# Neural Networks, Lattice Instantons, and the Anti-Integrable Limit

Paul C. Bressloff

*Department of Mathematical Sciences, Loughborough University of Technology, Loughborough, Leics, LE11 3TU, United Kingdom*
(Received 28 February 1995)

We prove the existence of localized ground states (lattice instantons) for the standard diffusive Haken model in any finite dimension. Our proof is based on singular perturbation theory in the limit of zero diffusive coupling (anti-integrable limit).

Exploiting certain similarities between pattern formation in synergetic systems and pattern recognition, Haken [1] has constructed a neural network model that implements a form of competitive gradient dynamics. The ground states of the system consist of strictly localized states in which only one neuron is excited and the remainder quiescent. In other words, the network dynamically realizes a winner-take-all strategy. An interesting extension of this type of network is the diffusive Haken model in which the introduction of a diffusive interaction between the neurons leads to a delocalization of the original model's ground states [2]. When there exists a balance between the effects of diffusion and localization, it is possible to obtain new ground states that are localized excitations (or bubbles) distributed over many neurons. Such ground states are more robust than those of the simple Haken model and also better reflect the kind of coherent structures found in neurobiological systems.

One of the major claims of Ref. [2] is that, for a network arranged on a $d$-dimensional square lattice with standard nearest neighbor diffusive coupling, localized ground states exist only when $d = 1$; such states cannot be sustained for $d > 1$ even for arbitrarily small diffusive coupling. This claim is based on numerical simulations of the lattice model together with some variational calculations of a continuum model obtained in the limit of large diffusive coupling. The continuum model possesses instantonlike solutions to the variational problem. In one dimension the instantons can be analyzed explicitly and are shown to correspond to localized ground states. Numerical analysis of higher-dimensional instantons suggests that they no longer represent ground states of the network when $d > 1$.

In this Letter, we prove analytically that the above claim is false. In other words, we show that the (standard) diffusive Haken model supports localized ground states in any finite dimension provided that the coupling is sufficiently small. Such a result also holds for more general choices of coupling such as next nearest neighbor. Our approach is based on singular perturbation theory about the limiting case of zero diffusive coupling, and is an extension of the work of Aubry and collaborators on the Frenkel-Kontorova model (see [3] and references therein).

The localized states are analogous to the breather solutions recently found for Hamiltonian networks [4]. Following Aubry we shall denote the limit of zero coupling the *anti-integrable* limit, although a more appropriate term here might be the *anticontinuum* limit. For our purposes, one of the important features concerning the qualitative behavior of a system near the anti-integrable limit is that it does not depend on the dimensionality of the lattice nor on the details of the coupling.

Consider a network of $N$ neurons arranged on a $d$-dimensional lattice. Denote the state of the $n$th neuron by $q_n \in R$ with $n = 1, \ldots, N$. (For a finite lattice we impose periodic boundary conditions.) In the standard diffusive Haken model, the state of the network at time $t$, $\mathbf{q}(t) = \{q_n(t)\}$, evolves according to the gradient dynamical rule

$$\dot{q}_n = -\frac{\partial V}{\partial q_n}, \tag{1}$$

where

$$V = \frac{\alpha}{2} \sum_{\langle m,n \rangle} (q_m - q_n)^2 - \frac{1}{2}D + \frac{1}{2}D^2 - \frac{1}{4}\sum_n q_n^4 \tag{2}$$

and

$$D = \sum_n q_n^2. \tag{3}$$

The first term on the right-hand side of Eq. (2) represents a nearest neighbor diffusive interaction with coupling constant $\alpha$ while the remaining terms correspond to the localizing potential of the original Haken model. The equations of motion are

$$\dot{q}_n = \alpha \sum_m W_{nm} q_m + (1 - 2D + q_n^2) q_n. \tag{4}$$

Here $\mathbf{W}$ is the coupling matrix on the lattice with $W_{nm} = \delta_{m,n+1} + \delta_{m,n-1} - 2\delta_{m,n}$ when $d = 1$, etc. Equation (4) is invariant under the transformation $\mathbf{q} \rightarrow -\mathbf{q}$. Moreover, $q_n(t) \geq 0$ for all $t > 0$ if $q_n(0) \geq 0$. The latter can be seen by setting $q_n = 0$ on the right-hand side of Eq. (4) and noting that the off-diagonal components of $\mathbf{W}$ are positive so that $\dot{q}_n \geq 0$ when $q_m \geq 0$ for all $m \neq n$. That is, $q_n$ cannot cross over to the negative real axis.

