Stability Analysis of Asynchronous States in Neuronal Networks with Conductance-Based Inhibition

Christian Leibold

Institute for Theoretical Biology, Humboldt-Universität zu Berlin, 10115 Berlin, Germany
(Received 17 June 2004; published 11 November 2004)

Oscillations in networks of inhibitory interneurons have been reported at various sites of the brain and are thought to play a fundamental role in neuronal processing. This Letter provides a self-contained analytical framework that allows numerically efficient calculations of the population activity of a network of conductance-based integrate-and-fire neurons that are coupled through inhibitory synapses. Based on a normalization equation this Letter introduces a novel stability criterion for a network state of asynchronous activity and discusses its perturbations. The analysis shows that, although often neglected, the reversal potential of synaptic inhibition has a strong influence on the stability as well as the frequency of network oscillations.

       PACS numbers: 87.18.Sn, 05.45.Xt

Mutual synchronization is found in a large variety of populations of biological oscillators [1]. Famous examples are the synchronous flashing of fireflies and crickets that chirp in unison. On the cellular level we know about synchronization of cardiac and circadian pacemakers as well as collective firing in neuronal networks. It is thus not surprising that, although the basics of synchronization of pulse-coupled oscillators can be considered as understood (cf. [2]), a more detailed evaluation of a specific system heavily depends on the peculiarities of its dynamics as well as the biophysical mechanism of interaction.

Theories of brain function attribute many fundamental roles to the (de)synchronization of a population of spiking neurons and suffer, in particular, from the gap between a wealth of experimental data and the unfeasiblity of an analytical treatment of collective phenomena in networks of elaborately modeled neurons. From a present perspective, one can distinguish two major classes of network models, the first of which is mainly dedicated to computer simulations of networks of realistically modeled neurons (cf. [3]), whereas the second one aims at an analytical understanding of network dynamics based on simplified neuron models [4–12]. The approach pursued in this Letter is based on the stochastic dynamics of a macroscopic activity variable and has the potential to combine advantages of both approaches mentioned above in that it can analytically infer the effects of complex single cell properties on network behavior. The dynamics has first been introduced by Gerstner [4] and, although mathematically equivalent, the present advancement extends its applicability to conductance-based integrate-and-fire neurons and hence provides an analytical understanding of the effect of changes [13] of synaptic reversal potentials on network activity. Reversal potentials are crucial especially for inhibitory-coupled networks.

A further focus is put on alternative numerical and analytical methods that permit efficient use of the activity dynamics that is defined through an integral equation. These methods may extend the applicability of the stochastic network theory also to pulse-coupled systems beyond neuroscience.

The primary assumption of the theory is that one considers a homogeneously interconnected network of identical cells with each single neuron receiving identical external input. Any particular neuron is *either* able to fire *or* is in a refractory state where it is unable to fire. After having fired a spike, a neuron becomes refractory for a period of $\tau_{\mathrm{ref}} \geq 0$.

In the present description, network dynamics is derived from the so-called "survivor function" $S(t, t')$ [4,9] that denotes the *conditional probability* that a neuron does *not* fire a spike in the interval $]t', t]$ given its last spike occurred at t'. The transition of cells into the refractory state is described by the firing probability density $p_F(t, t')$, which is defined as the relative decay of survivors, $p_F(t, t') = -S^{-1}(t, t') \frac{dS(t,t')}{dt}$. Since $S(t, t) = 1$, the dynamics of S can be specified in closed form, $S(t, t') = \exp[-\int_{t'}^{t} ds\, p_F(s, t')]$.

The population activity $A(t)$ of a network is defined to be the number of firing neurons per time. For a network of N neurons, taking the limit $N \to \infty$, $A(t)/N$ is the probability density of a cell to fire at time t. Using Bayes's theorem one henceforth constructs the compound probability density $\rho(t, t') = A(t')S(t, t')$ that a cell has fired its last spike at time $t' < t$. Being a compound probability density, ρ is normalized under the *proposition* that every cell in the network has at least fired once at some time $t' > -\infty$, which yields the normalization condition

$$N \equiv \mathcal{N}[A](t) = \int_{-\infty}^{t} dt'\, \rho(t, t'). \tag{1}$$

It is remarkable that the normalization condition (1) turns out to be sufficient to fully assess all dynamical aspects of the network activity. Nevertheless, for numerical reasons, we differentiate (1) with respect to t (see [4])

