  $T_c$  ( $\beta < \beta_c$ ), we introduce into the system an interface between the deconfinement and confinement phases. It will be located somewhere “between”  $z=L_z/2$  and  $L_z/2+1$ . Early calculations of the surface tension in an SU(3) gauge theory were performed along these lines (Kajantie *et al.*, 1990). The procedure involved an extrapolation  $\Delta\beta \rightarrow 0$  at the very end. The result, obtained for a time extension of  $N_\tau=2$ , was  $\sigma/T_c^3 = 0.24 \pm 0.06$ . This value is probably too large by an order of magnitude (see below), first because of possible finite-size effects due to the small extension in the time direction, and second because the extrapolation to  $\Delta\beta \rightarrow 0$  may be not safe. An external field can lead to suppression of fluctuations of the interface, if it is strong enough, causing the interface to look more rigid than it would be in the  $\Delta\beta \rightarrow 0$  limit.

Most of the work to date on interface calculations in SU(3) gauge theory has concentrated on time extensions of  $N_\tau=2$ , as the measurements are rather time consuming (Potvin and Rebbi, 1989; Huang *et al.* 1990; Kajantie

*et al.*, 1991; Grossmann *et al.*, 1992; Janke, 1992). Results for  $N_\tau=4$  can be found in Brower *et al.* (1992); Potvin and Rebbi (1991); Grossmann and Laursen (1993). In the following we report on work by Iwasaki *et al.* (1994), who extended the number of time slices up to  $N_\tau=6$ .

Iwasaki *et al.* applied the *histogram method*, introduced by Binder (1981 and 1982) for Ising-type systems, to an SU(3) gauge theory. The histogram method is based on the analysis of probability distributions  $P(\Omega)$  of order parameters  $\Omega$  in the vicinity of the phase transition. As we have seen in Sec. II.A, the probability distribution  $P(\Omega)$  develops a characteristic double-peak structure close to  $T_c$ . The structure grows more pronounced as the volume is increased. This behavior was identified as a characteristic signature for a first-order transition in a finite volume (see Fig. 3). The valley between the peaks corresponds to configurations with interfaces that are more suppressed the larger the volume.

As the suppression comes from the extra costs in interface free energy, it is plausible that the value of  $\sigma$  may be inferred from the position of the maxima and minimum of  $P(\Omega)$  as a function of volume. In the formulas below periodic boundary conditions are assumed, with a cylinder geometry satisfying ( $N_x, N_y \ll N_z$ ). Thus two interfaces will be created in the  $(x, y)$  plane with a total area of  $2A = 2N_x N_y a^2$ , where  $a$  is the lattice constant. The order parameter  $\Omega$  will be identified with the Polyakov loop.

The ansatz for the probability distribution  $P(\Omega)$ , which was chosen by Iwasaki *et al.* (1994), is then given as

$$P(\Omega) = P_1(\Omega) + P_2(\Omega) + P_m(\Omega)$$

with

$$P_i(\Omega) = c_i \exp(-f_i V/T) \exp[-(\Omega - \Omega_i)^2/d_i^2] \quad (i=1,2) \quad (3.72)$$

$$P_m(\Omega) = c_m \exp[-(f_1 V_1 + f_2 V_2)/T - \sigma 2A/T],$$

where the following notations are involved. The order parameters  $\Omega_1$  and  $\Omega_2$  denote the values in the confinement and deconfinement phases,  $f_1$  and  $f_2$  the corresponding free-energy densities, and  $V_1$  and  $V_2$  are the volumes occupied by each of the coexisting phases. The coefficients  $c_i$  and  $d_i$  depend on the volume. In Eq. (3.72) for  $P(\Omega)$  we recognize the superposition of the Gaussians around the characteristic peaks for each phase. The third term  $P_m(\Omega)$  gives the probability for finding the mixed phase. Here is the place where the interface tension enters. The probability  $P_m(\Omega)$  depends on  $\Omega$  via the volumes  $V_1$  and  $V_2$ , which are occupied by both phases such that  $\Omega V = \Omega_1 V_1 + \Omega_2 V_2$ .

The weight factors  $\exp(-f_i V/T)$  ( $i=1,2$ ) are proportional to the probabilities for the system to reside in the confinement or deconfinement phase. While  $f_1$  and  $f_2$  are degenerate at  $T_c$  in the infinite-volume limit, they are in general different in a finite volume. To determine  $\sigma$  in a finite volume  $V$ , one has to define  $\hat{\sigma}_V \equiv \sigma_V/T_c^3$ , in

such a way that the leading  $V_i$  dependence in the exponent cancels out. The cancellation is achieved if  $\hat{\sigma}_V$  is defined according to

$$\hat{\sigma}_V \equiv \frac{-N_\tau^2}{2N_x N_y} \log \frac{P_{\min}}{(p_{\max,1})^{\gamma_1} (p_{\max,2})^{\gamma_2}}. \quad (3.73)$$

Here  $p_{\max,1}$  and  $p_{\max,2}$  are the two maxima of  $P(\Omega)$ , while  $p_{\min}$  denotes the minimum between the two peaks. The powers  $\gamma_1$  and  $\gamma_2$  denote the weights of the contributions of  $\Omega_1$  and  $\Omega_2$  to  $\Omega$  at the minimum, i.e.,  $\Omega = \gamma_1 \Omega_1 + \gamma_2 \Omega_2$ . The infinite-volume limit of  $\hat{\sigma}_V$  is the value for the interface tension in units of  $T_c$ ,

$$\hat{\sigma} \equiv \sigma/T_c^3 = \lim_{V \rightarrow \infty} \hat{\sigma}_V. \quad (3.74)$$

The actual measurement of  $\hat{\sigma}_V$  is more subtle. The subtleties concern the choice of order parameter, the critical coupling in a finite volume, and the determination of  $p_{\max}$  and  $p_{\min}$ . Candidates for order parameters are the action density or the Polyakov loop. The Polyakov loop is a complex-valued observable. The distribution  $P(\Omega)$  develops four peaks near  $T_c$  corresponding to the confinement phase with  $\Omega_i = 0$  and three realizations of the deconfinement phase in the directions  $\exp(i2\pi n/3)$ ,  $n = 0, 1, 2$ . To reduce the numerical effort of obtaining high quality data, Iwasaki *et al.* have projected  $\Omega$  on the real axis by taking the absolute value  $\Omega_{\text{abs}}$  or by rotating it with  $\exp(i2\pi n/3)$  so that  $-\pi/3 < \arg \Omega \leq \pi/3$  and then taking the real part, leading to  $\Omega_{\text{rot}}$ . The dependence of the results on this choice was checked.

The choice of the critical coupling in a finite volume was a nontrivial issue. It should guarantee the correct infinite-volume limit of  $\beta_c$ . Here Iwasaki *et al.* adjusted  $\beta_c$  so that the peaks of the histograms for  $\Omega_{\text{rot}}$  had equal height.

Because the actual minimum and maxima of the histograms might be contaminated due to statistical fluctuations, the extrema were read off from third-order polynomial fits to the histograms in the vicinity of the extrema.

Only the leading volume dependence drops out of  $\hat{\sigma}_V$ , when it is calculated according to Eq. (3.73). Sub-leading corrections in  $1/V$  arise, for example, from fluctuations of  $\Omega$  in the bulk phases, capillary wave fluctuations of the interfaces, and zero modes corresponding to the translation of interfaces in the direction perpendicular to the interfaces. If the geometry of the lattice deviates from an idealized cylinder in the  $z$  direction, interfaces are no longer restricted to the  $xy$  plane, but sweep out in other directions as well. All of these finite-volume effects were taken into account by making an appropriate ansatz for the volume-dependent prefactors in the formula for  $p_{\min}/p_{\max}$ . For further details we refer the readers to Iwasaki *et al.* (1993).

An important point about the actual measurement of  $\sigma$  concerns the generation of histograms. In SU(3) gauge theory high-statistics histograms were available from the QCDPAX collaboration (Iwasaki *et al.*, 1991; Iwasaki, Kanaya, Yoshii 1992). They were used as in-

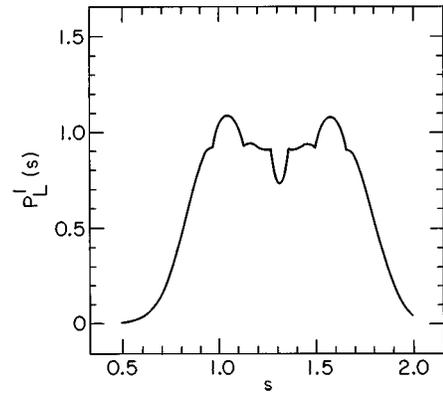


FIG. 6. Auxiliary multicanonical probability distribution as a function of the action density  $s$ , measured in dimensionless lattice units. From Berg and Neuhaus (1991).

put in the  $\sigma$  measurements of Iwasaki *et al.* (1994). If such histograms are not available for the quantities one is interested in, one can make use of a more advanced method, the so-called multicanonical algorithms, to generate such histograms.

#### b. Multicanonical algorithms

We briefly digress from the  $\sigma$  measurement to sketch the idea of a multicanonical updating procedure. This was developed by Berg and Neuhaus (1991) and first applied to two-dimensional  $q$ -state Potts models. For large volumes it is notoriously difficult to generate configurations corresponding to the valley between the two peaks. These configurations are suppressed by the large amount of surface free energy, which is just the quantity of interest. The reason is that usual local upgrading procedures, like the Metropolis or heat-bath algorithms, are designed for single-peaked distributions  $P(\Omega)$ . Configurations are selected with an importance sampling according to their Boltzmann weight  $\exp\{-\beta S\}$ . A precise calculation of the maxima and minimum of the double-peak structure of  $P(\Omega)$  requires frequent tunneling through the minimum, implying exponentially large autocorrelation times  $\tau \propto \exp\{2\sigma L^{d-1}\}$ .

The idea of the so-called multicanonical approach is to generate an auxiliary distribution  $P'(\Omega)$  (see Fig. 6) that does not suffer from the original problems, but that can be related to the canonical distribution in a controllable way. Figure 6 gives an example of the auxiliary distribution as a function of the action density  $s \equiv S/V$ , which is approximately flat between the two maxima. We have

$$P'_L(S) \propto n_L(S) P'_B(S). \quad (3.75)$$

The index  $L$  stands for the linear size  $L$ ,  $n_L(S)$  denotes the spectral density, and  $P'_B(S)$  has the form of a Boltzmann weight factor. In practice one starts with a guess for the parameters in the exponent of  $P'_B(S)$  such that  $P'_L(S)$  is approximately flat. Configurations with the multicanonical probability  $P'_B$  are generated with a generalized Metropolis or heat-bath updating procedure.

The original distribution  $P_L(S)$  [or  $P_L(\Omega)$ ] is then obtained from  $P'_L(S)$  by reweighting (Baumann and Berg, 1985; Ferrenberg and Swendsen, 1988, 1989).

The efficiency of the multicanonical algorithm has been well established (see, for example, Berg and Neuhaus, 1992; Janke, 1992). The algorithm has been generalized to an SU(3) gauge theory by Grossmann *et al.* (1992), where the time extent was restricted to  $N_\tau=2$ . For  $N_\tau=4$  Grossmann and Laursen (1993) used the histograms of the QCDPAX collaboration (Iwasaki *et al.*, 1991; Iwasaki, Kanaya, Yoshié *et al.* 1992) and Fukugita *et al.* (1989). Their  $N_\tau=4$  results for  $\sigma$  are compatible with the results of Iwasaki *et al.* (1994), which we discuss below.

The results of Iwasaki *et al.* (1994) are based on the histograms of the QCDPAX collaboration for lattices of temporal extent  $N_\tau=4$  and 6 and spatial sizes ranging from  $N_x N_y N_z = 12^2 \times 24$ ,  $24^2 \times 36$ ,  $20^3$ ,  $24^3$ , and  $36^2 \times 48$ . An extrapolation to the infinite-volume limit leads to

$$\sigma/T_c^3 = \begin{cases} 0.0292(22) & \text{for } N_\tau=4, \\ 0.0218(33) & \text{for } N_\tau=6. \end{cases} \quad (3.76)$$

The results do not depend on the actual choice of order parameter ( $\Omega_{\text{rot}}$  or  $\Omega_{\text{abs}}$ ). To assess the relevance of these numbers for the continuum limit, one has to check the scaling property, i.e., sending  $N_\tau \rightarrow \infty$  ( $a \rightarrow 0$ ) while keeping  $T_c$  in physical units fixed. It is obvious from Eq. (3.76) that  $\sigma/T_c^3$  violates scaling. The scaling violation is even more obvious if the  $N_\tau=2$  value is included [ $\sigma/T_c^3 = 0.103(7)$ ]. A similar violation of scaling has been observed for the latent heat  $\Delta L$ , when it is measured from the same data of the QCDPAX collaboration (Iwasaki *et al.*, 1991; Iwasaki, Kanaya, Yoshié *et al.*, 1992). Thus it is not surprising that the scaling violation approximately drops out if certain ratios of  $\Delta L$  and  $\sigma$  are considered. Such a ratio is given by

$$\alpha^2 = 16\pi\sigma^3 / (3\Delta L^2 T), \quad (3.77)$$

which enters the amount of supercooling and the average distance between nucleated bubbles in the early universe. Let us see what the results for  $\sigma/T_c^3$  of Iwasaki *et al.* (1994) imply for possible relics of the QCD transition in the early universe.

#### 4. Phenomenological implications for the early universe

The QCD transition happened at  $\sim 10^{-6}$  sec after the big bang, when the universe cooled down to a temperature of  $\sim 150$ – $200$  MeV. In the following we describe one possible scenario for the succeeding evolution, which has been frequently discussed in the literature (e.g., Kajantie and Kurki-Suonio, 1986; Fuller *et al.*, 1988; Applegate *et al.*, 1987; Applegate, 1991; Meyer *et al.*, 1991; Schramm *et al.*, 1992).

If the QCD transition is of first order and proceeds via bubble formation, the nucleation will not set in exactly at  $T_c$ , but start at some lower temperature  $T_i$  at time  $t_i$ . The plasma phase supercools until the gain in free energy due to the phase conversion can compensate the

costs in interface free energy between the coexisting phases. The degree of supercooling cannot be too large, as the cooling of the universe goes adiabatically slowly as viewed from the scale of QCD.

In the vicinity of the QCD transition two very different time scales enter the evolution: the QCD time scale of the order of  $1/T_c \sim 1 \text{ fm}/c \sim 0.33 \times 10^{-23}$  s and the Hubble time of the order of  $1/\chi = 0.36 \times 10^{-4} \text{ s} \sim 10^{19}/T_c$  with  $\chi = \sqrt{8\pi GB/3}$ , where  $G$  is the gravitational constant and  $B$  the vacuum energy density represented by the bag constant. The time dependence of temperature  $T(t)$  follows from Einstein's equations combined with QCD's equation of state [e.g., the bag model equation of state (Kajantie and Kurki-Suonio, 1986)]. It is given by

$$\frac{-dT}{T dt} = \frac{1}{2t} \frac{T^2}{M_{\text{Pl}}} \quad (3.78)$$

with  $M_{\text{Pl}}$  denoting the Planck mass, i.e., the expansion rate is  $10^{-19}$  times slower than typical time scales in QCD.

At time  $t_i$  and temperature  $T_i$  nucleation sets in, and bubbles of critical radius are formed. Their growth rate is slow compared to that of the shock waves they emit. The shock waves propagate with the velocity of sound  $v_{\text{sh}} = 1/\sqrt{3}$ . They reheat the plasma to some higher temperature  $T_i$ . The release of latent heat prevents further nucleation in regions of space that have been affected by shock waves. In such regions only bubbles will grow that have been created before. The creation and growth of bubbles continues, until the shock waves which are emitted from various nucleation centers collide.

At this time ( $t=t_{\text{PT}}$ ), when the entire universe has been affected by shock waves, nucleation shuts off. A new scale becomes important. It is the *average distance*  $R$  between hadronic bubbles at the time when the shock waves collide. This scale is a measure of the inhomogeneities in the baryon number density. During the coexistence of both phases most of the baryon number resides in the plasma phase. Since the hadronic bubbles are nucleated through random thermal or quantum processes, they contain on average no net baryon number. In thermal and chemical equilibrium the chemical potentials  $\mu_Q(B)$ ,  $\mu_h(B)$ , associated with the elementary baryon charge, are equal in both phases ( $h$  and  $q$ ). The ratio of the baryon number density in the quark ( $q$ ) and hadron ( $h$ ) phases,  $n_B(q)/n_B(h)$ , is approximately given by  $\exp(m_N c^2/T_c)$ , where  $m_N$  is the nucleon mass, since the light current quark masses are almost zero.

A competing scale to the scale of baryon number density fluctuations is the *neutron diffusion length*. The diffusion length of neutrons depends on the age of the universe. In the period where the age of the universe is 20  $\mu\text{s}$  to 1 s old, the weak interactions are fast enough to ensure a uniform neutron-to-proton ratio. The ratio of neutrons to protons is given by  $[\exp(m_n - m_p)/k_B T]$ . Protons and neutrons diffuse along with each other. Neutrons really diffuse, protons diffuse by converting into a neutron through the weak interaction, diffusing, and converting back. At the age of  $\sim 1$  sec, when the

temperature falls to  $\sim 1$  MeV, the weak and electromagnetic interactions decouple. The weak interaction drops out of equilibrium. The neutrons remain neutrons, the protons remain protons. The proton diffusion length is very short, because the proton is electrically charged. The diffusion length of neutrons is larger by several orders of magnitude. If the baryon number density fluctuations survive until the decoupling of weak interactions, the neutrons will diffuse out of regions that are rich in baryon number. Thus baryon-rich regions are transformed into neutron-poor (or proton-rich) regions, and vice versa, baryon-poor regions become rich in the number of neutrons. The ratio  $\exp[(m_n - m_p)/k_B T]$  becomes spatially inhomogeneous.

Neutron-rich and proton-rich regions constitute different initial conditions for primordial nucleosynthesis. An inhomogeneous universe uses neutrons less “efficiently” than a homogeneous one. (In proton-rich regions, two neutrons makes one alpha, while four neutrons are needed in a neutron-rich part ( $4n \rightarrow 2p + 2n \rightarrow 1\alpha$ ). It has been shown by Applegate *et al.* (1987) that primordial nucleosynthesis in an inhomogeneous universe leads to a distinctive pattern in the light-element abundances. In particular, the production of  ${}^7\text{Li}$  is increased as compared to the standard scenario without baryon number inhomogeneities.

Thus the alluring prediction of this scenario is that the observation of light-element abundances today provides a snapshot of the universe at its infancy (the time period of the QCD transition). Clearly it is a question of relative size of the competing scales, whether the inhomogeneities in the baryon number densities characterized by the mean interbubble separation  $R$  transform into variations in the  $n/p$ -ratio determined by the neutron diffusion length. If  $R$  is too large, neutron diffusion will not leave an effect. If  $R$  is too small, the baryon number inhomogeneities are washed out, as both neutrons and protons diffuse over these scales. In addition, back diffusion of neutrons into proton-rich regions must be taken into account. This process further reduces the range of length scales over which QCD inhomogeneities can affect the primordial nucleosynthesis.

Finally, let us indicate how the average interbubble distance  $R$  depends on the surface tension  $\sigma$ . (Notice that in this context  $\sigma$  belongs to one of the few quantities that can be calculated from first principles and does not depend on the “choice” of scenario.) The average separation  $R$  defined at the time when the shock waves collide is given by (Kajantie and Kurki-Suonio, 1986)

$$R = t_{\text{growth}} v_{\text{sh}}, \quad (3.79)$$

where  $v_{\text{sh}} = 1/\sqrt{3}$  is the velocity of the shock waves (as introduced above), and  $t_{\text{growth}}$  denotes the time interval between the onset of nucleation and its completion, when shock waves start colliding. The duration  $t_{\text{growth}}$  is known, once the fraction  $f$  of the universe that has been affected by shock-wave fronts goes to 1. This fraction is determined by the nucleation rate  $\Gamma$  per unit time and volume, according to

$$f(t) = \int_{t_i}^t dt' \Gamma(t') \frac{4\pi}{3} [v_{\text{sh}}(t' - t)]^3. \quad (3.80)$$

In classical nucleation theory, the rate  $\Gamma$  is proportional to  $T^4 \exp(-F_b/T)$ , where  $F_b$  is the free energy of a bubble of critical size or the minimal work for nucleating a bubble. Before nucleation the free energy  $F_{\text{before}} = -p(V + V')$ , where  $V$  and  $V'$  are the volumes of the coexisting phases. After nucleation we have  $F_{\text{after}} = -pV - pV' + \sigma A$ . For one spherical bubble it follows that  $F_{\text{after}} - F_{\text{before}} = -(p' - p) 4\pi/3 r^3 + \sigma 4\pi r^2$ . A bubble of critical radius satisfies  $\partial(F_{\text{after}} - F_{\text{before}})/\partial r = 0$ , leading to

$$F_b = \frac{16\pi}{3} \frac{\sigma^3}{(p_h - p_q)^2}. \quad (3.81)$$

It remains to express the difference in pressures of both phases via the latent heat  $\Delta L$ . We have  $\Delta L = T_c(\partial/\partial T)(p_q - p_h)$  such that in the vicinity of  $T_c$  it follows  $(p_h - p_q) = \Delta L(T_c - T)/T_c$  or

$$F_b = \alpha^2 T_c \left( \frac{T_c - T}{T_c} \right)^{-2} \quad (3.82)$$

with  $\alpha^2 = [16\pi\sigma^3/(3\Delta L^2 T_c)]$ . Notice that the surface tension and the latent heat enter the nucleation rate only in the combination of  $\alpha$ . The *degree of supercooling* is determined as  $(T_c - T_{\text{PT}})/T_c$ ,  $T_{\text{PT}}$  is defined as the temperature, where the nucleation shuts off or the fraction  $f(t)$  approaches 1. From  $\Gamma(T)$  and  $f(T)$  the degree of supercooling is finally obtained as (Iwasaki *et al.*, 1994)

$$(T_c - T_{\text{PT}})/T_c \propto \alpha \tilde{\chi}^{-1/2} \quad (3.83)$$

and the average separation of nucleation centers as

$$R \approx v_{\text{sh}} \pi^{1/3} e^{\tilde{\chi}/4} \tilde{\chi}^{-3/2} T_c^{-1} \alpha$$

with

$$\tilde{\chi} \equiv 4 \ln(T_c^2/C),$$

$$C = -\partial T/\partial t. \quad (3.84)$$

For the QCD transition in the early universe  $\tilde{\chi} \approx 173$ ,  $T_c \approx 150$  MeV,  $v_{\text{sh}} \approx 1/\sqrt{3}$ . Now we are ready to insert the value for  $\alpha$  which was obtained by Iwasaki *et al.* (1994) from a measurement of the surface tension and the value  $\Delta L$  taken from the QCDPAX collaboration (Iwasaki *et al.*, 1991; Iwasaki, Kanaya, Yoshé *et al.*, 1992). The degree of supercooling turns out to be as small as

$$(T_c - T_{\text{PT}})/T_c = 5.6(1.4) \times 10^{-4}, \quad (3.85)$$

and the average distance between hadronic bubbles

$$R = 22(5) \text{ mm}. \quad (3.86)$$

The value of  $R$  is clearly too small to lead to an observable relic in the present-day light-element abundances (see, for example, Meyer *et al.*, 1991).

One should keep in mind that the calculation of  $\sigma$  in the pure SU(3) gauge theory does not directly apply to the QCD transition in the early universe, although the

main conclusions probably would remain unchanged. Dynamical matter fields will not increase the degree of supercooling. In general matter fields weaken (or even completely wash out) a first-order transition.

### C. Including dynamical fermions

The ambiguity in studying the order of the deconfinement transition in the SU(3) gauge theory was caused by the finite volume. A finite-size scaling analysis has resolved the controversy. It is natural to attempt a finite-size scaling analysis in the presence of dynamical fermions, as well. Other artifacts can superimpose on finite-size effects. In Sec. III.C.1 below, we summarize some typical pitfalls in lattice calculations. In Sec. III.C.2 we report on the *finite-size scaling analysis* of the chiral transition by Fukugita, Mino, Okawa, and Ukawa (1990). The results are not yet conclusive. Possible effects of finite quark masses are disregarded, although finite masses amount to an ordering effect, which may well compete with the finite volume. Therefore we turn in Sec. III.C.3 to a *finite mass scaling analysis*. It serves to discriminate possible UV artifacts in the chiral transition. *Bulk transitions* are a further manifestation of UV effects. With the example of  $N_f=8$  flavor QCD we illustrate bulk transitions in Sec. III.C.4. Results for the more realistic cases of two and three flavors are summarized in Sec. III.C.5. The equation of state for two-flavor QCD is the topic of Sec. III.C.6. In Sec. III.C.7 we summarize the progress that has been achieved with *Wilson fermions*.

#### 1. Pitfalls on the lattice

(i) *Finite size effects*. Finite-size effects are specific to the Monte Carlo approach, which is frequently used in lattice calculations. They contaminate any numerical calculation performed in a finite volume, if the largest correlation length of the system is not small compared to the smallest linear size. In the case of first-order transitions we have presented a phenomenological approach (see Sec. II). The formulas of Sec. II for the scaling of the specific heat or the susceptibility as a function of linear size  $L$  hold in the *zero-external-field limit*  $h=0$ .

As we have seen in Sec. III.A, Monte Carlo calculations of QCD with dynamical fermions are necessarily performed at finite bare quark masses. The updating time is estimated to increase proportional to  $m_f^{-11/4}$ . The chiral limit  $m_f \rightarrow 0$  must be extrapolated. Thus a second pitfall is a contamination of data through finite-mass effects.

(ii) *Finite mass effects*. In Sec. II we stressed the similarity between the finite volume  $L^3$  and an external field  $h$ , and between an external field and finite quark masses. In the vicinity of  $T_c$  one may consider a *finite-mass scaling analysis* at vanishing  $1/L$  rather than a *finite-size scaling analysis* at vanishing  $h$ . For a moment let us assume that the condition  $1/L \rightarrow 0$  is satisfied to a sufficient accuracy. A mass scaling analysis then provides a tool to control the rounding of algebraic singularities in correlation functions due to finite quark

masses. The precondition is a second-order phase transition. (An analogous analysis may be performed for a first-order transition and rounding effects due to finite masses as well.) The need of extrapolation to zero masses is specific for exploring the chiral limit, when dynamical fermions are included.

(iii) *Artifacts due to the UV-cutoff*. A third class of pitfalls, which is known from the pure SU( $N$ ) gauge theory and is common to any latticized version of a continuum model, comprises artifacts due to the UV cutoff. Here we distinguish three manifestations. The most obvious occurs in the interpretation of any physical observable, when the lattice units are translated to physical units. In Sec. III.A we have already discussed the ambiguity in predicting  $T_c$  (MeV). Such an ambiguity occurs for QCD transitions in the presence of dynamical quarks, when the quark mass in lattice units, e.g.,  $ma=0.025$ , is translated into units of MeV, although the lattice is rather coarse grained and a mass splitting according to  $ma$  is not unique.

A less obvious and more subtle UV artifact is a change in the effective symmetry group of the lattice action as a function of the bare coupling. This artifact is a consequence of the “No-Go” theorems referring to lattice regularized actions for fermions. As mentioned in Sec. III.A, the lattice action with massless fermions has a global  $U(n) \times U(n)$  symmetry for  $n$  species of staggered fermions. Only in the continuum limit will the full  $SU(N_f) \times SU(N_f)$  flavor symmetry be restored with  $N_f=4n$ . Far outside the continuum region, the lattice action has only global  $U(1) \times U(1)$  symmetry for  $n=1$ .

A global  $U(1) \times U(1)$  symmetry may trigger a second-order phase transition with O(2) critical exponents, whereas the spontaneous breaking of the restored  $SU(4) \times SU(4)$  symmetry in the corresponding continuum limit is supposed to induce a first-order transition (Pisarski and Wilczek, 1984). Thus the order of the chiral transition may change when passing from strong to weak couplings. Such a possibility has been discussed by Boyd *et al.* (1992) and will be the topic of Sec. III.C.3.

If one expects an  $SU(2) \times SU(2)$  flavor symmetry for the continuum limit, one should also expect the continuum phase transition to be of second order. In that case only the critical exponents would change in passing from strong to weak couplings.

If one wants to describe an odd number of flavors in the continuum limit within the staggered fermion formulation, one usually “corrects” for the desired number of continuum flavors on the level of determinants. The correction is performed after integrating over the fermionic degrees of freedom, whereas the fermionic operator  $Q$  is that of the original lattice action. The flavor symmetry in the original action is broken and describes at least four flavors in the continuum limit. The representation of the prefactor of the determinant-term in Eq. (3.14) as  $N_f/4$  cannot be derived. It should be considered as a *prescription* when  $N_f$  is not an integer multiple of 4. In this case it is even more difficult to infer the symmetry that drives the phase transition and leads to a vanishing condensate above  $T_c$ .

The UV effects mentioned so far are modifications of the critical indices or the order of the chiral transition. A third manifestation of the UV cutoff are *bulk transitions*. Here the very transition itself is an artifact of the lattice. Bulk transitions are not restricted to QCD with dynamical fermions. They occur in pure  $SU(N)$  gauge theories as well. As an illustration we consider the example of  $N_f=8$ -QCD in Sec. III.C.4. For eight flavors the bulk transition is either dominant or superimposed on the usual finite-temperature transition of QCD.

## 2. Finite-size scaling analysis

Fukugita, Mino, Okawa, and Ukawa (1990) performed a finite-size scaling analysis of the chiral transition in the presence of fermions. In this analysis the fermionic action is given by Eqs. (3.12). The quark mass  $m_f$  is chosen as 0.025 for  $N_f=2$  and as 0.025 and 0.0125 for  $N_f=4$  [with  $l=N_f/4$  in Eq. (3.14)]. The chiral condensate  $\langle\bar{\chi}\chi\rangle$  and the Wilson line  $\Omega$  are taken as order parameters, although they strictly have this meaning only in the limiting cases  $m_f\rightarrow 0$  and  $m_f\rightarrow\infty$ , respectively. The expectation value of the chiral condensate is calculated according to Eq. (3.112) below. The Wilson line expectation value is defined as

$$\langle\text{Re}\Omega\rangle = \frac{\langle\sum_x \text{Re}\Omega(x)\rangle}{N_\sigma^3},$$

$$\text{Re}\Omega(x) = N^{-1} \text{Re Tr} \prod_{\tau=1}^{N_\tau} U_{x,\tau}^{(4)} \quad (3.87)$$

with  $N=N_\sigma^3\cdot N_\tau$ . Otherwise the notations are the same as those introduced in Sec. III.B.1. The expectation value  $\langle\cdots\rangle$  is calculated with the measure and the effective action of Eqs. (3.14). The temporal extent is fixed to  $N_\tau=4$ , while the spatial size is varied between  $N_\sigma^3=4^3$  and  $N_\sigma^3=12^3$ . The adopted algorithm is the  $R$  algorithm of Gottlieb *et al.* (1987a) with a time-step size of  $\Delta\tau=0.02$ . The accuracy of this algorithm has been checked against the exact hybrid Monte Carlo algorithm.

Figure 7 shows the time history of the Wilson line  $\text{Re}\Omega$  for  $N_f=2$  and  $m_f=0.025$  on an  $N_\sigma^3\times 4$  lattice for various spatial lattice extents. It illustrates that the “naive” criteria for identifying a first-order chiral transition are not unique for small lattices. The signatures of interest are signs of metastabilities and associated two-state signals.

The time evolution should be compared with Fig. 8, where the time history of  $\text{Re}\Omega$  is displayed for the four-flavor case at  $m_f=0.025$  on an  $N_\sigma\times 4$  lattice for several values of  $N_\sigma$ . In the  $N_f=4$  case, distinct flip-flops are seen for  $N_\sigma=6$  and 8, and a single event for  $N_\sigma=12$ . (The tunneling rate goes to zero in the infinite volume limit.) For  $N_\sigma=10$  no flip-flop is visible, but a metastability is observed as a two-state separation. When transitions between both phases occur via tunneling and tunneling events are rare, the phases are metastable over a long time. The system remains in the ordered phase over

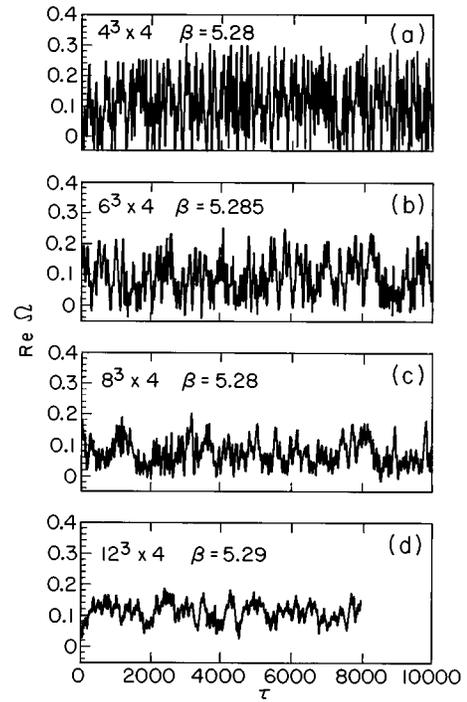


FIG. 7. Time history of the Wilson line  $\text{Re}\Omega$  for  $N_f=2$  and  $m_f=0.025$  on an  $N_\sigma^3\times 4$  lattice with  $N_\sigma=4, 6, 8,$  and  $12$  for (a), (b), (c), and (d). The time unit is set by one trajectory in the  $R$  algorithm, and  $\tau$  denotes the number of such trajectories. From Fukugita, Mino *et al.* (1990).

thousands of iterations starting from an ordered start and in the disordered phase for a random start.

A crucial ambiguity is hidden in what is called “long time.” A comparison of Figs. 7 and 8 shows a significant difference in time scales. The fluctuations in Fig. 7 look more irregular on a large scale. On a short time scale the interpretation may change. Consider Fig. 7 for the  $8^3\times 4$  lattice. If the runs had been stopped after 2000 iterations, the fluctuations (probably of statistical origin) could have been misinterpreted as a two-state signal. In the longer run of 10 000 iterations they look like a statistical fluctuation.

Figure 8 indicates for the  $10^3\times 4$  lattice how important a check of the dependence on initial conditions is (whether one uses a hot or a cold start). From a single run one can never exclude the possibility that a flip-flop will occur, if one waits for a long enough time. For large volumes, the metastabilities can be so pronounced that the system is in one phase over the entire simulation. The phase is interpreted as a truly stable phase and a first-order transition is easily overlooked. Thus the choice of the volume has to be optimized. It should be neither too small to see distinct flip-flops nor too large to see them at all.

Further indications of a first-order transition are taken from the finite-size scaling behavior of various susceptibilities. Susceptibilities are “magnetic” response functions that can be expressed in terms of an order parameter  $O$  according to

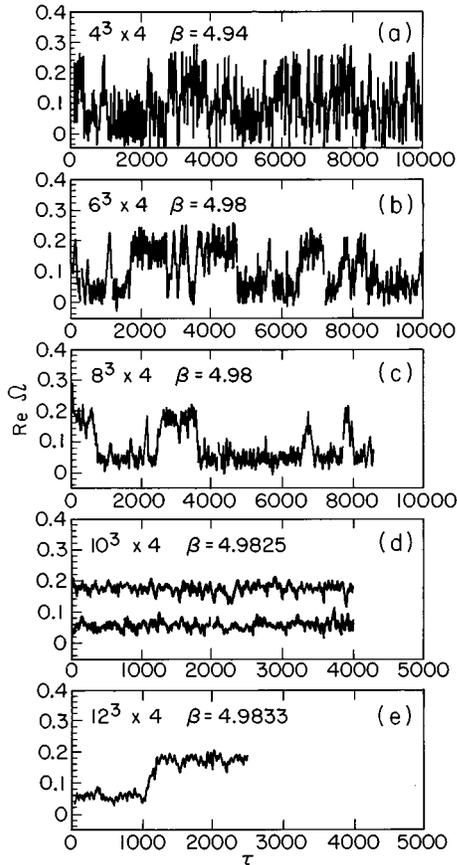


FIG. 8. Time history of the Wilson line as in Fig. 7, but with  $N_f=4$  and  $N_\sigma=4, 6, 8, 10,$  and  $12$  for (a), (b), (c), and (e). From Fukugita, Mino *et al.*, (1990).

$$\chi = V(\langle O^2 \rangle - \langle O \rangle^2). \quad (3.88)$$

Fukugita *et al.* investigated the finite-size scaling of  $\chi$  for different choices of the order-parameter field: the Wilson line  $\text{Re } \Omega$ , the chiral condensate  $\bar{\chi}\chi$  (not to be confused with the susceptibility itself), and the average plaquette  $P \equiv \text{Tr}U(\partial P)$ , where  $U(\partial P)$  denotes the product of  $U$ 's along the boundary of a plaquette.

From Sec. II and Table II we recall that the peak in the susceptibility  $\chi_{\max}$  should scale with the volume  $V$  according to  $\chi_{\max} = \text{const} + aV^p$ , where  $p=1$  for a first-order transition and  $p < 1$  for a second-order transition. The constant accounts for a contribution from the regular part of the free energy, which cannot be neglected as long as the volume is relatively small.

Runs with  $N_\sigma \geq 10$  showing no flip-flops are of no practical use for a finite-size scaling test with  $\chi$ , since the formulas are based on phase coexistence in the sample with a nonvanishing weight. A two-parameter fit for  $a$  and  $p$  of the ansatz for  $\chi_{\max}$  with various values of  $\text{const} (\geq 0)$  leads to an estimate of  $0.78 < p < 1.1$  for the Wilson line susceptibility in the case of four flavors. For comparison we mention the result for the pure  $\text{SU}(3)$  gauge theory with a first-order transition (Fukugita, Okawa, and Ukawa 1989, 1990), where the exponent  $p$  was estimated as  $0.86 < p < 1.0$ . The authors conclude from the estimate in the four-flavor case that it is con-

sistent with a first order transition.

Results for other susceptibilities lead to even larger bounds on  $p$ ; in particular, for the two-flavor case they are not conclusive.

For a further test of the order of the phase transition, we recall the definition of Binder's cumulant, Eq. (2.40). The minimum of Binder's cumulant vanishes for a second-order transition, but goes to a finite value ( $\geq 0$ ) in the case of a first-order transition. Fukugita *et al.* used a slightly different normalization,

$$V_L = 1 - \frac{1}{3} \frac{\langle P^4 \rangle}{\langle P^2 \rangle^2}. \quad (3.89)$$

Here  $P$  denotes the average plaquette. The minimum  $V_{L,\min}$  should approach  $2/3$  for a second-order transition as the volume increases. [For the explicit volume dependence, see Eq. (2.41)]. When  $V_{L,\min}$  is plotted as a function of  $1/V$ , the trend is obvious that  $V_{L,\min}$  does not approach  $2/3$  for four-flavor QCD. This is a further indication of a first-order transition. For two flavors, the deviation from  $2/3$  is smaller than for four flavors, although larger than in the pure gauge limit ( $N_f=0$ ). The difference of  $(V_{L,\min} - 2/3)$  goes almost to zero for  $N_f=0$ . This is rather surprising, as the transition is known to be of first order for zero flavors. Thus the analysis of Binder's cumulant turns out to be inconclusive as well.

It was suggested by Fukugita, Mino, Okawa, and Ukawa (1990) that the relatively large quark mass of  $m_f=0.025$  might interfere with the finite size of the system. Therefore the two-flavor simulations were repeated for a smaller mass value of  $m_f=0.0125$  (Fukugita *et al.*, 1991). The characteristic features were very similar to those for  $m_f=0.025$ , and the order of the two-flavor transition could not be determined there either. The favored possibilities were "no transition for finite  $m_f$ " or "a second-order transition in the  $m_f \rightarrow 0$  limit," but a first-order transition was not excluded.

This concludes our discussion of infrared artifacts. In the next two sections we turn to ultraviolet artifacts.

### 3. Finite-mass scaling analysis

In this section we discuss ultraviolet artifacts of phase transitions in more detail. We refer the reader to the work of Boyd *et al.* (1992) dealing with possible changes in critical indices and a change in the order of the chiral transition as a function of the bare coupling.

In the strong-coupling limit the staggered fermion action of QCD is given by Eqs. (3.12) with one difference. An additional coupling  $\gamma$  is introduced in the action, which is related to the anisotropy in the spatial and temporal lattice spacing in the weak-coupling limit. It should not be confused with the critical index  $\gamma$ , which will be considered later. The meaning of the notation should be clear from the context. The fermion operator  $Q$  is written as

$$Q_{xx'} = \sum_{\mu=1}^3 D_{xx';\mu} + \gamma D_{xx';\sigma} + m \delta_{xx'} \quad (3.90)$$

with  $D$  as in Eqs. (3.12). The temperature is varied by varying the coupling  $\gamma$  on lattices with a fixed number of time slices  $N_\tau$ . The chiral-symmetry-restoring transition is observed at some value  $\gamma(N_\tau)$  with  $\gamma(N_\tau) \rightarrow \infty$  for  $N_\tau \rightarrow \infty$ . The role of  $\gamma$  is analogous but not identical to that of the anisotropy parameter  $\xi$  [Eq. (3.19)] at weak couplings, which is directly related to the lattice spacings  $a_\beta$  and  $a_\sigma$  and hence to the temperature  $T = 1/(N_\tau a_\beta)$ . The particular role of the strong-coupling limit can be guessed from the special case in which  $\xi = 1$  and  $a_\sigma = a_\beta$ . An attempt to reach the high-temperature region in the strong-coupling ( $\beta = 0$ ) limit requires  $N_\tau < 1$ .

For vanishing quark masses the action (3.12) for one species of staggered fermions ( $n = 1$ ) is invariant under global  $U(1) \times U(1)$  transformations acting on even lattices sites according to

$$\chi' = e^{i\alpha}\chi, \quad \bar{\chi}' = e^{-i\beta}\bar{\chi}$$

and on odd sites

$$\chi' = e^{i\beta}\chi, \quad \bar{\chi}' = e^{-i\alpha}\bar{\chi}. \quad (3.91)$$

For finite mass  $m$  in Eq. (3.90) this symmetry is explicitly broken to  $U(1)$ .

Mean-field calculations of Damgaard *et al.* (1986) and Fäldt and Petersson (1986) and numerical calculations of Klaetke and Mütter (1990) for  $SU(2)$  suggest that the chiral QCD transition is of second order in the strong-coupling limit. Second-order transitions are immediately washed out in the presence of an external field, as we know from statistical physics. The singularities in thermodynamical functions will be rounded when the infinite-volume limit is taken at fixed, nonvanishing quark mass.

The critical index characterizing the “finite-mass scaling” at criticality is  $1/\delta$  (see Table II). It is defined as

$$\lim_{N_\sigma \rightarrow \infty} \langle \bar{\psi}\psi \rangle(m, t, N_\sigma) |_{t=0} \propto m^{1/\delta}, \quad (3.92a)$$

where the reduced temperature  $t = (T - T_c)/T_c$  is now replaced by

$$t = \frac{\gamma - \gamma_0}{\gamma_0}, \quad (3.92b)$$

and the order parameter

$$\langle \bar{\psi}\psi \rangle = \frac{1}{N_\sigma^3 N_\tau} \frac{\partial}{\partial m} \ln Z(m, \gamma) \quad (3.93)$$

is the chiral condensate of strong-coupling QCD. Here we have renamed the staggered fermion fields  $\chi, \bar{\chi}$  as  $\psi, \bar{\psi}$  to avoid confusion with the susceptibility  $\chi$  later on. It can be expressed in terms of expectation values of monomers, which are easily calculable on the lattice (see, for example, Karsch and Mütter, 1989).

Further recall the critical index  $\gamma$  (see Table II), which specifies the singular behavior of the susceptibility  $\chi$ , when  $T$  approaches  $T_c$  at zero field (here zero mass),

$$\chi(t) = \lim_{m \rightarrow 0} \lim_{N_\sigma \rightarrow \infty} \chi(t, m, N_\sigma) \propto t^{-\gamma},$$

$$\chi(t, m, N_\sigma) = \frac{\partial^2 \ln Z}{\partial m^2}. \quad (3.94)$$

Note the order of the limits. To measure  $\gamma$ , one first has to take the infinite-volume limit and next the zero-mass limit. (In the opposite order the order parameter would vanish in the broken phase due to tunneling events in the finite volume.) Equation (3.94) is still a zero-mass limit.

A finite-mass scaling analysis proceeds in complete analogy to a finite-size scaling analysis. The rounding and shifting effects on singularities are derived from a scaling ansatz for the free-energy density. The nonanalytic part of the free-energy density of a generic statistical ensemble in the presence of an external field (the finite quark mass) is written as

$$f(t, m) = b^{-1} f(b^{y_t} t, b^{y_h} m), \quad (3.95)$$

where  $b$  is an arbitrary scale factor as in Sec. II, and  $y_t$  and  $y_h$  are the thermal and magnetic critical exponents ( $y_t \equiv \lambda_1, y_h \equiv \lambda_2$  in our notation of Sec. II). Choosing the scale factor

$$b = m^{-1/y_h}, \quad (3.96)$$

we find that the free-energy density transforms to

$$f(t, m) = m^{1/y_h} f(tm^{-y_t/y_h}, 1). \quad (3.97)$$

Equation (3.97) implies the finite-mass scaling behavior of the order parameter  $\langle \bar{\psi}\psi \rangle(t, m)$  and the susceptibility  $\chi(t, m)$  in the vicinity of  $\gamma_0$  (the critical coupling)

$$\langle \bar{\psi}\psi \rangle(t, m) = m^{1/\delta} F(tm^{-y_t/y_h}) \quad (3.98)$$

and

$$\chi(t, m) = \frac{1}{\delta} m^{(1/\delta)-1} \left[ F(tm^{-y_t/y_h}) - \frac{y_t}{1-y_h} tm^{-y_t/y_h} F'(tm^{-y_t/y_h}) \right], \quad (3.99)$$

where  $F$  and  $F'$  are scaling functions. It follows that the peak of the finite-mass susceptibility occurs at

$$t_m = c m^{y_t/y_h} \quad (3.100)$$

and scales according to

$$\chi(t_0 = 0, m) \propto m^{(1/\delta)-1}. \quad (3.101)$$

Equation (3.101) should be compared with Eq. (2.30) for the finite-size scaling of the peak in the susceptibility. Equation (3.100) gives the shift in the critical coupling  $\gamma_0$  or in the reduced critical coupling  $t_0 \rightarrow t_m$  due to the finite quark mass.

With  $\delta = y_h/(1-y_h)$  it can be seen that the exponent  $\delta$  can be measured from the shift in  $t_m$  according to Eq. (3.100), or from Eq. (3.99) for the susceptibility, or from the mass dependence of the order parameter Eq. (3.98) at  $t = t_0 = 0$ .

For a second-order transition, the exponent  $\delta$  is characteristic for the universality class of the action. Thus a

measurement of  $\delta$  is of much interest for verifying that the restoration of the  $U(1) \times U(1)$  symmetry drives the phase transition at  $\gamma_0$ .

Boyd *et al.* (1992) have proposed a related quantity, the *chiral cumulant*, to measure  $\delta$  and  $t_0$  from finite-mass calculations. It has analogous properties to Binder's cumulant [cf. Eq. (2.40)] and is defined as

$$\Delta(t, m) = \frac{m\chi}{\langle \bar{\psi}\psi \rangle} = \frac{1}{\delta} - \frac{y_t x F'(x)}{y_h F(x)}, \quad (3.102)$$

with  $x \equiv t m^{-y_t/y_h}$ . It follows from Eqs. (3.98) and (3.99) that  $\Delta(0, m)$  gives  $1/\delta$ . The slope of  $\Delta$  increases with decreasing  $m$ ,

$$\Delta'(m) = \left. \frac{\partial \Delta}{\partial t} \right|_{t=0} \propto m^{-y_t/y_h}, \quad (3.103)$$

in such a way that the ratio  $\Delta$  itself has a  $\Theta$ -function shape for vanishing  $m$ ,

$$\lim_{m \rightarrow 0} \Delta(t, m) = \begin{cases} 1 & t > 0, \\ 1/\delta & t = 0, \\ 0 & t < 0. \end{cases} \quad (3.104)$$

Equation (3.104) follows from the definition of  $\Delta$  and Eqs. (3.98) and (3.99). When the ratios of  $\Delta(t, m)$  are plotted as a function of  $t$  for various values of  $m$ , the curves cross at the  $m=0$  critical point  $t_0$ . The crossing comes from the fact that  $\Delta$  increases with decreasing  $m$  for  $\gamma > \gamma_0$  and decreases with decreasing  $m$  for  $\gamma < \gamma_0$ . This behavior is analogous to that of Binder's cumulant ( $\langle O^4 \rangle / \langle O^2 \rangle^2$ ) where  $O$  stands either for the order parameter or for the internal energy  $E$ . In this way one may extrapolate the zero-mass critical point  $t_0$  from a series of finite-mass measurements (cf. Binder, 1981).

Such a determination of the zero-mass critical coupling is in principle free of an uncontrolled extrapolation. However, corrections originate in irrelevant terms, which may not be sufficiently suppressed in the vicinity of  $\gamma_0$ , and in contributions coming from the regular part of the free energy (see Sec. II). A further source that leads to interfering effects with the finite-mass scaling behavior is the finite lattice volume used in the Monte Carlo simulation. As we have seen above, the formulas (3.98), (3.99), (3.100), (3.104) hold in the infinite-volume limit. In general finite-mass and finite-volume effects are competing ordering effects. The relative sizes of both effects determine whether one is allowed to neglect one with respect to the other.

Let us compare the mass-scaling behavior of the peak in the susceptibility in the infinite-volume limit [Eq. (3.99)]

$$\chi_{\text{peak}} \propto m^{(1/\delta)-1} \quad \text{for } L \rightarrow \infty \quad (3.105)$$

with the finite-size scaling behavior in the zero-mass limit [Eq. (2.30)]

$$\chi_{\text{peak}} \propto L^{\gamma/\nu} \quad \text{for } m \rightarrow 0. \quad (3.106)$$

The correction coming from  $L < \infty$  as  $m \rightarrow 0$  is small compared to the rounding due to  $m > 0$  as  $L \rightarrow \infty$ , if

$$m > \text{const} \times L^{-b},$$

$$b = \frac{\gamma\delta}{\nu(\delta-1)}, \quad (3.107)$$

following from Eqs. (3.105) and (3.106). On the other hand, the quark mass has to be chosen sufficiently small to keep the contributions from the regular part of the free-energy density small. To keep the corrections small, the quark mass must be reduced proportional to  $1/\sqrt{N_\tau}$ , when  $N_\tau$  is increased in the contributions from the regular part.

The chiral condensate, the susceptibility, and  $\Delta$  have also been measured in a Monte Carlo simulation for strong-coupling QCD using the monomer-dimer algorithm (Karsch and Mütter, 1989). The lattice size  $N_\sigma^3 \times N_\tau$  was chosen as  $N_\sigma = 4, 8, 16$  and  $N_\tau = 4$ . The values for the bare fermion masses were  $m = 0.005, 0.01, 0.02, 0.04, \text{ and } 0.1$ . The results were in agreement with expectations: the chiral condensate vanished in the symmetric phase, when the mass approached zero. In the critical region  $\gamma = 2.3\text{--}2.4$  the finite-size effects were strong. For a smaller lattice a larger mass was necessary to keep finite-size effects negligible in accordance with condition (3.107). In Eq. (3.107) the critical indices  $\gamma, \delta, \nu$  are taken from three-dimensional  $Z(2)$  or  $O(N)$  models, for which  $b \sim O(2)$ . Hence the finite-size effects should be small for a  $16^3 \times 4$  lattice over the entire mass range.

The chiral condensate  $\langle \bar{\psi}\psi \rangle$  decreased smoothly in the transition region  $\gamma \sim 2.3\text{--}2.4$  as a function of  $\gamma$  without any signal of a first-order phase transition. The distribution  $P_m(\langle \bar{\psi}\psi \rangle)$  of the order parameter showed no sign of a double-peak structure (cf. Sec. II). We recall that the same lattice action as Eqs. (3.12) with (3.92) in the weak-coupling limit led to a first-order transition (Gavai *et al.*, 1990), where the continuum symmetry group  $SU(4) \times SU(4)$  of four-flavor QCD seemed to be sufficiently restored to induce the first-order transition.

The height of the peak in the susceptibility increased with decreasing  $m$ , but the statistics were not sufficient to measure  $\delta$  according to Eqs. (3.98) and (3.99). Instead  $\delta$  and the critical coupling  $\gamma_0$  were estimated from a simulation of  $\Delta$  with the result

$$\begin{aligned} 2.35 < \gamma_c < 2.4, \\ 0.18 < \frac{1}{\delta} < 0.25. \end{aligned} \quad (3.108)$$

The curves  $\Delta(\gamma)$  for different quark masses crossed in the vicinity of  $\gamma = 2.35$ . The result for  $\delta$  is in agreement with  $\delta$  for the three-dimensional  $O(2)$  spin model [ $\delta = 4.755(6)$ ]. It justifies the hypothesis that strong-coupling QCD in the staggered fermion formulation is in the same universality class as the  $3d$   $O(2)$  spin model. The reason is the  $U(1) \times U(1)$  remnant of the  $SU(N_f) \times SU(N_f)$  continuum symmetry for one species of staggered fermions.

A more subtle case is that of  $N_f = 2$  in the continuum, where the renormalization-group analysis of Pisarski

TABLE V. Critical exponents  $1/\beta\delta$  and  $1/\delta$  for two-flavor QCD in comparison with  $O(4)$ ,  $O(2)$ , and mean-field critical exponents.

	Two-flavor QCD	$O(4)$	$O(2)$	Mean field
$1/(\beta\delta)$	$0.77\pm 0.14$	$0.55\pm 0.02$	$0.60\pm 0.01$	0.67
$1/\delta$	$0.21 < 1/\delta < 0.26$	$0.208(2)\pm 0.003$	$0.2080(3)\pm 0.0003$	0.33

and Wilczek suggests a second-order transition with  $O(4)$  critical exponents. This hypothesis was recently questioned by Kocić and Kogut (1995) in an analysis of the three-dimensional Gross-Neveu model. Kocić and Kogut find mean-field scaling behavior rather than the expected universality of the two-dimensional Ising model. This casts some doubt on the conventional lore that actions for *composed* scalars (i.e., scalar mesons composed of quark bilinears) share the universality class with sigma models for *fundamental* scalars. In particular, the chiral transition in two-flavor QCD might have mean-field rather than  $O(4)$  critical exponents.

In Monte Carlo simulations with staggered fermions,  $l=n=N_f/4$  is set to 1/2 in the effective action, Eq. (3.14). No sign of a first-order transition has been found by Brown *et al.* (1990a, 1990b) for the two-flavor case. Thus a second-order transition is favored in the massless limit. The question now arises whether this continuous behavior is a remnant of the strong-coupling  $U(1)\times U(1)$  symmetry restoration or a restoration of the desired  $SU(2)\times SU(2)$  continuum symmetry with either  $O(4)$  or mean-field exponents. Karsch (1994) and Karsch and Laermann (1994) performed a finite-mass scaling analysis of critical exponents in two-flavor QCD to decide which of the three possibilities was realized:  $O(2)$ ,  $O(4)$ , or mean-field exponents. The exponent  $1/(\beta\delta) = y_t/y_h$  was read off from Eq. (3.100), expressed in terms of the critical and pseudocritical couplings  $g_c$  and  $g_{pc}$ , respectively,

$$6/g_{pc}^2(m_{qa}) = 6/g_c^2(0) + (m_{qa}N_\tau)^{1/\beta\delta}. \quad (3.109)$$

The exponent  $1/\delta$  was determined from the chiral cumulant Eq. (3.102) evaluated at  $x=t=0$ . The results are summarized in Table V and compared with  $O(4)$ ,  $O(2)$ , and mean-field values. The QCD result for  $1/(\beta\delta)$  is consistent with the  $O(2)$  and mean-field results, but slightly outside the error bars of the  $O(4)$  value. For  $1/\delta$  the QCD result is less consistent with mean-field values, but in good agreement with the  $O(4)$  and  $O(2)$  values.

DeTar (1995) has analyzed the scaling behavior of  $\langle \bar{\psi}\psi \rangle$  as function of a scaled temperature [see Eq. (3.98)] with  $O(4)$  critical exponents over a wide range of available data. The agreement is good apart from  $N_\tau=12$  data. The origin of this discrepancy has still to be clarified. Possible explanations are a shift of the crossover temperature as a function of  $N_\tau$ , erroneous extrapolations of lattice mass measurements, or the very scaling hypothesis with  $O(4)$  exponents. A clear identification of mean-field values in two-flavor QCD would certainly

be an important first step in gaining a deeper understanding of critical behavior in field theories with fermions.

The caveat concerning unwanted features of the strong-coupling symmetries also applies to the case of three flavors in the staggered fermion formulation. In the worst case, a crossover phenomenon might be the result of a second-order transition in the strong-coupling regime, which is immediately washed out when finite masses are included. The crossover phenomenon then would be an ultraviolet artifact indicating that the bare coupling  $g$  was still too strong.

Once a second-order phase transition has been identified, a measurement of critical indices is not a minor detail for its further characterization. As we have seen, at finite temperatures and fixed  $N_\tau$  the continuum limit coincides with the high-temperature limit. Thus one has to increase  $N_\tau$  to shift the critical coupling towards smaller values. The large-volume limit is necessary for the continuum limit at low temperatures. In Monte Carlo simulations *extrapolations* to zero mass, zero lattice spacing, and infinite volume are unavoidable. In the vicinity of a second-order phase transition, the critical exponents enter the extrapolation formulas. Thus their correct identification and knowledge of their precise values are needed for taking the right limits.

If finite-mass or finite-size scaling analysis is not practicable or applicable, an alternative fermion formulation, that of the Wilson fermions, should be explored to reestablish the type of phase transition that has been observed in the staggered formulation. In particular, when odd numbers of flavors are described in the staggered fermion formulation, it is difficult to ascertain how the trick of doubling and reducing the flavor degrees of freedom affects the effective symmetry of the lattice action which triggers the phase transition.

#### 4. Bulk transitions

Bulk transitions are phase transitions at zero temperature. Their very occurrence is a lattice artifact. Lattice systems are systems of statistical mechanics with their own dynamics. They do not care about a well-defined continuum limit. Generic phase transitions can occur at some critical coupling, while the temperature is zero. Their physical meaning depends on the context. They have no physical relevance for continuum QCD, if they do not “survive” the continuum limit. Thus one would like to ignore them completely, but one is not allowed to do so.

The phenomenon of bulk transitions in lattice gauge theory is known from the pure gauge sector. According to the conventional lore on four-dimensional lattice gauge theories, there should be no zero-temperature phase transition for non-Abelian  $SU(N)$  gauge groups separating the strong-coupling ( $g^2 \gg 1$ ) from the weak-coupling ( $g^2 \ll 1$ ) region. This is a desired feature, as it should guarantee that the continuum limit of lattice gauge theory includes both the confinement properties (proven on the lattice for strong couplings) and asymp-

otic freedom in the weak-coupling regime. A bulk transition may in principle destroy the confinement properties of the strong-coupling regime.

Monte Carlo calculations for the SU(2) and SU(3) gauge groups with fields in the fundamental representation have verified the conventional lore. For  $N \geq 4$  first-order bulk transitions have been found for the SU(4) lattice gauge theory by Moriarty (1981) and for the SU(5) theory by Creutz (1981). Creutz argues, however, that the confinement property is not lost at weak couplings in spite of the transition.

In SU(2) and SU(3) gauge theories one observes, instead, a rapid crossover phenomenon between the strong and weak coupling regimes. This is explained by a nearby critical point in the  $(\beta, \beta_A)$  plane, where  $\beta \equiv 2N/g^2$  as above and  $\beta_A$  denotes the coupling of a  $\text{Tr}_A U(\partial P)$  term in the adjoint representation of SU( $N$ ). For two colors we have a mixed SU(2)-SO(3) lattice action. A small  $\beta_A$  leads to a bulk transition (Bhanot and Creutz, 1981), which is absent for  $\beta_A = 0$ .

The interplay of bulk and thermal transitions in SU(2) and SU(3) gauge theories has been recently re-investigated by Gavai *et al.* (1994), Gavai and Mathur (1995), and Blum *et al.* (1995b). A bulk transition is signalled if the location of the phase boundary stays fixed in coupling parameter space, independently of the number of time slices. In contrast, continuum universality for the thermal transition requires a shift of the critical couplings towards smaller values, as  $N_\tau$  is increased (cf. Sec. III.A). For the mixed fundamental/adjoint SU(2) action, Gavai and Marthur (1995) found a shift towards weaker couplings of both the thermal and the bulk transition boundaries at  $N_\tau = 6$ . This result casts some doubt on the interpretation of the first-order phase boundary as a bulk transition.

For the SU(3) mixed fundamental/adjoint action, the bulk and thermal phase boundaries coalesce for  $N_\tau = 4$ , but split into two lines for small enough couplings and larger values of  $N_\tau$  ( $N_\tau = 6, 8$ ; Blum *et al.*, 1995c). The shift of the thermal transition line towards weaker couplings for increasing  $N_\tau$  supports the hypothesis of continuum universality.

The peculiar behavior that is sometimes found for the phase structure of QCD with dynamical fermions has been attributed to bulk transitions in mixed fundamental/adjoint actions. These actions arise as effective actions from the integration over fermions. It is still an open question whether the inclusion of dynamical Wilson fermion leads to a strong enough adjoint term in the SU(3) mixed action to explain the phase structure for Wilson fermions at large  $\kappa$  and  $N_\tau = 6$  (Rummukainen *et al.*, 1995).

Similarly the integration over eight dynamical flavors in the staggered fermion scheme has been conjectured to induce an adjoint term of the SU(2)-SO(3) mixed action in the effective action. This provides a possible explanation of the bulk transition, which is seen in  $N_f = 8$  QCD (for  $N_\tau \geq 8$  time slices; Brown *et al.*, 1992). We discuss the eight-flavor case now in more detail.

The action for two species of staggered fermions is given by Eqs. (3.12). Integration over the fermionic variables leads to

$$S = -\frac{1}{3} \beta \sum_p \text{Re Tr} U(\partial P) - \frac{N_f}{4} \ln \det(QQ^\dagger) \quad (3.110)$$

with

$$Q = D + ma,$$

where  $D$  is the Dirac operator acting on an SU(3) triplet field  $\phi$  according to

$$(D\phi)_x = \frac{1}{2} \sum_\mu \Gamma_\mu(x) [U_x^{\mu+} \phi_{x+\hat{\mu}} - U_{x-\hat{\mu}}^\mu \phi_{x-\hat{\mu}}]. \quad (3.111)$$

Here the number of continuum flavors  $N_f$  is chosen as 8, corresponding to  $l = N_f/4 = 2$  species of staggered fermions. The dynamical quark mass  $ma$  is chosen flavor independently as 0.015 throughout all simulations. The gauge coupling  $\beta = 6/g^2$  is varied between 4.5 and 5.0. The lattice size is  $16^3 \times N_\tau$  with  $N_\tau = 4, 8, 16$ , and 32.

One may wonder why the special case of eight light flavors is of any interest at all, as only two (or three) quark flavors are approximately massless in nature. Nevertheless the reason for us is a physical one. Usually the strength of the chiral transition is thought to grow with an increasing number of flavors. Earlier work on  $8^3 \times 4$  (Kogut *et al.*, 1985; Fukugita *et al.*, 1988),  $6^4$  (Kogut and Sinclair, 1988a),  $8^4$  (Kogut and Sinclair, 1988a; Fukugita *et al.*, 1988), and  $16^3 \times 4$  and  $\times 6$  (Ohta and Kim, 1991) lattices shows a strong first-order transition. A comparison of the transition for 2, 3, 4, and 8 flavors reveals a strengthening of the transition as the number of flavors increases (Ohta and Kim, 1991; also Gottlieb, 1991), where the range of time slices  $N_\tau$  lies between 4 and 8.

Usually this tendency is interpreted as a reflection of a physically plausible effect: The chiral transition becomes more pronounced the higher the number of flavors that drive the transition, whatever the “driving dynamics” in detail may be. Results about the  $N_f = 8$  transition expose this tendency as a possible lattice artifact and have given rise to further studies in effective models about the supposed flavor dependence of the chiral transition.

Let us first give a qualitative description of the diagram displayed in Fig. 9. It is partly conjectural and summarizes the results for the  $N_f = 8$  transition of Brown *et al.* (1992; see also Christ, 1992b; Dong and Christ, 1992). The solid line locates the bulk transition separating the parameter space into a weak-coupling ( $\beta \geq \beta_c = 6/g_c^2$ ) and a strong-coupling ( $\beta \leq \beta_c$ ) phase. For a temporal extent  $N_\tau \geq 8$  the transition becomes  $N_\tau$  independent and occurs at the same critical coupling  $\beta_c = 4.73(1)$  for  $N_\tau = 8$  and  $N_\tau = 16$  or at  $\beta_c = 4.64(1)$  for  $N_\tau = 8$  and  $\beta_c = 4.62(1)$  when the finite time-step interval  $\Delta\tau$  in the integration step is varied (see below). This behavior signals the bulk feature of the transition.

#### a. Identification of the lattice artifact

As mentioned above, in the usual finite-temperature transition with a correspondence in continuum field

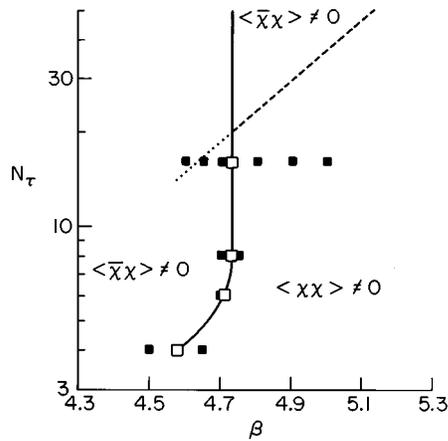


FIG. 9. Generic phase diagram in the  $N_\tau\beta$  plane for eight flavor QCD, where  $\beta=6/g^2$  and  $N_\tau$  is the temporal extent of the lattice. Units are dimensionless lattice units. The full line locates the bulk transition, the dashed line suggests a possible finite-temperature transition. Solid squares label parameter values where simulations have been actually performed, while open squares locate critical couplings. From Brown *et al.* (1992).

theory, the critical coupling (temperature) should scale with  $N_\tau$  such that the physical transition temperature  $T_c=[N_\tau a(g)]^{-1}$  remains constant (see Sec. III.A.). Thus it is the absence of a shift in  $\beta_c$  when  $N_\tau$  is increased which suggests a possible lattice artifact. The persistence of the transition on a symmetric ( $T=0$ )  $16^4$  lattice further supports the interpretation as a bulk transition.

While the strong-coupling phase shares features with the usual chirally broken phase at  $N_\tau=4,6$ , the structure of the weak-coupling phase (on the right-hand side of the bulk transition) is more complex and less obvious in its physical meaning. The weak-coupling phase is crossed by a hypothetical finite-temperature transition, which should be relevant for the continuum limit. It separates regions with chiral symmetry broken ( $\langle \bar{\chi}\chi \rangle \neq 0$ ) and chiral symmetry restored ( $\langle \bar{\chi}\chi \rangle = 0$ ) features. The conventional finite-temperature transition may be recovered on very large lattices ( $N_\tau > 32$ ).

The solid squares in Fig. 9 refer to parameters of actual simulations, while the open squares locate the critical couplings.

In what follows we sketch the criteria that led to the conjectures of Fig. 9.