       

Hence we can restrict our discussion to solutions in the domain $R_+^N = \{\mathbf{q} \in R^N \mid q_n \geq 0,\ n = 1,\ldots,N\}$. This is reasonable from the neural network perspective, since we are interested in detecting the dominant features of an input pattern whose components are all positive.

The network converges to one of the stationary states of the potential $V$ obtained by setting $\dot{q}_n = 0$ in Eq. (4), with $D$ determined self-consistently using Eq. (3). In the case of zero diffusive coupling ($\alpha = 0$), the lattice structure does not play a role and Eq. (4) has the solutions $q_n = 0$ or $q_n = \sqrt{2D - 1}$. Hence the stationary states $\mathbf{q} \in R^N$ can be divided into $N + 1$ classes determined by the number $m$ of excited neurons [2]. For a given $m$, $D = m/(2m - 1)$, and the corresponding potential is $V_m = -(m/4)(2m - 1)^{-1}$. The stationary states $m = 1$ are stable, whereas all other stationary states are either unstable ($m = 0$) or saddle points ($m > 1$). For $m = 1$ there exists a single excited neuron, $n_0$ say, such that $q_n = \delta_{n_0,n}$. Moreover, $D = 1$ and $V_1 = -1/4$. This is one of the $N$ strictly localized ground states of the network. Also note that there are two homogeneous stationary states given by the *vacuum* state $m = 0$ and the *dissipative* state $m = N$. The former satisfies $q_n = 0$ for all $n$ and $V_0 = 0$ and the latter has $q_n = 1/\sqrt{2N - 1}$ for all $n$ and $V_N = -N/(8N - 4)$. In the large $N$ limit the dissipative state becomes pointwise identical to the vacuum state but has lower energy, $V_\infty = -1/8$.

We wish to prove that there exists a continuation of the stable localized ground states when $\alpha$ increases from zero. We first give a heuristic argument for the existence of such states and then present an exact mathematical proof. Suppose that the initial state of the network is one of the strictly localized states $q_n(0) = \delta_{n,0}$. Then the potential at $t = 0$ satisfies $V < -1/8$ for $\alpha < 1/(4|W_{00}|) = 1/(8d)$. (We are assuming $N \to \infty$ and nearest neighbor coupling.) Since $dV/dt \leq 0$ for all $t$, the final state cannot be the dissipative state. Moreover, we expect this final state to be localized since $D < \infty$ for finite values of $V$ so that the equilibrium lattice configuration must decay sufficiently fast with distance on the lattice from the origin $n = 0$. One might expect the size of the basin of attraction of such localized equilibria to decrease as the dimensionality of the lattice increases; this suggests that the initial configurations used in the simulations of Ref. [2] were outside such basins of attraction when $d > 1$ so that the network converged to the dissipative state instead. (Alternatively, the values of $\alpha$ used were too large.) Note that for an infinite lattice the dissipative state is stable for arbitrarily small $\alpha$. Although the above argument supports the existence of localized states, we still need to establish whether or not they are stable, how they relate to the ground states at $\alpha = 0$, and how they decay with spatial location on the lattice. All of these issues can be tackled using perturbation theory near the anti-integrable limit, which is the basis of our proof presented below.