    0031-9007/04/93(20)/208104(4)$22.50

and also obtain a common integral-equation dynamics for the network activity

$$A(t) = \int_{-\infty}^{t} dt' \rho(t, t') p_F(t, t').$$ (2)

For low activity, and hence low p_F, numerical treatment of Eq. (2) causes normalization problems [9]. Here, this difficulty is overcome by a generalization of an idea from Eggert and van Hemmen [9]. One therefore assumes that there is a neuronal memory time τ_{mem}, where $p_F^{\infty}(t) = p_F(t, t')$ is independent of t' if $t - t' > \tau_{\text{mem}}$; i.e., the cells "forget" about the last spike after a period of *at most* τ_{mem}. Hence, Eq. (2) is approximated by $A(t) \approx \int_{t-\tau_{\text{mem}}}^{t} dt' \rho(t, t') p_F(t, t') + p_F^{\infty}(t) \int_{-\infty}^{t-\tau_{\text{mem}}} dt' \rho(t, t')$. The second integral on the right-hand side can be expressed by means of the normalization condition, $\int_{-\infty}^{t-\tau_{\text{mem}}} dt' \rho(t, t') = N - \int_{t-\tau_{\text{mem}}}^{t} dt' \rho(t, t')$. As compared to Eq. (2) now the integral becomes finite,

$$A(t) \approx N p_F^{\infty}(t) + \int_{t-\tau_{\text{mem}}}^{t} dt' \rho(t, t')[p_F(t, t') - p_F^{\infty}(t)],$$ (3)

with an integration interval of length τ_{mem}. Equation (3) is a major improvement to (2), since it permits robust straightforward numerics, also for paradigms with low activities. Simulations of large networks of neurons are thus equivalent to the integration of all local firing probabilities $p_F(t, t')$ within a memory interval $t' \in (t, t - \tau_{\text{mem}}]$ if one is solely interested in the dynamics of $A(t)$.

In order to investigate properties of inhibitory-coupled interneuronal assemblies, an appropriate cellular dynamics has to be specified. As elsewhere [4], a firing probability $p_F(t, t') = \Gamma[h(t, t')]$ being an exponential function ("gain function") $\Gamma(h) = \tau_F^{-1} \exp[\beta(h - \theta)]$ of the momentary membrane potential $h(t, t')$ is chosen, where θ denotes the threshold, β quantifies the noise level, and $\tau_F = 1$ ms is fixed and accounts for the right units.

The simplest model that decently approximates inhibitory synaptic interactions is a conductance-based integrate-and-fire dynamics, which serves as the model of choice for the rest of this Letter. The membrane time constant τ_M, the synaptic reversal potential E, and the capacitance C are constants, whereas the external current I is a time-dependent variable. So is the synaptic conductivity g that represents the input that is due to recurrent architecture of the network. All voltage variables are measured in units of the threshold θ; currents are measured in units of capacitance times threshold per time. For times $t \geq t' + \tau_{\text{ref}}$, and given the neuron has fired previously at time t', the solution to a conductance-based leaky integrator $C\dot{v} = -vC/\tau_M + g(E - v) + I$ is

$$v(t, t') = \int_{t'+\tau_{\text{ref}}}^{t} \frac{ds}{C}[I(s) + g(s, t')E]$$
$$\times \exp\left[-\int_{s}^{t} ds'[\tau_M^{-1} + C^{-1}g(s', t')]\right].$$ (4)

Then, the membrane potential $h(t, t') = v(t, t') + h^{\text{ref}}(t - t' - \tau_{\text{ref}})$ is a combination of Eq. (4) and a refractory kernel h^{ref} that models a phenomenon generally called after hyperpolarization, i.e., a negative, mostly potassium mediated potential that remains for a few milliseconds after the absolute refractory period τ_{ref}. For $t' < t < t' + \tau_{\text{ref}}$ one formally writes $h(t, t') = -\infty$, which indicates vanishing firing probability $p_F(t, t') = 0$. The integral kernel in Eq. (4) consists of a membrane part $\exp[-(t - s)/\tau_M]$ for $t \geq s$ and a synaptic part $\exp[-\int_{s}^{t} ds' C^{-1}g(s', t')]$, the latter reflecting the conductance-based integration of spikes. The conductivity $g(t, t') = \bar{g} \int_{-\infty}^{t} ds A(s) \kappa^{\text{syn}}(t - s)\Theta(s - t' - \tau_{\text{ref}})$ is assumed to be a linear mapping of the network activity A. Figure 1 demonstrates the quality and numerical robustness of the mean field approximation (3). For parameter values and model details, see [14].

