# Nonlinear Patterns in Urban Crime: Hotspots, Bifurcations, and Suppression* 

M. B. Short ${ }^{\dagger}$, A. L. Bertozzi ${ }^{\dagger}$, and P. J. Brantingham ${ }^{\ddagger}$


#### Abstract

We present a weakly nonlinear analysis of our recently developed model for the formation of crime patterns. Using a perturbative approach, we find amplitude equations that govern the development of crime "hotspot" patterns in our system in both the one-dimensional (1D) and two-dimensional (2D) cases. In addition to the supercritical spots already shown to exist in our previous work, we prove here the existence of subcritical hotspots that arise via subcritical pitchfork bifurcations or transcritical bifurcations, depending on the geometry. We present numerical results that both validate our analytical findings and confirm the existence of these subcritical hotspots as stable states. Finally, we examine the differences between these two types of hotspots with regard to attempted hotspot suppression, referencing the varying levels of success such attempts have had in real world scenarios.


Key words. bifurcations and instability, pattern formation, crime modeling
AMS subject classifications. $70 \mathrm{~K} 50,70 \mathrm{~K} 60,91 \mathrm{D} 99$
DOI. 10.1137/090759069

1. Introduction. The study of pattern formation in physical and mathematical systems has a long and interesting history. This general subject area is also quite diverse, examining biological (see, as a small sample, [37, 28, 27]), geological [33, 10, 2], and even sociological systems $[30,14]$, to name but a few. Though these various subjects and systems may seem completely unrelated, the mathematics describing the patterns in each are surprisingly similar. Consequently, a robust, powerful, and universal set of mathematical tools has been developed to study such systems, and the employment of these tools can lead to better understanding of pattern forming systems, regardless of their specific nature.

Recently, we set forth to develop a mathematical model to describe the spatio-temporal patterns of urban crime [35]. Using well-known criminological ideas regarding the way in which criminal events affect future crime risk in a location, and the way in which risk can spread from one area to another $[18,19,20,1]$, we constructed a model consisting of two coupled, nonlinear PDEs that may describe the formation and dynamics of crime "hotspots"-spatio-temporal clusters of high crime. Using a simple linear stability analysis of our model, we found that the homogeneous system can be unstable to disturbances of specific wavenumbers under certain parameter regimes, leading to hotspot formation. However, our previous work stopped there, with no investigation of the possibility of hotspots outside of this linearly unstable regime. This paper addresses this possibility by performing a weakly nonlinear analysis on our system and

[^0]developing amplitude equations for the model; this is a detailed follow-up to our paper [34], which presents only a few of the qualitative results of such an analysis. By investigating the possible bifurcations in the steady state solutions of our system both analytically and numerically, we indeed find that stable, "large" amplitude hotspots may exist even in the linearly stable regime.

The fact that these subcritical hotspots exist within our system is especially interesting when attempting to understand the outcome of hotspot suppression, typically by police executing a strategy known as hotspot policing, which has become dominant over the past two decades $[6,4,5,41,40]$. Recognizing that crime tends to form dense clusters in space and time, leaving some areas with little or no crime problem, police routinely target their limited resources at those locations experiencing high crime. That hotspot policing would be an improvement over random patrol is uncontroversial; it has been well known since the 1970s that random patrol has little measurable effect on crime [23]. However, questions have been raised about whether hotspot policing leads to lasting hotspot reductions, or simply the displacement of hotspots from one area to another [29, 3, 7]. The present research provides a formal theoretical foundation for understanding different potential outcomes from hotspot policing in relation to the classification of hotspots as either supercritical or subcritical [34].

The remainder of the paper is organized as follows. In section 2, we give a brief introduction to our crime model and the major results found in [35]. In section 3, we perform a weakly nonlinear analysis of our system in both the one-dimensional (1D) and two-dimensional (2D) cases, deriving some analytical results for the amplitude equations and bifurcations governing the hotspots exhibited by the system. In section 4, we compare these analytical results to numerical solutions. Finally, in section 5, we explore the possible results of hotspot suppression qualitatively and numerically using both the continuum and discrete models.
2. Background. We begin by reviewing the results of [35]. First, we developed an agentbased model of criminal activity that aims to reproduce the known phenomena of repeat and near-repeat victimization $[18,19,20,1]$, whereby crime risk becomes elevated in an area and its surroundings following an initial event there. This model couples the dynamics of moving, offending criminals on a 2D lattice (with lattice spacing $\ell$ ) with an underlying scalar field $A(\mathbf{x}, t)$ that we refer to as the attractiveness. As the name implies, the attractiveness field is a measure of how desirable any given location on the lattice $\mathbf{x}=(i, j)$ is as a target for criminal activity, with the numerical value of the field giving the stochastic rate of offending for the $n(\mathbf{x}, t)$ criminals at that location. The model evolves in discrete time, using a timestep $\delta t$, and during each timestep criminals may victimize their current location with probability

$$
\begin{equation*}
p_{v}(\mathbf{x}, t)=1-e^{-A(\mathbf{x}, t) \delta t} . \tag{2.1}
\end{equation*}
$$

If the criminal does in fact choose to commit a crime during this timestep, he is removed from the lattice. If, on the other hand, he does not commit a crime during this timestep, he will instead move to one of the four lattice points adjacent to his current location, selecting a particular neighbor $\mathbf{x}^{\prime}$ with probability

$$
\begin{equation*}
p_{m}\left(\mathrm{x}^{\prime}, t ; \mathbf{x}\right)=\frac{A\left(\mathrm{x}^{\prime}, t\right)}{\sum_{\mathbf{x}^{\prime \prime} \sim \mathrm{x}} A\left(\mathrm{x}^{\prime \prime}, t\right)}, \tag{2.2}
\end{equation*}
$$



Figure 1. Example output from the discrete system. These colormaps display high $A$ in red, $\bar{A}$ in green, and low A in blue to purple. In (a) is an example of stationary hotspots, where once hotspots form at a point, they tend to stay there indefinitely; in (b) is an example of transitory hotspots, where hotspots do not typically last forever, and will move about, deform, or disappear over time; and in (c) is an example of no hotspots, where no large spots are ever observed.
where the notation $\mathbf{x}^{\prime \prime} \sim \mathbf{x}$ indicates all of the sites neighboring site $\mathbf{x}$. In this way, criminals actively seek out areas of high $A$, where they are more likely to commit crimes.

The attractiveness field is composed of a static component $A^{0}$ (referred to as the "baseline" attractiveness) and a dynamic component $B(\mathbf{x}, t)$, such that $A(\mathbf{x}, t)=B(\mathbf{x}, t)+A^{0}$. After all of the criminal activity for the round is completed, the dynamic component of the attractiveness field is updated through the following mechanisms, meant to model repeat and near-repeat victimization. First, the dynamic attractiveness spreads spatially via a weighted averaging procedure between each site and its four neighbors. Next, it exhibits an exponential decay with rate $\omega$. Finally, for each criminal event that occurred at $\mathbf{x}$, the attractiveness is increased there by an amount $\theta$. So, if the number of events at site $\mathbf{x}$ during the current timestep is given by $E(\mathbf{x})$, then the attractiveness at the beginning of the next timestep will be

$$
\begin{equation*}
A(\mathbf{x}, t+\delta t)=\left[(1-\eta) B(\mathbf{x}, t)+\frac{\eta}{4} \sum_{\mathbf{x}^{\prime} \sim \mathbf{x}} B\left(\mathbf{x}^{\prime}, t\right)\right](1-\omega \delta t)+\theta E(\mathbf{x})+A^{0} \tag{2.3}
\end{equation*}
$$

Note that, by definition, $\eta \leq 1$. As a final step of the simulation, new criminals are generated on each point of the lattice at a rate $\Gamma$.

The discrete model thus described contains a number of parameters, and depending upon the choice of these parameters, the system may exhibit three general types of behavior: stationary (fixed in space) crime hotspots, transitory (moving about in space or appearing and disappearing in time) hotspots, or no hotspots at all; these three cases are illustrated in Figure 1.