The first step of our proof follows closely the analysis of the Frenkel-Kontorova model [3,5]. We consider an infinite lattice and let $l_\infty$ denote the space of configurations $\mathbf{q} = \{q_n,\ n \in Z\}$ with norm $\|\mathbf{q}\| = \sup_n |q_n|$. (For the moment we do not assume that $q_n \geq 0$ for all $n$.) We introduce a mapping $\mathcal{T}: \{q_n\} \to \{p_n\}$ on $l_\infty$ according to the implicit equation

$$A p_n - p_n^3 = \alpha \Delta_n(\mathbf{q}) = \alpha \sum_m W_{nm} q_m, \qquad n \in Z, \tag{5}$$

where $A = 2D - 1$ is assumed initially to be fixed such that $A \geq 0$. It is clear that for a given $\alpha$ and $D$ any fixed point of the mapping $\mathcal{T}$ corresponds to an equilibrium of Eq. (4). [In order for such an equilibrium to correspond to a stationary configuration of the potential $V$, one needs to solve the self-consistency condition for $D$, Eq. (3). We shall return to this issue later on.]

Define the subspace of configurations $\mathcal{B}(a) \subset l_\infty$ by $\mathcal{B}(a) = \{q_n \in [-a, a] \forall n \neq 0,\ q_0 \in [a, \sqrt{3} + a]\}$, where $a = \sqrt{A/3}$. Note that $\mathcal{B}(a)$ contains configurations at a distance smaller than $a$ from the localized configuration $q_n = \sqrt{A}\,\delta_{0,n}$, which is one of the stable ground states for $\alpha = 0$ and $A = 1$. Next impose the condition $\alpha \|\mathbf{W}\| (\sqrt{3} + 1) < 2a^2$. (We define the spectral norm for any bounded operator $\mathbf{W}$ on $l_\infty$ by $\|\mathbf{W}\| = \sup_{\mathbf{q} \neq 0} \|\mathbf{W}\mathbf{q}\|/\|\mathbf{q}\|$. For nearest neighbor coupling in $d$ dimensions we have $\|\mathbf{W}\| = 4d$.) This condition ensures that $\alpha |\Delta_n(\mathbf{q})| < F(a)$ with $F(q) \equiv Aq - q^3$. It follows that $\mathcal{T}$ determines a unique mapping on $\mathcal{B}(a)$ given by $p_n = z_+(\Delta_n(\mathbf{q}))$ for $n \neq 0$ and $p_0 = z(\Delta_0(\mathbf{q}))$. Here $z(\Delta), z_\pm(\Delta)$ are the three real roots of the cubic $F(q) = \alpha\Delta$ with $\alpha|\Delta| < F(a)$; see Fig. 1. That is (see page 17 of Ref. [6]),

$$z(\Delta) = 2a \cos\theta(\Delta), \tag{6}$$

$$z_\pm(\Delta) = -a \cos\theta(\Delta) \pm \sqrt{3}a \sin\theta(\Delta), \tag{7}$$

and

$$\theta(\Delta) = \frac{1}{3} \cos^{-1}\left(-\frac{\alpha\Delta}{2a^3}\right), \qquad \theta \in [0, \pi/3]. \tag{8}$$



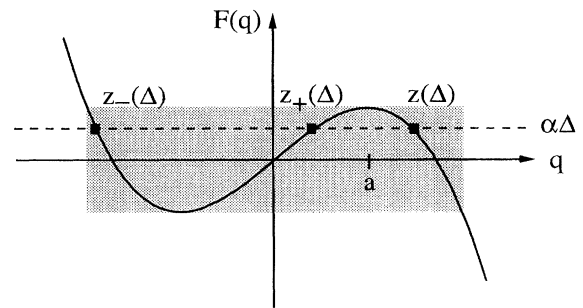FIG. 1. Construction of mapping $\mathcal{T}$ on $\mathcal{B}(a)$ with $a = \sqrt{A/3}$ and $\alpha\Delta$ restricted to the shaded region $\alpha|\Delta| < F(a)$.

Note that $\theta(\Delta)$ is a monotonically decreasing function of $\Delta$ over the given interval; if $\theta = \pi/3$, then $z(\Delta) = z_+(\Delta) = a$, $z_-(\Delta) = -2a$; if $\theta = \pi/6$, then $z(\Delta) = \sqrt{3}a$, $z_+(\Delta) = 0$, $z_-(\Delta) = -\sqrt{3}a$; and if $\theta = 0$, then $z_-(\Delta) = z_+(\Delta) = -a$, $z(\Delta) = 2a$.