Most analytical studies of oscillations in large-scale neuronal networks have been performed with integrate-and-fire neurons [7,11]. Since the reversal potential E for excitatory synapses resides at a clearly more depolarized level than the subthreshold dynamics, its nonlinear ef-
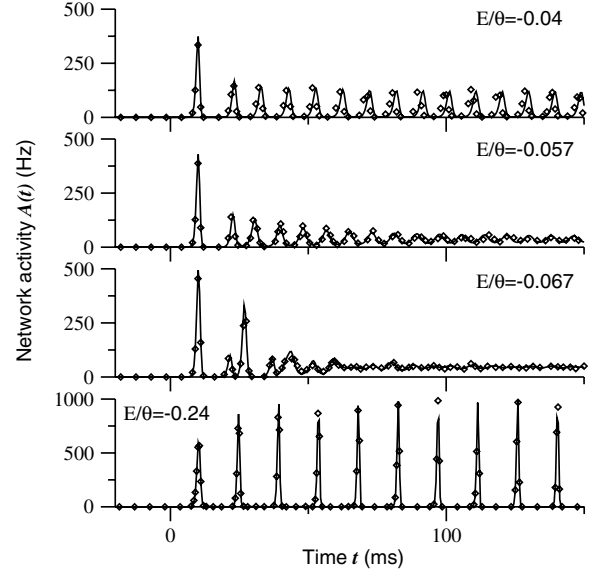

FIG. 1. Simulation of cellular network vs pool theory. Comparison of the network activity obtained from a simulation (diamonds) of 8000 fully interconnected neurons and theory [solid lines; Eq. (3)] for an external step current of amplitude $I/(C\theta) = 2.6$ kHz at $t = 0$ ms. For a constant value $NE\bar{g}/C = -0.4\theta$ kHz of the stationary inhibitory current, a reduction of the inhibitory reversal potential E yields a transition from a stable oscillation (top trace) to a stable asynchronous state (traces two and three) and back again to a stable oscillation (bottom).

fects on synaptic summation in purely excitatory coupled networks is thus generally omitted [4]. In networks with inhibitory connections, the reversal potential is closer to the resting potential and, as shown by Fig. 1, the stability of oscillations depends upon the value of the reversal potential, i.e., whether the synaptic current is conductance based or not. The critical influence of E upon both stability and frequency of network oscillations is now going to be further elaborated by means of a linear stability criterion that, in contrast to previous theories [4], is directly inferred from the normalization condition (1). As a result, mathematics is more concise.

Asynchronous states of the network are defined through $A(t) = A_0 = $ const. They can be obtained [4] through solving Eq. (1); i.e., they are the roots of the implicit equation $N = A_0 \int_{-\infty}^{t} dt' \exp[-\int_{t'}^{t} ds\, p_F(s, t')]$, where p_F is to be evaluated for constant activity A_0. Now it is argued that Eq. (1) can also be used for determining the stability of the asynchronous state. For $t > 0$ one therefore defines $\alpha(t) = A(t) - A_0$ to be the response to a perturbation at $t = 0$. The linearization of Eq. (1) then reads $\mathcal{N}[A_0 + \varepsilon\alpha] = \mathcal{N}[A_0] + \varepsilon \frac{d}{d\varepsilon} \mathcal{N}[A_0 + \varepsilon\alpha]_{\varepsilon=0}$. Because of $N \equiv \mathcal{N}[A_0 + \varepsilon\alpha] = \mathcal{N}[A_0]$, one finds

$$0 = \frac{d}{d\varepsilon} \mathcal{N}[A_0 + \varepsilon\alpha]_{\varepsilon=0}(t)$$
$$= \int_0^{\infty} d\hat{t}\, \alpha(\hat{t}) \frac{\delta \mathcal{N}[A](t)}{\delta A(\hat{t})} \bigg|_{A=A_0}. \qquad (5)$$

