International Journal of Pure and Applied Mathematics

Volume 19 No. 2 2005, 195-244

GEOMETRIC AND STOCHASTIC ANALYSIS OF REACTION-DIFFUSION PATTERNS

Jianhong Shen[1] §, Yoon Mo Jung[2]

[1,2]School of Mathematics
University of Minnesota
206 Church Street, S.E., Minneapolis, MN 55455, USA
[1]e-mail: jhshen@math.umn.edu
[2]e-mail: ymjung@math.umn.edu

Abstract: After Turing's ingenious work on the chemical basis of morphogenesis fifty years ago, reaction-diffusion patterns have been extensively studied in terms of modelling and analysis of pattern formations (in both chemistry and biology), pattern growing in complex laboratory environments, and novel applications in computer graphics. A fundamental question that remains unanswered in the literature is what one precisely means by (reaction-diffusion) *patterns*. Most patterns have only been discovered, identified, or explained by human vision and human intelligence. Inspired by the recent advancement in Mathematical Image and Vision Analysis (MIVA), the current paper develops both geometric and stochastic tools and frameworks for identifying, classifying, and characterizing common reaction-diffusion patterns and pattern formations. In essence, it presents a data mining theory for the scientific simulations of reaction-diffusion patterns, or various analytical tools for the automatic characterization of generic complex patterns by artificial intelligence.

AMS Subject Classification: 68T45, 92C15, 62H35
Key Words: reaction-diffusion, Gray-Scott model, turing instabilities, entropy, maturity, skewness, kurtosis, geometric measures, total curvatures, skeleton curves, singularities

Received: January 3, 2005

§Correspondence author

Dedicated to Gil Strang
on the occasion of his 70-th birthday.

1. Introduction and Motivation

1.1. Turing and Hopf Reaction-Diffusion Patterns

Fifty years ago in his seminal paper [64], Turing surprised the entire scientific community by theoretically justifying morphogen pattern formation in a *diffusive* chemical reaction environment. By intuition, as a smoothing and filtering process broadly applied even in image and vision analysis [1, 49], diffusion would wipe out any inhomogeneous spatial patterns like the stripes on zebra skins [43]. Turing, however, brilliantly identified Mother Nature's behavior as the counterintuitive consequence of the *non-commutativity* of linear operators [47], the very same mechanism underlying Heisenberg's uncertainty principle in quantum mechanics [28] (though Turing originally did not express it in this way).

Consider a planar two-species model as an example. Locally near a *stable* uniform kinetic equilibrium, let D denote the positive diffusivity matrix, and K the kinetic matrix whose eigenvalues both have *negative* real parts. Then the one-parameter family of (linearized) reaction-diffusion operators:

$$L_\mu = -\mu D + K, \qquad \mu \geq 0, \tag{1}$$

could lose stability for some positive μ, i.e., at least one eigenvalue of L_μ could cross the imaginary axis from the left (the parameter μ is proportional to the eigenvalues of $-\nabla^2$, the negative Laplacian). This can never happen if reaction commutes with diffusion: $[K, D] = KD - DK = 0$. For a non-commutative reaction-diffusion pair such as

$$K = \begin{bmatrix} 2 & 7 \\ -1 & -3 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}, \quad \text{one has } L_1 = \begin{bmatrix} 1 & 7 \\ -1 & -8 \end{bmatrix},$$

and $\det L_1 = -1$. As a result, the positive eigenvalue will stir up Turing instability.

Such linear instability, once stabilized by nonlinear kinetics, could sustain stable spatial or temporal patterns such as zebra stripes. In the current paper, we shall only work with planar two-species reaction-diffusion models in the common form of

$$\begin{cases} u_t = D_1 \Delta u + f(u, v), \\ v_t = D_2 \Delta v + g(u, v), \end{cases} \tag{2}$$

where f and g denote the nonlinear kinetics.

Reaction-diffusion patterns also frequently arise when Hopf instability occurs, i.e., when a focusing spiral loses its stability and a stable limit cycle is born from bifurcation. In fact, many simulation evidences (e.g., Roy Williams' comprehensive computational results obtained on Caltech's supercomputing facilities [68]) seem to show that reaction-diffusion patterns are more easily observed on the region of Hopf instability.

After Turing's pioneering work, exactly four decades elapsed before a remarkable experiment design by Castets et al [3] in 1990 eventually confirmed Turing patterns in laboratory. The tremendous challenge had lied in the design of reaction-diffusion systems like (2) in a controlled and analyzable manner in actual complex environments.

Coincidentally, almost at the same time, computer scientists started to pick up the reaction-diffusion mechanism for graphic designs. Around 1991, Turk [65] and Witkin and Kass [69] artistically applied reaction-diffusion systems like (2) for generating a variety of animal skin patterns in graphic applications. Like flow simulations based on the Navier-Stokes equation, such biochemical simulations of life phenomena surely expect a bright future.

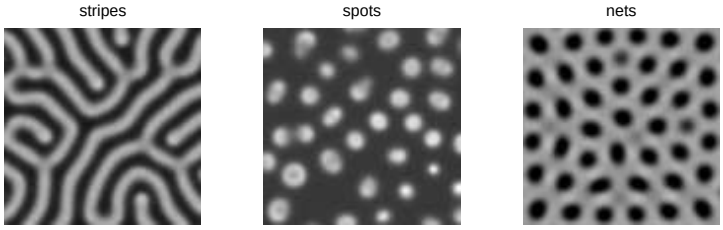

Figure 1: Some typical examples of reaction-diffusion patterns

1.2. Quest for the Meaning of "Pattern"

Despite all the afore mentioned remarkable historical works on the theoretical modelling, laboratory growing, or graphic simulations of reaction-diffusion patterns, some key elements have still been surprisingly missing, or blindly taken for granted for a long time. That is, for example, what does one mean precisely by reaction-diffusion *patterns*, and how to identify and classify them, and characterize their detailed properties using artificial intelligence (i.e., via computation) instead of human vision?

Human vision and intelligence have reached so superior a stage after mil-

lions of years of evolution, that their swift and effortless recognition of complex patterns starts to deceive most scientists. In the entire reaction-diffusion literature, for example, it has become household to identify and call some common patterns by "spots," "hexagons," or "stripes." But just as in the case of face or voice recognition, such identification and classification cannot be so trivial as they appear since our central nerve systems do carry out massive computation for these succinct descriptions.

As David Mumford put it in his most recent perspective article for ICM 2002 [40], *pattern theory*, a booming interdisciplinary field, is really about the mathematics of *perception*. Its mission is to reveal and model all the key elements and processes of human perception and intelligence, which are involved in a split-second action of pattern recognition. To achieve this goal, many mathematicians are working intensively on the proper mathematical frameworks for effectively learning, classifying, and analyzing general complex patterns (see for example, Grenander [27], Mumford [40], and Poggio and Smale [50]).

1.3. Pattern Analysis Meets Pattern Formation

On the other hand, the trend is growing stronger that pattern analysis theory and pattern formation research should be tightly coupled. This is mainly driven by the digital and information technologies, and the thirst of being more automatic and independent of human involvement in many areas of data formation and analysis (such as in numerous astronomical and medical applications). Take the simulation of Navier-Stokes fluids for example. Many researchers are striving to design efficient algorithms for automatically reporting and analyzing important flow features and patterns such as vortex sheets (see, for example, Thompson et al [63] and the references therein) from massive 4-D (space plus time) simulation data. Another remarkable example is molecular dynamics. Someday in the future when the complete folding processes of proteins or DNA chains are successfully simulated, could a computer win the Nobel Prize by automatically recognizing some crucial yet unidentified patterns (as analogous to the double-helix pattern discovered by James Watson and Francis Crick [67] based on Rosalind Franklin's X-ray images of DNA)?

It is guided by the very same spirit that the current work has been carried out. We attempt to apply novel tools recently developed in image and vision analysis, as well as in general pattern theory, to analyze and characterize generic Turing and Hopf reaction-diffusion patterns.

Since reaction-diffusion patterns are generated from the evolutionary partial differential equations often with initial random inputs, our tools are mainly

statistical as well as based on *geometric measures*. Statistical tools such as histograms, invariant moments, and the entropy, allow to efficiently differentiate and classify major spatial patterns. On the other hand, tools from geometric measures lead to robust schemes for computing important geometric features of each class of patterns, such as the average radius of a spots pattern and the total length of a stripes pattern, which can be very difficult to estimate even for human vision.

1.4. Contributions and Organization

We now highlight what we believe the most significant contributions of the current work.

1. Our approach combines systematically many important tools in modern Mathematical Image and Vision Analysis (MIVA), and could foster further broad applications of MIVA techniques in numerous data mining and analysis problems in physical, chemical, or biological simulations.

2. For the first time, generic reaction-diffusion patterns are analyzed both statistically and geometrically. Many novel notions and tools are developed based on the intrinsic characteristics of these patterns.

3. Although mainly focused on the patterns from the Gray-Scott model [48], most statistical and geometric tools developed here apply to more general reaction-diffusion models such as CIMA [30, 31] and Brusselator [51]. This is achieved by inventing the corresponding morphological models that are independent of the biochemical kinetics.

4. We introduce the notion of *entropy* from thermodynamics and statistical mechanics to monitor the dynamic behavior of pattern formations. Analysis of typical pattern data leads to the exciting statistical discovery that on average the entropy steadily increases, and converges when the target pattern is stabilized or *matures*.

While this inspirational connection to the *Second Law of Thermodynamics* awaits further explanation in the future, it enables us to introduce the important notion of "maturity" for developing patterns. The determination of maturity solves a long standing problem in the literature of reaction-diffusion simulations, namely, the problem of *stopping time*. To the best knowledge of the authors, no rigorous metric has ever been introduced before us for automatically stopping the evolution when patterns mature.

5. The geometric features identified by the current work could make substantial contributions to the solution of another long standing problem in bio-

chemical system identification, namely, to test modelling hypotheses and identify biochemical parameters such as diffusion coefficients.

Consider an idealized leopard for an oversimplified explanation. Suppose its spotty skin pattern is well modelled by the reaction-diffusion system (2). Let d and N denote the average radius and spots density (per square inch, say) separately. Suppose that predictive relations are available in the form of (e.g., Turing's formula for intrinsic chemical wavelength [64]):

$$d = \phi(D_1, D_2) \quad \text{and} \quad N = \psi(D_1, D_2),$$

where D's stand for the diffusivity coefficients (for illustration only, all the kinetic parameters involved are assumed known). Then the measurement of d and N could help identify the two diffusivity coefficients D_1 and D_2, which are often extremely difficult or costly to obtain for an on-going chemical reaction or a living leopard.

From surface features and patterns to their underlying driving-force parameters, such inverse problems clearly demonstrate the value of pattern analysis in physical, biological, or chemical systems.

The organization of the paper goes as follows. In Section 2, the working reaction-diffusion model of the current paper – the Gray-Scott system, is analyzed in terms of both Hopf and Turing instabilities. To our best knowledge, the existence of Turing instability appears for the first time in the literature.

Various statistical tools are developed in Section 3 for the monitoring of pattern formations and for the identification and classification of matured patterns. One of the most significant empirical discoveries is that pattern formations seem to obey the *Second Law of Thermodynamics*, namely, the entropy increases monotonically. Based on entropies, we are able to introduce the key quantities of pattern *maturity* and the *dynamic activity* of matured patterns. On the other hand, direct as well as geometry based histograms and their statistics successfully characterize and classify typical reaction-diffusion patterns. We also explain why most of the key statistical features are intrinsic and independent of the reaction-diffusion systems that generate them.

Pattern analysis based on geometric measure theory is presented in Section 4. Important geometric measures such as total areas, total variations, and total curvatures are employed to compute the key geometric features of common patterns, which include: (a) the average radius of a (matured) spot pattern, (b) the total number of spots in a given reaction-diffusion domain, (c) the average width of a stripe pattern, (d) the total length of *all* the stripes in a stripe pattern, and (e) the number of singular points (i.e. endpoints and Y-shaped branching points) in a stripe pattern. The estimation errors of the leading

term approximations are all intrinsically controlled by the signal-to-noise ratios (SNR, for spot and net patterns) or the aspect ratios (for stripe patterns).