The Jacobian of the mapping $\mathcal{T}$ is given by

$$J_{mn} = \frac{\partial p_m}{\partial q_n} = \frac{\alpha}{A - 3p_m^2} W_{mn} \,. \tag{9}$$

We wish to impose the condition $\|\mathbf{J}\| < K$ on $\mathcal{B}(a)$. Equation (9) and Fig. 1 show that this is equivalent to the requirement that the roots $z_+, z$ are sufficiently far from the turning point at $q = a$. More specifically, using Eqs. (6) and (7) we have $A - 3z_+(\Delta)^2 = A(\sqrt{3}\sin 2\theta - 2\sin^2\theta) > K^{-1}\|\mathbf{W}\|\alpha$ and $3z(\Delta)^2 - A = A(4\cos^2\theta - 1) > K^{-1}\|\mathbf{W}\|\alpha$. It turns out that the first condition is the stronger one and on using Eq. (8) together with the condition $|\Delta(\mathbf{q})| < \|\mathbf{W}\|(\sqrt{3} + 1)a$ for $\mathbf{q} \in \mathcal{B}(a)$ leads to an upper bound for $\alpha$. That is [8], $\|\mathbf{J}\| < K$ on $\mathcal{B}(a)$ provided that $\alpha < \alpha_c$ where $\alpha_c = K\rho_c A/\|\mathbf{W}\|$ and $\rho_c$ is the solution to

$$\rho_c = \frac{2}{3K(\sqrt{3} + 1)} \, |\cos 3\theta(\rho_c)| \,, \tag{10}$$

with the function $\theta(\rho) \in [\pi/6, \pi/3]$ defined by

$$\sqrt{3}\sin 2\theta(\rho) - 2\sin^2\theta(\rho) = \rho \,. \tag{11}$$

(We can restrict $\theta(\rho)$ to the interval $[\pi/6, \pi/3]$ since $|\cos 3\theta|$ is symmetric about $\pi/6$.) Note that the solution $\rho_c$ exists as a unique, monotonically decreasing function of $K$ that can be determined graphically. For example, $\rho_c \approx 0.238$ when $K = 1$. Hence, for a given $A > 0$, we have that $\mathcal{T}$ defines a unique contraction mapping on $\mathcal{B}(a)$ for $\alpha < \alpha_c$ and $K = 1$. An application of the Banach fixed point theorem [7] then shows that $\mathcal{T}$ has a unique fixed point $\mathbf{q}(\alpha; A)$, which is uniformly continuous in $\alpha$ such that $\lim_{\alpha \to 0} q_n(\alpha; A) = \sqrt{A}\delta_{n,0}$. It can also be proven that this continuation of the ground state is itself localized. More specifically, an application of Theorem 3 of Ref. [4] shows that there exist constants $h > 0$ and $0 < \mu < 1$ such that $|q_n(\alpha; A)| < \exp(h\alpha)\mu^{|n|}$ for all $n$.

One possible point of confusion is that the range of the mapping $\mathcal{T}$ is not restricted to the positive domain $q_n \geq 0$ for all $n$. For even if $q_n \geq 0$ for all $n$, it is possible that $\Delta_n(\mathbf{q})$ is negative. If $\Delta$ is negative then so is the root $z_+(\Delta)$ (see Fig. 1). However, the fixed point $\mathbf{q}(\alpha; A)$ of $\mathcal{T}$ does lie in the positive domain. We have essentially broken the reflection symmetry of the underlying lattice by choosing the subspace $\mathcal{B}(a)$ to be a neighborhood of $q_n = \delta_{n,0}$ rather than $q_n = -\delta_{n,0}$. The hidden symmetry ensures that there also exists a local continuation of the latter ground state given by $-\mathbf{q}(\alpha; A)$.