Second, we derived a continuum limit of the discrete model by converting the criminals into a number density $\rho(\mathbf{x}, t)$, taking expectation values for all probabilistic events, and letting $\ell, \delta t \rightarrow 0$ with the constraint $\ell^{2} / \delta t=D$, a diffusion coefficient ([35] presents all of the algebraic details). This hydrodynamic limit results in our PDE model, which is the major focus of this
paper and can be written in the dimensionless form

$$
\begin{align*}
& \frac{\partial A}{\partial t}=\eta \nabla^{2} A-A+A^{0}+\rho A  \tag{2.4}\\
& \frac{\partial \rho}{\partial t}=\vec{\nabla} \cdot\left[\vec{\nabla} \rho-\frac{2 \rho}{A} \vec{\nabla} A\right]-\rho A+\bar{A}-A^{0} \tag{2.5}
\end{align*}
$$

where $\eta$ is the same as in the discrete model, and the two remaining parameters $\bar{A}$ and $A^{0}$ can be found from various combinations of the six remaining parameters present in the discrete model. Inspecting these equations, we see that crimes occur locally at rate $\rho A$, and each such crime causes $A$ to increase. In addition, $A$ diffuses with dimensionless diffusion coefficient $\eta$ $(\leq 1)$ and decays exponentially to the baseline value $A^{0}$. Criminals exhibit diffusive motion with an advective bias up gradients of $\ln A$. Finally, criminals are subtracted from the system when they commit a crime and are added back at a constant rate $\bar{A}-A^{0} .{ }^{1}$ These equations exhibit a general reaction-diffusion form and are similar to models of other biological systems, such as the Keller-Segel chemotaxis model, which are well studied in the literature (see, for example, $[22,16,39,9,13,24,36,15,8,31])$.

The continuum system described by (2.4) and (2.5) may display two of the three behaviors from the discrete system: stationary hotspots or no hotspots. We believe that transitory hotspots are not seen in this continuum approximation since they are the result of statistical noise that is removed by considering only expectation values in our limit. We showed that the formation of hotspots in this system may arise as a result of a linear instability of the homogeneous steady state

$$
\begin{equation*}
A(\mathbf{x}, t)=\bar{A}, \quad \rho(\mathbf{x}, t)=\bar{\rho} \equiv 1-\frac{A^{0}}{\bar{A}} \tag{2.6}
\end{equation*}
$$

toward perturbations of certain wavenumbers $\mathbf{k}$, and that the dispersion relation could be written as

$$
\begin{align*}
\sigma(\mathbf{k})=-[1+ & \left.\bar{A}-\bar{\rho}+|\mathbf{k}|^{2}(1+\eta)\right] / 2  \tag{2.7}\\
& +\sqrt{\left[1+\bar{A}-\bar{\rho}+|\mathbf{k}|^{2}(1+\eta)\right]^{2} / 4-\left(\eta|\mathbf{k}|^{4}-(3 \bar{\rho}-\eta \bar{A}-1)|\mathbf{k}|^{2}+\bar{A}\right)}
\end{align*}
$$

The instability criterion, therefore, could be written as

$$
\begin{equation*}
A^{0}<A_{*}^{0}=\frac{2}{3} \bar{A}-\frac{1}{3} \eta \bar{A}^{2}-\frac{2}{3} \bar{A} \sqrt{\eta \bar{A}} \tag{2.8}
\end{equation*}
$$

In other words, if the baseline attractiveness is less than some critical value $A_{*}^{0}$, then the homogeneous state will be linearly unstable (exhibit some modes with a positive $\sigma$ ). Finally, we showed that the maximally unstable mode $\mathbf{k}_{\text {max }}$ is given by

$$
\begin{align*}
\left|\mathbf{k}_{\max }\right|^{2}=(1-\bar{A}) /(1-\eta)-\bar{\rho}(5-\eta) /(1-\eta)^{2} &  \tag{2.9}\\
& \quad+\sqrt{\eta(1+\eta)^{2} \bar{\rho}[(\bar{A}(3-\eta)-2)(1-\eta)+2 \bar{\rho}(3-\eta)]} / \eta(1-\eta)^{2} .
\end{align*}
$$

[^1]Note for future reference that, when $A^{0}=A_{*}^{0}$, the maximally growing mode can be greatly simplified to

$$
\begin{equation*}
\left|\mathbf{k}_{\max }\right|^{2} \equiv\left|\mathbf{k}_{*}\right|^{2}=\sqrt{\frac{\overline{\bar{A}}}{\eta}} . \tag{2.10}
\end{equation*}
$$

3. Weakly nonlinear analysis. Our goal now is to more deeply examine the continuum system of (2.4) and (2.5) and to move beyond the simple linear stability analysis outlined above, thus providing the technical details of the qualitative results presented in [34], and for an even wider range of possibilities. We will accomplish this by means of a weakly nonlinear analysis, using a standard perturbative expansion approach to derive amplitude equations for our system [12, 38, 17, 25, 11].

We begin by considering a parameter regime such that the homogeneous state is linearly unstable (or stable). Choosing $A^{0}$ as our control parameter (as suggested by our form of the stability criterion given in (2.8)), we define a new parameter $\epsilon$ via the equation

$$
\begin{equation*}
A^{0}=A_{*}^{0}-\epsilon \bar{A}, \tag{3.1}
\end{equation*}
$$

such that the homogeneous state will be linearly unstable for positive $\epsilon$ and linearly stable for negative $\epsilon$, as indicated by the linear stability criterion in (2.8) above. For the remainder of our analysis we will assume that $\epsilon$ is small, but we note here that in theory $\epsilon$ can take on any value between some $\epsilon_{\min }$ (where $A^{0}=\bar{A}$, the most it could ever physically be) and $\epsilon_{\max }$ (where $A^{0}=0$, the least it could ever physically be). These two values are given by

$$
\begin{gather*}
\epsilon_{\min }=-\frac{1}{3}-\frac{1}{3} \eta \bar{A}-\frac{2}{3} \sqrt{\eta \bar{A}},  \tag{3.2}\\
\epsilon_{\max }=\frac{2}{3}-\frac{1}{3} \eta \bar{A}-\frac{2}{3} \sqrt{\eta \bar{A}}, \tag{3.3}
\end{gather*}
$$

with the difference between these always being 1 . We further point out that $\epsilon_{\max }$ becomes negative for any $\sqrt{\eta \bar{A}}=\eta k_{*}^{2}>\sqrt{3}-1$, meaning that above this threshold the homogeneous state is incapable of being linearly unstable and our analysis is invalid in this regime.

Returning again to the results of the linear stability analysis, when we substitute (3.1) into (2.7) and expand for small $\epsilon$, we find that the growth rate for the $\mathbf{k}_{*}$ mode is given by

$$
\begin{equation*}
\sigma\left(\mathbf{k}_{*}\right)=\sigma_{*} \epsilon+O\left(\epsilon^{2}\right) \tag{3.4}
\end{equation*}
$$

where

$$
\begin{equation*}
\sigma_{*}=\frac{9 \eta\left|\mathbf{k}_{*}\right|^{2}}{\left(1+\eta\left|\mathbf{k}_{*}\right|^{2}\right)\left[2 \eta+\eta\left|\mathbf{k}_{*}\right|^{2}(3-\eta)\right]} \tag{3.5}
\end{equation*}
$$

Using this result, we see that we can define a new, slow time variable $T=|\epsilon| t$ that describes the dynamics of the system when near the stability transition; this means that the $\partial_{t}$ in (2.4) and (2.5) becomes $|\epsilon| \partial_{T}$. We use $|\epsilon|$ here to make our future results valid regardless of the sign of $\epsilon$, though this means that we must rewrite (3.1) as

$$
\begin{equation*}
A^{0}=A_{*}^{0}-\operatorname{sign}(\epsilon)|\epsilon| \bar{A} \tag{3.6}
\end{equation*}
$$

At this point we define a new spatial variable $\tilde{x} \equiv\left|\mathbf{k}_{*}\right| x$ (but continue to refer to $\tilde{x}$ as $x$ in the future for notational simplicity) and rewrite (2.4) and (2.5) as

$$
\begin{align*}
& |\epsilon| \frac{\partial A}{\partial T}=\eta\left|\mathbf{k}_{*}\right|^{2} \nabla^{2} A-A+A_{*}^{0}-\operatorname{sign}(\epsilon) \eta\left|\mathbf{k}_{*}\right|^{4}|\epsilon|+\rho A,  \tag{3.7}\\
& |\epsilon| \frac{\partial \rho}{\partial T}=\left|\mathbf{k}_{*}\right|^{2} \vec{\nabla} \cdot\left[\vec{\nabla} \rho-\frac{2 \rho}{A} \vec{\nabla} A\right]-\rho A+\eta\left|\mathbf{k}_{*}\right|^{4}-A_{*}^{0}+\operatorname{sign}(\epsilon) \eta\left|\mathbf{k}_{*}\right|^{4}|\epsilon| . \tag{3.8}
\end{align*}
$$