Section 5 contains a brief conclusion.

2. Hopf and Turing Patterns via the Gray-Scott Model

2.1. Gray-Scott Model

There are a number of well known planar two-species reaction-diffusion models in the literature, including the Brusselator model [51], and the CIMA (chlorite-iodide-malonic-acid) reaction [30, 31]. Throughout the current paper, the major example to be focused on is the Gray-Scott model, which was first studied phenomenologically in John Pearson's science paper [48], and has also been comprehensively computed using Caltech's supercomputing facilities by Roy Williams [68]. Both the kinetics and reported patterns of the Brusselator model are very similar.

The Gray-Scott model is a variant of the autocatalytic Selkov model of glycolysis and addresses the following hypothetical reactions involving two chemical morphogens:

$$U + 2V \to 3V \quad \text{and} \quad V \to P. \tag{3}$$

Here the product P is an inert precipitate, U is fed into the system by a constant feeding rate, and V removed by the same rate. After a step of partial dimensionless reduction, the reaction-diffusion equations for the concentrations u and v become:

$$\begin{cases} u_t = D_1 \Delta u - uv^2 + F(1-u), \\ v_t = D_2 \Delta v + uv^2 - (F+k)v, \end{cases} \tag{4}$$

where F corresponds to the feeding and extraction rate, and k the rate constant for the second reaction. Following Pearson [48], we shall also simulate the model with diffusivity coefficients:

$$D_1 = 2 \times 10^{-5} \quad \text{and} \quad D_2 = 10^{-5}.$$

There are two good reasons why this specific choice is a clever one. First, as clear in the next section, near the unique kinetic equilibrium that nurtures pattern formation, u acts as the inhibitor and v the activator. Then the choice is consistent with the famous mechanism of Turing pattern formation known as *long-range inhibition and short-range self-activation* [43, 46]. Secondly, for the

numerical simulation of the system, the well known CFL (Courant-Friedrichs-Levy) stability condition requires (for 2-D diffusion equations):

$$\frac{\Delta t}{(\Delta x)^2}\max(D_1, D_2) \le \frac{1}{4}.$$

Since the spatial grid size Δx is often in the order of $O(10^{-2})$, the diffusivity order $O(10^{-5})$ is therefore the highest possible for fast numerical marching in the order of $\Delta t = O(1)$.

It shall be assumed that the reaction-diffusion domain Ω is a Cartesian rectangle, with no feeding or extraction through its boundary. That is, the Neumann conditions are imposed: $\partial u/\partial \boldsymbol{n} = \partial v/\partial \boldsymbol{n} = 0$ along $\partial\Omega$.

Discussed next are the equilibrium analysis and Hopf and Turing instabilities. Related work has also appeared in [32]. The two major contributions here are: (a) A new pair of normalized parameters (α, δ) is introduced based on the raw kinetic parameters (F, k). The former leads to simpler solution representation and much better phase plane resolution; (b) We show for the first time that Turing instability does occur in the Gray-Scott model, which has never been identified either theoretically or computationally in the previous works [48, 32] due to nonlinear complexity.

2.2. Kinetic Equilibria

At kinetic equilibrium,

$$\begin{cases} 0 = -uv^2 + F(1-u), \\ 0 = +uv^2 - (F+k)v. \end{cases} \tag{5}$$

It is easy to see that $(u, v) = (1, 0)$ is the trivial equilibrium. Assume now $v \neq 0$. Define as in [32]

$$\delta = \frac{F+k}{F} = 1 + \frac{k}{F} > 1. \tag{6}$$

Then the addition of the two equilibrium equations leads to $u + \delta v = 1$, and the second equation gives

$$u(\delta v) = \alpha, \quad \text{with } \alpha = \delta(F+k) = F\delta^2 > 0. \tag{7}$$

Thus a nontrivial equilibrium is subject to the canonical quadratic system:

$$u + \delta v = 1 \quad \text{and} \quad u(\delta v) = \alpha, \tag{8}$$

in which u and δv are symmetric. Denote the two nontrivial equilibria by

$$\left(u_+ = \frac{1}{2} - \sqrt{\frac{1}{4} - \alpha}, \quad \delta v_+ = \frac{1}{2} + \sqrt{\frac{1}{4} - \alpha} \right), \quad \text{and} \tag{9}$$

$$\left(u_- = \frac{1}{2} + \sqrt{\frac{1}{4} - \alpha}, \quad \delta v_- = \frac{1}{2} - \sqrt{\frac{1}{4} - \alpha} \right). \tag{10}$$

They are valid as long as α is less than the critical value $\alpha_c = 1/4$, or equivalently,

$$k < k_c = \frac{\sqrt{F}}{2} - F = \frac{1}{16} - (\sqrt{F} - \frac{1}{4})^2, \tag{11}$$

which is an upsidedown parabola $k_c = k_c(f)$ with respect to $f = \sqrt{F}$ (see Figure 2).

2.3. Kinetic Stability and Hopf Bifurcation

The trivial kinetic equilibrium $(1, 0)$ is a stable node since by dropping the second order terms $uv^2 = O(v^2)$, the kinetic system decouples to

$$u_t = F(1 - u) \quad \text{and} \quad v_t = -(F + k)v.$$

For the other two nontrivial ones (9) and (10), notice that the kinetic Jacobian is

$$K = K(u, v) = \begin{bmatrix} -v^2 - F & -2uv \\ v^2 & 2uv - (F + k) \end{bmatrix}.$$

With the definitions of α and δ, and the equilibrium conditions

$$u(\delta v) = \alpha \quad \text{and} \quad (\delta v)^2 - (\delta v) + \alpha = 0,$$

the kinetic matrix simplifies to

$$\begin{aligned} &\delta^2 K = \begin{bmatrix} -\delta v & -2\alpha\delta \\ \delta v - \alpha & \alpha\delta \end{bmatrix}, \quad \text{with} \\ &\text{trace}(\delta^2 K) = \alpha\delta - \delta v \quad \text{and} \quad \det(\delta^2 K) = \alpha\delta(\delta v - 2\alpha). \end{aligned} \tag{12}$$

As a result, the signs of $\text{trace}(K)$ and $\det(K)$ are the same as

$$(\alpha\delta - \delta v) \quad \text{and} \quad (\delta v - 2\alpha). \tag{13}$$

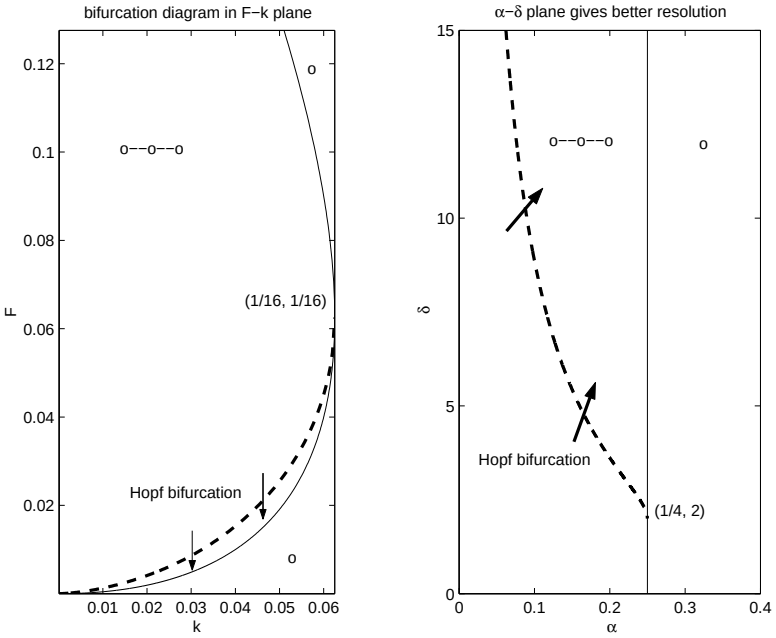

Figure 2: Bifurcation diagrams in both the $k-F$ plane and $\alpha-\delta$ plane. The solid lines indicate the onset of equilibrium bifurcation (with the little circles symbolically representing the number of equilibria). The dashed ones indicate the onset of Hopf bifurcation (with the arrows pointing toward the birth of stable limit cycles). It is clear that the $\alpha-\delta$ plane better resolves (or differentiates) these two different types of bifurcation

At (u_+, v_+) in (9), $\delta v_+ - 2\alpha > 0$ since

$$\delta v_+ > \frac{1}{2} \quad \text{and} \quad 2\alpha < \frac{1}{2}.$$

Thus $\det(K_+) > 0$ with $K_+ = K(u_+, v_+)$. Setting the trace to 0 yields $\alpha\delta - \delta v_+ = 0$, which reveals the onset of Hopf bifurcation:

$$\delta_c = \delta_c(\alpha) = \frac{\delta v_+}{\alpha} = \frac{1}{\alpha}\left(\frac{1}{2} + \sqrt{\frac{1}{4} - \alpha}\ \right).$$

For any fixed $\alpha < \alpha_c = \frac{1}{4}$, as δ increasingly crosses $\delta_c(\alpha)$, $\text{trace}(K_+)$ changes from negative sign to positive. As a result, (u_+, v_+) experiences a Hopf bifurcation (see Figure 2 and Figure 3).

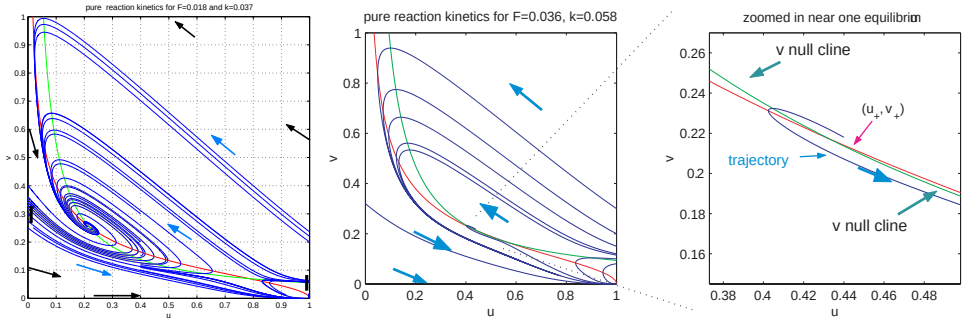

Figure 3: Stream lines of the kinetic reaction system

At (u_-, v_-), notice that

$$\delta v_- = \frac{\alpha}{u_-} < \frac{\alpha}{1/2} = 2\alpha.$$

implying that $\det K(u_-, v_-) < 0$, and (u_-, v_-) is a saddle point.

Figure 2 shows the kinetic bifurcation diagrams in both the $k-F$ parameter plane and the $\alpha-\delta$ plane. While the two bifurcations stay too close in the $k-F$ plane, they are well unfolded and resolved in the $\alpha-\delta$ plane. Figure 3 displays the streamlines of the kinetic system in two different sets of parameters. The leftmost one clearly outlines the stable spiral (u_+, v_+) before Hopf bifurcation, while the middle one depicts the streamlines when (u_+, v_+) loses stability. The rightmost is a zoomed-in version of the middle one.

2.4. Existence of Turing Instability

Turing instability of the Gray-Scott model has not been reported in any of the previous works [48, 32], largely due to the complexities caused by nonlinearity, including the way the kinetic parameters are involved in the equilibria.

To overcome the difficulty, we take a *demand-and-supply* based unconventional approach, instead of the traditional exhaustive analysis. That is, we preselect a set of eigenmodes for potential Turing instability (i.e., the demand), and then investigate whether there indeed exist reaction-diffusion parameters (i.e., the supply) to meet the instability demand.