### b. Basic observables

Basic observables are the chiral condensate and hadron masses. Of particular interest is their dependence on the bare quark masses. The chiral condensate of staggered fermion fields, defined as

$$\langle \bar{\chi}\chi \rangle = \frac{1}{3} \frac{1}{N_\sigma^3 N_\tau} \sum_x \langle \bar{\chi}_x \chi_x \rangle \quad (3.112a)$$

with a sum over all lattice sites  $x$ , is estimated by

$$\begin{aligned} \langle \bar{\chi}\chi \rangle &= \frac{1}{3} \frac{1}{N_\sigma^3 N_\tau} \left\langle \left\langle \sum_{xx'} h_x (D+m)_{xx'}^{-1} h_{x'} \right\rangle \right\rangle \\ &= \frac{1}{3} \frac{1}{N_\sigma^3 N_\tau} \left\langle \left\langle \sum_{xx'} h_x (m(DD^\dagger + m^2)^{-1})_{xx'} h_{x'} \right\rangle \right\rangle. \end{aligned} \quad (3.112b)$$

For each site  $x$ ,  $h_x$  is an independent, complex three-vector of Gaussian random numbers. This representation of  $\langle \bar{\chi}\chi \rangle$  follows from the effective action in terms of pseudofermionic fields  $h_x$ ; see Sec. III.A. Equation (3.112b) shows the nonlocality of the fermionic condensate. The expectation value  $\langle \langle \dots \rangle \rangle$  denotes an average over gauge fields and random three-vectors  $h_x$ .

The mass  $m$  entering Eqs. (3.112) should be identical with  $m$  of the effective action, as it enters the determinant in the path integral. Strictly speaking, the full calculations should be repeated for several quark mass values, if the quark condensate and the hadron propagators are to be checked on their quark mass dependence. The observables are then measured on new sets of equilibrated gauge-field configurations depending on the quark mass via the fermionic determinant. This way of proceeding costs an enormous amount of computing time.

In the actual measurements Brown *et al.* calculated the observables on the same set of configurations, generated for an action with  $ma=0.015$ , but with varying valence quark masses:  $m_{\text{val}}=0.004, 0.01, 0.025$ , and  $0.05$ . Valence quark masses enter the observables via the propagator, whereas sea-quark masses are used for the determinant in simulations with dynamical fermions. The mass entering Eqs. (3.112) should be identified with a valence quark mass.

A conclusive criterion for studying the chiral transition is the scaling behavior of the chiral condensate  $\langle \bar{\chi}\chi \rangle$  and the pion mass  $m_\pi$  as a function of quark mass  $m$  in the limit of  $m \rightarrow 0$ . Spontaneous chiral symmetry breaking is indicated if  $\langle \bar{\chi}\chi \rangle$  stays finite for  $m \rightarrow 0$ , while  $m_\pi(m)$  should vanish in the same limit. From chiral perturbation theory and partial conservation of axial vector current (PCAC) relations, one expects  $m_\pi^2$  to vanish linearly in the quark masses. As is argued by Brown *et al.*, the limit  $m_{\text{val}} \rightarrow 0$  for fixed  $m_{\text{sea}}$  may be conclusive as well.

The simulation method is the R algorithm of Gottlieb *et al.* (1987a). The R algorithm contains finite time-step errors of the order  $(\Delta\tau)^2$ , but requires half the number of Dirac propagator inversions per unit Monte Carlo time compared to the exact hybrid Monte Carlo method of Duane *et al.* (1987), when it used for the eight-flavor case.

### c. Determination of the critical coupling

An accurate determination of the critical coupling is essential for an identification of the bulk transition. Here one faces the usual dilemma encountered for strong first-order transitions. If the volume is chosen too large, strong metastabilities make both phases stable

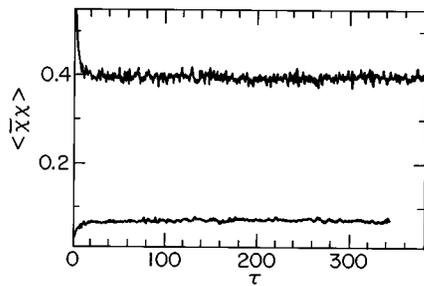


FIG. 10. Hot (upper curve) and cold (lower curve) start evolution of  $\langle \bar{\chi}\chi \rangle$  on a  $16^4$  lattice at  $\beta=4.65$ . The time unit is set by one trajectory in the  $R$  algorithm;  $\tau$  denotes the number of such trajectories. For further definitions, see the text. From Brown *et al.* (1992).

within a large range of couplings. This requires long Monte Carlo runs on large lattices. Metastabilities are less pronounced when the volume is small, but then the data are contaminated with strong finite-size effects.

This problem is solved in the simulations of Brown *et al.* by starting with a mixed phase configuration. Small changes in the coupling cause the system to evolve rapidly into one of the phases, which by itself is a typical signal for a first-order transition. Starting from a mixed phase configuration, the evaluation of  $\langle \bar{\chi}\chi \rangle$  is followed for several values of  $\beta$  to get upper and lower bounds on  $\beta_c$ .

The first order of the bulk transition is concluded from Figs. 10 and 11. Figure 10 shows the evolution of  $\langle \bar{\chi}\chi \rangle$  starting from hot and cold starts for  $\beta=4.65$  on a  $16^3 \times 16$  lattice. The first order is signalled by the persistence of two phases over a time scale considerably larger than the equilibration time. This criterion was mentioned in Sec. II as one of the naive criteria that are applicable for strongly first-order transitions. A further evidence comes from the jump in the order parameter seen in Fig. 10, which is interpreted as a tunneling event between coexisting phases in the transition region (see Sec. II). The evolution of the order parameter for the  $16^3 \times 32$  lattice at  $\beta=4.6$  is shown for a cold start. A cold

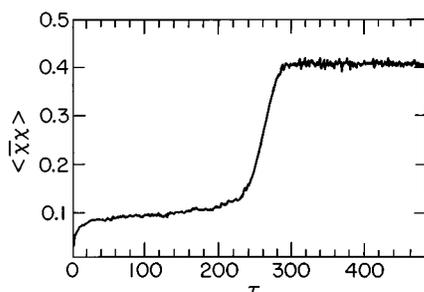


FIG. 11. Evolution of  $\langle \bar{\chi}\chi \rangle$  from a cold start on a  $16^3 \times 32$  lattice with  $\beta=4.60$ . The jump at  $\tau \approx 250$  is interpreted as tunneling from the metastable, weak-coupling phase to the stable, strong-coupling phase at this  $\beta$  value. The time unit is set by one trajectory in the  $R$  algorithm;  $\tau$  denotes the number of such trajectories. From Brown *et al.* (1992).

start corresponds to an “ordered” configuration typical of the weak-coupling phase (for a finite-temperature transition typical of the deconfinement phase). Since the critical coupling  $\beta_c=4.73$ , the value  $\beta=4.6$  belongs to the strong-coupling regime. The tunneling event that occurs at  $\tau \sim 250$  time units suggests that the enforced weak-coupling phase (via the initial configuration) becomes unstable at the strong-coupling value  $\beta=4.6$ .

#### *d. Indications of chiral symmetry restoration*

For  $N_\tau=4, 6$ , and 8 time slices, the weak-coupling phase (on the right-hand side of the solid line in Fig. 9) shows the typical features of chiral symmetry restoration. The condensate  $\langle \bar{\chi}\chi \rangle$  extrapolates linearly to zero as  $m_{\text{val}} \rightarrow 0$ . For  $N_\tau=16$  and/or 32, the strong-coupling phase has the typical features of a chiral symmetry broken phase.

Thus one would like to observe indications of chiral symmetry restoration in the weak-coupling phase. The value of  $\langle \bar{\chi}\chi \rangle$  becomes smaller when going from strong to weak couplings, but less small than expected from naive scaling arguments. It extrapolates to zero as  $m_{\text{val}} \rightarrow 0$ , but in a nonlinear way, while  $m_\pi^2$  extrapolates to a nonvanishing value. It is speculated that for  $N_\tau=16$  and  $\beta=4.65$  the weak-coupling phase is itself near a finite-temperature transition region.

For a larger value of  $\beta$  ( $\beta=5.0$ ), chiral symmetric behavior is manifest: hadronic screening lengths show parity doubling,  $\langle \bar{\chi}\chi \rangle$  extrapolates linearly to zero as  $m_{\text{val}} \rightarrow 0$ , and  $m_\pi \neq 0$  for  $m_{\text{val}}=0$ , varying little with  $m_{\text{val}}$ . For further (technical) details we refer the reader to the original reference (Brown *et al.*, 1992).

Our discussion may have demonstrated that the bulk transition does not merely replace the finite-temperature transition such that it could be ignored as a lattice artifact. It seems to be superimposed on the structure of a finite-temperature transition. For a smaller number of flavors ( $N_f=2,3,4$ ), no bulk transition has been observed. There it may be even more difficult to discriminate precursors of the bulk transition at  $N_f=8$  from continuum behavior.

Brown *et al.* (1992) attempt to explain the bulk transition as an outgrowth of the rapid crossover region seen in the pure SU(3) gauge theory for  $\beta=5.6$ . As the system goes from strong to weak couplings in this region, a strong deviation from the scaling behavior predicted by the perturbative  $\beta$  function is seen (Kennedy *et al.*, 1985). Adding light dynamical quarks to the pure gauge action, the crossover region narrows with increasing number of flavors. The sharper the crossover, the stronger the violation of perturbative scaling and the larger the increase in slope of  $N_\tau$  vs  $\beta=6/g^2$ . A plot of  $N_\tau$  vs  $\beta_c$  for  $N_f=0, 2, 4$ , and 8 flavors supports this view.

In concluding we summarize the phase structure that is expected for various temporal extents. For small values of  $N_\tau$  ( $N_\tau \leq 4$ ) one has to work with a coarse-grained lattice to reach the transition region. The scale is controlled by the lattice spacing. A single finite-

temperature transition is observed, but its relevance to a finite-temperature transition in the continuum is not obvious.

At intermediate values of  $N_\tau$  ( $4 \leq N_\tau \leq 16$ ) the corresponding critical couplings fall in the crossover region. A change occurs in what is called the relevant scale. At the end of the crossover region the scale becomes loosely related to the lattice spacing, but controlled by the continuum behavior. This rapid change of relevant scales is manifest in a bulk transition preventing any smooth change of  $N_\tau$  vs  $\beta$ ; both quantities are apparently unrelated.

For larger values of  $N_\tau$  ( $\geq 32$ ) (as well as a simultaneous extension of the spatial lattice size to mimic a finite-temperature box), the transition region is expected to lie in a coupling regime, where the lattice is rather fine grained. The bulk transition has disappeared, and a finite-temperature transition recurs. This time it will be related to the chiral transition of continuum QCD (taking for granted that it does recur).

Thus the  $N_f=8$  simulations—although far from modeling realistic QCD systems—allow an identification of lattice artifacts that may also influence more realistic lattice simulations ( $N_f=2,3$ ) in a weakened form.

## 5. Results for two and three flavors

The limiting cases we have discussed in the previous sections are not close to the realistic conditions of QCD. The pure gauge theory, the limit of four and eight nearly massless flavors, and the strong-coupling approximation may be regarded as tools for gaining some insight into the dynamical origin of the QCD transitions. They can further give some hints about the stability of QCD results with respect to variations of input parameters. Of particular interest is the role of the quark masses. For example, a crossover phenomenon for experimental quark masses may be understood as the result of mass values that are too small to sustain a first-order deconfinement transition and too large to sustain the chiral transition. Such a conclusion can be drawn if unrealistic mass values have been studied before.

Now let us turn to the cases of two and three flavors, which come closest to the experimental relation of two light [ $m_u \sim m_d \sim 5-7$  (MeV)] and one less light [ $m_s \sim 150-180$  (MeV)] flavor. Here we discuss the results of Brown *et al.* (1990a, 1990b), which are still representative for two light and one heavier flavor.

The simulations of Brown *et al.* are performed in the staggered fermion formulation. As we have argued above, the staggered formulation represents intrinsically only integer multiples of four continuum flavors. The projection on two or three flavors is enforced by writing  $l$  of the effective action in Eq. (3.14) as  $N_f/4$ . Although the local fermionic operator  $\mathcal{Q}$  describes  $N_f=4n$  ( $n$  integer) flavors, one allows  $N_f$  to take the desired continuum value and uses as a prescription for the effective action

TABLE VI. Mass parameters and results for two and three flavors. The quantity  $a$  represents the lattice constant and masses are given in lattice units.

$N_f$	$m_{u,d} \cdot a$	$m_s \cdot a$	Results of Brown <i>et al.</i> (1990a)
2	0.01	$\infty$	no transition
2	0.025	$\infty$	no transition
3	0.025	0.025	first-order transition
3	0.025	0.1	no transition

$$S_{\text{eff}} = \frac{1}{3} \beta \sum_p \text{Re Tr} U \left( \partial p - \frac{1}{4} N_{u,d} \ln \det(D + m_{u,d} a) - \frac{1}{4} N_s \ln \det(D + m_s a), \right) \quad (3.113)$$

where  $\beta = 6/g^2$ ,  $U(\partial p)$  denotes the product of  $U$ 's  $\in \text{SU}(3)$  along the boundary  $\partial p$  of a plaquette  $p$ ,  $N_{u,d}$  is the number of continuum up- and down flavors, that is, 2, and  $N_s$  is the number of strange flavors in the continuum,  $N_s=1$ . The Dirac operator  $D$  is given by Eq. (3.12). It describes four flavors of quarks in the continuum limit. Thus the representation of the prefactor  $n$  as  $N_f/4$  cannot be derived from an integration of a local action over fermionic degrees of freedom. The local staggered fermion action leads to  $N_f=4n$ , hence one should call Eq. (3.113) a “prescription.”

The condensate is given by Eqs. (3.112). The algorithm evolving the gauge fields with respect to the action Eq. (3.113) is the R algorithm of Gottlieb *et al.* (1987a) with a step size of  $\Delta\tau=0.0078$  for  $N_f=2$  and 0.01 for  $N_f=3$ . The lattice size is fixed to  $16^3 \times 4$ . Table VI lists the quark masses that were considered together with the results for the chiral transition. The results are based on Figs. 12(a)–12(d).

No transition for the two-flavor case is concluded from the time evolution of the ordered and disordered starts; see Fig. 12(a). No sign of metastability is seen, as the two starts mix together without clear tunneling events. In Fig. 12(b) the absence of a double-peak structure is taken as an indication for “no transition.”

Figure 12(c) shows the case of three light degenerate flavors. Here a two-state signal is visible. Over more than 2000 time units, the system stays in the ordered (disordered) phase depending on the starting condition. For two light and one heavy flavor [Fig. 12(d)]—the case that comes closest to realistic mass relations—the order parameter evolves similarly to Fig. 12(a). Clear signs of metastability and two-phase coexistence are absent. This result is in conflict with earlier conclusions of a clear transition on an  $8^3 \times 4$  lattice (Gavai *et al.*, 1989; Kogut and Sinclair, 1988b) and with some evidence for a first-order transition (Kogut and Sinclair, 1989).

At first glance the absence of first-order signals under realistic quark mass conditions have far-reaching consequences for phenomenological implications in heavy-ion collisions. Many predictions rely on the first-order nature of the chiral transition for three flavors (see Sec. V). From a practical point of view the alternative between a truly first-order transition and a crossover phenomenon

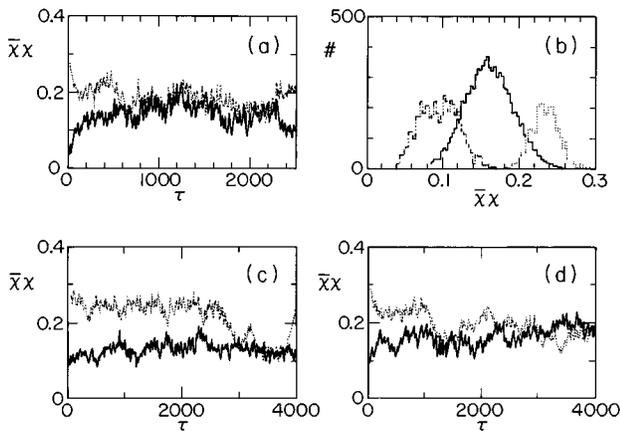


FIG. 12. Evolution of  $\langle \bar{\chi}\chi \rangle$  with various numbers of flavors, where  $a$  is the lattice constant and  $ma$  are masses in lattice units: (a)  $N_f=2$  with  $ma=0.01$  and  $\beta=5.265$ . Evolution of  $\langle \bar{\chi}\chi \rangle$  without clear tunneling events. Solid curve, ordered start; dotted curve, disordered start. (b)  $N_f=2$  with  $ma=0.01$ . Overlapping histograms from left to right,  $\beta=5.275, 5.265, 5.25$ . Counts are in units of trajectories in phase space. (c)  $N_f=3$  with  $ma=0.025$ ,  $\beta=5.132$ . Evolutions signal a first-order transition. Solid curve, ordered start; dotted curve, disordered start. (d)  $N_f=3$ ,  $\beta=5.171$  with nearly physical masses:  $m_{u,d}a=0.025$ ,  $m_s a=0.1$ . The evolution of  $\langle \bar{\chi}\chi \rangle$  signals no transition. From Brown *et al.* (1990a).

may not be distinguishable when the volume is small. The more sensible question to ask is whether the crossover is rapid enough to produce a sufficiently large gap in entropy densities over a small temperature interval. The preliminary answer seems to be positive, even for the  $N_f=2$  case.

The jump in entropy density occurs within a  $\beta$  interval that corresponds to a temperature interval of less than 10 MeV. In Sec. V.C we shall see that an entropy jump over a finite, but small, temperature range is *in principle* sufficient to induce multiplicity fluctuations beyond the statistical noise.

Although a translation from lattice units into physical units should be made with care for the considered  $\beta$  value of 5.171, it is of interest to estimate bounds on the critical quark masses in units of MeV. If we take  $a$  (MeV $^{-1}$ ) from Born *et al.* (1989), the set of masses  $m_u a = m_d a = 0.025$ ,  $m_s a = 0.1$  corresponds to  $m_u = m_d \sim 12$  MeV,  $m_s \sim 50$  MeV. These mass values give an upper bound on the critical quark masses at which the chiral transition changes from first order to second order.

In the case of three degenerate flavors  $N_f=3$ , two-state signals are observed for  $m_{u,d,s}=0.025$  (see Table VI), while no clear signals of metastability are seen for  $m_{u,d,s}=0.075$  (Gavai and Karsch, 1985; Gavai *et al.*, 1987), leading to an estimate for the critical lattice quark masses in physical units of  $12 \text{ MeV} \leq m_{u,d,s}^{\text{crit}} \leq 38 \text{ MeV}$ . These numbers should be compared with more recent results on critical quark masses obtained in Wilson's fermion scheme; see Sec. III.C.7. Lattice quark masses in physical units are related to the current quark masses in units of MeV by an unknown multiplicative renormalization factor.

Finally we come to the reliability of the results for two and three flavors. All simulations reported so far have been performed on a  $16^3 \times 4$  lattice. A temporal extent of four time slices leads to a transition region in the coupling range, where the lattice is rather coarse grained. For a pure gauge theory  $N_\tau \geq 10$  is necessary to reach the continuum region. The effect of fermions is to further lower the effective lattice spacing, so that an even larger temporal extent would be necessary to reach the range of asymptotic scaling.

The “distance” from the continuum limit is visible in the results of Brown *et al.* (1990a, 1990b) for hadron masses that were obtained in separate  $T=0$  simulations on a  $16^3 \times 24$  lattice with  $\beta=5.171$ ,  $m_{u,d}a=0.025$ , and  $m_s a=0.1$ . These calculations were performed to test the scaling properties in the considered coupling regime. The masses of two kaons that are degenerate in their flavor content in the continuum limit still differ by a factor of 2. The nucleon-over-rho mass ratio  $m_N/m_\rho$  is obtained as 1.5(1) in contrast to its physical value of 1.22. The ratio of  $m_K/m_\rho=0.46(1)$  is smaller than its physical value of 0.64. This suggests that the strange mass entering the lattice  $K$  meson is smaller than its physical value. Hence the first-order transition for three flavors disappears even *before* the strange mass adopts its physical value. The unphysical masses of the flavor partners can influence the transition dynamics in a way that is difficult to control.

Other UV artifacts due to the coarse-grained lattice may be hidden in the results. From the discussion in Sec. III.C.3 it cannot be excluded that the smooth behavior for the two-flavor case is a remnant of strong-coupling  $U(1) \times U(1)$  symmetry. Going to larger  $N_\tau$  values, it may happen that a first-order transition recurs. The same warning applies to the three-flavor case with one heavy and two light flavors, as the  $U(1) \times U(1)$  symmetry at strong couplings leads to a second-order transition independent of the number of fermionic flavors.

Concerning possible superimposed features from the bulk  $N_f=8$  transition,  $N_\tau=4$  is probably below the dangerous temporal sizes, where the associated coupling values fall in the crossover region from strong to weak couplings.

From experience with other simulations it is clear that the lattice size of  $16^3 \times 4$  is not large enough to be in the asymptotic *large-volume* region, where finite-size effects can be excluded.

More recent results for the two-flavor case of staggered fermions have been obtained for larger  $\beta$  values (i.e., smaller couplings), larger temporal extent  $8 \times 16^3$ , and/or smaller values of the bare quark masses. The bare quark masses are  $ma=0.00625$  (HTMCGC collaboration, see Bitar *et al.*, 1993; Gottlieb *et al.*, 1993, and Bernard *et al.*, 1995a),  $ma=0.004$  (Columbia group, see Mawhinney, 1993), and  $ma=0.008$  and  $0.016$  on  $24^3 \times 12$  lattices (Bernard *et al.*, 1995a, 1995b). The qualitative conclusions are the same. There is a rapid crossover for finite quark masses. As a sign of progress it should be mentioned that the pion mass has decreased to the order of  $T_c$  for the lowest value of the bare quark

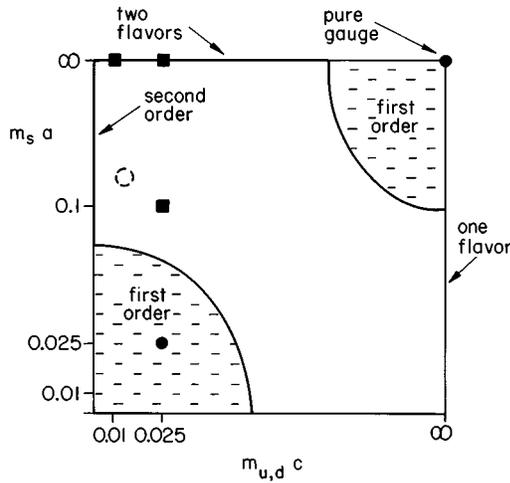


FIG. 13. Generic phase diagram, partly conjectural, for two and three flavors. Solid lines indicate supposed second-order transitions, shaded areas first order transitions; ●, mass parameters where the transition is seen; ■, where the transition is absent. The dashed circle locates the physical mass point. From Brown *et al.* (1990a).

mass. In earlier calculations (e.g., Brown *et al.*, 1990a, 1990b) it was too large for heavier mesons to decay into pions.

Another good indicator of the crossover phenomenon is the *baryon number susceptibility*. It is expected to be small in the low-temperature phase, since the baryon number can only be changed by creating or destroying a baryon, but large in the high-temperature phase, where it is sufficient to create or destroy single quarks to change the baryon number. Bernard *et al.* (1995a, 1995b) used the coupling dependence of the baryon number susceptibility  $\chi$  to determine the crossover coupling  $6/g_c^2$  for two-flavor QCD at given quark masses. They also plotted the baryon number susceptibility as a function of  $T$  (MeV) for various lattice sizes. The  $T$  dependence was obtained from  $T/m_\rho = 1/(N_\tau \cdot a \cdot m_\rho)$ , if the zero-temperature spectrum calculations of  $m_\rho a$  were extrapolated to the gauge coupling and the quark mass values of the thermodynamic simulations. Once the  $N_\tau$  dependence is under control, a plot of  $\chi(T)$  is of much interest for phenomenological applications. The baryon number susceptibility affects QCD's equation of state, hadronization processes, and heavy-ion collisions.

We conclude with a summary of conjectures and results for two and three flavors (based on the work of Brown *et al.*, 1990a, 1990b) in Fig. 13. Indicated are the presence or absence of the finite-temperature QCD transitions as a function of the quark masses  $m_s a$  and  $m_{u,d} a$  in lattice units. The solid line indicates second-order transitions, the shaded areas enclose mass values leading to first-order transitions. The concave shape of the critical boundary is hypothetical. The solid circles refer to mass points where a first-order transition is seen, the solid squares to points where no transition is found in the simulations of Brown *et al.* (1990a, 1990b). The dashed circle is located at the physical value of the

ratio of strange- to up-quark masses. Error bars for the location of the various points relative to the phase boundaries have been left out. The mass point  $(0,0)$  is the chiral limit of three flavors with an  $SU(3) \times SU(3)$  symmetry in the continuum limit and a first-order transition. The pure gauge theory in the upper right corner of the diagram ( $m_{u,d} = \infty$ ,  $m_s = \infty$ ) has an exact  $Z(3)$  symmetry both in the continuum and on the lattice for three colors. The transition is of first order.

The line ( $m_{u,d} = 0$ ,  $m_s$ ) has an  $SU(2) \times SU(2)$  continuum symmetry and a conjectured second-order transition above some tricritical strange-quark mass, where the line of first-order transitions ends. The universality class of QCD may change along the second-order transition line. There have been several proposals for candidates of the universality class: the  $O(4)$  model if  $m_{u,d} = 0$  and  $m_s$  is large; a tricritical  $\phi^6$  model in the vicinity of the tricritical point (Wilczek, 1992; Rajagopal and Wilczek, 1993a, 1993b; see Sec. IV.A.1); and an Ising model along the concave part when  $m_{u,d} \neq 0$  (Gavin *et al.*, 1994a). An identification of the appropriate universality class of QCD along the critical phase boundary is still in an exploratory stage. The precise location of the critical phase boundary is under debate. In particular, the location of the tricritical mass point on the ( $m_{u,d} = 0$ ) axis is as yet unknown. Both fermion schemes agree on the existence of a first-order transition region, but do not agree about its extent. In the staggered scheme the *physical* mass point lies outside, while in Wilson's scheme it lies inside the first-order region. Strong-coupling artifacts act in opposite directions in both schemes. As mentioned above, the smooth behavior in the staggered formulation may be a result of the  $U(1) \times U(1)$ -symmetry at strong couplings, while the first-order behavior in Wilson's formulation may be caused by light-fermion doublers at strong couplings (see Sec. III.C.7). Results for Wilson fermions will be summarized in Sec. III.C.7.

## 6. The equation of state for two-flavor QCD

Blum *et al.* (1995a, 1995b) have studied the energy density and pressure as a function of temperature for two light flavors in the staggered fermion formulation of QCD. The ir approach is based on fully nonperturbative ingredients along the same lines as in the pure gauge theory (cf. Sec. III.B.2). The generalization comes from the dependence of observables on the gauge coupling  $6/g^2$  and the bare light-quark mass  $m_q a$ . Thus the  $\beta$  function now has two components,

$$\beta(6/g^2, am_q) = \left( \frac{\partial(6/g^2)}{\partial \ln a}, \frac{\partial(am_q)}{\partial \ln a} \right). \quad (3.114)$$

For a number of points in  $(am_q, 6/g^2)$  space, the  $\beta$  function was extracted from data for  $(m_\pi a)$  and  $(m_\rho a)$  measured as functions of  $6/g^2$  and  $am_q$ . To find the change  $\delta(am_q)$  for a given change  $\delta(6/g^2)$  such that the physics remained the same, Blum *et al.* kept the mass ratio  $m_\pi/m_\rho$  fixed. An alternative way of finding two equations for the two unknown functions  $(am_q)(a)$  and

$(6/g^2)(a)$  is to fit  $m_\pi/m_\rho$  and  $m_\rho a$  as functions of  $m_q a$  and  $G$ , leading to  $a=a(6/g^2, m_q a)$ . The inverse function then yields the  $\beta$  function corresponding to a symmetric change of lattice spacings.

The nonperturbative  $\beta$  function enters the interaction measure according to

$$(\epsilon - 3p)a^4 = -2 \frac{\partial(6/g^2)}{\partial \ln a} (P_T - P_0) - \frac{\partial(am_q)}{\partial \ln a} (\langle \bar{\psi}\psi \rangle_T - \langle \bar{\psi}\psi \rangle_0). \quad (3.115)$$

Here  $P_{T,0}$  denote the average plaquette expectation values at temperature  $T$  or zero, respectively. They are calculated as

$$P_{T,0} = \frac{1}{2N_\sigma^3 N_\tau} \frac{\partial \ln Z_{T,0}(am_q, 6/g^2)}{\partial(G)}, \quad (3.116)$$

where  $\langle \bar{\psi}\psi \rangle_{T,0}$  are the corresponding light-quark condensates,

$$\langle \bar{\psi}\psi \rangle_{T,0} = \frac{1}{N_\sigma^3 N_\tau} \frac{\partial \ln Z_{T,0}(am_q, 6/g^2)}{\partial(am_q)}. \quad (3.117)$$

The pressure  $p$  is calculated by integrating either of the two relations

$$\frac{p}{T^4} = 2N_\tau^4 \int_{\text{cold}}^{6/g^2} d(6/g'^2) [P_{N_\tau}(6/g'^2, am_q) - P_0(6/g'^2, am_q)], \quad (3.118)$$

[cf. Eq. (3.61) in Sec. III.B.2] or

$$\frac{p}{T^4} = N_\tau^4 \int_{\text{cold}}^{a \cdot m_q} d(am_q)' [\langle \bar{\psi}\psi(6/g^2, (am_q)') \rangle_T - \langle \bar{\psi}\psi(6/g^2, (am_q)') \rangle_0]. \quad (3.119)$$

The lower limit for the integration should be chosen so that the contribution of the integrand is negligible. Here the subscript ‘‘cold’’ stands for the cold symmetric lattice at zero temperature. The energy density is then obtained from Eqs. (3.118) and (3.119) combined with Eq. (3.115). The results are shown in the two upper curves of Fig. 14 for two light bare-quark masses. The lower curves denote three times the pressure. Although there is no transition in the strict sense, there is still a rapid rise in the energy density over a temperature interval of the order of 10 MeV around a temperature of 150 MeV. A slower rise in the pressure is also seen. A shift of the transition interval to higher temperatures is observed for lower bare-quark masses. This qualitative feature seems to be at odds with observations in effective models, where finite light-quark masses delay the melting compared to the chiral limit; see Secs. IV.A.2 and IV.A.4.

The simulations of Blum *et al.* (1995a) were performed for temporal extensions of  $N_\tau=4$  or  $N_\tau=N_\sigma$  with  $N_\sigma=8, 12$ , or  $16$ . From the discussion of finite-size effects in case of the pure  $SU(N)$  gauge theories (Secs. III.B.1 and III.B.2) we expect that the temporal extent of  $N_{\tau=4}$  will not be large enough for us to interpret the

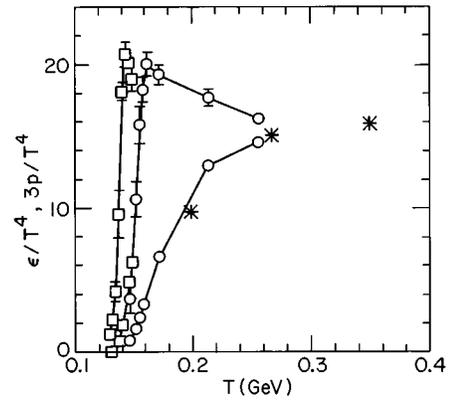


FIG. 14. Energy density (upper two curves) and three times the pressure (lower curve) vs temperature—ordinate quantities are normalized to  $T^4$  and are dimensionless—in the staggered fermion scheme of QCD with two light flavors [ $a \cdot m_q=0.025$  (octagons) and  $a \cdot m_q=0.1$  (squares)]. From Blum *et al.* (1995a).

lattice results as continuum physics. An extrapolation to the high-temperature region and a comparison with perturbative calculations would be a good test which has not yet been carried out. Thus the equation of state for two-flavor QCD is still in an exploratory stage.

## 7. Simulations with Wilson fermions

To confirm the conclusions of the last sections within the lattice approach, investigations using the Wilson formulation are indispensable. Staying within the same scheme it would be difficult to control the approximations of the desired continuum symmetries.

In the staggered fermion formulation, the chiral limit can be obtained by varying the bare mass  $ma$  in the Lagrangian to smaller values and extrapolating  $ma=0$  in the end. In contrast, the chiral limit in the Wilson fermion formulation must be determined as a one-dimensional submanifold  $\kappa_c(\beta)$  in the two-dimensional  $(\beta, \kappa)$  plane, where  $\beta=6/g^2$  denotes the inverse gauge coupling and  $\kappa$  the hopping parameters.  $\kappa_c(\beta)$  is the line of critical hopping parameters which characterize the chiral limit. The chiral limit can be defined as the vanishing of the pion mass on zero-temperature lattices. Another possibility is to determine  $\kappa_c(\beta)$  by the location of zeros in the fermion determinant. The two definitions are in general not equivalent, as the former involves an average over many gauge-field configurations, while the latter does depend on the configuration. Other definitions of  $\kappa_c$  on *finite-volume* lattices have been proposed by Bochicchio *et al.* (1985), Iwasaki *et al.* (1989) and Bitar *et al.* (1991).

In the following we discuss the QCD transitions in a space of four parameters: the number of flavors  $N_f$ , the number of time slices  $N_\tau$ , the coupling  $\beta=6/g^2$ , and the hopping parameter  $\kappa$ . The line of finite-temperature phase transitions/crossover phenomena from the confinement/chiral symmetry broken phase to the deconfinement/chiral symmetric phase will be denoted by  $\kappa_T(\beta)$ , while  $\kappa_d(\beta)$  stands for a line of bulk decon-

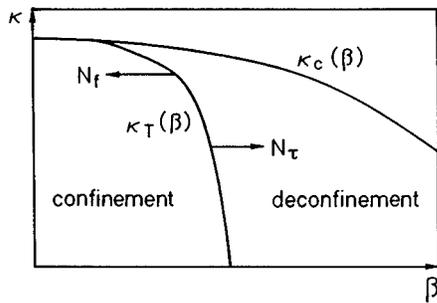


FIG. 15. Line of critical hopping parameters  $\kappa_c$  and thermal transition line  $\kappa_T$  vs  $\beta=6/g^2$ . The arrows indicate the shift of  $\kappa_T(\beta)$  towards weaker couplings for increasing  $N_\tau$  and stronger couplings for increasing  $N_f$ . The qualitative shape of the curves remains unchanged for  $N_f \leq 6$ .

finement transitions (further explanations follow below). It may be useful to visualize the change of parameters by a look at Fig. 15. Figure 15 indicates the shift of the thermal transition line  $\kappa_T(\beta)$  between the confinement and deconfinement phases as a function of  $N_\tau$  and  $N_f$ . Larger  $N_\tau$  shifts the transition towards the continuum limit. Note that  $\kappa_T(\beta)$  hits  $\kappa_c(\beta)$  only at rather strong couplings. The qualitative features hold for  $N_f \leq 6$ .

#### a. Difficulties with Wilson fermions

Fukugita *et al.* (1986) were the first to point out certain obstacles that may hamper the investigation of the confinement phase in the chiral limit with Wilson fermions. For  $N_f=4$ ,  $N_\tau=3$ , and spatial lattice sizes of  $5^3$  they performed simulations for  $3.0 < \beta < 6.0$  and  $0 < \kappa < 0.25$ . They defined the transition region of the deconfinement transition as the range of  $\beta$  over which the Polyakov loop and the gluon energy density varied rapidly. The result was that the transition line  $\kappa_T(\beta)$  did not meet the chiral line  $\kappa_c(\beta)$  down to  $\beta \sim 3.5$ . The lines were running almost parallel to each other to the strong-coupling regime without crossing. The warning of Fukugita *et al.* was that the line of critical hopping parameters may remain in the high-temperature phase for any finite lattice size. If this is indeed the case, it would be impossible to describe the confinement phase in the chiral limit within the Wilson formulation. The warning was extended to spectroscopic calculations with Wilson fermions, which might reflect high-temperature behavior at  $\kappa_c$ , if the spatial lattice size were not sufficiently large.

Results of Bitar *et al.* (1991) for  $N_f=2$  Wilson fermions on  $4 \times 8^3$  lattices pointed in the same direction. The hopping parameter was varied over the range  $0.12 \leq \kappa \leq 0.19$ , while the coupling region was  $4.5 < \beta < 6.0$ . The phase-transition line  $\kappa_T(\beta)$  was determined in this range, i.e. the transition/crossover coupling  $\beta_T$  lay between 5.12 and 5.13 for  $\kappa=0.17$ . If the chiral limit is defined as the vanishing of the pion mass on zero-temperature lattices, the chiral limit is nowhere reached in the low-temperature phase on the  $4 \times 8^3$  lattice in the above parameter range. The conclusion is, if

the line  $\kappa_T(\beta)$  ever reaches the chiral limit line  $\kappa_c(\beta)$ , it must be in the strong-coupling regime for these small lattice sizes.

Thermodynamics with Wilson fermions was reinvestigated by Iwasaki, Kanaya, Sakai, and Yoshié (1992, 1993). They varied the number of flavors between 2 and 18. The number of time slices was extended from  $N_\tau=4$  to  $N_\tau=6,8$  with spatial lattices of  $8^2 \times 10$  and to  $N_\tau=18$  with  $N_x N_y N_z = 18^2 \times 24$ . We distinguish between their results for  $N_f \geq 7$  and  $N_f < 7$ .

$N_f < 7$ . In the strong-coupling limit ( $\beta=0$ ) quarks were confined in the chiral limit. For  $N_f=6$  the critical hopping parameter was given as  $\kappa_c=0.25$ . The line  $\kappa_c(\beta)$  was defined as the values of  $\kappa$  where  $m_\pi^2$  vanished in the confinement phase with a linear extrapolation in terms of  $1/\kappa$ . The confining behavior was concluded from the fact that the number of iterations for the quark matrix inversion  $N_{\text{inv}}$  exceeded 10 000 for  $N_f=6$  with  $N_\tau=4$  at  $\kappa_c$ , while it was of  $O(100)$  for  $N_f \geq 7$  under the same conditions otherwise. The large value of  $N_{\text{inv}}$  was attributed to the existence of zero eigenvalues of the quark matrix.

The most interesting question concerns the existence of a crossing point of the  $\kappa_T(\beta)$  and the  $\kappa_c(\beta)$  lines at some  $\beta > 0$ . In the simulations of Iwasaki, Kanaya, Sakai, and Yoshié (1992) the answer is positive. For  $N_f=2$  and  $N_\tau=4$  the crossing point  $\beta_{cT}$  occurs at  $3.9 \leq \beta_{cT} \leq 4.0$  with  $\kappa_c \sim 0.222$ . The pion mass is consistent with zero at  $\beta_{cT}$ . The chiral transition is of second order or a crossover phenomenon. For  $N_f=6$  and  $N_\tau=4$ , the crossing point lies in the range  $0.2 \leq \beta_{cT} \leq 0.3$  with  $\kappa_c \sim 0.25$ , the chiral transition is of first order, and the value of  $m_\pi^2$  depends on the initial configuration.

For  $N_f=2$  and  $N_\tau=18$  the crossing point  $\beta_{cT}$  is only slightly shifted towards the continuum region,  $\beta_{cT} \sim 4.5-5.0$ . The unwelcome message is that one has to go to temporal extents  $N_\tau > 18$  to obtain a confining chiral limit closer to the continuum limit (i.e., for  $\beta > 5.0$ ).

$N_f \geq 7$ . Here we have to distinguish between a line  $\kappa_d(\beta)$  of bulk transitions that separates the chiral limit  $\kappa_c(\beta)$  from the confinement region, and the line  $\kappa_T(\beta)$  of finite-temperature transitions. In the strong-coupling limit ( $\beta=0$ ) the phase of deconfinement and chiral symmetry restoration is realized for  $\kappa > \kappa_d$  ( $\beta=0$ ). The transition is stable under an increase of  $N_\tau$  to  $N_\tau=18$ . Therefore the transition is called a bulk transition; see the discussion of bulk transitions for  $N_f=8$  Kogut-Susskind fermions in Sec. III.C.4. The  $\kappa_d(\beta)$  line of bulk transitions extends to  $\beta > 0$ . In addition there is the  $\kappa_T(\beta)$  line at  $\beta > 0$ , which reaches the  $\kappa_d(\beta)$  line without crossing the chiral limit. For  $N_f=7$  and 12, the quark mass is  $O(1)$  in units of  $a^{-1}$  at  $\kappa_d$ . If this behavior persists in the continuum limit, confinement is lost for  $N_f \geq 7$ . Recall that asymptotic freedom is lost in QCD if  $N_f \geq 17$ .

Peculiar behavior of the phase structure was found by the MILC collaboration (Blum *et al.*, 1994) for  $N_f=2$  at  $N_\tau=4$  or 6. At  $N_\tau=4$  the transition was smooth for light and heavy quark masses, but sharp for intermediate

masses with  $\beta \sim 5.0$ . This is at odds with the general expectation that the strength of the crossover should monotonically decrease with increasing masses. At  $N_\tau = 6$  a first-order phase transition with clear two-state signals was seen in simulations at  $\kappa = 0.17, 0.18,$  and  $0.19$ . It was interpreted as a bulk transition located in the vicinity of the thermal crossover (Blum *et al.*, 1994). Fortunately the lattice artifacts seem to disappear in a more recent simulation of the MILC collaboration (Bernard *et al.*, 1995b) at larger  $N_\tau$  ( $N_\tau = 8$ ) at  $6/g^2 = 5.3$  over a range of  $\kappa$  values up to  $\kappa_c \sim 0.168$ . The thermal crossover is now shifted to a slightly smaller  $\kappa_T \sim 0.167$ , in qualitative agreement with Fig. 14.

### b. 2+1 flavors and critical quark masses

The Tsukuba group (Iwasaki, Kanaya, Kaya, *et al.*, 1995) performed simulations on  $8^2 \times 10 \times 4$  and  $12^3 \times 4$  lattices for the physically interesting case of two light and one heavier flavor. The lattice parameters were translated to physical quark masses of  $m_{u,d} \sim 0$ ,  $m_s \sim 400$  MeV and  $m_{u,d} = m_s \sim 150$  MeV. The hopping parameters were chosen as  $\kappa^{(u,d)} = \kappa_c$  for the light quarks, while  $\kappa^{(s)}$  followed from measurements of  $m_\rho a, 2m_q a$  as functions of  $(1/\kappa - 1/\kappa_c)$  with  $\kappa_c(\beta)$  referring to the two-flavor case. The results of  $m_\rho a$  as function of  $N_f, N_\tau, \beta$  were used to set the scale. If the physical mass  $m_\rho$  of 770 MeV is identified with  $[m_\rho(\kappa_c)a]/a$ , it follows that there should be a lattice spacing of  $a \sim 0.8 \text{ GeV}^{-1}$  for  $\beta \leq 4.7$ . Once  $a$  ( $\text{GeV}^{-1}$ ) is known, one can find  $m_q a$  and the associated values of  $(1/\kappa^{(s)} - 1/\kappa_c)$  corresponding to  $m_q \sim 0$ ,  $m_s \sim 400$  MeV and  $m_{u,d,s} \sim 150$  MeV.

One should recall that  $m_\rho a$ , i.e., the  $\rho$  mass in lattice units, was measured at rather strong couplings ( $N_\tau = 4$ ) on small volumes with varying quark mass input  $m_q a$ . Keeping  $m_q$  fixed at 770 MeV assumes that  $m_q$  (MeV) is independent of  $m_q$  (MeV), which it is not. Furthermore the value of  $a$  ( $\text{GeV}^{-1}$ ) obtained from  $m_\rho a$  and  $m_\rho$  (GeV) at such strong couplings is certainly not universal, i.e., independent of the lattice observable. Thus the translation from lattice parameters  $\beta, \kappa^{(u,d,s)}$  to physical parameters  $a$  ( $\text{GeV}^{-1}$ ) and  $m_{u,d,s}$  (MeV) should be made with care.

The time histories of the plaquette expectation value show two-state signals with states depending on the initial configuration for both values of  $m_s$  ( $m_s = 150$  and  $400$  MeV). Future simulations with larger temporal extensions at weaker couplings will be needed to show whether the first-order signals survive the continuum limit or share the fate of the signals of the  $N_f = 2$ -simulations at  $N_\tau = 4, 6$  (Blum *et al.*, 1994), fading away on larger lattices. At strong couplings the fermion doublers in the Wilson formulation are still too light to decouple. It is known that the strength of the first-order transitions increases with the number of light flavors. Thus the first order at large  $m_s$  ( $\sim 400$  MeV) might be an UV artifact of the strong couplings caused by the small number of time slices.

At present the first-order behavior is in disagreement with the lattice results of the Columbia group (see Sec. III.C.5) and with results in effective continuum models (Meyer-Ortmanns and Schaefer, 1996; see Sec. IV.A.4).

It is of interest to compare the estimates for the *critical* quark masses up to which a clear first-order signal is seen with results obtained for staggered fermions and in effective models for QCD. For three degenerate flavors, the Wilson fermion scheme gives  $m_{u,d,s}^{\text{crit}} \geq 140$  MeV or  $m_{u,d,s}^{\text{crit}} a \geq 0.175(2)$ , as compared to 12–38 MeV for staggered fermions (Sec. V.C.5). For nondegenerate flavors it is  $m_{u,d}^{\text{crit}} \leq 3$  MeV,  $m_s^{\text{crit}} \leq 54$  MeV in the  $SU(3) \times SU(3)$  linear sigma model (see Sec. IV.A.4 below).

To sum up, although the translation to physical mass units may suggest continuum physics, it would be premature to call the lattice result of a first-order chiral transition for physical quark masses continuum physics of the “real world.”

### c. Improved actions

The experience with Wilson fermions in the original formulation has shown that the simulations are either inconclusive, due to lattice artifacts at strong couplings, or expensive due to volumes that should be even larger than for staggered fermions to avoid lattice artifacts. Thus it is natural to attempt an approach with an *improved action*. Improved actions are supposed to accelerate the approach to the continuum limit. One such choice has been adopted by the Tsukuba group (Iwasaki, Kanaya, Kärkkäinen, *et al.*, 1994; Iwasaki, Kanaya, Sakai, and Yoshié, 1995). So far the improvement concerns the gauge part of the full QCD action, while the fermionic part remains in Wilson’s formulation. The form of the improved pure gauge action was proposed by Iwasaki (1983). Besides the usual plaquette term it contains a  $1 \times 2$  loop with coefficients determined by a block-spin renormalization-group analysis. The preliminary results for full QCD with the renormalization-group improved action look promising. The improved gauge part seems to reduce (or remove) the lattice artifacts for  $N_f = 2$  and  $N_\tau = 4$  (6) observed by the MILC collaboration (mentioned above).

A quantitative measure of acceleration towards the continuum limit is obtained from the difference in (inverse) couplings  $\beta$ , which lead to approximately the same lattice spacing  $a$ . The improved action needs only  $\beta = 2.0$ , while the standard action needs  $\beta \sim 5.0$  for  $a^{-1} = 1.01 \text{ GeV}$  (if the  $\rho$  mass is calculated on an  $8^3 \times 16$  lattice and used as input to set the physical scale).

It is not surprising that, for small lattices without improved actions, the Wilson and staggered fermion schemes give different results. One manifestation of this discrepancy is given by an estimate of  $T_c$  in physical units for two flavors and four time slices. For Wilson fermions  $T_c$  is estimated as  $221 \pm 3$  MeV, whereas  $T_c = 142 \pm 6$  MeV for staggered fermions (Bitar *et al.*, 1991). A more recent estimate of  $T_c$  from Wilson fermions leads to  $T_c \sim 152_{-14}^{+18}$  MeV (Bernard *et al.*,

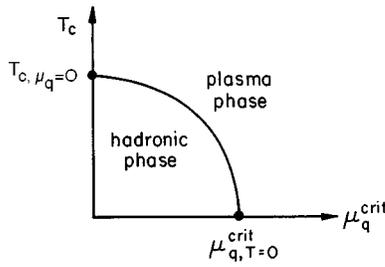


FIG. 16. Qualitative temperature dependence of the critical chemical quark potential  $\mu_q^{\text{crit}}$ .

1995b) as compared to  $140 \text{ MeV} \leq T_c \leq 160 \text{ MeV}$  for staggered fermions (DeTar, 1995). Future simulations will show whether other quantities show the same tendency for both fermion formulations to approach the same continuum behavior.

#### D. QCD at finite baryon density

The most convenient way of describing quark or nuclear matter at high density is to introduce a chemical potential  $\mu_q$  for quarks. From asymptotic freedom one would naively expect that chiral symmetry is restored and quarks are deconfined above a critical value of  $\mu_q$ ,  $\mu_q^{\text{crit}}$  and above a critical temperature  $T_c$ . A critical line  $(T_c, \mu_q^{\text{crit}})$  will separate the plasma and the hadron phase in a  $(T, \mu)$  diagram. The shape is indicated in Fig. 16. It is qualitatively reproduced in a bag model calculation at finite  $T$  and finite  $\mu_q$  (Cleymans *et al.*, 1986). Note in particular that chiral symmetry is always restored above the critical temperature for  $\mu_q=0$ ,  $T_{c, \mu_q=0}$ , and above the critical potential for  $T=0$ ,  $\mu_{q, T=0}^{\text{crit}}$ . In the bag model one considers an ideal gas of nucleons in the hadronic phase, and of quarks and gluons in the plasma phase. A real-valued non-negative energy difference at  $T=0$  and the critical chemical potential  $\mu_b^{\text{crit}}$  for baryons is then obtained if  $\mu_b^{\text{crit}}$  satisfies

$$M \leq \mu_b^{\text{crit}} \leq 3M/(2\sqrt{2}), \quad (3.120)$$

where  $M$  denotes the nucleon mass. The chemical potentials for quarks and baryons are related at the transition point via

$$3\mu_q^{\text{crit}} = \mu_b^{\text{crit}}. \quad (3.121)$$

Conceptually  $\mu_{q,b}^{\text{crit}}$  should be distinguished from threshold values  $\mu_{q,b}^t$ . In the bag model the threshold value for  $\mu_b$  is given by the nucleon mass  $M$ ,

$$\mu_b^t = M, \quad (3.122)$$

because a real-valued, nonvanishing baryon number density is only obtained for  $\mu_b > M$ . Note that  $\mu_b^{\text{crit}}$  within the bound of Eq. (3.120) is always close to  $\mu_b^t$ , so that the finite- $\mu$  phase transition at  $\mu_b^{\text{crit}}$  may not be dis-

tinguishable from a threshold effect at  $\mu_b^t$  in practical calculations. Threshold effects are in fact seen on the lattice (see the summary of results below). We turn now to the lattice approach to finite- $\mu$  (and finite- $T$ ) transitions.

A naive translation of the continuum expression onto the lattice leads to quadratic divergences of the internal energy density  $\epsilon$  in the continuum limit. The solution of this earlier difficulty is provided by introducing a chemical potential for quarks  $\mu a$  into the fermion matrix as  $e^{\mu a}$ , multiplying the forward links in the time direction, and  $e^{-\mu a}$  multiplying the backward links (Hasenfratz and Karsch, 1983).

For simplicity one may think of this as performing lattice simulations at finite density in the quenched approximation. It turns out that the quenched approximation leads to results that contradict intuitive expectations. Physical observables should be independent of  $\mu$  as long as  $\mu$  is smaller than some threshold value  $\mu_c$ . The threshold is related to the threshold for baryon production. In the confining phase, baryons made of three quarks will be produced that are bound together with an effective chemical potential  $3\mu$ . Once  $3\mu$  exceeds the value of the baryon mass, a finite baryon density will populate the system. In the quenched approximation the threshold comes out as  $\mu_c a = m_\pi^2/2$  even in the case of SU(3), where  $m_\pi$  is the pion mass. The threshold should depend on the baryon mass, and the baryon mass should stay finite in the chiral limit.

For a finite quark mass the chiral condensate shows some threshold behavior in the sense that  $\langle \bar{\psi}\psi \rangle$  stays constant for awhile and then goes to zero with increasing  $\mu$ , but  $\langle \bar{\psi}\psi \rangle$  vanishes in the chiral limit as soon as  $\mu > 0$ . This behavior also is at variance with the expectation. A critical density should be exceeded before chiral symmetry is restored.

Davies and Klepfish (1991) have argued that the strange threshold behavior  $a\mu_c = m_\pi^2/2$  is really an artifact of the quenched approximation and not of the finite volume or the strong coupling, as one might have suspected. Their argument is based on the fact that the gauge loops contributing to the pion propagator are unaffected by the chemical potential in the quenched approximation. The failure of the quenched approximation is shown for gauge couplings  $\beta = 6/g^2$  ranging from 0.0 to 6.2.

If the full fermion determinant  $\det M$  is included in lattice Monte Carlo simulations, the problem starts for SU( $N$ ) gauge theories with  $N \geq 3$ . In this case the determinant is a complex number for  $\mu \neq 0$ . Therefore the factor  $\det M \exp(-S_g)$  (where  $S_g$  denotes the pure gauge part of the action) can no longer be used as a probability for generating configurations in a Monte Carlo simulation as is usually possible when  $\det M$  is a positive real number. Nevertheless one can generate gauge-field configurations with some probability  $P(U)$ , which is frequently chosen as  $P(U) = |\det(M)| \exp(-S_g)$ . The expectation value for an observable  $O$  is then calculated as

$$\langle O \rangle = \left[ \int DU O \det(M) e^{-S_g} \right] / \left[ \int DU \det(M) e^{-S_g} \right] \\ = \langle O e^{i\phi} \rangle_P / \langle e^{i\phi} \rangle_P, \quad (3.123)$$

where  $\langle \dots \rangle_P$  denotes the average over configurations with respect to the measure  $DUP(U)$  and  $e^{i\phi}$  is the phase of the determinant. At  $\mu=0$  and  $\mu=\infty$ ,  $\langle e^{i\phi} \rangle_P=1$ . Therefore a simulation according to the above prescription is feasible for small and large values of  $\mu$ . For intermediate values of  $\mu$ ,  $\langle e^{i\phi} \rangle$  is very small. The phase  $\phi$  fluctuates violently from configuration to configuration. A report on how far one can push this method from the extreme values  $\mu=0$  and  $\mu=\infty$  can be found in Toussaint (1990); see also Vladikas (1988).

One attempt to minimize the large fluctuations of  $\phi$  is found in a detailed study of the grand canonical partition function (Gibbs, 1986; Barbour and Sabeur, 1990). It can be shown that the characteristic polynomial for the fermion propagator matrix is an expansion of the grand canonical partition functions for fixed quark number. Here it is nontrivial to measure the expansion coefficients  $\langle a_n \rangle$ . Monte Carlo measurements of the canonical partition functions within this expansion have been performed in an SU(3) gauge theory by Barbour *et al.* (1988), Barbour and Sabeur (1990), Barbour and Bell (1992), and Hasenfratz and Toussaint (1992) and are reviewed by Barbour (1992).

As an analytical alternative to the Monte Carlo simulations, lattice calculations have been performed at strong coupling in a  $1/d$  and  $1/g^2$  expansion (Bilić, Demeterfi, and Petersson, 1992; Bilić, Karsch, and Redlich, 1992). Here  $d$  denotes the spacetime dimension and  $g$  is the bare gauge coupling. The earlier pathological threshold behavior in Monte Carlo simulations (with a threshold proportional to  $m_\pi$  and vanishing for  $m_q \rightarrow 0$ ) does not, fortunately, occur in the analytical approach. Compared to the  $g=\infty$  limit, at finite coupling the threshold increases with  $1/g^2$ , while the baryon mass  $m_b$  decreases. The difference between the threshold and  $m_b$  becomes smaller at weaker couplings. This is to be expected, as the difference is a measure of the nuclear binding energy, which is large at strong couplings. Note that the threshold value is different from the baryon mass in contrast to Eq. (3.122), which was derived in an ideal-gas picture for nucleons.

Above a temperature of the order of 220 MeV, chiral symmetry is restored for any  $\mu_b \geq 0$ , in qualitative agreement with Fig. 16. Below this value, the chiral transition is of first order, but occurs at larger critical chemical potentials for increasing  $6/g^2$ , as one would expect from the considerations above. For  $\mu_b < \mu_b^t$  one is in the chiral symmetry broken QCD vacuum without nuclear matter (Bilić, Demeterfi, and Petersson, 1992).

A more detailed study of the threshold effect of zero temperature and finite  $\mu$  was performed by Bilić, Karsch, and Redlich (1992). The threshold behavior turned out to be flavor independent, in contrast to the strong flavor dependence of the ( $\mu=0$ , finite- $T$ ) transition. Apparently a threshold effect need not be a phase

transition. At strong coupling, however, the threshold effect seemed to coincide with the finite- $\mu$  chiral phase transition. Again the threshold value at strong coupling lay somewhat below the nucleon mass. The coincidence of the threshold effect with the finite- $\mu$  phase transition was likely an artifact of strong couplings. At the threshold value  $\mu^t$ , the baryon number density jumped from zero to the maximal occupation of lattice sites.

In the continuum limit at zero temperature one may expect to see a threshold effect at  $\mu^t$  separate from the phase transition at  $\mu^{\text{crit}} > \mu^t$ . In that case the threshold effect corresponds to a “transition” from the QCD vacuum to nuclear matter with broken chiral symmetry, and the phase transition restores chiral symmetry above the critical  $\mu_{q,b}^{\text{crit}}$ .

So far the lattice results confirm some qualitative features of Fig. 16, but they are seriously plagued with artifacts of strong couplings in the analytical approach and of intermediate couplings in the Monte Carlo approach.

Mean-field calculations of QCD at finite baryon density have been performed in a number of effective models for QCD in the *continuum*. See for example, Asakawa and Yazaki (1989), Lutz *et al.* (1992), or the review by Hatsuda (1992). Some of these calculations will be also mentioned in Sec. IV.

## IV. EFFECTIVE MODELS IN THE CONTINUUM

### A. Models for quark degrees of freedom

In Secs. II and III we described effective models on the lattice, where the  $Z(N)$  Potts model was of particular importance. As a common feature these models could be derived from lattice QCD in some coupling or temperature limit or in a renormalization-group approach. They suffer from the same shortcomings as the lattice approximation to the original model (full QCD). Their relevance for the continuum limit has first to be established. In this section we discuss effective models in the spacetime continuum and offer some remarks on dual Ginzburg-Landau models (on the lattice) in Sec. IV.B.2. Studies of the phase structure of QCD are frequently performed in a mean-field approach. The methods we have selected for more detailed discussion in this section go beyond the mean-field level in various ways.

We begin by illustrating the power of the renormalization group with the work of Wilczek (1992) and Rajagopal and Wilczek (1993a, 1993b), who consider a three-dimensional  $SU(2) \times SU(2)$  linear sigma model as a model for the chiral phase-transition region. In Sec. IV.A.2 we outline the approach of chiral perturbation theory, which is well established as a description of low-energy and low-temperature QCD. We discuss the results of Gerber and Leutwyler (1989). The leading term in a series of actions for chiral perturbation theory is a nonlinear  $SU(2) \times SU(2)$  sigma model in four dimensions. Although chiral perturbation theory fails to describe the transition region (and thus cannot predict the order of the phase transition), it allows an estimate of the latent heat in the case of a (hypothetical) first-order

transition (Leutwyler, 1992; see Sec. IV.A.3). In Sec. IV.A.4 we discuss the mass sensitivity of the chiral transition and the equation of state for physical meson masses (Meyer-Ortmanns and Schaefer, 1996).

### 1. QCD and dysprosium

In this section we discuss the work of Wilczek (1992) and Rajagopal and Wilczek (1993a, 1993b), who offer an effective description of QCD that is supposed to be valid just in the vicinity of the chiral transition. Their approach is based on three hypotheses.

(1) In the limit of two massless flavors, QCD is in the static universality class of an  $N=4$  Heisenberg ferromagnet when all other quark masses are infinite and  $T \sim T_c$ .

(2) In the limit of two massless flavors, QCD is in the static universality class of the  $\phi^6$  Landau-Ginzburg model when the strange-quark mass is near a (tricritical) value  $m_s^*$ , the remaining quark masses are infinite, and  $T \sim T_c$ .

(3) QCD with two massless flavors is in the dynamical universality class of an  $O(4)$  antiferromagnet.

These conjectures may be regarded as working hypotheses for lattice simulations of the chiral transition. Before we describe their predictive power, let us comment on the two basic assumptions.

(i) QCD is well approximated by the limit of two massless flavors ( $m_u = m_d = 0$ ).

(ii) The lightest excitations in the spectrum at zero temperature maintain their dominant role up to the transition region ( $T \sim T_c$ ). The transition of two-flavor QCD is expected to be of second order.

Possible objections could be the following.

(a) Even the lightest quark flavors are not massless. Their finite values lead to a pion mass of  $m_\pi \sim 135$  MeV. Unless the quark masses correspond to critical values (which is not likely), a second-order transition amounts to an inadequate idealization.

(b) From lattice calculations one knows that the chiral transition is of first order for four flavors, if the masses are not too large. For two light and one heavier flavor the transition is replaced by a crossover phenomenon.

(c) It is unlikely and implausible that a theory as intricate as QCD can be reduced to a model as simple as a scalar  $O(4)$  model.

(d) Finally the pions lose their dominant role in the transition region.

We shall argue against the first three objections and postpone to Sec. IV.A.2 arguments in favor of (d).

Reply to (a): The light quark masses have nonvanishing values, but their values are small compared to the energy scale of the critical temperature. Thus it makes sense to consider  $m_u, m_d$  as perturbations around the chiral limit, where the transition should be of second order, if  $N_f=2$ . The renormalization-group approach can account for small-finite mass values. Their “perturbing” effect on the critical behavior in the chiral limit can be parametrized with critical indices. Thus there is a predictable parametrization of the deviation from the

idealized limit. We recall from Sec. II that the mass plays a role comparable to that of the scaling field of the inverse volume in a renormalization-group analysis. Power-law singularities of a second-order transition will be rounded due to the finite volume and due to a finite mass. As we know from Sec. II, the rounding is specific for the second order, if the volume is sufficiently large. The deviation from the  $L = \infty$  limit can be predicted in a well-known way.

Similarly there is a good chance that the extrapolation from finite masses to the chiral limit is under control. It is under control if the deviations can be parametrized with the critical indices of a second-order transition. For a check of this assumption it would be sufficient to measure certain correlations (e.g., the specific heat) at different small, but finite, quark masses and compare the change in the rounding effects with the predictions of the renormalization-group analysis.

Reply to (b): In view of the physical values for the current quark masses, four light flavors are certainly less realistic than two. Typical signatures of a first-order transition (a nonvanishing latent heat, hysteresis effects, and abrupt changes in the order parameters) are absent for two flavors. Support for a second-order transition in two-flavor QCD is taken from lattice calculations of Gottlieb *et al.* (1987, 1989); Brown *et al.* (1990a, 1990b); Gottlieb (1991); Bernard *et al.* (1992). The rise in the specific heat and the critical slowing down are compatible with a second-order transition. As argued in Sec. III, lattice results are not yet fully conclusive. Hence they do not *prove* any of the conjectures about two-flavor QCD.

Reply to (c): Note that the hypotheses (1)–(3) refer to QCD only in the critical region ( $T \sim T_c$ ). There it may well be that the complicated substructure of QCD is not important for the transition dynamics; in which case one is free to replace QCD by the simplest model belonging to the same universality class. In condensed-matter physics many examples are known for which universality is used in the same spirit, such as the critical behavior of a binary mixture like isobutyric acid plus water, which shares the universality class of an Ising model in three dimensions.

What, then, is the right choice of universality class? In the case of two-flavor QCD, Wilczek’s proposal is the universality class of an  $O(4)$  Heisenberg ferromagnet. This proposal has been recently questioned by Kocic and Kogut (1995), claiming that QCD with its quark bilinear composite scalar mesons does not belong to the same universality class as sigma models with their fundamental scalars. Their warning is based on a study in a three-dimensional Gross-Neveu model that exhibits mean-field scaling rather than the expected two-dimensional Ising universality. Let us see how one is led to the idea of an  $O(4)$  Heisenberg ferromagnet in a natural way, and where the warnings of Kocic and Kogut would change the line of arguments.

Landau’s free energy was introduced as a framework for discussing the phase structure; see Sec. II.A, Eq.

(2.2). It is constructed as a power series in the order-parameter field. Landau's theory corresponds to a mean-field approximation. In general it leads to wrong predictions of characteristic singularities (e.g., the correlation length at  $T_c$ ). The physical reason are large scale fluctuations in the order-parameter field, although the average magnitude of the order parameter is small.

The theoretical breakthrough came with Wilson's renormalization-group concept (Wilson and Kogut, 1974). It leads to a systematic way of constructing an action in terms of the relevant degrees of freedom at  $T_c$  and answers the question of why the resulting action is representative in some sense. The action is constructed as a limiting (fixed-point) theory after a number of renormalization-group transformations. If the theory is exactly *scale invariant* at  $T_c$ , such a limiting theory exists. Scale invariance implies that models at different scales ( $|r| \ll \xi$ , i.e., scales much smaller than the correlation length  $\xi$  but larger than the microscopic scales, share the leading singularity structure.

In case of the chiral transition, these renormalization-group steps are not explicitly performed. The limiting theory in terms of pion fields (and their parity partners) is argued to arise out of such a procedure. More precisely only the zero modes of these fields are assumed to survive the iterated renormalization-group steps. The pions are the lightest modes at zero temperature, and the zero modes are the only modes that do not acquire a mass contribution  $\propto (2\pi nT)$  from the Matsubara sum. Thus the renormalized mass parameter  $\mu^2(T)$  in the resulting action should be understood as an effective mass of the zero modes, which vanishes as  $T$  approaches  $T_c$  from above.

An essential outcome of the renormalization-group approach is an explanation of *universality*. Universality defines in which sense the limiting theory is representative of a whole class of models, belonging to the same *universality class*. Once the order-parameter field is specified, the fixed-point theory in terms of these fields depends only on the dimensionality and the symmetry, which is assumed to be broken or restored at the transition. Models with the same underlying symmetry, order-parameter fields, and singularity structure in thermodynamic functions define a universality class.

Thus it is sufficient to find an order-parameter field, construct an action in terms of this field, and restrict the allowed terms by the requirement of chiral symmetry. The order-parameter field for the chiral transition should at least contain the pion multiplet. For two massless flavors, the QCD Lagrangian is invariant under the  $SU(2)_L \times SU(2)_R \times U(1)_B$  symmetry of independent  $SU(2)$  rotations of left- and right-handed fields and the vector baryon number symmetry. At the phase transition, the symmetry is assumed to be broken to  $SU(2)_{L+R} \times U(1)_{L+R}$ . One choice of an order parameter for the chiral transition is the quark bilinear

$$M_j^i = \langle \bar{q}_L^i q_{Rj} \rangle, \quad (4.1)$$

transforming under  $SU(2)_L \times SU(2)_R$  according to

$$M \rightarrow U^+ M V, \quad (4.2)$$

where  $U$  and  $V$  represent independent unitary transformations of the left- and right-handed quark fields. In the following the quark substructure of  $M_j^i$  will be disregarded. [This is the main point that is questionable according to Kocic and Kogut (1995)].