Next, we express $A$ and $\rho$ as expansions in our small parameter of the form

$$
\begin{gather*}
A(\mathbf{x}, T)=\bar{A}+\sum_{j=1}^{\infty}|\epsilon|^{\alpha j} A^{(j)}(\mathbf{x}, T),  \tag{3.9}\\
\rho(\mathbf{x}, T)=1-\frac{A_{*}^{0}}{\bar{A}}+\sum_{j=1}^{\infty}|\epsilon|^{\alpha j} \rho^{(j)}(\mathbf{x}, T), \tag{3.10}
\end{gather*}
$$

where $\alpha$ is a rational number that will depend upon the specific geometry in which we are interested; the reasoning behind the choice of $\alpha$ will be presented with each geometry we consider. We substitute these expansions into our differential equations and then separate the resulting equations by powers of $|\epsilon|$. We note that, upon doing this, (3.7) can be used to simply solve for a given $\rho^{(j)}(\mathbf{x}, T)$ algebraically in terms of lower order $\rho^{\left(j^{\prime}\right)}(\mathbf{x}, T)$ and $A^{\left(j^{\prime}\right)}(\mathbf{x}, T)$ and their derivatives, and that this result can then be substituted into (3.8). This leaves a series of fourth order differential equations to be solved that involve only the various $A^{(j)}(\mathbf{x}, T)$, each of which is of the form

$$
\begin{equation*}
\left(\nabla^{2}+1\right)^{2} A^{(j)}(\mathbf{x}, T)=f_{j}\left[A^{(1)}(\mathbf{x}, T)\right], \tag{3.11}
\end{equation*}
$$

where $f_{j}$ is a possibly nonlinear function. Regardless of $\alpha$, the first of these equations is always

$$
\begin{equation*}
\left(\nabla^{2}+1\right)^{2} A^{(1)}(\mathbf{x}, T)=0 . \tag{3.12}
\end{equation*}
$$

3.1. 1D case. In this geometry, we restrict our solution to a domain $x \in[0, L]$, where $L=2 n \pi$ for some integer $n>0$, and impose periodic boundary conditions for both $A(x, T)$ and $\rho(x, T)$. The solution to (3.12) in this geometry and for these boundary conditions is

$$
\begin{equation*}
A^{(1)}(x, T)=P(T) e^{i x}+\text { c.c. }, \tag{3.13}
\end{equation*}
$$

where $P(T)$ is the amplitude, which at this point is simply an integration constant, and "c.c." denotes the complex conjugate. Due to the inversion symmetry of this solution $(P \rightarrow-P$ is physically the same but just shifted), we expect a pitchfork bifurcation to occur here; we use $\alpha=1 / 2$ to reflect this. The first interesting equation therefore occurs at order $|\epsilon|$ :

$$
\begin{equation*}
\left(\nabla^{2}+1\right)^{2} A^{(2)}(x, T)=\frac{4\left(1-\eta^{2} k_{*}^{4}\right)}{\eta^{2} k_{*}^{6}}\left[P(T)^{2} e^{2 i x}+\text { c.c. }\right] . \tag{3.14}
\end{equation*}
$$

The particular solution to this equation, which is all we are after, is

$$
\begin{equation*}
A^{(2)}(x, T)=\frac{4\left(1-\eta^{2} k_{*}^{4}\right)}{9 \eta^{2} k_{*}^{6}}\left[P(T)^{2} e^{2 i x}+\text { c.c. }\right] . \tag{3.15}
\end{equation*}
$$

At order $|\epsilon|^{3 / 2}$, we find the equation

$$
\begin{equation*}
\left(\nabla^{2}+1\right)^{2} A^{(3)}(x, T)=f_{3,1}\left[P(T) ; \eta, k_{*}\right] e^{i x}+f_{3,3}\left[P(T) ; \eta, k_{*}\right] e^{3 i x}+\text { c.c.; } \tag{3.16}
\end{equation*}
$$

we do not reproduce the full expressions for $f_{3, j}$ here for the sake of simplicity. Note that (3.16) contains a secular term $\propto e^{i x}$. In order for the particular solution of (3.16) to fit the periodic boundary conditions chosen, this secular term must vanish, meaning that $f_{3,1}\left[P(T) ; \eta, k_{*}\right]=0$. Upon enforcing this constraint, rescaling $T$ back to $t$, and letting $|\epsilon|^{1 / 2} P(t) \equiv Q(t)$, we find the amplitude equation

$$
\begin{equation*}
\dot{Q}=\sigma_{*} \epsilon Q-C_{1}\left(\eta, k_{*}\right)|Q|^{2} Q, \tag{3.17}
\end{equation*}
$$

where

$$
\begin{equation*}
C_{1}\left(\eta, k_{*}\right)=\frac{-8+56 \eta k_{*}^{2}-31 \eta^{2} k_{*}^{4}-8 \eta^{3} k_{*}^{6}}{3 \eta^{2} k_{*}^{8}\left[2 \eta+\eta k_{*}^{2}(3-\eta)\right]} \tag{3.18}
\end{equation*}
$$

and $\sigma_{*}$ is given by (3.5) above. As expected, this is the standard form for a dynamical system exhibiting a pitchfork bifurcation, with the distinction between a supercritical and subcritical bifurcation determined by the sign of $C_{1}$. Upon inspection, it is found that $C_{1}$ will be negative for any $\eta k_{*}^{2} \lesssim 0.157$ (indicating a subcritical pitchfork bifurcation) and positive otherwise (indicating a supercritical pitchfork bifurcation). The steady state value $Q_{s}$ is either zero (the homogeneous case) or given by

$$
\begin{equation*}
Q_{s}= \pm \sqrt{\frac{\sigma_{*} \epsilon}{C_{1}\left(\eta, k_{*}\right)}} \tag{3.19}
\end{equation*}
$$

Finally, our solution for $Q$ is valid only to order $|\epsilon|$, so our solution for $A(x, T)$ is also valid only to this order and is given by

$$
\begin{equation*}
A(x, t)=\bar{A}+Q(t) e^{i x}+\frac{4\left(1-\eta^{2} k_{*}^{4}\right)}{9 \eta^{2} k_{*}^{6}} Q(t)^{2} e^{2 i x}+\text { c.c. } \tag{3.20}
\end{equation*}
$$

One can in general continue the expansion up to higher orders in $\epsilon$ by defining subsequent slow timescales $T_{j}$ for $j \geq 2$, each of which will modify $\partial_{t}$ by adding a term $|\epsilon|^{j} \partial_{T_{j}}$. One then continues with the above results and eliminates the secular terms at higher orders in the expansion, with the net result being amplitude equations that govern the various $P_{T_{j}}$. For example, the next order amplitude equation for the 1D case takes the form

$$
\begin{equation*}
\dot{Q}=\sigma_{*} \epsilon\left[1+\epsilon a_{1}\left(\eta, k_{*}\right)\right] Q-C_{1}\left(\eta, k_{*}\right)\left[1+\epsilon a_{2}\left(\eta, k_{*}\right)\right]|Q|^{2} Q-C_{2}\left(\eta, k_{*}\right)|Q|^{4} Q \tag{3.21}
\end{equation*}
$$

where $a_{1}, a_{2}$, and $C_{2}$ are the new corrections that arise as we move to the higher order (the exact formulas for these expressions are unimportant here). We will refer to this higher order amplitude equation later when discussing numerical simulations to help explain some of the results seen there.
3.2. 2D, radially symmetric case. We now consider solutions on a disk $r \in[0, R]$ with $R=\beta_{1, n}$, where $\beta_{1, n}$ is the $n$th root of the Bessel function $J_{1}(r)$; we enforce Neumann conditions on the boundary edge. For these boundary conditions in this geometry, the solution to (3.12) is

$$
\begin{equation*}
A^{(1)}(r, T)=P(T) J_{0}(r) . \tag{3.22}
\end{equation*}
$$

In this regime, therefore, the inversion symmetry of the 1D case is broken $(P \rightarrow-P$ is physically different here), so we expect a transcritical bifurcation rather than a pitchfork bifurcation. Hence, we choose $\alpha=1$, and the first interesting equation in our system is proportional to $|\epsilon|^{2}$ :

$$
\begin{align*}
&\left(\nabla^{2}+1\right)^{2} A^{(2)}(r, T)  \tag{3.23}\\
&\left.=\frac{9 \eta k_{*}^{2} P(T) \operatorname{sign}(\epsilon)-\left(1+\eta k_{*}^{2}\right)[2 \eta}{3 \eta^{2} k_{*}^{4}}+\eta k_{*}^{2}(3-\eta)\right] P_{T}(T) \\
&+\frac{2\left(1-\eta^{2} k_{*}^{4}\right)}{\eta^{2} k_{*}^{6}} P(T)^{2}\left[J_{0}^{2}(r)-J_{1}^{2}(r)\right] .
\end{align*}
$$