For numerical convenience, assume the reaction-domain is $\Omega = [0, 2] \times [0, 2]$, and the diffusion coefficients are $D_1 = 2 \times 10^{-5} = 2D_2$, as explained in Section 2.1. Then the eigenmodes of the Laplacian (with Nuemann boundary

conditions) on the domain are the tensor products

$$\phi_{n,m}(x,y) = \cos(n\pi\frac{x}{2})\ \cos(m\pi\frac{y}{2}), \quad n,m = 0,1,2,\cdots,$$

corresponding to eigenvalues

$$\mu_{n,m} = (n^2+m^2)(\frac{\pi}{2})^2, \quad n,m \geq 0.$$

Suppose it is demanded that all the following eigenmodes lose stability in the reaction-diffusion:

$$12 \leq \frac{\mu_{n,m}}{\pi^2} \leq 25, \tag{14}$$

which will at least include the following eigenmodes:

$$\begin{array}{ccccc} \cdot & \cdot & (4,6) & (4,7) & (4,8) \\ \cdot & (5,5) & (5,6) & (5,7) & (5,8) \\ (6,4) & (6,5) & (6,6) & (6,7) & (6,8) \\ (7,4) & (7,5) & (7,6) & (7,7) & \cdot \\ (8,4) & (8,5) & (8,6) & \cdot & \cdot \end{array} \tag{15}$$

The motivation is clear: when satisfied, it will roughly lead to *several* standing waves in both x and y directions, which is good for observation in simulations – neither too dense nor too loose. Since $\pi^2 = 9.8796\ldots$, the instability demand (14) is contained in the more workable interval

$$100 \leq \mu \leq 250. \tag{16}$$

Now consider the combined reaction-diffusion system when (u_+, v_+) is a *stable* equilibrium, i.e., when

$$\alpha < \alpha_c = \frac{1}{4} \quad \text{and} \quad \delta < \delta_c(\alpha) = \frac{1}{\alpha}\left(\frac{1}{2} + \sqrt{\frac{1}{4}-\alpha}\ \right), \tag{17}$$

as previously derived. In particular, observe that

$$\delta_c(\alpha) > \frac{1}{2\alpha} > 2 \ \text{ for } \ \alpha < \alpha_c, \quad \text{and} \quad \delta_c(\alpha_c) = 2.$$

Let $L_+ = K_+ - \mu D$ denote the linearized reaction-diffusion at (u_+, v_+), where as in (12),

$$\delta^2 K_+ = \begin{bmatrix} -w & -2\alpha\delta \\ w-\alpha & \alpha\delta \end{bmatrix}, \quad D = \begin{bmatrix} D_1 & 0 \\ 0 & D_2 \end{bmatrix}, \quad w = \delta v_+. \tag{18}$$

Notice that the signs of the kinetic matrix K_+ are in the form of $[-,-;+,+]$, the type in the reaction-diffusion literature known as *cross*-activator-and-inhibitor [43, 46]. Define $\lambda = \delta^2\mu$ and

$$L_\lambda = \delta^2 L_+ = \begin{bmatrix} -w - \lambda D_1 & -2\alpha\delta \\ w - \alpha & \alpha\delta - \lambda D_2 \end{bmatrix} \tag{19}$$

Then Turing instability occurs if and only if $\det(L_\lambda) < 0$.

Theorem 1. *Suppose that $D_1 = 2\times D_2 = 2\times 10^{-5}$, then Turing instability indeed occurs over the range (16) for any $(\alpha, \delta = 2)$ with $\alpha = \frac{1}{4} - \varepsilon^2$ and $0 < \varepsilon \le 0.003$.*

A careful check of the proof below confirms that by only using a three-digit decimal, 0.003 is the best value for the upper limit. Such narrow concentration of Turing's instability near the critical point $(\alpha = \frac{1}{4}, \delta = 2)$ in the parameter plane is well supported by Roy Williams' comprehensive computational results [68].

Proof. When $\delta = 2$, over the range $100 \le \mu \le 250$, we have $\lambda = \delta^2\mu \in [400, 1000]$. Define $\eta = \lambda D_2 \in [0.004, 0.01]$. Then under the specified conditions,

$$0.1 < \alpha < \frac{1}{4} \quad \text{and} \quad \eta < 0.1\alpha. \tag{20}$$

Also define $q = w - \alpha\delta$. Then

$$q = \frac{1}{2} + \sqrt{\frac{1}{4} - \alpha} - 2\alpha = \varepsilon + 2\varepsilon^2 < 0.0031. \tag{21}$$

Finally, by the definition of L_λ,

$$\begin{aligned} \det L_\lambda &= \alpha\delta(w - 2\alpha) - \lambda(\alpha\delta D_1 - wD_2) + \lambda^2 D_1 D_2 \\ &= (\alpha\delta + \lambda D_2)q - \alpha^2\delta(2 - \delta) - \alpha\delta\lambda D_2 + 2(\lambda D_2)^2 \\ &= (2\alpha + \eta)q - 2\alpha\eta + 2\eta^2 \\ &< 2.1\alpha q - 2\alpha\eta + 0.2\alpha\eta \qquad \text{by (20)} \\ &= (2.1q - 1.8\eta)\alpha \\ &\le (2.1 \times 0.0031 - 1.8 \times 0.004)\alpha < 0, \end{aligned}$$

which completes the proof. □

3. Statistical Analysis of Reaction-Diffusion Patterns

Our approach has been profoundly influenced by the recent progress in image, vision, and pattern analysis.

To analyze complex natural images such as wooden textures, satellite images, tissues or more general medical images, many statistical tools including statistical mechanics have been introduced and extended by image analysis experts. Spatial image patterns are often treated as Gibbs canonical ensembles or Markov random fields, whose short-range energies are defined through spatial filters and parametric or non-parametric learning. This framework was pioneered by Geman and Geman, Mumford, Zhu, and their collaborators [23, 70, 71, 41].

For reaction-diffusion patterns, the statistical approach is intrinsically justified. Although most such patterns are either spatially or temporally varying from the deterministic point of view, statistically they do seem stationary. Take some spots patterns for example. The sizes and locations of the spots often look quite arbitrary, and also some spots are constantly created while others being annihilated. But overall the spots clearly exhibit some statistical invariance, and remarkably resemble a box of gas molecules in thermal equilibrium. Consequently, statistics becomes a natural tool.

On the other hand, compared with general images with multiscale or multifrequency features, reaction-diffusion patterns resulting from nonlinear parabolic equations carry more geometric regularity. Therefore, nonlinear geometric filters such as the curvature could play a better role than conventional linear ones.

3.1. Entropy and Second Law of Thermodynamics

Ever since Shannon's famous paper on the mathematical theory of communication and information [56], the notion of *entropy* from statistical mechanics has been playing a universally crucial role in information theory, signal and image analysis [15], and pattern learning and analysis [70, 71]. Information entropy (or the negative Shannon information measure), is a natural metric for measuring the degree of "thermal" disorder in information.

The *Second Law of Thermodynamics*, on which the entire theory of statistical mechanics is supported, says that any adiabatic (i.e., Neumann or no flux exchange across the boundaries) dynamic process must result in an increment in entropy, which leads to the *maximum entropy principle* [24]. In this section, statistical evidences are presented to illustrate the remarkable *Second Law of*

Thermodynamics type behavior during the formation of reaction-diffusion patterns.

From now on, we shall only focus on the "activator" $v = v(x, y, t)$ (see the kinetic matrix in (18)) of the reaction-diffusion system:

$$\begin{cases} u_t = D_1 \Delta u - uv^2 + F(1-u), \\ v_t = D_2 \Delta v + uv^2 - (F+k)v, \end{cases} \tag{22}$$

with generic initial conditions $u(x, y, 0) = u_0(x, y)$ and $v(x, y, 0) = v_0(x, y)$, and Neumann adiabatic conditions $\partial u/\partial \boldsymbol{n} = 0$ and $\partial v/\partial \boldsymbol{n} = 0$ along the boundary $\partial\Omega$. The initial conditions of (22) are usually specified by random fields near the nontrivial uniform equilibrium:

$$u_0(x, y) = u_+ + n(x, y) \quad \text{and} \quad v_0(x, y) = v_+ + m(x, y),$$

where $n(x, y)$ and $m(x, y)$ denote white noises whose variances are well dominated by the signals (u_+^2, v_+^2) (i.e., with a large signal-to-noise ratio). As practiced in [48], one could also replace u_+ and v_+ in the initial data by some compactly supported data.

The value of the activator $v(x, y, t)$ is understood as the activation level at site (x, y). Although the system is formulated deterministically, with random fields as the initial inputs, as well as in the cellular level where diffusion becomes Brownian motion, the solution $v(x, y, t)$ is better treated as a random field, which has been well practiced in image analysis [23, 70, 71]. For each constant $\theta \in (0, \infty)$, define the θ-level set

$$\Gamma(\theta|t) = \{\ (x, y) \in \Omega \mid v(x, y, t) \equiv \theta\ \} \tag{23}$$

(notice that the first quadrant of the uv-plane is kinetically invariant, which makes it sufficient to only consider $v \in (0, \infty)$). The histogram of $v(x, y, t)$, or the empirical probability density, is defined as

$$p(\theta|t) = \frac{1}{|\Omega|} \lim_{\Delta\theta \to 0} \frac{|\cup_{\alpha \in (\theta, \theta+\Delta\theta)} \Gamma(\alpha|t)|}{\Delta\theta}, \tag{24}$$

where $|\cdot|$ denotes the Lebesgue measure in R^2. We then define the entropy of v at time t to be

$$S(t) = -\int_0^\infty p(\theta|t) \ln p(\theta|t) d\theta, \tag{25}$$

where the Boltzmann constant k_B in Gibbs' original definition has been set to 1 [24].

Our first remarkable statistical discovery indicates that the reaction-diffusion system appears to follow the *Second Law of Thermodynamics* (see Figure 4).

Statistical Discovery. The entropy $S(t)$ consists of two components $S_0(t) + S_a(t)$, where (see Figure 4):

1. the mean-field or *trend* component $S_0(t)$ is non-decreasing, and converges to some constant A as $t \to \infty$; and

2. the dynamically active component $S_a(t)$ is a rapidly fluctuating time series so that the signal-to-noise ratio $\lim_{t\to\infty} S_0(t)/\sigma(S_a(t)) \gg 1$, where $\sigma(\cdot)$ denotes the standard deviation.

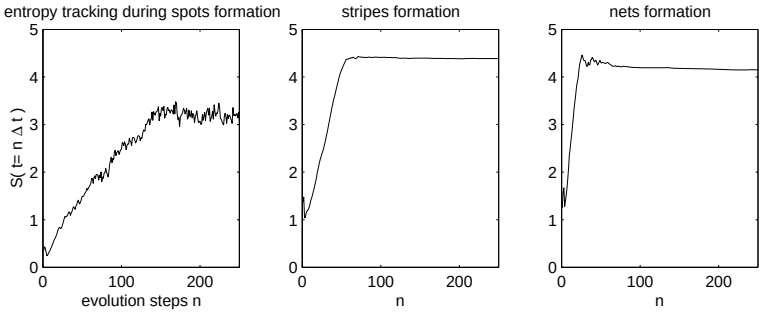

Figure 4: The entropy sequences during the formation of typical patterns

In Figure 4, the entropy curve for the spots pattern appears to be bumpier than the other two. This reflects the well known fact that the spots pattern often evolves more actively in computer simulations, with constant annihilation and creation.

While Figure 4 looks clear to human vision that the above two statements make perfect sense, it is yet not so trivial to be able to actually extract the two components $S_0(t)$ and $S_a(t)$ from $S(t)$. One has to apply tools from signal estimation theory such as Bayesian inference.