As long as  $M$  are general complex  $2 \times 2$ -matrices, the Lagrangian has too much symmetry [ $U(2) \times U(2)$ ]. A restriction of  $M$  to an  $SU(2)$  representation removes the additional  $U(1)$  symmetry. A possible choice for the  $SU(2)$  representation is  $O(4) \approx SU(2) \times SU(2)$ , where  $M$  is parametrized in terms of four real parameters  $(\sigma, \boldsymbol{\pi}) = \phi^a$  ( $a=1, \dots, 4$ ),

$$M = \sigma + i \boldsymbol{\pi} \boldsymbol{\tau}, \quad (4.3)$$

where  $\boldsymbol{\tau}$  denote the Pauli matrices. Thus  $M$  contains the pion multiplet along with the scalar meson  $\sigma$ . The action in terms of  $\phi^a$ , which is invariant under  $SU(2) \times SU(2)$ , is given as

$$S = \int d^3x \left\{ \frac{1}{2} \partial^j \phi^a \partial_i \phi_a + \frac{\mu^2}{2} \phi^a \phi_a + \frac{\lambda}{4} (\phi^a \phi_a)^2 \right\}. \quad (4.4)$$

The action takes the same form as the Landau-Ginzburg free energy  $F$  in Sec. II.A. It could be identified with  $F$  if the path integral with  $S$  of Eq. (4.4) were evaluated in the mean-field approximation. Note that  $S$  is an action in terms of zero modes. The Euclidean time dependence of  $\phi^a$  has been dropped. Equivalently the  $n \neq 0$  Matsubara modes of the original four-dimensional theory are neglected. [The treatment of the  $SU(2) \times SU(2)$  sigma model in the following section, IV.A.2, will differ in this respect, and time dependence there will be retained].

Here the action Eq. (4.4) should be understood as an effective action for the chiral transition. It coincides with the familiar linear sigma model of Gell-Mann and Levy (1960) up to the absence of nucleons and the dimension 3. Nucleon and quark fields are both omitted, as they refer to "microscopic substructures," which are claimed to be irrelevant for the transition. Moreover the action (4.4) agrees with the action of an  $N=4$  Heisenberg ferromagnet, which is believed to model a magnetic transition in dysprosium (Malmström and Geldart, 1980). This is the reason why Wilczek calls dysprosium an "analog computer" for QCD and explains the title of this section.

The  $O(4)$  model in three dimensions has been extensively studied in statistical physics (Bervillier, 1976; Hohenberg *et al.*, 1976; Lipatov, 1977; Baker *et al.*, 1978). It is known to have a second-order transition with  $\mu^2(T) \rightarrow 0$  as  $T \rightarrow T_c$  from above. In the following we summarize the results for critical exponents, taken from these references and Wilczek (1992).

As usual we distinguish between the response to a stimulus in  $T$  and to a stimulus in  $h$ , where  $h$  is an external field (see Table II). In this context the meaning of the external field is that of nonvanishing current quark masses.

### a. Response to a stimulus in $T$

The response of the  $O(4)$  “magnet” to a deviation from  $T_c$  is measured by the specific heat  $c$  and the quark pair correlation  $\xi$  in the limit of strictly vanishing quark masses ( $m_\pi=0$ ). The singular behavior of the specific heat is characterized by the index  $\alpha$ , and is given by

$$c(T) \xrightarrow{T \rightarrow T_c^\pm} A_\pm |T - T_c|^{-\alpha_\pm} + \text{less singular terms} \quad (4.5)$$

with  $\alpha = -0.19 \pm 0.06$ . The behavior of the quark pair correlation is characterized by the index  $\nu$  according to

$$\xi \xrightarrow{T \rightarrow T_c^\pm} B_\pm |T - T_c|^{-\nu_\pm} + \text{less singular terms} \quad (4.6)$$

with  $\nu = 0.73 \pm 0.02$ , where  $\xi$  is determined by

$$-\ln \langle \bar{q}(x)q(x)\bar{q}(0)q(0) \rangle \rightarrow \frac{|x|}{\xi} \quad (4.7)$$

at large distances  $|x|$ . The  $\pm$  signs refer to the approach of  $T_c$  from temperatures  $T > T_c (+)$  or  $T < T_c (-)$ . The exponents  $\alpha_\pm$  and  $\nu_\pm$  do not depend on the approach of  $T_c$ . The amplitudes  $A_\pm$  and  $B_\pm$  are universal and are known, as well.

Note that  $\alpha$  is negative. Thus there is only a cusp in the specific heat, no true singularity in the infinite-volume limit. This is compatible with the definition of a second-order transition (see Sec. II.A), but not so common. Usually more than one of the second derivatives of the thermodynamic potential diverges with power-law singularities.

We recall from Table II that the critical index  $\beta$  determines the behavior of the order parameter  $\langle |\phi| \rangle$  as  $T$  approaches  $T_c$  in the massless limit. We have

$$\langle |\phi| \rangle \sim \left| \frac{T - T_c}{T_c} \right|^\beta \quad \text{for } T - T_c < 0. \quad (4.8)$$

with  $\beta = 0.38 \pm 0.01$ .

The results for the critical indices  $\alpha$ ,  $\nu$ , and  $\beta$  should be checked by lattice simulations of the condensate correlations Eq. (4.7) and the specific heat ( $T \partial S / \partial T$ ). In principle this check can be performed, although an extrapolation to the zero-mass limit and the struggle with critical slowing down are unavoidable.

### b. Response to a stimulus in $m$

The simplest mass term that can be added to the action and that is invariant under vector  $SU(2)$  transformations and discrete  $P$  and  $T$  transformations, is given by  $m\sigma$ , where  $m$  denotes the common quark mass. Table II in Sec. II tells us how we should translate the compressibility in a liquid/gas system and the susceptibility in a ferromagnet into order parameters for the chiral transition. The corresponding quantity is the variation of the condensate as a function of  $m$ ,

$$\frac{\partial \langle \bar{q}q \rangle}{\partial m} \rightarrow C_\pm |T - T_c|^{-\gamma_\pm} + \text{less singular terms.} \quad (4.9)$$

The approach to  $T_c$  is characterized by the critical exponent  $\gamma$  (with  $\gamma_+ = \gamma_-$ ), and the amplitudes  $C_\pm$  are universal.

The second exponent  $\delta$  concerns the response of the condensate to a change in  $m$  at  $T_c$ ,

$$\langle \bar{q}q \rangle \rightarrow m^{1/\delta} + \text{less singular terms, } T = T_c. \quad (4.10)$$

Here the results for the  $O(4)$  model are  $\gamma = 1.44 \pm 0.04$  and  $\delta = 4.82 \pm 0.05$ .

### c. The critical equation of state

Wilczek’s first hypothesis can be further used to predict the temperature dependence of the pion and the sigma masses in the vicinity of  $T_c$ . These functions can be calculated as a function of the underlying quark masses. The important relation is a critical equation of state, which includes the results for  $\gamma$ ,  $\beta$ , and  $\delta$  of the previous sections as special cases. An equation of state in a liquid/gas system can be formulated as a relation between  $p$ ,  $V$ , and  $T$ ,  $p = p(V, T)$ . In a ferromagnet such a relation translates to  $H = H(M, T)$  with the notations of Table II, while in QCD it reads  $m = m(\langle \bar{q}q \rangle, T)$ , and in the linear sigma model  $m = m(\langle \sigma \rangle = \langle |\phi| \rangle, T)$ . In the “magnetic” language of Brézin *et al.* (1973) the critical equation of state is given as

$$\frac{H}{M^\delta} = \kappa_1 g(\kappa_2 t |M|^{-1/\beta}) \equiv f(t/M^{1/\beta}), \quad (4.11)$$

where  $t = (T - T_c)/T_c$ ,  $g$  is a universal function, and  $\kappa_1, \kappa_2$  are nonuniversal constants. For the  $O(4)$  model the function  $f$  has been calculated by Brézin *et al.* (1973) to order  $\varepsilon^2$  in the  $\varepsilon$  expansion. Without derivation we state the expressions for  $m_\pi^2$  and  $m_\sigma^2$  in terms of the function  $f$  at small  $t$  and  $H(\equiv \langle \sigma \rangle)$  (Rajagopal and Wilczek, 1993a, 1993b):

$$m_\pi^2 = M^{\delta-1} f(x), \quad (4.12a)$$

$$m_\sigma^2 = M^{\delta-1} \left( \delta f(x) - \frac{x}{\beta} f'(x) \right) \quad (4.12b)$$

with  $x \equiv t/M^{1/\beta}$ . Note that the relations hold at finite temperature. Thus we have to specify which mass is meant by  $m_\pi^2, m_\sigma^2$ . The definitions used by Rajagopal and Wilczek (1993a, 1993b) are

$$m_\sigma^{-2} = \int d^3x G_{00}, \quad m_\pi^{-2} \delta_{ij} = \int d^3x G_{ij}, \quad (4.13)$$

where  $G_{\alpha\beta}(x) \equiv \langle \phi(x)_\alpha \phi(0)_\beta \rangle - \langle \phi_\alpha \rangle \langle \phi_\beta \rangle$  and  $\phi_0 \equiv \sigma$ ,  $\phi_i = \pi_i$ ,  $i = 1, 2, 3$ .

If the masses are defined according to Eq. (4.13), they are related to the spatial correlation functions in the equilibrium theory. They are sensitive to static screening lengths. “Static” correlation lengths should be seen in contrast to dynamic length scales, which are relevant in the context of the dynamic universality class of an  $O(4)$  Heisenberg antiferromagnet.

In the *chiral limit above  $T_c$*  ( $t > 0$  and  $H \rightarrow 0$ ) the criti-

cal equation of state gives the expected degeneracy between  $m_\pi^2$  and  $m_\sigma^2$ ,

$$m_\sigma^2 = m_\pi^2 = ct^\gamma \quad \text{for } x \rightarrow \infty. \quad (4.14)$$

In the *chiral limit below*  $T_c$  ( $t < 0$  and  $H \rightarrow 0$ ) it follows from Eq. (4.12) that  $m_\pi^2 \propto H$ , i.e.,  $m_\pi^2 \propto m$ , the familiar proportionality from zero temperature. On the other hand, Eq. (4.12a) follows from  $m_\pi^2 = H/M$ , which reads in a literal translation according to Table II

$$m_\pi^2 = \frac{m}{\langle \bar{q}q \rangle}. \quad (4.15)$$

This is less familiar in view of the  $T=0$  PCAC relation

$$f_\pi^2 m_\pi^2 = 2m \langle \bar{q}q \rangle. \quad (4.16)$$

Equations (4.15) and (4.16) do not contradict each other, if  $f_\pi$  is identified with  $f_\pi = \langle 0 | \sigma | 0 \rangle$ . The last relation is compatible with a definition of  $f_\pi$  via the axial current of the linear sigma model at zero temperature. If we had started with Eq. (4.16) and used  $f_\pi = \langle 0 | \sigma | 0 \rangle$ , the resulting relation at  $T \sim T_c$  might not be considered as derived from the  $T=0$  relation (4.16). This point has been emphasized by Rajagopal and Wilczek (1993a, 1993b). The masses  $m_\pi^2$  in Eqs. (4.15) and (4.16) cannot be the same as the notation suggests. In Eq. (4.15)  $m_\pi$  is the mass in the  $(3+1)$ -dimensional Lorentz-invariant  $T=0$  theory. In Eq. (4.16) it resembles a screening mass in a 3D theory.

The result for the sigma mass at temperature  $T < T_c$  in the limit of vanishing quark mass reads

$$\left( \frac{\beta m_\sigma^2}{M^{\delta-1}} \right)^{-1} \rightarrow c_1 + c_2 \left( \frac{H}{M^\delta} \right)^{-\epsilon/2}. \quad (4.17)$$

Here  $\beta$  denotes the critical exponent of the ‘‘magnetization,’’  $c_1$  and  $c_2$  are constants, and  $\epsilon$  is the remnant of the  $\epsilon$  expansion.

To summarize: The critical equation of state (4.11) may serve as a working hypothesis for lattice simulations. Its various limits are predictions for the scaling behavior of  $m_\pi^2$  [Eq. (4.15)],  $m_\sigma^2$  [Eq. (4.17)], or  $m_\sigma^2 = m_\pi^2$  [Eq. (4.17)] in the chiral symmetric phase. In particular, the quantitative predictions for the critical indices  $\beta, \gamma, \delta$  referring to the condensate as a function of  $m$  and  $T$  are ready for numerical tests.

#### d. The role of the strange quark mass

We come now to the second hypothesis concerning the role of the strange-quark mass in the chiral transition. So far we have implicitly assumed that the remaining quark masses  $m_s, m_c, m_b,$  and  $m_t$  are infinite. This is certainly justified for the charm, bottom, and top quark masses, which are large compared to the chiral transition temperature, but the strange-quark mass is just of the order of the transition temperature. Thus it can influence thermodynamic quantities in a nontrivial way. Such an influence is already visible in lattice results. If  $m_s$  is infinite, the chiral transition seems to be of second order. We have quoted the references above. If it is

zero, renormalization-group arguments predict a first-order transition (Pisarski and Wilczek, 1984). Numerical simulations for three light flavors verify this conjecture (see, for example, Gottlieb, 1991). Hence a critical value  $m_s^*$  should exist at which the second-order transition changes into a first-order transition. Usually such an end point, where the order changes from first into second, is called a *tricritical point*. The physical value of  $m_s$  is unlikely to coincide with  $m_s^*$ , but it may be close by. Therefore it is tempting to describe the realistic mass parameters as a perturbation around the idealized tricritical limit. If such an ansatz is justified, the deviations from tricriticality are under control.

According to the second hypothesis, a simple model that shares the universality class of QCD with two massless flavors of quarks,  $T \sim T_c$  and  $m_s$  near  $m_s^*$ , is the  $\phi^6$  Landau-Ginzburg model. Its action reads

$$S = \int d^3x \left\{ \frac{1}{2} (\nabla \phi)^2 + \frac{\mu^2}{2} \phi^2 + \frac{\lambda}{4} (\phi^2)^2 + \frac{\kappa}{6} (\phi^2)^3 - H\sigma \right\}. \quad (4.18)$$

The field  $\phi$  is the same as in Eq. (4.4). The explicit symmetry breaking due to  $(-H\sigma)$  has been added to account for finite masses  $m_{u,d}$ . The  $\phi^6$  term arises as follows. The effect of a finite strange-quark mass is to renormalize the mass and coupling  $\mu^2$  and  $\lambda$  in Eq. (4.4). For example, one contribution to the renormalization of  $\lambda$  comes from a  $K$ -meson exchange between two pions. The ‘‘amount’’ of renormalization depends on  $m_s$ . The effect of a mass and coupling renormalization is a shift in  $T_c$  as long as  $\lambda$  stays positive. If  $\lambda$  is negative the model becomes unstable, and a  $\phi^6$  term is needed for stabilization. It is easily checked that, for  $\lambda < 0$  and fixed, the minimum of the free energy jumps discontinuously from zero to a finite value  $|\lambda|/(2\kappa)$  when  $\mu^2 = \lambda^2/(4\kappa)$ . (For  $\lambda > 0$  and fixed, the minimum moves continuously from zero to positive values, when  $\mu^2$  goes through zero, as one enters the broken phase.) Hence the value of  $m_s^*$  can be defined as the strange-quark mass for which the renormalized coupling  $\lambda$  vanishes. At this point the second order of the  $\lambda > 0$  region changes into first order ( $\lambda < 0$  region).

Singularities of thermodynamic functions are universal near tricritical points. Thus one may again employ results from statistical physics. Tricritical exponents of the  $\phi^6$  model have been calculated by Lawrie and Sarbach (1984). Of particular interest is the result for  $\alpha$ , which they find to be  $\alpha = 1/2$ . Note that now  $\alpha > 0$  [in contrast to the three-dimensional  $O(4)$  model], indicating a true divergence of the specific heat when  $T \rightarrow T_c$  and  $m_s \rightarrow m_s^*$ . In lattice calculations  $m_s$  can be tuned to small values. A qualitative change in the shape of the cusp in the specific heat would be a hint of the presence of a tricritical point nearby. This illustrates the predictive power of the second hypothesis of Rajagopal and Wilczek (1993a, 1993b) and Wilczek (1992).

The third hypothesis concerns the chiral transition and its relation to dynamic universality classes. So far we have dealt with static properties of equilibrium QCD. The third hypothesis may be relevant for off-equilibrium situations in heavy-ion collisions. Therefore we postpone discussion of it to Sec. V.D.3.

Lattice simulations in full QCD can measure critical and tricritical indices in the vicinity of  $T_c$  (and  $m_s \sim m_s^*$ ). If these results disagree with the predictions of the above conjectures, it would be for several reasons. One possible reason is the size of the subleading terms in Eqs. (4.6), (4.9), and (4.10). The subleading terms are only small so long as the parameters  $T$  and  $m_s$  are in the close vicinity of the critical values  $T_c$  and  $m_s^*$ . Corrections from subleading terms do not invalidate the hypothesis. The ansatz would fail if the physical values for  $m_u$ ,  $m_d$ , and  $m_s$  perturbed the idealized situation of two-flavor QCD too strongly to reveal the underlying substructure of the chiral limit. Furthermore, while  $SU(2) \times SU(2)$  is a sensible choice for the symmetry group, it is not unique. If  $SU(2) \times SU(2)$  were replaced by  $SU(3) \times SU(3)$ , it would become impossible to adapt the hypothesis to the change in the symmetry group. Along with the second order of the transition, the universality arguments would be lost, which were so welcome for justifying a replacement of QCD by scalar  $O(4)$  models.

Finally it is a question of size whether corrections due to nonzero modes and heavier mesons are negligible at  $T_c$ . In the next section we present an attempt to include heavier mesons, in a four-dimensional  $SU(2) \times SU(2)$  nonlinear sigma model. The results cast some doubt on the very ansatz for the three-dimensional action in terms of  $\phi \equiv (\sigma, \boldsymbol{\pi})$ .

## 2. The chiral transition in chiral perturbation theory

It is tempting to extrapolate results from low temperatures to the transition region. Pions, as the lightest hadrons, are most easily excited at low temperatures. One may expect that they remain the only relevant degrees of freedom up to  $T_c$ . Current quark masses are small compared to the scale of the transition temperature. Thus the naive conclusion would be that their effect, if any, is negligible. Based on these simplifying assumptions one might study the chiral limit of an  $O(4)$  model as an effective description of the chiral transition region ( $T \sim T_c$ ). The results of Gerber and Leutwyler (1989) suggest that this line of argument is too naive. We present their results in this section.

Chiral perturbation fails to describe the transition region, but leads to reliable predictions for the condensate and other thermodynamic quantities at small temperatures. The influence of finite quark masses on the  $T$  dependence of  $\langle \bar{q}q \rangle$ ,  $\varepsilon$ , and  $p$  can be analyzed in the low-temperature region. The effect of heavier mesonic modes may be estimated in a dilute-gas approximation for somewhat higher temperatures. Non-negligible effects of finite quark masses and heavier mesons show up

in the marginal validity range of chiral perturbation theory. Their extrapolations to the transition region can be summarized as follows.

(i) Below  $T = 150$  MeV the nonzero quark masses reduce the temperature dependence of the condensate by roughly a factor of 2.

(ii) The effect of massive states on the energy density  $\varepsilon(T)$  is even more significant. In the chiral limit the energy stored in the massive states reaches the order of the energy stored in the pions when  $T \sim 130$  MeV. Nevertheless the massive states may be diluted at this temperature. The main part of the energy is the rest energy of the massive states. At  $T \sim 200$  MeV the mean distance between two particles is reduced to  $d = 0.9$  fm. This estimate is based on the density formula for a free gas. The approximation of a dilute free gas of massive states is then no longer justified. Below  $T \sim 100$  MeV massive states may be neglected.

(iii) At  $T \sim 200$  MeV the dilute-gas approximation for the massive modes predicts a melting of the condensate, even if the pions are completely ignored. This sheds some light on the idealization of the predominant role of pions in the chiral transition (which we anticipated at the end of the previous section).

(iv) The extrapolated value of  $T_c$  decreases from  $\sim 190$  MeV to  $\sim 170$  MeV when heavier mesons are included in the chiral limit.

(v)  $T_c$  decreases from  $\sim 240$  MeV to  $\sim 190$  MeV when heavier mesons are included at finite quark masses. While heavier mesons accelerate the melting of the condensate, nonvanishing quark masses delay the melting by  $\sim 20$  MeV.

Let us briefly recall why chiral perturbation theory is a suitable framework for describing QCD at low temperatures. Chiral perturbation theory is an expansion in small momenta. Its applicability to QCD is based on the fact that strong interactions are weak at low energies. This is a consequence of chiral symmetry. The chiral symmetry of QCD in the massless limit implies that the interaction strength is proportional to the square of the energy if the energy is small. At low temperatures, the properties of the hadron gas are determined by the lightest excitations. The lightest hadrons are the pions. At low temperatures the average momenta of the pions are small. Thus strong interactions between the pions may be treated *perturbatively* in the framework of chiral perturbation theory.

To be specific, we consider the  $SU(2)_R \times SU(2)_L$  chiral symmetry. The particle content is given by the three pion components  $\pi^0$ ,  $\pi^-$ , and  $\pi^+$ . A fourth component corresponding to the  $\sigma$  mode is frozen due to the nonlinear realization of the symmetry. The pion field is described by a matrix field  $U(x) \in SU(2)$ . It transforms under global chiral rotations according to

$$U(x) \rightarrow V_R U(x) V_L^\dagger, \quad (4.19)$$

where  $V_R, V_L \in SU(2)$  and  $U(x)$  is parametrized as  $U(x) = \exp\{i\boldsymbol{\tau}\boldsymbol{\varphi}(x)/f\}$ . Here  $\boldsymbol{\tau}$  denotes the Pauli matrices

ces,  $f$  will later be identified with the pion constant, and the components of  $\varphi$  represent the three components of the pion field.

If one allows only for small four-momenta of the pions, the field  $U(x)$  is slowly varying. It is then convenient to expand the Lagrangian in powers of derivatives of the fields  $\partial_\mu U$ , equivalent to a power series in external momenta  $p$ . The ansatz is given as

$$L_{\text{eff}} = L^{(0)} + L^{(2)} + L^{(4)} + \dots \quad (4.20)$$

The upper index counts the number of derivatives  $\partial_\mu$ , hence  $L^{(0)}$  does not depend on  $U$  and can be dropped. Lorentz invariance forbids odd powers in the derivatives. First we consider the chiral limit with a massless pion field. The form of  $L^{(2n)}$  is then completely determined by the symmetry requirement, i.e., invariance under  $SU(2)_R \times SU(2)_L$  chiral transformations. This leads to

$$L^{(2)} = \frac{f^2}{4} \text{Tr} \partial_\mu U^+ \partial^\mu U. \quad (4.21)$$

Note that this term is just the familiar nonlinear  $SU(2) \times SU(2)$  sigma model. The interaction between the pion fields is described in a loop expansion with  $\varphi$  fields propagating in the loops. The only place where the finite temperature enters is in the boundary conditions, which lead to modified propagators. The power counting rules for graphs associated with the free energy are modified compared to the rules for chiral perturbation theory in the loop expansion at zero temperature. The only modification consists in having temperature replace the external momenta. There is a one-to-one correspondence between the order in the low-temperature expansion and the loop expansion of chiral perturbation theory.

The expansion parameter is  $T^2/(8f^2)$ . Since an increasing number in these loops corresponds to a higher order in the low-temperature expansion,  $L^{(2n)}$ ,  $n=1,2,\dots$  is to be expanded in powers of  $\varphi$  to the order needed for the desired accuracy in the low- $T$  expansion. Up to terms of order  $\varphi^6$  we find, for example, for  $L^{(2)}$

$$L^{(2)} = \frac{1}{2} \partial_\mu \varphi \partial^\mu \varphi + \frac{1}{6f^2} (\varphi \partial_\mu \varphi) \times (\varphi \partial^\mu \varphi) - \frac{1}{6f^2} (\partial_\mu \varphi \partial^\mu \varphi) \varphi^2 + O(\varphi^6). \quad (4.22)$$

Note that the chiral symmetry fixes  $L^{(2)}$  up to one constant  $f$ , which may be identified with  $f_\pi$  in the chiral limit. At the order of  $p^4$ , two further dimensionless couplings enter the effective Lagrangian. They have to be fixed from experiment.

Away from the chiral limit, the mass terms for the pion field induce an explicit breaking of the chiral symmetry in the effective Lagrangian. In the chiral limit the guiding principle for constructing  $\mathcal{L}$  is an expansion in powers of external momenta. It can be maintained in the presence of mass terms if the quark masses are small compared to the relevant physical scales of the theory.

The full effective Lagrangian is Taylor expanded in powers of  $m$ . The derivative expansion can be superimposed on the Taylor expansion, term by term. In this way additional terms involving powers of  $m$  are generated, which break the chiral symmetry explicitly. The power counting in external momenta of the chiral limit goes through if  $m_\pi$  is taken to be of the order of  $p$ . Since the pion mass square  $M^2$  to lowest order in chiral perturbation theory is given as  $M^2 = m_\pi^2 = (m_u + m_d)B$  ( $B$  being the  $T=0$  condensate, and  $m_\pi$  being the physical pion mass), the quark masses are counted as quantities of order  $p^2$ . Inclusion of mass terms then leads to the following modifications in  $L^{(2)}$ :

$$L^{(2)} = \frac{f^2}{4} \{ \text{Tr}(\partial_\mu U^+ \partial^\mu U) - \text{Tr}[M^2(U + U^+)] \}. \quad (4.23)$$

$L^{(4)}$  now depends on four input parameters  $l_1, \dots, l_4$  (Gasser and Leutwyler, 1984). Thus the main effect of finite quark masses is an increase in the number of input parameters. Once the effective action is determined and the input parameters are fixed from experiment, the thermodynamics of the pion gas can be derived in the usual way.

The partition function is given as the path integral over all pion field configurations that are periodic in the Euclidean time direction,  $U(\mathbf{x}, x_4 + \beta) = U(\mathbf{x}, x_4)$ :

$$Z = \int \mathcal{D}U \exp \left\{ - \int_{\mathbb{R}^3 \times S^1} d^4x (L^{(2)} + \dots) \right\}, \quad (4.24)$$

where  $L^{(2)}$  is given by Eq. (4.23). The couplings of  $L^{(2n)}$  are treated as temperature independent. The order parameter, energy density  $\varepsilon$ , pressure  $p$ , and entropy density  $s$  follow from the free-energy density  $\mathcal{F}$  in the standard way [ $\mathcal{F} = -1/\beta V \ln Z$  in the large-volume limit, where  $Z$  is given by Eq. (4.24)].

To obtain a low-temperature expansion of  $Z$ , one expands  $L_{\text{eff}}$  in powers of  $\varphi$ .

Let us illustrate the power counting in the expansion parameters for the case in which the free-energy density is to be determined including terms of order  $T^8$ . This is the accuracy achieved by Gerber and Leutwyler (1989).

To obtain  $\mathcal{F}$  to the order of  $T^8$ , one must expand  $L$  to order  $p^8$ , with contributions coming from  $L^{(2)}$ ,  $L^{(4)}$ ,  $L^{(6)}$ , and  $L^{(8)}$ . Tree graphs are generated from all terms  $L^{(2)}, \dots, L^{(8)}$ ; one-loop graphs contain vertices of  $L^{(4)}$  and  $L^{(6)}$ , two-loop graphs contain vertices of  $L^{(2)}$  and  $L^{(4)}$ , and three-loop graphs are generated by  $L^{(2)}$ . Thus  $L^{(2)}$  has to be specified in terms of  $\varphi$  including powers of the order  $\varphi^6$ . Divergent integrals are treated in dimensional regularization.

The leading contribution of the Lagrangian  $L^{(2)}$  contains a free-field term (i.e., a term quadratic in  $\varphi$  with mass  $M^2$ , the lowest-order contribution to the pion mass). Further terms of the Lagrangian are kept as perturbations depending on their importance at low energies. In this way chiral perturbation theory also provides a systematic expansion for thermodynamic quantities. It goes beyond the mean-field level, where one is frequently stuck in the treatment of effective models.

In the *chiral limit* the free-energy density  $\mathcal{F}$  takes the general form

$$\mathcal{F} = \sum_{m,n=0,1,\dots} c_{mn} (T^2)^m (T^2 \log T)^n + \mathcal{O}(e^{-M_e/T}). \quad (4.25)$$

The correction of the order  $\exp\{-M_e/T\}$  stands for the contribution of particles that remain massive in the chiral limit, and  $M_e$  denotes the lightest of these masses.

A similar expression can be derived in the presence of finite quark masses when a common small expansion parameter  $\lambda$  is introduced such that  $T \rightarrow \lambda T, m \rightarrow \lambda^2 m$ . The formula for  $\mathcal{F}$  then holds in powers of  $\lambda$  and  $\log \lambda$  with  $\lambda$ -independent coefficients  $c_{mn}$ , which are now non-trivial functions of  $(m_\pi/T)$ .

#### a. The free-field case

Let us first ignore the interactions between the pions and consider  $L^{(2)}$ . It leads to the following familiar expressions for the free-energy density  $\mathcal{F}$ , the pressure  $p$ , and the condensate  $\langle \bar{q}q \rangle$  for an ideal gas of pions with mass  $m_\pi$ :

$$\mathcal{F} = \varepsilon_0 - p = \varepsilon_0 + 3T \int \frac{d^3p}{(2\pi)^3} \ln\{-e^{-E/T}\} + \dots, \quad (4.26)$$

$$E = \sqrt{\mathbf{p}^2 + m_\pi^2}.$$

Here  $\varepsilon_0$  is the vacuum energy density at  $T=0$ . In the chiral limit the result for  $\mathcal{F}$  is

$$\mathcal{F} = \varepsilon_0 - \frac{\pi^2}{30} T^4 + \dots \quad (4.27)$$

with the  $T^4$  law for the pressure of a free gas. The condensate is obtained as

$$\langle \bar{q}q \rangle = \frac{\partial \mathcal{F}}{m} = \frac{\partial \varepsilon_0}{\partial m} + \frac{3}{2} \frac{\partial m_\pi^2}{\partial m} \int \frac{d^3p}{(2\pi)^3} \frac{1}{E(e^{E/T} - 1)}. \quad (4.28)$$

The first term on the right-hand side gives the condensate at temperature  $T=0$ . Using the PCAC relation,

$$f_\pi^2 m_\pi^2 = -m \langle 0 | \bar{q}q | 0 \rangle \{1 + \mathcal{O}(m)\}, \quad (4.29)$$

the condensate in the chiral limit is given by

$$\langle \bar{q}q \rangle = \langle 0 | \bar{q}q | 0 \rangle \left\{ 1 - \frac{T^2}{8f_\pi^2} + \dots \right\}, \quad (4.30)$$

where  $f = f_\pi$  up to terms of order  $m$ , which are counted as order  $(T^2)$ . Thus the difference shows up in  $T^4$  terms. Equation (4.30) shows the melting of the condensate as the temperature increases. Simultaneously the pion density  $n_\pi$  grows according to

$$n_\pi = 0.365 T^3 \quad (4.31)$$

in the massless limit. Physically the increase in number density can be made responsible for the melting process.

For  $m \neq 0$ , the integral in Eq. (4.28) becomes a function of  $m_\pi/T$ .

#### b. Including interactions up to three loops

Let us see next how the results for  $p$ ,  $\varepsilon$ , and  $\langle \bar{q}q \rangle$  change when interactions are included. It is clearly beyond the scope of this review to go into the details of the three-loop calculation. Nor shall we explain, how the various constants in the effective Lagrangian and the logarithmic scales are fixed from experimental observables (the pion mass, the pion decay constant, and various scattering lengths).

#### c. Results in the chiral limit without heavier mesons

The low-temperature expansion for the pressure is given as

$$p = \frac{\pi^2}{30} T^4 \left\{ 1 + \frac{T^4}{36f^4} \ln \frac{\Lambda_p}{T} + \mathcal{O}(T^6) \right\}. \quad (4.32)$$

Here  $\Lambda_p = 275 \pm 65$  MeV, and  $m_\pi$ ,  $f_\pi$ , and the  $I=0$ ,  $D$ -wave (i.e.,  $L=2$ ) scattering length were used as experimental input data.

The low-temperature expansion for the energy density  $\varepsilon = T(\partial p / \partial T) - p$  follows as

$$\varepsilon = \frac{\pi^2}{10} T^4 \left\{ 1 + \frac{T^4}{108f^4} \left( 7 \ln \frac{\Lambda_p}{T} - 1 \right) + \mathcal{O}(T^6) \right\}. \quad (4.33)$$

The general expression for the condensate can be written as

$$\langle \bar{q}q \rangle = \langle 0 | \bar{q}q | 0 \rangle \left\{ 1 + \frac{c}{f^2} \frac{\partial p}{\partial m_\pi^2} \right\} \quad (4.34)$$

with

$$c = -f^2 \frac{\left( \frac{\partial m_\pi^2}{\partial m} \right)}{\langle 0 | \bar{q}q | 0 \rangle}$$

[cf. Eqs. (4.28), (4.30) above]. Hence one needs to know  $p = p(m_\pi^2)$  (Leutwyler, 1988). The temperature-independent term  $c$  has to be determined;  $c=1$  in the chiral limit. The result for  $\langle \bar{q}q \rangle$  including terms of order  $T^6$  takes the form

$$\langle \bar{q}q \rangle = \langle 0 | \bar{q}q | 0 \rangle \left\{ 1 - x - \frac{1}{6} x^2 - \frac{16}{9} x^3 \ln \left( \frac{T}{\Lambda_q} \right) + \mathcal{O}(T^8) \right\}$$

$$x = \frac{T^2}{(8f^2)}, \quad \Lambda_q = 470 \pm 110 \text{ MeV}. \quad (4.35)$$

The expansion parameter  $x$  shows that the temperature scale is determined by  $\sqrt{8}f \sim 250$  MeV. The scale  $\Lambda_q$  is fixed in an analogous way to  $\Lambda_p$ . Both scales  $\Lambda_p$  and  $\Lambda_q$  can be expressed in terms of a single scale  $\Lambda_b$ , which is fixed from  $\pi\pi$  scattering. Thus to describe the behavior of the pressure  $p$  to order  $T^8$  and of  $\langle \bar{q}q \rangle$  to order  $T^6$  one needs only two constants,  $f_\pi$  and  $\Lambda_b$ . This does not leave much room for phenomenological fitting and demonstrates the predictive power of chiral perturbation theory.

The results for  $\langle \bar{q}q \rangle$  and the energy density  $\varepsilon$  are displayed in Fig. 17.

It is tempting to read off a transition temperature of  $T_c \sim 190$  MeV, if the three-loop curve is extrapolated to a temperature where the condensate vanishes. This extrapolation suffers from two error sources. The error bars on the three-loop curve refer only to the uncertainty in the constants  $f$  and  $\Lambda_b$ , and not to the finite-temperature corrections of Eq. (4.35). Contributions of massive states to  $\mathcal{F}$  have been left out so far. They are no longer negligible when  $T$  exceeds 150 MeV. Below we shall discuss the influence of massive states further.

The energy density in Fig. 18 grows with  $T$  up to  $T \sim 180$  MeV. The decrease of  $\varepsilon/T^4$  for higher temperatures indicates the invalidity of the expansion scheme.

#### d. Nonvanishing quark-masses

Nonvanishing quark-masses change the general expression for the pressure, where certain constants of the chiral limit now depend on  $m_\pi$  also via  $m$ . The result for the quark condensate is shown in Fig. 19. The effect of finite quark masses is that it takes longer for the condensate to melt. The extrapolated transition temperature is  $\sim 240$  MeV (which lies far outside the validity range of the approximation). The error bars indicate the uncertainties in four input parameters, which have to be fixed from experiment for  $m > 0$ .

A comparison of Figs. 17 and 19 shows that the effect of finite quark masses is substantial even inside the validity range of the chiral expansion. The reason why the tiny quark masses may cause such changes in the final results is that they are confined to pions. It is the pion mass ( $\sim 140$  MeV) that enters the Boltzmann factors in the thermodynamic formulas for the hadronic phase.

#### e. Influence of heavier mesons

In an expansion around the chiral limit, quark masses must be treated as perturbations when they are incorporated in an expansion of chiral perturbation theory. A perturbative treatment of  $m_u$  and  $m_d$  is certainly justified, as the up- and down-quarks are light compared to the energy scale of  $T_c$ . Mesons containing strange quarks are accordingly heavier than the pions. Frequently they are argued to be negligible in the thermodynamics, as massive states are suppressed by their Boltzmann factors  $\exp\{-M/T\}$ . This argument fails for higher temperatures, in particular in the vicinity of the transition.

One way of including (part of the) heavier mesons is to extend the symmetry group from  $SU(2) \times SU(2)$  to  $SU(3) \times SU(3)$  (see Sec. IV.A.4). In the framework of chiral perturbation theory this means that  $m_s$  should also be treated perturbatively.

A further alternative is to treat all massive mesons (other than pions) on an equal footing and consider their number density and energy density in a dilute-gas approximation (Gerber and Leutwyler, 1989). In a dilute-gas approximation the interactions between the mesons are neglected. This justifies an ansatz for the change in the free-energy density according to

$$\Delta\mathcal{F} = -\frac{1}{2} T^4 \sum_i h_0\left(\frac{T}{M_i}\right). \quad (4.36)$$

The sum extends over all mesons tabulated in the particle data booklet apart from pions. It is clear that many small contributions may accumulate to a non-negligible amount. The change  $\Delta\mathcal{F}$  is due to a change in the pressure. The term  $1/2 T^4 h_0(T/M_i)$  gives the pressure of a free boson gas of particles with mass  $M_i$ . The function  $h_0$  can be expressed in terms of Bessel functions. The induced change in the condensate (due to the mass  $M_i$ ) follows from Eq. (4.28) with  $m_\pi$  replaced by  $M_i$ . It is given as

$$\Delta\langle\bar{q}q\rangle = T^2 \sum_i h_1\left(\frac{M_i}{T}\right) \left(\frac{\partial M_i}{\partial m^2}\right) M_i \quad (4.37)$$

where

$$T^2 h_1 = \int \frac{d^3p}{(2\pi)^3} \frac{1}{E} \frac{1}{e^{E/T} - 1}$$

or equivalently

$$\Delta\langle\bar{q}q\rangle = -\langle 0|\bar{q}q|0\rangle \frac{T^2}{f^2} \sum_i h_1\left(\frac{T}{M_i}\right) \frac{M_i m}{m_\pi^2} \frac{\partial M_i}{\partial m}. \quad (4.38)$$

Here we have used the lowest-order PCAC relation for  $f_\pi^2 m_\pi^2$ . The derivative  $\partial M_i/\partial m$  can be estimated in the nonrelativistic quark model, where it equals  $N_i$ , i.e., the number of valence quarks of type  $u$  or  $d$ . Equation (4.38) explains the effect of massive states on the melting of the condensate, accelerating the melting and lowering the transition temperature. In Fig. 20 the melting of the chiral condensate is compared for the three-loop pion contribution and the superposition of pions plus massive states. The extrapolated transition temperature is around 190 MeV, smaller than  $T_c$  without heavier mesons, but  $\sim 20$  MeV larger than in the chiral limit with heavier mesons. The shaded area reflects the uncertainty of  $(m \partial M_i/\partial m)$  entering the estimate for the contribution of massive states [Eq. (4.38)]. The acceleration of the melting process due to the massive states is more pronounced for realistic quark masses than for vanishing ones. In the chiral limit the melting of the condensate is mainly enforced by the pions, which play a more distinguished role, and thus the condensate is less sensitive to heavier mesons.

At temperatures  $T > 150$  MeV pions have lost their dominant role in melting the condensate. If lattice calculations or other approximation schemes support a value of  $T_c > 150$  MeV, it is challenging to elucidate just how the massive modes drive  $\langle\bar{q}q\rangle$  to zero. It would seem to be very worthwhile to improve on the dilute-gas approximation. A suggestion due to Gerber and Leutwyler (1989) is to treat the mutual interactions between the massive states in some bootstrap model, modified and adapted to QCD (see, for example, Hagedorn, 1985, 1983).

More recently the temperature dependence of the normalized light-quark condensate has been derived by

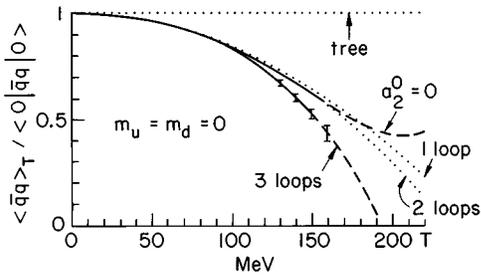


FIG. 17. Quark condensate, normalized to the zero-temperature value, as a function of temperature in the chiral limit. From Gerber and Leutwyler (1989).

Bunatian and Wambach (1994) in a nonperturbative approach. The approach is known from phonon-phonon interactions in condensed-matter physics (e.g., Luttinger and Ward, 1960; Fulde and Wagner, 1971) and may be more appropriate to the transition region. In the chiral limit the critical temperature is identified as the point where the residue of the single-pion propagator becomes purely imaginary. The pion interactions are described by the Weinberg Lagrangian (Weinberg, 1966, 1967, 1979). The thermodynamic potential for the interacting pion gas is determined in the Hartree approximation. After using the Gell-Mann-Oakes-Renner relation, Bunatian and Wambach find that the normalized light-quark condensate follows from  $\partial\Omega/\partial m_\pi^2$ , where  $\Omega$  denotes the difference in free-energy densities at  $T=0$  and  $T>0$ . The results confirm chiral perturbation theory up to temperatures  $<100$  MeV, but are clearly different in the transition region, where chiral perturbation theory is expected to fail.

### 3. Clausius-Clapeyron equation for QCD

We recall the Clausius-Clapeyron equation for a liquid/vapor system in statistical physics. It was derived in the last century by Clapeyron (1834) and Clausius (1850). The Clausius-Clapeyron equation relates the discontinuity in the order parameter to the latent heat when the phase transition is of first order. For a liquid/gas system the order parameter is the specific volume or the volume per molecule, and the latent heat equals the entropy gap per molecule. The ratio of the discontinuities in the specific volume and the entropy is determined by the dependence of the critical temperature on pressure during the transition. The relation is given as

$$\frac{\partial T_c}{\partial p} = \frac{v_{\text{gas}} - v_{\text{fluid}}}{s_{\text{gas}} - s_{\text{fluid}}} = \frac{\text{disc}v}{\text{disc}s}, \quad (4.39)$$

where disc stands for discontinuity.

With the help of Table II, Eq. (4.39) can be translated to QCD, where it reads

$$\frac{\partial T_c}{\partial m_q} = \frac{T_c \text{disc}\langle \bar{q}q \rangle_T}{\text{disc}\epsilon}. \quad (4.40)$$

The pressure is replaced by the current quark mass  $m_q$ . The conjugate variable of  $m_q$  is the order param-

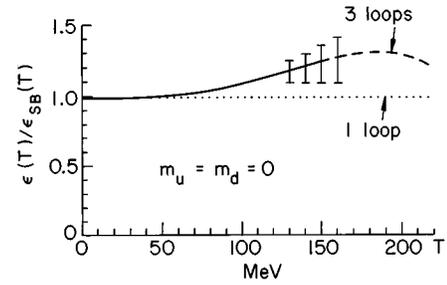


FIG. 18. Energy density of the pion gas (normalized to the corresponding Stefan-Boltzmann value  $\epsilon_{\text{SB}}$ ) in the chiral limit. From Gerber and Leutwyler (1989).

eter  $\langle \bar{q}q \rangle$  at temperature  $T$ . The chiral condensate is determined as a response of the partition function to a change in the quark mass,

$$\langle \bar{q}q \rangle_T = -\frac{1}{V} \frac{\partial}{\partial m_q} \ln Z. \quad (4.41)$$

We sketch the derivation of Eq. (4.40) and show how it leads to a bound on the latent heat in a hypothetical first-order chiral transition. The derivation is due to Leutwyler (1992). First note that the discontinuity in the energy density  $\epsilon = (T\partial p/\partial T - p)$  can be expressed by the gap in the entropy density

$$\text{disc}\epsilon = T_c \text{disc} \frac{\partial p}{\partial T}. \quad (4.42)$$

From Eq. (4.41) with  $\ln Z = V(p - \epsilon_0)$ , where  $\epsilon_0$  is the vacuum energy density, it follows that

$$\langle \bar{q}q \rangle_T = \frac{\partial \epsilon_0}{\partial m_q} - \frac{\partial p}{\partial m_q} = \langle 0 | \bar{q}q | 0 \rangle - \frac{\partial p}{\partial m_q}. \quad (4.43)$$

If we expand the pressure in the vicinity of  $T_c$  according to

$$p = p_c + (T - T_c) \left. \frac{\partial p}{\partial T} \right|_{T_c} + \dots, \quad (4.44)$$

$p$  depends on  $m_q$  via  $p_c$  and  $T_c$ ,

$$\frac{\partial p}{\partial m_q} = \frac{\partial p_c}{\partial m_q} - \frac{\partial T_c}{\partial m_q} \left. \frac{\partial p}{\partial T} \right|_{T=T_c}. \quad (4.45)$$

Next we insert Eq. (4.45) into Eq. (4.43) and apply “disc” on both sides. The result is

$$\text{disc}\langle \bar{q}q \rangle_T = \frac{\partial T_c}{\partial m_q} \text{disc} \left. \frac{\partial p}{\partial T} \right|_{T=T_c}. \quad (4.46)$$

While  $p$  is continuous at  $T_c$  (as in the case of a first-order transition),  $\partial p/\partial T$  may jump. From Eqs. (4.42) and (4.46) we finally obtain Eq. (4.40).

Let us see what the Clausius-Clapeyron relation implies for the change in the critical temperature  $\Delta T_c$  when the physical values for the current quark masses  $m_u$  and  $m_d$  are replaced by their chiral limits. As  $m_{u,d}$  are small compared to  $T_c$ , we may expand

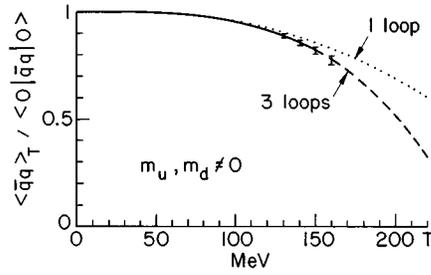


FIG. 19. Quark condensate, normalized to the zero-temperature value, for nonzero quark mass ( $m_\pi=140$  MeV). From Gerber and Leutwyler (1989).

$$\Delta T_c = m_u \frac{\partial T_c}{\partial m_u} + m_d \frac{\partial T_c}{\partial m_d}, \quad (4.47)$$

where  $\partial T_c / \partial m_{u,d}$  is given by Eq. (4.40). It remains to replace the quark masses in Eq. (4.47) by means of the Gell-Mann-Oakes-Renner relation (Gell-Mann *et al.*, 1968),  $m_u \langle \bar{u}u \rangle_{T=0} + m_d \langle \bar{d}d \rangle_{T=0} \sim f_\pi^2 m_\pi^2 \sim 20 \text{ MeV/fm}^3$ . The result is

$$\frac{\Delta T_c}{T_c} = \left| \frac{\text{disc} \langle \bar{q}q \rangle_{T>0}}{\langle \bar{q}q \rangle_{T=0}} \right| \left| \frac{f_\pi^2 m_\pi^2}{\text{disc } \epsilon} \right|. \quad (4.48)$$

This relation restricts the size of the latent heat  $\text{disc } \epsilon$  by the sensitivity  $\Delta T_c / T_c$  of the critical temperature to the quark masses. An upper bound on the discontinuity in  $\epsilon$  corresponds to an upper bound on the (hypothetical) gap in the order parameter.

To get an estimate for this bound, we take  $T_c \sim 140$  MeV as a lower bound on the chiral transition temperature. The number density  $n$  of the hadron gas grows with temperature. At  $T_c \sim 140$  MeV  $n$  increases by  $\sim 30\%$  in passing from the physical quark masses to the chiral limit ( $m_{u,d} \rightarrow 0$ ) (see Leutwyler, 1992). This estimate is based on a dilute-gas approximation, where the change in the particle abundances is only induced by the change in their masses as a function of  $m_{u,d}$ . If the onset of the phase transition is determined by the average volume per particle, a smaller transition temperature has to compensate for the increased density to get the same average volume. A 30% increase in the density  $n$  induces a decrease of 5% in  $T_c$ , i.e.,  $\Delta T_c / T_c \sim 5\%$  under the above conditions. For the maximal jump of  $\langle \bar{q}q \rangle_{T_c}$  one thus finds

$$\text{disc } \epsilon < 0.4 \frac{\text{GeV}}{\text{fm}^3}. \quad (4.49)$$

This value is in good agreement with the bound on the latent heat obtained in a large- $N$  approximation to the  $SU(3) \times SU(3)$  linear sigma model (Meyer-Ortmanns and Schaefer, 1996), which will be described in the next section. For comparison we mention the gap in the gluonic energy density in a pure  $SU(3)$  gauge theory (Laermann *et al.*, 1995). Laermann *et al.* find discontinuity  $\epsilon_{\text{gluonic}} / T_c^4 \sim 1.5$  or  $\text{disc } \epsilon_{\text{gluonic}} \sim 0.9 \text{ GeV/fm}^3$  for  $T_c = 260$  MeV.

For a larger value of discontinuity  $\epsilon$ , say

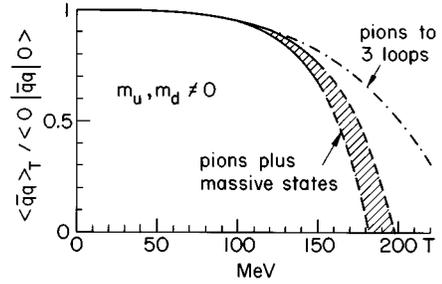


FIG. 20. Temperature dependence of the quark condensate (normalized to the zero-temperature value) for nonzero quark masses. The shaded area represents the superposition of the contributions generated by pions and by massive states. From Gerber and Leutwyler (1989).

$\sim 2 \text{ GeV/fm}^3$ , the Clausius-Clapeyron relation implies  $\Delta T_c / T_c < 1\%$ . Results of chiral perturbation theory indicate that such a small change in the critical temperature is rather unlikely. The condensate and the energy density are rather sensitive to the pion mass  $m_\pi$ .

Chiral perturbation theory cannot predict the order of the phase transition, as it loses its validity in the transition region, but it does provide a useful upper bound on the latent heat. This bound should be noticed in particular by phenomenologists, who like a large latent heat for visible effects in heavy-ion collisions. Scenarios based on such an assumption may be incompatible with the underlying QCD.

#### 4. Mass sensitivity of the chiral transition

In this section we discuss two topics. The first is the mass sensitivity of the chiral transition in the  $SU(3) \times SU(3)$  linear sigma model. We give estimates for *critical meson/quark masses* at which the first-order chiral transition becomes of second order and turns into a smooth crossover phenomenon for larger masses. The second topic is the equation of state for physical meson masses.

##### a. Mass sensitivity

The mass sensitivity of the order of the chiral transition is important for any realistic predictions of heavy-ion collisions. To estimate the effect of finite masses on the chiral transition we recall the analogy to a ferromagnet or a liquid/gas system. Quark masses are analogous to the pressure in a liquid/gas system or an external magnetic field in a ferromagnet. If the pressure exceeds a critical value, the first-order transition from a liquid to a gas ceases to occur, and is replaced by a smooth crossover between the liquid and the gas phases. Similarly a first-order transition in a  $Z(3)$  Potts model becomes a second-order transition for a critical value of the external field and disappears beyond this strength.

When the analogy to a statistical system is translated to an effective model for QCD, the question concerning the effect of finite masses can be posed in the following way. Are the physical meson masses too small for the

chiral transition to maintain its first order? In Sec. III we have described an attempt to answer this question within the lattice approach (see Fig. 13). Here we study the same question on the mesonic level. The bare-quark masses  $m_s a$  and  $m_{u,d} a$  of Fig. 13 should be replaced by meson masses with and without strange-quark content.

A first attempt to estimate the critical phase boundary between mass parameters of first-order transitions and crossover phenomena was made by (Meyer-Ortmanns *et al.*, 1992) in the linear  $SU(3) \times SU(3)$  sigma model. More recently Gavin, Goksch, and Pisarski (1994a) addressed a similar question in the same sigma model within a mean-field approach. In the remainder of this section we shall focus on the results of a finite-mass analysis of Meyer-Ortmanns and Schaefer (1996) in a large- $N_f$  approximation. As a quantitative measure of the distance (in mass parameter space) between the physical and critical meson masses, we consider the ratio of the associated critical to physical light-quark masses  $m_{u,d}^{\text{crit}}/m_{u,d}$ . If this ratio turns out to be much smaller than 1, physical masses lie deeply in the crossover region, and it is difficult to imagine any signatures specific for the phase conversion from the plasma to the hadron phase. The most attractive possibility is a ratio of the order of 1. In that case the physical masses are almost “critical,” and effects due to a large correlation length should be visible. Nonuniversal features of the sigma model are then negligible, and the reduction of QCD to an effective model in the same universality class is an allowed simplification.

Here we would like to recall that sigma models may not be the right candidates for sharing a universality class with QCD (Kocić and Kogut, 1995). As long as the deviations from ideal second-order behavior are perturbatively small, however, they are theoretically under control (see, for example, the finite-mass scaling analysis of Boyd *et al.*, 1992, Rajagopal and Wilczek, 1993a, 1993b, or Kocić *et al.*, 1993). A comparison of critical indices in the effective model and on the lattice serves as a check of the conjectured universality class.

### b. Choice of the model

In our application we assume that the restoration of spontaneously broken  $SU(3) \times SU(3)$  symmetry is the driving mechanism for the chiral phase transition. Deviations in the spectrum from the idealized octet of pseudoscalar Goldstone bosons are parametrized by terms that break the  $SU(3) \times SU(3)$  symmetry explicitly. The assumption in our simplification is that only mesons associated with the  $SU(3) \times SU(3)$  multiplets are important for the phase transition. The criterion is chiral symmetry (rather than the size of the meson masses, otherwise one should include  $\rho$  mesons or others as well). The reason why we have chosen  $SU(3) \times SU(3)$  rather than  $SU(2) \times SU(2)$  is to account for the influence of the strange-quark mass on the thermodynamics. With  $SU(3) \times SU(3)$  we also include some of the heavier mesons. From Sec. IV.A.2 we know that heavier mesons are non-negligible in the transition region.

We have chosen the sigma model as an effective model for the low-temperature phase of QCD ( $T \leq T_c$ ). In the low-temperature phase, quarks are confined to hadrons, and chiral symmetry is spontaneously broken. The meson spectrum reflects some remnants of this symmetry breaking. In the transition region the use of the model becomes questionable; the model certainly fails above  $T_c$  as a description of the plasma phase.

We use the sigma model in a similar spirit to that in which we used the  $O(4)$  model in Sec. II.A.1. Its action is constructed in terms of QCD’s chiral order-parameter field  $\phi$ , where  $\phi$  now is a complex  $3 \times 3$  matrix, parametrized as

$$\phi = \frac{1}{\sqrt{2}} \left( \sum_{l=0}^8 (\sigma_l + i \pi_l) \lambda_l \right). \quad (4.50)$$

Here  $\lambda_l$  denote the Gell-Mann matrices,  $\pi_l$  are the pseudoscalar mesons, and  $\sigma_l$  are the scalar mesons. The mesonic order-parameter field is bilinear in the left-handed and right-handed quark fields  $\phi_{ij} = \langle \bar{q}_i^L q_j^R \rangle$ . In the sigma model the quark structure is ignored by construction for all temperatures  $T \geq 0$ . In terms of  $\phi$  the Lagrangian reads

$$\begin{aligned} L = \int d^4x & \left\{ \frac{1}{2} \text{Tr}(\partial_\mu \phi \partial_\mu \phi^\dagger) \right\} - \frac{1}{2} \mu_0^2 \text{Tr}(\phi \phi^\dagger) \\ & + f_1 (\text{Tr} \phi \phi^\dagger)^2 + f_2 \text{Tr}(\phi \phi^\dagger)^2 \\ & + g(\det \phi + \det \phi^\dagger) - \epsilon_0 \sigma_0 - \epsilon_8 \sigma_8. \end{aligned} \quad (4.51)$$

Note that there are two independent quartic terms with couplings  $f_1$  and  $f_2$ . The determinantal terms are cubic in the components of  $\phi$ ,  $g$  is the “instanton” coupling that takes care of the right  $\eta$ - $\eta'$  mass splitting, and  $\mu_0^2$  is the coupling of the quadratic term. The external field  $\epsilon_0$  gives a common mass to the (pseudo)scalar meson octet, while  $\epsilon_8$  accounts for the right mass splitting inside the (pseudo)scalar meson octet.

### c. Tree-level parametrization at zero temperature

The parameters  $\mu_0^2, f_1, f_2, g, \epsilon_0, \epsilon_8$  of the Lagrangian (4.51) should be chosen such that the model reproduces the experimental values of the (pseudo)scalar meson masses. The parametrization of the sigma model is not unique (Chan and Haymaker, 1973; Meyer-Ortmanns *et al.*, 1992; Gavin *et al.*, 1994a). Here we are interested in a tuning of meson masses in terms of a few parameters. Suitable parameters are the external fields  $\epsilon_0, \epsilon_8$ . These induce finite quark masses according to

$$\begin{aligned} -\epsilon_0 &= \alpha(2\hat{m} + m_s), \\ -\epsilon_8 &= \beta(\hat{m} - m_s), \end{aligned} \quad (4.52)$$

where  $\hat{m} \equiv (m_u + m_d)/2$ ,  $\alpha$  and  $\beta$  are constants. Equation (4.52) follows from an identification of terms in the Lagrangians for quarks and mesons, which transform identically under  $SU(3) \times SU(3)$ . The meson masses are determined for given  $\epsilon_0, \epsilon_8$ , once the couplings  $\mu_0^2, f_1, f_2, g$  are specified and the condensates  $\langle \sigma_0 \rangle_{T=0}$ ,

TABLE VII. Tree-level parametrization of the SU(3)×SU(3) linear sigma model (input data taken from experiment).

Input	$f_1$	$f_2$	$g$ [MeV]	$f_\pi$ [MeV]	$\epsilon_0$ [GeV <sup>3</sup> ]	$\epsilon_8$ [GeV <sup>3</sup> ]		
$\mu_0^2$ [GeV <sup>2</sup> ]								
$5.96 \times 10^{-2}$	4.17	4.48	-1812.0	94	0.0265	-0.0345		
Output (all masses are in units of [MeV])								
Mass values	$m_\pi$	$m_K$	$m_\eta$	$m_{\eta'}$	$m_{\sigma_\pi}$	$m_{\sigma_K}$	$m_{\sigma_\eta}$	$m_{\sigma_{\eta'}}$
Realistic	129.3	490.7	544.7	1 045.5	1 011.6	1 031.2	1 198.0	749.5
Experimental	138.0	495.7	547.5	957.8	980 if $\sigma_\pi \equiv a_0$	1 322.0 if $\sigma_K \equiv K_0^*$	1 476.0 if $\sigma_\eta \equiv f_0(1476)$	975 if $\sigma_{\eta'} \equiv f_0(975)$

$\langle \sigma_8 \rangle_{T=0}$  are calculated for given  $\mu_0^2, f_1, f_2, g$ . Thus we vary the quark and meson masses by varying  $\epsilon_0$  and  $\epsilon_8$ . The chiral limit is obtained for  $\epsilon_0 = 0 = \epsilon_8$ . The couplings  $\mu_0^2, f_1, f_2, g$  are then determined from the mass input in the chiral limit, i.e.,  $m_\pi = m_K = m_\eta = 0$ ,  $m_{\eta'} = 850$ ,  $m_{\sigma_{\eta'}} = 800$ ,  $m_{\sigma_\eta} = 600$ , where all masses are in units of (MeV) and  $f_\pi = 94$  MeV. Next we keep  $\mu_0^2, f_1, f_2, g$  fixed to their values in the chiral limit and change  $\epsilon_0, \epsilon_8$ . The choice  $\epsilon_8 = 0, \epsilon_0 \neq 0$  leads to an SU(3)-symmetric case with only one order-parameter field  $\sigma_0$ , for which the numerics considerably simplifies. Meson masses with almost experimental values are induced for  $\epsilon_0 = 0.0265$  GeV<sup>3</sup>,  $\epsilon_8 = -0.0345$  GeV<sup>3</sup>.

The results are listed in Table VII and compared to the experimental values.

In this way we have constructed a mapping

$$\{m_{u,d}, m_s\} \leftrightarrow (\epsilon_0, \epsilon_8) \leftrightarrow \{m_{\text{Meson}}^2\} \quad (4.53)$$

between quark and meson masses.

It remains to translate the meson condensates at zero temperature to the light- and strange-quark condensates. In the same way as we obtained Eq. (4.52), we find here

$$\begin{aligned} \langle \bar{q}q \rangle &= \frac{\epsilon_0}{2\hat{m} + m_s} \langle \sigma_0 \rangle + \frac{\epsilon_8}{2(\hat{m} - m_s)} \langle \sigma_8 \rangle, \\ \langle \bar{s}s \rangle &= \frac{\epsilon_0}{2\hat{m} + m_s} \langle \sigma_0 \rangle - \frac{\epsilon_8}{\hat{m} - m_s} \langle \sigma_8 \rangle. \end{aligned} \quad (4.54)$$

Equations (4.54) are derived at zero temperature. We take these relations as temperature independent and use them to determine  $\langle \hat{q}q \rangle(T), \langle \hat{s}s \rangle(T)$  from the measured values for  $\langle \sigma_0 \rangle(T), \langle \sigma_8 \rangle(T)$ ; see below.

*d. Critical meson masses in a mean-field calculation*

Although the method is crude, it is instructive to get a first estimate for critical meson/quark masses. Later the results will be compared with estimates from a large-*N* approach. In a mean-field calculation the full effective potential is replaced by the classical part in terms of two constant background fields  $\sigma_0, \sigma_8$ . For simplicity we consider here only the SU(3)-symmetric case, where  $\epsilon_8 = 0 = \sigma_8, \epsilon_0 \neq 0, \sigma_0$  denotes a constant background field. The effect of a finite (high) temperature in a mean-

field calculation is renormalization of the quadratic term in the Lagrangian. Thus a finite temperature can be mimicked by tuning  $\mu_0^2$  while keeping the other couplings  $f_1, f_2, g$  fixed. For a critical field  $\epsilon_0^{\text{crit}}$ , the first-order transition simply disappears, and so does the cubic term in  $U_{\text{class}}$ . At  $\epsilon_0^{\text{crit}}$   $U_{\text{class}}$  starts with a term proportional to  $(\sigma_0 - \sigma_0^{\text{crit}})^4$ , where  $\sigma_0^{\text{crit}}$  is the minimum of  $U_{\text{class}}$  for critical values  $\mu_0^{2\text{crit}}, \epsilon_0^{\text{crit}}$ . Thus we have

$$U_{\text{class}}(\sigma_0) = -\frac{1}{2} \mu_0^2 \sigma_0^2 + \frac{2g}{3\sqrt{3}} \sigma_0^3 + \left(f_1 + \frac{f_2}{3}\right) \sigma_0^4 - \epsilon_0 \sigma_0, \quad (4.55a)$$

$$U_{\text{class}}|_{\text{crit}}(\sigma_0) = \frac{1}{4!} \left. \frac{\partial^4 U_{\text{class}}}{\partial \sigma_0^4} \right|_{\text{crit}} (\sigma_0 - \sigma_0^{\text{crit}})^4 + o(\sigma_0^5). \quad (4.55b)$$

Here  $|_{\text{crit}}$  means “evaluated at critical parameters.” Note that  $U_{\text{class}}$  in Eq. (4.55a) takes the same form as a free-energy functional for a liquid/gas system. It supports the analogy between a liquid/gas system and the chiral transition in QCD as mentioned in the dictionary of correspondences of Table II. The vanishing of the first three derivatives in Eq. (4.55b) determines  $\sigma_0^{\text{crit}}, \mu_0^{2\text{crit}}$ , and  $\epsilon_0^{\text{crit}}$  as functions of  $f_1, f_2$ , and  $g$ . In the physical case of  $\epsilon_0 = 0.0265$  [GeV<sup>3</sup>],  $\epsilon_8 = -0.0345$  [GeV<sup>3</sup>] we obtain

$$\frac{m_{u,d}^{\text{crit}}}{m_{u,d}} \sim 0.03 \pm 0.02. \quad (4.56)$$

Such a small ratio of 3% for the critical to physical light-quark masses would mean that the chiral phase transition is easily washed out by tiny quark masses, and for physical quark masses one is left with a rather smooth crossover phenomenon.

Our interest in the mean-field result is the order of magnitude of this ratio. Recall that a first-order transition can have different origins. One such origin is a cubic term in the classical part of the potential. A second is a  $\phi^6$  term which may be needed for stabilization of the free energy when the quartic coupling picks up a negative sign due to renormalization effects. For two or more independent relevant couplings, a further first-order transition can be a so-called *fluctuation-induced* transition. Since the linear SU(3)×SU(3) sigma model con-

tains two such couplings,  $f_1$  and  $f_2$ , the chiral transition may be mainly fluctuation induced. This hypothesis has been recently discussed by Gavin *et al.* (1994a, 1994b); see also Jacobsen and Amit (1981), Paterson (1981), Shen (1993).

If the order of magnitude of the ratio (4.56) changes beyond the mean-field level, it casts some doubt on the simplified description of Eq. (4.55a) and favors the hypothesis of a fluctuation-induced transition.

#### e. The large- $N_f$ approach

Ultimately we are interested in the temperature dependence of the meson condensates  $\langle \sigma_0 \rangle(T)$ ,  $\langle \sigma_8 \rangle(T)$ . The condensates are determined as the minima of the constrained free-energy density under the constraint on the average values

$$\frac{T}{V} \int_0^{1/T} d\tau \int d^3x \sigma_{0,8}(\mathbf{x}, \tau) = \bar{\sigma}_{0,8}, \quad (4.57)$$

where  $\bar{\sigma}_{0,8}$  are prescribed values for the background fields, while the same average value should vanish for  $\sigma_\ell$ ,  $\ell=1, \dots, 7$  and  $\pi_\ell$ ,  $\ell=1, \dots, 8$ . The relation to the full partition function is given by

$$\begin{aligned} Z &= \int d\sigma_0 \int d\sigma_8 \exp\{-\beta V U_{\text{eff}}(\sigma_0, \sigma_8)\} \\ &= \int d\sigma_0 \int d\sigma_8 \int \mathcal{D}\phi \text{ constraint} \times e^{-\int d^4x \mathcal{L}[\phi]}, \end{aligned} \quad (4.58)$$

with ordinary integrals for the  $\sigma_{0,8}$  integrations.

As a next step we convert the quartic part of the interaction to quadratic form by applying a matrix version of the Hubbard-Stratonovich transformation (Stratonovich, 1958; Hubbard, 1959; Frei and Patkós, 1990). It reads

$$\begin{aligned} &\exp\{-\beta[f_1(\text{Tr}\phi' \phi'^+)^2 + f_2 \text{Tr}(\phi' \phi'^+)^2]\} \\ &= \text{const} \int_{c-i\infty}^{c+i\infty} \mathcal{D}\Sigma(x) \exp\{\text{Tr}\Sigma^2 + 2\epsilon \text{Tr}(\Sigma \phi' \phi'^+) \\ &\quad + 2\alpha \text{Tr}(\phi' \phi'^+) \text{Tr}\Sigma\}, \end{aligned} \quad (4.59a)$$

where  $\phi'(x)$  is an  $N \times N$  matrix field and

$$\begin{aligned} \epsilon^2 &= \beta f_2, \\ 2\epsilon\alpha + 3\alpha^2 &= \beta f_1. \end{aligned} \quad (4.59b)$$

The path integral over the auxiliary field  $\Sigma(x)$  is evaluated in a saddle-point approximation. We replace  $\Sigma(x)$  by  $\text{sad} \times \text{diag}(1,1,1)$ . The saddle-point approximation corresponds to the leading term in a  $1/N$  expansion in an  $O(N)$  model (Coleman *et al.*, 1974). The  $SU(3) \times SU(3)$  linear sigma model reduces to an  $O(18)$  model for  $g=f_2=0$ . The  $N=18$  mesonic modes correspond to  $N_f=3$  flavors ( $N=2N_f^2$ ). Therefore we call our approximation ‘‘large- $N_f$ .’’