As before, we will need to eliminate any secular term proportional to $J_{0}(r)$ on the righthand side of (3.23) so that our solution will respect the Neumann boundary conditions imposed. In order to do so, we take advantage of the fact that the Bessel functions can be used as an orthogonal basis for expanding other functions, so we are free to write the $J_{0}^{2}(r)-J_{1}^{2}(r)$ portion on the right as a sum of Bessel functions to the first power, one of which will be $J_{0}(r)$. With the definition that

$$
\begin{equation*}
q=\frac{2 \int_{0}^{R} r J_{0}(r)\left[J_{0}^{2}(r)-J_{1}^{2}(r)\right] d r}{R^{2} J_{0}^{2}(R)}, \tag{3.24}
\end{equation*}
$$

we see that setting the secular term to zero (and rescaling $T$ to $t$ and letting $|\epsilon| P(t) \equiv Q(t)$ ) is equivalent to the amplitude equation

$$
\begin{equation*}
\dot{Q}=\sigma_{*} \epsilon Q+C_{3}\left(\eta, k_{*}\right) Q^{2}, \tag{3.25}
\end{equation*}
$$

where

$$
\begin{equation*}
C_{3}\left(\eta, k_{*}\right)=\frac{6 q\left(1-\eta k_{*}^{2}\right)}{k_{*}^{2}\left[2 \eta+\eta k_{*}^{2}(3-\eta)\right]} . \tag{3.26}
\end{equation*}
$$

As expected, we find that in the 2D, radially symmetric case, our system will undergo a transcritical bifurcation. Interestingly, the constant $C_{3}$ will always be positive for any value of $\eta k_{*}^{2}$ for which the above analysis is valid. That is, $C_{3}$ would only be negative if $\eta k_{*}^{2}>1$, but the maximum value of $\eta k_{*}^{2}$ for which linear instability is at all possible (which must be the case for our analysis to work) is $\eta k_{*}^{2}=\sqrt{3}-1<1$, as discussed previously. Hence, there is really only one qualitatively distinct bifurcation diagram in this case. The steady state value $Q_{s}$ in this case is either zero (the homogeneous case) or

$$
\begin{equation*}
Q_{s}=-\frac{\sigma_{*} \epsilon}{C_{3}\left(\eta, k_{*}\right)} . \tag{3.27}
\end{equation*}
$$

As in the 1D case, our amplitude equation is valid only to order $|\epsilon|$, so our equation for $A(r, t)$ in this case is

$$
\begin{equation*}
A(r, t)=\bar{A}+Q(t) J_{0}(r) \tag{3.28}
\end{equation*}
$$

As stated above, in this geometry there is a physical difference between positive $Q$ and negative $Q$ solutions, with the former corresponding to a solution that exhibits a bump in $A$ at the origin (hereafter referred to as the "bump solution") and the latter corresponding to a solution that has a ring of high $A$ at the outer edge of the domain (hereafter referred to as the "ring solution"). Our theoretical results from (3.27) state that the steady state bump solution will exist only for negative $\epsilon$ and that it will be unstable, and that the steady state ring solution will exist only for positive $\epsilon$ and be stable.
3.3. Fully 2D case. For the fully 2D system, the first order equation (3.12) (along with suitable boundary conditions) admits solutions of the form

$$
\begin{equation*}
A^{(1)}(\mathbf{x}, T)=\sum_{j=1}^{N} P_{j}(T) e^{i \mathbf{q}_{j} \cdot \mathbf{x}}+\text { c.c. } \tag{3.29}
\end{equation*}
$$

for any $N$, so long as $\left|\mathbf{q}_{j}\right|=1$. We will limit our discussion here to the simple cases of rolls, squares, and hexagons, however, as they display simple periodicity. Rolls $(N=1)$, though, are just the 1D patterns discussed previously extended into a second dimension, so we need not perform any further analysis on them here.
3.3.1. Squares. We begin our analysis with squares: $N=2, \mathbf{q}_{1}=\mathbf{x}, \mathbf{q}_{2}=\mathbf{y}$, a domain $x \in\left[0, L_{x}\right], y \in\left[0, L_{y}\right]$, where $L_{x}=2 n \pi$ and $L_{y}=2 m \pi$ for some integers $n, m>0$, and periodic boundary conditions for both $A(x, T)$ and $\rho(x, T)$. The first order solution is therefore

$$
\begin{equation*}
A^{(1)}(\mathbf{x}, T)=P_{1}(T) e^{i x}+P_{2}(T) e^{i y}+\text { c.c. } \tag{3.30}
\end{equation*}
$$

Equation (3.30) displays the same inversion symmetry as the 1D case; hence we expect a pitchfork bifurcation here as well. Consequently, the mathematics in this case follow almost exactly as in the 1D case above, so we omit the small details. In the end, we arrive at the second order solution

$$
\begin{align*}
A^{(2)}(\mathbf{x}, T)=\frac{4\left(1-\eta^{2} k_{*}^{4}\right)}{9 \eta^{2} k_{*}^{6}}[ & P_{1}(T)^{2} e^{2 i x}+P_{2}(T)^{2} e^{2 i y}  \tag{3.31}\\
& \left.\quad+9 P_{1}(T) P_{2}(T) e^{i(x+y)}+9 P_{1}(T) P_{2}(T)^{*} e^{i(x-y)}+\text { c.c. }\right]
\end{align*}
$$

and amplitude equation for $Q_{1}$ (as in 1D, $Q_{1}=\sqrt{|\epsilon|} P_{1}$ )

$$
\begin{equation*}
\dot{Q}_{1}=\sigma_{*} \epsilon Q_{1}-C_{1}\left(\eta, k_{*}\right)\left|Q_{1}\right|^{2} Q_{1}+C_{S}\left(\eta, k_{*}\right)\left|Q_{2}\right|^{2} Q_{1} \tag{3.32}
\end{equation*}
$$

Here, $C_{1}\left(\eta, k_{*}\right)$ is as given in (3.18) above and

$$
\begin{equation*}
C_{S}\left(\eta, k_{*}\right)=\frac{6\left(8-8 \eta k_{*}^{2}-9 \eta^{2} k_{*}^{4}+8 \eta^{3} k_{*}^{6}\right)}{\eta^{2} k_{*}^{8}\left[2 \eta+\eta k_{*}^{2}(3-\eta)\right]} \tag{3.33}
\end{equation*}
$$

the amplitude equation for $Q_{2}$ is symmetric with (3.32), with the subscripts 1 and 2 switched.
The steady states for squares $\left(Q_{1}=Q_{2} \equiv Q_{s}\right)$ are therefore $Q_{s}=0$ or

$$
\begin{equation*}
Q_{s}= \pm \sqrt{\frac{\sigma_{*} \epsilon}{C_{1}\left(\eta, k_{*}\right)-C_{S}\left(\eta, k_{*}\right)}} . \tag{3.34}
\end{equation*}
$$

For all valid values of $\eta k_{*}^{2}, C_{S}>0$, and for all values of $\eta k_{*}^{2} \lesssim 0.7, C_{S}>\left|C_{1}\right|$. Hence, squares almost always developed through a subcritical pitchfork bifurcation from the homogeneous state, only developing through a supercritical pitchfork over the narrow range $0.7 \lesssim \eta k_{*}^{2}<$ $\sqrt{3}-1$. In either case, the roll steady state is unstable to squares. This is because the roll steady state is simply the square system with $Q_{1}=Q_{s}$ from (3.19) and $Q_{2}=0$. This fixed point is always unstable along the direction of increasing $Q_{2}$, however, since the eigenvalue in this direction is $\lambda_{2}=\sigma_{*} \epsilon\left(1+C_{S} / C_{1}\right)$, which is always positive. That is, either $C_{1}, \epsilon>0$ (a supercritical roll), where $\lambda_{2}$ is clearly positive since $C_{S}>0$, or $C_{1}, \epsilon<0$ (a subcritical roll), where $\lambda_{2}$ is positive since $C_{S}>\left|C_{1}\right|$.
3.3.2. Hexagons. The final patterns we will examine are hexagons: $N=3, \mathbf{q}_{1}=\mathbf{x}$, $\mathbf{q}_{2}=-\frac{1}{2} \mathbf{x}+\frac{\sqrt{3}}{2} \mathbf{y}, \mathbf{q}_{3}=-\frac{1}{2} \mathbf{x}-\frac{\sqrt{3}}{2} \mathbf{y}$, a domain $x \in\left[0, L_{x}\right], y \in\left[0, L_{y}\right]$, where $L_{x}=4 n \pi$ and $L_{y}=4 m \pi / \sqrt{3}$ for some integers $n, m>0$, and periodic boundary conditions for both $A(x, T)$ and $\rho(x, T)$. The first order solution is therefore

$$
\begin{equation*}
A^{(1)}(\mathbf{x}, T)=P_{1}(T) e^{i x}+P_{2}(T) e^{i(-x / 2+\sqrt{3} y / 2)}+P_{3}(T) e^{i(-x / 2-\sqrt{3} y / 2)}+\text { c.c. } \tag{3.35}
\end{equation*}
$$

The lack of inversion symmetry in this case, along with the resonance $\mathbf{q}_{1}+\mathbf{q}_{2}+\mathbf{q}_{3}=\mathbf{0}$, will generally cause our amplitude equations to display a quadratic nonlinearity, as seen in the radially symmetric system above, leading to a transcritical bifurcation. Correspondingly, we choose $\alpha=1$ in this case and follow a derivation similar to the radially symmetric case. We find that the amplitude equation for $Q_{1}$ (recall that $Q_{1}=|\epsilon| P_{1}$ in this case) is given by

$$
\begin{equation*}
\dot{Q}_{1}=\sigma_{*} \epsilon Q_{1}+C_{H}\left(\eta, k_{*}\right) Q_{2}^{*} Q_{3}^{*}, \tag{3.36}
\end{equation*}
$$

where

$$
\begin{equation*}
C_{H}\left(\eta, k_{*}\right)=\frac{6\left(1-\eta k_{*}^{2}\right)}{k_{*}^{2}\left[2 \eta+\eta k_{*}^{2}(3-\eta)\right]} ; \tag{3.37}
\end{equation*}
$$

the amplitude equations for $Q_{2}$ and $Q_{3}$ are given by cyclic permutations of the indices 1,2 , and 3 in (3.36).