Employed here is the deterministic version of Bayesian estimation [11, 12, 39] – Tikhonov regularization [11, 17, 66]. Due to the monotonicity consideration on S_0, we apply Rudin-Osher-Fatemi's total variation (TV) restoration model [54, 4, 6] to extract the two components. Define an error functional for any estimator $g(t)$ by

$$e[g(t)\big|S(t)] = \int_0^T |g'(t)|_\varepsilon dt + \frac{\lambda}{2}\int_0^T (S(t) - g(t))^2 dt, \tag{26}$$

where T is the total running time (e.g., $T = 250\Delta t$ for Figure 4), and $|a|_\varepsilon = \sqrt{a^2 + \varepsilon^2}$ the regularized absolute value for some fixed small constant. The least square fitting term implies that Gaussian type fluctuations are assumed. As in Bayesian estimation, the parameter λ is inversely proportional to the fluctuation variance, and is either tunable or estimated. We then define the mean-field (or trend) entropy S_0 and the dynamically active (or fluctuating) entropy S_a by

$$S_0(t) = \text{argmin}_{g(t)} e[g(t)\big|S(t)] \quad \text{and} \quad S_a(t) = S(t) - S_0(t).$$

It is well known [4, 17, 36] that the decomposition does exist and is unique.

The extraction results for the examples in Figure 4 are ilustrated on Figure 5. Notice that the dynamically active component have been plotted in the dB unit: $10\log_{10}|S_a(t)|$, as customary in signal analysis and processing [45]. The dynamic difference is manifest in the S_a plotting between the active spots pattern and relatively quieter stripes and nets patterns (assuming $\Delta t = 1$ in Figure 4, we have taken the TV regularizer $\varepsilon = 0.01$ and $\lambda = 1$).

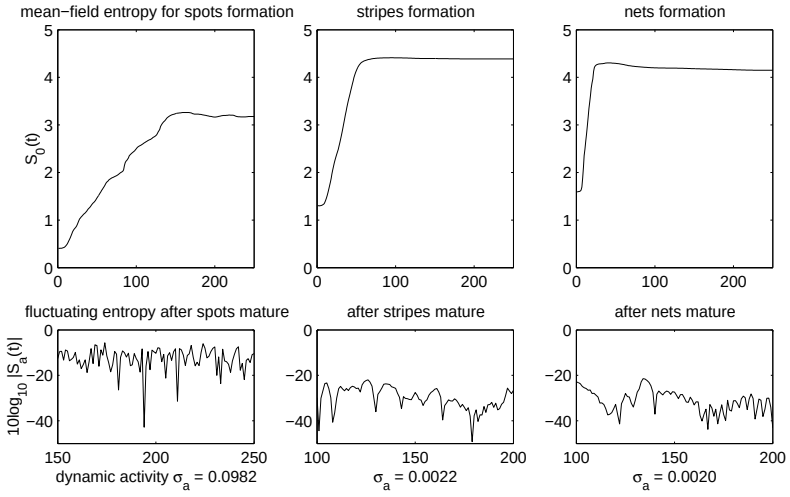

Figure 5: Extraction of the mean-field entropy $S_0(t)$ and the dynamically active entropy $S_a(t)$ by the model in equation (26). Notice that S_a has been plotted in the dB unit: $10\log_{10}|S_a(t)|$, as customary in signal processing. It clearly shows that the spots pattern is more dynamically active

3.2. Pattern Maturity $r(t)$ and Dynamic Activity σ_a

The main goal of this section is to define two important quantities based on the entropies: maturity $r(t)$ and dynamic activity σ_a.

Assume the fluctuation component $S_a(t)$ has a zero mean $E[S_a(t)] \equiv 0$. Then the trend $S_0(t) = E[S(t)]$ is precisely the mean field of the entropy time series. In particular, by approximating the expectation operator by empirical averaging, ones obtains

$$A = \lim_{t\to\infty} S_0(t) = \lim_{t\to\infty} \frac{1}{h} \int_{t-h}^{t} S(t)dt, \tag{27}$$

where h could be any prefixed window length. We shall call A the *saturated* entropy of the pattern formation. Let β be a positive weight constant. The *ripeness index* or *maturity* is then defined by

$$r(t) = r(v(t)) = e^{-\beta(A-S_0(t))^2}. \tag{28}$$

Notice that $r(t) \in (0, 1]$, and $\lim_{t\to\infty} r(t) = 1$. Thus the closer $r(t)$ gets to 1, the more matured the pattern appears to be.

Note. One could also try Gibbs' form $e^{-\beta(A-S_0(t))}$ since the mean entropy $S_0(t)$ has been assumed non-decreasing. The quadratic Gaussian form is adopted here since numerical extraction may never get the exact monotonicity. Also notice that the weighting constant β corresponds to the reciprocal of the temperature in Gibbs' form and of the variance in the Gaussian form.

The ripeness index is a valuable measure for deciding a good stopping time T_s in computer simulation after the pattern has become "ripe" enough. For example, one could preset a maturity level $r_0 \in (0, 1)$ close to 1, and define the stopping time by

$$T_s = T_s(r_0) = \sup\{t > 0 \mid \text{for any } s < t,\ r(s) < r_0\ \}. \tag{29}$$

Then arises a controversial issue regarding the actual computation of the ripeness index $r(t)$. On one hand, its definition (28) crucially depends on the saturated entropy A, which is obtained from $S_0(t)$ by letting $t \to \infty$. On the other hand, the major goal of introducing $r(t)$ is to offer an effective strategy for automatic termination way ahead of the infinity.

The issue could be resolved by two approaches. First, if one is experimenting a system with fixed reaction-diffusion parameters, but with different initial conditions for instance, the saturated entropy A could be learned and well estimated from a single full run of a particular but generic set of initial conditions.

Alternatively, one could modify the definitions in (28) and (29) so that they become progressive or *causal*, independent of any "future" development. This can be done as follows. Suppose the targeted maturity level is $r_0 = 1 - \varepsilon^2$ with $\varepsilon \ll 1$. Then (29) implies that for any $t_1, t_2 : T_s < t_1 < t_2$ (under the assumption that the mean entropy is non-decreasing),

$$r(t_i) \geq r_0 = 1 - \varepsilon^2, \qquad i = 1, 2.$$

On the other hand, by (28), to the first order,

$$r(t_i) = e^{-\beta(A - S_0(t_i))^2} \simeq 1 - \beta(A - S_0(t_i))^2.$$

Thus, again to the first order,

$$\beta(A - S_0(t_i))^2 \leq \varepsilon^2, \qquad i = 1, 2,$$

which leads to

$$|S_0(t_2) - S_0(t_1)| \leq b\varepsilon, \qquad b = \frac{2}{\sqrt{\beta}}. \tag{30}$$

This inspires a causal definition of the stopping time: with a being any fixed time duration,

$$T_s = T_s(\varepsilon) = \sup\{t > a \ : \ |S_0(t) - S_0(t-a)| > \varepsilon \ \}. \tag{31}$$

Compared with the original definition, it is *causal* and independent of future computation. In real simulation, one may add a further delay of some fixed length h so that numerical error could be diminished and the pattern can get securely matured up to the precision ε. That is, one takes $T_s(\varepsilon) + h$ as the actual stopping time.

The fluctuation entropy $S_a(t)$, though behaving like noise, does reflect important dynamic properties of matured patterns. Weak fluctuations often correspond to static *standing waves*, while stronger ones are associated with *travelling waves* like spontaneous spots. We thus define the *dynamic activity* for any *matured pattern* to be

$$\sigma_a = \lim_{t \to \infty} E[S_a(t)^2]^{\frac{1}{2}}. \tag{32}$$

As before, S_a has been assumed to have zero mean. In application, the expectation operator is replaced by empirical averaging, and the limit is approximated at any moment $t > T_s + h$,

$$\sigma_a^2 \simeq \frac{1}{h} \int_{t-h}^{t} S_a^2(\tau) \, d\tau, \qquad \text{for some time interval } \ h.$$

For the typical patterns in Figure 5 for instance, $\sigma_a = 0.0982,\ 0.0022,\ 0.0020$ separately for the matured spots, stripes, and nets patterns. The difference is as big as nearly 50 times.

3.3. Patterns of Histograms via Invariant Moments

So far, we have only defined key quantities for monitoring the maturity and activity of *developing* patterns. In the coming two sections, we discuss how to differentiate and classify the *matured* patterns.

Let us first briefly review some basic elements of pattern classification [19]. Let X denote a set of target objects to be classified. A feature map of dimension n associates to each object $x \in X$ a feature vector in R^n:

$$x \to \boldsymbol{F}(x) = (f_1(x), \cdots, f_n(x)).$$

For any subset $E \subset R^n$, the collection of objects $\boldsymbol{F}^{-1}(E) \subset X$ is called an E-pattern class. A finite and exclusive patten classification based on the feature map $\boldsymbol{F}$ is a finite partition of R^n:

$$R^n = E_1 \cup E_2 \cup \cdots \cup E_N, \qquad \text{for some} \quad N,$$

and its induced partition of X into different pattern classes:

$$X = X_1 \cup \cdots \cup X_N, \qquad \text{with} \;\; X_k = \boldsymbol{F}^{-1}(E_k), \;\; k = 1 : N.$$

The main challenges of pattern classification reside in [70, 71, 50]:

1. how to establish a suitable feature map $\boldsymbol{F}$ which is neither too coarse (with substantial information loss) nor over-redundant; and

2. how to properly partition the feature space based on either supervised or unsupervised learning.

In the current application of reaction-diffusion patterns, X is the collection of all matured patterns, and the main challenge is to properly define features and patterns. On the other hand, pattern classification could conveniently benefit from *supervised* learning, as supervised by human vision and popular classification name tags such as *spots* and *stripes*.

Our first characterizing scheme is directly based on the histogram defined in Section 3.1:

$$p_v(\theta) = \lim_{\Delta\theta \to 0} \frac{|\{(x,y) \;:\; \theta \le v(x,y,t) \le \theta + \Delta\theta\}|}{\Delta\theta |\Omega|}, \tag{33}$$

where it has been assumed that $t \ge T_s$ (after the pattern matures) so that $p_v(\theta)$ is effectively independent of t (i.e., the stationary assumption).

Let M denote the collection of all continuous probability measures supported on $(0, \infty)$. Then the histogram could be considered as a feature map:

$$p : X \to M : \quad v \to p_v,$$

which maps any matured pattern v to its histogram. From 2-D to 1-D, it already achieves a significant amount of data compression by being blind to the spatial organization of each individual activation level $\Gamma(\theta)$ (see (23), with t dropped as well for matured patterns).

The histograms of three typical matured patterns corresponding to spots, stripes, and hexagonal type nets are plotted in Figure 6. It is patent to human vision that the histograms have clearly distinguished the three. The stripes histogram is bimodal and almost symmetric, while those of the spots and nets are both monomodal but tilted toward opposite ends.

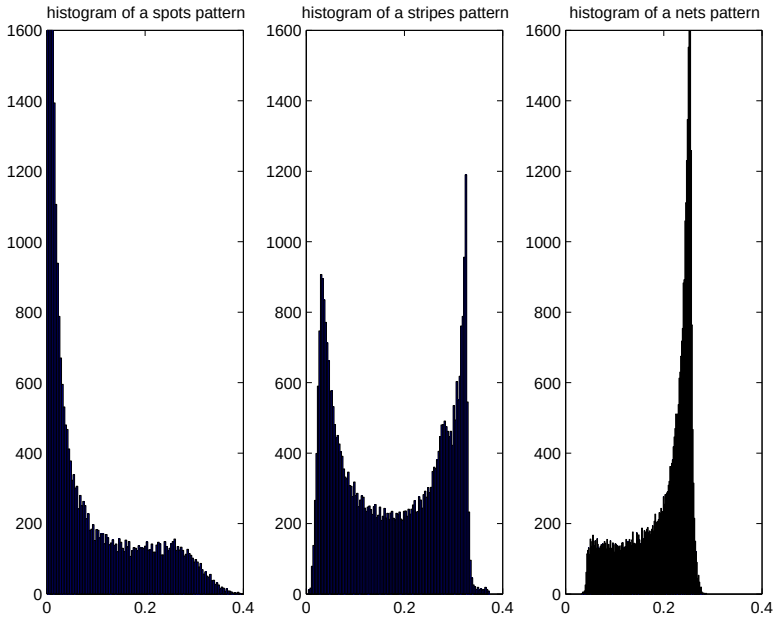

Figure 6: The histograms of three typical matured patterns

This inspires further data compression and more efficient feature representation. To characterize the features of a probability measure (such as symmetry, tilting, and resemblance to Gaussians, etc.), one could project them onto different templates such as canonical monomials or orthogonal polynomials. Here we work with two high order *invariant moments* – *skewness* and *kurtosis*, which have been frequently applied in language modelling and image analysis [40].