The second step in our proof involves checking that the continuation $\mathbf{q}(\alpha, A)$ satisfies the self-consistency condition (3) for an appropriate choice of $A(\alpha)$. Substitution into Eq. (3) gives an implicit equation for $A$ in terms of $\alpha$

$$A = 2\sum_n q_n^2(\alpha; A) - 1 \,. \tag{12}$$

It immediately follows that $A = 1$ when $\alpha = 0$ as expected. Moreover, the right-hand side of Eq. (12) is finite for all $\alpha < \alpha_c$ since $\mathbf{q}(\alpha; A)$ decays exponentially with distance. An application of the implicit function theorem then shows that there exists a local continuation $A = A(\alpha)$ with $A(0) = 1$ and $0 < A(\alpha) < 1$. Introduce the matrix $\mathbf{M}$ that is obtained by linearizing Eq. (4) about the fixed point $\mathbf{q}(\alpha; A)$; namely, $\mathbf{M} = [\mathbf{1} - \mathbf{J}]\mathbf{D}$, where $D_{mn} = (A - 3q_m^2)\delta_{m,n}$. The matrix $\mathbf{M}$ is invertible with $\mathbf{M}^{-1} = \mathbf{D}^{-1}[\mathbf{1} - \mathbf{J}]^{-1}$ since $\|\mathbf{J}\| < 1$ at the fixed point. [Note that $\|\mathbf{J}\| < 1$ provided that $\alpha < \rho_c A(\alpha)/\|\mathbf{W}\|$ where $\rho_c$ is the solution to Eq. (10) for $K = 1$.] Differentiating Eq. (12) with respect to $\alpha$, we find that the solution $A(\alpha)$ satisfies [8]

$$\frac{dA}{d\alpha} = \frac{4\langle \mathbf{q}|\mathbf{M}^{-1}\mathbf{W}|\mathbf{q}\rangle}{4\langle \mathbf{q}|\mathbf{M}^{-1}|\mathbf{q}\rangle + 1} \,, \tag{13}$$

where $\langle \mathbf{q}|\mathbf{M}^{-1}|\mathbf{q}\rangle = \sum_{n,m} q_m(\alpha)M_{mn}^{-1}q_n(\alpha)$ with $\mathbf{q}(\alpha) = \mathbf{q}(\alpha; A(\alpha))$, etc. Direct substitution shows that $dA/d\alpha |_{\alpha=0} = -4$ [the minus sign arising from the denominator on the right-hand side of Eq. (13)].

Thus, for sufficiently small $\alpha$ there exists a localized state $\mathbf{q}(\alpha) = \mathbf{q}(\alpha; A(\alpha))$ with $\lim_{\alpha \to 0} q_n(\alpha) = q_n(\alpha; 1) = \delta_{n,0}$ for all $n$. Given the energy gap between the dissipative state $(V = -1/8)$ and the ground state at $\alpha = 0$ $(V = -1/4)$, it follows that $\mathbf{q}(\alpha)$ has lower energy than the dissipative state for small $\alpha$. Indeed, differentiating Eq. (2) with respect to $\alpha$ for $\mathbf{q} = \mathbf{q}(\alpha)$ gives $dV/d\alpha = \langle \mathbf{q}|\mathbf{W}|\mathbf{q}\rangle/2$ (since $\mathbf{q}$ is a stationary state of $V$). It follows that $|\Delta V| \leq \alpha\|\mathbf{W}\|[1 + A(\alpha)]/4$. Finally, we need to check the stability of this localized state. Linearizing Eq. (4) about $\mathbf{q}(\alpha)$ with $D$ satisfying Eq. (3) shows that this state is stable if and only if the matrix $\hat{M}_{mn} = M_{mn} + 4q_n(\alpha)q_m(\alpha)$ is positive semidefinite. In other words, we require $|\langle \mathbf{q}|\mathbf{M}|\mathbf{q}\rangle| \leq 4\langle \mathbf{q}|\mathbf{q}\rangle^2 = [1 + A(\alpha)]^2$. Using spectral properties of the matrix $\mathbf{M}$ one finds that $|\langle \mathbf{q}|\mathbf{M}|\mathbf{q}\rangle| \leq 3A(\alpha)[1 + \|\mathbf{J}\|]\langle \mathbf{q}|\mathbf{q}\rangle$ [8]. Since $A(\alpha) < 1$, it follows that $\mathbf{q}(\alpha)$ is stable provided that $\|\mathbf{J}\| < 1/3$. The latter condition holds if $\alpha < \rho_c A(\alpha)/(3\|\mathbf{W}\|)$ with $\rho_c$ the solution to Eqs. (10) and (11) for $K = 1/3$, which has the numerical value $\rho_c \approx 0.621$. [We note that the condition $\|\mathbf{J}\| < 1/3$ ensures that the denominator in Eq. (13) is nonzero and hence $A(\alpha)$ exists [8].] Finally, using $\|\mathbf{W}\| = 4d$ and the approximation $A(\alpha) \approx 1 - 4\alpha$ shows that there exists a stable localized ground state if $\alpha < 1/(4 + 20d)$. This is in fact an underestimate since, from our heuristic energy argument, we expect these localized states to persist at least for all $\alpha < 1/(8d)$. [From continuity arguments we do not expect there to be any other stable low energy stationary states, so we identify each continuation $\mathbf{q}(\alpha)$ as a ground state.] This completes our proof.