The valid hexagonal steady states ( $Q_{1}=Q_{2}=Q_{3} \equiv Q_{s}$ ) are $Q_{s}=0$ or

$$
\begin{equation*}
Q_{s}=-\frac{\sigma_{*} \epsilon}{C_{H}\left(\eta, k_{*}\right)} . \tag{3.38}
\end{equation*}
$$

Note that $C_{H}$ is positive for all valid values of $\eta k_{*}^{2}$ and is in fact just $C_{3}$ (see (3.26)) without the factor of $q$, meaning that the bifurcation diagram for hexagons will be quite qualitatively similar to that of the radially symmetric state for very small $|\epsilon|$ and amplitudes. Therefore, and for future reference, we will continue to denote the positive $Q$ solutions as "bumps" and


Figure 2. $1 D$ system with $\eta k_{*}^{2}=0.4$. In (a) is a bifurcation diagram for the system where dashed lines represent unstable branches and solid lines are stable branches; numerical results are in black, and analytic results are in red (with circles). We find very good agreement for essentially all $\epsilon$ values. In (b) are plots of the numeric (black) and analytic (red with circles) solutions for $A_{\text {amp }}(t)$ with $\epsilon=0.01$ and $Q(0)=0.01$; there is very good agreement here as well.
the negative $Q$ solutions as "rings" even in the hexagonal geometry. However, unlike the radially symmetric system, the hexagonal ring solution is not stable but is instead a saddle point, meaning that the only stable hexagonal steady state predicted by this approximation is $Q_{s}=0$ for $\epsilon<0$. However, the quadratic nonlinearity of the hexagons dominates over the cubic nonlinearity of the rolls and squares near onset; hence, hexagons are the preferred pattern for our model at small $|\epsilon|$. This is not surprising, as it is a generic feature of systems that lack inversion symmetry [12], as ours does.
4. Numerical results. As a verification of our analytical results above, we numerically solve our model system in various geometries. For the dynamical system, we use a fully implicit Newton-Raphson-based solver; for the steady state solutions, we use a Newton-Raphson-based relaxation method. For each case, we look at a quantity we will refer to as simply the "amplitude" of $A$, which we define as

$$
\begin{equation*}
A_{\mathrm{amp}}(t)=\sqrt{\frac{1}{|\mathbb{D}|} \int_{\mathbb{D}}[A(\mathbf{x}, t)-\bar{A}]^{2} d \mathbf{x}} \tag{4.1}
\end{equation*}
$$

where $\mathbb{D}$ is the domain of the simulation, and $|\mathbb{D}|$ is its size. Our measure is, therefore, essentially a root mean square (RMS) measure of the attractiveness field's deviation from the homogeneous steady state. Finally, for all simulations in this section, we employ Neumann boundary conditions.
4.1. 1 D case. In this geometry, our domain $\mathbb{D}$ is $x \in\left[0, \pi / k_{*}\right]$. The first case we explore is a supercritical system, in which $\eta k_{*}^{2}=0.4\left(\eta=0.1\right.$ and $\left.k_{*}=2\right)$. The two plots in Figure 2 summarize the results here. Figure 2(a) shows a bifurcation diagram for our system as derived by computing the steady state value $A_{\text {amp }}(\infty)$ as a function of $\epsilon$, plotting both the analytical and numerical results. We find that there is good agreement in this case for essentially all $\epsilon$ values. Figure 2(b) plots the analytic and numerical solutions for $A_{\text {amp }}(t)$ using $\epsilon=0.01$ and $Q(0)=0.01$; there is very good agreement here as well.


Figure 3. $1 D$ system with $\eta k_{*}^{2}=0.1$. In (a) is a bifurcation diagram for the system where dashed lines represent unstable branches and solid lines represent stable branches; numerical results are in black, analytic results from (3.19) are in red (with circles), and a higher order analytic solution as in (3.21) is in blue (with squares). There is good agreement between the numerics and both analytic solutions for smaller $\epsilon$ values along the unstable branch, but only the higher order analytic solution predicts the existence of the large amplitude branch seen in the numerics. In (b) are plots of the numeric (black) and analytic (red circles and blue squares) solutions for $A_{\text {amp }}(t)$ with $\epsilon=-0.001$ and varying $Q(0)$; the dashed line represents the analytic steady state value for this $\epsilon$. The lower line corresponds to $Q(0)=0.0028$, with the analytic solution from (3.19) in red (with circles); there is good agreement here between the two. The upper line corresponds to $Q(0)=0.0032$, with the higher order analytic solution in blue (with squares). The agreement between these two is reasonable, though the analytic solution predicts a higher steady state value than the numerics.

The next case we explore is a subcritical system, in which $\eta k_{*}^{2}=0.1\left(\eta=0.1\right.$ and $\left.k_{*}=1\right)$, with the results shown in Figure 3. Referring to Figure 3(a), the numerical solutions (black) display the small amplitude, unstable branch predicted by the theory above, and the numerics match the theory (red with circles) well at small $\epsilon$. However, there is also a stable, large amplitude branch in the bifurcation diagram that is not predicted by the theory above. As alluded to before, however, if we continue our analytic solution to the next higher order in $\epsilon$ as in (3.21) (blue with squares), we can predict the location of the secondary bifurcation where the upper and lower branches meet. Note, however, that this higher order amplitude equation is not necessarily valid along the upper branch seen in the numerics, which explains the substantial deviation seen there. Figure 3(b) shows the evolution of $A_{\text {amp }}(t)$ using $\epsilon=-0.001$ and two different values for $Q(0)$. The first value is $Q(0)=0.0028$, which is just slightly below the unstable branch, so we expect our analytic results above (red with circles) to be close to the numerical results (black). However, the second $Q(0)$ is 0.0032 , which is slightly above the unstable branch, so our results above cannot be used. Instead, we compare with the higher order analytic result (blue with squares), and find reasonably good agreement until $t \approx 800$. After this time the validity of (3.21) is clearly lost, as the true solution exhibits oscillatory behavior before settling down to the steady state, which is something our amplitude equation could not predict.
4.2. 2D, radially symmetric case. In this geometry, $\mathbb{D}$ is $r \in\left[0, \beta_{1,1} / k_{*}\right]$. Furthermore, we have chosen a convention whereby bump solutions are shown with a positive $A_{\text {amp }}$ value, while ring solutions are shown with a negative one. Simulations use $\eta k_{*}^{2}=0.2(\eta=0.01$ and $k_{*}=2 \sqrt{5}$ ), and the results are shown in Figure 4. In Figure 4(a), we see our bifurcation


Figure 4. Radially symmetric $2 D$ system with $\eta k_{*}^{2}=0.2$. In (a) is a bifurcation diagram for the system where dashed lines represent unstable branches and solid lines represent stable branches; numerical results are in black, and analytic results are in red (with circles). We find good agreement between the two for smaller $\epsilon$ values, though the numerics display a large amplitude stable branch that the analytic solution does not. In (b) are plots of the numeric (black lines) and analytic (red with circles) solutions for $A_{\text {amp }}(t)$ with $|\epsilon|=0.01$ and varying $Q(0)$. The horizontal dashed line indicates the analytical unstable steady state for $\epsilon<0$. The lower line corresponds to $Q(0)=-0.01, \epsilon>0$, and the middle line corresponds to $Q(0)=0.3, \epsilon<0$; the agreement is good in these two cases. The upper line corresponds to $Q(0)=0.36, \epsilon<0$. This is above the unstable branch, so it grows to the large amplitude stable branch, which is not available from our analytical results.
diagram for this geometry, which exhibits a transcritical bifurcation near the origin (black) that matches the theory (red with circles) well at small $|\epsilon|$. However, the numerics also display a large amplitude, stable bump solution that our theory does not predict. Unlike the subcritical 1D case above, we do not extend to higher order approximations here. This large amplitude branch indicates that both the bump and ring steady state solutions are stable and available at positive $\epsilon$ values, with the bump also being available in both a stable form and an unstable form over some range of negative $\epsilon$ values. Figure 4(b) shows the evolution of both the numeric (black) and analytic (red with circles) $A_{\text {amp }}(t)$ using $|\epsilon|=0.01$ and three different values for $Q(0)$. The first and lowest value is a ring with $Q(0)=-0.01$ (and positive $\epsilon$ ), which compares well with the analytic results. The second, intermediate value is a bump with $Q(0)=0.3$ (and negative $\epsilon$ ), which is just slightly below the unstable branch, so we expect our analytic result above to work reasonably well in this case, and it does. However, the final value is a bump with $Q(0)=0.036$ (and negative $\epsilon$ ), which is slightly above the unstable branch, so our analytic results above cannot be used. Numerically, though, we see that the solution grows until it reaches the stable, large amplitude branch.