The skewness of a given continuous probability distribution $p(\theta)$ (supported

on $(0,\infty)$) is the 3rd order normalized moment:

$$s[p]=\int_0^\infty\left(\frac{\theta-m}{\sigma}\right)^3 p(\theta)\,d\theta,$$

where m and σ stand for the mean and standard deviation. Similarly, the kurtosis is the 4th order normalized moment:

$$k[p]=\int_0^\infty\left(\frac{\theta-m}{\sigma}\right)^4 p(\theta)\,d\theta.$$

For the Gaussian class $N(m,\sigma)$ for instance, the skewness always vanishes while the kurtosis is 3. Both the skewness and the kurtosis are invariant under linear transforms of the associated random variables, which makes them ideal and robust for pattern analysis (see for example Mumford's ICM 2002 paper [40] on the role of kurtosis in language and natural pattern analysis. For more discussion on invariant moments such as cumulants, see for example the recent work by Rota and Shen [52]).

For a given matured pattern v, we are able to define two invariant moments $s_v=s[p_v]$ and $k_v=k[p_v]$. Thus a 2-D feature map is eventually established for the collection of matured patterns:

$$\begin{aligned}\text{matured R.-D. pattern } v \Rightarrow \quad & \text{histogram } p_v\\ \Rightarrow \quad & \text{invariant moments } \boldsymbol{F}_v=(s_v,k_v).\end{aligned}\tag{34}$$

This is certainly a highly *lossy* information compression scheme going from 2-D functions to 2-component vectors. It is however highly efficient as shown in Figure 7 – the skewness and kurtosis indeed well differentiate the common patterns.

As an example of supervised learning (by human vision), one could now easily partition the skewness-kurtosis plane to match the supervised classification of the most common patterns of spots, stripes, and nets. For instance, for the current model, one could even try simple linear decision rules solely based on the skewness, such as:

$$\begin{aligned}E_{\text{sp}}&=\{(s,k)\mid s>0.8\},\\ E_{\text{st}}&=\{(s,k)\mid -0.8<s\le 0.8\},\\ E_{\text{net}}&=\{(s,k)\mid s\le -0.8\}.\end{aligned}$$

Generally, by assuming properly the distribution type (in the skewness-kurtosis half-plane) for each pattern, one could obtain the decision boundaries from

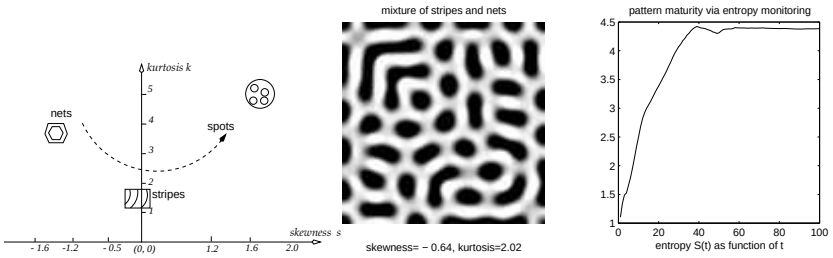

Figure 7: Typical matured patterns in the skewness-kurtosis feature plane. Some representative values of (s_v, k_v) from simulations: spots (1.80, 5.19) and (1.69, 4.70); stripes (-0,03, 1.47) and (-0.24, 1.51); and nets (-1.33, 3.74) and (-1.19, 3.26). For a matured mixture of stripes and nets as shown below, $(s_v, k_v) = (-0.64, 2.02)$, which lies between the regions of stripes and nets. The plane also hints a natural direction from nets to stripes, and eventually to spots, which is nicely confirmed by the comprehensive simulation results of Roy Williams [68]

sufficiently many simulation data, and pattern analysis methods such as cross-validation and maximum likelihood estimation (MLE) [19]. For example, in the case of identical isotropic Gaussians and MLE, pattern classification is tightly connected to Voronov partitioning of the skewness-kurtosis half-plane.

On the other hand, one should also question whether it makes perfect sense to partition all matured patterns of a given reaction-diffusion system into three or several specific categories. This is because some matured patterns are really mixtures, like the one displayed at the bottom of Figure 7. In such situations, even human vision would find it difficult to make a binary decision. Therefore, the decision boundaries must be understood as transitional gray areas instead of clear-cut binary ones. In this sense, the (s, k) half-plane itself is already a good representation.

3.4. Statistical Pattern Analysis via Geometric Features

Histograms of activation levels are ignorant of geometric organization. In this section, geometric information is to be incorporated into the statistical schemes of pattern analysis.

In modern image and vision analysis, linear filters have become a common tool for representing and analyzing various spatial features such as orientations, multiple scales or resolutions, and local spatial wave numbers. Some well known examples include the Gabor filters and many family of linear filters in wavelets

design [16, 60, 61, 62, 33, 70, 71].

Compared with general image patterns such as textures and clusters of natural scenes, reaction-diffusion patterns are relatively *better* organized due to the nature of the system (i.e., controlled by a system of partial differential equations). Thus geometric regularity seems to be more important and pertinent than other common image features. This observation motivates the application of nonlinear geometric filters such as *curvatures*. The curvature feature has been actively employed in modern image and vision analysis, such as in *mean curvature motions* [21], and image denoising, deblurring, inpainting, and vision modeling [8, 7, 11, 9, 53, 58].

For a smooth image $v = v(x, y)$, the *naturally signed* curvature is defined to be

$$\kappa_v = \kappa_v(x, y) = \nabla \cdot \left[\frac{\nabla v}{|\nabla v|} \right] = \mathrm{div}(\boldsymbol{n}), \qquad \boldsymbol{n} = \frac{\nabla v}{|\nabla v|}, \tag{35}$$

which is well defined at all regular points with nonzero gradients. Notice that κ_v is *morphologically invariant*: $\kappa_{g(v)} = \kappa_v$ for any monotonic function g, which should be the case since curvature is geometric and only belongs to level sets. To avoid the singularity on homogeneous regions where the gradients almost vanish, κ_v is often regularized to [53, 9]

$$\kappa_v = \kappa_v(x, y) = \nabla \cdot \left[\frac{\nabla v}{|\nabla v|_\varepsilon} \right], \quad \text{with } |a|_\varepsilon = \sqrt{a^2 + \varepsilon^2}, \tag{36}$$

for some small regularization constant ε. This regularized curvature is only approximately morphologically invariant but well defined everywhere.

On the other hand, the curvature information on homogeneous regions is less significiant than where activation levels change rapidly, e.g., along the "edges" of spots or stripes. Therefore, instead of the *pure* curvature, we consider its weighted version:

$$\psi_v(x, y) = |\nabla v| \kappa_v = |\nabla v| \, \nabla \cdot \left[\frac{\nabla v}{|\nabla v|} \right] = \Delta v - D^2 v(\boldsymbol{n}, \boldsymbol{n}), \tag{37}$$

where Δv denotes the Laplacian, and $D^2 v$ the Hessian bilinear form. The ε-regularization technique (36) applies to ψ_v as well. Note that the weighted curvature ψ_v is no longer morphologically invariant due to the contrast weighting $|\nabla v|$.

To establish robust comparison schemes among different patterns, we shall work with the statistically normalized matured patterns, as in the definitions of skewness and kurtosis. Define the empirical mean and variance of v to be

$$m = \frac{1}{|\Omega|} \int_\Omega v(x, y) dx dy = \int_0^\infty \theta p_v(\theta) d\theta,$$

$$\sigma^2 = \frac{1}{|\Omega|}\int_\Omega (v-m)^2 dxdy = \int_0^\infty (\theta - m)^2 p_v(\theta) d\theta.$$

Then we investigate the statistics of the normalized weighted curvature:

$$\phi_v = \psi_{\frac{v-m}{\sigma}} = \frac{1}{\sigma}\psi_v = \frac{1}{\sigma}(\Delta v - D^2 v(\boldsymbol{n}, \boldsymbol{n})). \tag{38}$$

The histogram of ϕ_v shall be denoted by $q_v(\phi)$, and Figure 8 displays the histograms for some typical matured patterns. It is quite clear to human vision that these geometry based histograms successfully differentiate different patterns.

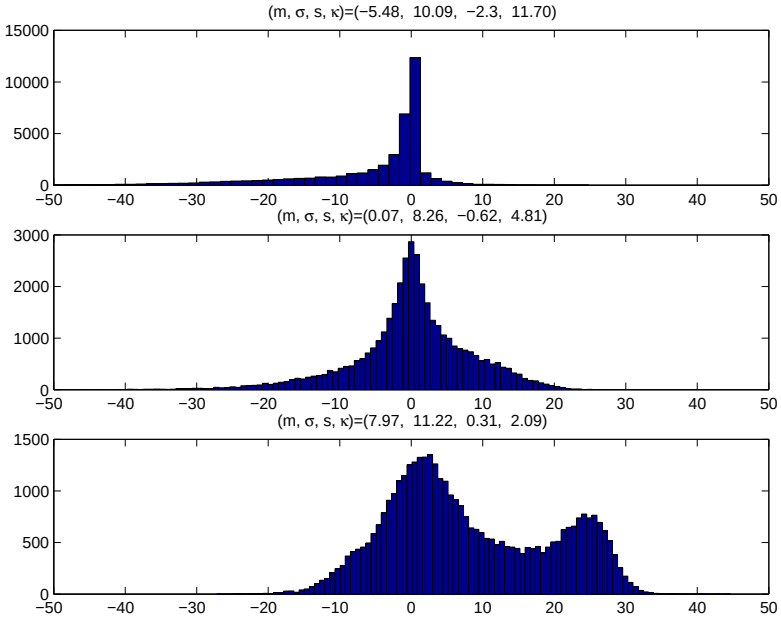

Figure 8: Histograms of the normalized weighted curvatures ϕ_v's (see equation (38)) for typical patterns of spots, stripes, and nets (for a good axis scale, the histograms are actually for $h \times \phi_v$ where $h = 0.01$ is the pixel size). The title lines give the corresponding information on the mean, standard deviation, skewness, and kurtosis

Unlike the ones in the proceeding section, geometry based histograms cannot be translated along its horizontal axis, since zero curvature has its intrinsic meaning. The peaks at zero in the weighted curvature histograms in Figure 8 for example, reflect the fact that most sites in a matured reaction-diffusion

pattern are almost homogeneous. In particular, the mean itself is a valuable statistical feature containing crucial pattern information.

Thus we define a feature map based on the first four moments

$$v \Rightarrow \boldsymbol{F}_g = \boldsymbol{F}_g(q_v(\phi)) = (m_g, \sigma_g, s_g, k_g),$$

i.e., the mean, standard deviation, skewness, and kurtosis, where the symbol g stands for geometry. From the empirical data shown in Figure 8, it is clear that the mean and kurtosis are outstandingly effective in differentiating the patterns. Thus one could further simplify the above four-component feature map to a two-component one:

$$v \Rightarrow \boldsymbol{F}_g = (m_g, k_g).$$

Figure 9 shows the typical matured patterns on the (m_g, k_g) half-plane. This new feature plane also clearly resolves different patterns. The discussion on pattern quantization based on supervised learning is similar to the proceeding section.