Equation (5) implies a condition for the possible perturbation responses α. The asynchronous state A_0 is stable if, for $t \to \infty$, all allowed perturbation responses α converge to zero. Combining (1) and (5) with $\rho = AS$, we have

$$0 = \int_0^{\infty} d\hat{t}\, \alpha(\hat{t}) \Big[\Theta(t - \hat{t}) S(t, \hat{t}) - A_0 \int_{-\infty}^{t} dt' S(t, t')$$
$$\times \int_{t'}^{t} ds \frac{\delta p_F(s, t')}{\delta A(\hat{t})} \Big]. \qquad (6)$$

Because of constant synaptic input an asynchronous state is stationary. Thus, Eq. (6) becomes a convolution $0 = \int_0^t d\hat{t}\, \alpha(\hat{t}) \Lambda(t - \hat{t})$ of the perturbation response α with a linear normalization kernel Λ. As a proof and in order to specify Λ, one considers that stationarity means $p_F(t, t') = p_F(t - t', 0)$ or $S(t, t') = S(t - t', 0)$, respectively. Moreover, taking into account causality $[\Theta(t - \hat{t})]$ and one-spike memory $[\Theta(\hat{t} - t')]$ one introduces a susceptibility χ, such that $\frac{\delta p_F(t, t')}{\delta A(\hat{t})} = \Theta(t - \hat{t})\Theta(\hat{t} - t')\Theta(t - t')\chi(t - \hat{t}, t - t')$. After some basic calculus one obtains a linear kernel $\Lambda(t) = \Theta(t)S(t, 0) - A_0\Phi(t)$, where $\Phi(t) = \int_t^{\infty} ds\, S(s, 0) \int_0^t ds'\, \chi(t - s', s - s')$. The Laplace-transformed version of Eq. (6) therefore reads $0 = \mathcal{L}[\Lambda](z)\mathcal{L}[\alpha](z)$, where $\mathcal{L}[f](z) := \int_0^{\infty} dt\, f(t) \times \exp(zt)$. As a consequence, nonvanishing perturbation responses $\mathcal{L}[\alpha](z)$ are forbidden, except for $\mathcal{L}[\Lambda](z) = 0$. In other words, the perturbation response is a linear

combination of exponential functions $\exp[-zt]$, where $z = \lambda + i\omega$ is a complex root of $\mathcal{L}[\Lambda]$. Moreover, stability of the asynchronous state means that *all* roots have a positive real part, i.e., $\lambda > 0$. Then, the imaginary parts ω of the roots determine the oscillation frequency of α if $|\lambda|$ is sufficiently small, e.g., $|\lambda| < |\omega|/(2\pi)$.

In the case of the conductance-based integrate-and-fire model with a noisy threshold, we find the susceptibility to be $\chi(t, t + x) = 0$ for $x < \tau_{\text{ref}}$ or $t < 0$, whereas for $x \geq \tau_{\text{ref}}$ and $t > 0$ Eq. (4) yields $\chi(t, t + x) = \beta\Gamma[h(t + x, 0)]\Xi(t, t + x)$, where

$$\Xi(t, t + x) = \int_{\tau_{\text{ref}}}^{t} ds \frac{\delta h(t + x, 0)}{\delta g(s, 0)} \frac{\delta g(s, 0)}{\delta A(x)}$$
$$= \bar{g} \int_x^{t+x} ds \frac{e^{(x-s)/\tau_{\text{syn}}} F_{A_0}(t + x)}{e^{(t+x-s)/\tau_M} F_{A_0}(s)}$$
$$\times [E - (h - h^{\text{ref}})(s, 0)] \qquad (7)$$

and $F_{A_0}(s) = \exp(A_0 \frac{\bar{g}\tau_{\text{syn}}}{C} e^{-(s-\tau_{\text{ref}})/\tau_{\text{syn}}})$. The roots of $\mathcal{L}[\Lambda]$ for conductance-based integrate-and-fire neurons are obtained numerically and are plotted in Fig. 2 for the four parameter regimes of Fig. 1. The imaginary part of the root approximates the oscillation frequency of the perturbation response. Positive but small real parts λ may suffice to maintain long-lasting oscillatory perturbation responses from noisy input; see Fig. 2 (diamonds).