### 4.3. Fully 2D case.

4.3.1. Squares. In this geometry, $\mathbb{D}$ is $x \in\left[0, \pi / k_{*}\right], y \in\left[0, \pi / k_{*}\right]$. The first case illustrated is a supercritical system, in which $\eta k_{*}^{2}=1 / \sqrt{2}>0.7\left(\eta=0.01\right.$ and $\left.k_{*}=10 / 2^{1 / 4}\right)$. The two plots in Figure 5 summarize the results here. Figure 5(a) shows the bifurcation diagram for this system at very small $\epsilon$ values, and we find that there is good agreement at the smallest $\epsilon$ values. Figure 5(b) plots the analytic and numerical solutions for $A_{\text {amp }}(t)$ using $\epsilon=10^{-6}$ and $Q(0)=10^{-3}$; there is very good agreement here, as $\epsilon$ is very small in this case.

The next case is a subcritical system, in which $\eta k_{*}^{2}=0.2\left(\eta=0.01\right.$ and $\left.k_{*}=2 \sqrt{5}\right)$, with


Figure 5. Fully $2 D$, square system with $\eta k_{*}^{2}=1 / \sqrt{2}$. In (a) is a bifurcation diagram for the system where dashed lines represent unstable branches and solid lines represent stable branches; numerical results are in black, and analytic results are in red (with circles). We find good agreement for the smallest $\epsilon$ values. In (b) are plots of the numeric (black) and analytic (red with circles) solutions for $A_{\text {amp }}(t)$ with $\epsilon=10^{-6}$ and $Q(0)=10^{-3}$; there is very good agreement here.


Figure 6. Fully $2 D$, square system with $\eta k_{*}^{2}=0.2$. In (a) is a bifurcation diagram for the system where dashed lines represent unstable branches and solid lines represent stable branches; numerical results are in black, and analytic results are in red (with circles). There is good agreement between the numeric and analytic solutions for smaller $\epsilon$ values along the unstable branch, but the actual solution also displays a large amplitude, stable branch. In (b) are plots of the numeric (black) and analytic (red with circles) solutions for $A_{\text {amp }}(t)$ with $\epsilon=-0.001$ and varying $Q(0)$; the dashed line represents the analytic steady state value for this $\epsilon$. The lower line corresponds to $Q(0)=0.025$, and there is good agreement here between the two. The upper line corresponds to $Q(0)=0.027$, and the solution in this case rapidly grows to the large amplitude steady state.
the results shown in Figure 6. Referring to Figure 6(a), the numerical solution (black) displays the small amplitude, unstable branch predicted by the theory above, and the numerics match the theory (red with circles) well at small $\epsilon$. However, there is also a stable, large amplitude branch in the bifurcation diagram that is not predicted by the theory, but which is similar to that seen in cases above. Figure 6(b) shows the evolution of $A_{\text {amp }}(t)$ using $\epsilon=-0.001$ and two different values for $Q(0)$. The first value is $Q(0)=0.025$, which is just slightly below the unstable branch, so we expect our analytic results above (red with circles) to be close to the numerical results (black), and they are. However, the second $Q(0)$ is 0.027 , which is slightly


Figure 7. Fully $2 D$, hexagonal system with $\eta k_{*}^{2}=0.2$. In (a) is a bifurcation diagram for the system where dashed lines represent unstable branches and solid lines represent stable branches; numerical results are in black, and analytic results are in red (with circles). We find good agreement between the two for smaller $\epsilon$ values, though the numerics display both a large amplitude stable branch and secondary instabilities along the ring solution that the analytic solution does not. In (b) are plots of the numeric (black lines) and analytic (red with circles) solutions for $A_{\text {amp }}(t)$ with $|\epsilon|=0.001$ and varying $Q(0)$. The horizontal dashed line indicates the analytical unstable steady state for $\epsilon<0$. The lower line corresponds to $Q(0)=-1.25 \times 10^{-4}, \epsilon>0$, with good agreement until $t \approx 2500$ (the ring solution is a saddle point in this geometry). The middle line corresponds to $Q(0)=6.125 \times 10^{-3}, \epsilon<0$, and the agreement is relatively good in this case. The upper line corresponds to $Q(0)=6.375 \times 10^{-3}, \epsilon<0$, which is above the unstable branch, so it grows to the large amplitude stable branch.
above the unstable branch, so our analytic results cannot be used; we find here that the system rapidly approaches the large amplitude steady state seen in the bifurcation diagram.
4.3.2. Hexagons. In this geometry, $\mathbb{D}$ is $x \in\left[0,2 \pi / k_{*}\right], y \in\left[0,2 \pi / \sqrt{3} k_{*}\right]$. As in the radially symmetric geometry, we have chosen a convention whereby bump solutions are shown with a positive $A_{\text {amp }}$ value, while ring solutions are shown with a negative one. Simulations use $\eta k_{*}^{2}=0.2\left(\eta=0.01\right.$ and $\left.k_{*}=2 \sqrt{5}\right)$, and the results are shown in Figure 7. In Figure 7(a), we see our bifurcation diagram for this geometry, which exhibits a transcritical bifurcation near the origin (black) that matches the theory (red with circles) well at small $|\epsilon|$. However, the numerics display both a large amplitude stable bump solution and secondary bifurcations along the ring branch that our theory above cannot predict. As in the radially symmetric case, the large amplitude branch indicates that the bump solution is stable and available at all positive $\epsilon$ values, in addition to some range of negative $\epsilon$ values. The ring branch follows an interesting series of secondary bifurcations as $\epsilon$ varies: the ring first breaks up into separate spots that grow in amplitude, the spots then begin moving toward each other to form a rectangle, and the spots of the rectangle eventually merge into the roll solution (the final continuous portion displayed for this branch is in fact just the roll solution). Of course, all of these solutions are unstable and, therefore, unlikely to be observed as steady states in any dynamical simulation.

Figure $7(\mathrm{~b})$ shows the evolution of both the numeric (black) and analytic (red with circles) $A_{\text {amp }}(t)$ using $|\epsilon|=10^{-3}$ and three different values for $Q(0)$. The first and lowest value is a ring with $Q(0)=-1.25 \times 10^{-4}$ (and positive $\epsilon$ ), which compares well with the analytic results until $t \approx 2500$. Due to the saddle point nature of the ring, the numerics eventually
begin to diverge from the analytic solution and end up at the large amplitude, stable state (albeit shifted from the standard bump solution due to the initial conditions). The second, intermediate value is a bump with $Q(0)=6.125 \times 10^{-3}$ (and negative $\epsilon$ ), which is just slightly below the unstable branch, so we expect our analytic result above to work reasonably well in this case, and it does. However, the final value is a bump with $Q(0)=6.375 \times 10^{-3}$ (and negative $\epsilon$ ), which is slightly above the unstable branch, so our analytic results above cannot be used. Numerically, though, we see that the solution grows until it reaches the stable, large amplitude branch.
5. Hotspot suppression. Now that we know our system may exhibit two qualitatively different types of crime hotspots (supercritical and subcritical) it is natural to question what differences may exist, if any, between the behavior of these two classes of pattern with regards to hotspot suppression. As mentioned in the introduction, "hotspot policing" is a law-enforcement strategy whereby more police resources are focused on areas currently believed to be within a hotspot in an effort to disrupt and destroy said hotspot. Field studies conducted to test the effectiveness of this strategy reveal that in some instances the hotspots seem to be destroyed, while in others they seem to simply be displaced. The 2D analyses we have performed above seem to offer an explanation as to why these two very different responses to suppression occur (refer to Figures 4 and 7). First, imagine a crime hotspot that exists within a linearly stable parameter regime $(\epsilon<0)$; the hotspot is therefore subcritical. If the police presence is enough to drive the attractiveness of the hotspot below the critical unstable branch of the bump solution in the bifurcation diagram, the system will tend to naturally drop down to the homogeneous state once suppression is relaxed, destroying the hotspot in question utterly. However, imagine now that the hotspot in question exists within the linearly unstable regime $(\epsilon>0)$ and is therefore supercritical. Any effort to suppress the bump solution will simply lead to the attractiveness being displaced to the surrounding area, i.e., a ring-like solution. Of course, in a fully 2 D system, the ring solution will not be stable and will break up into separate spots, leading to a system that looks similar to the original one, but with the hotspots shifted to nearby regions. In this case, then, the hotspot policing will have simply led to a displacement of the hotspot to nearby areas rather than its destruction.