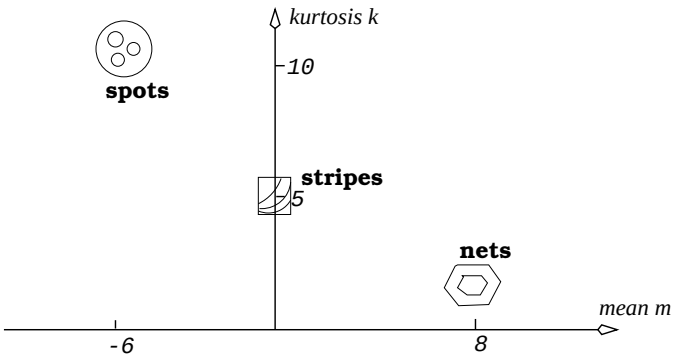

Figure 9: Typical matured patterns in the geometric mean-kurtosis (m_g, k_g) half-plane

3.5. Universality of Statistical Features via Qualitative Models

In this section, we explain why the statistical features observed for some typical patterns from the Gray-Scott model are universal for general reaction-diffusion patterns.

3.5.1. Histogram Features of Spots, Stripes and Nets

A matured *spots* pattern typically consists of isolated and almost circular islands in a darker sea background. A *nets* pattern, on the other hand, could be qualitatively treated as the reversal of the regular spots pattern, in the sense that it consists of dark holes in a brighter foreground. This explains why the histograms of the two are approximately the reversals of each other (see Figure 6).

Thus it suffices to model the histogram features of the spots and stripes patterns only.

Consider a normalized disk model for a single spot in R^2:

$$v(x,y) = g(r^2) = g(x^2+y^2),$$

where the rotational generator $g(x)$ satisfies

$$g(x) \in C^1[0,\infty), \quad g(0)=1,\ g(+\infty)=0, \quad \text{and } g \text{ decreases monotonically} \tag{39}$$

(notice that due to the squared argument r^2, $\nabla v(0,0) = g'(0)(0,0) = (0,0)$, implying a smooth spot center). For example, $g(r^2) = e^{-\frac{r^2}{2\sigma^2}}$ is a Gaussian spot, and

$$g(r^2) = \begin{cases} (1-r^2)^2, & r \le 1\,, \\ 0, & r > 1 \end{cases} \tag{40}$$

is a compactly supported spot. For the convenience of superimposing many isolated spots, we shall now work with a typical spot model (39) whose compact support is the unit disk, which implies that $g'(1)=0$. In addition, assume the domain Ω is exactly the unit disk as well. Let $x = f(\theta)$ denote the inverse function of $\theta = g(x)$ with $x \in [0,1]$ and $\theta \in [0,1]$, which is well defined since g is monotonically decreasing. Then by the definition of the empirical histogram,

$$\begin{aligned} p_v(\theta) &= \lim_{\Delta\theta\to 0} \frac{|\{(x,y) : \theta \le g(r^2) < \theta + \Delta\theta\}|}{|\Omega|\ \Delta\theta} \\ &= \lim_{\Delta\theta\to 0} \frac{|\{(x,y) : f(\theta+\Delta\theta) \le r^2 < f(\theta)\}|}{|\Omega|\ \Delta\theta} \\ &= \frac{1}{|\Omega|} \lim_{\Delta\theta\to 0} \frac{\pi(f(\theta) - f(\theta+\Delta\theta))}{\Delta\theta} = -\frac{\pi}{|\Omega|} f'(\theta) = -\frac{\pi}{|\Omega|}\frac{dr^2}{d\theta}. \end{aligned}$$

As a result,

$$p_v(0^+) \propto -f'(0^+) = -1/g'(1^-) = +\infty,$$

which well explains the high peak at the left end of the histogram in a matured spots pattern (see Figure 6). Figure 10 displays the predicted histogram for the spot model (40).

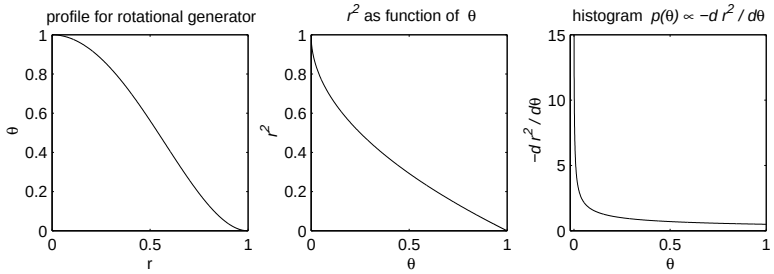

Figure 10: The histogram of a model spot: $v(x,y) = g(r^2) = g(x^2+y^2)$, with $g(s) = (1-s)^2$ for $s \in [0,1]$, and 0 otherwise. This model histogram qualitatively matches that of the spots pattern in Figure 6

In a realistic pattern when N spots emerge almost independently against a uniform darker background, the histogram is simply

$$p_v(\theta) = \frac{p_1(\theta)|\Omega_1| + \cdots + p_N(\theta)|\Omega_N|}{|\Omega|}, \tag{41}$$

where $\Omega_1 \cup \cdots \cup \Omega_N$ partitions Ω into effective influence regions associated to the N spots. When the profiles of the spots are very similar, we have $p_v(\theta) \approx p_1(\theta)$. In general situations where the sizes and profiles of the spots do vary slightly, formula (41) could be thought of a smoothing (or filtering) process of the histograms of individual spots.

We now explain the histogram features for the stripes pattern. Consider a *canonical* ideal stripe along the y-direction on the square $Q = [0,1] \times [0,1]$:

$$v(x,y) = g(x) \;\text{ with }\; \begin{cases} g(0) = 0, \quad g(1) = 1; \quad \text{and} \\ g'(0) = 0, \quad g'(1) = 0, \quad \text{and} \quad g' > 0. \end{cases}$$

Furthermore, assume that g is symmetric: $g(x) = g(1-x)$ about the stripe "edge" $x \equiv 1/2$. For example, one could take

$$g(x) = \frac{1-\cos(\pi x)}{2} = \sin^2 \frac{\pi}{2} x.$$

Let $x = f(\theta)$ denote the inverse function of $\theta = g(x)$. Then by definition,

$$p_v(\theta) = f'(\theta) = \frac{1}{g'(x)},$$

which well explains the blowups at the two ends of the histogram of a generic stripes pattern (see Figure 6), since

$$p_v(0^+) = \frac{1}{g'(0^+)} = +\infty \quad \text{and} \quad p_v(1^-) = \frac{1}{g'(1^-)} = +\infty.$$

As an example, consider $\theta = g(x) = \frac{1-\cos(\pi x)}{2}$. Since $g'(x) = \frac{\pi}{2}\sin(\pi x)$, one has

$$p_v(\theta) = \frac{2}{\pi}\frac{1}{\sqrt{1-(2\theta-1)^2}} = \frac{1}{\pi}\frac{1}{\sqrt{\theta(1-\theta)}}\,.$$

Thus for this ideal model, the empirical histogram is precisely the beta distribution $B(\frac{1}{2}, \frac{1}{2})$ on (0, 1), which faithfully models the bimodal pattern of the actual stripes histogram in Figure 6.

3.5.2. Feature Interpretation for the Geometric Histograms

We now qualitatively explain why the geometric histograms of different patterns are noticeably concentrated on different regions (see Figure 8).

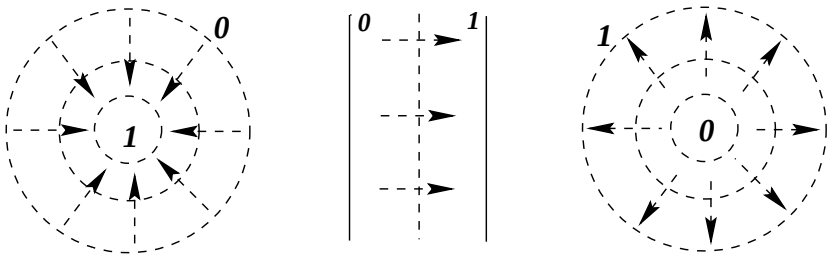

Figure 11: The normal vector fields near the "edges" of different patterns. The normal field is contractive for a spots pattern, expansive for a nets pattern, and almost constant for an ideal stripe pattern, leading to negative, positive, and almost zero divergences (i.e., curvatures)

Consider the normal fields $\boldsymbol{n} = \nabla v/|\nabla v|$ of the patterns of spots, stripes, and nets. Near the "edges" where $|\nabla v|$ is relatively large, the normal fields are typically depicted by Figure 11. For a spots pattern, the normal field $\boldsymbol{n}$ is contractive and points toward the peak at the center. Therefore, its divergence $\kappa_v = \mathrm{div}(\boldsymbol{n})$ (i.e., curvature) must be negative by the divergence theorem. This explains why in Figure 8, the geometric histogram for $\phi_v = |\nabla v|\kappa_v$ associated

to the spots pattern is mostly concentrated on the negative axis. The situation reverses for a nets pattern. For an ideal *local* stripe, the normal field is close to be constant which leads to almost zero curvatures.

On the other hand, for all the three patterns, the peaks at $\phi_v(= |\nabla v|\kappa_v) = 0$ are due to the fact that away from the transitionary edges, v is almost constant and $|\nabla v|$ close to zero.

4. Geometric Measures of Reaction-Diffusion Patterns

In this section, we employ tools from geometric measure theory to answer questions like: what is the average spot size in a spot pattern? how many are there on a square inch on average? and what is the aspect ratio of the stripes in a stripe pattern? Many of these tools have been recently revived in mathematical image and vision analysis.

4.1. Total Variation and Curvatures

The broad applications of the total variation (TV) measure (e.g., [9, 10, 7, 8, 12, 59]) in image analysis processing started from Rudin, Osher, and Fatemi's celebrated work on TV based image denoising and deblurring [53, 54]. From the applied harmonic analysis point of view, total variation and its sibling Besov norms are also naturally connected to wavelets based image denoising and regularization mechanisms [11, 14, 18, 36].

The total variation TV[u] or $\int_\Omega |Du|$ of an integrable function $v \in L^1(\Omega)$ generalizes the regular Sobolev norm $\int_\Omega |\nabla u| dxdy$, and is defined in the distributional sense [26]:

$$\mathrm{TV}[u] = \int_\Omega |Du| = \sup_{\boldsymbol{g} \in C_0^1(\Omega, B^2)} \int_\Omega u(\nabla \cdot \boldsymbol{g}) dxdy, \tag{42}$$

where $C_0^1(\Omega, B^2)$ denotes all compactly supported C^1 vectorial functions that take values inside the unit disk in R^2:

$$B^2 = \{\boldsymbol{g} = (g_1, g_2) \ : \ |\boldsymbol{g}| = \sqrt{g_1^2 + g_2^2} \leq 1\}.$$

Applying (42) to any open subset of Ω, one could define rigorously the set function $\int_\bullet |Du|$ as a *finite* Radon measure on Ω, provided that TV[u] $< \infty$.

Unlike the Sobolev norm in $W^{1,1}(\Omega)$, the total variation of the indicator function χ_E of a Borel set $E \subset \Omega$ is well defined (though can be ∞), and is finite

for any Lipschitz domain with finite 1-D Hausdorff measure of its boundary. In the latter case they are in fact identical. Thus total variation leads to a more general definition of the perimeter of a Borel set $E \subset \Omega$ by:

$$\mathrm{Per}(E|\Omega) = \int_\Omega |D\chi_E|.$$

Any function $u \in L^1(\Omega)$ with finite total variation is said to have bounded variation (BV). The space of all such functions is denoted by $\mathrm{BV}(\Omega)$, which is a Banach space under the BV-norm:

$$\|u\|_{\mathrm{BV}} = \|u\|_{L^1} + \int_\Omega |Du|.$$

It is the space of $\mathrm{BV}(\Omega)$ that has been recently widely applied as an approximate but efficient image model for general images [36, 9, 12].