At this point, it is a straightforward exercise to plot stability regions of asynchronous states. It is done so in Fig. 3 for a variation of the external current I and the synaptic reversal potential E, while keeping $\bar{g}E$ constant. One therefore defines a principal frequency f_0 to be $1/(2\pi)$ times the imaginary part of the complex root
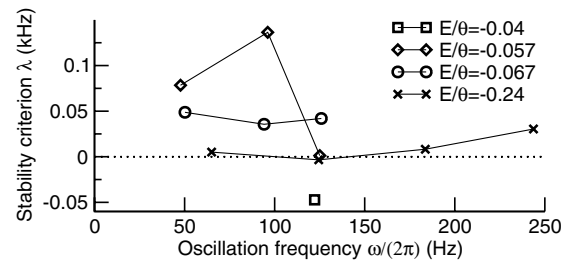

FIG. 2. Complex roots of the linear normalization kernel $\mathcal{L}[\Lambda]$ determine stability and oscillation frequency. Parameters are those of Fig. 1. For $E/\theta = -0.04$ there is a single root with a negative real part λ. Its imaginary part (122 Hz) roughly corresponds to the frequency of stable oscillation. In the case of $E/\theta = -0.057$ (diamonds) there is a root with a positive but very small real part at an oscillation frequency of 125 Hz. At $E/\theta = -0.067$ (circles) the real part of the root at about 125 Hz increases, whereas the real parts of the roots at lower frequencies decrease, which yields a mixture of different oscillatory components; see Fig. 1 (third trace). A further decrease to $E/\theta = -0.24$ while keeping $\bar{g}E$ constant, again yields a root with a negative real part and therefore an unstable asynchronous state. Roots with $\lambda > 0.15$ kHz are not shown.
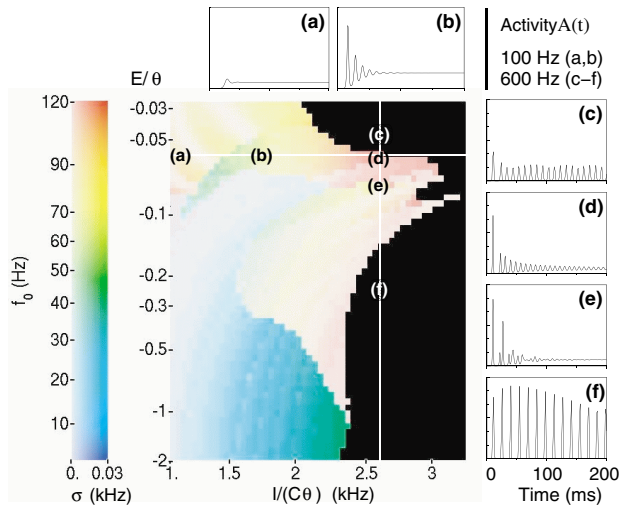
FIG. 3 (color). Stability and time course of the relaxation to asynchronous states. Regions of unstable asynchronous state are plotted black, stability regions are color-coded. Regions of low oscillation index σ (see the text) are depicted through grayish colors; high oscillation indices are marked with bright colors. Insets (a)–(f) are obtained by numerical integration of Eq. (3). For high values of $|E|$ oscillations are in a "current-based" limit and show only little dependence on E. Note that the E axis has a logarithmic scale. The vertical solid line denotes the value $I/(C\theta) = 2.6$ kHz, the horizontal line $E/\theta = -0.057$.

with the smallest real part λ and introduces an *oscillation index* $\sigma = A_0/R |\sum_{n=1}^{R} \exp[-(\lambda_n + i\omega_n)/f_0]|$ that measures the degree of monoperiodicity and the duration of the perturbation responses, given the asynchronous state is *stable* and $\mathcal{L}[\Lambda]$ has R roots $\lambda_n + i\omega_n$.

For a constant current $I/(C\theta) = 2.6$ kHz, the theory predicts a switch from an unstable asynchronous state to a stable one and back by continuously decreasing the reversal potential E. For small values of $|E|$, minuscule changes have dramatic effects on stability. Low values of the oscillation index σ (grayish colors) moreover indicate that oscillatory perturbation responses either decay rapidly [Fig. 3(a)] or consist of two interfering frequencies [Fig. 3(e)]. Abrupt transitions of principal frequencies f_0 (colors) indicate that two roots have a similar real part λ; cf. Figs. 2 (circles) and 3(e).