The above hypothetical scenarios have been verified in computer simulations of the radially symmetric 2 D continuous system and the fully 2 D system for both the continuous and the discrete crime models. To do so, we choose a combination of parameters that are known to make the homogeneous steady state either linearly unstable or stable, whichever is desired. Then, we run simulations as described above in section 4 with initial conditions set to give a bump solution at the origin (considered to be the center of the field in the fully 2D case). We allow the simulation to run until a time $t_{s}$ when it seems to have reached a steady state, at which point we begin the suppression. This is accomplished by first defining an instantaneous damping field $d(\mathbf{x})$ in the following way:

$$
\begin{equation*}
d(\mathbf{x})=\frac{1}{2}\left[1-\tanh \left[\kappa\left(A\left(\mathbf{x}, t_{s}\right)-A_{\text {cutoff }}\right)\right]\right] \tag{5.1}
\end{equation*}
$$

where $\kappa$ sets the width of the transition region between total suppression and no suppression and $A_{\text {cutoff }}$ sets the attractiveness value above which suppression is desired. This damping


Figure 8. Suppression in the radially symmetric $2 D$ system with $\eta k_{*}^{2}=0.316, \kappa=3$. The curves show $A(r, t)$ as it evolves following the suppression that occurs at $t=0$, and the horizontal dashed line represents $A_{\text {cutoff }}=1$. Shown in (a) is the case $\epsilon=0.4$, and the suppression of the bump drives the system to the ring solution, which persists after the suppression is removed. Shown in (b) is the case $\epsilon=-0.02$, and the suppression of the bump drives the system to a temporary ring structure that decays to homogeneity once the suppression is removed.
field is meant to represent police presence, which is concentrated almost exclusively in the areas of high attractiveness (hotspots). We assume that this presence has two effects. First, the damping field will reduce the crime rate in areas where there is a large police presence ( $d \simeq 1$ ). Second, the police presence will prevent burglars from beginning their search in these same areas. Mathematically, then, our PDE system is modified to

$$
\begin{align*}
& \frac{\partial A}{\partial t}=\eta \nabla^{2} A-A+A^{0}+d \rho A,  \tag{5.2}\\
& \frac{\partial \rho}{\partial t}=\vec{\nabla} \cdot\left[\vec{\nabla} \rho-\frac{2 \rho}{A} \vec{\nabla} A\right]-d \rho A+d\left(\bar{A}-A^{0}\right) . \tag{5.3}
\end{align*}
$$

Note that this damping field remains unchanged between any two successive $t_{s}$ values. In other words, the police may remain within an area for some time even after the crime there has been reduced. This is reasonable in the sense that in the real world, police do not have instantaneous information about what areas are most attractive and must instead rely on where events have occurred in the recent past when deciding where to allocate resources. Therefore, there is an inherent lag between the information possessed by the criminals and that possessed by the police. The typical timescale for this lag in the real world may be on the order of weeks to months [26], which is enough time for new hotspots to emerge [32, 5]. Of course, this damping method is only one of many possible choices, some of which are explored in [21]. However, we suspect that for a large range of damping models, especially those with an appreciable temporal lag between criminal events and decisions on where police resources should be allocated, the basic outcomes described here will remain; subcritical spots may be destroyed, but supercritical spots will simply move.

Results for the radially symmetric case are shown in Figure 8, and the hypothetical scenarios play out as anticipated. In the supercritical case, suppression of the bump drives the system to the ring solution, which, due to its stability, remains even after suppression is


Figure 9. Suppression results for the fully $2 D$ system with $\eta k_{*}^{2}=0.2, \epsilon=-0.05, \kappa=5$, and $A_{\text {cutoff }}=5.72$. The top row shows results from the PDE system, while the bottom row shows results from the discrete system with equivalent parameters. These colormaps display high $A$ in red, $\bar{A}$ in green, and low $A$ in blue to purple. Shown in (a) is the system configuration right before suppression is first implemented. Soon after implementation, the central hotspot has disappeared entirely, but no further spots have emerged (b). Eventually the suppression is lifted, and the system begins to adopt the homogeneous steady state (c).
relaxed (Figure 8(a)). Suppression of the subcritical bump initially sends the system to a ring-like state as well, since the suppression by definition will cause the origin to have very low $A$ values, leaving the outer edge as the only place for criminal activity to occur. However, once the suppression is removed, the ring's instability causes it to decay to the homogeneous state, and the original hotspot is now destroyed (Figure 8(b)).

Figure 9 illustrates the effects of hotspot suppression in a fully 2D, subcritical system with periodic boundary conditions. Before suppression (Figure 9(a)), we see that our initial condition has led to a stable hotspot in the center of the field in both the continuum and discrete cases, though the discrete case also displays some quasi-hotspots near the edges of the domain due to random fluctuations that push the system at least temporarily above the unstable branch. Once suppression is introduced (Figure 9(b)), the hotspot dies away rather quickly, leaving an area of very low $A$ in the center where the police presence remains and a faint ring near the domain edges. Critically, though, we do not see the emergence of new hotspots. Finally, when the next $t_{s}$ is reached (Figure 9(c)), there is actually no suppression needed since no hotspots remain, and the "coldspot" in the center returns to the


Figure 10. Suppression results for the fully $2 D$ system with $\eta k_{*}^{2}=0.2, \epsilon=0.05, \kappa=5$, and $A_{\text {cutoff }}=6.12$. The top row shows results from the PDE system, while the bottom row shows results from the discrete system with equivalent parameters. These colormaps display high $A$ in red, $\bar{A}$ in green, and low $A$ in blue to purple. Before suppression is first implemented, the system displays a number of hotspots (a). Soon after the implementation of suppression, the original hotspots vanish, but the attractiveness of the neighboring regions correspondingly increases, leading to a transient, ring-like structure that surrounds the location of the original central hotspot (b). By the time the next suppression time $t_{s}$ has arrived, a new steady state featuring hotspots near the original ones has been achieved (c).
homogeneous value soon after the police leave the area. As predicted, the suppression was effective in eradicating the hotspot in the subcritical case.

Figure 10 illustrates the effects of hotspot suppression in a fully 2D, supercritical system with periodic boundary conditions. Before suppression (Figure 10(a)), we see not only that our initial condition has led to a hotspot in the center of the field but that a number of other hotspots have developed near the edge due to the linear instability of the system. Once suppression is introduced (Figure 10(b)), the original hotspots disappear quickly. However, we see in the continuum case especially that the eradication of these spots has simply pushed the system into a different nonhomogeneous configuration, with a temporary ring-like structure surrounding the area where the central hotspot was located. Finally, by the time the next $t_{s}$ has arrived (Figure 10(c)), the system has reached a new steady state that exhibits hotspots in areas near where the original spots were. So, the suppression was ineffective in eradicating the supercritical hotspots and merely led to their displacement.
6. Conclusions. Through a weakly nonlinear analysis of our coupled PDE system ((2.4) and (2.5)), we have shown that in both the 1D and 2D cases, our system may exhibit stable hotspots in both the supercritical and subcritical regimes. The existence of the subcritical hotspots offers another mechanism for crime pattern formation, in addition to the linear instability discussed in our previous work.

Importantly, these distinct hotspot mechanisms may help explain the varying measures of success that police agencies have when attempting to suppress hotspots. In the supercritical case, suppression of a hotspot seems to simply displace the spot to neighboring regions, as the bump solution gives way to the ring solution, which either will be a new stable state (in the radially symmetric case) or will then break up into separate bumps (in the fully 2D case); this is illustrated in Figures 8(a) and 10. In the subcritical case, on the other hand, the suppression of the hotspot below the unstable bump solution branch of the bifurcation diagram (Figures 4 and 7) should destroy it completely, as the only other stable state available in this regime is the homogeneous one; this is illustrated in Figures 8(b) and 9.

As a corollary to this argument, we point out that the existence of these large amplitude branches introduces the possibility of hysteresis into the system. That is, if the parameters of the system are slowly varying with time (as social or economic conditions vary, perhaps), what was once a peaceful city ( $\epsilon<0$, homogeneous state) may experience a sudden burst of crime once the stability threshold is passed (as $\epsilon>0$ ), rather than the crime slowly increasing as the parameters move further into the unstable regime (as would happen if the system were supercritical). Once in this linearly unstable state, police attempts at suppression may only have the effect of displacing crime hotspots. Furthermore, even if the parameters then change in such a way that $\epsilon$ begins to decrease, this high level of crime may persist until the relevant parameters are even lower than when the initial outbreak occurred; though once the stability threshold is passed (so that $\epsilon<0$ ), police suppression should help in destroying hotspots.