The geometric nature is reflected in the celebrated co-area formula of Fleming and Rishel [22] and De Giorgi [25]:

$$\mathrm{TV}[u] = \int_\Omega |Du| = \int_R \mathrm{length}(u \equiv \lambda) d\lambda, \tag{43}$$

where $u \equiv \lambda$ denotes the λ-level set, assuming that u is regular. For general BV functions whose individual level sets are less meaningful, the co-area formula is given by

$$\mathrm{TV}[u] = \int_\Omega |Du| = \int_R \mathrm{Per}(u < \lambda|\Omega) d\lambda,$$

where $u < \lambda$ represents the Borel set $E_\lambda = \{(x, y) \in \Omega : \ u(x, y) < \lambda\}$. Thus total variation provides a natural way to collectively sum up the lengthes of *all* the level sets, which has made it so powerful in image processing and vision analysis.

Although general smoothing operators like heat diffusion are not continuous in $\mathrm{BV}(\Omega)$, they are often continuous in terms of the value of the TV measure. That is, suppose $u_\varepsilon \in W^{1,1}(\Omega)$ or $C^\infty(\Omega)$ is a family of mollifications of $u \in \mathrm{BV}(\Omega)$ (often spatially adaptive when Ω is bounded), so that $u_\varepsilon \to u$ in L^1 as $\varepsilon \to 0$. Then,

$$\begin{aligned} &\mathrm{TV}[u_\varepsilon - u] \to 0 \ \text{ is generally untrue,} \\ &\qquad\qquad \text{while } \ \mathrm{TV}[u_\varepsilon] \to \mathrm{TV}[u] \ \text{ could be well achieved.} \end{aligned} \tag{44}$$

If the sequence is not from a mollification design, then generally only the *lower semi-continuity* can be guaranteed [26]. This observation is very useful for computing the total variation norm using ordinary schemes such as finite differences (also see later sections).

For BV functions with smooth level sets, such as the indicators of smooth domains, one could further investigate high order geometric measures. The m-th order total curvature is defined by

$$K_m[u] = \int_\Omega \kappa^m |Du| = \int_R \left(\int_{u\equiv\lambda} \kappa^m ds \right) d\lambda, \tag{45}$$

where ds denotes the arc length element along the level lines, and $\kappa = \nabla \cdot (\nabla u/|\nabla u|)$. For instance, Euler's elastica measure $E[u] = a\mathrm{TV}[u] + bK_2[u]$ has numerous applications in image analysis ever since Mumford first introduced it into computer vision [38, 44, 35, 57, 5, 20, 55]. Theoretical difficulties on analyzing high order geometric measures still remain open (see for example, the remarkable work by the De Giorgi school Bellettini, Dal Maso, and Paolini [2]).

4.2. Binary Segmentation of Reaction-Diffusion Patterns

When one asks questions like "what is the average radius of the spots or the average width of the stripes," it has already been assumed that the spots and stripes are isolated "objects" with clear-cut boundaries. The human vision system is able to carry out this so called *segmentation* task unconsciously but effortlessly even for complex images. Its mathematical modelling and algorithm design, however, have turned out to be highly nontrivial and fundamentally important in a number of fields. For example, the celebrated Mumford-Shah segmentation model and its variations have been proven very powerful in segmenting complicated image scenes with low textures [42, 13]. Another well known method is David Marr's zero-crossing theory [34].

Reaction-diffusion patterns are better behaved compared with general natural images. Thus in the current work, we adopt a simpler segmentation scheme which is directly based on binary thresholding. Let $v_2(x, y)$ be the binary image perceived by human vision. Then we model v_2 by

$$v_2(x, y) = \begin{cases} 1, & v \geq \theta_c, \\ 0, & v < \theta_c, \end{cases}$$

where θ_c is the critical activation level for binary thresholding. An effective threshold θ_c has to depend on the image $v(x, y)$ itself: $\theta_c = \theta_c[v]$.

In what follows, we take θ_c to be the mean value:

$$\theta_c = \theta_c[v] = \frac{1}{|\Omega|}\int_\Omega v dx dy = \int_0^\infty \theta p_v(\theta) d\theta.$$

From the statistics in the proceeding section, the histograms for typical matured stripe patterns are almost symmetric, and the two spikes correspond to the dark valleys and bright peaks in $v(x, y)$ (Figure 6). Thus the mean provides a natural cut-off for the boundaries. For a typical spots pattern v, the mean is tilted toward the dark background, and the associated binary segmentation often produces "fatter" bright spots than human perception. The situation reverses for a typical nets pattern. However, due to the well known reaction-diffusion mechanism of short-range activation and long-range inhibition, the transition is often rapidly realized. As a result, the mean based binary segmentation cannot be too fat or slim. Figure 12 shows two examples of mean based binary segmentation of matured reaction-diffusion patterns. From now on, we shall only work with the binary images v_2.

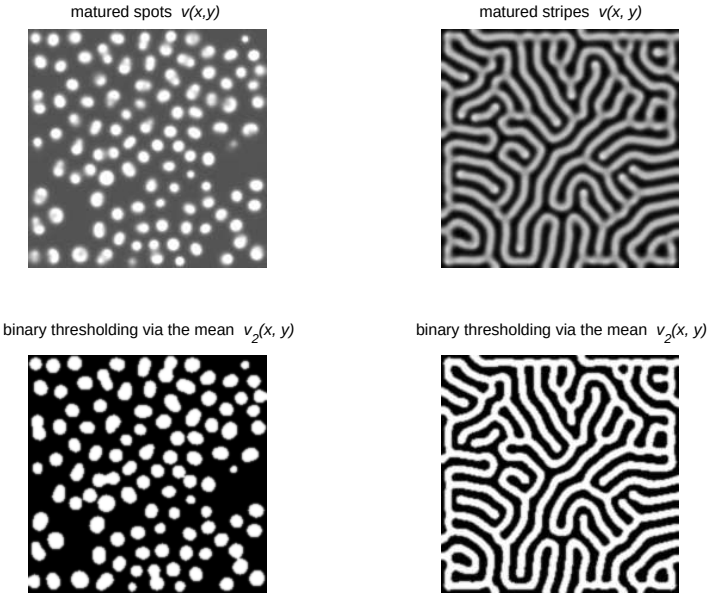

Figure 12: Two examples of mean based binary segmentation of matured reaction-diffusion patterns

4.3. Geometric Analysis of Spots and Nets

As clear from Figure 12, most spots are almost circular. Consequently consider the following approximate model of spots distribution. Suppose there are totally N non-overlapping circular white spots in a binary image v_2 with radii

$$r_1, r_2, \cdots, r_N, \text{ subject to certain unknown random distribution.} \tag{46}$$

Definition. (Average Radius) The average radius $\bar{r}$ of a binary spots pattern v_2 is defined to be:

$$\bar{r} = \langle r \rangle = \frac{r_1 + r_2 + \cdots + r_N}{N}. \tag{47}$$

According to the definition, it seems natural to first compute all the individual radii and then have them averaged. Such computation is generally very costly due to the individual segmentation, identification, and information extraction. The beauty of the geometric measure approach is that it carries out all the tasks at once.

Theorem 2. *Suppose the circular spots do not overlap in a binary image v_2. Then the average radius is given by* $\bar{r} = \dfrac{1}{2\pi N} \displaystyle\int_\Omega |Dv_2|$.

Proof. Let $D_k \subset \Omega$ denote the disk region corresponding to the k-th circular spot with radius r_k, $k = 1 : N$. Then we must have

$$v_2(x, y) = \sum_{k=1}^{N} \chi_{D_k}(x, y).$$

Since D_k's have been assumed non-overlapping,

$$\int_\Omega |Dv_2| = \sum_{k=1}^{N} \int_\Omega |D\chi_{D_k}| = \sum_{k=1}^{N} \mathrm{Per}(D_k) = 2\pi \sum_{k=1}^{N} r_k.$$

This leads to the conclusion. □

For simplicity it has been assumed that there are no incomplete spots along the boundary of the reaction-diffusion domain Ω, which is at least a very good approximation since the population of incomplete spots is a negligible fraction. In Figure 12 for example, the binary image v_2 contains only two incomplete spots along the boundary, out of more than 100 ones.

The major issue with Theorem 2 is that in reality N is unknown either. How to count the number without getting down to the individual level? This is again made possible by suitable geometric measures.

To start, suppose that $r_k = \bar{r} + \delta r_k, k = 1 : N$, with

$$\langle \delta r \rangle = 0 \quad \text{and} \quad \sigma^2 = \langle (\delta r)^2 \rangle \ll 1. \tag{48}$$

Then,

$$\frac{1}{\pi N} \int_\Omega v_2 dxdy = \frac{1}{\pi N} \sum_{k=1}^N \int_\Omega \chi_{D_k} dxdy = \frac{1}{N} \sum_{k=1}^N r_k^2,$$

which is precisely the second empirical moment $\langle r^2 \rangle$ for the random radii $(r_k)_{k=1}^N$. Since $\sigma^2 = \langle r^2 \rangle - \bar{r}^2$, we have

$$\frac{1}{\pi N} \int_\Omega v_2 dxdy = \bar{r}^2 + \sigma^2.$$

Combined with Theorem 2, it gives

$$\frac{[\pi N]^{-1} \int_\Omega v_2 dxdy}{[2\pi N]^{-1} \int_\Omega |Dv_2|} = \frac{\bar{r}^2 + \sigma^2}{\bar{r}} = \bar{r}(1 + \varepsilon^2), \tag{49}$$

where $\varepsilon = \sigma/\bar{r}$. Therefore, we have proven the following theorem.

Theorem 3. *Suppose the signal-to-noise ratio (SNR) $\varepsilon^{-1} = \bar{r}/\sigma \gg 1$, then the average radius is*

$$\bar{r} = \frac{2 \int_\Omega v_2 dxdy}{\int_\Omega |Dv_2|} (1 + O(\varepsilon^2)). \tag{50}$$

Compared with Theorem 2, Theorem 3 does not require the *a priori* information of N, as long as the SNR is high, or equivalently, the variation of the radii is small.

The combination of Theorem 2 and equation (49) also gives a formula for N.

Theorem 4. *Following the assumptions of Theorem 3, the total number N of spots is given by*

$$N = \frac{\left[\int_\Omega |Dv_2| \right]^2}{4\pi \int_\Omega v_2 dxdy} (1 + O(\varepsilon^2)). \tag{51}$$

Note that the computation of the TV measure can be benefited from the discussion in the equation line (44). In all the numerical examples in the current section, we employ the heat equation as the mollifier $u(x, y; t)$:

$$\begin{cases} u_t = \Delta u, & (x, y) \in \Omega; \text{ with} \\ \dfrac{\partial u}{\partial \boldsymbol{n}}\Big|_{\partial\Omega} = 0, & u(x, y, 0) = v_2(x, y), \end{cases}$$

and approximate TV measure $\int_\Omega |Dv_2|$ by the Sobolev norm

$$\int_\Omega |\nabla u(x, y, \tau)| dxdy$$

for some small time step τ (often several numerical steps). The computation of the latter is easily achieved by uniform finite differences.

As an example, we have applied the formulae in Theorem 3 and Theorem 4 to the binary spots pattern v_2 in Figure 12. Our numerical results show that

$$\bar{r} = \frac{A}{33} \quad \text{and} \quad N = 103,$$

where the average radius has been compared with the side length A of the square domain. Thus the average diameter of the spots is about 1/16 of A. On the other hand, our manual counting of all the spots gives $N = 107$, which has even taken into account automatic splitting (benefiting from human visual intelligence) of seemingly merged spots. Thus Theorem 3 and Theorem 4 do perform remarkably well.

4.4. Geometric Analysis of Stripe Patterns

In this section, theory and a number of interesting tools are developed to study some salient geometric features of stripe patterns.