It is well-known that the value of the inhibitory reversal potential is highly dependent on age, synaptic location, and even activity [13]. From this perspective, the present results raise important questions on how these changes influence stability and frequency of network oscillations, especially during maturation.

The present theory is restricted by the assumption of homogeneity of both the network topology and the input. Since it is known [15] that heterogeneities easily disturb synchronization, a thorough adaption of the reversal potential can be even more crucial. The results obtained for the homogeneous case can then serve as an estimate of

network behavior at the edge of (de)synchronization. More generally, the theory provides a useful tool for investigating the effect of cellular parameters on network dynamics. This is possible also for systems with arbitrary-dimensional cellular dynamics and probably also for experimentally obtained survivor functions $S(t, t')$.

[1] A. T. Winfree, J. Theor. Biol. **16**, 15 (1967); R. E. Mirollo and S. H. Strogatz, SIAM J. Appl. Math. **50**, 1645 (1990).

[2] N. Kopell and G. B. Ermentrout, in *Handbook on Dynamical Systems: Toward Applications*, edited by B. Fiedler (Elsevier, Amsterdam, 2002).

[3] R. D. Traub, J. G. R. Jefferys, and M. A. Whittington, *Fast Oscillations in Cortical Circuits* (MIT Press, Cambridge, MA, 1999); R. Maex and E. De Schutter, J. Neurosci. **23**, 10503 (2003).

[4] W. Gerstner, Phys. Rev. E **51**, 738 (1995); W. Gerstner, Neural Comput. **12**, 43 (2000).

[5] H. R. Wilson and J. D. Cowan, Biophys. J. **12**, 1 (1972); H. R. Wilson and J. D. Cowan, Kybernetik **13**, 55 (1973).

[6] D. J. Amit, *Modeling Brain Function* (Cambridge University Press, Cambridge, 1989); P. S. Churchland and T. T. Sejnowski, *The Computational Brain* (MIT Press, Cambridge, MA, 1992); E. T. Rolls and A. Treves, *Neural Networks and Brain Function* (Oxford University Press, New York, 1998).

[7] L. F. Abbott and C. van Vreeswijk, Phys. Rev. E **48**, 1483 (1993)

[8] H. S. Seung, Proc. Natl. Acad. Sci. U.S.A. **93**, 13 339 (1996).

[9] J. Eggert and J. L. van Hemmen, Phys. Rev. E **61**, 1855 (2000); J. Eggert and J. L. van Hemmen, Neural Comput. **13**, 1923 (2001).

[10] N. Brunel, J. Comput. Neurosci. **8**, 183 (2000).

[11] C. van Vreeswijk and D. Hansel, Neural Comput. **13**, 959 (2001); D. Hansel and G. Mato, Neural Comput. **15**, 1 (2003).

[12] O. Shriki, D. Hansel, and H. Sompolinsky, Neural Comput. **15**, 1809 (2003).

[13] C. Rivera *et al.*, Nature (London) **397**, 251 (1999); P. Andersen *et al.*, J. Physiol. **305**, 279 (1980); K. J. Staley, B. L. Soldo, and W. R. Proctor, Science **269**, 977 (1995).

[14] Integral kernels are $\kappa^{\mathrm{syn}}(t) = \Theta(t - \Delta) \exp[-(t - \Delta)/\tau_{\mathrm{syn}}]$ and $h^{\mathrm{ref}}(t) = -\gamma\Theta(t) \exp(-t/\tau_\gamma)$, where $\Theta(t) = 1$ for $t \geq 0$ and $\Theta(t) = 0$ for $t < 0$. Default parameters are $\beta^{-1} = 0.04\theta$, $\tau_F = 1$ ms, $\tau_{\mathrm{ref}} = 1$ ms, $\tau_M = 20$ ms, $E\bar{g}/C = -0.04\theta/(N$ ms$)$, $\tau_{\mathrm{syn}} = 3$ ms, $\Delta = 1$ ms, $\gamma/\theta = 14$, $\tau_\gamma = 3$ ms, and $\tau_{\mathrm{mem}} = 5\tau_M$.

[15] J. A. White *et al.*, J. Comput. Neurosci. **5**, 5 (1998); D. Golomb, X. J. Wang, and J. Rinzel, J. Neurophysiol. **72**, 1109 (1994).