The advantage of the large- $N_f$  approximation is that we end up with an effectively free field theory. The only

remnant of the interaction is hidden in the dispersion relation

$$\omega_Q^2 = p^2 + \text{sad} + \mu_0^2 + m_Q^2. \quad (4.60)$$

Here  $Q=1, \dots, 8$  labels the particle multiplets,  $m_Q^2$  are the mass squares defined by the quadratic terms in the fluctuating fields, and  $\text{sad}$  is the contribution from the auxiliary field.

The final expression for  $U_{\text{eff}}(\sigma_0, \sigma_8, \text{sad})$  contains a classical part  $U_{\text{class}}$ , a part coming from the quadratization  $U_{\text{saddle}}$ , a zero-point energy part that is dropped, and a thermal part  $U_{\text{th}}$ ,

$$\begin{aligned} U_{\text{eff}}(\sigma_0, \sigma_8, \text{sad}) &= U_{\text{class}} + U_{\text{saddle}} + U_{\text{th}} \\ U_{\text{th}} &= \frac{1}{\beta} \sum_{Q=1}^8 g(Q) \int \frac{d^3p}{(2\pi)^3} \ln(1 - e^{-\beta\omega_Q}), \end{aligned} \quad (4.61)$$

where  $g(Q)$  are the multiplicities of the multiplets, e.g.,  $g(1)=3$  for three pions, etc. The full expression for  $U_{\text{eff}}$  can be found in Meyer-Ortmanns and Schaefer (1996).

The expression for  $U_{\text{eff}}(\sigma_0, \sigma_8, \text{sad})$  is evaluated in a high-temperature expansion and—alternatively—fully numerically. The high-temperature expansion for a free field theory is standard. Since we are interested in the low-temperature phase ( $T \leq T_c$ ), the most we can expect from a high-temperature expansion are qualitative results. The numerical evaluation of  $U_{\text{eff}}$  looks quite straightforward, but it is hampered by imaginary parts in the effective potential. Although the final maximum in the saddle point and the minima in  $\sigma_0, \sigma_8$  turn out to lie in the region of real-valued  $U_{\text{eff}}$ , the routines encounter imaginary parts in intermediate steps. The choice of initial values for the saddle point,  $\sigma_0, \sigma_8$  has to be optimized to make the numerical evaluation tractable.

#### f. Results for critical quark/meson masses

In the large- $N_f$  approximation the chiral transition is washed out for an average pseudoscalar octet mass  $\geq 203$  [MeV] if the ratio of  $m_s/m_{u,d}$  is kept fixed at its realistic value 18.2. This gives an upper bound on the first-order transition region. For the corresponding critical quark masses we find

$$\begin{aligned} m_{u,d}^{\text{crit}} &\leq 2.96 \pm 0.85 \text{ MeV}, \\ m_s^{\text{crit}} &\leq 54 \pm 15.4 \text{ MeV}, \end{aligned} \quad (4.62a)$$

or a ratio of

$$m_{u,d}^{\text{crit}}/m_{u,d} \sim 0.26 \pm 0.08. \quad (4.62b)$$

In the  $SU(3)$ -symmetric case with three degenerate flavors, the common critical pseudoscalar mass is only  $\leq 51$  MeV, and  $m_{u,d}^{\text{crit}} \leq 0.9 \pm 0.14$  MeV. Thus the critical mass values depend on the direction in mass parameter space. Our values are clearly below the estimates for Wilson fermions ( $m_s^{\text{crit}} \geq 400$  MeV for  $m_{u,d} \sim 0$ , and  $m_{u,d,s}^{\text{crit}} \geq 140$  MeV in the degenerate case). They are of the same order as the estimates for staggered fermions ( $m_{u,d}^{\text{crit}} \leq 12$  MeV,  $m_s^{\text{crit}} \leq 50$  MeV) for the  $N_f=2+1$  case,

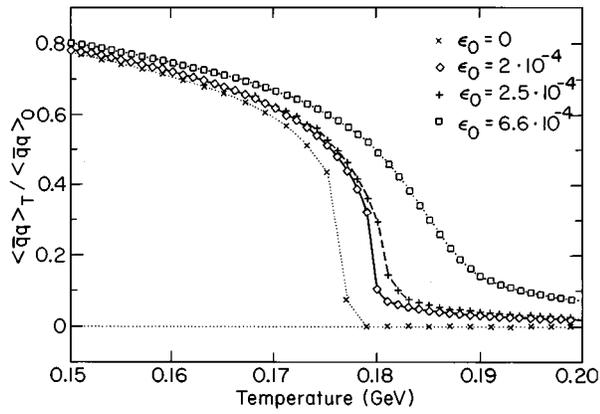


FIG. 21. The light-quark condensate normalized to its value at zero temperature  $\langle \bar{q}q \rangle_T / \langle \bar{q}q \rangle_0$  as a function of  $T$  in the SU(3)-symmetric case. The external field  $\epsilon_0$  is introduced as an explicit symmetry-breaking term in the action to induce nonvanishing meson/quark masses. The weakening of the first-order transition is obvious, when  $\epsilon_0$  [GeV<sup>3</sup>] is varied between  $\epsilon_0=0$  ( $\times$ ),  $2 \times 10^{-4}$  ( $\diamond$ ),  $2.5 \times 10^{-4}$  ( $+$ ), and  $6.6 \times 10^{-4}$  ( $\square$ ). From Meyer-Ortmanns and Schaefel (1996).

but below the bound for  $N_f=3, 12 \text{ MeV} \leq m_{u,d,s}^{\text{crit}} \leq 38$  (MeV), see Secs. III.C.5 and III.C.7. Thus our ratio  $m_{u,d}^{\text{crit}}/m_{u,d} \sim 30\%$  lies between the mean-field value of  $\sim 3\%$  and the lattice estimates. The intermediate value for the large- $N_f$  result is plausible, as “large- $N$ ” goes beyond “mean-field” due to the resummation of a certain subclass of Feynman diagrams (Jain, 1993), whereas lattice Monte Carlo simulations include all quantum fluctuations at once. It should be kept in mind, however, that a direct comparison between lattice quark masses in physical units and current quark masses (as we are using here) is questionable (see the argument in Sec. III.C.5 above).

A ratio of 30% is certainly not large enough for predicting visible remnants of a nearby second-order chiral transition. There is some hope that the ratio gets closer to 1 if further fluctuations are included in the effective model and the true nature of the chiral transition is fluctuation induced.

We have measured the weakening of the first-order transition in the SU(3)-symmetric case ( $\epsilon_8=0$ ). The gap in the light-quark condensate above  $T=177$  MeV slowly decreases for finite meson masses. It has disappeared for an external field strength of  $\epsilon_0=6.6 \times 10^{-4}$  GeV<sup>3</sup> (see Fig. 21). Similarly the barrier height between the coexisting minima in the effective potential decreases from  $1.4 \times 10^{-4}$  GeV/fm<sup>3</sup> in the chiral limit to  $2.1 \times 10^{-6}$  GeV/fm<sup>3</sup> for  $\epsilon_0=2 \times 10^{-4}$  GeV<sup>3</sup> and  $\epsilon_8=0$ .

We want to conclude with an interesting speculation of Gavin *et al.* (1994a) about the universality class of the SU(3)  $\times$  SU(3) sigma model. It is based on the observation that the mass of  $\sigma_{\eta'}$  vanishes for a specific choice of tree-level parameters  $f_1, f_2, g, \mu_0^2$ . If the renormalization-group flow were to drive the couplings towards these values in the vicinity of  $T_c$ , the light mass

could lead to large correlation volumes. If the  $\sigma_{\eta'}$  were the only light mode, the universality would be Ising-like, in contrast to  $O(4)$ .

#### g. An upper bound on the latent heat

For physical meson masses we find a sharp crossover phenomenon in the light-quark condensate between  $T=181.5$  and  $192.5$  MeV, in which  $\langle \bar{q}q \rangle_T$  decreases to 50% of  $\langle \bar{q}q \rangle_{T=0}$  over a temperature interval of  $\Delta T=10$  MeV. The strange-quark condensate  $\langle \bar{s}s \rangle_T$  stays almost constant up to a temperature of  $\sim 200$  MeV; see Fig. 22.

Similar results are found for the energy density and the entropy density (Fig. 23). Both quantities behave smoothly as a function of  $T$  over the entire temperature range up to large errors in the transition region. From these errors we obtain an upper bound on a finite latent heat  $\Delta L$  of

$$\Delta L \leq 0.2 \text{ GeV/fm}^3, \quad (4.63)$$

which is compatible with our data. The value is in agreement with Leutwyler’s bound of  $0.4 \text{ GeV/fm}^3$  (Leutwyler, 1992), which was obtained from Clausius-Clapeyron relations in the framework of chiral perturbation theory; see Sec. IV.A.3. Further note that this upper bound is only 10% of the value predicted by the naive bag model equation of state. To our knowledge it is still unclear whether  $0.2 \text{ GeV/fm}^3$  latent heat is sufficiently large to induce measurable signatures in heavy-ion experiments.

Furthermore we have calculated the difference  $\epsilon(T_2)/T_2^4 - \epsilon(T_1)/T_1^4 \equiv \Delta \epsilon / \langle T_c^4 \rangle$ . Here  $T_1=181.5$  MeV and  $T_2=192.6$  MeV denote the temperatures at which the rapid crossover sets in and ends, respectively. In a first-order transition with a finite discontinuity,  $T_1=T_2=T_c$ . Hence we can compare our value for  $\Delta \epsilon / \langle T_c^4 \rangle = 0.29$  with the ratio for the gluonic energy density  $\Delta \epsilon_{\text{gluonic}}/T_c^4$  on the lattice in a pure SU(3) gauge theory. The lattice result is  $\epsilon_{\text{gluonic}}/T_c \sim 1.5$  or  $\text{disc} \epsilon_{\text{gluonic}} \sim 0.9 \text{ GeV/fm}^3$  for  $T_c=260$  MeV as mentioned above. Our bound on the chiral contribution to the total gap in the energy density over  $T_c^4$  is smaller by almost an order of magnitude than the gluonic contribution. This should be taken as a warning not to jump to conclusions from our results concerning relativistic heavy-ion experiments. So far we have investigated only the contribution from chiral symmetry to the full equation of state for a hot hadron gas. The contribution from gluonic degrees of freedom has been completely left out. It is a question of relative size, which effects in the full deconfinement/chiral transition are dominant.

It seems to us very worthwhile to elaborate alternative inclusions of heavier mesons in the chiral transition region. From the work of Gerber and Leutwyler (1989) it has become clear that heavier mesons are increasingly important when  $T$  exceeds 100 MeV. I regard our parametrization of the scalar and pseudoscalar meson masses in the SU(3)  $\times$  SU(3) linear sigma model as just one ansatz for including a portion of the heavier mesons.

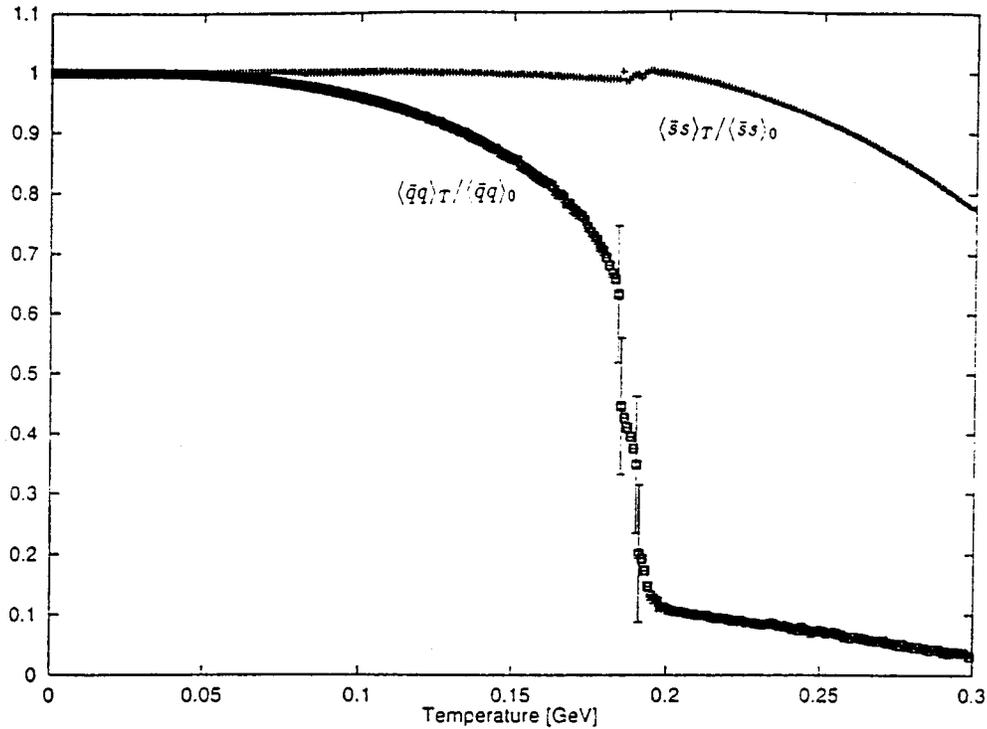


FIG. 22. Light ( $\langle \bar{q}q \rangle$ ) and strange ( $\langle \bar{s}s \rangle$ ) quark condensates normalized to their corresponding values at zero temperature as functions of temperature. The crossover behavior is most rapid in the range  $181.5 \leq T \leq 192.6$  [MeV]. From Meyer-Ortmanns and Schaefer (1996).

This portion is determined by the assumed underlying chiral symmetry. A further assumption is that deviations from the broken  $SU(3) \times SU(3)$  symmetry can be parametrized by two external fields, breaking the symmetry explicitly. Such an ansatz seems to be justified in the sense that the predictions of the pseudoscalar meson masses agree reasonably with experimental values.

An alternative point of view is the following.  $SU(2) \times SU(2)$  is the true symmetry with only pions as idealized Goldstone bosons and one external field to account for the finite pion mass, while all other mesons—

scalar, pseudoscalar, and vector mesons—are treated on an equal footing. Gerber and Leutwyler (1989) take this point of view when they describe all mesons (apart from the pion) in a dilute-gas approximation, in order to study their influence on the chiral phase transition. The *quantum virial expansion* of thermodynamic quantities in a hot pion gas provides a further alternative. Heavier mesons enter the second virial coefficient via their contribution to experimentally measured phase shifts of  $\pi\pi$  scattering (Welke *et al.*, 1990). A more general framework is the generalized Beth-Uhlenbeck approach (see, for example, Schmidt *et al.*, 1990), in which the scattering phase shifts of  $\pi\pi$  scattering are replaced by thermodynamic  $T$  matrix elements, characterizing the interacting pion gas at finite temperature and density. When the quantum virial expansion or the generalized Beth-Uhlenbeck approach are applied to a pion gas, the underlying chiral symmetry of QCD plays a less prominent role in the description than it does in our treatment in the  $SU(3) \times SU(3)$  linear sigma model.

## B. Models for gluonic degrees of freedom

### 1. A network of strings

The color flux-tube models of Patel (1984a, 1984b) were developed to describe the deconfinement phase transition. Quark degrees of freedom can be included as well. Let us first consider the case of a pure  $SU(2)$  or  $SU(3)$  gauge theory. We recall from Sec. III that the spontaneous breaking of the global  $Z(2)$  or  $Z(3)$  sym-

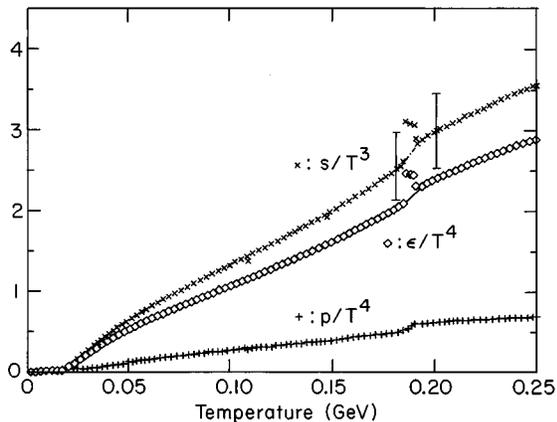


FIG. 23. Large- $N_f$ -approximation:  $\times$ , Entropy density  $s$  over  $T^3$ ;  $\diamond$ , energy density  $\epsilon$  over  $T^4$ ;  $+$ , pressure  $p$  over  $T^4$ . The meson masses have been chosen at their physical values. Ordinate quantities are dimensionless.

metry is associated with a phase transition from the confinement to the deconfinement phase. The order parameter is the expectation value of the Wilson line. It vanishes when the free energy to find an isolated test quark in the system grows to infinity. It is different from zero in the deconfinement phase.

Effective  $Z(N)$  spin models for the deconfinement transition have been presented above (Sec. III). They can be derived from the  $SU(N)$  lattice gauge theory in the high-temperature limit and share the essential extra symmetry of QCD. In the limit of infinitely heavy masses, this is the global  $Z(N)$  symmetry. For high temperatures the gauge fields are almost frozen to unity matrices. Lowering the temperature increases the disorder. One big cluster of aligned spins breaks up into several clusters. Below the transition point the system is, if we use “magnetic” language, completely disordered. The formulation in terms of spin systems is natural for  $Z(N)$  models. Intuitively, however, it is much less clear how this picture should be translated to the original  $SU(N)$  degrees of freedom.

Patel’s description is complementary to the above model. It starts at  $T=0$  and follows the evolution of the system as the temperature is increased. Symmetry breaking is no longer the driving force for the system to undergo a phase transition. It is replaced by entropy production.

Below we present a heuristic description. The picture is based on the flux-tube model of the deconfinement transition. Flux tubes connect quarks and antiquarks in strong-coupling expansions on the lattice. Thus one would not expect that they leave some remnant in the continuum limit. On the other hand, they are also the ingredients of phenomenological continuum descriptions like string models for hadrons. We add a few remarks about the yo-yo string in the end.

Several properties of flux tubes or strings have to be specified first. Strings in Patel’s models are characterized by three parameters: the string tension  $\sigma$ , the string width  $w$ , and the rigidity parameter  $a$ . [The notation of the parameter  $a$  suggests its actual meaning as an (effective) lattice constant.] We assume that there is a constant energy per unit length along the string; this is  $\sigma$ . The string has a constant width  $w$ . It has a certain resistance against bending. It has to go at least a distance  $a$  apart, before it can change its orientation. The distance  $a$  is of the order of 1 fm. On the lattice  $a$  naturally coincides with the lattice constant, at which the strings bend at right angles. Here the constant  $a$  will not be tuned to zero in the end. Its role resembles that of a lattice constant in models of condensed-matter physics (models in continuum spacetime, where the constant  $a$  is given by the physical lattice constant). In Patel’s flux-tube theory the parameters  $a$ ,  $\sigma$ , and  $w$  must be fixed from experimental input.

Two further assumptions about flux tubes must be specified to get a well-defined model. Flux tubes can terminate only on quarks. Their interaction occurs at baryonic vertices. The constants  $\sigma, w, a$  are treated as temperature independent. The only driving force of the

phase transition is the increasing entropy of the flux tubes when the temperature is turned on.

#### a. The $SU(2)$ case

Based on these ingredients let us see how far heuristic arguments can lead us in a pure  $SU(2)$  gauge theory. The only allowed flux-tube structures are closed loops differing only in size and shape, i.e., in the length of the string. Physically these loops may be interpreted as glueballs. The partition function is written as

$$Z = \sum_{\text{loops}} N(\text{loops}) e^{-E(\text{loop})/T}. \quad (4.64)$$

The sum runs over all loops of fixed length  $la$  (in physical units). The energy of such a loop is given as  $\sigma la$ . The combinatorial prefactor  $N$  gives the number of loops of length  $la$ . The dependence of  $N$  on the rigidity parameter  $a$  is essential, as it provides the possibility of a phase transition. How often strings like to bend is a function of temperature and rigidity.

Consider random walks without backstepping of length  $l$  (in lattice units) on a lattice of spacing  $a$ . Backstepping should be forbidden for a physical string. This number is given as  $(2d-1)^l$  in  $d$  dimensions. The constraint that the walks should perform closed loops leads to a power-law correction in  $l$  to  $N$ . It will not be specified further, because it is irrelevant in the large- $l$  limit. The partition function is then proportional to

$$Z \sim \sum_l \exp \left[ -l \left( \frac{\sigma a}{T} - \ln 5 \right) \right] \quad (4.65)$$

in three dimensions. The dimension is chosen as 3 since the string model is constructed to provide an alternative description to the three-dimensional  $Z(N)$  spin model for the deconfinement transition. As  $T$  increases, the average length of the loops becomes larger. A phase transition is signalled if  $Z$  diverges. This happens for large  $l$  if  $(\sigma \cdot a/T = \ln 5)$ , that is, at

$$T_c = \sigma a / \ln 5. \quad (4.66)$$

The effective string tension  $\sigma_{\text{eff}}$  defined via  $\exp\{-la\sigma_{\text{eff}}/T\}$  vanishes continuously at  $T_c$ ,

$$\sigma_{\text{eff}} = \sigma - \frac{T \ln 5}{a} \xrightarrow{T \rightarrow T_c} 0, \quad (4.67)$$

while the average length of the loops diverges. Both features suggest a second-order phase transition.

To see its relation to the deconfinement transition, we have to probe the system at  $T_c$  with a static  $\bar{q}q$  pair. If it costs a finite amount of energy to isolate both quarks at an infinite distance, we have reached the deconfinement phase. Equivalently deconfinement is manifest if the  $q$  and  $\bar{q}$  are a finite distance apart, but only rather loosely correlated. Color screening requires that a flux tube connect the quarks of our probe. With increasing temperature the flux reorients itself more and more often. It oscillates between its end points; see Fig. 24.

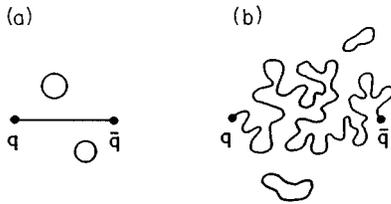


FIG. 24. Flux tubes between static quarks in pure SU(2) gauge theory: (a) at low temperatures; (b) at high temperatures.

Since one quark knows about its partner only via the connecting string, this information is lost if the string becomes infinitely long. This is just what happens at  $T_c$ , where it costs no extra free energy to create an infinitely long flux tube. Thus at  $T \geq T_c$  both quarks ( $q$  and  $\bar{q}$ ) are effectively independent of each other and free. This means deconfinement.

To get a quantitative estimate of the transition temperature, we must fix the string tension and the rigidity parameter in physical units, or the product of both in lattice units. In the strong-coupling approximation of the lattice theory, the lowest-lying  $O^+$  glueball at  $T=0$  is given by a square loop

$$m_{O^+} = 4\sigma a. \tag{4.68}$$

The string parameters are assumed to be temperature independent, hence one can take this  $T=0$  result to fix  $T_c$ ,

$$T_c = m_{O^+} / 4 \ln 5. \tag{4.69}$$

Lattice Monte Carlo calculations give a slightly larger value.

*b. The SU(3) case*

Next let us consider the pure SU(3) gauge theory. Two new features must be accounted for. Quarks and antiquarks are no longer in equivalent representations of SU(3). Thus a direction is associated with a string, indicating whether it terminates in a quark or an antiquark. (The quarks and antiquarks are test quarks in the pure gauge theory.) Furthermore, a flux-tube representation of a baryon or antibaryon requires that a string be able to bifurcate at a vertex  $v$ ; see Fig. 25. This reflects the previous assumption that flux tubes interact only at baryonic vertices. The diagrammatic rules for allowed string structures in the SU(3) case follow from the allowed vertices, shown in Fig. 26. This excludes closed loops made up of an odd number of links.

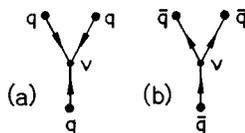


FIG. 25. String bifurcation for a baryon (a) and an antibaryon (b).



FIG. 26. Allowed vertices for the SU(3) gauge theory.

The ansatz for the partition function can be chosen analogously to the SU(2) case. The sum over loops has to be replaced by a sum over more complicated topological structures which resemble nets. A nonvanishing energy  $v$  is associated with each vertex to respect the additional bifurcation degree of freedom. At zero temperature the vacuum is filled with closed loops made out of nets of strings; see Fig. 27. These are virtual “glueballs.”

As the temperature is increased, the size and density of these structures grow. In contrast to the SU(2) case, the phase transition is not induced by a divergence in the length of strings. A new qualitative feature enters, the *connectivity* of the network. A phase transition occurs if an infinite network is generated, infinite in the sense that each string is connected with the entire volume of the lattice. Between any given pair of flux-tube segments one can find a flux-tube path on the lattice that connects them.

The definition of “critical temperature” can be made more precise. This type of phase transition is well known in condensed-matter physics as a percolation transition. Examples of percolation transitions are gelation transitions in the context of polymer chains (Flory, 1941a, 1941b). To find a quantitative measure for the connectivity of the network, one has to distinguish between relevant and irrelevant links of the network. If relevant links are removed, the infinite network is destroyed. In the same sense irrelevant links are not essential for connectivity. The fraction  $f$  of relevant links depends on the temperature and the specific underlying dynamics. It can be easily estimated that  $2/3 < f < 1$ . The lower bound is realized at  $T_c$ , where the network is minimally connected. The upper bound means maximal connectivity.

It is challenging to look at the *deconfinement transition as a percolation transition*. Again we have to answer, first, what happens to a static test quark when it is put into the network below or above the percolation transition. Below  $T_c$  the free energy of an isolated quark is infinite, because the flux tube originating from this quark may not terminate on any closed loop that is a finite distance apart. It is excluded by the diagrammatically forbidden flux-tube structures. Instead the string attached to the quark has to fuse with the string of an

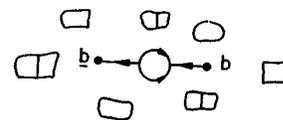


FIG. 27. SU(3) vacuum at low temperatures filled with glueballs.

antiquark of the test probe. When both  $q$  and  $\bar{q}$  can be considered as isolated or free, to separate them that much costs an infinite amount of free energy, since we have assumed a constant energy per unit length of the string.

Above  $T_c$  the distance between the  $q$  and  $\bar{q}$  need no longer be infinite for both to lose their correlation (to behave effectively as free test probes). Already at a finite separation the correlation is weak due to the presence of the infinite network to which they are attached. Now the clusters of closed loops are connected, and the corresponding flux-tube structures for a  $q\bar{q}$  pair at finite distance are allowed (see Fig. 28). The costs in energy due to the long, dense, and connected flux tubes are compensated by the gain in entropy. Thus the free energy of a configuration in which  $q$  and  $\bar{q}$  test quarks have lost any information about each other is finite. The  $q$  and  $\bar{q}$  are effectively free. We have ended in the deconfinement phase.

For a crude estimate of  $T_c$  one has to specify the amount of extra energy due to bifurcation points.

In our heuristic argument it remains to explain the first order of the deconfinement phase transition in the SU(3) case. Either one refers to the percolation transition where the transition is known to be of first order, or one proves directly for the colored network of the QCD flux that it costs a finite interface free energy at  $T_c$  for both phases to coexist. Such a coexistence may be visualized as a hole in the network. Roughly speaking, when the loose ends at the boundary of the hole are tied to the network again, it costs interface energy proportional to the area and leads to a reduction in entropy.

### c. Inclusion of matter fields

Dynamical quarks can be easily introduced into flux-tube models, since the flux tube picture does not rely on any symmetry argument, and the extra global  $Z(N)$  symmetry is explicitly broken in the presence of dynamical quarks. Recall that the limit of a pure gauge theory may be considered as the  $m \rightarrow \infty$  limit of full QCD. Thus we are interested in the effect when the quark mass is lowered or a  $Z(N)$  symmetry-breaking field is switched on.

Second-order phase transitions in ferromagnets are known to disappear when the magnets are exposed to an external magnetic field, regardless of its strength. Similarly the second-order transition of the pure SU(2) gauge theory ceases to occur: there is always a finite probability of an infinitely long string's breaking up due to the creation of a  $q\bar{q}$  pair out of the vacuum. The probability is proportional to  $le^{-2m/T}$ ; it is finite, even if the quark mass is very large, as long as  $l \rightarrow \infty$ . (Infinitely long strings turned out to be responsible for the second order phase transition in this picture.) The expected change is that the phase transition from the confinement to the deconfinement region is replaced by a crossover from tightly to weakly bound  $q\bar{q}$  pairs.

The SU(3) case is more subtle. Again it is known from statistical physics that first-order transitions are

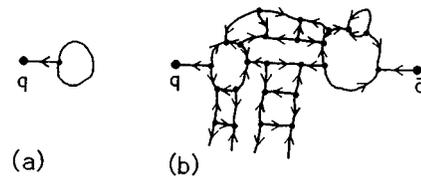


FIG. 28. Flux-tube structures: (a) disallowed structure below  $T_c$ , (b) allowed structure above  $T_c$ .

stable with respect to small perturbations of an external field. Hence we still expect for large enough quark masses a deconfinement transition, and it is still the Wilson line expectation value which would indicate the transition by a discontinuity at  $T_c$ . However,  $\langle L \rangle$  will fail to remain a good order parameter when the quark mass is further lowered. It was sensitivity to  $Z(N)$  symmetry which qualified  $\langle L \rangle$  as an order parameter in the  $m \rightarrow \infty$  case. Now, where the symmetry is explicitly broken independently of the phase, we have to look for new criteria that tell us the phase of the system at a given temperature. This is the screening property, which is different in the high- and low-temperature phases. At low temperatures, color charges are screened due to breakage of strings. At high temperatures, flux tubes do not break, but get attached to large networks of strings, at least for very heavy masses. For lighter masses, quarks can also break strings at high temperature. (We see that the clear distinction between low- and high- $T$  phases may get lost depending on the value of the quark mass.)

The probability of string breakage at high temperature is determined by the energy costs of popping a  $q\bar{q}$  pair out of the vacuum and the competing gain in entropy by breaking a link. Lighter masses facilitate breaking of links. Less surface energy is necessary to stabilize a hole in an infinitely connected network. Clearly there should be a critical mass at which the surface energy vanishes at  $T_c$  and the number of broken flux tubes exceeds the critical number of irrelevant links. Once relevant links are broken, by definition, the network ceases to exist.

In Patel's models one has a heuristic understanding of how the transition temperature should scale as a function of quark masses. As the quark mass decreases, string breakage is facilitated. It becomes more difficult to generate an infinitely connected network. Thus  $T_c$  should increase with decreasing  $m$ . The opposite tendency is observed in numerical simulations.

The above line of argument suggests that the transition will always disappear in the  $m \rightarrow 0$  limit. The energy cost for the creation of a  $q\bar{q}$  pair is zero. Recall, however, that it is precisely in the  $m \rightarrow 0$  limit where chiral symmetry comes into play and generates dynamical masses at low temperatures. In the chiral limit, scalar mesons are believed to obtain their finite masses exclusively from chiral symmetry breaking. In the preceding sections we have investigated just the opposite effect: perturbing around the chiral limit, turning on an external field by increasing  $m$  from zero towards realistic values.

In the chiral limit the strength of the first-order transition is thought to increase for an increasing number of flavors. On the other hand, in the same chiral limit, there is a gain in entropy due to easy string breakage as a remnant of the deconfinement transition. It is difficult to estimate whether a critical number of flavors exists that will enable one of the competing effects to win.

The combination of two tendencies may result in a crossover phenomenon at intermediate masses. The quark masses are so small that the deconfinement transition has disappeared, and an infinitely connected network can no longer be formed, yet they are too large for effects of chiral symmetry breaking to be important, and the chiral transition is washed out as well. It remains to be seen whether realistic quark masses occur in a window where both transitions are replaced by crossover phenomena.

As long as chiral symmetry is not implemented in the massless limit, such competing effects cannot be resolved by Patel's models, even if matter fields are included.

Our presentation so far might have suggested that a description of the deconfinement transition in terms of color flux tubes gets stuck on the heuristic level. This is not the case. A three-dimensional action can be formulated in terms of occupation number variables (Patel, 1984b). These variables live on a lattice with lattice constant  $a$  and take on values in the positive integers. The flux-tube variables are associated with the links, the quarks and vertex variables with the sites of the lattice. The input parameters for the action are the string tension  $\sigma$ , the lattice constant  $a$ , the quark mass  $m$ , and the vertex coupling  $v$ . If Gauss's law is implemented, the model is equivalent to a three-dimensional  $XY$  model with nearest-neighbor coupling in a uniform magnetic field  $h$  and a magnetic field  $p$ . Mean-field calculations in various limiting cases of this action support the heuristic arguments (Patel, 1984b). The model can be extended to include several numbers of flavors and chemical potentials.

Let us summarize the virtues of the color flux-tube model. It abandons symmetry breaking as a driving force of the phase transition, at least as a primary force. This is desirable as an alternative way of understanding the transition dynamics in mass parameter regions, where the symmetry concept is questionable. Here the deconfinement transition has been described in analogy to a gas/liquid transition at finite temperature for different values of the pressure. The pressure plays the role of the quark masses. The order parameter is the density or the volume. Its change is discontinuous when the temperature is increased, as long as the pressure is low enough. In the QCD case,  $\langle L \rangle$  jumps at the transition as long as the quarks are heavy enough. As in the QCD case with nonzero masses, there is no symmetry in the gas/liquid system that could easily distinguish between both phases. In a  $(p, T)$  diagram the liquid and gas phases may be smoothly connected as well. For analogies we refer the reader to Table II.

The deconfinement transition here resembles a gelation transition, in which the connectivity of the color flux-tube network is an essential characteristic of  $SU(3)$ . A similar intuitive understanding of the chiral transition is still lacking when it is described in the framework of the sigma model. For the chiral transition it is less clear what the competing effects to string breakage or network formation could be.

It seems rather worthwhile to incorporate chiral symmetry properties in color flux-tube models or vice versa to supplement sigma models with flux tubes for the gluonic degrees of freedom.

We conclude with a few remarks on *phenomenological string models* for hadrons. Experimentalists describe the phenomenology of heavy-ion collisions with Monte Carlo codes that are based on relativistic microscopic models. String formation and fragmentation are essential ingredients. These models share some features with cellular automata. One tries to find a set of rules for string formation and breakage (fragmentation) such that observed particle multiplicities and  $\langle p_T \rangle$  spectra are reproduced for given initial conditions. An example of such a string model is the VENUS code of Werner (1989), utilizing the concept of yo-yo strings. Strings are formed via *color exchange*.

One may suspect that yo-yo strings in VENUS share only the name, and so far little else, with Patel's color flux-tube strings. The yo-yo string is elastic, its elasticity characterized by the string tension. Closed loops are forbidden; the flux can terminate only on colored partons (quarks, diquarks etc.). The end-points are characterized by their flavor content. Yo-yo strings are allowed to stretch, break, and participate in a collective motion. They cannot bend or bifurcate. There is no space for thermodynamic concepts, for a transient plasma in the early stage of the collision, and no way to model a phase transition in the present formulation.

Patel's strings are neither one-dimensional rigid tubes connecting the colored end points, nor arbitrarily pliable. The rigidity constant is essential for modulating the entropy and initiating the phase transition. The range of possible applications of VENUS-type models may be extended if the strings are not restricted to yo-yos, but allowed to close, bend, and bifurcate.

## 2. Dual Ginzburg-Landau models

The main idea in dual Ginzburg-Landau models is to explain the confinement mechanism by a dual Meissner effect. Let us recall the Meissner effect in a superconductor. The magnetic field is expelled from the embedding superconducting medium and concentrated in a flux tube between magnetic monopoles of opposite charge. The constant energy per unit flux tube causes the interaction between monopole charges to increase linearly for large distances. In the superconducting state of metal, the electrons generate Cooper pairs. The superconducting medium is characterized by an order parameter  $\psi$ , and  $|\psi|^2$  represents the density of Cooper pairs. The order parameter vanishes in the normal conducting

state where Cooper pairs are absent. The transition to the superconducting phase is triggered by the condensation of Cooper pairs.

A dual Meissner effect may be an appealing explanation for confinement. In the dual Meissner effect the role of the (chromo)electric and (chromo)magnetic fields is reversed. The chromoelectric field between the chromoelectric charges of quarks and antiquarks is expelled from the embedding confinement vacuum. The chromoelectric flux tube leads to a linear rise in the energy between quarks at large distances. The transition to the confinement phase is triggered by a condensation of chromomagnetic monopoles.

The interest in models based on such a heuristic explanation has recently revived. The dual Meissner effect seems to be realized in compact lattice QED (Polyakov, 1975; Banks *et al.*, 1977; DeGrand and Toussaint, 1980; Stack and Wensley, 1992). The main task in QCD is to identify the magnetic monopoles. This identification should be possible ('t Hooft, 1981) after an Abelian projection. An Abelian projection of QCD amounts to a partial gauge fixing. There are many ways of extracting an Abelian theory out of non-Abelian QCD. Popular choices are a variety of unitary gauges and the maximal Abelian gauge (Kronfeld *et al.*, 1987; Suzuki and Yotsuyanagi, 1990, 1991; Suzuki, 1993). The maximal Abelian gauge turns out to be a particularly useful way of demonstrating the dual Meissner effect. If the “non-Abelian part” of the gauge freedom is fixed, the original SU(3) gauge symmetry is reduced to the maximal Abelian torus group  $U(1) \times U(1)$ . QCD becomes an effective theory of abelian “electric” charges, “photons,” and “magnetic” monopoles with respect to  $U(1) \times U(1)$ . The photon fields are the regular part of the Abelian field, the magnetic monopoles the pointlike singular part. The monopole part is isolated from the U(1) field following a procedure proposed by DeGrand and Toussaint (1980). 't Hooft's conjecture about the dominant role of monopoles seems to be realized in the maximal Abelian gauge. The conjecture may be verified in several steps.

(i) Typical indicators of confinement are the string tension, the Wilson loop, and the Polyakov loop. It is natural to test whether their expectation values from full lattice QCD can be reproduced by the Abelian variables alone. This is indeed what is observed. The phenomenon is called *Abelian dominance* (Suzuki and Yotsuyanagi, 1990, 1991; Hioki *et al.*, 1991; Suzuki, 1993; Matsubara *et al.*, 1994).

(ii) More specifically one would like to know which part of the Abelian residual variables is responsible for the observed Abelian dominance. Here it turns out that the monopole contributions alone determine the string tension and the Polyakov loop in an SU(2) and SU(3) pure gauge theory (Shiba and Suzuki, 1994; Kitahara *et al.*, 1995, 1995b; Ejiri *et al.*, 1995; Matsubara *et al.*, 1995; Miyamura, 1995). The similar behavior of Abelian and conventional Polyakov loop expectation values is seen not only in the maximal Abelian gauge, but also in a number of different U(1) projections.

(iii) Once the relevant degrees of freedom have been identified, the next question concerns their dynamics. Which monopoles are important in the confinement and deconfinement phases? What drives the phase transition? Here some partial answers are known. In the confinement phase one has a long monopole in each configuration, while long monopole configurations are absent in the deconfinement region (Kitahara *et al.*, 1995a, 1995b). This is easily understood from the main difference between low- and high-temperature QCD. The difference in the monopole dynamics comes from the bound on the maximal monopole extension due to the large or small number of time slices. The driving mechanism for the deconfinement transition is supposed to be the balance of the entropy and energy of maximal extended monopoles (see the preceding section for a similar explanation in the flux-tube model). Further studies are, however, necessary to clarify the calculation of the entropy of monopoles (Kitahara *et al.*, 1995a, 1995b). The difference in the order of the deconfinement transition for SU(2) and SU(3) remains to be explained.

(iv) The observation of Abelian dominance and the identification of long monopoles as the most important field configurations in the confinement mechanism suggest construction of an effective action in terms of these monopoles. Effective monopole actions have been considered in compact QED and SU(2) QCD (Shiba and Suzuki, 1994a, 1994b, 1995). Extended monopoles can be implemented as well by performing a block spin transformation on the dual lattice. The results obtained so far indicate an interesting scaling behavior, giving some hope that the effective action actually describes continuum physics.

The results in dual Ginzburg-Landau models look rather encouraging. The dependence on the finite volume deserves further study on larger lattices to make sure that the dynamics of Abelian  $U(1) \times U(1)$  monopoles of lattice QCD survive the continuum limit. The *gauge dependence* of the proposed confinement mechanism is another open question. So far most of the features of the dual Meissner effect have been observed in the maximal Abelian gauge. Recently Chernodub *et al.* (1995) have studied the pure SU(2) gauge theory in different Abelian projections. Their results suggest the conclusion that the topological configurations relevant for confinement do depend on the gauge. *Monopole condensation* as a mechanism for confinement appears as a specific feature of the maximal Abelian gauge. So-called minopoles replace monopoles in the “minimal” Abelian gauge, and stringlike topological objects may be important as well. Such a translation of the confinement mechanism between different gauges deserves further investigation. Remarks on unitary gauges can be found in Matsubara *et al.*, (1995) and Suzuki and Yotsuyanagi, (1991).

### 3. Some further approaches

We conclude this section about effective models with a look at some possible future directions. Descriptions

including quark degrees of freedom have been left out so far. Moreover the string networks of the preceding section cannot be regarded as representative of the variety of models that have been proposed for gluonic degrees of freedom.

An important class of models for the gluonic sector are *instanton descriptions* of the QCD vacuum (Shuryak, 1984). The idea is that the most important (nonperturbative) contributions to the QCD partition function come from instantons and anti-instantons at short distance scales. Dyakonov and Mirlin (1988) and Kanki (1988) have considered the instanton vacuum at finite temperature. Within a variational approach they have shown that the “evaporation” of instantons and anti-instantons out of the vacuum leads to chiral symmetry restoration as the temperature is increased.

Igenfritz and Shuryak (1989) and Nowak *et al.* (1989) have extended the analysis of instanton models to the inclusion of quark degrees of freedom. The qualitative conclusion of these authors is a crossover phenomenon rather than a chiral transition, when finite quark masses are included. Their results were obtained within mean-field-type calculations.

Frequently it is the symmetry of the underlying QCD action that is taken as a guiding construction principle for the effective action. Applying this principle to the gluonic sector, one may look for the analog of the chiral condensate and the role it plays in the deconfinement (and chiral) transition. A natural analog is the gluon condensate  $\langle F_{\mu\nu}F^{\mu\nu} \rangle$ , if  $F_{\mu\nu}$  denotes the QCD field-strength tensor.

With regard to the gluon condensate we recall the basic symmetry of QCD related to the condensate. At the classical level the QCD Lagrangian with massless quarks is invariant under scale transformations  $x^\mu \rightarrow e^\lambda x^\mu$  with some real constant  $\lambda$ . The corresponding (classically conserved) current is the dilatation current  $S_\mu = T_{\mu\nu}x^\nu$ , where  $T_{\mu\nu}$  is the energy-momentum tensor. On the quantum level, even massless QCD is no longer scale invariant. It is broken by the scale or trace anomaly of QCD,

$$\partial_\mu S^\mu = T^\mu_\mu = \frac{\beta(g)}{2g} F_{\mu\nu}F^{\mu\nu}. \quad (4.70)$$

Here  $\beta(g)$  denotes the beta function of QCD, and  $g$  is the coupling. The perturbative contribution to one-loop order is given as  $(-g^3/16\pi^2)(11 - 2/3 N_f)$ .

Schechter (1980) has proposed an effective action for both mesonic and gluonic degrees of freedom. In the limit of zero quark masses this action is invariant under the  $SU(N_f)_L \times SU(N_f)_R$  chiral transformation ( $\times U(1)_V$  baryon number); it reproduces the axial  $U(1)$  anomaly (the chiral anomaly) and in addition the trace anomaly of QCD.

For definiteness we state the result for the effective Lagrangian, as used by Campbell *et al.* (1990):

$$L = L_m + L_g, \quad (4.71)$$

$$L_m = \frac{1}{16\pi^2} f\pi^2 (\chi/\chi_0)^2 \text{Tr} \partial_\mu U \partial^\mu U^+ - c (\chi/\chi_0)^3 \text{Tr} (m_q (U + U^+)) - \frac{1}{2} m_0^2 \phi_0^2 (\chi/\chi_0)^4, \\ L_g = \frac{1}{2} \partial_\mu \chi \partial^\mu \chi + B \left[ \frac{1}{4} \chi^4 + \chi^4 \ln(\chi/e^{1/4} \chi_0) \right].$$

The following notations are involved. Gluonic degrees of freedom are represented by a scalar gluonium field  $\chi$  with scale dimension 1. It is assumed to take a nonvanishing vacuum expectation value  $\chi_0 = \langle 0|\chi|0 \rangle$ , and thus scale invariance is spontaneously broken on the effective level. The matter fields are represented by  $U$ ,

$$U(x) = \exp \left( i \sum_{i=0}^8 \lambda_i \phi_i(x) / f_\pi \right). \quad (4.72)$$

The fields  $\phi_i$ ,  $i=0, \dots, 8$  denote the nine pseudoscalar meson fields. The light-quark mass matrix is given by  $m_q = \text{diag}(m_u, m_d, m_s)$ . Up to prefactors depending on  $\chi$  and  $\chi_0$ , the matter part of the Lagrangian is the familiar nonlinear  $SU(3) \times SU(3)$  sigma model (first term of  $L_m$ ), plus a term proportional to the quark masses' breaking chiral symmetry explicitly, plus a third term ( $\propto \phi_0^2$ ) accounting for the ninth pseudoscalar meson. All terms are multiplied by prefactors correcting for the right scale dimensions.

The specific form of the potential for the gluonic part  $L_g$  was proposed by Schechter (1980). This term was designed to guarantee the trace anomaly according to

$$\partial_\mu S^\mu = -B \chi^4, \quad (4.73)$$

where  $B$  is the vacuum energy density of the pure gauge sector. Equation (4.73) leads to the identification

$$\langle 0 | (\beta/2g) F_{\mu\nu} F^{\mu\nu} | 0 \rangle = -B \chi_0^4 X. \quad (4.74)$$

Campbell *et al.* (1990) used the Lagrangian (4.71) as the starting point for their thermodynamic considerations.

The matter part of the Lagrangian is easily modified. Patkós (1991) used the linear version of the  $SU(3) \times SU(3)$  sigma model for  $L_m$  to calculate the surface tension of the first-order chiral transition under the inclusion of gluonic degrees of freedom.

Kusaka and Weise (1992) have chosen the Nambu-Jona-Lasinio (NJL) Lagrangian with  $SU(2)_L \times SU(2)_R$  symmetry. An interesting relation can be read off from their Lagrangian. Due to the prefactors  $(\chi/\chi_0)^3$  in Eq. (4.71), the quark condensate is not only determined by the meson condensate  $\langle 0|U + U^+|0 \rangle$ , but knows about the gluon condensate as well. Such a relation reads, in the effective model of Kusaka and Weise (1992),

$$\langle \bar{u}u \rangle = \langle \bar{d}d \rangle = -\frac{1}{N_f G} \left( \frac{\chi_0}{\chi_m} \right)^2 \langle \sigma \rangle. \quad (4.75)$$

Here  $\chi_0 = \langle 0|\chi|0 \rangle$  denotes the mean-field vacuum expectation value as above,  $\chi_m$  minimizes the potential term of  $L_g$ , and  $G$  is the four-quark coupling.

When this relation is extrapolated to temperatures larger than zero, an interesting feature arises. Chiral symmetry may be restored ( $\langle \bar{u}u \rangle_{T \rightarrow 0} = 0$ ), either because the mesonic part  $\langle \sigma \rangle$  has melted or because the gluon condensate  $\chi_0$  vanishes. This leads to an inequality between the chiral ( $T_{\text{ch}}$ ) and the deconfinement ( $T_d$ ) transition temperatures

$$T_{\text{ch}} \leq T_d. \quad (4.76)$$

The interplay between both condensates can be studied this way. To question of “who drives whom” in the phase transition may be answered in these models.

The chiral phase transition has been investigated as a function of temperature and/or density in a variety of Nambu–Jona-Lasinio *models* (Hatsuda and Kunihiro, 1985, 1987; Bernard and Meissner, 1988; Kunihiro, 1989, 1991; Klimt *et al.*, 1990; Lutz *et al.*, 1992). In general the calculations were performed within a mean-field approach. The common conclusion was that under realistic conditions for coupling strengths (or quark masses) the chiral transition is replaced by a crossover phenomenon.

An extended NJL model has been derived by Bijlens *et al.* (1993) as a low-energy approximation to QCD. Starting from basic QCD, they integrated out gluon fields or alternatively gluon fields and quark fields within a certain approximation in the path-integral formulation. The resulting extended NJL model includes various low-energy models in appropriate limit cases, in particular, the chiral quark model of Manohar and Georgi (1984), which extends the linear sigma model of our previous sections.

Let us finally discuss the *chiral quark model* in a little more detail. Originally this model was proposed by Manohar and Georgi as an effective model for quarks, gluons, and Goldstone bosons to explain the success of nonrelativistic quark models. Gocksch (1991) has considered a version of the chiral quark model in which gluonic degrees of freedom are dropped and the Goldstone part is realized as the linear  $SU(2) \times SU(2)$  sigma model. The Lagrangian is given by

$$L = \frac{1}{2} [(\partial_\mu \sigma)^2 + (\partial_\mu \boldsymbol{\pi})^2] + \frac{\mu^2}{2} (\sigma^2 + \boldsymbol{\pi}^2) + \frac{\lambda}{4!} (\sigma^2 + \boldsymbol{\pi}^2)^2 + c\sigma + \bar{\Psi}[\not{\partial} + g(\sigma - i\gamma_5 \boldsymbol{\tau} \boldsymbol{\pi})]\Psi. \quad (4.77)$$

The notation is obvious from earlier notations in the text. Note that the Goldstone bosons ( $\boldsymbol{\pi}$ ) and the  $\sigma$  channel are treated on an equal footing with the quark degrees of freedom. This accounts for the prominent role the pions and the sigma meson play among the low-mass hadrons in the chiral transition. They cannot be described as states of weakly interacting quarks. (The well-known price of this treatment is double counting. For example, there is a pseudoscalar  $\bar{q}q$  bound state as well as the “fundamental” pion state.)

Gocksch performed mean-field calculations at finite temperature based on the Lagrangian (4.77). In this way he was able to reproduce lattice results for static had-

ronic screening lengths and the quark number susceptibility (Gocksch *et al.*, 1988; Gottlieb *et al.*, 1987c). There are different ways that chiral symmetry may be restored above the transition. One possibility is parity doubling of the known hadron spectrum (McLerran, 1986; DeTar and Kunihiro, 1989). Parity partners should be degenerate in their mass above the chiral transition, but still in the confinement phase.

The more standard symmetry restoration is realized through massless quarks. The results of Gocksch indicate that the lattice data for the  $\rho$  meson and nucleon screening masses can be understood in terms of nearby massless propagating quarks, while the  $\pi$  and  $\sigma$  modes cannot be explained this way. This favors the realization of chiral symmetry restoration in the “standard” way, i.e., via massless quarks with degenerate  $m_\pi$  and  $m_\sigma$  masses in the chiral symmetric phase.

The chiral quark model (4.77) gives a nice example of an effective model that allows explicit comparison with lattice results and is able to explain them. Such an agreement justifies the reduction of the ansatz for the effective Lagrangian. In this particular case it confirms the prominent role of the sigma and pions. It remains unclear whether double counting in the chiral quark model has any contaminating effect on the results.

Interesting directions for future work are combinations of “chiral” and “gluonic” effective models to study the interplay between both types of transitions. Dual Ginzburg-Landau models may be promising candidates for describing *both* aspects, confinement and chiral symmetry. For first attempts see, for example, Miyamura (1995). It should be verified that the  $U(1) \times U(1)$  Abelian monopoles survive the continuum limit and maintain their dominant role in the presence of dynamical quarks. An interesting question then is, how does the confinement mechanism in terms of Abelian monopole condensation translate to different conditions for gauge fixing?

Unlike QCD there is no unique favorite action on the effective level. Therefore one should study the variation of the results under different reductions of the underlying full QCD. This applies to generic effective models.

## V. RELATIVISTIC HEAVY-ION COLLISIONS

### A. Scales and observables

The experimental possibilities for testing the QCD predictions of finite-temperature phase transitions from normal hadronic matter to a quark-gluon plasma are limited if we face the transition temperature of  $\approx 10^{12}$  K. Most probably, the QCD transition occurred  $10^{-6}$  seconds after the big bang when the universe cooled down to the transition temperature. It is very questionable whether any remnants of the transition can be seen today. Promising alternatives to reproducing the QCD transition in some kind of little bang in laboratory experiments are heavy-ion collisions at ultrarelativistic energies. In such collisions a large amount of the initial

kinetic energy would be concentrated in a short space-time interval and form a fireball of matter which could reach thermal equilibrium.

In the last decade, it has become technically feasible to create matter with energy densities 10–100 times that of ordinary nuclear matter (see, for example, Jacob and Satz, 1982 or Schmidt and Schukraft, 1993). The available energy for particle production is specified in terms of  $\sqrt{s}$ , where  $\sqrt{s}$  is the total center-of-mass energy. In 1992 the AGS accelerator at Brookhaven National Laboratory (BNL) produced very heavy-ion beams of  $^{197}\text{Au}$  with momenta of  $11.4/c$  GeV per nucleon (Barrette, 1994). At the CERN SPS accelerator, 200 GeV/ $n$  are reached for beams up to  $^{32}\text{S}$ —still “light” heavy ions. Since November 1994, a Pb beam of 160 GeV/ $c$  per nucleon has been available at the SPS accelerator. Future experiments are planned starting in 2006, in which a value of  $\sqrt{s}=6300$  GeV/ $n$ - $n$  ( $n$ - $n$  stands for nucleon-nucleon) for lead should be reached at the CERN Linear Hadron Collider. At the Relativistic Heavy-Ion Collider of BNL  $\sqrt{s}=200$  GeV/ $n$ - $n$  will be reached in 1999.

Given the value of  $\sqrt{s}$  and the mass number  $A$  of colliding ions, one would like to estimate the initial temperatures that can be reached in such a collision. The basic observable is the multiplicity per unit rapidity  $dN/dY$  of secondary hadrons which are emitted in the collision. If one extrapolates the known relation between  $\sqrt{s}$  and  $(dN/dY)_p$  in proton-proton collisions to central nucleus ( $A$ )-nucleus ( $A$ ) collisions, the relation is given as (Satz, 1990a)

$$\left(\frac{dN}{dY}\right)_{AA} = 0.8A^\alpha \ln \sqrt{s} \quad (5.1)$$

with  $\alpha \geq 1.1$ . This leads to multiplicity densities of about 480 at midrapidity (including neutrons and neutral pions) at the Brookhaven AGS accelerator (Barrette *et al.*, 1995) for Au+Au, and to an estimate of 790 for the maximal  $dN/dY$  (when all hadrons are included) for Pb+Pb at the SPS accelerator at CERN (Margetis *et al.*, 1995). At CERN’s Linear Hadron Collider, for  $\sqrt{s}=6300$  GeV/ $n$ - $n$  multiplicity densities of 1500–2500 are expected according to Eq. (5.1). The multiplicity density of the final-state hadrons can be related to the initial-state energy density  $\varepsilon$  either with (approximate) energy conservation (free flow) or entropy conservation (see Sec. V.B.1). For free flow the relation is given by

$$\varepsilon = \{(dN/dY)_A m_T\} / (\pi R_A^2 \tau). \quad (5.2)$$

Here  $R_A$  is the nuclear radius,  $m_T$  the transverse mass, and  $\tau$  the equilibration time. With  $\tau \approx 1$  fm/ $c$  and  $m_T \approx 0.5$  GeV, Eq. (5.2) leads to an average initial energy density of 1.5–2.5 GeV/ $\text{fm}^3$  for central Pb-Pb collisions at the SPS accelerator and 4.6–7.8 GeV/ $\text{fm}^3$  at the Linear Hadron Collider. For Si+Pb-collisions at the AGS accelerator of BNL, the initial energy density has been estimated as 0.6 GeV/ $\text{fm}^3$  (Stachel and Young, 1992).

The energy density is translated to a temperature according to an ideal-gas relation  $\varepsilon \propto T^4$ , where the pro-

portionality constant depends on the number of included degrees of freedom. For three massless quark flavors, the initial temperature for Pb-Pb collisions at CERN’s SPS accelerator ranges from 170 to 190 MeV (free-flow assumption) or 160–190 MeV (isentropic expansion). If we compare these values to estimates from lattice calculations including dynamical fermions, AGS and SPS energies could produce temperatures slightly above  $T_c$ . Pronounced signatures of a plasma via thermal radiation can only be expected if the initial temperature  $T_i$  is well above  $T_c$ . More recent estimates of  $T_i$  by Geiger (1992a, 1992b), Kapusta *et al.* (1992), and Shuryak (1992) give more optimistic values for  $T_i$ , predicting that  $T_i \sim 500$ –600 MeV should be reached (see also below).

In applying these formulas, we have implicitly made use of the fact that energy, entropy, and temperature are well-defined notions in describing the collision. This assumption will be discussed later on.

At this point, let us pause for a comparison with the observation of a phase transition under daily life conditions, the boiling of water. Usually one does not appreciate all the well-defined experimental conditions. The fluid container has a fixed volume, the fluid is at rest and in thermal equilibrium. Calibrated thermometers are at hand, whose Hg column grows linearly in the considered temperature range. The phase transition is easily seen as conversion from the liquid to the vapor phase, while the temperature stays constant. Moreover, heating sources are available such that tuning of the temperature does not pose a problem. The temperature may be tuned adiabatically or quenched as one likes.

It is not surprising that at temperatures where hadronic matter gets dissolved into its components, the experimental conditions change drastically. In collider experiments the volume is not fixed in the evolution of the plasma to the hadron gas. At high collision energies the nuclei interpenetrate each other at the collision and recede as Lorentz-contracted pancakes, leaving a hot vacuum with secondaries of the collisions between them (in the central rapidity region). After hadronization this hot area is primarily a pion gas.

At Linear Hadron Collider energies the volume at the transition is estimated to be 5–8 times larger than the initial volume; for Pb-Pb collisions the critical volume is of the order of 800–1200  $\text{fm}^3$ . After the transition the system continues expanding until freezeout, where interactions can be neglected and/or the mean free path of particles becomes of the order of the size of the system. (Both definitions of “freeze-out” lead to different estimates for freezeout volumes). At the Linear Hadron Collider, freezeout volumes could be of the order of  $10^4$  to  $10^5$   $\text{fm}^3$  (Satz, 1990a).

Next we come to various time scales involved in heavy-ion collisions. Note that the total duration of a “little bang” is only of the order of  $10^{-23}$  sec. (The characteristic length scale of 1 fm is a very short distance for light to pass by.) Time scales are the equilibration time  $\tau_0$ , the freezeout time, the delay caused by a possible first-order phase transition, the conversion rate of one

phase into the other, and the expansion rate. All of them are between say 0.1 and 100 fm/c. The largest value refers to an extreme delay of freezeout due to strong supercooling, the smallest one corresponds to a more recent estimate for the equilibration time (Shuryak, 1992). The equilibration time is the time it takes until the system reaches a stage of local thermodynamic equilibrium, when it has passed an intermediate preequilibrium stage after the bang. “Local equilibrium” means that energy, entropy, pressure, and temperature can be locally defined. In contrast to our water-boiling experiment, the temperature is—after it is defined at all—a function of space and time  $T(\mathbf{x}, t)$  in a volume whose geometry is not even fixed, but depends on the impact-parameter value  $b$ .

Volumes of the order of thousands of fm<sup>3</sup> and time scales of some tens of fm/c can be measured with pion interferometry, especially a time delay due to a transition (see Sec. V.C.4). The lifetime of the fireball is estimated to be 10–20 fm/c for Si+Pb(Au) collisions at 14.6 GeV/n-n from pion interferometry measurements at the Brookhaven AGS accelerator (see, for example, Stachel, E814 Collaboration, 1994).

Temperature scales that should be distinguished are the initial temperature at the onset of local equilibrium  $T_i$ , the transition temperature  $T_c$ , possible values of superheating/supercooling effects, and the decoupling temperature  $T_d$  (of the order of the pion mass).

Candidates for thermometers are thermal photon and dilepton spectra (Shuryak, 1978; see also Gyulassy, 1984). The differential cross section of dilepton production as a function of the invariant lepton pair mass  $M$  and the rapidity  $Y$ ,  $(d^2\sigma/dM^2 dY)_{Y=0}$ , is predicted to scale according to  $\exp(-M/T)$ , where  $T$  is the temperature of the emitting system. Unfortunately these thermometers like to hide in the background of other dileptons. Thermal dileptons are difficult to identify in the total dileptonic yield (see Sec. V.C.2). Pions also have been proposed as thermometers for measuring the freezeout temperature (Brown, Stachel, and Welke, 1991). Pions from the decay of the  $\Delta(33)$  resonance have a characteristic  $p_T$  distribution that is very different from that of primary pions. The ratio of  $\Delta(33)$  resonances to nucleons sensitively depends on the temperature via Boltzmann factors. Thus a measurement of the  $p_T$  distribution permits the identification of decay pions and a measurement of the temperature-dependent ratio of  $\Delta(33)$  resonances to nucleons.

If it is difficult to measure the temperature, let us see how we can tune it. The initial temperature changes with the initial energy density ( $\varepsilon \propto T^4$  for an ideal gas), and that depends on  $\ln\sqrt{s}$ . Thus in principle one could vary  $s$  for a given nucleus  $A$ . Due to the  $\ln\sqrt{s}$  dependence this would require a large variation in the incident-beam energy, accompanied by a considerable loss in the luminosity, if  $\sqrt{s}$  is reduced (Ludlam and Samios, 1988; Satz, 1990a). For asymmetric collisions between small and large nuclei, the energy can be increased by going from peripheral to central collisions, i.e., by varying the impact parameter. For symmetric collisions the realistic

possibility that remains is the variation of  $A$ , i.e., the type of nucleus itself. Going from  $SS$  to  $UU$  collisions roughly gives a gain in energy density by a factor of 7–8 at the price that the volume changes as well. Going from fixed-target to collider experiments one expects a factor of 2–4 increase of the initial temperature  $T_i$ . The estimate of  $T_i$  at the AGS and SPS accelerator is about 190–200 MeV; at the Relativistic Heavy-ion Collider  $T_i$ 's will be reached between 400 and 500 MeV, at Linear Hadron Collider at CERN it may be 600–900 MeV.

Last but not least, to complete our comparison to a phase transition under “normal” conditions, we have to find observable signatures of the QCD transition. The goal is to identify observables that could reflect almost constant pressure or temperature over an interval where energy and entropy densities ( $\varepsilon$  and  $s$ ) change rapidly, where the phase conversion takes place. Typical observables at our disposal in heavy-ion collisions are multiplicity distributions in rapidity space and average transverse momenta  $\langle p_T \rangle$ . It turns out (see Sec. V.C.1) that  $\langle p_T \rangle$  values can be a measure of the initial pressure and temperature, while  $(dN/dY)$  distributions depend on the initial values of  $\varepsilon$  and  $s$ . Thus a  $T$ - $\varepsilon$  diagram corresponds to a  $\langle p_T \rangle$ - $dN/dY$  plot.

There are a variety of other signatures that are in principle sensitive to a transient plasma and to the transition dynamics as well. They are discussed in Secs. V.C and V.D.

Let us summarize the main complications we have to face in heavy-ion experiments. The basic assumptions that the fireball is large enough and long-lived enough to reach thermal equilibrium, and that thermodynamic concepts apply have to be checked and justified. Due to expansion dynamics, different competing time and length scales are involved. The spacetime expansion from initial thermalization until freezeout has to be traced back from the final observables, if one wants to compare the signatures with predictions of static, microscopic equilibrium quantities. Usually one traces back using a hydrodynamic description.

Hydrodynamics is a useful computational tool in estimating bulk features like leptonic or hadronic particle yields, multiplicity fluctuations, orders of lifetimes, etc. Note also that hydrodynamics describes off-equilibrium situations although it is based on local equilibrium conditions. The system is expanding and cooling and out of global equilibrium. In this sense, hydrodynamics is also a conceptual framework in which to treat nonequilibrium situations in such a way that all the information of equilibrium QCD is not lost. This information includes the equation of state, elementary cross sections, structure functions, and other derived quantities of equilibrium thermodynamics which we have partly outlined above.

There are a few other concepts that are useful in treating off-equilibrium aspects of heavy-ion collisions. We shall discuss the determination of the nucleation rate of hadronic bubbles in the plasma according to Langer's approach in condensed-matter physics (Langer, 1969; Csernai and Kapusta, 1992a, 1992b; Sec. V.D.1), an alternative process for phase conversion

(large domain coarsening) in Sec. V.D.2 (Borrill and Gleiser, 1995), a calculation of transport coefficients based on a combination of linear response theory with hydrodynamic concepts (Sec. V.D.3), and the concept of dynamical universality applied to systems far from (Rajagopal and Wilczek, 1993a, 1993b; Sec. V.D.4).

The effect of a phase transition on the hydrodynamical flow depends on the type of phase-transition dynamics. If the transition is of first order but proceeds smoothly, close to equilibrium, the effect is just to slow down the expansion. Discontinuities in thermodynamic quantities would be reflected in shocklike discontinuities of the fluid. If the transition is of first order but involves metastable states—a supercooled plasma or a superheated hadron gas—deflagrations or detonations may evolve with the possible effect of large multiplicity fluctuations (Sec. V.C.5).

Such explosive processes are one source of entropy production during the evolution. Other sources are dissipation effects and the freezeout transition. All of them are estimated to produce little extra entropy, so that the approximation of entropy conservation during the evolution seems to be justified (Blaizot and Ollitrault, 1990, and references therein).

Although dissipation effects on entropy production may be small, pure glue is rather viscous. It is amusing to estimate the viscosity of a gluon gas at tera degrees in comparison to “normal” gases. For  $T$  close to  $T_c$  the shear viscosity is of the order of the  $\Lambda$ -QCD scale. Thus it is  $10^{16}$  times the viscosity of a classical gas under “normal” conditions (where it is  $10^{-5}$  kg/m s; Hosoya and Kajantie, 1985).

The main ingredients in a hydrodynamic description are the initial conditions (Sec. V.B.1) and the equation of state (Sec. V.B.2). Before we go into detail, let us quote some numbers from Cleymans *et al.* (1986) to show that the basic condition for applying a hydrodynamic description is not violated too much: The mean free path of a particle in a medium has to be much smaller than the size  $L$  of the medium. The mean free path of a quark at an initial energy density of  $2.7 \text{ GeV/fm}^3$  for  $UU$  collisions is supposed to be  $0.22 \text{ fm}$ , as compared to a diameter of  $15 \text{ fm}$  for a uranium nucleus. The corresponding quark density is  $9 \text{ quarks/fm}^3$ . Groups presently working on three-dimensional hydrodynamics for relativistic heavy-ion collisions are those of Venugopalan *et al.* (1994), Waldhauser *et al.* (1992), and Bravina *et al.* (1993).

We close this section with a warning concerning the following subsections. The reader will not find any weighing of the evidence for the transition to be of either first or second order. We leave this question open until the end.

## B. The hydrodynamic framework

The derivation of the hydrodynamic equations can be found in various textbooks (e.g., Landau and Lifshitz, 1959), and their adaptation to heavy-ion collisions is described in the work of Cooper and Frye (1974) or in

reviews by Cleymans *et al.* (1986) or Blaizot and Ollitrault (1990). For completeness we sketch the main steps in deriving an appropriate form for heavy-ion collisions. The adaptation to heavy-ion collisions amounts to a suitable choice of coordinates and a set of initial conditions based on experimental observations.