Acknowledgments. The authors would like to thank George Tita, Lincoln Chayes, Paul Jones, and the Los Angeles Police Department for helpful discussions.

## REFERENCES

[1] L. Anselin, J. Cohen, D. Cook, W. Gorr, and G. Tita, Spatial analyses of crime, in Criminal Justice 2000, Vol. 4, D. Duffee, ed., National Institute of Justice, Washington, DC, 2000, pp. 213-262.
[2] M. D. Betterton, Theory of structure formation in snowfields motivated by penitentes, suncups, and dirt cones, Phys. Rev. E, 63 (2001), 056129.
[3] K. J. Bowers and S. D. Johnson, The Role of Publicity in Crime Prevention: Findings from the Reducing Burglary Initiative, Home Office, London, 2003.
[4] A. A. Braga, The effects of hot spots policing on crime, Ann. Amer. Acad. Political Social Sci., 578 (2001), pp. 104-125.
[5] A. A. Braga, The crime prevention value of hot spots policing, Psicothema, 18 (2006), pp. 630-637.
[6] A. A. Braga, D. L. Weisburd, E. J. Waring, L. Green-Mazerolle, W. Spelman, and F. GaJAEWSKI, Problem-oriented policing in violent places: A randomized controlled experiment, Criminology, 37 (1999), pp. 541-580.
[7] P. J. Brantingham and P. L. Brantingham, Anticipating the displacement of crime using the principles of environmental criminology, Crime Prevention Stud., 16 (2003), pp. 119-148.
[8] M. Burger, M. Di Francesco, and Y. Dolak-Struss, The Keller-Segel model for chemotaxis with prevention of overcrowding: Linear vs. nonlinear diffusion, SIAM J. Math. Anal., 38 (2006), pp. 1288-1315.
[9] H. M. Byrne and M. R. Owen, A new interpretation of the Keller-Segel model based on multiphase modelling, J. Math. Biol., 49 (2004), pp. 604-626.
[10] P. Y. Chan and N. Goldenfeld, Steady states and linear stability analysis of precipitation pattern formation at geothermal hot springs, Phys. Rev. E, 76 (2007), 046104.
[11] M. Cross and H. Greenside, Pattern Formation and Dynamics in Nonequilibrium Systems, Cambridge University Press, Cambridge, UK, 2009.
[12] M. C. Cross and P. C. Hohenberg, Pattern formation outside of equilibrium, Rev. Modern Phys., 65 (1993), pp. 851-1112.
[13] M. del Pino and J. Wei, Collapsing steady states of the Keller-Segel system, Nonlinearity, 19 (2006), pp. 661-684.
[14] J. M. Epstein and R. L. Axtell, Growing Artificial Societies: Social Sciences from the Bottom Up, The Brookings Institution, Washington, DC, 1996.
[15] C. Escudero, The fractional Keller-Segel model, Nonlinearity, 19 (2006), pp. 2909-2918.
[16] M. A. Herrero and J. J. L. Velázquez, Chemotactic collapse for the Keller-Segel model, J. Math. Biol., 35 (1996), pp. 177-194.
[17] R. B. Hoyle, Pattern Formation: An Introduction to Methods, Cambridge University Press, Cambridge, UK, 2006.
[18] S. D. Johnson and K. J. Bowers, The burglary as clue to the future: The beginnings of prospective hot-spotting, European J. Criminology, 1 (2004), pp. 237-255.
[19] S. D. Johnson And K. J. Bowers, Domestic burglary repeats and space-time clusters: The dimensions of risk, European J. Criminology, 2 (2005), pp. 67-92.
[20] S. D. Johnson, K. J. Bowers, and A. Hirschfield, New insights into the spatial and temporal distribution of repeat victimisation, British J. Criminology, 37 (1997), pp. 224-241.
[21] P. A. Jones, P. J. Brantingham, and L. R. Chayes, Statistical models of criminal behavior: The effects of law enforcement actions, Math. Models Methods Appl. Sci., to appear.
[22] E. F. Keller and L. A. Segel, Initiation of slime mold aggregation viewed as an instability, J. Theoret. Biol., 26 (1970), pp. 399-415.
[23] G. L. Kelling, The Kansas City Preventative Patrol Experiment: A Technical Report, Police Foundation, Kansas City, MO, 1974.
[24] S. Luckhaus and Y. Sugiyama, Large time behavior of solutions in super-critical cases to degenerate Keller-Segel systems, M2AN Math. Model. Numer. Anal., 40 (2006), pp. 597-621.
[25] P. Manneville, Instabilities, Chaos, and Turbulence, Imperial College Press, London, 2004.
[26] M. H. Moore and A. A. Braga, Measuring and improving police performance: The lessons of compstat and its progeny, Policing Internat. J. Police Strategies Management, 26 (2003), pp. 439-453.
[27] S. V. Petrovskil and H. Malchow, A minimal model of pattern formation in a prey-predator system, Math. Comput. Modelling, 29 (1999), pp. 49-63.
[28] T. Reichenbach, M. Mobilia, and E. Frey, Mobility promotes and jeopardizes biodiversity in rock-paper-scissors games, Nature, 448 (2007), pp. 1046-1049.
[29] T. A. Reppetto, Crime prevention and displacement phenomenon, Crime \& Delinquency, 22 (1976), pp. 166-167.
[30] T. C. Schelling, Dynamic models of segregation, J. Math. Sociol., 1 (1971), pp. 143-186.
[31] T. Senba, Type II blowup of solutions to a simplified Keller-Segel system in two dimensional domains, Nonlinear Anal., 66 (2007), pp. 1817-1839.
[32] L. Sherman and D. Rogan, Deterrent effects of police raids on crack houses: A randomized controlled experiment, Justice Quart., 12 (2006), pp. 755-782.
[33] M. B. Short, J. C. Baygents, J. W. Beck, D. A. Stone, R. S. Toomey, and R. E. Goldstein, Stalactite growth as a free-boundary problem: A geometric law and its platonic ideal, Phys. Rev. Lett., 94 (2005), 018501.
[34] M. B. Short, P. J. Brantingham, A. L. Bertozzi, and G. E. Tita, Dissipation and displacement of hotspots in reaction-diffusion models of crime, Proc. Natl. Acad. Sci. USA, 107 (2010), pp. 3961-3965.
[35] M. B. Short, M. R. D’Orsogna, V. Pasour, P. J. Brantingham, G. Tita, A. Bertozzi, and L. Chayes, A statistical model of criminal behavior, Math. Models Methods Appl. Sci., 18 (2008), pp. 1249-1267.
[36] Y. Sugiyama, Global existence in sub-critical cases and finite time blow-up in super-critical cases to degenerate Keller-Segel systems, Differential Integral Equations, 19 (2006), pp. 841-876.
[37] A. M. Turing, The chemical basis of morphogenesis, Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci., 237 (1952), pp. 37-72.
[38] M. van Hecke, P. C. Hohenberg, and W. van Saarloos, Amplitude equations for pattern forming systems, in Fundamental Problems in Statistical Mechanics, Vol. VIII, H. van Beijeren and M. H. Ernst, eds., North-Holland, Amsterdam, 1994, pp. 245-278.
[39] J. J. L. Velázquez, Well-posedness of a model of point dynamics for a limit of the Keller-Segel system, J. Differential Equations, 206 (2004), pp. 315-352.
[40] D. Weisburd and J. E. Eck, What can police do to reduce crime, disorder, and fear?, Ann. Amer. Acad. Political Social Sci., 593 (2004), pp. 42-65.
[41] D. Weisburd, L. A. Wyckoff, J. Ready, J. E. Eck, J. C. Hinkle, and F. Gajewski, Does crime just move around the corner? A controlled study of spatial displacement and diffusion of crime control benefits, Criminology, 44 (2006), pp. 549-592.


[^0]:    *Received by the editors May 11, 2009; accepted for publication (in revised form) by M. Silber March 17, 2010; published electronically May 21, 2010. This work was supported by NSF grants BCS-0527388, DMS-0907931, and DMS-0968309; ARO MURI grant 50363-MA-MUR; and the Department of Defense.
    http://www.siam.org/journals/siads/9-2/75906.html
    ${ }^{\dagger}$ Department of Mathematics, UCLA, Los Angeles, CA 90095 (mbshort@math.ucla.edu, bertozzi@math.ucla.edu).
    ${ }^{\ddagger}$ Department of Anthropology, UCLA, Los Angeles, CA 90095 (pjb@anthro.ucla.edu).

[^1]:    ${ }^{1}$ The choice of the notation $\bar{A}$ here is due to the fact that, at steady state, this quantity is indeed the spatially averaged value of $A(\mathbf{x})$, regardless of the other parameters or whether hotspots are displayed or not.