In conclusion, we have used perturbation theory near the anti-integrable limit (small diffusive coupling) to prove the existence of stable localized ground states for any finite dimension $d$. Although we considered only nearest neighbor coupling, our proof actually applies to any choice of coupling matrix $\mathbf{W}$ provided that the coupling decays exponentially with distance on the lattice. Our analysis raises a number of interesting questions. First, do such states persist to arbitrarily large values of the coupling $\alpha$? The instanton solutions obtained in the continuum limit $\alpha \to \infty$ [2] suggest that the answer is yes for $d = 1$ and no for $d > 1$. Therefore, when $d \geq 2$ is there some form of phase transition at a critical value of $\alpha$ beyond which localized ground states no longer exist? (That is, in some sense, the converse of Anderson localization in disordered systems [9].) Another interesting question concerns how the size of the basin of attraction of a localized state varies with $\alpha$ and the dimensionality of the lattice. One can reinterpret the results of the computer simulations of Ref. [2] as demonstrating that the basin of attraction of localized ground states is much smaller for $d > 1$ than $d = 1$. This would then justify looking at nonstandard forms of diffusive coupling in order to make such localized states more accessible [2]. Finally, we believe that the anti-integrable limit could prove a powerful tool in studying self-organization and coherent structures in other lattice or network models. Note, for example, that our analysis of the diffusive Haken model [Eq. (4)] for fixed $A$ can be directly applied to prove the existence of breathers or localized periodic solutions of the discrete nonlinear Schrödinger equation $i\dot{\psi}_n = \alpha \sum_m W_{nm}\psi_m + |\psi_n|^2\psi_n$. Assuming a solution of the form $\psi_n = \exp(-i\omega t)\phi_n$, $\omega > 0$, leads to a time-independent Schrödinger equation identical in form to the equilibrium condition in Eq. (4) with $\omega = A$. The application of the anti-integrable limit to such a system was previously suggested in Ref. [4].

[1] H. Haken, *Synergetic Computers and Cognition* (Springer, Berlin, 1991).

[2] M. Schmutz and W. Banzhaf, Phys. Rev. A **45**, 4132 (1992).

[3] S. Aubry, Physica (Amsterdam) **71D**, 196 (1994).

[4] S. Aubry and R. S. Mackay, Nonlinearity **7**, 1623 (1994).

[5] S. Aubry and G. Abramovici, Physica (Amsterdam) **43D**, 199 (1990).

[6] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).

[7] E. Krezsig, *Introductory Functional Analysis with Applications* (Wiley, New York, 1989).

[8] P. C. Bressloff (to be published).

[9] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).