Hydrodynamic equations describe the evolution of a gas (or fluid) in space and time. The gas is specified by a local temperature, pressure, energy, entropy, and velocity. The equations result from constraints of energy and momentum conservation and other conserved quantities like the baryon number. If we first neglect dissipative effects (viscosity, thermal conductivity), the energy-momentum tensor of a relativistic perfect fluid in motion with velocity  $u^\mu$  is obtained by a Lorentz boost from its rest frame as

$$T^{\mu\nu} = (\varepsilon + p)u^\mu u^\nu - g^{\mu\nu}p. \quad (5.3)$$

The equations for energy-momentum and baryon number conservation are

$$\partial_\mu T^{\mu\nu} = 0 \quad (5.4a)$$

and

$$\partial_\mu J^\mu = \partial_\mu (n_B u^\mu) = 0, \quad (5.4b)$$

respectively, where  $n_B(\mathbf{x}, t)$  is the local baryon number density. Using Eq. (5.3), contraction of Eq. (5.4a) with  $u_\nu$  leads to

$$u^\nu \partial_\nu \varepsilon + (\varepsilon + p) \partial_\nu u^\nu = 0, \quad (5.5)$$

where  $\varepsilon$  denotes the energy density and  $p$  the pressure. An analogous equation can be derived for the entropy density  $s$ , which can be converted to a temperature equation in the baryon free case. Contracting Eq. (5.4a) with  $(g_{\nu\rho} - u_\nu u_\rho)$  leads to the second hydrodynamic equation,

$$(\varepsilon + p)u^\tau \partial_\tau u_\lambda - \partial_\lambda p + u_\lambda u^\tau \partial_\tau p = 0. \quad (5.6)$$

The next step is to choose coordinates adapted to a plasma evolution in cylinder geometry, where the  $z$  axis is commonly identified with the beam axis. Now one can express Eqs. (5.5) and (5.6) in coordinates  $z$  and  $t$ ; the four-velocity  $u^\mu$  of the matter is written as

$$u^\mu = \frac{1}{\sqrt{1 - v_z^2 - v_r^2}} (1, v_z, v_r, 0). \quad (5.7)$$

For vanishing radial velocity  $v_r$  (which is frequently used as an approximation), the remaining components of  $u^\mu$  are parametrized according to

$$u^\mu = (\cosh \theta, \sinh \theta, 0, 0) \quad (5.8a)$$

where  $\theta$  is the fluid rapidity, defined via

$$\theta = \arctan v_z. \quad (5.8b)$$

A more convenient choice of variables are the spacetime rapidity  $\eta$  defined as

$$\eta = \frac{1}{2} \ln \frac{t+z}{t-z} \quad (5.9)$$

and the proper time  $\tau$

$$\tau = \sqrt{t^2 - z^2}$$

with the inverse transformations

$$\begin{aligned} t &= \tau \cosh \eta, \\ z &= \tau \sinh \eta. \end{aligned} \quad (5.10)$$

Note that the fluid rapidity coincides with the spacetime rapidity in the case of  $v_z = z/t$ .

In terms of these new coordinates, the hydrodynamic equations (5.5) and (5.6) for a longitudinal motion are

$$\begin{aligned} \tau \frac{\partial \varepsilon}{\partial \tau} + \tanh(\theta - \eta) \frac{\partial \varepsilon}{\partial \eta} + (\varepsilon + p) \left[ \frac{\partial \theta}{\partial \eta} + \tanh(\theta - \eta) \tau \frac{\partial \theta}{\partial \tau} \right] \\ = 0 \end{aligned} \quad (5.11a)$$

and

$$\begin{aligned} \frac{\partial p}{\partial \eta} + \tanh(\theta - \eta) \tau \frac{\partial p}{\partial \tau} + (\varepsilon + p) \left[ \tau \frac{\partial \theta}{\partial \tau} + \tanh(\theta - \eta) \frac{\partial \theta}{\partial \eta} \right] \\ = 0. \end{aligned} \quad (5.11b)$$

This set is complemented by a third equation, which follows from baryon number conservation (Kajantie *et al.*, 1983),

$$\tau \frac{\partial n_B}{\partial \tau} + n_B \frac{\partial \theta}{\partial \eta} + \tanh(\theta - \eta) \left( \frac{\partial n_B}{\partial \eta} + n_B \tau \frac{\partial \theta}{\partial \tau} \right) = 0. \quad (5.12)$$

Now we have three equations (5.11)–(5.12) for four unknown functions: the energy density  $\varepsilon$ , the pressure  $p$ , the fluid rapidity  $\theta$ , and the baryon number density  $n_B$ , all of them being functions of  $\eta$  and  $\tau$  in a longitudinal expansion. Thus, in order to find solutions of the set (5.11)–(5.12), we have to supply one additional equation and to specify the initial conditions. The additional equation is an *equation of state* relating  $\varepsilon$  and  $p$  or  $T$  and  $s$ . (Other combinations are possible as well.) As an example, we will discuss the bag model equation of state; see Sec. V.B.2. Several proposals have been made for the initial conditions. Here we sketch only the Bjorken-Shuryak expansion scenario. We discuss the longitudinal solutions of Eqs. (5.11)–(5.12) for this choice. The solutions considerably simplify in this special case, which may be one reason why they are frequently used in the hydrodynamic treatments of matter evolution in heavy-ion collisions.

Radial solutions of Eqs. (5.11) are rarefaction waves propagating from the boundary into the fluid with the velocity of sound. They differ in an essential way from Bjorken's scaling solution as they are independent of proper time  $\tau$ . Radial solutions enter measurements of enthalpy and pressure, see Sec. V.C.1.

### 1. Bjorken-Shuryak expansion scenario

In the Bjorken-Shuryak scenario (Shuryak, 1978; Bjorken, 1983), several experimental observations are taken into account. The phenomenon of "nuclear transparency" leads to a separate treatment of the central

rapidity and the fragmentation region. In nuclear transparency a large fraction of the incoming energy is carried away by two receding nucleons in a nucleon-nucleon collision at high energy. Similarly, in a nucleus-nucleus collision the baryon contents of the colliding nuclei interpenetrate at the collision and recede as two Lorentz-contracted pancakes after the collision. The central rapidity region refers to the fluid of quanta contained in the region between the receding pancakes. In the hadronic phase it consists mostly of pions. Thus it should be a good approximation to neglect the baryon number. Setting  $n_B = 0$  leads to a first simplification of the set of hydrodynamic equations. Separate treatment of the central and fragmentation regions is justified only if both regions are well separated in phase space. Experimental conditions should be checked to guarantee this, otherwise an analysis in this picture is not adequate (Blaizot and Ollitrault, 1990). At future colliders (the Relativistic Heavy-ion Collider and Linear Hadron Collider) these conditions may be satisfied.

Secondly, pronounced space-time correlations are observed in particle production in the sense that particles with large longitudinal momenta are produced at a late time, those with low momenta promptly, in the center-of-mass system. This is nothing but the twin paradox. Particles live longer when their velocities are higher. In Bjorken and Shuryak's ansatz the effect of time dilatation is incorporated in the boundary conditions. Consider an ensemble of particles that are produced at  $z=0=t$ . If it is only the proper time  $\tau$  that determines the moment of disintegration, all particles that measure the same  $\tau$  in their rest frame constitute an initial condition at  $\tau = \tau_0$ . That is, the initial condition refers to a hyperbola  $\sqrt{t^2 - z^2} = \tau_0$  of constant proper time  $\tau_0$ . The spacetime rapidity or light-cone variable  $\eta = 1/2 \ln t + z/t - z$  specifies the position on this hyperbola. Two distinct positions are related via a Lorentz boost in the  $z$  direction. Since particles in the fluid element are supposed to move as free particles, their velocity component  $v_z$  is given by  $z/t$ . The physics of a  $z$  slice of a fluid element at time  $t$  is equivalent to the physics of a  $z'$  slice at time  $t' = z'/v_z$ . This is the scaling property in Bjorken's scaling ansatz. For  $v_z = z/t$ , the spacetime rapidity  $\eta$  equals the rapidity  $Y$ ,

$$Y = \frac{1}{2} \ln \frac{1+v_z}{1-v_z} = \frac{1}{2} \ln \frac{E+p_z}{E-p_z}, \quad (5.13)$$

if the four-momentum  $p$  is parametrized as  $p = (E, p_z, p_t)$ , where  $p_t$  is the transverse momentum.

A third feature observed in proton-proton collisions is the plateau structure of inclusive cross sections when they are plotted as functions of  $Y$ . A plateau for central values of  $Y$  is also expected for nucleus-nucleus collisions. At least the particle multiplicity depends only weakly on  $Y$  for central rapidities (Bjorken, 1983). Accordingly, a further simplifying assumption seems to be justified. The local thermodynamic quantities like  $\varepsilon, p, T, s$  depend only on  $\tau_0$ , but not on  $\eta(\tau_0)$ , when the hydrodynamic expansion commences. Thus the initial

condition is invariant under Lorentz boosts in the  $z$  direction. The dynamics preserves Lorentz covariance, which is most easily seen from the tensor equation (5.4a). Therefore we shall look for solutions  $\varepsilon(\tau), s(\tau), T(\tau)$  of Eqs. (5.11) depending merely on  $\tau$ . Inserting  $v_z = z/t$  in the collective four-velocity of the fluid leads to

$$\begin{aligned} u^\mu &= \frac{1}{\sqrt{1-v_z^2-v_r^2}} (1, v_z, v_r, 0) = (t/\tau, z/\tau, 0, 0) \\ &= (\cosh \eta, \sinh \eta, 0, 0). \end{aligned} \quad (5.14)$$

A comparison with Eq. (5.8a) shows that the fluid rapidity can be identified with the space-time rapidity  $\eta$ , which furthermore coincides with the rapidity  $Y$ . With  $\eta = \theta$ , Eqs. (5.11) simplify to

$$\begin{aligned} \tau \frac{\partial \varepsilon}{\partial \tau} + \varepsilon + p &= 0 \\ \frac{\partial p}{\partial Y} &= 0. \end{aligned} \quad (5.15)$$

The baryon number is set to zero in the following considerations. The entropy equation for  $s(\tau)$  simplifies to

$$\tau \frac{\partial s}{\partial \tau} + s = 0, \quad (5.16)$$

where the relations  $u^\eta = 0$  and  $(u^\tau = u^t \cosh \eta - u^z \sinh \eta)$  have been used. The solution is

$$\frac{s}{s_0} = \frac{\tau_0}{\tau}. \quad (5.17)$$

As can be seen, upon integration over  $d^3x = \tau dY d^2x$ , the entropy per given rapidity interval remains constant as long as the hydrodynamic equations can be applied. This need not hold throughout all stages of the expansion, especially not close to freezeout or in the intermediate period, where the plasma converts to the hadronic phase in one or another way. Let us assume that it holds approximately. Then the important feature of Eq. (5.29) is that it allows us to infer the entropy density in the initial state (more precisely  $s_0 \tau_0$ ) from an observation in the final state (the pion multiplicity); see Sec. V.C.1. Under the same assumptions as above, the temperature equation simplifies in  $(\tau, \eta)$  coordinates to

$$c_s^2 + \tau \partial_\tau \ln T = 0 \quad (5.18)$$

for  $\mu = 0$ , where  $c_s$  denotes the velocity of sound and, by definition,

$$c_s^2 = \frac{\partial p}{\partial \varepsilon}. \quad (5.19)$$

Integration of Eq. (5.18) gives

$$T = T_0 (\tau_0 / \tau) c_s^2. \quad (5.20)$$

For a massless free gas, the speed of sound is  $1/\sqrt{3}$  in units where  $c = 1$ . Thus the temperature drops more slowly than the entropy density. In fact, the predicted decrease may be too slow, since transverse expansion has been neglected so far.

Finally, we have to solve Eq. (5.15) for the energy density as a function of proper time. One possibility is to use an equation of state (in principle, it should be *the* equation of state of QCD) to eliminate the pressure in Eq. (5.15). The result is that the energy density decreases not merely because of the expanding volume in proper time, but also due to the pressure exerted by the gas of the covolume. We come back to the equation of state in the next section. Note that the velocity in the transverse direction  $v_r$  has been neglected so far. We shall sometimes abbreviate the solutions in the Bjorken-Shuryak scenario as Bjorken's scaling solution.

## 2. The bag model equation of state

Although the bag model leads to a crude description of the equation of state for QCD, we devote a short subsection to it, since it is often used in combination with the hydrodynamic equations and leads to quantitative predictions in the end. In the MIT bag model, the basic features of QCD—confinement and asymptotic freedom—are effectively incorporated via bags (Chodos *et al.*, 1974; Johnson, 1975). In the hadronic phase, quarks and gluons are allowed to move freely or with perturbatively small interactions inside small volumes of space inside the bags. Outside the bags, the quarks are forbidden to move as free particles. The vacuum outside the bags is given a constant energy density  $B$  (the bag constant), which keeps the quarks and gluons confined to the bags. During the phase transition, latent heat is necessary to liberate the color degrees of freedom. It turns out to be proportional to  $B$ . Its original value ( $0.145 \text{ GeV}^4$ ; De Grand *et al.*, 1975) was based on fitting the mass spectrum at  $T = 0$  and low density in the MIT bag model. The effective value for the “vacuum pressure”  $B$ , which should be used in the quark gluon plasma phase at a baryonic matter density of  $n_B = 1/\text{fm}^3$ , is  $0.5 \text{ GeV}/\text{fm}^3$  (Shuryak, 1988).

An additive shift  $B$  in the energy density of the plasma due to the vacuum energy is obtained if  $\ln Z$  of an otherwise free gas of quarks and gluons is shifted by  $-BV/T$ , that is,

$$T \ln Z(\text{plasma phase}) = \text{free-gas contribution} - BV/T. \quad (5.21)$$

The free-gas contribution follows from the usual expression for a free gas of particles and antiparticles with mass  $m$ , chemical potential  $\mu$ , and degeneracy factor  $g$ . In the large-volume limit it is given by

$$\begin{aligned} \ln Z(T, \mu, V) &= \frac{gV}{6\pi^2 T} \int_0^\infty dK \frac{K^4}{(K^2 + m^2)^{1/2}} \\ &\times \left[ \frac{1}{\exp\{[(K^2 + m^2)^{1/2} - \mu]/T\} \pm 1} \right. \\ &\left. + \frac{1}{\exp\{[(K^2 + m^2)^{1/2} + \mu]/T\} \pm 1} \right], \end{aligned} \quad (5.22)$$

where the  $+$  sign refers to fermions and the  $-$  sign to

bosons. Adding up the various contributions from bosons (gluons), fermions (quarks and antiquarks), and the vacuum, one is led to

$$T \ln Z = \frac{1}{6} N_c N_f V \left( \frac{7}{30} \pi^2 T^4 + \mu_q^2 T^2 + \frac{1}{2\pi^2} \mu_q^4 \right) + \frac{\pi^2}{45} N_g V T^4 - B V. \quad (5.23)$$

Here  $N_c$  is the number of colors,  $N_f$  the number of flavors,  $N_g$  the number of gluons, and  $\mu_q$  the chemical potential due to quarks. The standard thermodynamic relations Eqs. (2.1) lead to the following expressions for the energy density, pressure, and entropy density in the plasma phase:

$$\varepsilon_p = 111aT^4 + B, \quad (5.24a)$$

$$p_p = 37aT^4 - B, \quad (5.24b)$$

$$s_p = 148aT^3, \quad (5.24c)$$

where  $a = \pi^2/90$ . These expressions hold for  $\mu = 0$ ,  $N_c = 3$ ,  $N_f = 2$ ,  $N_g = 8$ . The general expression for  $\varepsilon_p$  is

$$\varepsilon_p = \frac{N_c N_f}{\pi^2} \left( \frac{7\pi^4}{60} T^4 \right) + \frac{\pi^2}{15} N_g T^4 + B. \quad (5.25)$$

From Eqs. (5.24a) and (5.24b) we can easily read off the bag model equation of state in the plasma phase as

$$p = \frac{1}{3}(\varepsilon - 4B), \quad (5.26)$$

which remains valid for  $\mu \neq 0$ .

Similarly, expressions for  $\varepsilon$ ,  $p$ , and  $s$  are obtained in the hadron phase, when it is described as a free gas of the lightest mesons and baryons, i.e., pions, nucleons, and antinucleons, where the baryonic contribution is sometimes omitted. For pions analytic expressions for  $\varepsilon$ ,  $p$ , and  $s$  can be derived in terms of modified Bessel functions following from Eq. (5.22). Contributions of the nucleon-antinucleon gas can be calculated numerically. Heavier mass particles are often omitted for moderate temperatures ( $T \leq 250$  MeV), although the restriction to pions is rather questionable above  $T_c \geq 150$  MeV (see Sec. IV.A.2). Here we state the result for the limit of a gas of *massless* pions:

$$\varepsilon_h = 9aT^4, \quad (5.27a)$$

$$p_h = 3aT^4, \quad (5.27b)$$

$$s_h = 12aT^3, \quad (5.27c)$$

so that

$$\varepsilon_h = 3p_h \quad (5.28)$$

is the bag-model equation of state in the hadron phase. The finite  $T$  (and finite  $\mu$ ) transition occurs, when the following Gibbs criteria are satisfied:

$$p_h = p_p = p_c, \quad (5.29a)$$

$$T_h = T_p = T_c, \quad (5.29b)$$

$$\mu = 3\mu_q = \mu_c. \quad (5.29c)$$

The indices  $h, p$  stand for the hadron and plasma phases,  $c$  for the critical value, and  $\mu$  is the chemical potential associated with nucleons. Pressure balance at  $T_c$  relates  $T_c$  to the bag constant

$$T_c = (B/34a)^{1/4}. \quad (5.30)$$

The latent heat, determined as the gap in the energy densities  $\varepsilon_p - \varepsilon_h$  at the transition, is  $4B$  in this model, e.g., for an effective bag constant of  $0.5 \text{ GeV/fm}^3$  the latent heat is  $2 \text{ GeV/fm}^3$ .

Almost by construction the bag model leads to a first-order transition at finite  $T$  and vanishing  $\mu$ . There is a finite gap in energy and entropy densities, while the pressure is continuous. Note that in the mixed phase, the velocity of sound  $c_s$  vanishes, since  $c_s = dp/d\varepsilon|_{T_c} = 0$ .

This is true only so long as  $\mu = 0$ . An exception is a second-order transition at  $\mu_c \neq 0$ , but  $T_c = 0$  for a particular value of  $B$ , which we shall not consider further here (see Cleymans *et al.*, 1986).

Equations (5.26) and (5.28) can easily be combined to a single equation by using  $\theta$  functions as projections on the distinct phases above and below  $T_c$ . Similarly it is not difficult to formulate an equation of state for a second-order transition by smoothly interpolating the steplike behavior of the bag model equation of state. Such an ansatz has been proposed by Blaizot and Ollitrault (1990) for  $s(T)$ . Although the interpolation is *ad hoc*, it provides a useful check of how sensitively phenomenological implications depend on the order of the transition. The difficulty is to *derive* such an equation for a second-order transition within an effective model. The  $O(4)$  model in three dimensions allows a second-order transition, but is supposed to describe only the low-temperature phase of QCD ( $T \leq T_c$ ) or the immediate vicinity above  $T_c$ , where pions and sigma mesons have not yet dissolved into their constituents.

The bag model does not provide an adequate description of the transition region. Even at the transition point, the plasma and the hadron phases are treated as noninteracting gases differing only in the degrees of freedom and the vacuum energy. An increasing number of hints warns against a naive counting of the number of modes, treating them the same as in the limiting cases of high and low temperatures. In particular, the number of degrees of freedom of a hadron gas is a delicate problem, if the change of hadron masses as a function of temperature and density is respected. In this case a counting of pion and nucleon degrees of freedom is certainly insufficient. In the vicinity of  $T_c$  the confinement/deconfinement properties should be implemented in a slightly more sophisticated way than with a single parameter  $B$ . Nevertheless the bag model is frequently used for temperatures  $T \sim T_c$ . The reason is probably its very tractable analytic form compared to (preliminary) plots of numerical simulations.

At high or low temperatures the bag model is more adequate; the gluonic sector may be treated as a gas of noninteracting glueballs at low  $T$  and of gluons at high  $T$ .

Various improvements on QCD's equation of state have been proposed within the framework of the bag model. Finite quark masses and perturbative QCD corrections to the partons inside the bag can be taken into account (Kapusta, 1979; Shuryak, 1979), leading to corrections of  $\epsilon$ ,  $p$ , and  $s$ . Strange quarks, finite masses for pions, and higher-mass hadrons can also be included.

The ultimate goal is an equation of state from lattice QCD with dynamical quarks, which is merely based on nonperturbative ingredients (see Sec. III.B). The numerical data should then be presented in a feasible parametric form to facilitate their handling.

In the next section we shall see what we can learn directly from heavy-ion experiments about QCD's equation of state.

### C. Signatures sensitive to the nature of the phase transition

Here we described theoretical and experimental tools that are sensitive to the nature of the transition, in particular to its order. The sensitivity holds at least in principle. A dependence on the transition dynamics is sometimes hidden in the spacetime expansion. The signatures reveal characteristic features like time delays, nucleation of bubbles, strong correlations, or large fluctuations in an indirect way.

Let us start with direct experimental tests of the equation of state. We sketch the possibilities for measuring thermodynamic quantities.

#### 1. Thermodynamic observables

The basic observables at our disposal in heavy-ion collisions are the rapidity distributions of final-state particles and their transverse momentum distributions. An extraction of the equation of state requires measurable observables that are related to  $\epsilon(T, \mu)$ ,  $s(T, \mu)$ , or  $p(T, \mu)$ . The signals of a first-order transition in a finite volume may be qualitatively very similar to those for a sharp crossover phenomenon, consisting in a rapid rise in the effective number of degrees of freedom over a small range of temperatures, say less than 10 MeV. These numbers are revealed in  $s/T^3$  or  $\epsilon/T^4$ ; see, for example, Eqs. (5.24) of the bag model equation of state. Thus we have to identify the observables that are related to  $s$ ,  $\epsilon$ ,  $p$ , and  $T$ . Roughly speaking, temperature and pressure are measured by the average transverse momentum  $\langle p_T \rangle$  (as usual under certain restrictive conditions), energy and entropy by the particle multiplicity distribution in rapidity space.

We consider the baryon-free case,  $\mu=0$ . Four quantities  $s$ ,  $\epsilon$ ,  $p$ , and  $T$  have to be determined. An equation of state is a relation between any two of these four variables [ $\epsilon(p)$ ,  $s(T)$ ,  $s(\epsilon)$ ]. It can be obtained from experiments (and compared with theoretical predictions) if two relations are used as experimental input. For example, if  $\epsilon(T)$  and  $p(T)$  are known,  $s(T)$  follows from  $dp/dT=s$ , and  $\epsilon(p)$  is the equation of state. Alternatively, if the initial condition  $s_0(\tau_0)$  is fixed from a measurement of  $dN/dY$ , then  $s(\tau)$  and  $T(\tau)$  and thus

$s(T)$  are known from the hydrodynamic equations, and  $p(T)$  follows as an integral over  $s$ . Suppose that the second relation is provided by a measurement of  $\langle p_T \rangle$  as a function of  $s^{1/3}$ , where  $\langle p_T \rangle$  is related to  $s/\epsilon$  (see below) and  $s^{1/3} \propto T$ . The knowledge of  $(s/\epsilon)(T)$  yields  $\epsilon(T)$ ; thus all relations are known (Blaizot and Ollitrault, 1990).

In the following subsections we explain the relations between  $dN/dY$  and  $s_0$  or  $\epsilon_0$ ,  $\langle p_T \rangle$  and  $\epsilon/s$  or  $T$ , as well as  $\langle p_T \rangle$  and  $p$  or  $(\epsilon+p)$ . We will show why  $\langle p_T \rangle$  vs  $\langle dN/dY \rangle$  diagrams are roughly equivalent to  $(T$  vs  $s)$  or  $(T$  vs  $\epsilon)$  diagrams. In the case of a phase transition a flattening of the  $\langle p_T \rangle$  distribution is expected, which was originally proposed by van Hove (1982) as a possible signature for a phase transition.

#### a. Entropy measurements

The basic step is to identify the entropy density with the particle density in the final state. The final-state particles consist mostly of pions. Pions are nearly massless bosons, and thus their entropy  $S$  is approximately proportional to their number  $N$ . For the densities we have

$$s = \alpha n \quad (5.31)$$

with  $\alpha=3.6$  for a free gas of massless pions. The rapidity distribution is obtained from the particle density by an appropriate integration. When the four-volume element is expressed in terms of the spacetime rapidity  $\eta$  and the proper time  $\tau$ , it follows for  $N$

$$N = \int n d^4x = \int n \tau d\tau d\eta d^2x, \quad (5.32)$$

or for the number of particles per unit rapidity at a fixed final proper time  $\tau_f$

$$\frac{dN}{d\eta} = \int n(\tau_f, \eta, x) \cdot \tau_f d^2x. \quad (5.33)$$

From Eqs. (5.31) and (5.17) we have

$$\begin{aligned} \frac{dN}{d\eta} &= \int d^2x \tau_f \frac{1}{\alpha} s(\tau_f, \eta, x) \\ &= \int d^2x \tau_0 \frac{1}{\alpha} s_0(\tau_0, \eta, x). \end{aligned} \quad (5.34)$$

If the initial entropy density is taken to be independent of the transverse coordinates  $x$ , we find that the final rapidity distribution of multiplicity is proportional to the initial entropy density

$$\frac{dN}{d\eta} = \tau \frac{1}{\alpha} \cdot s_0(\tau_0, \eta) A_T, \quad (5.35)$$

where  $A_T = \pi R^2$  is the transverse size of the nucleus. For example, the multiplicity per unit rapidity may be about 150 for a central collision of  $^{16}\text{O}$  with  $R \sim 3$  fm on a heavy nucleus [Blaizot and Ollitrault (1990)]. If the initial (=equilibration) time is estimated as 1 fm/c, it implies a value of  $\sim 20 \text{ fm}^{-3}$  for the initial entropy den-

sity. We shall see next what is further needed to extrapolate the initial temperature and the initial energy density.

### b. Energy measurements

To derive an expression similar to Eq. (5.35) for the initial energy density  $\epsilon_0$ , we need a relation between  $s_0$  and  $\epsilon_0$  and  $\epsilon_0$  and  $T_0$  to eliminate  $T_0$ . From the thermodynamic relation ( $\epsilon + p = Ts$ ) and the bag model equation of state we find

$$(\epsilon_0 - B)(1 + c_s^2) = T_0 s_0. \quad (5.36)$$

In Eq. (5.36) the factor  $1/3$  for a massless free gas has been replaced by  $c_s^2$ . Next we use Eq. (5.24a) of the bag model in the plasma phase (assuming that the initial temperature is high enough for the system to be in the plasma phase),

$$\epsilon_0 - B = g T_0^4. \quad (5.37)$$

Again the exponent can be generalized to  $1 + 1/c_s^2$ . The degeneracy factor  $g$  in the case of two massless  $u$  and  $d$  quarks, their antiquarks and gluons is given by

$$g = \frac{\pi^2}{30} \left( 2 \times 2 \times 2 \times 3 \times \frac{7}{8} + 2 \times 8 \right) = 111a. \quad (5.38)$$

The different factors correspond to the spin, quark, antiquark, color, flavor, and gluon degrees of freedom;  $7/8$  arises from the Fermi-Dirac statistics. When  $T_0$  is eliminated in Eq. (5.37) via (5.36), and (5.35) is solved for  $s_0$ , we obtain

$$\epsilon_0 = B + g \left[ \frac{\alpha}{A_T (1 + c_s^2) g \tau_0} \frac{dN}{d\eta} \right]^{1 + c_s^2}. \quad (5.39)$$

Equation (5.39) gives the promised relation between the final-state rapidity distribution and the initial energy density  $\epsilon_0$ .

### c. Transverse momentum distributions

A second important class of observables comprises average transverse momentum distributions. First we consider the relation between the average transverse momentum  $\langle p_T \rangle$  and  $E/S$ . For a thermalized fluid at rest, the momentum distribution is isotropic. For an ultrarelativistic fluid the total momentum is equal to the energy, so that the average transverse momentum is proportional to the energy per particle [Blaizot and Ollitrault (1990)],

$$\langle p_T \rangle = \frac{\pi}{4} E/N. \quad (5.40)$$

If we replace  $E$  and  $N$  by the corresponding densities, it remains to express the particle density  $n$  in terms of the entropy density via Eq. (5.31),

$$\langle p_T \rangle = \frac{\alpha \pi}{4} \frac{\epsilon}{s}. \quad (5.41)$$

This is already a relation of the type we are looking for. Note that  $\epsilon/s$  scales approximately with  $T$  and  $s$  with

$T^3$ . Measuring  $\langle p_T \rangle$  as a function of  $s_0^{1/3}$  by varying  $dN/dY$  accordingly could in principle provide a check of a theoretically predicted relation between  $\epsilon/s$  and  $T$ .

When  $\epsilon/s$  is calculated from the bag model equation of state and plotted against  $(s_0/s_c)^{1/3}$ , it displays a feature that turns out to be characteristic of generic  $\langle p_T \rangle$  distributions: a flattening in the region where the initial entropy density  $s_0$  is chosen to have the critical value  $s_c$ , accompanied by a linear rise is for small ratios  $(s_0/s_c)^{1/3}$  (low temperatures) where  $\epsilon/s$  and  $s^{1/3}$  are proportional to each other.

Unfortunately Eq. (5.41) is too simple to be true. As the QCD plasma fluid undergoes a transverse and a longitudinal expansion, the simple relation between  $\langle p_T \rangle$  and the initial temperature via  $s_0$  gets lost. The decoupling temperature  $T_d$  at freezeout and the time at which the transverse expansion sets in enter as new scales. The essential point is whether decoupling and hadronization happen before or after the transverse expansion becomes important. The results of Blaizot and Ollitrault (1990) are the following. If the system decouples before the transverse expansion sets in,  $\langle p_T \rangle$  is not sensitive to the equation of state, but characterizes a fluid at rest at the decoupling temperature. In the opposite case (decoupling after transverse expansion),  $\langle p_T \rangle$  is roughly that of a thermal distribution with a three-dimensional expansion at decoupling. As an effect of longitudinal expansion,  $s_0$  has to be replaced in the naive expression (5.41) in an appropriate way. The longitudinal expansion diminishes the available energy for transverse expansion.

The value of  $\langle p_T \rangle$  is no longer proportional to  $\epsilon/s$ . The modified  $\langle p_T \rangle$  diagram as a function of the initial entropy density can be compared with the diagram resulting from a bag model equation of state combined with a hydrodynamic expansion. The qualitative feature of each is the same: a flattening of  $\langle p_T \rangle/T_c$  for initial entropy values lying in the transition region. The physical reason for this similarity is reduced pressure on individual particles if the entropy content is distributed among more degrees of freedom in a transient plasma. A smaller value for the pressure leads to a reduced average transverse momentum.

### d. Enthalpy and pressure

The average transverse momentum  $\langle p_T \rangle$  can be predicted from solutions of the hydrodynamic equations. The transverse velocity component is parametrized in terms of rapidity variables. The average transverse momentum is written in terms of the average transverse rapidity  $Y_T$  (due to thermal motion) and the fluid rapidity  $\theta$  according to

$$\langle p_T \rangle = m_\pi \sinh(Y_T + \theta). \quad (5.42)$$

Enthalpy measurements can be performed via  $\langle p_T \rangle$ , if the fluid rapidity  $\theta$  is determined as function of the initial enthalpy according to (Cleymans *et al.*, 1986)

$$\theta_R = \frac{c_s}{1 + c_s^2} \ln \frac{\epsilon_0 + p_0}{\epsilon + p}, \quad (5.43)$$

where  $\theta_R$  is a solution of the set of longitudinal hydrodynamic equations, specialized to rarefaction waves  $\theta_R$ . Solving Eq. (5.42) for  $\theta$  with  $\theta$  given by Eq. (5.43), we find for large  $p_T/m_\pi$  that  $\langle p_T \rangle$  is proportional to  $e^{\theta_R}$ , thus

$$\langle p_T \rangle \sim \frac{m_\pi}{2} \left( \frac{\epsilon_0 + p_0}{\epsilon + p} \right)^{c_s/(1+c_s^2)}. \quad (5.44)$$

Now we must express the right-hand side of Eq. (5.44) in terms of measurable quantities. We write the initial enthalpy as  $T \cdot s_0$ , eliminate  $T_0$  in favor of  $\epsilon_0$  with an equation of state, and replace  $s_0$  and  $\epsilon_0$  via Eqs. (5.35) and (5.39).

*Barometers* are provided by  $\langle p_T \rangle$  distributions as a function of the pressure. The fluid rapidity is now considered as a function of the final energy density  $\epsilon_f$ . The analogous relation to Eq. (5.43) reads (Gyulassy, 1984)

$$\theta(\epsilon_f) = \frac{c_s}{1+c_s^2} \ln \left( \frac{p_0}{p_f} \right), \quad (5.45)$$

leading to

$$\langle p_T \rangle = \frac{m_\pi}{2} e^{Y_T} \left( \frac{p_0}{p_f} \right)^{c_s/(1+c_s^2)}. \quad (5.46)$$

Equation (5.45) is derived from a set of hydrodynamic equations in which the longitudinal expansion has been neglected for simplicity. Equation (5.46) explains why  $\langle p_T \rangle$  distributions are called barometers. If the speed of sound and the pressure  $p_f$  in the final state are known, a measurement of  $\langle p_T \rangle$  allows one to estimate the initial pressure  $p_0$ . Equation (5.46) gives  $\langle p_T \rangle$  for a hadron gas, which is described as a massless ideal pion gas with rarefaction waves in the transverse direction, but without longitudinal expansion.

The transverse flow is overestimated as long as the longitudinal expansion is neglected. If  $p_0/p_f$  or  $\epsilon_0/\epsilon_f$  in Eq. (5.46) are replaced by the pion rapidity density  $dN/dY$ , the resulting  $\langle p_T \rangle$  values are actually too large as compared to experimentally measured values. The fluid feels transverse pressure gradients when the rarefaction wave arrives at a fluid element. The arrival is somewhat delayed, since the wave moves inwards with the speed of sound. Thus some energy has already gone into longitudinal expansion, and less energy is available for transverse motion. Formally this effect can be included by an appropriate form factor, which should be multiplied on the right-hand sides of Eqs. (5.46) and (5.41) [see Gyulassy (1984), Baym (1984a, 1994b), Blaizot and Ollitrault (1990) for further details]. The terminology of calling  $\langle p_T \rangle$  a “barometer” for heavy-ion collisions goes back to Gyulassy (1984). Gyulassy further identifies thermometers and seismometers in connection with dilepton production and large multiplicity fluctuations; see Secs. V.C.2 and V.C.5 below.

To summarize, bulk quantities like  $\epsilon(T)$ ,  $p(T)$ ,  $s(T)$ , derived quantities like the velocity of sound, the enthalpy, the energy per degree of freedom  $\epsilon/s$  (Redlich and Satz, 1986), and QCD’s equation of state are calculable on the lattice. The equation of state including dy-

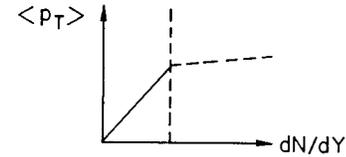


FIG. 29. Hypothetical correlation between the average transverse momentum  $\langle p_T \rangle$  and the multiplicity of final-state particles in a given rapidity interval  $dN/dY$ . The vertical line separates the areas of hadronic phase (left) and the transition region (right).

namical fermions and a finite chemical potential should finally replace the bag model equation of state, which entered the derivation of various relations in the preceding section. Equations (5.35) and (5.44) are a remarkable example of relations between observables that are directly accessible in experiments and on the lattice. Initial energy and entropy are related to the pion multiplicity per unit rapidity. A variation of  $dN/dY$  amounts to a variation of  $s_0$  or  $\epsilon_0$ . On the other hand,  $\langle p_T \rangle$  can be a measure of the freezeout temperature and is sensitive to the pressure. Thus a flattening of  $\langle p_T \rangle$  in a  $\langle p_T \rangle$  vs  $dN/dY$  diagram could be nothing else but a reflection of a slow change of temperature and pressure during a rapid rise in energy and entropy densities, i.e., a rapid crossover or a first-order transition in a finite volume. The typical shape of a  $\langle p_T \rangle$  vs  $dN/dY$  diagram, which is schematically drawn in Fig. 29 has actually been seen at the Tevatron (Alexopoulos *et al.*, 1990). Figure 30, shows the average transverse momentum  $\langle p_T \rangle$  of  $K^\pm$  and  $\pi^\pm$  in  $p\bar{p}$  collisions as a function of the charged-particle multiplicity per unit rapidity. Unfortunately a sharp crossover or a phase transition are not the only explanations for the shape. The flattening could be just a kinematic effect (lack of available energy). Kinematic constraints like energy conservation require that  $\langle p_T \rangle$  go to zero at the boundaries of the allowed rapidity interval (Schmidt and Schukraft, 1993). Rescattering effects or minijet production are further alternatives to explain an increase followed by a flattening effect (Eskola and

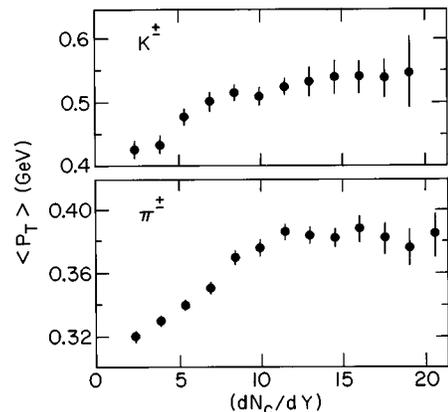


FIG. 30. The average transverse momentum  $\langle p_T \rangle$  of  $K^\pm$  and  $\pi^\pm$  in  $p\bar{p}$  collisions as a function of the charged-particle multiplicity per unit rapidity. From Satz (1990a).

Lindfors, 1990; Satz, 1990a, 1990b). The multiplicity increases only slowly with the incident kinetic energy. The relevant variable has been identified as  $s_0 t_0 / s_c R_0$ , where  $t_0$  is the proper time, at which the hydrodynamical evolution starts,  $s_0$  is the entropy density at that time,  $R_0$  is the initial radius, and  $s_c$  the entropy density at the critical temperature (Blaziot and Ollitrault, 1990). This variable should be larger than 1 to see an effect of the transverse expansion. Thus it is not sufficient that the initial entropy density  $s_0$  be larger than the critical value  $s_c$ , which one might naively expect. The inverse initial radius  $1/R_0$  is proportional to  $1/B$ , where  $B$  is the mass number of the smaller of the colliding nuclei. Therefore the use of heavier nuclei in the collision may not help, but make things worse. For a review about the available data on  $\langle p_T \rangle$  distributions we refer the reader to Schmidt and Schukraft (1993).

## 2. Dileptons and real photons

### a. Dileptons

We assume that the initial energy densities of a collision between nuclei  $A$  of the projectile and nuclei  $B$  of the target are high enough so that a quark-gluon plasma is initially generated. Since the size of the system produced in such a collision is not too large and the electromagnetic cross section is small, all dileptons can escape, especially those from the hottest dense state of the plasma. Dileptons will be emitted at all stages of the evolution, from the initial plasma phase through an intermediate mixture of plasma and hadron phases to the hadron phase at freezeout. Thus, in contrast to hadronic yields, dileptons are sensitive to the whole spacetime history of the evolution. The relevant kinematic variables for dileptonic cross sections are the invariant mass  $M$  of the dilepton pair and its transverse momentum  $p_T$ . A typical observable is the differential multiplicity  $R$ . This is the multiplicity of dileptons per invariant mass squared  $M^2$ , transverse momentum  $p_T$  and unit rapidity interval  $\Delta Y$ ,

$$R = \frac{dN}{dM^2 d^2 p_T dY} = \int_V d^4 x \frac{dN}{d^4 x d^4 p}. \quad (5.47)$$

Here  $R$  is a spacetime integral over the multiplicity rate  $dN/d^4 x d^4 p$ , which is the production rate at a given four-momentum  $p$  of the dilepton pair at a spacetime point  $(\mathbf{x}, t)$ . The quantity  $dN/d^4 x d^4 p$  depends on the temperature  $T$  via  $(\mathbf{x}, t)$ , and therefore on the phase of the system, since the temperature is a function of the spacetime coordinate of the evolution. The phase (plasma or hadron) determines the dominant production mechanisms contributing to the elementary cross sections and the structure functions in  $dN/d^4 x d^4 p$ . The structure functions reflect the medium in which the elementary processes take place (quarks inside a pion or in a heat bath of other partons). Collective effects of the plasma may also influence the production rate for a given value of  $(\mathbf{x}, t)$ .

In a given phase, several sources are usually responsible for dilepton production. We are mostly interested

in thermal collisions reflecting thermodynamic properties of the supposed heat bath environment. Thermal collisions are certainly not the only way to produce dileptons. Other mechanisms are Drell-Yan production, preequilibrium production in the plasma phase, and dilepton generation from pions via  $\rho$ -resonance decays in the hadronic phase. Thermal dileptons result from parton collisions in a medium which can be characterized by a local temperature. They are specific for a plasma or a hadron gas that can be described with thermodynamic concepts; they are absent from hadron-hadron collisions.

Finally, we want to find out the mass range  $\Delta M$  in which the differential multiplicity  $R$  is dominated by thermal dileptons. In this range  $R$  should be sensitive to the kinetics of the phase transition and in particular to its order. Let us assume that  $R(\mathbf{x}, t)$  has been calculated as a function of  $T(\mathbf{x}, t)$ . The spacetime integral in Eq. (5.47) can be performed if the missing relation  $T = T(\mathbf{x}, t)$  is known. Such a relation is provided by the hydrodynamic approach. So far it is the only framework which is detailed enough to predict particle spectra according to Eq. (5.47). Further simplifying assumptions are made. These include Bjorken's scaling ansatz and the neglect of transverse flow, which lead to the conservation of entropy. As we have seen above, the hadronic multiplicity of the final state can be related to the initial conditions under the assumption of entropy conservation. Once the entropy density  $\Sigma$  is known as function of time, the time dependence of  $T$  follows immediately if  $s(T)$  is known. This is where the equation of state enters. Further assumptions about the kinetics of the phase transition are necessary to justify the application of hydrodynamic concepts throughout all stages of the evolution.

We will now outline the predictions of Cleymans *et al.* (1987) for thermal dilepton spectra, which are derived for different scenarios of the phase transition. The reason we have selected this reference event though it is not the most recent one in this field is that it explicitly addresses the influence of the order of the phase transition on the dilepton production rate. Although the ingredients do not represent the latest state of the art, the representation is suited for illustrating which features of dileptonic yields are sensitive at all to the dynamics of the phase transition. Later we comment on areas of possible improvements and extensions.

Consider a particle production rate that is differential in  $x$  and  $p$ . The general expression for the thermal rate of lepton pairs (here written for  $\mu^+ \mu^-$  rather than  $e^+ e^-$ ) in the independent-particle approximation of kinetic theory (Kapusta and Gale, 1987) is given by

$$\begin{aligned} \frac{dN}{d^4 x d^4 p} &= \int \frac{d^3 q_1}{2E_1 (2\pi)^3} f(q_1) \int \frac{d^3 q_2}{2E_2 (2\pi)^3} f(q_2) \\ &\times \int \frac{d^3 q_+}{2E_+ (2\pi)^3} \int \frac{d^3 q_-}{2E_- (2\pi)^3} \\ &\times |M(p_1 \bar{p}_2 \rightarrow \mu^+ \mu^-)|^2 \delta(p - q_1 - q_2). \end{aligned} \quad (5.48)$$

Here  $M$  is the matrix element for the process of lepto-production from particles  $p_1$  and  $\bar{p}_2$  with momenta  $q_1$  and  $q_2$ ,  $p$  is the momentum of the lepton pair and  $q_+$  and  $q_-$  are the momenta of  $\mu^+$  and  $\mu^-$ , respectively. The statistical distributions  $f(q_1)$  and  $f(q_2)$  measure the probability of finding particles  $p_1$  and  $\bar{p}_2$  with momenta  $q_1$  and  $q_2$  in the given medium. The dileptons do not receive such factors, since we assume that rescattering of the electromagnetically interacting particles can be neglected. The cross section is small, and the plasma volume is assumed not to be too large. In terms of cross section  $\sigma$ , Eq. (5.48) reads

$$\frac{dN}{d^4x d^4p} = \int \frac{d^3q_1}{(2\pi)^3} \frac{d^3q_2}{(2\pi)^3} v_{q_1q_2} \sigma(p_1 \bar{p}_2 \rightarrow \mu^+ \mu^-) \times f_1(q_1) f_2(q_2) \delta(p - q_1 - q_2) \quad (5.49)$$

upon integration over the lepton momenta;  $v_{q_1q_2}$  denotes the relative velocity of particles  $p_1$  and  $\bar{p}_2$ :

$$v_{q_1q_2} = \frac{\sqrt{(q_1q_2)^2 - m_{p_1}^4}}{E_1 E_2}. \quad (5.50)$$

In the plasma phase the lowest order process for dilepton production is the same as in Drell-Yan production: two quarks annihilate via a virtual photon to yield a lepton pair

$$q\bar{q} \rightarrow \gamma^* \rightarrow \mu^+ \mu^-. \quad (5.51)$$

The momentum distribution functions  $f_1$  and  $f_2$  of the quarks  $q$  and antiquarks  $\bar{q}$  are given by the Fermi-Dirac distributions:

$$f_q = \frac{6}{e^{(uq-\mu)/T} + 1}, \quad f_{\bar{q}} = \frac{6}{e^{(u\bar{q}+\mu)/T} + 1}. \quad (5.52)$$

The chemical potential  $\mu$  is set to zero in the end,  $u$  is the local four-velocity of the plasma fluid element in the fixed laboratory frame, and  $q(\bar{q})$  is the four-momentum of the quark (antiquark). For the hydrodynamic scenario, a longitudinal expansion will be used with vanishing transverse velocity of the plasma. With this in mind, it is convenient to express the product  $u \cdot q$  in terms of the transverse invariant mass  $M$ , the rapidity  $Y$  of the lepton pair, and the plasma spacetime rapidity  $\theta$ :

$$u \cdot q = M_T \cosh(\theta - Y), \quad (5.53a)$$

where  $M_T$  is the transverse mass, and

$$M_T = \sqrt{M^2 + p_T^2}, \quad (5.53b)$$

$$Y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \quad (5.53c)$$

$$\theta = \operatorname{arctanh} v, \quad (5.53d)$$

with  $v = z/t$  denoting the collective fluid velocity. If the transverse mass is large compared to the temperature, the exponentials are sharply peaked around  $\theta = Y$ . In this case no final  $\theta$  dependence is left. Upon integration over the dilepton momentum, the result for the *dilepton rate in the plasma phase* at a spacetime point

$\underline{x} = (\tau, x_T, y)$  ( $y$  being the local spacetime rapidity of a fluid element) and four-momentum  $\underline{p} = (M_T \cosh Y, p_T, M_T \sinh Y)$  is given by

$$\frac{dN_P}{d^4x d^4p} = \frac{\alpha^2}{4\pi^4} \left[ 1 + \frac{2m^2}{M^2} \right] \times \left[ 1 - \frac{4m^2}{M^2} \right]^{1/2} e^{-E/T} K_P(\underline{p}, T, \mu) \sum_i e_i^2. \quad (5.54)$$

The index  $P$  stands for plasma phase, the quantities  $e_i$  are the charges of the quarks,  $\alpha$  is the electromagnetic coupling, constant,  $m$  stands for the lepton mass and  $K_P$  is a function that depends on  $\underline{x}$  via  $T$ , which we will not specify further at this stage; it is characteristic for the plasma phase.

In the *hadronic phase* it is the invariant mass  $M$  of the dilepton pair that determines which process makes the leading contribution to the elementary cross section  $\sigma$  in Eq. (5.49). For small-mass pairs (small compared to the  $\rho$  peak), bremsstrahlung-type emission of soft virtual photons is important. For masses  $M$  well above the order of  $T_c$ , processes of the type  $h\bar{h} \rightarrow \gamma^* \rightarrow \mu^+ \mu^-$  play a role, where  $h\bar{h}$  are hadrons others than pions. Their thermal production should be suppressed due to the relatively low temperatures, unless the in-medium masses are changed due to effects of chiral symmetry restoration, which may be dramatic! The only process that has been considered thus far in Cleymans *et al.*, (1987) is  $\pi^+ \pi^- \rightarrow \rho \rightarrow \mu^+ \mu^-$ . The electromagnetic cross section  $\sigma(M)$  is modified by the strong interactions of the pions, leading to

$$\sigma_\pi(M) = F_\pi^2(M) \left( 1 - \frac{4m_\pi^2}{M^2} \right)^{1/2} \sigma(M), \quad (5.55a)$$

where

$$\sigma(M) = \frac{4\pi}{3} \frac{\alpha^2}{M^2} \left[ 1 + \frac{2m_\pi^2}{M^2} \right] \left[ 1 - \frac{4m_\pi^2}{M^2} \right]^{1/2}. \quad (5.55b)$$

The pion form factor  $F_\pi$  is calculated in the vector-meson-dominance approximation, in which only the  $\rho$  pole is kept in the sum over all  $\rho$ -like resonances.

Furthermore, the functions  $f_1$  and  $f_2$  in Eq. (5.49) have to be replaced by Bose-Einstein distributions. The result for the *dilepton rate in the hadron phase* is given by

$$\frac{dN_H}{d^4x d^4p} = \frac{\alpha^2}{48\pi^4} \left[ 1 + 2 \frac{m_\pi^2}{M^2} \right] \left[ 1 - 4 \frac{m_\pi^2}{M^2} \right]^{3/2} |F_\pi(M^2)|^2 \times \exp\{-E/T\} \cdot K_H(p, T), \quad (5.56)$$

where  $K_H$  is a slightly different function from  $K_P$  in Eq. (5.54).

#### b. Integration over the spacetime history

The assumption of Bjorken's scaling solution for the hydrodynamical expansion simplifies the evaluation of the spacetime integration. The appropriate representa-

tion of the volume element is  $d^4x = \tau d\tau dy d^2x_T$ , where  $\tau$  is the proper time coordinate in a comoving reference frame with the fluid,  $y$  is the spacetime rapidity (denoted as  $\eta$  above),  $x_T$  are the transverse coordinates. The transverse velocity is neglected, the transverse distributions of thermodynamic quantities are taken as step functions, and the energy density  $\varepsilon$ , pressure  $p$ , entropy density  $s$ , and temperature  $T$  depend then only on the time  $\tau$ . The dileptonic yields per invariant mass, transverse momentum and unit rapidity interval  $dY$  are given by

$$\frac{dN_{Ph}(\text{hydro})}{dM^2 d^2p_T dY} = \frac{1}{2} \int_{\tau_i}^{\tau_f} \tau d\tau \int_{y_{\min}}^{y_{\max}} dy \frac{dN_{Ph}}{d^4x d^4p} [T(\tau)]. \quad (5.57)$$

The index  $Ph$  indicates the dependence on the phase,  $\tau_f - \tau_i$  is the duration of a certain phase or a mixture of phases, and  $y$  varies between the minimal ( $y_{\min}$ ) and maximal ( $y_{\max}$ ) spacetime rapidity of the beam. The *equation of state* enters the cooling law  $T(\tau)$ , and the phase duration is sensitive to the dynamics of the transition.

Cleymans *et al.* (1987) have discussed the following scenarios:

(i) A first order transition described by a bag model equation of state proceeding in equilibrium or with supercooling and subsequent superheating.

(ii) A second order transition with an equation of state taken from lattice Monte Carlo results (Redlich and Satz, 1986).

As we saw in Sec. V.B.2, the bag model equation of state leads to a  $T^3$  dependence of the entropy density and is typical for an ideal gas. In the baryonless plasma phase for two massless flavors, the entropy density is given by  $s_P = 4 \times 37 \pi^2 T^3 / 90$ . In the hadron phase described as an ideal gas of massless pions we have  $s_H = 2 \pi^2 T^3 / 15$ . In the case of the second-order transition  $s(T)$  has been read off from numerical data of Redlich and Satz (1986), which were obtained using lattice QCD. Using the scaling solution for isentropic flow [ $s(\tau)\tau = \text{const}$ , see Sec. V.B.1], we find the relations  $T(\tau)$ . When the scaled temperature  $T/T_0$  is plotted as a function of the scaled eigentime  $\tau/\tau_0$  (where  $T_0$  and  $\tau_0$  refer to the initial values chosen as  $T_0 = 284$  MeV and  $\tau_0 = 1$  fm/c), the characteristic difference between a first-order transition proceeding in equilibrium and a second-order transition is that the average temperature is higher in the quark-gluon plasma phase and lower in the hadronic phase for a second-order transition. Having  $T(\tau)$  at hand, the integration in Eq. (5.57) can be carried out. Depending on the scenario, four production rates are distinguishable:

(i) The dilepton production rate in the pure plasma phase produced in the time interval between  $\tau_0$  and  $\tau_P$ , where  $T(\tau)$  is given as  $T_0(\tau_0/\tau)^{1/3}$  for the bag model equation of state.

(ii) The dilepton production rate in the pure hadronic phase produced during the interval between  $\tau_H$  and the freezeout time (where the application of hydrodynamic concepts is already questionable).

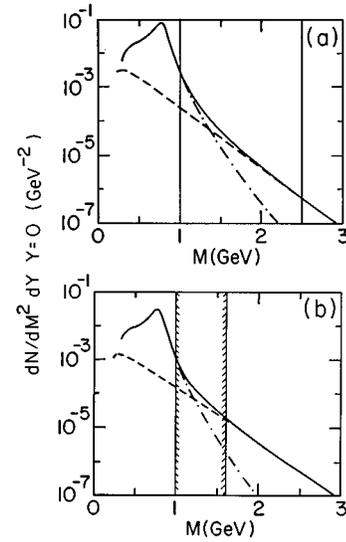


FIG. 31. Production rate of muon pairs vs the invariant mass  $M$  of the dilepton pair calculated (a) for a first-order transition and (b) for a second-order transition. The dashed-dotted line indicates the hadron contribution; the dashed line shows the quark contribution; the solid line is the sum of the quark and hadron contributions. The two vertical lines separate regions in which different phases make the dominant contribution, from left to right: the hadronic phase, the interference region, the quark-gluon plasma phase. From Cleymans *et al.* (1987).

(iii) If there is a coexistence of phases as is expected in the first-order case, there is a mixed phase at temperature  $T_c$ , where the plasma and hadron phases contribute  $dN_P/dM^2 d^2p_T dY$  and  $dN_H/d^2M d^2p_T dY$ , respectively. Both rates are calculated separately. Their relative weight is described by a factor  $f$ , which is the fraction of the entropy in the plasma phase. Its value follows from  $s(\tau)\tau = \text{const}$  and the constraint

$$s(\tau) = f(\tau)s_P + [1 - f(\tau)]s_H. \quad (5.58)$$

The entropy densities  $s_P$  and  $s_H$  remain constant for constant  $T = T_c$ .

The ratio  $s_P/s_H$  of entropy densities also gives an estimate for the time scales  $\tau_H$  and  $\tau_P$ , as  $\tau_H/\tau_P = s_P/s_H = 37/3$  if we assume for a moment that the naive counting of degrees of freedom is appropriate close to  $T_c$ . This ratio is frequently quoted in the literature, although the number three in the hadronic phase (resulting from three pions) is most likely incorrect. The large difference in entropy densities between the two phases explains the long duration time of the mixed phase. It takes time to rearrange the effective degrees of freedom since they have to be reduced by an order of magnitude for the above counting when the conversion to the hadronic phase sets in.

We display the results of Cleymans *et al.* (1987) only for the cases of (a) a first-order transition with Maxwell construction and (b) a second-order transition. Upon integration over the transverse momenta, the resulting dileptonic production rates  $dN/dM^2 dY$  at  $Y=0$  GeV $^{-2}$  are shown in Figs. 31(a) and 31(b). The main difference is the extent of the interference region in

which both phases contribute to dilepton production. It is considerably smaller for a second-order transition. In this case, the production mechanisms change instantaneously at  $T_c$ , while those of both phases would be at work in the coexistence phase of a first-order transition.

The width of the interference region does not yet provide a signal that is suitable for experiments to infer the phase-transition dynamics. The sudden change in the production mechanism is reflected in the average transverse momentum  $\langle p_T \rangle$  of the dilepton pair, or even more clearly in its derivative with respect to  $M$ . The average  $\langle p_T \rangle$  can be derived from the differential multiplicity  $R$ . A plot of  $d\langle p_T \rangle/dM$  as a function of  $M$  for both transition scenarios shows a peak in the  $d\langle p_T \rangle/dM$  distribution for a second-order transition. Thus a peak in the  $d\langle p_T \rangle/dM$  distribution is a possible candidate for a signal of a second-order phase transition (keeping in mind all approximations that have been applied so far). As outlined by Cleymans *et al.* (1987), such a sudden change in the slope of  $\langle p_T \rangle$  as a function of  $M$  can be also produced by a higher-mass resonance like the  $\rho'$  (1600). To disentangle the origins of bumps in the  $\langle p_T \rangle$  spectra, one should vary the initial temperature. At a higher initial temperature, the contribution from the plasma phase to the dilepton rate increases. A peak structure in the  $\langle p_T \rangle$  spectrum due to a second-order transition will be evident for low initial temperatures, while higher initial temperatures are necessary to see a similar structure in a first-order scenario.

Yet another possibility may lead to a sudden change in  $\langle p_T \rangle(M)$ . Dilepton rates are calculated in such a way that they do not automatically interpolate smoothly at  $T_c$ , although they should do so in the case of a second-order transition, where no discontinuities should be seen in thermodynamic observables. Cleymans *et al.* (1987) have investigated the influence of smoothly interpolating the rates between both phases. The bump structure, however, survived the interpolation between  $dN_P/d^4x d^4p$  and  $dN_H/d^4x d^4p$ .

In general, it is necessary to calculate the electromagnetic current-current correlation  $W(x,p)$  (whose space-time integral is proportional to  $dN/dM^2 dp_T dY$ ) in terms of one effective model that replaces the bag model in the case of a second-order transition. We mentioned this point already in Sec. V.B.2.

So far we have been concerned with thermal production rates and specific features reflecting the transition dynamics in the rate dependence of the invariant mass  $M$ . In heavy-ion experiments it is obviously an objective of high priority to guarantee that thermal dileptons can be distinguished from various background sources. In the order of increasing masses, we have Dalitz pairs, soft gluon bremsstrahlung, hadronic resonance decays, pre-equilibrium production, and Drell-Yan production in the very early stage of the collision.

Ruuskanen (1990) has derived an analytic expression for the integrated thermal production rate  $R$ . The characteristic feature of this expression concerns the  $T$  and  $M_T$  dependence compared to other production mechanisms [ $M_T = (M^2 + p_T^2)^{1/2}$  denotes the transverse mass,

which is used instead of the transverse momentum  $p_T$ ]. The dependence of the thermal rates on the initial temperature  $T_i$  is strong, proportional to  $T_i^6$  compared to the  $T_i^3$  dependence of a Drell-Yan rate. Recall that  $T_i$  at equilibration time  $\tau_i$  can be related to the hadronic multiplicity in the final state according to

$$(\tau_i T_i^3)^2 \propto (dN_\pi/dY)^2, \quad (5.59)$$

if an isentropic expansion scheme is assumed. This allows an identification of thermal dileptons via their strong (quadratic) dependence on pion multiplicity fluctuations; Drell-Yan production rates show only a linear dependence. At fixed multiplicity (i.e., fixed  $\tau_i T_i^3$ ), the thermal production rates are expected to show an approximate power-law behavior proportional to  $M_T^{-6}$  in the range  $M_T/T_i < 5.5 < M_T/T_c$  (Ruuskanen, 1990) with a  $\langle p_T^2 \rangle$  dependence proportional to  $M^2$ , and an exponential decrease in the dilepton production rate for large  $M_T/T_i$ , with  $\langle p_T^2 \rangle$  showing a linear dependence on  $M$ . For an initial temperature of 600 MeV, the window of the  $M_T^{-6}$  dependence of the transverse-mass spectrum lies between 1 and 3 GeV. Its lower end has to face a background from the hadronic phase, its upper end the Drell-Yan background.

Thus the first goal of a measurement of dilepton rates in a collision of heavy-ion nuclei is to identify a window in the transverse-mass spectrum that shows the behavior of a thermal rate produced in the plasma phase. It is easier to filter out the  $M_T^{-6}$  dependent part if high initial temperatures are reached in the experiment. For detailed calculations of the different backgrounds and their competition with thermal rates we refer the reader to, for example, Ruuskanen (1992), Weldon (1991), Kajantie and Miettinen (1981), and to McLerran and Toimela (1985). For more recent predictions of the dilepton spectrum at energies of the CERN Linear Hadron Collider, see Redlich *et al.* (1992). The effect of various equations of state on the transverse expansion and the dilepton production rate is discussed in Kataja *et al.*, (1992). The background of dileptons from  $\rho$  decays is usually considered undesirable from the point of view of thermal dilepton identification. Recently it has been shown by Heinz and Lee (1992) that such dileptons may contain valuable information on the lifetime of the hadronic phase; that is why they are called “fast clocks” for heavy-ion collisions. The overall judgement is that the identification of thermal rates in the different phases is difficult but not completely unrealistic. Yet it seems to be an order of magnitude more difficult to infer specific features of the transition itself from structures of  $R$  or  $\langle p_T \rangle$  distributions as functions of  $M$ .

To summarize, the approximations that enter predictions of dilepton production rates, along the lines we have indicated above, are:

(i) Uncertainties in the background to thermal dileptons. The Drell-Yan background is the one that is best understood. Corrections due to higher orders in QCD have been assumed to be small, in fact they are large in the usual deep-inelastic structure-function approach.

[For recent results on Drell-Yan production at collider energies including  $O(\alpha_s^2)$  contributions we refer the reader to van Neerven (1995).] Deviations from the linear dependence of the structure functions on the mass number  $A$  (entering the Drell-Yan cross section for colliding nuclei) are assumed to be negligible in the relevant mass range ( $M=2-6$  GeV). Probably the least understood rates are those that are produced shortly after the collision but before the onset of thermal equilibrium. Effective models are needed to describe nonequilibrium properties of the evolution of the initial state. In the fluxtube model of Kajantie and Matsui (1985) and Bialas and Czyz (1984, 1985), dilepton rates depend on the strength of the color field that is formed after the color exchange between the colliding nuclei. The new production mechanism relies on plasma oscillations of quarks and antiquarks. In the mass range of a few GeV, which is of special interest for thermal production, the worst perspective is that both the thermal and the Drell-Yan productions—including all their specific signatures—are buried under the yields of plasma oscillations. For further discussions, see also Bialas *et al.* (1988) and Bialas and Blaizot (1985).

(ii) The applicability of hydrodynamic concepts make it possible to evaluate the integral over the spacetime history of production rates at a given point in phase space. This is marginal in several aspects. The volumes are not so large compared to the mean free paths, at least not close to freeze-out, where by definition the system is so diluted that interactions can be neglected. The scale of the equilibration time is estimated in a purely phenomenological way. Thus it cannot really be argued which time  $\tau_0$  is to be used for the onset of hydrodynamic expansion. The adiabatic expansion guarantees the conservation of entropy and leads to a simple relation between entropy and time. During the phase conversion this condition may be violated.

(iii) The bag model equation of state has been used close to  $T_c$ , although strong deviations from the ideal-gas behavior are expected there. Lattice Monte Carlo calculations should do their best to include the effect of dynamical fermions as realistically as possible and to derive an equation of state for temperatures close to  $T_c$ . In deriving the equation of state that was used by Cleymans *et al.* (1987), which led to a second-order transition, the fermions were treated in the hopping parameter expansion; the parameters involved are known not to be close enough to the continuum limit. A more realistic equation of state obtained from lattice calculations with dynamical fermions can be found, for example in, Christ (1992), although it should not be considered as *the* realistic equation of state for QCD (see Sec. III.B.2). Given the hydrodynamic approach and the equation of state, further simplifications were made in deriving the results we have shown. We have neglected effects due to

(i) A realistic nuclear geometry

(ii) The impact parameter dependence. In particular, the initial temperature profile depends on a nonzero impact parameter  $b_0$ .

(iii) The transverse-coordinate dependence of temperature at all stages of the evolution.

(iv) The transverse flow of the hadron/plasma fluid. When transverse flow is included, hadronic rates can be reduced by an order of magnitude due to the reduced lifetime of mixed and hadronic phases, see Fig. 9 of Ruuskanen (1990).

(v) Last, but not least, the possible temperature and density dependence of the masses that are involved in dilepton production. Meson masses can be expressed in terms of the chiral condensate  $\langle\bar{\psi}\psi\rangle$ ; they should be sensitive to the order of the transition. Either  $\langle\bar{\psi}\psi\rangle$  vanishes smoothly when approaching  $T_c$  from below, or two values of  $\langle\bar{\psi}\psi\rangle$  coexist at  $T_c$ , a large value and a small one. Finally we refer the reader to the work by Geiger and Kapusta (1993), Shuryak and Xiong (1993), and Ruuskanen (1992) for more recent references in this field.

When dileptons are produced from vector-meson decays, the dilepton rates are also sensitive to the phase transition dynamics via vector-meson mass changes in the medium both as a function of  $T$  and  $\mu$ . We have skipped this possible manifestation of the transition dynamics, as we are not aware of any specific prediction to date.

### c. Real photons

Real photons should be understood in contrast to virtual photons in dilepton production. Like dileptons, they indicate the electromagnetic response to the plasma evolution. They can provide clean signals in the sense that they escape from different stages of the system with rather small interaction cross sections. According to their production conditions, the three major contributions to real photons are *direct photons* from partonic processes in the initial state, *thermal photons* from both phases, and *decay photons* produced at a late stage of the evolution. At large transverse momenta  $K_T \geq 3$  GeV, the lowest order ( $\alpha_s$ ) contributions to direct photons are the gluon-to-photon Compton scattering off (anti)quarks  $q(\bar{q})g \rightarrow q(\bar{q})\gamma$ , quark-antiquark annihilation to photon and gluon processes  $q\bar{q} \rightarrow \gamma g$ , and bremsstrahlung (with quark fragmentation into the photon). At very high collision energies, the main contribution to direct photons comes from the fragmentation of jets (Gupta, 1990). At low transverse momenta  $K_T \leq 3$  GeV, decays of neutral mesons  $\pi^0$ ,  $\eta$ ,  $\eta'$ , and  $\omega$  make a large contribution. This contribution is particularly difficult to estimate, but will be measured (Stachel, 1993). The uncertainty in the estimate arises from a lack of knowledge of the  $p_T$  distributions of these mesons in nuclear collisions. For very small transverse momenta ( $K_T \leq 10-20$  MeV), hadronic bremsstrahlung is important (Ruuskanen, 1992).

A possible window for observing thermal photons could be in the 2–3 GeV region (Ruuskanen, 1992). Thermal photons can be produced in both phases. In earlier calculations, only the plasma contribution was considered (see, for example, Hwa and Kajantie, 1985). Thermal photons in the plasma phase are produced via

the same elementary processes as direct photons; the difference in the final thermal-photon rate compared to the direct-photon production rate is a replacement of the product of structure functions by the square of the entropy. The contribution from the hadronic phase has been calculated by Kapusta, Lichard, and Seibert (1991, 1992), with the result that “a hadron gas shines as brightly as a quark-gluon plasma” at temperatures between 150–200 MeV.

In the 1 GeV region, thermal-photon production rates are clearly below those of decay photons. This does not necessarily imply that thermal rates cannot be measured. An electromagnetic calorimeter with high energy resolution and fine granularity has been proposed for the Relativistic Heavy-ion Collider at Brookhaven; it aims to measure the ratio of (No. of thermal photons)/(No. of  $\pi_0$ ) as a function of  $p_T$  to within  $\sim 3\text{--}5\%$  (Stachel, 1993). In view of Sec. V.D.3, we anticipate that an accurate measurement of this ratio is rather valuable for verifying certain remnants of a hypothetical second-order chiral transition. At large transverse momenta ( $\geq 3$  GeV), direct photons dominate. The chances for thermal photons improve at very large multiplicities  $dN_\pi/dY$  of pions, because thermal photons are proportional to  $(dN_\pi/dY)^2$ , while photons from pion decay vary with  $(dN_\pi/dY)$ . In the intermediate region, whether or not thermal photons have a chance of being identified depends sensitively on the shape of the hadronic  $p_T$  spectra, which are not well understood. Real photons have certain advantages over dileptons. Prompt real photons originate from the Compton scattering of gluons and quarks. Gluons equilibrate very rapidly compared to quarks. Therefore the prompt-photon measurement provides a direct (and the only direct) access to the very early initial phase of the collision, which is not accessible with dilepton production rates. The signal-to-noise ratio is expected to be larger than for the dilepton continuum.

Thus it seems to us that future experiments with large  $dN_\pi/dY$  and high resolution detectors should offer promising opportunities to look for signatures of a transient plasma and even for certain substructures in thermal-photon distributions that could be conclusive in view of the phase-transition dynamics. The procedure to calculate the thermal production rate of real photons is the same as that for dileptons. Once the elementary processes are identified and emission rates are calculated in the stationary plasma or hadronic phase, the result has to be combined with the spacetime expansion. In a first approximation one can use Bjorken’s similarity flow and correct for transverse flow and nuclear geometry in later steps. For such calculations we refer the reader to Ruuskanen (1992) and to references therein. High multiplicities ( $\geq 4000$ ) and short equilibration times favor thermal-photon production. We shall have to wait and see whether the relativistic heavy-ion collider or the linear hadron collider do indeed reach such multiplicities.

### 3. Strangeness production

In heavy-ion collisions, an enhancement of certain particle ratios involving strange quarks is predicted com-

pared to these ratios in hadron-hadron or hadron-nucleus collisions. Ratios enhanced by a factor of 2 for  $K^+/\pi^+$  have in fact been observed in experiments at the AGS accelerator at Brookhaven and SPS accelerator at CERN. A theoretical explanation in terms of a transient quark-gluon plasma is not compulsory. Rescattering processes in a hot hadronic gas can also change the ratios in the same direction.

The idea that strangeness could be a possible signature for a quark-gluon plasma relies on the following arguments. Strangeness production in the plasma should be facilitated for two reasons:

(i) Independently of the assumed baryon density, the threshold energies for strangeness production in the plasma phase are much lower than those for the production of strange baryons or mesons in the hadron phase. In the hadron phase, strange mesons or baryons are typically made in a collision of two nonstrange hadrons. The reaction with the lowest threshold energy already requires 671 MeV ( $p+n \rightarrow \Lambda^0 + K^+ + n$ ). In the plasma, the threshold for  $s\bar{s}$  production is equal to the rest mass of  $s\bar{s}$ , about 300 MeV. For a given temperature, the density of noninteracting strange quarks is higher than the density of noninteracting kaons; these will be discussed later as an example. In a plasma an  $s$  quark has two spin and three color degrees of freedom. If it is bound in a  $K^-$  or a  $\bar{K}^0$  in the hadron phase, only two degrees of freedom remain. Also, the kaon mass is  $\sim 494$  MeV, compared to the current  $s$ -quark value of  $180 \pm 30$  MeV.

(ii) The second reason applies to a baryon-rich environment, which can be found in collisions with high baryon stopping or in the rapidity range of fragmentation regions. The Pauli exclusion principle will prohibit the creation of  $u\bar{u}$  and  $d\bar{d}$  pairs instead of strange pairs. The light quarks have to supply the large Fermi energy represented by the chemical potential  $\mu_B$ , while the  $s$  quarks are suppressed only by their finite mass. Moreover  $\bar{u}$  and  $\bar{d}$  quarks have a high probability to recombine with  $u$  and  $d$  quarks to form gluons. For strange quarks the recombination to gluons is less likely. Due to the volume expansion and the decreasing temperature of the fireball, the process  $gg \rightarrow s\bar{s}$  is soon out of thermal equilibrium, in other words, it is too slow to proceed in the inverse direction. In a background that is rich in  $u$  and  $d$  quarks (large  $\mu$ ), the number of (massless)  $\bar{u}$  or  $\bar{d}$  quarks in a free gas is given by

$$n(\bar{u}) = 6 \int \frac{d^3p}{(2\pi)^3} \frac{1}{\exp\{(p+\mu)/T\} + 1}. \quad (5.60)$$

This number should be compared to the number of strange quarks

$$n(s) = 6 \int \frac{d^3p}{(2\pi)^3} \frac{1}{\exp\{(p^2 + m_s^2)^{1/2}/T\} + 1}. \quad (5.61)$$

To lowest order in perturbative QCD,  $s\bar{s}$  pairs are created in collisions of two gluons and by annihilation of light quark and antiquark pairs (Rafelski and Müller, 1982). Rafelski and Müller have shown that gluonic production is the dominating mechanism of strangeness

production. For a small window of the quark mass equal to 0.5–1 times the temperature, the equilibration time is similar to the expected lifetime of the plasma phase.

For dileptons it has been sufficient (see the previous section, Sec. V.C.2) to calculate the differential rate per given spacetime volume. Due to the very small electromagnetic cross section the dileptons of high invariant mass escape from the plasma without further interaction.

Strange quarks are kept in the subsequent evolution. This is the essential difference of strange production compared to dilepton production. In various ways strange quarks are incorporated as hadronic constituents during the conversion of the plasma into the hadron phase. After completion of the conversion, strangeness is confined to hadrons. The strange hadrons continue to react via strangeness creation, annihilation or exchange reactions until freeze out. Thus it is not sufficient to calculate the formation rate of strangeness per unit time and volume as if the predicted rate would directly correspond to observed multiplicities (as for dilepton production).

The rate of change  $dn/dt$  in the density  $n$  of strange particles is of primary interest now. Later we shall see that, in the case of a static plasma,  $dn/dt$  is equal to the rate  $R$  if  $n$  vanishes, that is, at the start of the production. The change goes to zero when the equilibrium value of  $n$  ( $n_{\text{eq}}$ ) has been reached. It is a question of time whether  $n_{\text{eq}}$  can be reached at all. The two competing time scales are the equilibration time  $\tau_{\text{eq}}$ , defined by  $n(\tau_{\text{eq}}) = n_{\text{eq}}$ , and the lifespan of the phase in which strangeness production is considered.

At this point, one can already suspect that the appropriate “choice” of the dynamical scenario for the phase transition is even more essential here than it was in the case of dilepton production. Time-scale estimates have more far-reaching consequences in strangeness production. The contribution of a thermal rate is not only a question of size in the sense that it could be certainly identified only if it is large enough compared to other backgrounds. Certain rare multistrange hadron species may completely fail as candidates for quark-gluon plasma signatures if their equilibration time is just too long compared to the lifetime of the phase in which they could have been otherwise produced.

We shall now describe in more detail the differential equation that determines the equilibration time for a given species.

We shall not judge the chances of using strangeness enhancement as a signature for a transient plasma phase. Moreover, we do not see a way of specifying features that could be conclusive in terms of the underlying phase-transition dynamics. We merely want to give a flavor of the complexity involved in calculating equilibration times. The ratio of  $K^-/\pi^-$  serves only as an example. We shall describe the work of Kapusta and Mekjian (1986) and indicate places for further extensions or alternative treatments.

The preconditions that are taken for granted are the following: (a) the rate constant  $R$  for producing strange

quarks in the plasma phase is so large that the density quickly attains its chemical equilibrium value  $n_s^{\text{eq}}$ , and (b) the system expands and cools so rapidly that the strangeness abundance—once strangeness has been produced—is not lost to a significant extent due to annihilation into gluons. Both conditions ensure that the facilitated strangeness production in a plasma environment has a chance to become manifest in possible plasma signals.

To obtain numbers for particle ratios at the very end, we have to specify the following.

(i) The kinematical approach and the geometry. Bjorken’s ansatz of a longitudinal hydrodynamic scaling expansion is used. Possible effects due to viscosity, heat conduction, and transverse expansion are neglected. Viscosity and heat conduction would prolong the decrease of temperature as a function of time; transverse expansion would accelerate it. These effects may or may not approximately cancel.

(ii) The rapidity range. The central rapidity range has been chosen to be the region in which the baryon number approximately vanishes. Hence the calculations are performed at  $\mu_B = 0$ . This may not be a favorable choice, in view of the differing sensitivities of  $\bar{u}, \bar{d}$  and  $s, \bar{s}$  quarks with respect to  $\mu_B$ , but it simplifies the discussion considerably.

(iii) The equation of state. Once again the bag model equation of state is used.

(iv) The gross features of the phase transition. Corresponding to the bag model, the phase transition is assumed to be of first order. Two possibilities are considered: a smooth transition via the Maxwell construction and a rapid transition with supercooling and subsequent reheating.

(v) The observable to measure for the strangeness contents of each phase. For the hadron phase  $K^-$ -mesons will be considered.

(vi) The elementary reactions entering the cross sections in the rate formulas.

(vii) An ansatz for the differential rate equation that determines the equilibration time.

Let us start at the end and describe the rate equation as was used in Kapusta and Mekjian (1986). For an expanding plasma, the rate of change in the density of strange quarks  $n_s$  is given by

$$\frac{dn_s(t)}{dt} = R_p(T(t)) \left\{ 1 - \left[ \frac{n_s(t)}{n_s^{\text{eq}}(T(t))} \right]^2 \right\} - \frac{n_s(t)}{t}. \quad (5.62)$$

Here  $n_s^{\text{eq}}$  is the equilibrium density of strange quarks, which depends on time via the temperature  $T$ . The rate  $R_p$  is defined as the number of specific reactions per unit time and volume producing  $s\bar{s}$  pairs in the plasma phase. It corresponds to  $\int d^4p dN/(d^4x d^4p)$ , which also entered the dilepton rate in Sec. V.C.2. Similarly to Eq. (5.48), the rate is calculated in the independent-particle limit, that is, according to the ansatz

$$R(a_1 + a_2 \rightarrow X) = \frac{1}{1 + \delta_{a_1, a_2}} \int \frac{d^3 k_1}{(2\pi)^3} f(\mathbf{k}_1) \int \frac{d^3 k_2}{(2\pi)^3} \times f(\mathbf{k}_2) \sigma(a_1 + a_2 \rightarrow X; \mathbf{k}_1, \mathbf{k}_2) v_{\text{rel}}, \quad (5.63)$$

in which the occupation probabilities  $f(\mathbf{k})$  are chosen as Boltzmann weights

$$f(\mathbf{k}) = \exp\{- (\sqrt{k^2 + m_a^2})/T\}, \quad (5.64)$$

where  $m_a$  is the degenerate mass of particle  $a_1$  and  $a_2$ . This ansatz neglects possible final-state interactions due to the influence of Pauli blocking. Here it is consistent with considering a baryon-free rapidity region. The prefactor  $(1/1 + \delta)$  prevents double counting when gluon pairs annihilate to  $s\bar{s}$  pairs,  $v_{\text{rel}}$  denotes the relative velocity between the particles  $a_1$  and  $a_2$  and is given by a formula analogous to Eq. (5.50). Finally,  $\sigma$  denotes the elementary cross sections which contribute to  $R$  in a given phase. Note also that the spacetime dependence enters  $f$  via the temperature  $T$ .

Some further comments on the rate equation (5.62) are in order. The last term in Eq. (5.62) is a dilution term due to the volume expansion. The particle density will decrease roughly like the entropy density, that is, as  $1/t$  when the volume increases linearly with  $t$ . The first two terms of Eq. (5.62) characterize the rate equation for a nonexpanding source. Restricting the right-hand side of Eq. (5.62) to the static part, we note that the linearized rate equation implies for small deviations from equilibrium

$$\tau_{\text{eq}} = n_{\text{eq}}/(2R), \quad (5.65)$$

that is, a high rate  $R$  implies a short equilibration time.

The cooling law as a function of time depends on the chosen scenario. In the plasma phase the bag model equation of state implies a decrease according to

$$T(t) = T_0(t_0/t)^{1/3}. \quad (5.66)$$

Other quantities entering the rate equation (5.62) are the equilibrium density  $n_s^{\text{eq}}$  of strange particles and the rate  $R_p$  ( $gg$  or  $q\bar{q} \rightarrow s\bar{s}$ ) for producing strange quarks in the plasma phase. The density  $n_s^{\text{eq}}$  is approximated by its Maxwell-Boltzmann limit. The rate  $R_p$  is taken from Müller and Rafelski (1982) and is based on the lowest order QCD cross section, for which  $s\bar{s}$  pairs are created via  $gg \rightarrow s\bar{s}$ ,  $u\bar{u} \rightarrow s\bar{s}$ , and  $d\bar{d} \rightarrow s\bar{s}$ . (It remains to be checked that suppressions due to higher orders in  $\alpha_s$  are not compensated by large combinatorial factors in the graphical expansion.) The rate  $R_p$  depends on  $\alpha_s$  and  $m_s$  as parameters, both of which vary with the renormalization scheme, the energy scale, and the temperature.

Finally, in order to solve the rate equation for  $n_s(t)$  in the plasma phase, one has to specify the initial condition. Since the net strangeness of two cold nuclei is zero, one can choose to set  $n_s(t_0) = 0$ , where  $t_0$  denotes the formation time.

The dependence of the rate on this initial condition should be weak, if the final rate is close to its equilibrium value, otherwise it has to be checked carefully. The conserved charge is the difference of strange and anti-

strange particle numbers. Only as long as strange and anti-strange particles do not interact, their charges are conserved separately. Thus if the  $s$  quark density represented by the  $K^-$  mesons is to be considered separately, as in Kapusta and Mekjian (1986), one relies on the treatment of quarks and mesons as effectively ideal gases in both phases. The treatment as ideal gases is implemented by the choice of equation of state. As lattice results have shown, the ideal-gas picture is certainly not correct in the vicinity of  $T_c$ —in particular, it fails if long-range correlations are enforced by a second-order transition. We stress this point, because in the presence of interactions between  $s$  and  $\bar{s}$  quarks it is not clear what a sensible measure for the strangeness contents of each phase could look like.

Next we turn to the rate equations in the intermediate phase between the onset of hadronization and its completion. The terms in the rate equation depend on the scenario. Kapusta and Mekjian (1986) considered two alternatives, the Maxwell scenario and a phase conversion with superheating and supercooling effects.

In the first case, the transition proceeds adiabatically. The system stays in local equilibrium, and Maxwell's condition for coexisting phases is assumed to be fulfilled. The nucleation rate of hadronic bubbles is large compared to the expansion such that a fraction  $[1 - f(t)]$  of the hadron phase is created immediately after the temperature has reached its critical value  $T_c$ . This happens at time  $t_1$  given by

$$t_1 = (T_0/T_c)^3 t_0. \quad (5.67)$$

Equation (5.67) follows from entropy conservation and the equation of state (see Sec. V.C.2, where it was derived). Thereafter, the temperature stays constant until the transition to the hadronic phase is completed. Both phases contribute to the entropy density according to

$$s(t) = f(t)s_p(T_c) + [1 - f(t)]s_h(T_c), \quad (5.68)$$

where  $s_p$  and  $s_h$  are the entropy densities in the plasma and the hadron phases, respectively. Combined with the time dependence of  $s(t)$ , this implies that the time dependence of the fraction  $f(t)$  is

$$f(t) = \frac{37}{34} (T_0/T_c)^3 \frac{t_0}{t} - \frac{3}{34}. \quad (5.69)$$

All plasma is converted to the hadron phase by a time  $t_2$ , where  $t_2$  follows from  $f(t_2) = 0$ :

$$t_2 = \frac{37}{3} t_1. \quad (5.70)$$

The second scenario (superheating and supercooling) would be realized if an interface between the hadron and the plasma phase were to cost excessive energy, that is, if the surface tension between the plasma and hadronic bubbles is large. (As we showed in Sec. III, this is not likely to be the case.) The onset of the phase transition is delayed until the temperature has dropped to a value  $T_p < T_c$ . At  $T_p$  the system is assumed to transform instantaneously to the hadron phase. The hadron phase is superheated to a value  $T_h > T_c$ , while the plasma was

supercooled to  $T_p < T_c$ . The entropy should not decrease (during the transition it may increase), and the energy is conserved. These assumptions, together with the bag model equation of state, lead to a prediction of the time dependence of  $T$  in the various time intervals.

In view of the final strangeness contents the interesting feature about the time evolution of temperature is that the system spends quite a long time in the mixed phase or in the overheated hadron phase, depending on the scenario and the initial temperature  $T_0$ .

Even the very ansatz for the rate equations depends on the scenario. The common part is that which describes the approach to the equilibrium density of strange particles in a static situation. When matter is expanding, loss terms in the density have to be included; these depend on the presence or absence of a phase mixture. In particular, a coexistence of phases as occurs in Maxwell's construction leads to an additional complication. For the lifespan of the mixed phase the volume occupied by the hadron phase expands while the fraction of plasma phase shrinks. This structure has to be superimposed on the common volume expansion.

In both scenarios, the rate changes  $dn/dt$  in  $K^-$  mesons are followed over the period between the onset of the hadron phase until all of the plasma has been converted. In Maxwell's scenario, the temperature stays constant at  $T_c$ , while in the supercooling/reheating scenario it drops from  $T_h > T_c$  to  $T_c$ .

Finally, to solve the rate equations for the density of  $K^-$  mesons, we have to determine the rate constants  $R_h(T(t))$  and  $R_h(T_c)$  in the *hadronic* phase. Here the possible contributions simplify due to the restriction of a baryon-free environment. The relevant processes are  $\pi^+ + \pi^- \rightarrow K^+ + K^-$ ,  $\pi^- \pi^0 \rightarrow K^- K^0$ , and  $\pi^0 \pi^0 \rightarrow K^+ K^-$ . The corresponding cross sections must be inserted: these are inferred from experiments.

After all these specifications we are now ready to solve the rate equations for  $n_s(t)$ . In heavy-ion collisions, densities are not directly observable as functions of time. Instead of densities, particle ratios, in our case  $K^-/\pi^-$  ratios, are suitable observables. The volume drops out, leaving

$$K^-/\pi^- \propto n_s/n_{\pi^-}. \quad (5.71)$$

Here  $n_s$  is the density of strange particles at freeze-out, hidden to equal fractions in  $K^-$  and  $\bar{K}^0$  mesons, and  $n_{\pi^-}$  is the density of  $\pi^-$  mesons. For massless pions and under the assumption of an ideal gas, the pion number density is proportional to the pion entropy density  $s \propto T^3$ , which is almost equal to the total entropy density of the hadronic gas. This leads to

$$K^-/\pi^- \propto n_s/s_h. \quad (5.72)$$

This ratio should be compared to  $n_s(\text{plasma})/s_p$ , the ratio of the number of strange quarks in the plasma phase to the plasma entropy density. The question as to what would be a good measure for the strangeness contents of the plasma has been raised for some time. Rather than  $n_s/s$ , the ratios  $n_s/n_d$  ( $n_d$  being the number of down quarks) were considered earlier in Rafelski

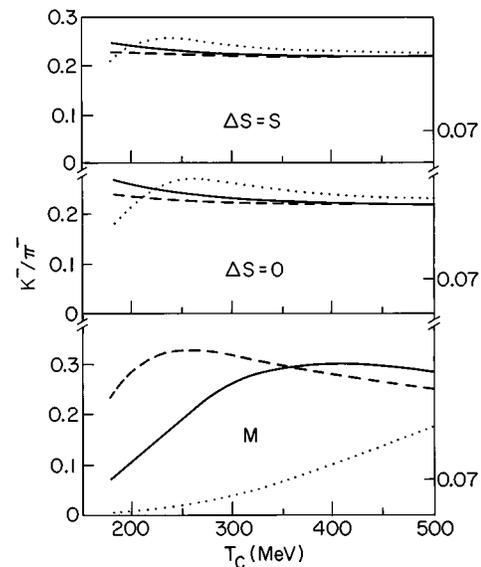


FIG. 32. Sensitivity of the ratios  $K^-/\pi^-$  at  $T_c$  (when the phase conversion is completed) to the chosen scenario— $\Delta S=S$ , 0, or Maxwell (M)—and the input parameters. For further explanations see the text. From Kapusta and Mekjian (1986).

(1982). However, the quantity which should replace  $n_{\pi^-}$  in the plasma phase is  $s$  and not  $n_d$ , since all  $u, d$  and gluonic degrees of freedom (“summed up” in the entropy) must hadronize into pions.

Plots of  $n_s/s$  and  $n_s/n_{\text{eq}}$  as a function of time can be found in Kapusta and Mekjian (1986) for three scenarios: Maxwell construction, the supercooling/reheating scenario without extra entropy production ( $\Delta S=0$ ), and with doubling of the entropy at  $T_c$  ( $\Delta S=S$ ). The curves display some expected features: the higher the initial temperature, the faster  $n_s$  reaches its equilibrium value. For the Maxwell scenario  $n_s/s$  and  $n_s/n_{\text{eq}}$  are continuous functions of time. After completion of the phase transition, the final ratio of  $K^-/\pi^-$  is shown in Fig. 32 as function of the initial temperature  $T_0$  for three scenarios and three sets of input parameters corresponding to the dashed, solid and dotted curves. Input parameters are the strong coupling  $\alpha_s$ , the strange-quark mass  $m_s$ , two cross sections entering the rates of  $K\bar{K}$  annihilation, and the proper formation time  $t_0$  of the plasma. Obviously in the supercooling/reheating scenarios with  $\Delta S=0$  or  $\Delta S=S$ , the  $K^-/\pi^-$  ratios are little sensitive to the initial temperature  $T_0$  and to the input parameters, and even to the degree of supercooling and the related entropy production. The value falls in the range 0.21–0.25. The value of chemical equilibrium is in this range, at 0.217. The  $K^-/\pi^-$  values, predicted here for heavy-ion collisions, should be compared to the same ratios in  $\bar{p}p$  collisions, in which  $K^-/\pi^-$  is  $0.07 \pm 0.02$ .

In the Maxwell scenario, such an enhancement by a factor about 3 is seen only for an “optimistic” choice of parameters (dashed curves in Fig. 32) for all values of  $T_0 \geq 200$  MeV. For a “realistic” choice (solid lines in Fig. 32) it is seen only for  $T_0 > 250$  MeV, and for a “pessi-

mistic" set (dotted lines in Fig. 32) only for large values of  $T_0$ . The physical reason for the higher sensitivity of the Maxwell scenario to the input parameters is likely to be the continuous time evolution throughout all stages of the evolution. There are no jumps in thermodynamic quantities at  $T_c$  to cut off the history of the plasma evolution.

The enhancement of  $K^-/\pi^-$  by a factor of 3 compared to  $\bar{p}p$  collisions at the same beam energy is to be interpreted as an indirect sign for a transient plasma phase. Without the first-order phase transition the lifetime of the hadron phase would be shorter, so short that there would not be sufficient time to saturate the  $K^-/\pi^-$  ratio to near its equilibrium value. Thus, in contrast to dileptons,  $K^-$  mesons are not a direct probe of the high-temperature plasma phase; their abundance is influenced by the phase transition and the subsequent evolution in the hadron phase.

This example of an estimate of the strangeness enhancement may have illustrated the urgency of calculating from first principles the equation of state, the latent heat, the surface tension, the temperature dependence of couplings and masses, and the order of the phase transition. Theoretical calculations may help to find the most realistic scenario. In particular, we have seen the simplifying power of the bag model equation of state, which does not justify its use. For strong correlations between particles close to  $T_c$ , the very indicator itself may fail as a measure for the strangeness contents of both phases. Also the gain and loss terms in the rate equations, which describe the dynamical expansion, are sensitive to *details* of the phase transition scenario.

The approach of Kapusta and Mekjian (1986) can be generalized in various aspects, which we mention rather briefly below. For our limited goal of identifying specific signatures for the phase transition, a detailed discussion of a more general framework would only serve to support further the overall impression: it is extremely difficult to separate the underlying phase-transition dynamics from strangeness production.

Koch *et al.* (1986b) have considered the strangeness production in the *plasma* phase within a framework, where the baryochemical potential is nonzero. The rate equation is considered in an alternative form, including a linear loss term in  $n_s/n_s^{\text{eq}}$ . The volume expansion is investigated in the cases of two different time dependences of  $V$  and  $T$ , one following from the bag model equation of state and one chosen by hand to simulate a more rapid expansion. In addition, Koch *et al.* have checked the influence of Pauli blocking on final states by replacing the momentum distributions of strange quarks in a suitable way.

A more general set of rate equations in the *hadronic* phase has been considered in Koch *et al.* (1986a). The underlying physical picture is similar to the Maxwell scenario of Kapusta and Mekjian (1986). It is more general in including baryons and antibaryons in addition to pions and kaons. A wider spectrum of possible reactions in the hadron phase has to be taken into account: strangeness production and exchange reactions, and an-

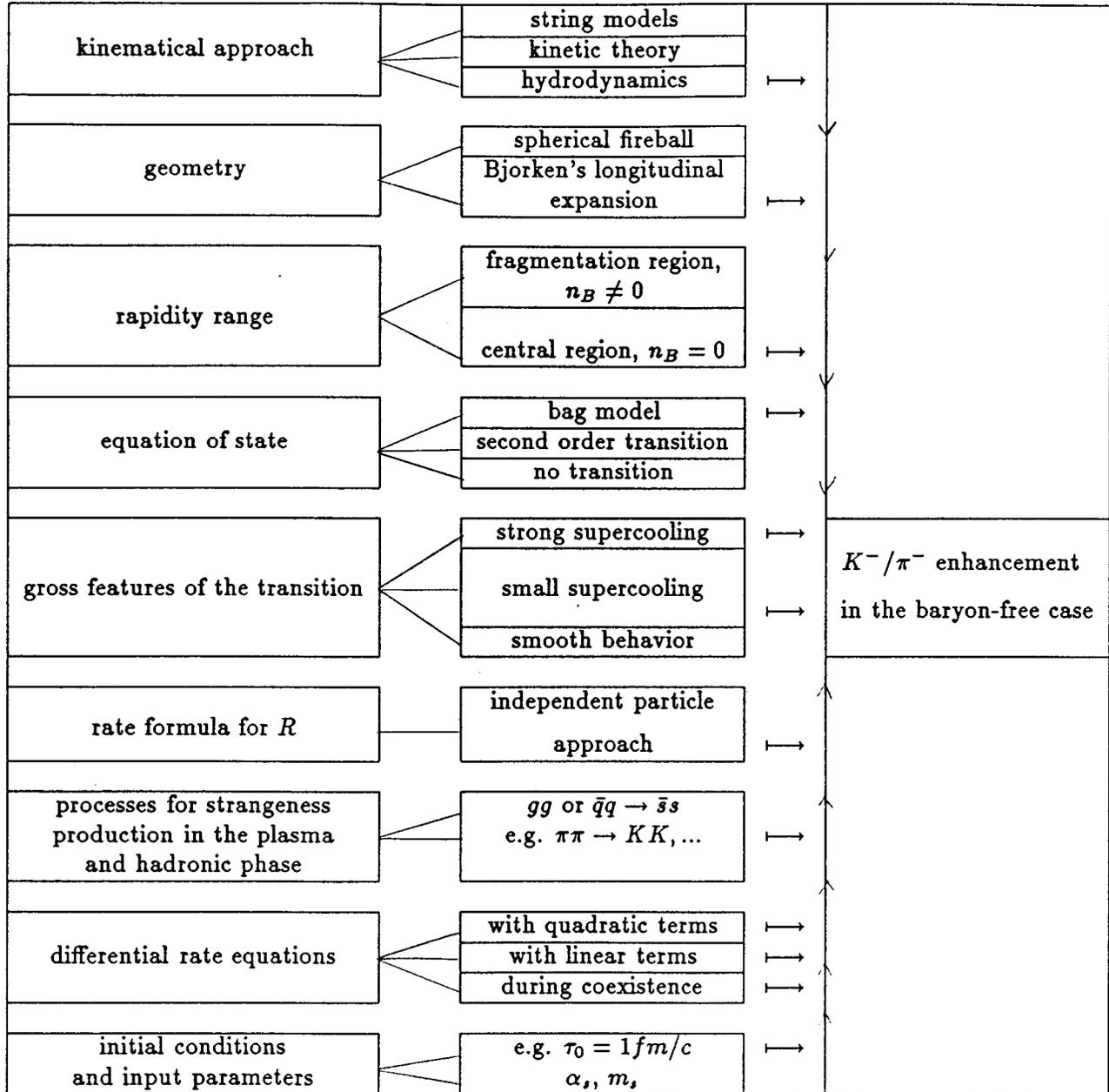
nihilation processes. The abundances not only of  $\pi$  and  $K$  mesons, but also of  $N, Y, \Xi, \Omega$ , and their antiparticles, have to be followed in their evolution. An additional term appears in the rate equation for the mixed phase that is specific for the chosen hadronization process (here the so called fragmentation-recombination scenario). Needless to say, the way in which the strangeness abundance of the plasma is modified during the phase conversion depends on the choice of the hadronization process. Unfortunately there is a whole variety of fragmentation models at one's disposal.

We conclude with a short and incomplete summary of the experimental status. The ratios of  $K^+/\pi^+$  and  $K^-/\pi^-$  have been measured in the E 802 experiment (Abbott *et al.*, 1990; Abbott *et al.*, 1991) at the AGS accelerator at Brookhaven and in the NA34 experiment at the SPS accelerator at CERN (Van Hecke, 1991). Both experiments operate in a baryon rich region of phase space, thus we cannot compare the results with the theoretical predictions of Kapusta and Mekjian (1986) for a baryon-free environment. In both experiments the same trends are seen: an enhancement of  $K^+/K^-$  ratios roughly by a factor of 2 when going from hadron-hadron to hadron-nucleus and nucleus-nucleus collisions, whereas the  $K^-/\pi^-$  ratio increases less strongly or stays approximately constant. It should be mentioned that the ratios  $K^+/\pi^+$  and  $K^-/\pi^-$  are not the only measure for strangeness enhancement. Another quantification is the strangeness suppression factor  $\lambda = 2s\bar{s}/(u\bar{u} + d\bar{d})$ , evaluated by Wroblewski (1985) for  $pp$  collisions and applied to S+S collisions at 200 GeV/nucleon (Seyboth *et al.*, 1992; Foley *et al.*, 1992). With this criterion one finds a factor-of-2 strangeness enhancement compared to  $pp$  and  $pA$  collisions.

This does not mean that a plasma has been observed. For many strange particles, the equilibrium values in a hot extended hadron phase are similar to those in a quark-gluon plasma. Thus the enhanced particle ratios do not uniquely signal a transient plasma. If there is another way to get strangeness into equilibrium by adjusting parameters in a hadronic cascade model, one will never be able to distinguish this possibility from a transient plasma. A system in equilibrium has lost the memory of how it got there. There may be exceptions that provide a unique signal of a plasma (Rafelski, 1991). Enhanced multistrange (anti)baryon production (antihyperons  $\bar{\Lambda}, \bar{\Xi}^-$ ) have been proposed as special tests for a plasma environment (Eggers and Rafelski, 1991). A hot hadron gas cannot overcome their large mass barriers.

In Table VIII we recapitulate the ingredients that have led to the prediction of  $K^-/\pi^-$  enhancement in the baryon-free case. Some freedom of choice is given for the geometry, the rapidity range, and the initial conditions. This freedom corresponds to the possibility of experimental tuning of certain parameters such as the initial energy density. In contrast to that, one is truly not allowed to choose between various versions for the kinematical approach, the equation of state, the gross features of the transition, the rate formula, the contributing

TABLE VIII. Flow diagram of “ingredients” for predicting the  $K^-/\pi^-$  enhancement in the very end (for special notations see the text).



elementary processes, and the ansatz for the rate formula or the differential rate equation. These ambiguities are due to uncertainties in the theoretical description and shed some light on the reliability of predicted numbers in the very end. For example the choice of a strong supercooling scenario must be regarded as *ad hoc*.

#### 4. Pion interferometry

Interferometry is nowadays a well-known technique. It was developed by Hanbury-Brown and Twiss (1954) almost 40 years ago in astrophysics, as a tool to measure the size of various stellar objects in the visible and radio frequency ranges. Figure 33 shows a source that is assumed to emit identical particles from positions  $P_1$  and

$P_2$ . The particles are later observed at positions  $P_3$  and  $P_4$ . Both emission points may contribute to both observation points, even if the particles are noninteracting, but have small relative momenta. The reason lies in the symmetrization (antisymmetrization) of the quantum-mechanical wave function in the case of bosons (fermions). A correlation function is constructed from the number of counts at  $P_3$  and  $P_4$ . The particles detected in astrophysical interferometry experiments are photons.

The overwhelming majority of particles radiated in heavy-ion collisions are pions. Thus we describe pion interferometry as a particularly useful tool in estimating the lifetime and the final-state size of the source that radiates the pions. It is evident from the discussion in

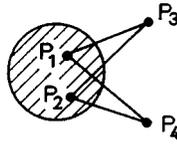


FIG. 33. Source emitting photons at points  $P_1$  and  $P_2$ , which are registered at  $P_3$  and  $P_4$ .

the previous sections that we need some measurements delivering information about the equation of state. Production rates of dileptons and strange particles strongly depend on the duration of a certain phase. Three extensive quantities must be measured to obtain an equation of state for a static gas with no conserved charge. In heavy-ion collisions, the total energy and entropy can always be estimated from experiment (see Sec. V.C.1). If we know the volume and the fraction of collective energy at a given time of the collision, we have a single value in an equation of state corresponding to the conditions at that time. The correlation function measured in pion interferometry could give an estimate for the size at freezeout. The main reason that we are reporting on this experimental tool is the possibility of measuring the prolongation of lifetime due to a phase transition.

Let us see how we can infer the spacetime structure of the source from correlations in momentum space. The correlations refer to two-particle correlations in the pion's momenta distributions. They are defined as

$$C(\mathbf{p}, \mathbf{q}) = P(\mathbf{p}, \mathbf{q}) / [P(\mathbf{p})P(\mathbf{q})], \quad (5.73)$$

where  $P(\mathbf{p}, \mathbf{q})$  denotes the probability of finding two pions with three-momenta  $\mathbf{q}$  and  $\mathbf{p}$  in the same event. In terms of rates,  $C(\mathbf{p}, \mathbf{q})$  is given as

$$C(\mathbf{p}, \mathbf{q}) = \frac{d^6 N}{d^3 p d^3 q} \bigg/ \left( \frac{d^3 N}{d^3 p} \frac{d^3 N}{d^3 q} \right). \quad (5.74)$$

Thus  $C$  is measurable as a ratio of two-pion to one-pion inclusive yields. Its width is a measure of the inverse source size. The weaker the correlation, the smaller its width, and the larger is the source that emits the pions.

Two-pion correlations can arise from different origins. The only ones that we will discuss in detail are Bose-Einstein correlations due to the quantum mechanical symmetrization of the outgoing wave functions of identical bosons. In the same way, electrons in a metal are anticorrelated in their spatial distribution due to the fact that they occupy only a finite volume in momentum space.

Other causes of correlations due to final-state interactions are hadronic and Coulomb interactions. Coulomb interactions lead to positive correlations between particles of opposite charge if the relative momentum is small. It has yet to be seen to what extent their contribution is averaged out due to a comparable number of pairs with equal and opposite charges.

A third source for correlations comes from resonance decays. Pions are detected that are not directly emitted from the hadronic gas at freeze-out, but which result from decays of heavier mesons. Such pions may con-

taminate the information regarding the source size.

Bose-Einstein correlations are the only ones that contribute when radiating source is completely chaotic. If we follow the evolution of the plasma through the phase transition to the final state of a hadron gas, it consists of up to  $\sim 90\%$  pions in the baryon-free region at freeze-out, when the pion gas is already rather dilute. [Present experiments deal with equal nucleon and pion numbers at the AGS accelerator, Brookhaven, and at the SPS accelerator, CERN; the ratios used are 1:7 for S+S and 1:5 for S+Ag. At this stage, the baryon-free region is still an idealization.] The assumption of an incoherent source for a dilute pion gas seems to be a good approximation to start with.

We shall therefore focus on the case of a chaotic pion source. The probability  $P(\mathbf{p}, \mathbf{q})$  of measuring two pions with momenta  $\mathbf{p}$  and  $\mathbf{q}$  in the same event is given as (Pratt, 1984)

$$P(\mathbf{p}, \mathbf{q}) = P(\mathbf{p})P(\mathbf{q}) + \int d^4 x d^4 y g(\mathbf{K}, x)g(\mathbf{K}, y) \exp\{ik(x-y)\}, \quad (5.75a)$$

where

$$P(\mathbf{p}) = \int d^4 x g(\mathbf{p}, x), \quad (5.75b)$$

and where

$$\mathbf{K} = (\mathbf{p} + \mathbf{q})/2 \quad (5.75c)$$

is the average momentum of the pion pair, and

$$k = p - q \quad (5.75d)$$

denotes the relative four-momentum. The function  $g(\mathbf{K}, x)$  gives the probability of emitting a pion of momentum  $\mathbf{K}$  from a spacetime point  $x$ ;  $g(\mathbf{K}, x)$  is called the emission function. Thus in the simplest case, where the source is incoherent and only Bose-Einstein correlations relate the wave functions of emitted particles, the correlation function essentially measures the expectation value of  $\cos(kr)$  weighted by the product of emission functions ( $r = x - y$  is the relative distance of the pions at their source). In other words, the measurable correlation in momentum space is the Fourier transform of the emission functions. (A general ansatz for the two-particle correlation function associated with final-state interactions in rapidity space can be written in a similar form.) The probability  $P(\mathbf{p}, \mathbf{q})$  depends on six degrees of freedom. Under a further constraint on the spacetime evolution of the source or on its symmetry properties, it may be possible to determine the emission function  $g$  from a measurement of  $P$ ;  $g$  depends originally on seven degrees of freedom. The final spacetime distribution described by the product of emission functions is a result of the spacetime evolution of the plasma through a possible phase transition to the hadronic phase. The spatial distributions of the pions at the moment of freezeout constitute the radiating pion source in which we are interested. Hence the ansatz for  $g(\mathbf{K}, x)$  depends on two

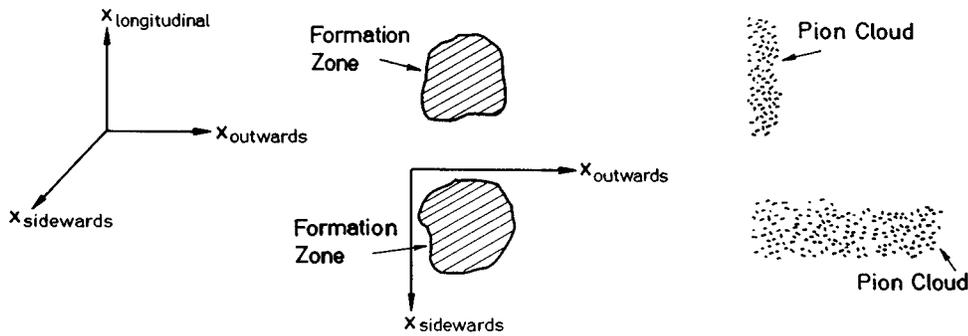


FIG. 34. Illustration of pion formation zones (shaded area) emitting pions in the sideward and outward directions; upper part, pion clouds (dotted area) without a phase transition; lower part, with a transition. From Bertsch (1989).

ingredients: the *equation of state* and the *dynamical concept* relating the spacetime dependence to thermodynamic quantities.

Let us recall why a phase transition is expected to delay the breakup. The moment of freezeout or breakup  $\tau_B$  is determined by the degree of dilution. Equivalently, it is given by the value of its entropy density  $s(\tau_B)$ . For fixed initial energy density and equilibration time  $\tau_0$ , the initial entropy density for a plasma is considerably higher than that for a hot pion gas to start with. The total entropy is assumed to be approximately conserved during the expansion. Thus it takes longer to reach the same low entropy density at freezeout if one starts in the plasma phase. A phase transition must have occurred at an intermediate stage.

In Fig. 34 (from Bertsch, 1989) we visualize how the prolonged lifetime is expected to influence the effective spatial dimension of the source. The beam axis is chosen in the  $z$  direction and labeled  $x_{\text{Longitudinal}}$ . The two transverse directions are  $x_{\text{sideways}}$  and  $x_{\text{outwards}}$ , which points towards the detector. This coincides with the direction of the main pion stream only if the pions are viewed in a frame in which they are emitted perpendicular to the beam. The pion cloud has the shape of a cylinder. Fig. 34 shows a snapshot of the pion sources at freezeout for both scenarios: a plasma at the beginning or a hot and dense pion gas. In the latter case, the high density of pions in the formation zone leads to a rescattering. Thus the transverse dimensions at freezeout would clearly exceed the size of the formation zone if the formation zone were at rest. The longitudinal expansion of the source modifies the picture in the following way. The collective flow of the pions singles out a preferred direction. Recall that in interferometry measurements only pions with almost the same momentum are detected. For a moving source, the probability of detecting such pairs in a certain direction is reduced. Only pions which come from the same side of the cylinder as the detector have a chance of arriving at the detector with nearly the same momenta. This reduces the transverse dimensions of the pion cloud considerably in spite of the rescattering effects (Pratt, 1984; Bertsch, 1989).

A *first-order phase transition* that proceeds via nucleation of the plasma to the hadronic phase will alter the picture. The plasma lives only a very short time, so its transverse expansion may be neglected. In the intermediate phase, a possible scenario is that the plasma will

break up into droplets of hadrons. The conversion to the hadronic phase is slow, because the rearrangement of degrees of freedom takes time. The pions are emitted over a longer period at a smaller rate, hence they have less chance to rescatter. The result is a *small sideward* and a *large outward dimension* of the cloud at the moment of last interaction. This is illustrated in Fig. 34. The hope of identifying a (strong) first-order phase transition via pion interferometry is essentially based on this scheme: to measure a large source size in the outward direction at freezeout.

Pratt (1986) derived a functional dependence of the emission function  $g$  for two equations of state (the bag model equation of state to simulate a phase transition and an ultrarelativistic pion gas without a phase transition). As a dynamical concept he used the hydrodynamical equations for two geometries, a spherically symmetric exploding plasma and Bjorken's scaling solution. For Bjorken's scenario, an analytic expression can be obtained, and for the spherically exploding plasma the hydrodynamic equations were solved for different initial conditions. The change in the dynamics caused by a phase transition is more pronounced the longer the mixed phase lasts—the pressure stays constant while the energy is absorbed in latent heat. Qualitatively, the following features can be seen in a plot of  $C(K, k)$  as a function of  $k$ , the relative momentum. Higher average momentum  $K$  leads to a broader correlation, corresponding to a smaller effective source size. This reflects the collective expansion of the system. One effect of a phase transition is a reduction of the explosive velocity. The other effect is only apparent in  $C$  when the relative momentum is chosen parallel to the average momentum. For larger values of  $K$ , the correlations drop to smaller values, indicating a larger size or a longer lifetime of the source. The strengths of the signals depend, however, on the initial conditions chosen for the spherically exploding fireball. They may fade away under realistic conditions. In order to get information on a transient phase transition in a cylindrical geometry, the orientation of the relative momentum  $k$  with regard to the average momentum  $K$  has to be chosen in an appropriate way. The large outward dimension shows up only in correlations with  $\mathbf{k}$  parallel to  $\mathbf{K}_{\text{per}}$  ( $\mathbf{K}_{\text{per}}$  being the transverse component of the average momentum).

The difference between sideward and outward dimensions in a cylindrical geometry has been numerically in-

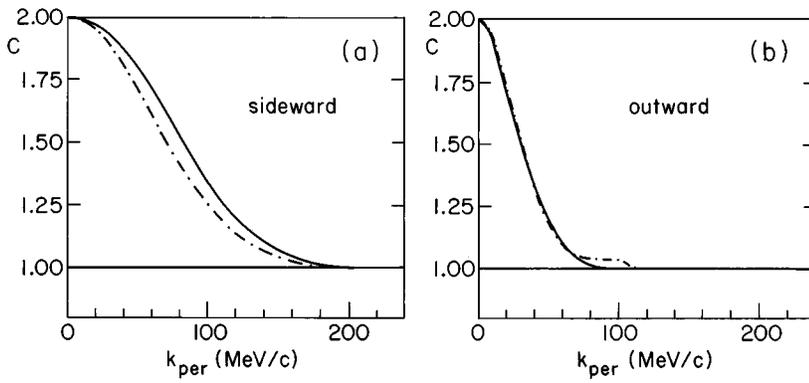


FIG. 35. Momentum correlations: (a) sideward transverse momentum correlation between pions with  $k_{\text{perp}0}=0$  and  $k_{\text{longitudinal}}=0$ ; (b) outward transverse momentum correlation between pions, with  $k_{\text{perp}s}=0$  and  $k_{\text{longitudinal}}=0$ . Solid lines are the cascade results, dashed lines correspond to a semianalytic parametrization. From Bertsch, Gong, and Tohyama, 1988.

investigated by Bertsch, Gong, and Tohyama (1988). This calculation is based on a cascade description of the hadronic phase, especially in the period close to freezeout, when the system is too diluted to justify a hydrodynamic approach. On the other hand it turns out that a hydrodynamic expansion in the sense of a collective transverse and longitudinal expansion velocity may be reinstated from a more microscopic starting point, if realistic cascade descriptions are used (Stachel, 1993). In the cascade approach of Bertsch, Gong, McLerran *et al.* (1988) individual pions are followed along their classical trajectories. The plasma is assumed to break up into “blobs” that emit and reabsorb pions at a certain rate. Entropy conservation is no longer assumed, but entropy is produced during the phase conversion. The freezeout time is also predicted in this approach. The breakup is described as occurring instantaneously, i.e. without supercooling delay.

Again one has to specify an ansatz for the emission function. The classical treatment of pion motions leads to an emission function which is too singular to display interference effects. In order to implement interference effects, the momentum dependence of the source function is smoothed. Bertsch, Gong, and Tohyama (1988) devised the following parametrization:

$$g(x, \mathbf{K}) = f(\mathbf{K}_{\text{per}}) \tau e^{-(\tau/\tau_0)^2} e^{-(r_{\text{per}}/R)^2} e^{-(Y_r - Y_{\mathbf{K}})^2/Y_0^2}, \quad (5.76)$$

The spacetime coordinates are  $x$ , the proper time  $\tau = \sqrt{t^2 - z^2}$ ,  $r_{\text{per}}$  denoting the transverse position, and  $Y_r \equiv \tanh^{-1}(z/t)$  is the rapidity. The average three-momentum  $\mathbf{K}$  is written in terms of the transverse components  $\mathbf{K}_{\text{per}}$  and the momentum rapidity  $Y_{\mathbf{K}} \equiv \tanh^{-1}(\mathbf{K}_z/E)$ . The amplitude  $f$  only depends on  $\mathbf{K}_{\text{per}}$ . The values for  $\tau_0$ ,  $R$ , and  $Y_0$  are chosen as  $\tau_0 = 9$  fm/c,  $R = 3.3$  fm, and  $Y_0 = 0.76$ . Arguments for this choice can be found in the original paper. The result for the correlation  $C(\mathbf{p}, \mathbf{q})$  of finding a pion pair with momenta  $\mathbf{p}$  and  $\mathbf{q}$  is obtained from Eq. (5.76) as an integral over the transverse positions. The integral remaining is given as

$$C(\mathbf{p}, \mathbf{q}) = 1 + e^{-(k_{\text{per}}R)^2/2} \int d\tau d\tau' dY dY' \frac{4\tau\tau'}{\tau_0^4 \pi Y_0^2} \times \exp\left\{-\frac{\tau^2 + \tau'^2}{\tau_0^2}\right\} \exp\{-[(Y_{\mathbf{K}} - Y)^2 + (Y_{\mathbf{K}} - Y')^2]/Y_0^2\} \cdot \cos[(\mathbf{p}_z - \mathbf{q}_z)(z - z') - (E_{\mathbf{p}} - E_{\mathbf{q}})(t - t')]. \quad (5.77)$$

Here  $k_{\text{per}} \equiv |(\mathbf{p} - \mathbf{q})_{\text{per}}|$  has a sideward and an outward component. The result is a Gaussian in the sideward direction;  $C$  measures the transverse size of the source, which is the same in sideward and outward directions as long as the time dependence of the distribution is neglected. If the transverse relative momentum is chosen in the outward direction (i.e.,  $\mathbf{k}_{\text{perp}s} = 0$ ,  $\mathbf{k}_{\perp 0} \neq 0$ ), the Gaussian in  $\mathbf{k}_{\text{perp}0}$  is modified by the  $\mathbf{k}_{\text{perp}0}$  dependence of  $E_{\mathbf{p}}$  and  $E_{\mathbf{q}}$ , and the effective size in the outward direction is sensitive to the time correlation as is seen in Eq. (5.77). The numerical results for sideward and outward correlations are displayed in Figs. 35(a) and 35(b).

The reduced correlation for outward (compared to sideward) relative momenta reflects the effectively larger size of the source in this direction, as was anticipated above (see Fig. 34).

Expression (5.77) for the pion correlation partly motivates an ansatz which is sometimes used in real interferometry experiments (see, for example, Schmidt and Schukraft, 1993). This is a threefold Gaussian in  $\mathbf{k}_{\text{perp}0}$ ,  $\mathbf{k}_{\text{perp}s}$ , and  $\mathbf{k}_L$  (the longitudinal component of  $\mathbf{k}$  instead of the rapidity  $Y_{\mathbf{K}}$ ). The possibility of having different length scales  $R_{T_{\text{out}}}$ ,  $R_{T_{\text{side}}}$ , and  $R_{T_L}$  in all three directions is realized [see Eq. (5.78) below].

So far, we have reported on pion interferometry as a tool for estimating the *lifetime* of a pion radiating source. Naturally, the question arises whether it is also possible to resolve spatial inhomogeneities of bubble structures in the mixed phase. This question has been recently addressed by Wieand *et al.* (1992). The answer is, in principle, affirmative. Pion interferometry may be a tool for resolving the grain structure of the hadronic phase. Clearly, the bubble structure should not be too fine grained, or it will not be resolvable for the detectors. In practice this does not seem to be a realistic pos-

sibility. The theoretical predictions have not yet reached a stage at which it is clear what should be verified experimentally.

In reviewing the experimental status of interferometry measurements, we limit ourselves to a few aspects of applications in heavy-ion experiments. The primary observables are the double-pion and single-pion inclusive distribution functions  $N(p_1, p_2)$  and  $N(p)$  that determine the correlation  $C$  according to Eq. (5.73). In Bjorken's geometry,  $C$  depends on five independent variables, the transverse and longitudinal components of the average momenta  $K_{\text{transv}}$ ,  $K_L$  or the rapidity  $Y$ , and the sideward, outward and longitudinal components of the relative momenta ( $k_{\text{per } s}$ ,  $k_{\text{per } 0}$ ,  $k_L$ , respectively). An ansatz which is frequently used by experimentalists is a simple Gaussian in the relative momentum of the pion pair. A comparison to the analytic expression, Eq. (5.77), may indicate the gap between theory and experiment. The parametrization used in the NA35 experiment at the CERN SPS accelerator (Baechler *et al.*, 1992) is a Gaussian in all three components of the relative momentum:

$$C(\mathbf{k}) = 1 + \lambda \exp\{-k_{\text{per } s}^2 R_{\text{per } s}^2 / 2\} \times \exp\{-k_{\text{per } 0}^2 R_{\text{per } 0}^2 / 2\} \exp\{-k_L^2 R_L^2 / 2\}, \quad (5.78)$$

where  $k_{\text{per } s}$  is orthogonal to  $K_{\text{transv}}$  and  $K_L$ ,  $k_{\text{per } 0}$  is parallel to  $K_{\text{transv}}$  and orthogonal to  $K_L$ , and  $k_L$  is orthogonal to  $K_{\text{transv}}$  and parallel to  $K_L$ . Furthermore,  $\lambda$  is the coherence, or *chaoticity parameter*, which is assumed to absorb the Coulomb corrections of final-state interactions. It is unity for an incoherent boson source. The parameters  $R_{\text{per } s}$ ,  $R_{\text{per } 0}$ , and  $R_L$  are the observables that are determined from a fit. Note that the simple Gaussian ansatz of Eq. (5.78) describes only spatial correlations. The time correlation is more complicated if a phase transition occurs. The parameters  $R$  include some information on the spacetime evolution of the system (Ferenc *et al.*, 1992). The data for  $R_L$  (fm) as a function of rapidity  $Y$  as seen by an observer at fixed  $Y_{\text{obs}} = 2.5$  are consistent with a longitudinally expanding source. This result is not trivial. Predictions for a static source are quite different (Makhlin and Sinyukov, 1988). The longitudinal radius  $R_L$  can be related to the freezeout or decoupling time  $\tau_f$ . The relation between  $R_L$  and  $\tau_f$  for a scaling expansion is given as (Makhlin and Sinyukov, 1988)

$$R_L = (2T_f / \langle m_T \rangle)^{1/2} \tau_f \quad (5.79)$$

in a comoving frame, where  $T_f$  is the freezeout temperature and  $\langle m_T \rangle$  the average transverse pion mass. From the measured values of  $R_L$ ,  $\tau_f$  is estimated to be in the range  $4.5 < \tau_f < 6.3$  fm/c for  $100 < T_f < 200$  MeV (Ferenc *et al.*, 1992). The difference between the outward and sideward extensions ( $R_{\text{per } 0} - R_{\text{per } s}$ ) should be a measure of the duration of pion emission, as was argued above. The source dimension in the outward direction has two contributions (see Fig. 34): the effective geometric depth (which is comparable to the sideward extension and can be larger than the geometrical size of the

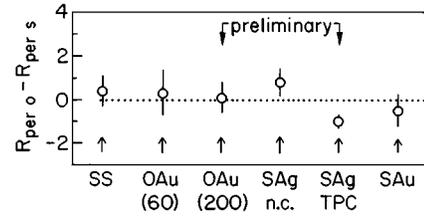


FIG. 36.  $R_{\text{per } 0} - R_{\text{per } s}$  [fm] for the indicated collisions. From Ferenc *et al.* (1992).

formation zone, due to rescattering effects) and the effective prolongation due to a reduced particle emission rate in the case of a phase transition. In Fig. 36 from Ferenc *et al.* (1992), it is seen that the difference in the transverse dimensions is negligible.

For a specification of the rapidity intervals, see Ferenc (1992). A more careful estimate for the duration of particle emission is  $\Delta\tau_f \leq 2$  fm/c. This rules out a long-lived source ( $\geq 10$  fm/c) as a result of a first-order phase transition with supercooling delay. Finally, the transverse size at decoupling time  $R_{\tau_f}$  turns out to be 1.5–2 times larger than the geometric size if the geometric size is estimated as the transverse size of the projected transverse density distribution. This is compatible with a high-density formation zone, from which the pions first rescatter before they freeze out. The freezeout volume becomes independent of the formation volume in this case; it merely depends on the particle density.

The statistics of the data still have to be improved in order to check the reliability of the projections of  $R_{\text{per}}$  on  $R_{\text{per } s}$  and  $R_{\text{per } 0}$ . One should also keep in mind that the results for the different length scales could be an artifact of the simplified ansatz (5.78) for  $C$ . It is unlikely that just one parameter  $\lambda$  can absorb the full complexity of different origins of pion correlations. Another complication may arise from rescattering from the surrounding target spectator matter (Schmidt and Schukraft, 1993), a “dirt effect” that we have not mentioned so far. The relative weights of the contributions to pion correlations of different origins depend on the impact parameter. One would like to ignore these effects, unfortunately one is not allowed to do so.

For comparison, we also mention results of pion interferometry measurements with the AGS-machine at Brookhaven National Laboratory (Stachel, 1994; Xu, 1994). The data were taken measured for central 14.6 GeV/n Si+Al- and Si+Pb collisions. At freezeout the results were as follows. In the center-of-mass frame, the pions are emitted from a source with transverse radius  $R_{\text{per}} = 6.7$  fm and longitudinal radius  $R_L = 5.0$  fm (on average). The source is nearly spherically symmetric. If this transverse radius is compared to the initial transverse size of the system  $R_{\text{per}}(\text{Si}) = 2.9$  fm, one finds a transverse expansion of the system by a factor of 2.3. Assuming an expansion velocity of  $v/c = 0.3$  (0.2), the time scale of the expansion follows as 10 (15) fm/c. The uncertainty in the expansion velocity is due to resonance decays. The pion interferometry results lead to a freezeout volume of  $V = 2400$  fm<sup>3</sup>. A sort of consistency

check is provided by the estimate for the volume, if freezeout is defined to occur when the average distance from a given nucleon to the nearest pion is  $d = \sqrt{\sigma_{\pi n}/\pi}$ , where  $\sigma_{\pi n}$  is the pion-nucleon cross section. For a freezeout temperature of 140 MeV and  $\sigma_{\pi n} = 62$  mb, one finds  $V = 2750$  fm<sup>3</sup> at freezeout, in reasonable agreement with the interferometry estimate.

In these measurements, the estimates for the radii are not obtained from a direct fit of a certain functional form to the data, but by use of a dynamical model [realized in the event generator RQMD—relativistic quantum molecular dynamics (Sorge *et al.*, 1989)] with known spacetime characteristics of the source. Two-particle correlations are constructed from the RQMD-generated single-particle distributions in such a way that the Bose-Einstein effect is imposed (Xu, 1994). The realistic experimental parameters are imposed as input data in the RQMD code.

Although there is no indication of a first-order transition in these experiments, the important conclusion is that the hadron gas does have enough time to be in thermal (and chemical) equilibrium at freezeout. This supports one of the basic underlying assumptions in applying thermodynamic concepts.

Signals of a strong first-order transition should not be missed by pion interferometry. An extension of lifetime by an order of magnitude should be visible for a certain projection of the pair's relative momentum. Nonappearance of such signals is less conclusive. Specific signals for a second-order transition with a large correlation length have not yet been invented. Experimental groups that are at present dealing with pion, kaon and rho interferometry are the E802 collaboration (Abbot *et al.*, 1992; Akiba *et al.*, 1993), the E814 collaboration (see Xu, 1994; Stachel, 1994) at Brookhaven, and the NA35 collaboration (Seyboth, 1992; Ferenc, 1992), the NA44 collaboration (Sarabura *et al.*, 1992), and the WA80 collaboration (Peitzmann *et al.*, 1992, 1993a, 1993b) at CERN. A chance for future high-statistics experiments may lie in kaon interferometry. The yield of kaons is considerably lower than that of pions, but kaons are less plagued by contamination due to resonance decays (Ferenc, 1992).

## 5. Multiplicity fluctuations

The possible signatures of a first-order transition that we discussed in the preceding sections are based on a prolonged duration of the matter evolution. Multiplicity fluctuations rely on additional entropy production, which might be generated in explosive processes during the conversion of the plasma to hadronic matter. Several candidates for hadronization processes have been under discussion: nucleation in the bulk, “boiling by cooling” (van Hove, 1983), or an adiabatic procedure, in which the system stays in local thermal equilibrium, and pressure and temperature remain constant during the conversion. These scenarios do not lead to strong multiplicity fluctuations.

In this section, we describe alternatives to the adiabatic scenarios. These are hadronization processes via

*deflagrations* or *detonations*. They are compatible with a strong first-order phase transition and a liberation of a large amount of latent heat, supercooling and/or superheating effects during the transition. One of the proposed physical pictures is the following (van Hove, 1985). Energy and entropy densities change drastically over a small temperature and pressure interval in the vicinity of  $T_c$ . Therefore, the proposed hadronization mechanism has to provide a fast liberation of the large entropy and energy of the plasma, and a large energy has to flow into the hadron phase in an efficient way. The main part of hadronization occurs at the phase transition. Hadronization via evaporation of hadronic bubbles before the transition is negligible in this picture. Close to the transition temperature, the color fields (induced by the color charges of the plasma) collapse to flux tubes creating something like a network of strings (see Patel, 1984a and 1984b, and Sec. IV.B.1). Due to string formation, stretching, and breaking, plasma droplets will be stopped in their longitudinal expansion or will break up in a few more droplets. In the next stage, the main part of the hadronization is assumed to proceed via ejection of hadrons through the surface of the plasma blobs with the kinematics and thermodynamics of a deflagration. Deflagration is said to occur only at the outer surface, and not in the interior (in the “bulk”) of the plasma droplets. The plasma blobs themselves are small in size (a few fermi at most).

Let us see which constraints from energy-momentum and positive entropy production are obtained for deflagration or detonation scenarios. Again we use a hydrodynamic description and follow Cleymans *et al.* (1986) to sketch the derivation of kinematic constraints. Deflagrations start at the surface of the plasma, detonations in its interior, to form hadronic bubbles. In the case of deflagrations, the front separating both phases moves more slowly inwards than the hadrons escape outwards. In the rest frame of quark matter, the hadrons and the front move in opposite directions. In detonations, the velocities of the front and the hadrons are both directed outwards, but at different speeds. Usually, the speed of the hadrons is assumed to be lower than the speed of the front. When the speed of the hadrons is higher, the relative velocity of the hadrons with respect to the front is negative. This explosion is called an *eruption* and is very unlikely, as we shall see.

Let us go to the rest frame of the front. Pictorially, in this frame the front “eats up” the plasma that wants to pass it. In the case of deflagrations the hadron velocity exceeds the velocity of quark matter, leading to

$$\tanh\theta_h > \tanh\theta_q. \quad (5.80)$$

Here  $\theta_h$  and  $\theta_q$  are the fluid rapidities in the hadron and quark phases, respectively. In usual detonations the quark matter is “eaten up” faster than the hadronic bubbles are formed ( $\theta_h < \theta_q$ ); in eruptions quark- and hadronic-matter velocities have opposite sign in the rest frame of the front [see Fig. 37(a)–37(c)].

Thus the difference between deflagrations and (usual) detonations lies in the relative magnitude of their veloci-

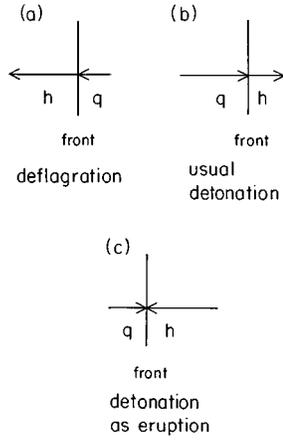


FIG. 37. Velocities of quark ( $q$ ) and hadronic ( $h$ ) matter for three possible scenarios viewed in the rest frame of the front: (a) deflagration; (b) usual detonation; (c) detonation as eruption.

ties in the rest frame of the separating front. To derive constraints for pressure, energy, and entropy densities, we go to the rest frame of the front, in which the flow of quark and hadronic matter is steady. Derivatives with respect to proper time  $\partial/\partial\tau$  vanish. Furthermore, we restrict the interface separating both phases to a (1+1)-dimensional planar interface. The interface is chosen to be infinitesimally thin. Energy-momentum conservation  $\partial_\mu T^{\mu\nu}=0$  across the front leads to

$$\begin{aligned} T^{zz} &= \text{const}, \\ T^{z0} &= \text{const} \end{aligned} \quad (5.81)$$

in  $(t, z)$  coordinates. The energy-momentum tensor is taken to be that of a perfect fluid (apart from the place of discontinuity). The velocity is parametrized in terms of the fluid rapidity  $\theta$  as usual, according to  $u_0 = \cosh \theta$  and  $u_z = \sinh \theta$ . Continuity of  $T^{zz}$  and  $T^{z0}$  across the front leads to

$$\begin{aligned} (\varepsilon_h + p_h) \sinh^2 \theta_h + p &= (\varepsilon_q + p_q) \sinh^2 \theta_q + p_q, \\ (\varepsilon_h + p_h) \sin \theta_h \cos \theta_h &= (\varepsilon_q + p_q) \sin \theta_q \cos \theta_q \end{aligned} \quad (5.82)$$

for a perfect fluid. The subscripts on the energy density  $\varepsilon$  and pressure  $p$  indicate hadronic and quark matter. Solving for  $\tanh^2 \theta_{h,q}$  and implementing the inequalities for  $\theta_h$  and  $\theta_q$ , Eq. (5.82) gives

$$\varepsilon_q - p_q > \varepsilon_h - p_h \text{ for deflagrations} \quad (5.83)$$

and

$$\varepsilon_q - p_q < \varepsilon_h - p_h \text{ for detonations.}$$

Imposing the bag model equation of state (Sec. V.B.2) with  $c_s = 1/\sqrt{3}$ , this means that

$$\begin{aligned} \varepsilon_q + 2B &> \varepsilon_h \text{ for deflagrations, and} \\ \varepsilon_q + 2B &< \varepsilon_h \text{ for detonations.} \end{aligned} \quad (5.84)$$

The inequality in Eq. (5.84) for detonations shows that a strong supercooling is necessary in order for  $\varepsilon_q$  to be small enough to satisfy Eq. (5.84). (The gap between the

energy densities of both phases may be unrealistically large in the bag model, but it is obvious that some supercooling is necessary.) The next constraint on the allowed values of  $\varepsilon$  and  $p$  comes from the second law of thermodynamics. The entropy flux into the front must not be larger than that out of the front. The condition of entropy increase across the discontinuity leads to  $s_h^2/s_q^2 \geq (\varepsilon_h + p_h)(\varepsilon_h + p_q)/[(\varepsilon_q + p_q)(\varepsilon_q + p_h)]$ . For a perfect fluid, where dissipation effects are absent, entropy is strictly conserved inside each phase [ $\partial_\mu(su^\mu)=0$ ]. It is also approximately conserved across the front (equality sign above), if the discontinuity in  $\varepsilon$  is infinitesimally small (van Hove, 1983).

In general, however, deflagrations and detonations are irreversible processes and do produce entropy. The equation for  $s_h^2/s_q^2$  can be exclusively expressed in terms of  $\varepsilon_h$  and  $\varepsilon_p$  by using the thermodynamic relation  $Ts = \varepsilon + p$  and the equation of state. Finally, the velocities must be smaller than the velocity of light. This means

$$0 \leq \tanh^2 \theta_{h,q} \leq 1. \quad (5.85)$$

The combined constraints for  $s_h^2/s_q^2$  and condition (5.85) limit the values for  $\varepsilon_q$  and  $\varepsilon_h$  that are compatible with detonations or deflagrations to a small window. Thus reliable results for  $\varepsilon_q$  and  $\varepsilon_h$  and the equation of state could lead to an exclusion of these scenarios.

Deflagration and detonation solutions of the hydrodynamic equations have been constructed in a numerical analysis by Gyulassy *et al.* (1984). The plasma expansion is neglected in the initial state, and the baryon number is set to zero. The constants from energy conservation and entropy increase are evaluated as indicated above, and the equation of state is taken from the bag model. The results are as follows.

Detonations and deflagrations use the latent heat in a different way. Detonations need strong supercooling and subsequent strong superheating of the hadron phase. Deflagrations are possible under less extreme conditions. They are compatible with mild supercooling. The calculations have been extended from discontinuities along a single surface to several surfaces, which means bubble deflagration rather than surface deflagration. A surface deflagration shock was considered by van Hove (1985). It seemed to move too slowly into the plasma to provide a realistic possibility of plasma conversion to the hadron phase.

Entropy-producing processes can lead to enhanced rapidity-density fluctuations and to peaks in  $dN/dY$  of the final-state hadrons. These peaks may be isolated or overlapping, depending on the resolution  $\delta Y$  and the size of a typical scale  $\Delta Y$  over which  $N$  fluctuates. Large transverse momenta  $\langle p_T \rangle$  are another possible effect. Multiplicity fluctuations are sometimes called seismometers (Gyulassy, 1984). Experimental hints for ‘‘crests’’ of plasma bubbles are taken from cosmic-ray events (Iwai *et al.*, 1976 and 1982). The observations are at least compatible with an explanation of deflagration processes in QCD plasma conversions (Gyulassy *et al.*, 1984; van Hove, 1985).

Let us assume that the QCD transitions are truly sharp crossover phenomena. The question arises whether there is still a chance for observing deflagrations. The answer seems to be in the affirmative, if we look at the derivation of deflagrations, for example, in van Hove (1985). The essential ingredients are a large variation of  $\varepsilon$  and  $s$  over a small interval of  $T$  and  $p$ . One ingredient is the small value of  $p$  compared to  $Ts = \varepsilon + p$ . For a crossover phenomenon, the transition interval is defined as the region of rapid variation of  $dp/dT$ . Van Hove (1985) studied the implications for deflagrations in both cases, a first- and a second-order transition. The results were quite similar. Deflagrations may be realized in heavy-ion collisions in spite of a smooth crossover only if it is sharp enough. This result is not surprising. Nucleus-nucleus collisions are performed in a finite spacetime volume. From numerical simulations we know that a strictly first-order transition and a sharp crossover phenomenon are practically indistinguishable in a finite volume, if the measurement is performed for a single size. Similarly, deflagrations may be practically the same for first-order transitions and crossover phenomena in a finite volume. The finite volume constitutes part of the physical boundary conditions in heavy-ion collisions. Experimental results will not expose finite-volume roundings as artifacts of approximations to a hypothetical infinite volume world.

## 6. Intermittency analysis

Experimental manifestations of the QCD transitions are frequently discussed under the assumption that they are of first order. As we have seen in Secs. III and IV, there are indications that the finite-temperature transition resembles more closely a crossover phenomenon. The realistic set of quark masses may be close to a “critical” set in mass parameter space, which could lead to a second-order transition. A standard physical picture of critical phenomena in spin systems is a diverging correlation between spins, leading to clusters of aligned spins of arbitrary length scales. The  $3d$   $Z(2)$  spin system is supposed to share the universality class of an  $SU(2)$  pure gauge theory; an  $O(4)$  ferromagnet is assumed to share the class of two-flavor QCD.

Thus the question arises how a picture of clusters of aligned spins in color or isospin space can be transferred to typical observables in heavy-ion collisions, such as particle multiplicities. One of the possible answers lies in the *intermittent behavior* in rapidity distributions. (For a particular realization of a second-order transition in the case of a hot pion gas, see also Sec. V.D.3 below.)

The concept of intermittency was originally introduced in studies of turbulent behavior of fluids (Zel'dovich *et al.*, 1987). In general, intermittency can be defined as the appearance of structure in random media. In the context of particle physics, it has to do with large fluctuations of charged-particle density in small regions of phase space. In heavy-ion collisions, intermittency refers to certain moments of rapidity distributions. Other

distributions, such as energy, pseudorapidity, and azimuthal angles, can be considered as well (Ochs and Wosiek, 1988).

Let us first consider a toy model that is universal enough for applications in different areas. Consider a set of  $N$  balls distributed in a box of total size  $R$  into  $M$  cells of size  $L$  ( $M=R/L$ ). In the case of rapidity distributions, they correspond to  $N$  particles per given unit of some kinematic range. A cell in a box is a rapidity interval in a certain available rapidity range. For the Ising model, the balls are the spins on a spacetime lattice  $R$  subdivided into cells of size  $L$ . We are interested in the ball distributions if the resolution of the lattice is made finer and finer ( $M \rightarrow \infty$ ). There are many ways in which to realize this limit. One possibility is to keep  $R$  constant and to let  $L$  go to zero. This limit is of interest for rapidity distributions, where  $L$  stands for the bin size  $\delta Y$  of the rapidity interval. Another realization is the thermodynamic limit, with  $R \rightarrow \infty$  for fixed  $L$ . This limit is usually taken when critical phenomena are discussed in the infinite volume limit.

The same type of limit is of interest in a second-order finite temperature transition in spacetime continuum, which is described by a model of lattice gauge theory. One may expect that intermittency in the limit of  $R \rightarrow \infty$ , with  $L$  fixed, has a correspondence in the limit of  $R$  fixed,  $L \rightarrow 0$ .

We define the  $\ell$ th normalized moment  $f_\ell$  for a given distribution of  $N$  balls as

$$f_\ell(M) = \left[ \frac{1}{M} \sum_{m=1}^M K_m^\ell \right] / \left[ \left( \frac{1}{M} \sum_{m=1}^M K_m \right)^\ell \right]. \quad (5.86)$$

Here  $K_m$  denotes the number of balls in the  $m$ th cell. First, we keep  $N$  fixed and vary  $M$ . In the extreme case of an equidistribution of  $N/M$  balls in each box, it is easily seen that  $f_\ell(M)$  is independent of the grain size, such that

$$f_\ell(M) = 1 \quad \forall \ell. \quad (5.87)$$

In the other extreme, a strong fluctuation (where all balls are concentrated in one box), we find a logarithmic dependence of  $\ln f_\ell$  on the “resolution size”  $L$ :

$$\ln f_\ell(M) = -(\ell-1) \ln L + (\ell-1) \ln R. \quad (5.88)$$

Such behavior of a given distribution with  $N$  balls is called *intermittent*.

More generally, intermittent behavior is attributed to average values of  $f_\ell$ . The weighted average  $\langle \dots \rangle$  of the  $\ell$ th moments runs over an ensemble of configurations in a  $d$ -dimensional volume  $R^d$  divided into  $M=(R/L)^d$  cells of equal size  $L^d$ . In a multiplicity measurement it could be an average over all events; in the Ising model it is the thermodynamic average with Boltzmann weights.

Intermittency is called the property of fluctuations around some average distribution that lead to a power law behavior of  $\langle f_\ell \rangle$  in the number of cells  $M$ , or equivalently if

$$\ln \langle f_\ell(L) \rangle = -\lambda_\ell \ln L + g_\ell(R). \quad (5.89)$$

Here  $\lambda_\ell > 0$  are constant and  $g_\ell(R)$  is independent of  $L$ ,  $L$  being the linear cell size. The constants  $\lambda_\ell$  are called intermittency indices, and are a measure of the strength of intermittent behavior.

Results for  $\lambda_\ell$  and  $g_\ell$  in case of a d-dimensional Ising model can be found in Satz (1989). The essential point at which criticality enters the derivation of intermittency in an Ising model is scale invariance at  $T_c$ . While the balls of the toy model are the same independently of the partition of the lattice, one has first to argue why one may choose the same type of variables on all length scales in an Ising model. This is justified by self-similarity as  $T \rightarrow T_c$ , if  $1 \ll L \ll \xi$  is satisfied,  $\xi$  being the correlation length. For a given  $\ell$ , the intermittency indices can be expressed in terms of the more familiar critical exponents. Like critical exponents, they are universal for all models belonging to the same universality class.

The mere divergence of a correlation length  $\xi$  is not sufficient for introducing intermittent behavior. In a 1d Ising model, where  $\xi$  has an *essential* singularity rather than a power-law singularity as in the two- and three-dimensional cases, the normalized moments  $f_\ell$  are bounded from above for all  $\ell$  (Hajduković and Satz, 1992).

We turn now to intermittent behavior in heavy-ion experiments. To identify this behavior in rapidity distributions, we consider moments  $C_\ell$  that have been introduced by Bialas and Peschanski (1986), according to

$$C_\ell = \frac{1}{M} \sum_{m=1}^M (M p_m)^\ell. \quad (5.90)$$

Here  $M$  is the number of intervals of size  $\delta Y$  in a given rapidity interval  $\Delta Y = M \delta Y$ ,  $p_m$  ( $m=1, \dots, M$ ) denotes the probability of finding particles in any of these rapidity intervals  $dp_M$ . The total distribution  $P(p_1, \dots, p_M) dp_1 \cdots dp_M$  of probabilities for finding particles in the intervals  $dp_1 \cdots dp_M$  is normalized such that

$$p_1 + \cdots + p_M = 1, \quad \int dt \int dp_1 \cdots dp_M P(p_1, \dots, p_M) = 1. \quad (5.91)$$

The variable  $t$  stands for additional kinematic variables such as the energy of the collision. The average moment  $\langle C_\ell \rangle$  is obtained as the sum over all configurations in rapidity space, weighted by the probability distribution  $P(\dots) dp$

$$\langle C_\ell \rangle = \int dt \int dp_1 \cdots dp_M P(p_1, \dots, p_M; t) \times \frac{1}{M} \sum_{m=1}^M (M p_m)^\ell. \quad (5.92)$$

In terms of rapidity variables, genuine intermittent behavior is signalled, if

$$\ln \langle C_\ell \rangle = \phi_\ell \ln(\Delta Y / \delta Y) = \phi_\ell \ln M \quad (5.93)$$

in the limit of  $\delta Y \rightarrow 0$ , in other words a logarithmic dependence on the resolution size in the limit of increasing

resolution, which is equivalent to the definition in Eq. (5.89). Small values for  $\delta Y$  correspond to a fine resolution in rapidity space, and  $\Delta Y$  is the full considered rapidity interval. The lower bound on  $\delta Y$  is given by the experimental resolution. The intermittency indices  $\phi_\ell$  vary between  $0 < \phi_\ell \leq \ell - 1$ . They are a measure for the strength of the intermittent behavior. Intermittency occurs if self-similar fluctuations exist on all scales  $\delta Y$ . In the case of rapidity distributions, intermittency should be seen in contrast to dynamical fluctuations.

In *real experiments* we have to deal with finite-size systems. Statistical fluctuations around the probability distribution  $dp_1 \cdots dp_M P(p_1, \dots, p_M; t)$  will always provide a noisy background for fluctuations of dynamical origin that we are trying to identify. Unless the multiplicity in the events is very high, the probability  $p_m$  of finding particles in the rapidity interval  $m$  is different from the *measured fraction*  $K_m/N$  of the total multiplicity  $N$ . As a filter of dynamical fluctuations, Bialas and Peschanski (1986) have proposed to consider scaled factorial moments  $\langle F_\ell \rangle$  of the distribution  $Q$  that is actually measured in the experiment

$$\langle F_\ell \rangle = M^\ell \sum_{K_1 \cdots K_M} Q(K_1 \cdots K_M) \frac{1}{M} \times \sum_{m=1}^M \frac{K_m(K_m-1) \cdots (K_m-i+1)}{N(N-1) \cdots (N-i+1)}, \quad (5.94a)$$

where

$$Q(K_1, \dots, K_M) = \int dt \int dp_1 \cdots dp_M P(p_1, \dots, p_M; t) \times B(p_1, \dots, p_M; K_1, \dots, K_M), \quad (5.94b)$$

and  $B$  is the Bernoulli distribution. The statement is that the scaled moments  $\langle C_\ell \rangle$  of a probability distribution  $P$  are equal to the scaled factorial moments  $\langle F_\ell \rangle$  of the experimental distribution  $Q$ . It can be shown that  $\langle F_\ell \rangle = \langle C_\ell \rangle$ . Notice the advantage of using  $\langle F_\ell \rangle$  instead of  $\langle C_\ell \rangle$ . In Eq. (5.94a),  $\langle F_\ell \rangle$  is exclusively expressed in quantities which are directly measured. The average in Eq. (5.94a) can be evaluated by absorbing the weights  $Q(\dots)$  in the selection of events.

Further information about the origin of intermittency is contained in the  $\ell$ -dependence of  $\phi_\ell$ . Equivalently, one can use the  $\ell$ -dependence of anomalous fractal dimensions  $d_\ell$  that are related to  $\phi_\ell$  according to

$$d_\ell = \phi_\ell / (\ell - 1). \quad (5.95)$$

At least two physical mechanisms are known that may lead to real intermittent behavior in heavy-ion collisions (real in contrast to fluctuations that are induced by two-particle correlations). One of these mechanisms is *the QCD transition*, if the correlation length diverges at  $T_c$ , or, adapted to the finite volume, the correlation length spans the typical volume of the collision. In this case,  $d_\ell$  should be approximately independent of  $\ell$ . Other origins are *self-similar cascades*, which are at

tempts to model the evolution of the plasma to the hadronic phase. Here  $d_\ell$  will depend on  $\ell$  in general. Unfortunately, the  $\ell$ -dependence of  $d_\ell$  does not allow a unique identification of the origin of intermittency. The independence of  $d_\ell$  on  $\ell$  does not guarantee a phase transition, nor does its dependence exclude a transition. According to the work by Bialas and Peschanski (1986), it may be possible to tune the parameters of a self-similar cascade in such a way that  $d_\ell$  does not depend on  $\ell$ . On the other hand, multifractal phase transitions are known which lead to an  $\ell$ -dependent  $d_\ell$  (although the latter possibility is not predicted by QCD).

We add a remark on dynamical fluctuations due to two-particle correlations. In this case, the scaled moments first rise with decreasing bin size, and then saturate to some constant value below a certain resolution size  $\delta Y_0$ ;  $\delta Y_0$  also gives the typical size of the dynamical fluctuations. Two-particle correlations can result from resonance decays or Bose-Einstein correlations. For resolutions  $\delta Y > \delta Y_0$ , two-particle correlations are mainly responsible for apparent intermittency effects. At least in  $e^+e^-$ -reactions, the observed intermittent behavior can be fully explained by two-particle correlations within a certain particle-production scheme. For heavy-ion collisions, the particle-production processes in the matter evolution are less well understood, but at least the contribution from Bose-Einstein correlations can be tested experimentally—see, for example Schmidt and Schukraft (1993).

We summarize what a measurement of the anomalous fractal dimension  $d_\ell$  can tell us. Assuming that the contribution from two-particle correlations can be subtracted, the following alternatives can be distinguished:

(i) If  $d_\ell \sim 0$ , the fluctuations are of purely statistical origin. A first-order transition with a correlation length smaller than the size of the system would be compatible with such a scenario.

(ii) If  $d_\ell$  depends strongly on  $\ell$ , it seems rather unlikely that the QCD transition itself is responsible for inducing such multifractal behavior. More likely, there is a cascade mechanism at work whose vertices should be chosen to reproduce the observed  $\ell$ -dependence of  $d_\ell$ . For a cascading process one expects  $d_\ell$  to depend linearly on  $\ell$  (Bialas and Peschanski, 1988).

(iii) If  $d_\ell$  depends weakly on  $\ell$ , but is clearly different from zero, a second-order transition provides an explanation for intermittency, but there may be other explanations.

We conclude with a short glance at the experimental status. Intermittency has been found in  $e^+e^-$  reactions, hadron-hadron and heavy-ion collisions—see for example, Abreu *et al.* (1990), Buschbeck and Lipa (1989), Åkesson *et al.* (1990), respectively. Qualitatively, different experiments agree about the result that intermittency has been observed and that different mechanisms are at work, as the observed behavior is not uniform. Quantitatively, they disagree. In Fig. 38 we display the dependence of the anomalous fractal dimension on  $\ell \neq p$  at  $\sqrt{s} = 20$  GeV (taken from Bialas and Hwa 1991). Obviously, the  $\ell$ -dependence of  $d_\ell$  weakens when pass-

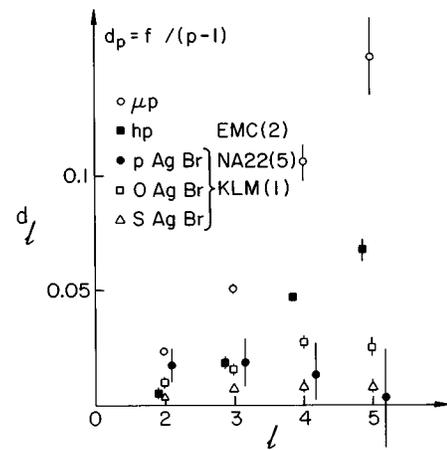


FIG. 38. Anomalous dimensions of pseudorapidity spectra at  $s^{1/2} \sim 20$  GeV. After Bialas and Hwa (1991).

ing from lepton-hadron to hadron-hadron and nucleus-nucleus collisions. In Bialas and Hwa, (1991), this is interpreted as a result of self-similar cascade mechanisms in the first two cases, but as a phase transition in the S-Ag, Br central collisions. The data for Ag, Br collisions, which are consistent with an almost constant non-zero value of  $d_\ell$ , suffer from low statistics (Schmidt and Schukraft, 1993). Another tendency seen in Fig. 38 is a decrease in  $d_\ell$  when the process becomes more complex, through this decrease is less strong for small values of  $\ell$ . Thus it may be more than an artifact of generally small values of  $d_\ell$  that  $d_\ell$  is almost independent of  $\ell$  for heavy-ion collisions.

Less encouraging are the high statistics data taken at the CERN SPS accelerator (Bloomer *et al.*, 1992) for  $^{32}\text{S}+\text{S}$  and  $^{32}\text{S}+\text{Au}$  collisions. The slope of  $\ln\langle F_2 \rangle$  vs  $(-\ln\delta\eta)$  ( $\eta$  being the pseudorapidity) is consistent with a value of zero or less than zero. Thus one of the conclusions of the authors is that there is “no need for new physics” in the sense of fluctuations of unknown dynamical origin.

Another systematic trend seems to be well established by the data. As we have mentioned above, two-particle correlations may look like intermittency signals as long as the resolution is not sufficiently high. In fact, for lighter projectiles like oxygen or silicon, the data for intermittency indices  $\phi_2$  can be well fitted as a function of the average particle density  $\langle \rho \rangle$ . Such a scaling law has been proposed by Seibert (1990) and is based on the assumption that only two-particle correlations are responsible for the observed intermittent effects. The optimistic news is that, for heavier projectiles such as those in sulfur-induced reactions such an explanation is not sufficient. For heavier projectiles, there seems to be a chance of seeing collective effects. These could be due to a nearly second-order transition or to a cascade mechanism during the matter evolution. Last, but not least, detector effects should be kept in mind as a possible source for contaminating the data. For further details on the experimental status we refer the reader to the review of Schmidt and Schukraft (1993) and references therein.

Thus the main task in an intermittency analysis is to disentangle fluctuations of different origins: geometry (impact-parameter dependence), statistics (finite number of particles), resonance decays or Bose-Einstein correlations, hadronization along cascade mechanisms, and a phase transition with a large correlation length. Recently work by Campi and Krivine (1995) questions any connection between the observation of intermittent behavior in the mass distributions of finite nuclei and a second-order phase transition. In any case, one has to face a superposition of several backgrounds on the type of multiplicity fluctuations in which one is primarily interested. Once an interesting structure is filtered from the backgrounds, the  $\ell$ -dependence of intermittency indices may indicate the specific dynamical origin of this structure. The situation resembles that for dilepton rates, where the main practical problem is identification of thermal rates. Once the thermal rates are identified, the remaining task is to look for conclusive structures of the transition scenario (see Sec. V.C.2).

#### D. Theoretical concepts for off-equilibrium situations

##### 1. Nucleation rate of hadronic bubbles

Many phenomenological implications assume the scenario of a first-order transition, which is often described by a bag model equation of state. The scenarios differ in the assumptions about entropy production during the transition, the amount of supercooling, and the production of latent heat.

In this section we describe an approach of Csernai and Kapusta (1992a and 1992b) for calculating the nucleation rate of hadronic bubbles in the plasma phase. In principle, such a calculation can select a realistic scenario. For example, it should tell us whether or not a Maxwell construction of two-phase equilibrium is realistic in view of the total duration of the plasma evolution to the final hadronized state, a time scale is of the order of 10–20 fm/c. This requires a nucleation rate sufficiently high that the latent heat is conducted fast enough from the bubble surfaces for the bubbles to grow. Naturally, transport properties of the latent heat are determined by properties of the “conducting” medium surrounding the bubble. Relevant properties are the thermal conductivity, and the shear and bulk viscosities, if a hydrodynamical description is used. In our case of an expanding plasma on its way to hadronization, the plasma viscosity will turn out to determine the nucleation rate of the hadronic bubbles.

The prediction of a definite value for the nucleation rate and of the time delay due to a first-order transition is rather ambitious and involves, unavoidably, a number of assumptions and approximations. Deviations from equilibrium have to be admitted during the phase transition, which need not be perturbatively small. One has to deal with an expanding system of many particles interacting according to the excitation spectrum of QCD.

Quite generally, the nucleation rate of one phase out of another (in this case the hadronic phase out of the plasma phase) has the form

$$I = I_0 e^{-\Delta F/k_B T}, \quad (5.96)$$

where  $\Delta F$  is the activation energy to form a bubble of critical size, and  $k_B$  is Boltzmann’s constant. This is the usual expression for the rate of a thermally activated process. More precisely, it is the probability to form a (hadronic) bubble of critical size per unit time and unit volume. The activation energy  $\Delta F$  can be expressed more generally as the difference in the free energies between two configurations with and without a condensed part of the new phase (it need not have the form of a droplet). The delicate part is the prefactor  $I_0$ , which describes some sort of quantum and/or thermal fluctuations. It sensitively depends on the context and the approximation scheme.

A specific realization of Eq. (5.96) is the famous formula of Callan and Coleman (1977) for the decay rate of the “false” vacuum. The  $\Delta F$  in the exponent is given by the Euclidean action  $S_4$  (in four dimensions) of the field theory considered, evaluated at a solution  $\phi$  of the Euclidean equation of motion satisfying  $\phi \rightarrow 0$  for  $x^2 + t^2 \rightarrow \infty$  as boundary condition. The prefactor depends on  $S_4$  and a ratio of determinants depending on some zero-temperature (effective) potential  $V$ . The occurrence of determinants in the prefactor is a common feature of formulas for decay rates when fluctuations are taken into account in the Gaussian approximation. Although Callan and Coleman’s formula is sometimes mentioned in the context of phase transitions in the early universe, it is not appropriate for calculating a decay rate at finite temperature in a situation out of equilibrium, such as that encountered in heavy-ion collisions.

A suitable framework is provided by Langer’s approach (Langer, 1969), which has been generalized to relativistic field theory and applied to QCD by Csernai and Kapusta (1992a). We briefly summarize the main ingredients in Langer’s formalism. The zero-temperature infinite volume potential in Callan and Coleman’s formula is replaced by a so called coarse-grained free energy, which is the appropriate quantity for describing a phase coexistence in a finite volume. Langer considers an effectively classical system of  $N$  degrees of freedom, represented by a set of  $N$  coordinates and their conjugate momenta  $\eta_i$ , where  $i = 1, \dots, 2N$ . These variables need not describe the system on a microscopic level, but enter the coarse-grained free energy as some kind of collective coordinates. Whereas the Euclidean action  $S_4$  (mentioned above) depends on microscopic variables, the coarse-grained free energy is formulated in terms of ‘mesoscopic’ variables, which are collective coordinates.

The  $\Delta F$  term in Eq. (5.96) keeps its meaning as the change in (coarse-grained) free energy due to a formation of one droplet of critical radius  $R_*$ . The prefactor  $I_0$  of the general formula accounts for (thermal) fluctuations in a specific way. It is written as a product of two terms

$$I_0 = \frac{\kappa}{2\pi} \Omega_0, \quad (5.97)$$

a statistic factor  $\Omega_0$ , which is a measure of phase-space volume around the saddle point, and a dynamical factor  $\kappa$ . While  $\Omega_0$  will be seen to depend on equilibrium quantities,  $\kappa$  knows about off-equilibrium features of the dynamical expansion. The dynamics may be described by hydrodynamical equations of motion. Thus we no longer restricted to the regime of  $O(4)$ -invariant Euclidean equations of motion in equilibrium situations. A direct manifestation of off-equilibrium features is dependence of  $\kappa$  on the thermal conductivity in a liquid-gas system or on the shear and bulk viscosities ( $\rho$  and  $\eta$ , respectively) in the case of QCD, [see Eq. (5.116) below]. This is clearly a generalization compared to the approach of Callan and Coleman, however Langer's formulation does not apply to generic off-equilibrium systems. The *basic assumption* is that the conversion proceeds via nucleation. The nucleation rate comes out as time independent by construction. The phase transition is described as the motion of a probability current  $\rho$  in phase space. It starts at  $\{\eta_0\}$ , the metastable phase, passes most likely a saddle point  $\{\bar{\eta}\}$ , where the current is assumed to be stationary

$$\partial_t \rho = 0, \quad (5.98)$$

and ends in the stable-phase configuration  $\{\eta_f\}$ . The configuration  $\{\bar{\eta}\}$  differs from  $\{\eta_0\}$  by just one critical droplet of new phase. To allow a stationary current across the saddle in phase space, there has to be a heat bath in the background, which replenishes the metastable phase at the same rate as it is lost in the phase-conversion process. The assumption in Langer's derivation is that such a heat bath exists, and is in equilibrium prior to each interaction with the converting phase. The probability for a transition between two configurations  $\{\eta\}$  and  $\{\eta'\}$ , induced by thermal fluctuations of the heat bath, is determined by the same Hamiltonian  $F$  that enters the dynamical part of the system which transforms into the new phase.

The dynamical part is assumed to be governed by equations of motion of the form

$$\frac{\partial \eta_i}{\partial t} = - \sum_{j=1}^{2N} M_{ij} \frac{\partial F}{\partial \eta_j}. \quad (5.99)$$

In general,  $F$  plays the role of a Hamiltonian; in an application to a hydrodynamical system  $F$  has the specific form of a coarse-grained free energy. Later,  $F$  will contain information that is specific for the underlying QCD. The matrix  $M$  is the mobility matrix—in our application, the coefficients  $M_{ij}$  can be read off from a comparison with the equations of relativistic fluid dynamics. Equation (5.99) enters the time development of the distribution function  $\rho(\{\eta\}, t)$  associated with the variables  $\eta_i$  and gives the probability density over configurations  $\{\eta\}$ . For equilibrium configurations,  $\rho$  is proportional to the Boltzmann weight  $\exp[-F\{\eta\}/T]$ , where  $F$  are the costs in (coarse-grained) free energy of generating such a configuration. Skipping some steps in the derivation, we see that the time evolution of  $\rho$  can be finally written in the form of a continuity equation:

$$\frac{\partial \rho\{\eta\}}{\partial t} = - \sum_i \frac{\partial J_i}{\partial \eta_i}$$

with

$$J_i = - \sum_j M_{ij} \left( \frac{\partial F}{\partial \eta_j} \rho + T \frac{\partial \rho}{\partial \eta_j} \right), \quad (5.100)$$

where  $F$  is the Hamiltonian and  $T$  is the temperature. The probability current across the saddle is identified with the decay rate. The prefactor  $\kappa$  of Eq. (5.97) gives the initial exponential growth rate of a bubble that has just exceeded its critical size  $R_*$ . It is defined as

$$\kappa(t) = \frac{d}{dt} \ln[R(t) - R_*]. \quad (5.101)$$

To obtain  $\kappa$  from the time evolution of  $\rho$ , Eq. (5.101) is solved for *small* deviations  $v_i = \eta_i - \bar{\eta}_i$  from the saddle  $\{\bar{\eta}\}$ . Note that this is a further assumption entering the derivation of  $\kappa$ . The deviations from the saddle must be small for the saddle-point approximation to make sense. For details about the statistical prefactor  $\Omega_0$  we refer the reader to the original literature (Langer, 1969; Langer and Turski, 1973; Turski and Langer, 1980).

Let us turn next to the application in QCD (Csernai and Kapusta, 1992a, 1992b). We have to specify the collective variables  $\eta_i$ , the coarse-grained free energy  $F$ , the hydrodynamic equations of motion, and their energy momentum tensor  $T_{\mu\nu}$ . The mobility matrix  $M$  and the stationary configuration  $\{\bar{\eta}\}$  must be identified.

The collective variables  $\eta_i$  are chosen to be energy density  $e(\mathbf{r})$  and flow momentum  $\mathbf{P}(\mathbf{r})$  at positions  $\mathbf{r}$  in the system.

The choice of a coarse-grained free energy  $F$  for QCD is more subtle. The hydrodynamic approach is suitable for describing nucleation rates in a fluid, since droplet formation occurs on a semimacroscopic level, where a large number of particles contribute to a single droplet. The concept of the coarse-grained free energy is introduced when hydrodynamics is derived from microscopic kinetic theory (Jackson, 1960; Van Kampen, 1964). The macroscopic system is divided into cells of a given volume. Specific densities ( $e$  and  $\mathcal{P}$  in our case) are assigned to each cell. If the partition function of the microscopic system is evaluated under these cellular constraints on the microscopic variables, one obtains an effective action (in more field-theoretical language), named the *coarse-grained free energy* over  $k_B T$ . (Note that this concept is very similar to the *constrained effective potential* used in Sec. IV.A.4. In the constrained effective potential, the entire system is considered as one cell subjected to a constraint on the vacuum expectation value of the order parameter field.)

The underlying QCD dynamics enters the interaction part  $f(e)$  of the coarse-grained free energy  $F$  for a small volume of nonuniform energy density  $e$ . The grain size of  $F$  is obviously essential for the approximation to make sense. If the grain size is too small, the use of hydrodynamic variables is not meaningful; if it is too large, the coarse-grained free energy cannot resolve the

substructure of the fluid. This happens if phase separation occurs within single cells, in which case  $F$  would be a convex function of the local average energy density  $e$ . Similarly, the constrained effective potential becomes convex in the large-volume limit. Thus an ansatz for  $f(e)$  that leads to a nonconvex shape of  $F$  as function of  $e$  will indirectly implement the right grain size, which is sensitive to nonequilibrium properties such as droplet growth and phase separation. Note that  $F$  is nonconvex, but real; the nonconvex shape is not an artifact of a perturbative approximation scheme, in which a thermodynamic potential is evaluated in the infinite-volume limit. The very grain structure is implemented in the choice of hydrodynamic variables, implicitly involving averages over cells.

As an ansatz for  $f(e)$ , Csernai and Kapusta (1992a) chose a fourth-order polynomial in  $(e - e_0)$ , where  $e_0$  is the location of the top of the potential barrier. The coefficients depend on the energy densities in the hadron and plasma phases, and on the curvature  $f''_0$  of  $f(e)$  at  $e_0$ . In the full ansatz for  $F$ , one is left with two free parameters,  $f''_0$  and the coefficient  $K$  of the derivative term  $(\nabla e)^2$ . Both parameters can be expressed in terms of the interface tension  $\sigma$  and the interface thickness  $2\xi_0$ . Here  $\xi_0$  denotes the correlation length defined at the top of the barrier by  $\xi_0^2 \equiv -K/f''_0$ . Note that both quantities ( $\sigma$  and  $\xi_0$ ) can be calculated from first principles, that is, from the underlying QCD Lagrangian.

The *equations of motion* are the equations of relativistic fluid dynamics,  $\partial_\mu T^{\mu\nu} = 0$ . In the baryon-free region of the quark-gluon plasma, we have to deal with a relativistic system. Without a net baryon number, there is no distinguishable reference frame in which to define the flow four-velocity  $u_\mu$  of the matter. Moreover, the pressure is not small compared to the energy. Therefore  $u_\mu$  is defined as the velocity of the collective energy flow. The *stress-energy-momentum tensor* for a relativistic dissipative fluid depends on the shear and bulk viscosities  $\eta$  and  $\zeta$ , respectively. It should be noticed that for a vanishing dissipation part the growth rate of bubbles would vanish. Similarly, it has been noticed by Langer and Turski (1973) that the growth rate in a liquid-gas transition vanishes near the critical point, if the thermal conductivity is set to zero. The bubbles can only grow if the latent heat is conducted from the bubble surface.

We skip further details of the derivation and come to the results. The exponential suppression factor in Eq. (5.96) depends on the change in free energy  $\Delta F$ , if a hadronic bubble has formed in the QCD plasma at  $T < T_c$  due to a thermal fluctuation. This is given by a sum of a volume and a surface term, according to

$$\Delta F = \frac{4\pi}{3}(f_h - f_p)R^3 + 4\pi R^2\sigma, \quad (5.102)$$

where the indices  $h$  and  $p$  indicate the corresponding phases. The critical radius  $R_*$  is given by

$$R_*(T) = \frac{2\sigma}{p_h(T) - p_p(T)}, \quad (5.103)$$

where  $p_h$  and  $p_p$  denote the pressure in the hadron and the plasma phase, respectively. (Bubbles with  $R < R_*$  collapse, those with  $R > R_*$  expand.)

The *statistical prefactor*  $\Omega_0$  can be expressed in terms of the surface tension  $\sigma$ , the bubble radius  $R$ , the correlation length  $\xi_p$  in the plasma phase, and the volume  $V$ , according to

$$\Omega_0 = \frac{2}{3\sqrt{3}} \left(\frac{\sigma}{T}\right)^{3/2} \left(\frac{R}{\xi_p}\right)^4 V. \quad (5.104)$$

The *dynamical prefactor*  $\kappa$  that determines the exponential growth of a fluctuation  $\nu$  of a droplet of critical size is derived to be

$$\kappa = \frac{4\sigma(\zeta + 4\eta/3)}{(\Delta w)^2 R_*^3}, \quad (5.105)$$

where  $\Delta w$  is the enthalpy difference between the two phases. Note that  $\kappa$  is determined in terms of equilibrium quantities  $\sigma$ ,  $R_*$ , and  $w$ , which are in principle accessible in lattice calculations, and nonequilibrium quantities, which are the transport coefficients  $\eta$  and  $\zeta$ . For an ultrarelativistic gas, the bulk viscosity  $\zeta$  is much smaller than the shear viscosity  $\eta$ , which has been estimated to leading order in perturbative QCD by Baym *et al.* (1990). For two flavors, the result is

$$\eta = \frac{2.11T^3}{\alpha_s^2 \ln(1/\alpha_s)}, \quad (5.106)$$

where  $\alpha_s$  is the strong coupling constant at finite temperature.

Eqs. (5.97), (5.104), and (5.105), with  $\Delta F$  from Eq. (5.102), imply that the nucleation rate

$$I = \frac{4}{\pi} \left(\frac{\sigma}{3T}\right)^{3/2} \frac{\sigma(\zeta_p + 4\eta_p/3)R_*}{\xi_p^4 (\Delta w)^2} e^{-\Delta F/T}. \quad (5.107)$$

The input parameters are chosen as follows:  $\sigma = 50$  MeV/fm<sup>2</sup>,  $\alpha_s = 0.23$ ,  $\xi_p = 0.7$  fm,  $\eta_p = 14.4$  T<sup>3</sup> and  $B^{1/4} = 235$  MeV for the bag constant that enters the enthalpy difference  $\Delta w$ . The pressure and the enthalpy difference between the hadronic and the plasma phases are calculated from the bag model equation of state. We would like to add a comment on the input. More recent results on the surface tension in pure SU(3) gauge theory show that the estimate for  $\sigma$  of 50 MeV/fm<sup>2</sup> that has been used by Csernai and Kapusta (1992a) is too large by an order of magnitude, see for example, Sec. III.B.3, or Kajantie (1992). It would also be quite interesting to update the equation of state, replacing the bag model values by more recent lattice results, since the bag model equation of state fails in particular in the transition region.

The nucleation rate  $I$  naturally defines a *nucleation time* if  $I$  is multiplied by the volume of a bubble of critical size:

$$\tau_{nuct}^{-1}(T) = \frac{4\pi}{3} R_*^3(T) I(T). \quad (5.108)$$

This time scale is a measure for the delay due to supercooling. It should be distinguished from the total time

duration of the phase transition. It just refers to bubbles of critical size, but neglects bubble growth. When  $\tau$  is plotted as a function of  $(T/T_c)$ , one sees that 5% of supercooling is necessary in order for the nucleation to set in, when the above input parameters are used. The use of updated input parameters may considerably reduce the degree of supercooling.

A hint on the self-consistency of the approximation scheme is provided by the temperature dependence of the critical bubble radius  $R_*$ , which is determined by Eq. (5.103). The radius diverges at  $T_c$ , but rapidly reduces to typical hadronic scales, and even below. At  $T/T_c=0.9\%$ ,  $R_*\sim 0.73$  fm. This is comparable to the correlation length in the plasma phase  $\xi_P$  and is of the order of the surface thickness. For such values of  $R_*$ , the approximation scheme breaks down. Therefore the question arises whether the phase transition has been completed before or after these low temperatures are reached. The quantity of interest is the *time dependence of temperature*. In the early universe,  $T$  drops so slowly with time that the phase transition is most likely over before the approximations lose their validity.

For heavy-ion collisions, the time dependence of  $T$  has been estimated by Csernai and Kapusta (1992b). The  $T(t)$  dependence in the bulk phase is known, once an equation of state is combined with an ansatz for the expansion scenario, such as Bjorken's scaling ansatz. The time dependence of  $T$  during the phase conversions is a more subtle problem. In the Maxwell construction,  $T$  is independent of  $t$  right from the beginning. To find  $T(t)$  for a given nucleation rate  $I$ , one starts from a dynamical equation that couples  $T(t)$  to the fraction of space that has been converted to the hadronic phase.

The result for  $T(t)$  clearly shows supercooling, superheating effects, and a time delay in the temperature decrease compared to a Maxwell scenario. When the temperature has decreased to  $\sim 0.8T_c$ , the nucleation is sufficient for reheating the system up to  $\sim 0.95T_c$ , at which point further nucleation is stopped. From that temperature on, the previously created bubbles grow further until the conversion to the hadronic phase is complete. Compared to the Maxwell scenario, the time of completion is delayed by 11 fm/c, leading to an increase in volume at that completion time. The increase is linear in  $t$  in Bjorken's ansatz. If the entropy density at completion is only determined by the equation of state in the pure hadronic phase, but independent of the conversion procedure, a delay of 11 fm/c implies 30% of extra entropy generation to keep the entropy density fixed. This is not negligible in an extrapolation from the final pion multiplicity to the initial temperature. It has to be verified whether or not the amount of extra entropy generation is an artifact of the bag model equation of state.

Finally, we note that the critical bubble radius in the late period of the phase conversion is about 1 fm (not too large in view of the finite size of the system and not too small for the approximation to hold).

Now let us return to the question posed in the beginning of this section: Is a Maxwell construction a realistic

scenario? The time dependence of temperature during the transition, based on the *derived* nucleation rate  $I$  [Eq. (5.107)], comes out not so different from a Maxwell construction, which is usually taken as an *ad hoc* assumption. It remains to be seen whether the similarity can be maintained if one improves on various approximations, and the most recent input data are used for the surface tension, the surface thickness, and the enthalpy difference.

In Langer's formalism for deriving the nucleation rate, an embedding heat bath was assumed to guarantee a stationary flow in probability space or a time-independent nucleation rate. Langer's description does not apply to a situation far out of equilibrium. The remnants of off-equilibrium are contained in the viscosity coefficients of the nucleation rate. It may well be that such a heat bath is not realized at all stages of phase conversion in heavy-ion collisions. Some time dependence of the nucleation rate is natural, at least at the onset and the completion of nucleation. A calculation of time-dependent rates requires a different formalism.

Boyanovsky and Aragao de Carvalho (1993, and 1994) analyzed the thermal activation via sphaleron transitions in a (1+1)-dimensional field theory, in which they used a real-time formalism. They obtained a time-dependent nucleation rate from the real-time evolution of the initial density matrix along the unstable direction in configuration space. The rate is rather sensitive to the initial state, a trait typical of an off-equilibrium process. The initial state is metastable and was formed after a period of *rapid supercooling*.

One may think of the phase transition in heavy-ion collisions in a similar way. The initial state for the conversion process is a metastable state arising from a rapid cooling, rather than being in local equilibrium with a heat reservoir. The time evolution of the conversion may come out quite differently, if the heat bath is not replenished during the conversion process. A comparison of nucleation rates calculated in one or the other way would be quite useful for assessing the estimates of time delay and extra entropy production, which should be visible in heavy-ion collisions.

Recent results from the lattice and from studies in effective models suggest that the transition is weakly first order—if it is first order at all. In this case, nucleation may completely fail as a conversion mechanism from the plasma to the hadron phase. Large domain coarsening might replace nucleations as  $T$  approaches  $T_c$  from above. Nonequilibrium aspects in weak first-order transitions are discussed in the next section.

## 2. Large domain coarsening

Different nonequilibrium aspects may be relevant for the transition dynamics. Here we distinguish only between *nucleation* and *large domain coarsening* or *spinodal decomposition*. Nucleation of hadronic bubbles in a first-order QCD transition was the topic of the preceding section. The physical assumption there was that the system was still in a homogeneous plasma phase, as  $T_c$  was reached from above. Some supercooling below  $T_c$

was necessary, until bubbles of critical size could be formed and the phase conversion could proceed via nucleation. This scenario is likely to be a realistic description for a *strong first-order* transition. The formal manifestation for the underlying physical assumption is the saddle-point evaluation of the path integral in Langer's formulation, as we saw above. The saddle-point approximation is justified as long as the fluctuations about the homogeneous metastable ground state are small. The system should be well localized in the metastable minimum as the temperature drops below  $T_c$ .

The nonequilibrium dynamics of the phase conversion may drastically change if the transition is *weakly first order* (Borrill and Gleiser, 1995). In spite of a potential barrier, the system may no longer be well localized in the metastable minimum, before the temperature drops below  $T_c$ . Large amplitude fluctuations are allowed, and subcritical bubbles can form *above*  $T_c$ . At  $T_c$  the system is in a completely mixed phase of high- and low-temperature "components." The phase mixing is assumed *not* to be an artifact of the finite volume. (Recall that spontaneous symmetry breaking, strictly speaking, occurs only in the infinite-volume limit. Here the volume is chosen large enough that the tunneling rate between the different phases is negligible.) If the system starts at  $T_c$  in a mixed phase and then cools below  $T_c$ , the phase separation may evolve by domain coarsening in a way that resembles spinodal decomposition rather than nucleation.

Borrill and Gleiser (1995) studied the phase mixing at  $T_c$  as a function of the strength of the first-order transition. The framework involves discretized scalar field theories in 3+1 dimensions. The homogeneous part of the free-energy density contains a quadratic, a cubic, and a quartic term in a real scalar field. The couplings in the quadratic and cubic part are assumed to be temperature dependent. The precise form is motivated by an effective potential for the electroweak phase transition, but the results are applicable to QCD as well. The strength of the first-order transition is tuned by the quartic coupling  $\lambda$ , keeping the cubic coupling fixed. Small values of  $\lambda$  correspond to a high barrier in the effective potential and a strong first-order transition; large values of  $\lambda$  correspond to a weak first-order transition. The system is always considered at fixed temperature  $T_c$ . It is prepared so as to be initially well localized in the high-temperature phase. It is then evolved by a discretized Markovian Langevin equation with white stochastic noise to mimic the coupling with a thermal bath.

Let us denote the fractions of the total volume in the high- and low temperature phases divided by the total volume a  $f_0$  and  $f_+$ , respectively. The fraction of the total volume in each phase,  $f_0(t)$  or  $F_+(t)$ , is then measured as a function of time, until the final ensemble-averaged equilibrium values  $f_{0,+}^{\text{eq}}$  are reached. The qualitative results of Borrill and Gleiser are the following. For small values of  $\lambda$  (strong first-order transitions), the system remains localized in the initial metastable minimum,  $f_0^{\text{eq}} \sim 1$ . For large enough  $\lambda$  (sufficiently weak

first-order transitions),  $f_0^{\text{eq}}$  approaches 0.5, corresponding to a complete phase mixing at  $T_c$ .

A further interesting observation is made, if the equilibrium fractional population difference at  $T_c$

$$\Delta F_{\text{eq}} = f_0^{\text{eq}} - f_+^{\text{eq}} \quad (5.109)$$

is plotted as a function of  $\lambda$ . The curve suggests a second order phase transition as a function of  $\lambda$  (the control parameter of the strength of the first-order finite- $T$  transition). There seems to exist a critical strength of the first-order transition (obtained for  $\lambda = \lambda_c$ ), so that for larger values of  $\lambda$  (weaker transitions),  $\Delta F$  soon approaches zero, corresponding to a perfect thermal mixing  $f_0^{\text{eq}} = 0.5 = f_+^{\text{eq}}$ . For smaller values of  $\lambda$  (stronger transitions),  $\Delta F_{\text{eq}}$  remains close to 1, and the system is trapped in its homogeneous initial state. Therefore, Borrill and Gleiser proposed,  $\Delta F_{\text{eq}}$  can be used as an order parameter for the "transition" from strong to weak first-order transitions:  $\Delta F_{\text{eq}}$  is a quantitative measure for how strong the finite- $T$  transition is.

The prediction of complete thermal mixing at  $T_c$  relies on a cooling process, which is slow compared to other equilibration scales in the system. Keeping  $T$  fixed at  $T_c$ , this assumption was implicitly built into the derivation. The narrow transition region between the two areas of strong and weak first-order transitions may be related to the fact that the order parameter for the finite- $T$  transition  $\langle \phi(t) \rangle$  is not conserved. For conserved order parameters in binary mixtures and long internal equilibration times, the transition between the two regimes is smooth (Borrill and Gleiser, 1995, and references therein).

To conclude, we summarize some possible implications for QCD. The results of Borrill and Gleiser have shown that the nonequilibrium transition dynamics may change from nucleation to large domain coarsening in passing from strong to weak first-order transitions. The interpolation between both schemes need not be smooth, and may be a phase transition itself. In QCD it is likely that the finite-temperature transitions are rather weakly first order (if they are of first order at all), or rapid crossover phenomena, which may be indistinguishable from a practical point of view. The role of the control parameter of the strength of the transition will be played by the current quark masses. If the cooling of the plasma is slow compared to equilibration time scales of amplitude fluctuations about the metastable ground state, the QCD transition may proceed via large domain coarsening rather than supercooling with subsequent nucleation.

Observable signatures in heavy-ion collisions will in general depend on the dominant process during the phase conversion. Large domain coarsening may result in similar effects as those expected for second-order transitions, while the extreme case of strong supercooling would lead to explosive effects.

In view of the results of Borrill and Gleiser, a good quantitative estimate of the strength of the first-order/crossover phenomenon no longer appears as a minor detail, once the order of the transition is determined. It

is particularly important if the regimes of nucleation and spinodal decomposition are separated by a narrow range in the control parameter of the potential barrier. Thorough estimates of time scales for the cooling process and relaxation phenomena during the phase conversion are needed to turn the speculations of today into predictions for observable signatures.

### 3. Transport coefficients

As we saw in Sec V.D.1, dissipation properties such as the shear and bulk viscosities of the plasma will influence the nucleation rate of hadronic bubbles, if the transition proceeds via nucleation. Viscosity and thermal conductivity are further sources of entropy production during the evolution from the initial to the final state. Although it is sometimes claimed that the amount of extra entropy is negligible compared to the entropy generated during a hypothetical supercooling scenario it is worth calculating; it may well be that there is no supercooling at all.

Corresponding to this extra entropy, dissipation leads to an increase in the final rapidity density of observed particles. The entropy as a function of time enters various production rates, thus an extra entropy will influence predictions of final rates as well.

Transport processes result from gradients in thermodynamical parameters. Gradients in the temperature lead to a thermal current, in the velocity field to a friction force, and in color distributions to a color current. They are irreversible and are typical nonequilibrium phenomena that have to be treated with nonequilibrium methods. We have discussed at length the results from equilibrium thermodynamics. Typically, they refer to the existence of phase transitions, critical parameters and the ingredients for the equation of state. In order to incorporate all of this knowledge in nonequilibrium effects, the equilibrium has to play the role of an expansion point, from which small perturbations are allowed. Thus it is not surprising that transport coefficients are calculated within some linear response approach. The linear response refers to the response of the energy-momentum tensor to the thermodynamic forces (friction forces, temperature gradients, etc.).

Two approaches are known: the *kinetic* and the *phase space approach*. The kinetic approach emphasizes particle aspects. Transport processes arise as the result of collisions between particles (partons or hadrons, depending on the phase). In the vicinity of  $T_c$ , kinetic theory is expected to fail. A hydrodynamic treatment seems to be more adequate, and is incorporated in the phase-space approach. In principle, the phase-space approach applies to all temperatures. The transport coefficients are expressed in terms of the expectation values of retarded Green's functions. Following the work of Zubarev (1974), Hosoya *et al.* (1984) elaborated the phase-space approach for the hydrodynamic regime (although it is not necessarily restricted to that). The hydrodynamic description is apparent in the choice of variables, the locally defined temperature and velocity fields.

As usual, this assumes that the mean-free time of constituent particles is much shorter than the typical relaxation time of the system.

Only linear terms in gradients of inhomogeneous distributions are kept. The system has to be in a late stage of preequilibrium expansion rather close to equilibrium to satisfy this condition. Under these restrictions, one obtains as an expression for the shear viscosity

$$\eta = - \int d^3x' \int_{-\infty}^t dt_1 \exp\{\varepsilon(t_1 - t)\} \times \int_{-\infty}^{t_1} dt' \langle T_{12}(\mathbf{x}, t) T_{12}(\mathbf{x}', t') \rangle_{\text{ret}} \quad (5.110)$$

in the limit of  $\varepsilon \rightarrow 0$ , where  $T_{\mu\nu}$  denote the components of the energy-momentum tensor. The formulas for the heat conductivity  $\chi$  and the bulk viscosity  $\zeta$  are very similar; they mainly differ in the components of  $T_{\mu\nu}$ . The expectation value  $\langle \dots \rangle_{\text{ret}}$  stands for a retarded Green's function in *Minkowski* time. It is related to the thermal expectation value  $\langle \dots \rangle_0$  with respect to the thermal equilibrium distribution according to

$$\langle T_{\mu\nu}(\mathbf{x}, t) T_{\mu\nu}(\mathbf{x}', t') \rangle_{\text{ret}} = -i\theta(t' - t) \times \langle [T_{\mu\nu}(\mathbf{x}, t), T_{\mu\nu}(\mathbf{x}', t')] \rangle_0. \quad (5.111)$$

Equation (5.111) is the expression in which we are finally interested. On the other hand, expectation values that are easily accessible on the lattice are thermal Green's functions in *Euclidean* time  $\tau$ . They are expressed via thermal expectation values

$$\langle T_{\mu\nu}(\mathbf{x}, \tau) T_{\mu\nu}(0, 0) \rangle_0, \quad 0 < \tau < \beta = \frac{1}{T}. \quad (5.112)$$

The procedure of analytic continuation from Euclidean to Minkowski time is standard and can be found in textbooks—see, for example, Kadanoff and Baym (1962). Here we only sketch the procedure.

In order to relate the retarded Green's function  $G_R(\mathbf{x}, t)$  to the thermal function  $G_\beta(\mathbf{p}, \tau)$ , one proceeds via their Fourier transforms  $\tilde{G}_R(\mathbf{p}, p_0)$  and  $\tilde{G}_\beta(\mathbf{p}, \omega_n)$ , respectively. They have almost identical representations in terms of the spectral functions  $\rho(\mathbf{p}, \omega)$ . We have

$$\tilde{G}_\beta(\mathbf{p}, \omega_n) = \int d\omega \frac{\rho(\mathbf{p}, \omega)}{i\omega_n - \omega}. \quad (5.113)$$

Here  $\tilde{G}_\beta(\mathbf{p}, \omega_n)$  is defined only on a discrete (but infinite) set of Matsubara frequencies  $\omega' = \omega_n$  (due to the periodicity in Euclidean time). Next, one performs an analytic continuation of the Fourier coefficients  $\tilde{G}_\beta(\mathbf{p}, \omega_n)$  for all (nonreal)  $\omega'$ . If there is no essential singularity at  $\omega' = \infty$ , the unique continuation is

$$\tilde{G}_\beta(\mathbf{p}, \omega') = \int d\omega \frac{\rho(\mathbf{p}, \omega)}{\omega' - \omega}. \quad (5.114)$$

The spectral function is then given by the discontinuity of  $\tilde{G}_\beta(\mathbf{p}, \omega')$  across the real axis. Once we know  $\rho$ , the Fourier transform of the retarded Green's function can be calculated according to

$$\tilde{G}_R(\mathbf{p}, p_0) = \int d\omega \frac{\rho(\mathbf{p}, \omega)}{p_0 + i\varepsilon - \omega} \quad (5.115)$$

with the same  $\rho$  as in Eq. (5.113). The retarded Green's function  $G_R(\mathbf{x}, t)$  then follows as a Fourier transform of Eq. (5.116). Thus in the ideal case we are done if we are able to extract the spectral function from certain correlators in Euclidean space. The spectral function is essential not only for analytic continuation; it also allows an identification of the physical degrees of freedom. In a given channel  $C$ ,  $\rho_c$  has a pole for each physical degree of freedom with quantum numbers  $C$ , and a branch cut for every allowed multiparticle state with the same  $C$ .

Practically, it is very difficult to determine  $\rho$  in lattice calculations.  $G_\beta(\mathbf{x}, \tau)$  is known only for a finite (often small) number of points  $\tau_i$ . Therefore,  $\tilde{G}_\beta(\mathbf{p}, \omega_n)$  cannot be calculated as integral  $\int^\beta d\tau e^{i\omega_n\tau} (\dots)$ , but the spectral representation of  $\tilde{G}_\beta(\mathbf{p}, \omega_n)$  is just the quantity of interest.

What can be achieved in spite of these difficulties has been shown by Karsch and Wyld (1987), who pointed out the possibilities and limitations of calculating transport coefficients in lattice gauge theory. The approximation for extracting the spectral function is to make an ansatz for  $\rho$  involving certain unknown parameters, to calculate  $\tilde{G}_\beta(\mathbf{p}, \omega_n)$  and  $G_\beta(\mathbf{p}, \tau)$ , and to fix the unknown parameters by fitting  $G_\beta(\mathbf{p}, \tau)$  with Monte Carlo data.

Kubo-type formulas [see Eq. (5.110)] were used with an energy-momentum tensor of pure SU(3) lattice gauge theory. The ansatz for the zero-momentum spectral function is given as

$$\rho(\omega) = \frac{A(1 - e^{-\beta m})}{\pi} \left( \frac{\gamma}{(m - \omega)^2 + \gamma^2} - \frac{\gamma}{(m + \omega)^2 + \gamma^2} \right). \quad (5.116)$$

This involves three parameters:  $A$ ,  $\gamma$ , and  $m$ . The spectral function of a free field theory contains just two,  $A$  and  $m$ , thus  $\gamma$  represents the interactions. Let us label the different transport coefficients  $\alpha_i$  with index  $i$ . Correspondingly, the spectral functions  $\rho_i(\omega)$  needed for the retarded Green's functions  $G_{R_i}(t)$  in the associated Kubo formulas depend on parameters  $A_i$ ,  $\gamma_i$  and  $m_i$ . The transport coefficients read, in terms of  $A_i$ ,  $\gamma_i$ , and  $m_i$  (Karsch and Wyld, 1987)

$$\alpha_i = 2A_i(1 - e^{-\beta m_i}) \frac{2\gamma_i m_i}{(\gamma_i^2 + m_i^2)^2}. \quad (5.117)$$

This expression can be easily reformulated in dimensionless quantities that are directly measurable on the lattice:  $m_\ell = ma$ ,  $\gamma/m$ ,  $1/T = N_\tau a$ , and  $A_\ell = Aa^5$ ,  $a$  being the lattice constant. The goal is to determine  $A_\ell$ ,  $m_\ell$ , and  $\gamma/m$  from fits of  $G_\beta^{(i)}(\tau)$ . For example,  $G_\beta^{(i)}(\tau)$  for the shear viscosity is given in terms of space-space off-diagonal correlations according to

$$G_\beta^{(\eta)}(\tau) = \sum_x \langle T_{12}(0,0) T_{12}(\mathbf{x}, \tau) \rangle. \quad (5.118)$$

This quantity was evaluated on an  $8^3 \times 4$  lattice. Thus the time distance  $\tau$  in lattice units is restricted to the three values 0, 1, and 2; it cannot be larger than 2. In fact, only  $G^{(\eta)}(\tau)$  could be measured up to distance 2. For the other correlators  $G^{(s)}$  and  $G^{(x)}$ , the signal is already lost in the noise at  $\tau=2$ .

The measurements were performed above and below the critical coupling of the deconfinement transition, which is at  $6/g^2 = 5.68$  for the given lattice size. The mass parameters  $m_i$  are more easily accessible than  $A_i$  and  $\gamma_i$ . They follow from the exponential decay of the correlations. The interesting result is that the correlations drop rapidly above and below  $T_c$ . Especially above  $T_c$ , there still seem to be massive modes in the system with a large effective mass. Thus there is no indication of an abrupt change from a free glueball gas below  $T_c$  to a free gluon gas above  $T_c$ . Only at low temperatures do the effective masses approach the lowest values for the glueball masses obtained from independent calculations (Berg and Billoire, 1983); in the vicinity of  $T_c$  they differ. Yet one should keep in mind the typical caveats for any lattice calculation. In particular, the extraction of masses from short-distance properties could lead to artificially large values. For the interaction parameter  $\gamma$ , Karsch and Wyld obtain an upper bound, resulting in an upper bound for the shear viscosity  $\eta$  of  $\eta/T^3 < 9.5$ . This bound is in agreement with analytic estimates in the collision-time approximation of the kinetic approach (see Hosoya and Kajantie, 1985).

We shall now summarize the lattice calculation of transport coefficients in the phase-space approach. The original goal of deriving nonequilibrium properties in Minkowski space from Euclidean field theory in thermal equilibrium is divided into two steps. Based on a linear response approach, the nonequilibrium properties are first expressed in terms of equilibrium quantities in Minkowski space. The second step is the analytic continuation between Minkowski and Euclidean spacetime. The essential ingredient is the spectral function  $\rho$ . A basic obstacle to deriving it from Euclidean time correlators measured on the lattice is the rather limited set of time points. At best, one can only guess an ansatz for the spectral function that is compatible with the time correlator. In addition, the correlators  $G_\beta$  of the matrix elements of  $T_{\mu\nu}$  turn out to drop so rapidly that the small values for  $G_\beta$  are lost in the noise for large distances. Large distances are needed to project on lowest excitation states, to justify the very ansatz for  $G_\beta$ . Thus the values for  $\tau$ , the time distance in lattice units, are too small both in number and size. From a realistic point of view, it seems to be impossible to reveal the rich analytic structure of the spectral function from a lattice approach. The main part of it will be hidden in the Monte Carlo noise. Further discussions of transport coefficients in the lattice approximation and comparisons to the results of kinetic theory can be found in Horsley and Schoenmaker (1986) and Schoenmaker and Horsley (1988).

Heat conductivity and the viscosities are *relaxation phenomena* of the *energy-momentum* with respect to temperature gradients, friction forces, etc. Their impact

on the phase transition concerns the phase conversion (see the preceding section). A direct manifestation of the phase transition should be visible in a different relaxation phenomenon associated with color degrees of freedom. It is noteworthy that one of the most challenging properties of a quark-gluon plasma—its *color conductivity*—presents itself as a transport phenomenon. In analogy with the Mott transition from an electric insulator to a conductor, the transition from the hadronic color insulator to the plasma color conductor should be manifest in the transport coefficient of color conductivity, which drops to zero either smoothly or suddenly, when  $T_c$  is approached from above.

Selikhov and Gyulassy (1993) found a surprisingly small value for the color conductivity, which is determined by the ratio of the static color electric and magnetic screening masses,  $m_e$  and  $m_m$ , respectively. [Their assumption is that  $m_m \ll m_e$ . This assumption is not borne out by lattice simulations in SU(2) and SU(3) gauge theories, in which  $m_m \sim m_e$  is found for temperatures from  $T_c$  up to  $2\text{--}3 T_c$  (Boyd, 1995).] The color conductivity coefficient is calculated from the linear response of the system to a weak external field. The reason for the rapid damping of collective color modes is rapid color diffusion. Color diffusion is characterized by a specific *relaxation time scale*  $t_c$ , which measures the precession frequency of the color of a parton in a fluctuating background field. This time scale turns out to be much smaller than a typical *momentum relaxation scale*  $t_p$ , which has been used in (classical) non-Abelian transport theory (see, for example, Heinz, 1986; Baym *et al.*, 1990; Eskola and Gyulassy, 1993). The small value of  $t_c$  explains the poor color conductivity.

Color transport phenomena are still in their infancy and deserve further attention in future work. One may think about constructing a device for a “color ammeter.”

We conclude with some remarks on a promising new direction in transport theory, which is suited for applications in heavy-ion collisions. It has been studied by Zhang and Willets (1992) and is based on a closed time-path Green’s function technique (Schwinger, 1961; Keldysh, 1964 1965) combined with the loop expansion scheme of Cornwall, Jackiw, and Tomboulis (1974). This theoretical framework leads to transport equations that include the generalized Boltzmann equations in the special case of the quasiclassical limit. The virtue of this approach is a unified description of equilibrium and off-equilibrium systems. The deviations from equilibrium in the off-equilibrium situations considered are not restricted to a small size, whereas small deviations are necessary for justifying a linear-response approach.

In particular, the framework is well suited for incorporating the phase transition. The order parameter (chosen here to be the chiral condensate of the Nambu–Jona-Lasinio model) is a *dynamical variable* that enters the generalized Boltzmann equations. The Boltzmann equations reduce to the Vlasov equation in the Hartree approximation. The Vlasov dynamics control the collision dynamics, particularly medium effects of chiral

symmetry breaking in transverse flow. The transverse flow is sensitive directly to the chiral condensate.

A further numerical elaboration of these microscopic transport equations could lead to predictions of observable signatures for chiral symmetry restoration. Such predictions should be more reliable (in the sense of being more realistic) than any predictions based on equilibrium thermodynamics. Ultimately, full QCD should be substituted for the Nambu–Jona-Lasinio model used in the analysis of Zhang and Willets. A less ambitious extension would incorporate three rather than two flavors, or vary the effective microscopic model for QCD to include gluonic degrees of freedom as well.

#### 4. Dynamical universality and disoriented chiral condensates

Most of the proposed experimental consequences of a QCD transition at finite temperature are based on the assumption of a first-order transition scenario. As a manifestation of a second-order transition we discussed intermittent behavior in Sec. V.C.6. In this section we describe further consequences of a second-order transition under an additional assumption of rapid cooling from the high- to the low-temperature phase. The underlying theory is the three-dimensional  $O(4)$  model of Sec. V.A, which is supposed to describe two-flavor QCD in the vicinity of  $T_c$  and predicts a second-order transition in the chiral limit. At criticality, long-range order in an  $O(4)$  model leads to large clusters of aligned spins, or, translated to the particle contents of the sigma model, large clusters in three-dimensional space, where the ratio of charged to neutral pions is fixed. The question arises whether the cluster size is sufficiently large to have observable effects. Due to the finite quark masses, the second order of the transition is an idealization anyway (unless the masses take tricritical values). Although the up and down masses are only a few MeV, the induced zero-temperature pion mass is  $\sim 135$  MeV and increases further with temperature (Rajagopal and Wilczek, 1993a, 1993b). As the lightest particles in the spectrum, the pions have the largest correlation length. Thus the largest correlation length is smaller than  $(135 \text{ MeV})^{-1}$ , which is not large compared to  $T_c^{-1}$  [if we use  $T_c \sim 140$  MeV for two flavors (Bernard *et al.*, 1992)]. The number of pions arising from a typical cluster of “aligned  $O(4)$  spins” in isospin space can be estimated from the energy stored in a typical correlation volume in the vicinity of  $T_c$  (Rajagopal and Wilczek, 1993a, 1993b). The estimate is based on lattice results and indicates that not more than one or two pions belong to the same correlation volume. Thus the hope of seeing clusters of hundreds of pions with a fixed ratio of neutral to charged pions seems to be gone. The upper bounds on the maximal size of a correlation volume and the energy stored in this volume are based on equilibrium thermodynamics. They apply to a phase transition in heavy-ion collisions, if the system always remains close to local equilibrium, particularly during the phase transition.

The situation may change drastically, if the transition proceeds far out of equilibrium via a *quench*. Let us re-

call what a quench is in spin systems. The system starts at high temperature in a completely disordered state. The temperature is then instantaneously turned to zero. An example of this is a bar of iron that is plunged into ice water, rather than being slowly cooled through the transition point. The high-temperature spin configuration is far out of equilibrium when the thermal fluctuations are suddenly switched off. The equilibrium configuration at zero temperature would be a well ordered ground state. In condensed-matter physics it is well known what happens to a high-temperature spin configuration after a quench (Halperin *et al.*, 1974; Hohenberg and Halperin, 1977). The system evolves according to the zero-temperature equations of motion, which determine the universal dynamics of the long-wavelength modes near  $T_c$ . Theories that are described by the same equations of motion for the long-wavelength modes define a *dynamical universality class*. In contrast to *static* universality classes, dynamical universality classes are specified not only by the space dimension and the number of order-parameter components, but also by those quantities that are conserved. The important result of condensed-matter physics (Bray, 1990; Newman *et al.*, 1990) is that the linear size of a correlation volume  $L(t)$  grows with time according to  $L(t) \sim t^p$ , where  $p$  is an exponent that depends on the characteristic parameters of the dynamical universality class. It is crucial, from the viewpoint of the application to QCD, that  $L(t)$  does not depend on equilibrium correlation lengths. Following this digression to condensed-matter physics, let us return to QCD.

The interesting suggestion of Rajagopal and Wilczek (1993a; 1993b) is that QCD belongs to the dynamical universality class of an  $O(4)$  antiferromagnet in the vicinity of  $T_c$ . The implications of this hypothesis are discussed below.

The question of interest is the specific way in which different regions of space relax to the ground state at low temperature. In the case of an  $O(4)$  model, the ground state at low  $T$  is characterized by small oscillations around the  $\sigma$ -direction in isospin space. The vacuum expectation value of the  $O(4)$  field is  $\langle \phi \rangle = \langle \sigma, \boldsymbol{\pi} \rangle = (f_\pi, \mathbf{0})$ . If the cooling is adiabatically slow, it is likely that different clusters of aligned spins relax independently to the ground state configuration. In contrast, if the cooling proceeds via a quench, the long-wavelength modes relax much more slowly to the equilibrium configuration. This allows the formation of large clusters of misaligned spins characterized by a “misaligned” condensate value  $\langle \boldsymbol{\pi} \rangle \neq \mathbf{0}$ . Such regions of misaligned condensates have been called *disoriented chiral condensates* by Bjorken *et al.* (1993). They speculated about events in which the pion yield comes in clusters of pions aligned in a single direction in isospin space over a large fraction of the collision volume. When these regions relax *coherently* to the true ground state with  $\langle \sigma \rangle \neq 0$ , a specific radiation of pions is emitted. This radiation is proposed as an observable signature in heavy-ion collisions. The “misaligned” high-temperature vacuum is determined by oscillations around some fixed direc-

tion  $(\sigma, \boldsymbol{\pi})$  in classical field space. The radiated pions correspond to a coherent configuration in that direction. Pions originating from one such cluster of misaligned condensate have a fixed ratio of charged to neutral pions. The probability distribution for a given charge ratio  $R$  is predicted as (Rajagopal and Wilczek, 1993a; 1993b)

$$\text{Prob}(R) = \frac{1}{2} R^{-1/2}, \quad (5.119)$$

where  $R$  is the ratio of neutral pions to the total number of pions.

Correlations in space are not directly accessible in heavy-ion collisions. Thus a further relation is needed to translate correlations in space to measurable correlations in rapidity space. Such a relation is provided by Bjorken’s scaling ansatz. In Bjorken’s description, different positions in space become different positions in rapidity space. Given  $z$  and  $t$  we know the rapidity  $Y = 1/2 \ln[(1+v)/(1-v)]$  (see Sec. V.B.1).

Thus an observable effect of a quenched QCD transition could be constant ratios of neutral pions to this total number of pions over large volumes in rapidity space. Fluctuations in pion multiplicities of this type should be distinguished from fluctuations originating in exploding plasma blobs, where the criterion to form a blob is a genuine nucleation process rather than a fixed direction in isospin space (see Sec. V.C.5).

Generic spin systems and an  $O(4)$  model as the effective description for QCD share similarities in view of dynamical universality, but also differ in several aspects, which we list in the following.

(i) At high temperatures, the  $O(4)$  vector will not really fluctuate in arbitrary directions from cluster to cluster due to the “external field”. The nonvanishing pion mass singles out a preferred direction even at high temperatures, assuming that the model has not lost its validity at high temperatures. The analysis of Bray (1990) and Newman *et al.* (1990) has to be extended to account for the effect of finite quark masses (see Rajagopal and Wilczek, 1993a; 1993b).

(ii) The sudden quench as a model for instantaneous cooling is certainly an idealization of the realistic cooling process. In Bjorken’s scaling ansatz,  $T$  drops according to  $T = T_0(\tau_0/\tau)^{c_s^2}$ , more slowly than the corresponding entropy density if transverse expansion is neglected ( $c_s$  is the speed of sound in the medium).

(iii) Strictly speaking it is not possible to talk about a well ordered ground state at  $T=0$ . The plasma is expanding until freezeout, at which point the temperature is larger than zero. After freezeout it no longer makes sense to call some parameter “temperature.”

(iv) Furthermore, a description in terms of field configurations is not convenient when individual particle aspects become more relevant. In the very end detectors register single particles. Even before freezeout the language of kinetic theory is more adequate.

(v) The equations of motion at zero temperature have to be Lorentz invariant in the QCD case. The proposal

is to use the  $T=0$   $O(4)$  linear sigma model to guarantee Lorentz-invariant equations of motion.

(vi) The plasma is expanding, whereas the ferromagnet is static.

In spite of the complications in heavy-ion collisions, there is at least a chance that a fast cooling amplifies the role of long-wavelength pionic modes during the phase conversion of the quark-gluon plasma. However, for realistic coupling strengths in the sigma model, their relaxation to the low- $T$  equilibrium configuration goes too fast to allow the formation of large regions of “misaligned” vacuum (Gavin *et al.*, 1994b). Gavin *et al.* performed a numerical simulation of the zero-temperature  $O(4)$  linear sigma model in the presence of an external field  $H$ , with couplings in the Lagrangian that induce physical values for  $f_\pi=92.5$  MeV,  $m_\pi=135$  MeV, and  $m_\sigma=600$  MeV. In particular, the quartic coupling is strong ( $\lambda=20$ ). In contrast, simulations at weak couplings corresponding to unrealistically light masses ( $m_\pi=0.3$  MeV and  $m_\sigma=1.8$  MeV) reproduce the formation of domains with a field slowly varying about some nonzero value  $\langle\pi\rangle$ . The “collapse” of these domains to the true ground state is accompanied by characteristic pion radiation, where the number of pions in a typical clump could be large enough for detection.

Under realistic conditions, the cooling in heavy-ion collisions may not be fast enough to sufficiently amplify the long-wavelength modes, and the wavelengths of the lightest modes may not be long enough for the formation of disoriented chiral condensates. For further references to disoriented chiral condensates, see Krzywicki (1994), Huang, (1995), Rajagopal (1995), and references therein.

A useful tool that helps in assessing the chances of observing the predicted pion radiation is an energy budget for the chiral and the deconfinement transition (Wilczek, 1994). The energy density that is stored in the chiral vacuum (associated with chiral condensation) should be compared with the energy density of the quark-gluon plasma for temperatures  $T\geq T_c$ . The decimation of QCD’s degrees of freedom to pions and the description of the cooling process as a quench are certainly idealizations. Within the idealized limits we have seen the predictive power of the renormalization group approach. As we argued in Secs. IV.A.1 and IV.A.2, the basic open question is how important are the heavier modes in comparison to the light modes associated with spontaneous chiral symmetry breaking. One measure for their importance is their contribution to the total energy density. The accumulated effects of heavier modes can bury the interesting (predictable) structures arising from the singularities in the chiral two-flavor transition. A quench does not guarantee, but improves the chance for seeing relics of the chiral transition.

Support from “experiment” is taken from a particular class of cosmic-ray events called Centauros (Rajagopal and Wilczek, 1993a, 1993b). In Centauros, isospin invariance seems indeed to be violated. The number of

charged pions strongly exceeds the number of neutral ones. For further discussion of Centauros see, for examples, Lattes *et al.* (1980).

## VI. SUMMARY AND CONCLUSIONS

The ultimate question of the order of the finite-temperature QCD transition for three colors, two light and one heavier flavor, with physical values for the quark masses, is open. Closest to this physical case come lattice simulations with two light and one heavier flavor in the fermion schemes, with staggered and Wilson fermions. At present their results are in disagreement, predicting a crossover phenomenon in the staggered formulation (Sec. III.C.5) and a first-order transition in the Wilson formulation (Sec. III.C.7). Both results may be still artifacts of the strong couplings inherent in the simulations.

Partial answers as to the order of the QCD transitions are known in limiting cases. The first [second] order of the deconfinement transition in the SU(3) [SU(2)] Yang-Mills theory seems to be well established (Secs. III.B.1 and II.B.2). General agreement is also found concerning the second order of the chiral transition in the limit of two massless flavors when dynamical fermions are included. An open question is the universality class of two-flavor QCD [whether it is  $O(2)$ ,  $O(4)$ , or mean field]. Consistent results in different fermion formulations are found for the first order of the chiral transition in the case of  $N_f\geq 3$  flavors with sufficiently light masses (Secs. III.C.5 and III.C.7).

An interesting topic for future studies is a determination of the critical strange-quark mass  $m_s^*$  at which the first order of the chiral transition changes into second order. For larger values of  $m_s$  it stays of second order for  $m_{u,d}=0$ , and disappears otherwise. The location of the “critical” phase boundary between mass regions of first-order transitions and crossover phenomena could tell us how far the physical quark masses are from “critical” quark masses. It would be appealing, both from theoretical and experimental points of view, if they were close. The powerful renormalization-group approach then becomes applicable for finite mass-scaling predictions. Remnants of a nearby second-order transition should lead to large correlation lengths and pronounced effects in relative heavy-ion collider experiments (Secs. III.C.5, III.C.7, and IV.A.4).

Lattice calculations start from first principles as they simulate QCD in its discretized version. They are *the* nonperturbative tool for studying the phase-transition region, where many expansion schemes break down. But they are plagued with artifacts from the finite volume, the finite lattice constant, and unphysical mass parameters. Infrared artifacts may amount to a harmless rounding of an infinite-volume singularity or a misleading double-peak structure, suggesting a first-order transition where it is truly of second order (Sec. II.A.2). Ultraviolet artifacts may result in a “harmless” change of critical exponents [ $O(2)$  rather than  $O(4)$ ] (Sec.

III.C.3), but also in bulk transitions being a *pure* lattice artifact without any impact on continuum physics (Sec. III.C.4).

Progress on the lattice is not only a question of powerful machines (such as special-purpose computers). Progress in the past came also from a better control of ultraviolet and infrared artifacts with finite-size and finite-mass scaling analysis, using refined criteria developed in statistical physics. Extrapolations to the infinite-volume and zero-mass limits are unavoidable in numerical simulations. Therefore, a precise measurement of critical exponents is worth the effort, assuring that the right extrapolation formulas are used (Sec. III.C.5).

The earlier discrepancies between predictions from staggered and Wilson fermions shrink. Wilson fermions seem to have a chiral limit in the confinement phase with broken chiral symmetry at low temperatures, if the lattice is large enough. The lattice size must be even larger than for staggered fermions in order to reach the continuum limit, too large for practicable simulations, unless improved actions accelerate the approach to the continuum limit. Work in this direction is in progress (Sec. III.C.7).

The main contributions from the Euclidean lattice formulation of QCD to physical applications in relativistic heavy-ion collider experiments should be reliable predictions of *in-equilibrium* properties. To these belong the equation of state, the tension of interfaces between coexisting phases, and the velocity of sound. A derivation of the equation of state under the inclusion of dynamical fermions is still in an exploratory stage. The dependence on the lattice size has not yet been analyzed, and the simulations are performed for 2 rather than 2+1 flavors, but the ingredients are fully nonperturbative (Sec. III.C.6). This represents important progress compared to earlier calculations that used perturbative and nonperturbative input data in an inconsistent way (see also Sec. III.B.2).

From results in the pure gauge sector, it has become clear that strong deviations from the phenomenological bag model equation of state must be expected in the transition region. Deviations from the ideal-gas behavior are pronounced up to temperatures of the order of 2–3  $T_c$  (Sec. III.B.2). So far, many phenomenological applications are based on the bag model equation of state. An update is called for, once the lattice results are well established.

With a reliable calculation of interface tensions, one could exclude certain scenarios, which are under discussion for RHIC experiments, right from the beginning. Thus a small value for the interface tension is incompatible with a strong supercooling scenario (Sec. III.B.3). So far lattice calculations have mainly focused on temperature driven transitions at zero baryon density, the reasons for which have been explained in Sec. III.D. Nevertheless, inclusion of a finite chemical potential is needed for a realistic description of heavy-ion experiments. An idealization of a baryon-free region in the collisions can be hardly justified.

Studies of the phase structure of QCD in effective models in the spacetime continuum may give useful hints about the underlying physical mechanisms. Results in dual Ginzburg-Landau models suggest that Abelian  $U(1) \times U(1)$  monopoles are the most important field configurations in the confinement mechanism (Sec. IV.B.2) in the maximal Abelian gauge. Confinement would then have an explanation as a dual Meissner effect, but the type of most relevant topological field configurations does depend on the gauge.

*Nonequilibrium* aspects enter the evolution of the initially created hot and dense “fireballs” in relativistic heavy-ion collider experiments and in the transition dynamics during the phase conversion. Most of the non-equilibrium approaches in the past have studied *small* deviations from in-equilibrium properties. Field-theoretical descriptions of nonequilibrium phenomena are still in their infancy. Recently, they became increasingly important for the electroweak transition in the early universe. These tools may have fruitful applications to QCD as well. For example, it is still unclear whether the conversion from the plasma to the hadron phase proceeds via nucleation or large-domain coarsening (Secs. V.D.1 and V.D.2).

For phenomenological discussions of QCD transitions in laboratory experiments, theoretical concepts are sometimes missing. Thus it seems to be rather difficult to derive the equilibration time of the hot initial state from first principles. Formulas are missing for particle production rates in media with *long-range* correlations. Such correlations exist at  $T_c$ , if the transition is almost of second order. An interpolation between nuclear structure functions and thermal Bose-Einstein or Fermi-Dirac distributions would be desirable to close the gap between the initial off-equilibrium and the final in-equilibrium medium.

Altogether the judgement of the experimental situation is the following. Even from a very conservative point of view, extended and dense systems have been observed that display some collective features. Some of them can be explained with a transient plasma, but alternative mechanisms explain the observed features as well. To date, the question of the order of the transition has been of secondary importance in heavy-ion experiments. Once the creation of a plasma phase may be taken for granted, one can focus on substructures exhibiting details of the transition dynamics (Sec. V).

Our aim was to list the most important signatures that are, at least in principle, sensitive to the transition, and, in case it is hopeless to see an effect, to argue why it is hopeless (see, for example, Sec. V.C.3).

In Table IX we summarize signatures in various experimental tools that may be indicative of a first- or second-order transition and display alternative explanations in a separate column. Crosses (×) stand for an affirmative answer, which generally should be taken cautiously, and “/” stands for “not known.” Explanations of further notations can be found in the corresponding sections.

TABLE IX. Summary of chapter V.

Observable experimental tool	Signature	Experimentally observed?	First order or sharp crossover	Idealized second order explanations	Alternative explanations	Section
$\langle p_T \rangle$ distribution of charged particles	Flattening of $\langle p_T \rangle$ vs $dN/dY$	x	x	/	e.g., lack of phase space superpositions	V.C.1
Diplet spectra	Peak in $\frac{d\langle p_T \rangle}{dM}$ in thermal rates	/	x size of interference region is larger	x smaller	of various back-grounds	V.C.2
Strangeness production	Enhancement of $(K^+/\pi^+)$ , $(K^-/\pi^-)$	x	no specific prediction although strong dependence on the scenario may be hidden		initial- and final-state interactions in a purely hadronic gas	V.C.3
Pion interferometry	Asymmetry between outwards and sideways extensions of pion source at freeze-out	/	x pronounced prolongation of lifetime	/	e.g., contamination due to resonance decays	V.C.4
$dN/dY$	Strong multiplicity fluctuations	x	x deflagrations during the phase conversion	/	e.g., fluctuations of nondynamical origin, insufficient statistics	V.C.5
Intermittency	$d_l \sim$ independent of $l$	x	/	x self-similarity	e.g., cascade mechanisms, Bose-Einstein correlations,...	V.C.6
$R = \frac{\#\pi^0}{\#(\pi^+ + \pi^-)}$	Large clusters of charged pions or neutral pions	x (Centauros)	/	x clusters in isospin space	/	V.D.3
“Color conductivity”	Change in “color resistivity”	/	x discontinuity	x continuous behavior	/	V.D.2

We conclude with the following *conjecture*. For physical quark masses, the deconfinement and the chiral transitions are replaced by rapid crossover phenomena. The rapid change is seen in the energy and entropy densities over a temperature interval of the order of 10 MeV, and in variations of the order-parameter expectations values such as the Polyakov loop and the light-quark condensate. It would be interesting to investigate on the phenomenological side when a change in one of these quantities might be considered as rapid enough to produce multiplicity fluctuations in rapidity space beyond the statistical noise, or clear deviations in dilepton production rates from the standard rates without a crossover. Is a temperature interval of 10 MeV resolvable from an experimental point of view? Finding a quantitative answer to these questions is a challenge for further studies.

#### ACKNOWLEDGMENTS

The author wishes to thank J. Kapusta and E. Shuryak for an invitation to the workshop on Strong Interactions at Finite Temperature at the ITP in Santa Barbara in August 1993, and for discussions with the participants of this workshop. I would also like to thank J. Cleymans, A. Patkós, H.-J. Pirner, K. Redlich, and H. Satz for dis-

cussions. In particular I am indebted to G. Boyd, F. Karsch, and J. Stachel for critical comments on parts of the manuscript.

#### REFERENCES

- Abbott, T., *et al.*, E802 Collaboration, 1990, Phys. Rev. Lett. **64**, 847.  
 Abbott, T., *et al.*, E802 Collaboration, 1991, Phys. Rev. Lett. **66**, 1567.  
 Abbott, T., *et al.*, E802 Collaboration, 1992, Phys. Rev. Lett. **69**, 1020.  
 Abreu, P., *et al.*, DELPHI Collaboration, 1990, Phys. Lett. B **247**, 137.  
 Aizu, K., 1970, Phys. Rev. B **2**, 754.  
 Akemi, K., M. Fujisaki, M. Okuda, Y. Tago, Ph. de Forcrand, T. Hashimoto, S. Hioki, O. Miyamura, T. Takaishi, A. Nakamura, and I. O. Stamatescu, 1993, Phys. Rev. Lett. **71**, 3063.  
 Åkesson, T., *et al.*, NA34 Collaboration, 1990, Phys. Lett. B **252**, 303.  
 Akiba, Y., *et al.*, E802 Collaboration, 1993, Phys. Rev. Lett. **70**, 1059.  
 Albanese, M., *et al.*, APE Collaboration, 1987, Phys. Lett. B **192**, 163.  
 Alexopoulos, T., *et al.*, E735 Collaboration, 1990, Phys. Rev. Lett. **64**, 991.

- Altmeyer, R., K. D. Born, M. Göckeler, R. Horsley, E. Laermann, and G. Schierholz, 1993, Nucl. Phys. B **389**, 445.
- Appelquist, T., and J. Carazzone, 1975, Phys. Rev. D **11**, 2856.
- Applegate, J. H., 1991, Nucl. Phys. A **527**, 195c.
- Applegate, J. H., C. J. Hogan, and R. J. Sherrer, 1987, Phys. Rev. D **35**, 1151.
- Asakawa, M., and K. Yazaki, 1989, Nucl. Phys. A **504**, 668.
- Bacilieri, P., *et al.*, 1988, Phys. Rev. Lett. **61**, 1545.
- Baechler, J., *et al.*, NA35 Collaboration, 1992, Nucl. Phys. A **544**, 531c.
- Bak, P., S. Krinsky, and D. Mukamel, 1976, Phys. Rev. Lett. **36**, 52.
- Baker, G., B. Nickel, and D. Meiron, 1978, Phys. Rev. D **17**, 1365.
- Banerjee, B., and R. V. Gavai, 1992, Phys. Lett. B **293**, 157.
- Banks, T., and A. Casher, 1980, Nucl. Phys. B **169**, 103.
- Banks, T., R. Myerson, and J. B. Kogut, 1977, Nucl. Phys. B **129**, 493.
- Banks, T., and A. Ukawa, 1983, Nucl. Phys. B **225** [FS9], 145.
- Barber, M. N., 1983, in *Phase Transitions and Critical Phenomena*, Vol. 8, edited by C. Domb and J. L. Lebowitz (Academic, New York), p. 145.
- Barbour, I., 1992, Nucl. Phys. B Proc. Suppl. **26**, 22.
- Barbour, I., and A. Bell, 1992, Nucl. Phys. B **372**, 385.
- Barbour, I., C. Davies, and Z. Sabeur, 1988, Nucl. Phys. B **215**, 567.
- Barbour, I., and Z. Sabeur, 1990, Nucl. Phys. B **342**, 269.
- Barrette, J., 1994, Nucl. Phys. A **566**, 411c.
- Barrette, J., *et al.*, E814 Collaboration, 1992, Phys. Rev. C **45**, 819; **46**, 312.
- Barrette, J., *et al.*, E877 Collaboration, 1995, Phys. Rev. C **51**, 3309.
- Batrouni, G. G., G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, 1985, Phys. Rev. D **32**, 2736.
- Baumann, B., and B. Berg, 1985, Phys. Lett. B **164**, 131.
- Baym, G., 1984a, Nucl. Phys. A **418**, 525c.
- Baym, G., 1984b, Phys. Lett. B **138**, 18.
- Baym, G., J. P. Blaizot, W. Czyz, B. L. Friman, and M. Soyeur, 1983, Nucl. Phys. A **407**, 541.
- Baym, G., H. Monien, C. J. Pethick, and D. G. Ravenhall, 1990, Phys. Rev. Lett. **64**, 1867.
- Berezin, F. A., 1966, *The Method of Second Quantization*, Pure and Applied Physics, Vol. 24 (Academic, New York).
- Berg, B. A., and A. Billoire, 1983, Nucl. Phys. B **221**, 109.
- Berg, B. A., and T. Neuhaus, 1991, Phys. Lett. B **267**, 249.
- Berg, B. A., and T. Neuhaus, 1992, Phys. Rev. Lett. **68**, 9.
- Bernard, C., 1974, Phys. Rev. D **9**, 3312.
- Bernard, C., M. C. Ogilvie, M. A. DeGrand, C. DeTar, S. Gottlieb, A. Krasnitz, R. L. Sugar, and D. Toussaint, 1992, Phys. Rev. D **45**, 3854.
- Bernard, C., *et al.*, 1995a, Nucl. Phys. B Proc. Suppl. **42**, 448.
- Bernard, C., T. Blum, T. A. DeGrand, C. DeTar, S. Gottlieb, U. Heller, L. Kärkkäinen, K. Rummukainen, B. Sugar, and D. Toussaint, 1995b, Nucl. Phys. B Proc. Suppl. **42**, 451.
- Bernard, V., and U. G. Meissner, 1988, Phys. Rev. D **38**, 1551.
- Bertsch, G., 1989, Nucl. Phys. A **498**, 173c.
- Bertsch, G., M. Gong, L. McLerran, V. Ruuskanen, and E. Sarkkinen, 1988, Phys. Rev. D **37**, 1202.
- Bertsch, G., M. Gong, and M. Tohyama, 1988, Phys. Rev. C **37**, 1896.
- Bervillier, C., 1976, Phys. Rev. B **14**, 4964.
- Bhanot, G., and M. Creutz, 1981, Phys. Rev. D **24**, 3212.
- Bialas, A., and J. P. Blaizot, 1985, Phys. Rev. D **32**, 2954.
- Bialas, A., and W. Czyz, 1985, Z. Phys. C **28**, 255.
- Bialas, A., and W. Czyz, 1984, Phys. Rev. D **30**, 2371.
- Bialas, A., W. Czyz, A. Dyrek, and W. Florkowski, 1988, Nucl. Phys. B **296**, 611.
- Bialas, A., and R.C. Hwa, 1991, Phys. Lett. B **253**, 436.
- Bialas, A., and R. Peschanski, 1986, Nucl. Phys. B **273**, 703.
- Bialas, A., and R. Peschanski, 1988, Nucl. Phys. B **308**, 857.
- Bijnens, J., C. Bruno, and E. de Rafael, 1993, Nucl. Phys. B **390**, 501.
- Bilić, N., K. Demeterfi, and B. Petersson, 1992, Nucl. Phys. B **377**, 651.
- Bilić, N., F. Karsch, and K. Redlich, 1992, Phys. Rev. D **45**, 3228.
- Billoire, A., 1992, Int. J. Mod. Phys. C **3**, 913.
- Billoire, A., S. Gupta, A. Irbäck, R. Lacaze, A. Morel, and B. Petersson, 1990, Phys. Rev. B **42**, 6743.
- Billoire, A., R. Lacaze, and A. Morel, 1990, Nucl. Phys. B **340**, 542.
- Billoire, A., R. Lacaze, and A. Morel, 1992, Nucl. Phys. B **370**, 773.
- Binder, K., 1981, Z. Phys. B **43**, 119.
- Binder, K., 1982, Phys. Rev. A **25**, 1699.
- Binder, K., 1987, Rep. Prog. Phys. **50**, 783.
- Binder, K., and D. P. Landau, 1984, Phys. Rev. B **30**, 1477.
- Bitar, K. M., *et al.*, 1991, Phys. Rev. D **43**, 2396.
- Bitar, K. M., *et al.*, 1993, Nucl. Phys. B Proc. Suppl. **30**, 315.
- Bitar, K. M., A. D. Kennedy, and P. Rossi, 1990, Phys. Lett. B **234**, 333.
- Bjorken, J. D., 1983, Phys. Rev. D **27**, 140.
- Bjorken, J. D., K. L. Kowalski, and C. C. Taylor, 1993, in *Observing Disoriented Chiral Condensates, Proceedings of the workshop on Physics at Current Accelerators and the Supercollider*, Argonne (1993), hep-ph/9309235, and SLAC preprint SLAC-PUB-6109 (April 1993).
- Blaizot, J.-P., and J.-Y. Ollitrault, 1989, Phys. Lett. B **217**, 386.
- Blaizot, J.-P., and J.-Y. Ollitrault, 1990, in *Quark-Gluon Plasma*, edited by R. C. Hwa (World Scientific, Singapore), p. 393.
- Blöte, H. W. J., and M. P. Nightingale, 1982, Physica **112A**, 405.
- Bloomer, M. A., *et al.*, WA80 Collaboration, 1992, Nucl. Phys. A **544**, 543c.
- Blum, T., T. A. DeGrand, C. DeTar, S. Gottlieb, A. Hasenfratz, L. Kärkkäinen, D. Toussaint, and R. L. Sugar, 1994, Phys. Rev. D **50**, 3377.
- Blum, T., S. Gottlieb, L. Kärkkäinen, and D. Toussaint, 1995a, Phys. Rev. D **51**, 5153.
- Blum, T., S. Gottlieb, L. Kärkkäinen, and D. Toussaint, 1995b, Nucl. Phys. B Proc. Suppl. **42**, 460.
- Blum, T., C. DeTar, U. M. Heller, L. Kärkkäinen, and D. Toussaint, 1995c, Nucl. Phys. B Proc. Suppl. **42**, 457.
- Bochicchio, M., L. Maiani, G. Martinelli, G. Rossi, and M. Testa, 1985, Nucl. Phys. B **262**, 331.
- Borgs, C., and R. Kotecký, 1990, J. Stat. Phys. **61**, 79.
- Borgs, C., R. Kotecký, and S. Miracle-Sole, 1991, J. Stat. Phys. **62**, 529.
- Born, K. D., E. Laermann, N. Pirsch, T. F. Walsh, and P. M. Zerwas, 1989, Phys. Rev. D **40**, 1653.
- Borrill, J., and M. Gleiser, 1995, Phys. Rev. D **51**, 4111.
- Boyanovsky, D., and C. Aragao de Carvalho, 1993, Phys. Rev. D **48**, 5850.
- Boyanovsky, D., and C. Aragao de Carvalho, 1994, Phys. Rev. D **49**, 2769.

- Boyd, G., 1995, private communication.
- Boyd, G., J. Engels, F. Karsch, E. Laermann, C. Legeland, M. Lutgemeier, and B. Petersson, 1995, *Phys. Rev. Lett.* **75**, 4169.
- Boyd, G., J. Fingberg, F. Karsch, L. Kärkkäinen, and B. Petersson, 1992, *Nucl. Phys. B* **376**, 199.
- Braaten, E., and R. D. Pisarski, 1990a, *Phys. Rev. Lett.* **64**, 1338.
- Braaten, E., and R. D. Pisarski, 1990b, *Phys. Rev. D* **42**, 2156.
- Braaten, E., and R. D. Pisarski, 1990c, *Nucl. Phys. B* **337**, 569.
- Braaten, E., and R. D. Pisarski, 1990d, *Nucl. Phys. B* **339**, 310.
- Braaten, E., and R. D. Pisarski, 1992a, *Phys. Rev. D* **45**, 1827.
- Braaten, E., and R. D. Pisarski, 1992b, *Phys. Rev. D* **46**, 1829.
- Brahm, D. E., and S. D. H. Hsu, 1991a, Caltech preprint CALT-68-1762, HUTP-91-A064, 1991.
- Brahm, D. E., and S. D. H. Hsu, 1991b, Caltech preprint CALT-68-1705, HUTP-91-A063, 1991.
- Bravina, L., L. P. Csernai, P. Levai, and D. Strottman, 1993, in *Proceedings of the 2nd International Workshop on Heavy Ion Physics at the AGS*, MIT 93, MITLNS-2158, p. 329.
- Bray, A. J., 1990, *Phys. Rev. B* **41**, 6724.
- Brézin, E., J. C. Le Guillou, and J. Zinn-Justin, 1976, in *Phase Transitions and Critical Phenomena*, Vol. 6, edited by C. Domb and M. S. Green (Academic Press, New York), p. 125.
- Brézin, E., D. J. Wallace, and K. G. Wilson, 1973, *Phys. Rev. B* **7**, 232.
- Brower, R., S. Huang, J. Potvin, and C. Rebbi, 1992, *Phys. Rev. D* **46**, 2703.
- Brown, F. R., F. P. Butler, H. Chen, N. H. Christ, Z. Dong, W. Schaffer, L. Unger, and A. Vaccarino, 1990a, *Phys. Rev. Lett.* **65**, 2491.
- Brown, F. R., F. P. Butler, H. Chen, N. H. Christ, Z. Dong, W. Schaffer, L. Unger, and A. Vaccarino, 1990b, *Phys. Lett. B* **251**, 181.
- Brown, F. R., H. Chen, N. H. Christ, Z. Dong, R. D. Mawhinney, W. Schaffer, and A. Vaccarino, 1992, *Phys. Rev. D* **46**, 5655.
- Brown, F. R., N. H. Christ, Y. Deng, M. Gao, and T. J. Woch, 1988, *Phys. Rev. Lett.* **61**, 2058.
- Brown, G. E., J. Stachel, and G. M. Welke, 1991, *Phys. Lett. B* **253**, 19.
- Bunatian, G., and J. Wambach, 1994, *Phys. Lett. B* **336**, 290.
- Buschbeck, B., and P. Lipa, 1989, *Mod. Phys. Lett. A* **4**, 1871.
- Cahn, J. W., and J. E. Hilliard, 1958, *J. Chem. Phys.* **28**, 258.
- Callan, C., R. Dashen, and D. J. Gross, 1976, *Phys. Lett. B* **63**, 334.
- Callan, C. G., Jr., and S. Coleman, 1977, *Phys. Rev. D* **16**, 1762.
- Callaway, D. J. E., and A. Rahman, 1982, *Phys. Rev. Lett.* **49**, 613.
- Callaway, D. J. E., and A. Rahman, 1983, *Phys. Rev. D* **28**, 1506.
- Campbell, B. A., J. Ellis, and K. A. Olive, 1990, *Nucl. Phys. B* **345**, 57.
- Campi, X. and H. Krivine, 1995, preprint IPNO/TH 95-13, IPN Orsay, Paris.
- Celik, T., J. Engels, and H. Satz, 1983a, *Phys. Lett. B* **125**, 411.
- Celik, T., J. Engels, and H. Satz, 1983b, *Phys. Lett. B* **129**, 323.
- Challa, M. S., D. P. Landau, and K. Binder, 1986, *Phys. Rev. B* **34**, 1841.
- Chan, L.-H., and R. W. Haymaker, 1973, *Phys. Rev. D* **7**, 415.
- Chernodub, M. N., M. I. Polikarpov, and A. I. Veselov, 1995, *Phys. Lett. B* **342**, 303.
- Chodos, A., R. L. Jaffe, K. Johnson, C. B. Thorn, and V. F. Weisskopf, 1974, *Phys. Rev. D* **9**, 3471.
- Christ, N., 1992a, *Nucl. Phys. A* **544**, 81c.
- Christ, N., 1992b, *Nucl. Phys. B Proc. Suppl.* **26**, 217.
- Clapeyron, B. P. E., 1834, *J. l'Ecole Polytechnique (Paris)* **14**, 153.
- Clausius, R., 1850, *Ann. Phys. (Leipzig)* **79**, 368, 500.
- Cleymans, J., J. Fingberg, and K. Redlich, 1987, *Phys. Rev. D* **35**, 2153.
- Cleymans, J., R. V. Gavai, and E. Suhonen, 1986, *Phys. Rep.* **130**, 217.
- Coleman, S., R. Jackiw, and H. D. Politzer, 1974, *Phys. Rev. D* **10**, 2491.
- Cooper, F., and G. Frye, 1974, *Phys. Rev. D* **10**, 186.
- Cornwall, J., R. Jackiw, and E. Tomboulis, 1974, *Phys. Rev. D* **10**, 2428.
- Creutz, M., 1981, *Phys. Rev. Lett.* **46**, 1441.
- Creutz, M., L. Jacobs, and C. Rebbi, 1983, *Phys. Rep.* **93**, 201.
- Csernai, L. P., and J. Kapusta, 1992a, *Phys. Rev. D* **46**, 1379.
- Csernai, L. P., and J. Kapusta, 1992b, *Phys. Rev. Lett.* **69**, 737.
- Damgaard, P. H., N. Kawamoto, and K. Shigemoto, 1986, *Nucl. Phys. B* **264**, 1.
- Davies, C. T. H., and E. G. Klepfish, 1991, *Phys. Lett. B* **256**, 68.
- DeGrand, T. A., and C. E. DeTar, 1983, *Nucl. Phys. B* **225**, 590.
- DeGrand, T. A., and C. E. DeTar, 1986, *Phys. Rev. D* **34**, 2469.
- DeGrand, T. A., R. L. Jaffe, K. Johnson, and J. Kiskis, 1975, *Phys. Rev. D* **12**, 2060.
- DeGrand, T. A., and D. Toussaint, 1980, *Phys. Rev. D* **22**, 2478.
- Deng, Y., 1989, *Nucl. Phys. B Proc. Suppl.* **9**, 334.
- DeTar, C. E., 1985, *Phys. Rev. D* **32**, 276.
- DeTar, C. E., 1995, *Nucl. Phys. B Proc. Suppl.* **42**, 73.
- DeTar, C. E., and T. Kunihiro, 1989, *Phys. Rev. D* **39**, 2805.
- Dong, Z., and N. H. Christ, 1992, *Nucl. Phys. B Proc. Suppl.* **26**, 314.
- Duane, S., 1985, *Nucl. Phys. B* **257** [FS14], 652.
- Duane, S., A. D. Kennedy, B. J. Pendleton, and D. Roweth, 1987, *Phys. Lett. B* **195**, 216.
- Dyakonov, D. I., and A. D. Mirlin, 1988, *Phys. Lett. B* **203**, 299.
- Eggers, H. C., and J. Rafelski, 1991, *Int. J. Mod. Phys.* **6**, 1067.
- Ejiri, S., S. Kitahara, Y. Matsubara, and T. Suzuki, 1995, *Nucl. Phys. B Proc. Suppl.* **42**, 481.
- Engels, J., J. Fingberg, and M. Weber, 1990, *Nucl. Phys. B* **332**, 737.
- Engels, J., J. Fingberg, F. Karsch, D. Miller, and M. Weber, 1990, *Phys. Lett. B* **252**, 625.
- Engels, J., F. Karsch, and K. Redlich, 1995, *Nucl. Phys. B* **435**, 295.
- Eskola, K. J., and M. Gyulassy, 1993, *Phys. Rev. C* **47**, 2329.
- Eskola, K. J., and J. Lindfors, 1990, *Z. Phys. C* **46**, 141.
- Espinosa, J. R., M. Quiros, and F. Zwirner, 1992, *Phys. Lett. B* **291**, 115.
- Fäldt, G. and B. Petersson, 1986, *Nucl. Phys. B* **264**, 197.
- Ferenc, D., 1992, private communication.
- Ferenc, D., *et al.*, NA35 Collaboration, 1992, in *Quark Matter '91: 9th International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions* [*Nucl. Phys. A* **544**, 531c].
- Ferrenberg, A. M., and R. H. Swendsen, 1988, *Phys. Rev. Lett.* **61**, 2635.
- Ferrenberg, A. M., and R. H. Swendsen, 1989, *Phys. Rev. Lett.* **63**, 1658.
- Feynman, R. P., 1981, *Nucl. Phys. B* **188**, 479.
- Fisher, M. E., 1974, *Rev. Mod. Phys.* **46**, 597.

- Fisher, M. E., and A. N. Berker, 1982, *Phys. Rev. B* **26**, 2507.
- Flory, P. J., 1941a, *J. Am. Chem. Soc.* **63**, 3083.
- Flory, P. J., 1941b, *J. Am. Chem. Soc.* **63**, 3091.
- Foley, K. Y., E810 Collaboration, 1992, in *Proc. Quark Matter '91: 9th International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions* [Nucl. Phys. A **544**, 335c].
- Frei, Z., and A. Patkós, 1989, *Phys. Lett. B* **229**, 102.
- Frei, Z., and A. Patkós, 1990, *Phys. Lett. B* **247**, 381.
- Fucito, F., and A. Vulpiani, 1982, *Phys. Lett. A* **89**, 33.
- Fukugita, M., H. Mino, M. Okawa, and A. Ukawa, 1990, *Phys. Rev. Lett.* **65**, 816.
- Fukugita, M., H. Mino, M. Okawa, and A. Ukawa, 1991, *Nucl. Phys. B Proc. Suppl.* **20**, 258.
- Fukugita, M., S. Ohta, and A. Ukawa, 1986, *Phys. Rev. Lett.* **57**, 1974.
- Fukugita, M., S. Ohta, and A. Ukawa, 1988, *Phys. Rev. Lett.* **60**, 178.
- Fukugita, M., M. Okawa, and A. Ukawa, 1989, *Phys. Rev. Lett.* **63**, 1768.
- Fukugita, M., M. Okawa, and A. Ukawa, 1990, *Nucl. Phys. B* **337**, 181.
- Fukugita, M., and A. Ukawa, 1985, *Phys. Rev. Lett.* **55**, 1854.
- Fulde, P., and H. Wagner, 1971, *Phys. Rev. Lett.* **27**, 1280.
- Fuller, G. M., G. J. Mathews, and C. R. Alcock, 1988, *Phys. Rev. D* **37**, 1380.
- Gasser, J., and H. Leutwyler, 1984, *Ann. Phys. (USA)* **158**, 142.
- Gavai, R. V., M. Grady, and M. Mathur, 1994, *Nucl. Phys. B* **423**, 123.
- Gavai, R. V., S. Gupta, A. Irbäck, F. Karsch, S. Meyer, B. Petersson, H. Satz, and H. W. Wyld, 1990, *Phys. Lett. B* **241**, 567.
- Gavai, R. V., and F. Karsch, 1985, *Nucl. Phys. B* **261**, 273.
- Gavai, R. V., and F. Karsch, 1989, *Phys. Lett. B* **233**, 417.
- Gavai, R. V., F. Karsch, and B. Petersson, 1989, *Nucl. Phys. B* **322**, 738.
- Gavai, R., and M. Mathur, 1995, *Nucl. Phys. B, Proc. Suppl.* **42**, 490.
- Gavai, R. V., J. Potvin, and S. Sanielevici, 1988, *Phys. Rev. D* **37**, 1343.
- Gavai, R. V., J. Potvin, and S. Sanielevici, 1987, *Phys. Rev. Lett.* **58**, 2519.
- Gavin, S., A. Gocksch, and R. D. Pisarski, 1994a, *Phys. Rev. D* **49**, R3079.
- Gavin, S., A. Gocksch, and R. D. Pisarski, 1994b, *Phys. Rev. Lett.* **72**, 2143.
- Gavin, S., and M. Gyulassy, 1988, *Phys. Lett. B* **214**, 241.
- Geiger, K., 1992a, *Phys. Rev. D* **46**, 4965.
- Geiger, K., 1992b, *Phys. Rev. D* **46**, 4986.
- Geiger, K., and J. Kapusta, 1993, *Phys. Rev. Lett* **70**, 1920.
- Gell-Mann, M., and M. Levy, 1960, *Nuovo Cimento* **16**, 705.
- Gell-Mann, M., R.J. Oakes, and B. Renner, 1968, *Phys. Rev.* **175**, 2195.
- Gerber, P., and H. Leutwyler, 1989, *Nucl. Phys. B* **321**, 387.
- Gibbs, P., 1986, *Phys. Lett. B* **182**, 369.
- Ginsparg, P., 1980, *Nucl. Phys. B* **170** [FS1], 388.
- Gocksch, A., 1991, *Phys. Rev. Lett.* **67**, 1701.
- Gocksch, A., U. Heller, and P. Rossi, 1988, *Phys. Lett. B* **205**, 334.
- Goldberg, H., 1983, *Phys. Lett. B* **131**, 133.
- Gonzalez-Arroyo, A., and M. Okawa, 1987, *Phys. Rev. Lett.* **58**, 2165.
- Gottlieb, S., 1991, *Nucl. Phys. B Proc. Suppl.* **20**, 247.
- Gottlieb, S., A. Krasnitz, U. M. Heller, A. D. Kennedy, J. B. Kogut, R. L. Renken, D. K. Sinclair, R. L. Sugar, D. Toussaint, and K. C. Wang, 1993, *Phys. Rev. D* **47**, 3619.
- Gottlieb, S., W. Liu, R. L. Renken, R. L. Sugar, and D. Toussaint, 1987a, *Phys. Rev. D* **35**, 2531.
- Gottlieb, S., W. Liu, R. L. Renken, R. L. Sugar, and D. Toussaint, 1987b, *Phys. Rev.* **35**, 3972.
- Gottlieb, S., W. Liu, D. Toussaint, R. L. Renken, and R. L. Sugar, 1987c, *Phys. Rev. Lett.* **59**, 1881.
- Gottlieb, S., W. Liu, R. Renken, R. Sugar, and D. Toussaint, 1989, *Phys. Rev. Lett.* **40**, 2389.
- Gottlieb, S., W. Liu, R. L. Renken, R. L. Sugar, and D. Toussaint, 1991, *Phys. Rev. D* **41**, 622.
- Griffiths, R. B., 1965, *J. Math. Phys.* **8**, 478.
- Gross, D. J., R. D. Pisarski, and L. G. Yoffe, 1981, *Rev. Mod. Phys.* **53**, 43.
- Grossmann, B., and M. L. Laursen, 1993, *Nucl. Phys. B* **408**, 637.
- Grossmann, B., M. L. Laursen, T. Trappenberg, and U.-J. Wiese, 1992, *Phys. Lett. B* **293**, 175.
- Gupta, S., 1990, *Phys. Lett. B* **248**, 453.
- Gupta, S., and H. Satz, 1992, *Phys. Lett. B* **283**, 439.
- Gyulassy, M., 1984, *Nucl. Phys. A* **418**, 59c.
- Gyulassy, M., K. Kajantie, H. Kurki-Suonio, and L. McLerran, 1984, *Nucl. Phys. B* **237**, 477.
- Hagedorn, R., 1983, *Riv. Nuovo Cimento* **6**, 1.
- Hagedorn, R., 1985, in *Quark Matter '84, Lecture Notes in Physics* **211** (Springer, Berlin), p. 53.
- Hajduković, D., and H. Satz, 1992, "Does the one-dimensional Ising model show intermittency?," preprint CERN-TH.6674/92 and BI-TP 92/43.
- Halperin, B. I., P. C. Hohenberg, and S.-K. Ma, 1974, *Phys. Rev. B* **10**, 139.
- Hanbury-Brown, R., and R. Q. Twiss, 1954, *Philos. Mag.* **45**, 633.
- Hasenfratz, P., and F. Karsch, 1983, *Phys. Lett. B* **125**, 308.
- Hasenfratz, P., F. Karsch, and I. O. Stamatescu, 1983, *Phys. Lett. B* **133**, 221.
- Hasenfratz, A., and D. Toussaint, 1992, *Nucl. Phys. B* **371**, 539.
- Hatsuda, T., 1992, in *Quark Matter '91: 9th International Conference on Ultra-Relativistic Nucleus-Nucleus Collision* [Nucl. Phys. A **544**, 27c].
- Hatsuda, T., and T. Kunihiro, 1985, *Progr. Theor. Phys.* **74**, 765.
- Hatsuda, T., and T. Kunihiro, 1987, *Phys. Lett. B* **198**, 126.
- Heinz, U. 1986, *Ann. Phys. (N.Y.)* **168**, 148.
- Heinz, U., and K. S. Lee, 1992, *Nucl. Phys. A* **544**, 503c.
- Herrmann, H. J., and F. Karsch, 1991, Eds., *Fermion Algorithms, Workshop on Fermion Algorithms* (World Scientific, Singapore).
- Hioki, S., S. Kitahara, S. Kiura, Y. Matsubara, O. Miyamura, S. Ohno, and T. Suzuki, 1991, *Phys. Lett. B* **272**, 326.
- Hoek, J., 1990, *Nucl. Phys. B* **339**, 732.
- Hohenberg, P. C., A. Aharony, B. I. Halperin, and E. Siggia, 1976, *Phys. Rev. B* **13**, 2986.
- Hohenberg, P. C., and B. I. Halperin, 1977, *Rev. Mod. Phys.* **49**, 435.
- 't Hooft, G., 1981, *Nucl. Phys. B* **190**, 455.
- Horsley, R., and W. Schoenmaker, 1986, *Phys. Rev. Lett.* **57**, 2894.

- Hosoya, A., M.-A. Sakagami, and M. Takao, 1984, *Ann. Phys. (USA)* **154**, 229.
- Hosoya, A., and K. Kajantie, 1985, *Nucl. Phys. B* **250**, 666.
- Huang, S., J. Potvin, C. Rebbi, and S. Sanielevici, 1990, *Phys. Rev. D* **42**, 2864; **43**, 2056. (E).
- Huang, Z., 1995, preprint Berkeley, LBL-36740 and hep-ph/9501366.
- Hubbard, J., 1959, *Phys. Rev. Lett.* **3**, 77.
- Hüfner, J., Y. Kurihara, and H. J. Pirner, 1988, *Phys. Lett. B* **215**, 218.
- Hwa, R. C., and K. Kajantie, 1985, *Phys. Rev. D* **32**, 1109.
- Iacobson, H. H., and D. J. Amit, 1981, *Ann. Phys. (USA)* **133**, 57.
- Ilgenfritz, E. M., and E. V. Shuryak, 1989, *Nucl. Phys. B* **319**, 511.
- Imry, Y., 1980, *Phys. Rev. B* **21**, 2042.
- Iwai, J., I. Ohta, Y. Muraki, T. Ogata, S. Toyoda, R. Ihara, F. Fumuro, T. Yanagita, Y. Takahashi, and J. N. Capdevielle, 1982, *Nuovo Cimento A* **69**, 295.
- Iwai, J., N. Suzuki, and Y. Takahashi, 1976, *Prog. Theor. Phys.* **55**, 1537.
- Iwasaki, Y., 1983, Tsukuba Preprint UTHEP-118, 1983.
- Iwasaki, Y., 1995, *Nucl. Phys. B Proc. Suppl.* **42**, 96.
- Iwasaki, Y., K. Kanaya, L. Kärkkäinen, K. Rummukainen, and T. Yoshié, 1994, *Phys. Rev. D* **49**, 3540.
- Iwasaki, Y., K. Kanaya, S. Kaya, S. Sakai, and T. Yoshié, 1995, *Nucl. Phys. B Proc. Suppl.* **42**, 499.
- Iwasaki, Y., K. Kanaya, S. Sakai, and T. Yoshié, 1992, *Phys. Rev. Lett.* **69**, 21.
- Iwasaki, Y., K. Kanaya, S. Sakai, and T. Yoshié, 1993, *Nucl. Phys. B Proc. Suppl.* **30**, 327.
- Iwasaki, Y., K. Kanaya, S. Sakai, and T. Yoshié, 1995, *Nucl. Phys. B Proc. Suppl.* **42**, 502.
- Iwasaki, Y., K. Kanaya, T. Yoshié, T. Hoshino, T. Shirakawa, Y. Oyanagi, S. Ichii, and T. Kawai (QCDPAX Collaboration), 1991, *Phys. Rev. Lett.* **67**, 3343.
- Iwasaki, Y., K. Kanaya, T. Yoshié, T. Hoshino, T. Shirakawa, Y. Oyanagi, S. Ichii, and T. Kawai (QCDPAX Collaboration), 1992, *Phys. Rev. D* **46**, 4657.
- Iwasaki, Y., Y. Tsuboi, and T. Yoshié, 1989, *Phys. Lett. B* **220**, 602.
- Jackson, J. L., 1960, *Phys. Fluids* **3**, 928.
- Jacob, M., and H. Satz, 1982, Eds., *Quark Matter Formation and Heavy Ion Collisions* (World Scientific, Singapore).
- Jain, V., 1993, *Nucl. Phys. B* **394**, 707.
- Janke, W., 1992, *Int. J. Mod. Phys. C* **3**, 1137.
- Johnson, K., 1975, *Acta Phys. Polonica B* **6**, 865.
- Kadanoff, L. P., and G. Baym, 1962, *Quantum Statistical Mechanics* (Benjamin/Cummings, Menlo Park, CA).
- Kajantie, K., 1992, *Int. J. Mod. Phys. C* **3**, 1137.
- Kajantie, K., L. Kärkkäinen, and K. Rummukainen, 1990, *Nucl. Phys. B* **333**, 100.
- Kajantie, K., L. Kärkkäinen, and K. Rummukainen, 1991, *Nucl. Phys. B* **357**, 693.
- Kajantie, K., and H. Kurki-Suonio, 1986, *Phys. Rev. D* **34**, 1719.
- Kajantie, K., and T. Matsui, 1985, *Phys. Lett. B* **164**, 373.
- Kajantie, K., and H. I. Miettinen, 1981, *Z. Phys. C* **9**, 341.
- Kajantie, K., R. Raitio, and P. V. Ruuskanen, 1983, *Nucl. Phys. B* **222**, 152.
- Kanki, T., 1988, Osaka University preprint OSGE 86-24.
- Kapusta, J., 1979, *Nucl. Phys. B* **148**, 461.
- Kapusta, J., and C. Gale, 1987, *Phys. Rev. C* **35**, 2107.
- Kapusta, J., P. Lichard, and D. Seibert, 1991, *Phys. Rev. D* **44**, 2774.
- Kapusta, J., P. Lichard, and D. Seibert, 1992, *Nucl. Phys. A* **544**, 485c.
- Kapusta, J., L. McLerran, and D. K. Srivastawa, 1992, *Phys. Lett. B* **283**, 145.
- Kapusta, J., and A. Mekjian, 1986, *Phys. Rev. D* **33**, 1304.
- Kärkkäinen, L., P. Lacock, D.E. Miller, B. Petersson, and T. Reisz, 1992, *Phys. Lett. B* **282**, 121.
- Kärkkäinen, L., P. Lacock, D. E. Miller, B. Petersson, and T. Reisz, 1994, *Nucl. Phys. B* **418**, 3.
- Kärkkäinen, L., P. Lacock, B. Petersson, and T. Reisz, 1993, *Nucl. Phys. B* **395**, 733.
- Karsch, F., 1982, *Nucl. Phys. B* **205**[FS5], 285.
- Karsch, F., 1990, *Nucl. Phys. B Proc. Suppl.* **15**, 157.
- Karsch, F., 1992, private communication.
- Karsch, F., 1993, in *QCD 20 Years Later*, edited by P. M. Zerwas and H. A. Kastrup (World Scientific, Singapore), Vol. 2, p. 717 and references therein.
- Karsch, F., 1994, *Phys. Rev. D* **49**, 3791.
- Karsch, F., 1995, *Nucl. Phys. A* **590**, 367c.
- Karsch, F., and E. Laermann, 1994, *Phys. Rev. D* **50**, 6954.
- Karsch, F., and K.-H. Mütter, 1989, *Nucl. Phys. B* **313**, 541.
- Karsch, F., and H. W. Wyld, 1987, *Phys. Rev. D* **35**, 2518.
- Kataja, M., J. Letessier, P.V. Ruuskanen, and A. Tounsi, 1992, *Z. Phys. C* **55**, 153.
- Keldysk, L. V., 1964, *Zh. Eksp. Teor. Fiz.* **47**, 1515.
- Keldysk, L. V., 1965, *Sov. Phys. JETP* **20** 1018.
- Kennedy, A. D., J. Kuti, S. Meyer, and B. J. Pendleton, 1985, *Phys. Rev. Lett.* **54**, 87.
- Kitahara, S., Y. Matsubara, and T. Suzuki, 1995a, *Prog. Theor. Phys.* **93**, 1
- Kitahara, S., Y. Matsubara, and T. Suzuki, 1995b, *Nucl. Phys. B (Proc. Suppl.)* **42**, 511.
- Klaetke, J.-U. and K.-H. Mütter, 1990, *Nucl. Phys. B* **342**, 764.
- Klimt, S., M. Lutz, and W. Weise, 1990, *Phys. Lett. B* **249**, 386.
- Ko, C. M., 1995, private communication.
- Ko, C. M., and M. Asakawa, 1994, *Nucl. Phys. A* **566**, 447c.
- Koch, P., B. Müller, and J. Rafelski, 1986a, *Z. Phys. A* **324**, 453.
- Koch, P., B. Müller, and J. Rafelski, 1986b, *Phys. Rep.* **142**, 167.
- Kocić, A., and J. B. Kogut, 1995, *Phys. Rev. Lett.* **74**, 3109.
- Kocić, A., J. B. Kogut, and M.-P. Lombardo, 1993, *Nucl. Phys. B* **398**, 376.
- Kogut, J. B., 1983, *Rev. Mod. Phys.* **55**, 775.
- Kogut, J. B., H. Matsuoka, M. Stone, H. W. Wyld, S. Shenker, J. Shigemitsu, and D. K. Sinclair, 1983, *Phys. Rev. Lett.* **51**, 869.
- Kogut, J. B., J. Polonyi, H. W. Wyld, and D. K. Sinclair, 1985; *Phys. Rev. Lett.* **54**, 1475.
- Kogut, J. B., and D. K. Sinclair, 1988a, *Nucl. Phys. B* **295** [FS21], 465.
- Kogut, J. B., and D. K. Sinclair, 1988b, *Phys. Rev. Lett.* **60**, 1250.
- Kogut, J. B., and D. K. Sinclair, 1989, *Phys. Lett. B* **229**, 107.
- Kogut, J., and L. Susskind, 1975, *Phys. Rev. D* **11**, 395.
- Kronfeld, A. S., M. L. Laursen, G. Schierholz, and U.-J. Wiese, 1987, *Phys. Lett.* **198** B, 516.
- Kronfeld, A. S., G. Schierholz, and U. -J. Wiese, 1987, *Nucl. Phys. B* **293**, 461.

- Krzywicki, A., 1994, "Disoriented Chiral Condensates," talk at 29th Rencontres de Moriond, LPTHE-ORSAY-94-43, and hep-ph/9405244.
- Kunihiro, T., 1989, Phys. Lett. B **219**, 363.
- Kunihiro, T., 1991, Nucl. Phys. B **351**, 593.
- Kusaka, K., and W. Weise, 1992, Z. Phys. A **343**, 229.
- Laermann, E., G. Boyd, J. Engels, F. Karsch, C. Legeland, M. Lütgemeier, and B. Petersson, 1995, Nucl. Phys. B Proc. Suppl. **42**, 120.
- Landau, L. D., and E. M. Lifshitz, 1958, *Statistical Physics* (Pergamon, Oxford).
- Landau, L. D., and E. M. Lifshitz, 1959, *Course of Theoretical Physics Vol. 6, Fluid Mechanics* (Pergamon, London).
- Landsman, N. P., 1989, Nucl. Phys. B **322**, 498.
- Langer, J. S., 1969, Ann. Phys. (USA) **54**, 258.
- Langer, J. S., and L. A. Turski, 1973, Phys. Rev. A **8**, 3230.
- Lattes, C. M. G., Y. Fujimoto, and S. Hasegawa, 1980, Phys. Rep. **65**, 151.
- Lawrie, I. D., and S. Sarbach, 1984, in *Phase Transitions and Critical Phenomena*, Vol. 9, edited by C. Domb and J. L. Lebowitz (Academic, New York), p. 2.
- Lee, J., and J. M. Kosterlitz, 1991, Phys. Rev. B **43**, 3265.
- Leutwyler, H., 1988, Nucl. Phys. B (Proc. Suppl.) **4**, 248.
- Leutwyler, H., 1992, Phys. Lett. B **284**, 106.
- Lipatov, A., 1977, Sov. Phys.-JETP **72**, 411.
- Ludlam, T., and N. P. Samios, 1988, Z. Phys. C **38**, 353.
- Luttinger, J. M., and J. C. Ward, 1960, Phys. Rev. **118**, 1417.
- Lutz, M., S. Klimt, and W. Weise, 1992, Nucl. Phys. A **542**, 521.
- Ma, S.-K., 1976, *Modern Theory of Critical Phenomena* (Benjamin, Reading, MA).
- Makhlin, A. N., and Yu. M. Sinyukov, 1988, Z. Phys. C **39**, 69.
- Malmström, G., and D. Geldart, 1980, Phys. Rev. B **21**, 1133.
- Manohar, A., and H. Georgi, 1984, Nucl. Phys. B **234**, 189.
- Margetis, S., *et al.*, NA49 Collaboration, 1995, Nucl. Phys. A **590**, 335c.
- Matsubara, Y., S. Ejiri, and T. Suzuki, 1994, Nucl. Phys. B Proc. Suppl. **34**, 176.
- Matsubara, Y., S. Ilyar, T. Okude, K. Yotsuji, and T. Suzuki, 1995, Nucl. Phys. B, Proc. Suppl. **42**, 529.
- Matsui, T., and H. Satz, 1986, Phys. Lett. B **178**, 416.
- Mawhinney, R. D., 1993, Nucl. Phys. B (Proc. Suppl.) **30**, 331.
- McLerran, L. D., and T. Toimela, 1985, Phys. Rev. D **31**, 545.
- McLerran, L. D., 1986, Rev. Mod. Phys. **58**, 1021.
- Metropolis, N. A., M. N. Rosenbluth, A. H. Rosenbluth, E. Teller, and J. Teller, 1953, J. Chem. Phys. **21**, 1087.
- Meyer, B. S., C. R. Alcock, G. J. Mathews, and G. M. Fuller, 1991, Phys. Rev. D **43**, 1079.
- Meyer-Ortmanns, H., 1989, in *Proceedings of the XXIIIrd International Symposium on the Theory of Elementary Particles*, edited by E. Wieczorek (Ahrenshoop, GDR).
- Meyer-Ortmanns, H., H.-J. Pirner, and A. Patkós, 1992, Phys. Lett. B **295**, 255.
- Meyer-Ortmanns, H., and B.-J. Schaefer, 1996, Phys. Rev. D **53**, 6586.
- Michel, L., 1980, Rev. Mod. Phys. **52**, 617.
- Miyamura, O., 1995, Nucl. Phys. B Proc. Suppl. **42**, 538.
- Montvay, I., and G. Münster, 1994, *Quantum Fields on the Lattice*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge).
- Moriarty, K. J. M., 1981, Phys. Lett. B **106**, 130.
- Mott, N. F., 1968, Rev. Mod. Phys. **40**, 677.
- Müller, B., and J. Rafelski, 1982, Phys. Rev. Lett. **48**, 1066.
- Narison, S., 1989, *QCD Spectral Sum Rules*, Lecture Notes on Physics, Vol. 26 (World Scientific, Singapore).
- Newman, T. J., A. J. Bray, and M. A. Moore, 1990, Phys. Rev. B **42**, 4514.
- Nielsen, H. B., and M. Ninomiya, 1981a, Nucl. Phys. B **185**, 20.
- Nielsen, H. B., and M. Ninomiya, 1981b, Phys. Lett. **105B**, 219.
- Nienhuis, B., and M. Nauenberg, 1975, Phys. Rev. Lett. **35**, 477.
- Nowak, M. A., J. J. M. Verbaarschot, and I. Zahed, 1989, Nucl. Phys. B **325**, 581.
- Ochs, W., and J. Wosiek, 1988, Phys. Lett. B **214**, 617.
- Ohta, S., and S. Kim, 1991, Phys. Rev. D **44**, 504.
- Parisi, G., and Y.-S. Wu, 1981, Sci. Sin. **24**, 483.
- Patel, A., 1984a, Nucl. Phys. B **243**, 411.
- Patel, A., 1984b, Phys. Lett. B **139**, 394.
- Paterson, A. J., 1981, Nucl. Phys. B **190** [FS3], 188.
- Patkós, A., 1991, Nucl. Phys. B **365**, 243.
- Peitzmann, T., *et al.*, WA80 Collaboration, 1992, Z. Phys. C **53**, 225.
- Peitzmann, T., *et al.*, WA80 Collaboration, 1993a, in *Borlange Quark Matter* (Borlange, 1993), p. 519.
- Peitzmann, T., *et al.*, WA80 Collaboration, 1993b, Z. Phys. C **59**, 127.
- Petcher, D. N., and D. H. Weingarten, 1981, Phys. Lett. B **99**, 333.
- Pfeuty, P., and G. Toulouse, 1977, *Introduction to the Renormalization Group and Critical Phenomena* (Wiley, New York).
- Pisarski, R. D., and F. Wilczek, 1984, Phys. Rev. D **29**, 338.
- Polonyi, J., and K. Szlachanyi, 1982, Phys. Lett. B **110**, 395.
- Polyakov, A. M., 1975, Phys. Lett. B **59**, 82.
- Potvin, J., and C. Rebbi, 1989, Phys. Rev. Lett. **62**, 3062.
- Potvin, J., and C. Rebbi, 1991, Nucl. Phys. B, Proc. Suppl. **20**, 317.
- Pratt, S., 1984, Phys. Rev. Lett. **53**, 1219.
- Pratt, S., 1986, Phys. Rev. D **33**, 1314.
- Privman, V., and M. E. Fisher, 1983, J. Stat. Phys. **33**, 385.
- Rafelski, J., 1982, Phys. Rep. **88**, 331.
- Rafelski, J., 1991, Phys. Lett. **262B**, 333.
- Rafelski, J., and B. Müller, 1982, Phys. Rev. Lett. **48**, 1066.
- Rajagopal, K., 1995, "The chiral phase transition in QCD: Critical phenomena and long-wavelength pion oscillations," to appear in *Quark-Gluon Plasma 2*, edited by R. Hwa, (World Scientific, Singapore) and hep-ph-9504310.
- Rajagopal, K., and F. Wilczek, 1993a, Nucl. Phys. B **399**, 395.
- Rajagopal, K., and F. Wilczek, 1993b, Nucl. Phys. B **404**, 577.
- Redlich, K., R. Baier, H. Nakkagawa, and A. Niégawa, 1992, Nucl. Phys. A **544**, 511c.
- Redlich, K., and H. Satz, 1986, Phys. Rev. D **33**, 3747.
- Redlich, K., H. Satz, and J. Cleymans, 1992, Nucl. Phys. A **544**, 509c.
- Reisz, T., 1992, Z. Phys. C **53**, 169.
- Reisz, T., and B. Petersson, 1991, Nucl. Phys. B **353**, 757.
- Rothe, H., 1992, *Lattice Gauge Theories* (World Scientific, Singapore).
- Rummukainen, K., 1995, Nucl. Phys. B Proc. Suppl. **42**, 505.
- Ruuskanen, P. V., 1990, in *Quark-Gluon Plasma*, edited by R. Hwa (World Scientific, Singapore), p. 519.
- Ruuskanen, P. V., 1992, Nucl. Phys. A **544**, 169c.
- Sarabura, M., *et al.*, NA44 Collaboration, 1992, in *Quark Matter '91: 9th International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions* [Nucl. Phys. A **544**, 125c].
- Satz, H., 1989, Nucl. Phys. B **326**, 613.

- Satz, H., 1990a, in *Proceedings Large Hadron Collider Workshop*, edited by G. Jarlskog and D. Rein (CERN, Geneva), Vol. 1, p. 188.
- Satz, H., 1990b, Phys. Lett. B **242**, 107.
- Schechter, J., 1980, Phys. Rev. D **21**, 3393.
- Schmidt, H. R., and J. Schukraft, 1993, J. Phys. G **19**, 1705.
- Schmidt, M., G. Röpke, and H. Schulz, 1990, Ann. Phys. (USA) **202**, 57.
- Schoenmaker, W., and R. Horsley, 1988, Nucl. Phys. B Proc. Suppl. **4**, 318.
- Schramm, D. N., B. Fields, and D. Thomas, 1992, Nucl. Phys. A **544**, 267c.
- Schwinger, J., 1961, J. Math. Phys. **2**, 407.
- Seibert, D., 1990, Phys. Rev. D. **41**, 3381.
- Seiler, E., 1982, *Gauge Theories as a Problem of Constructive Quantum Field Theory and Statistical Mechanics* (Springer, Heidelberg).
- Selikhov, A. V., and M. Gyulassy, 1993, Phys. Lett. B **316**, 373.
- Seyboth, P., *et al.*, NA35 Collaboration, 1992, in *Quark Matter '91: 9th International Conference on Ultra-Relativistic Nucleus-Nucleus Collisions* [Nucl. Phys. A **544**, 293c].
- Shaposhnikov, M. E., 1992, Phys. Lett. B **277**, 324; **282**, 483 (E).
- Shen, Y., 1993, Phys. Lett. B **315**, 146.
- Shiba, H., and T. Suzuki, 1994a, Phys. Lett. B **333**, 461.
- Shiba, H., and T. Suzuki, 1994b, Nucl. Phys. B Proc. Suppl. **20**, 236.
- Shiba, H., and T. Suzuki, 1995, Phys. Lett. B **343**, 315.
- Shuryak, E. V., 1978, Phys. Lett. B **78**, 150.
- Shuryak, E. V., 1979, Phys. Rep. **61**, 71.
- Shuryak, E. V., 1984, Phys. Rep. **115**, 151.
- Shuryak, E. V., 1988, *The QCD Vacuum, Hadrons and the Superdense Matter* (World Scientific, Singapore).
- Shuryak, E. V., 1992, Phys. Rev. Lett. **68**, 3270.
- Shuryak, E. V., and L. Xiong, 1993, Phys. Rev. Lett. **70**, 2241.
- Sorge, H., H. Stöcker, and W. Greiner, 1989, Nucl. Phys. A **498**, 567c.
- Stachel, J., 1993, private communication.
- Stachel, J., E814 Collaboration, 1994, Nucl. Phys. A **566**, 183c.
- Stachel, J., and G. R. Young, 1992, Annu. Rev. Nucl. Part. Sci. **42**, 537.
- Stack, J. D., and R. J. Wensley, 1992, Nucl. Phys. B **371**, 597.
- Stanley, H. E., 1971, *An Introduction to Phase Transitions and Critical Phenomena* (Oxford University Press, Oxford).
- Stratonovich, R. L., 1958, Sov. Phys. Dokl. **2**, 416.
- Susskind, L., 1977, Phys. Rev. D **16**, 3031.
- Suzuki, T., 1993, Nucl. Phys. B Proc. Suppl. **30**, 176.
- Suzuki, T., and H. Shiba, 1995, Nucl. Phys. B Proc. Suppl. **42**, 282.
- Suzuki, T., and I. Yotsuyanagi, 1990, Phys. Rev. D **42**, 4257.
- Suzuki, T., and I. Yotsuyanagi, 1991, Nucl. Phys. B Proc. Suppl. **20**, 236.
- Svetitsky, B., 1986, Phys. Rep. **132**, 1.
- Svetitsky, B., and L. G. Yaffe, 1982a, Phys. Rev. D **26**, 963.
- Svetitsky, B., and L. G. Yaffe, 1982b, Nucl. Phys. B **210** [FS6], 423.
- Toledano, J.-C., 1981, Ferroelectrics **35**, 31.
- Toussaint, D., 1988, "Introduction to algorithms for Monte Carlo simulations and their application to QCD"; lectures presented at the Symposium on New Developments in Hardware and Software for Computational Physics (Buenos Aires), unpublished.
- Toussaint, D., 1990, Nucl. Phys. B Proc. Suppl. **17**, 248.
- Trappenberg, T., and U. J. Wiese, 1992, Nucl. Phys. B **372**, 703.
- Trincherro, R. C., 1983, Nucl. Phys. B **227**, 61.
- Turski, L. A., and J. S. Langer, 1980, Phys. Rev. A **22**, 2189.
- van Hecke, H., NA34 Collaboration, 1991, Nucl. Phys. A **544**, 27c.
- van Hove, L., 1982, Phys. Lett. **118** B, 138.
- van Hove, L., 1983, Z. Phys. C **21**, 93.
- van Hove, L., 1985, Z. Phys. C **27**, 135.
- van Kampen, N. G., 1964, Phys. Rev. A **135**, 362.
- van Leeuwen, J. M. J., 1975, in *Fundamental Problems in Statistical Mechanics*, Vol. 3, edited by E. D. G. Cohen (North-Holland, Amsterdam), p. 81.
- van Neerven, W. L., 1995, Int. J. Mod. Phys. A **10**, 2921.
- Venugopalan, R., M. Prakash, M. Kataja, and V. Ruuskanen, 1994, Nucl. Phys. A **566**, 473c.
- Vladikas, A., 1988, Nucl. Phys. B Proc. Suppl. **4**, 322.
- Waldhauser, R., D. H. Rischke, U. Katscher, J. A. Maruhn, H. Stöcker, and W. Greiner, 1992, Z. Phys. C **54**, 459.
- Wegner, F. J., 1972, Phys. Rev. B **5**, 4529.
- Weinberg, S., 1966, Phys. Rev. Lett. **17**, 616.
- Weinberg, S., 1967, Phys. Rev. Lett. **18**, 188.
- Weinberg, S., 1979, Physica A **96**, 327.
- Weldon, H. A., 1991, Phys. Rev. Lett. **66**, 293.
- Welke, G. M., R. Venugopalan, and M. Prakash, 1990, Phys. Lett. B **245**, 137.
- Werner, K., 1989, Phys. Rev. Lett. **62**, 2460.
- Wieand, K. L., S. E. Pratt, and A. B. Balantekin, 1992, Phys. Lett. B **274**, 7.
- Wilczek, F., 1992, Int. J. Mod. Phys. A **7**, 3911.
- Wilczek, F., 1994, Nucl. Phys. A **566**, 123C.
- Wilson, K., 1974, Phys. Rev. D **14**, 2455.
- Wilson, K., and J. Kogut, 1974, Phys. Rep. C **12**, 75.
- Witten, E., 1979, Nucl. Phys. B **156**, 269.
- Wroblewski, A., 1985, Acta Phys. Pol. B **16**, 379.
- Xu, N., E814 Collaboration, 1994, Nucl. Phys. A **566**, 585c.
- Zel'dovich, Ya. B., *et al.*, 1987, Usp. Fiz. Nauk **152**, 3 [Sov. Phys. Usp. **30**, 353 (1987)].
- Zhang, W.-M., and L. Wilets, 1992, Phys. Rev. C **45**, 1900.
- Zubarev, D. N., 1974, *Nonequilibrium Statistical Thermodynamics* (Plenum, New York).