4.4.1. Skeleton Curves: Definitions and Properties

Recall that a $C^{(m)}$ diffeomorphism $\phi : U \to V$ between two open sets in R^2 (or more generally, between two differential manifolds) is a $C^{(m)}$ map $\boldsymbol{q} = \phi(\boldsymbol{p})$ whose inverse $\boldsymbol{p} = \psi(\boldsymbol{q}) : V \to U$ exists and is also $C^{(m)}$.

Definition. (Skeleton Curves) A *compact* set $\Gamma \subset R^2$ is said to be a $C^{(m)}$ *skeleton curve* if for any $\boldsymbol{p} \in \Gamma$, there is an open neighborhood $U \subset R^2$, and

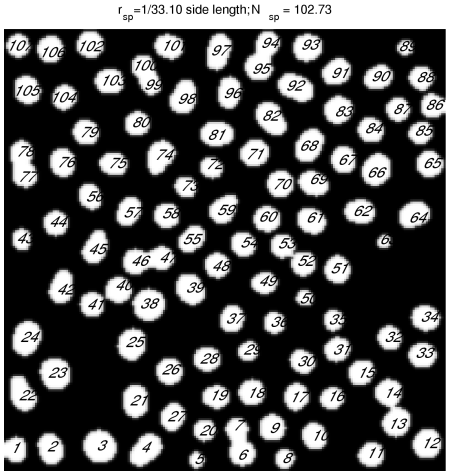

Figure 13: Estimating the average radius $\bar{r}$ and total number N of spots for a segmented binary spots pattern by Theorem 3 and Theorem 4. Our results are $\bar{r} = A/33$ and $N = 103$ where A denotes the side length of the domain. The time consuming manual counting as marked on top of the image gives $N = 107$, which has even taken into account automatic splitting of some "merged" pairs relying on human visual intelligence (e.g., pairs 6 and 7, 46 and 47, 77 and 78, 94 and 95, and 99 and 100)

a $C^{(m)}$ diffeomorphism $\phi : U \to R^2$, under which $\phi(\boldsymbol{p}) = (0, 0)$ and $\Gamma \cap U$ is mapped onto either the x-axis $R^1 \times 0$, the half axis $R^+ \times 0$, or the standard *convex* Y-junction:

$$Y_\theta = R^- \times 0 \bigcup 0 \times R^- \bigcup R^+(\cos\theta, \sin\theta), \quad \text{with } 0 \le \theta < \pi/2. \tag{52}$$

Here $R^+ = [0, \infty)$ and $R^- = (-\infty, 0]$ (see Figure 14).

Explanation is needed for the notion of *convex* Y-junctions. First, the word convexity refers to that the convex hull of Y_θ for all $0 < \theta < \pi/2$ is the whole plane R^2. It is in this sense that any Y_θ with $\theta \in (\pi/2, 2\pi)$ is non-convex. Remarkably, visual inspection of any generic matured stripe pattern shows that all the Y-junctions seem to be convex (see Figure 12 for example). The convex assumption will be important for Theorem 8.

Secondly, the value $\theta = 0$ is permitted in the definition to include T-junctions, i.e., structures that are diffemorphic to $R \times R^-$. T-junctions are crucial in human and machine vision [44, 29].

Finally, the notion of convexity is invariant under diffeomorphism since

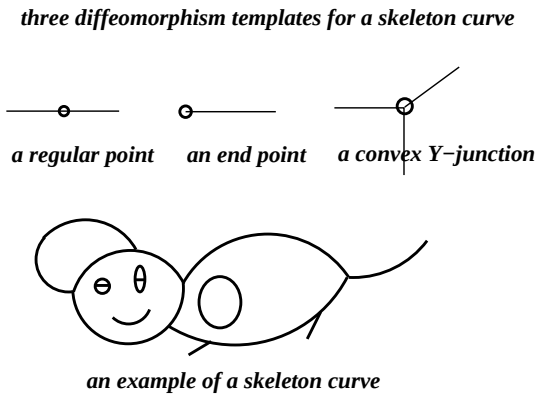

Figure 14: Skeleton curves: the three diffeomorphism templates and an example

locally a diffeomorphism coincides with its non-degenerate Jacobian, and convexity as a linear concept is invariant under invertible linear transforms.

For convenience, we shall denote by Γ° all the *regular* points on a skeleton curve Γ which are locally diffeomorphic to $R \times 0$, and Γ^* all the *singular* points on Γ which are either diffeomorphic to $R^+ \times 0$ or a convex Y-junction Y_θ.

The diffeomorphisms involved are all assumed at least $C^{(3)}$, so that curvatures can be properly defined and easily controlled.

Theorem 5. *Let Γ be a $C^{(m)}$ compact skeleton curve. Then (a) the singular set Γ^* is finite; and (b) the regular set Γ° has finitely many connected components, each of which is $C^{(m)}$ diffeomorphic to R^1 or S^1.*

Proof. By definition, a singular point, whether an end point or a Y-junction, must be isolated. If Γ^* is infinite, it must have at least one limit point $\boldsymbol{q}$ in Γ since Γ is compact. Now that $\boldsymbol{q}$ can no longer be singular, it must be regular. But by definition, a regular point has at least one neighborhood that contains only regular points, which leads to contradiction. This proves (a).

By definition, the regular set Γ° is a 1-D $C^{(m)}$ submanifold in R^2. According to the well known classification theorem for 1-D connected manifolds [37], each connected component of Γ° must be diffeomorphic to either S^1 (in the compact case) or R^1 (non-compact case). Since the skeleton curve is compact and each *regular piece* is isolated, the number of components diffeomorphic to S^1 must be finite. To prove (b), therefore, it suffices to show that the number of connected components diffeomorphic to R^1 is finite as well. For convenience, such components are said to be *open*.

Let C denote any open component. It can be naturally oriented and parameterized by R^1: $\boldsymbol{p} = \boldsymbol{p}(t), -\infty < t < \infty$. Let $\boldsymbol{q}_+ \in R^2$ be one of the limit points of $\boldsymbol{p}(t)$ as $t \to +\infty$, and $\boldsymbol{q}_-$ as $t \to -\infty$. They must exist and belong to Γ since Γ is compact. Notice that $\boldsymbol{q}_\pm$ cannot be regular since otherwise locally near $\boldsymbol{q}_\pm$, Γ is diffeomorphic to $R^1 \times 0$ in R^2. Thus $\boldsymbol{q}_\pm \in \Gamma^*$. Let $|\Gamma^*|$ denote the cardinality of the singular set, which is finite by (a). Then by tracking the unique unordered pair of $\boldsymbol{q}_\pm$ for each open component, one could easily show that the total number of open components is bounded by $3 \times |\Gamma^*|/2$. This completes the proof. □

As a result of the endpoint-tracking technique in the proof, we also conclude that

Corollary 1. *Let C denote any open connected component of the regular set Γ°. Then its closure in R^2 is diffeomorphic to either a compact interval $[a, b] \times 0$ in R^2, or the unit circle S^1 with at most one corner.*

In combination, the theorem and corollary imply that the total length $L = |\Gamma|$ of a skeleton curve Γ is well defined and must be finite.

Definition. (Skeleton Stripe Pattern) Let Γ be any compact skeleton curve and $d > 0$ a positive parameter. Then the open neighborhood

$$(\Gamma)_d = \{\boldsymbol{p} : \operatorname{dist}(\boldsymbol{p}, \Gamma) < \frac{d}{2}\} \tag{53}$$

is call a *skeleton stripe pattern*, and d its width. For convenience, Γ is called the *skeleton* of the pattern.

The main challenge is the inverse problem: for a given open domain which is the skeleton stripe pattern built upon some unknown skeleton curve, how to efficiently compute its width d and length L?

4.4.2. Geometric Measures for Width, Length, and Singular Sets

As for spots patterns, we start with simple stripe models for developing and analyzing proper geometric measures.

Let C denote a $C^{(m)}$ simple curve in R^2, parametrized by its arc-length: $\boldsymbol{r} = \boldsymbol{r}(s)$ with $0 \le s \le L$, where L is the total length. As a special case of skeleton curves, let $(C)_d$ denote its open neighborhood defined by (53). Assume that the aspect ratio $L/d \gg 1$, which is tantamount to letting $d \to 0$ since L will be fixed in the analysis.

Let T and N denote the unit tangent and normal of C (see Figure 15). Define a map from $D = [0, L] \times R$ to R^2 by:

$$(s, t) \to \boldsymbol{p}\,(s, t) = \boldsymbol{r}\,(s) + tN(s). \tag{54}$$

By the implicit function theorem, there exists some $b > 0$, so that $\boldsymbol{p}\,(s, t)$ is a diffeomorphism from $D_b = [0, L] \times (-b/2, b/2)$ to its image in R^2. For any $d < b$, $(C|d) = \boldsymbol{p}\,(D_d)$ is called a *tubular neighborhood* of C in differential topology [37]. Due to its explicit formula, the tubular neighborhood $(C|d)$ is much easier to work with than $(C)_d$ defined in (53). On the other hand, as $d \to 0$, they well match each other (except at the two ends). The definition of $(C)_d$, however, is more general since it belongs to general metric spaces.

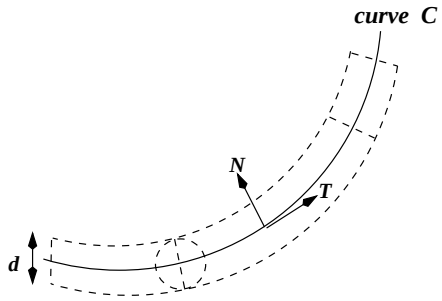

Figure 15: The tubular neighborhood $(C|d)$ of a simple curve segment C

Let $C(t)$ denote the parametric curve $\boldsymbol{p}\,(s, t), 0 \le s \le L$ for any fixed t. Then

$$\boldsymbol{p}_s(s, t) = T(s) - t\kappa T(s) = (1 - t\kappa)T(s),$$

where κ is the curvature. Thus the length $|C(t)|$ is give by

$$|C(t)| = \int_0^L |1 - t\kappa| ds \quad \text{leading to}$$

$$|\,|C(t)| - L| \le |t| \int_0^L |\kappa(s)| ds = |t| K_a, \tag{55}$$

where K_a denotes the total *absolute* curvature of C.

Now consider the total area of the tubular neighborhood $(C|d)$. Notice that

$$\boldsymbol{p}_t(s, t) = N(s) \quad \text{and} \quad \boldsymbol{p}_s(s, t) = (1 - t\kappa)T(s).$$

Therefore,

$$\text{area}(C|d) = \int_{-d/2}^{d/2}\int_0^L |\boldsymbol{p}_t \times \boldsymbol{p}_s|\, dsdt = \int_{-d/2}^{d/2}\int_0^L |1 - t\kappa|\, dsdt$$
$$= \int_{-d/2}^{d/2} |C(t)|\, dt. \quad (56)$$

Combined with equation (55), it leads to

$$|\text{area}(C|d) - Ld| \le \frac{K_a}{4} d^2. \quad (57)$$

Also notice that both equations (55) and (56) can be made exact if it is assumed that

$$\frac{d}{2} \le \min_{0\le s\le L} \frac{1}{|\kappa(s)|}, \quad \text{the lower bound of all curvature radii along } C.$$

Then $|t\kappa(s)| \le 1$ for any $t \in [-d/2, d/2]$, and

$$|C(t)| = L - tK \quad \text{and} \quad \text{area}(C|d) = Ld, \quad (58)$$

where $K = \int_C \kappa(s)ds$ is the total *signed* curvature of C.

Finally, let $v_2(\boldsymbol{p}) = v_2(x, y)$ denote the indicator function of the tubular domain $(C|d)$, whose closure is assumed to be inside a reaction-diffusion domain $\Omega \subset R^2$. Then by equation (55)

$$\int_\Omega |Dv_2| = \left|C\left(\frac{d}{2}\right)\right| + \left|C\left(-\frac{d}{2}\right)\right| + 2d = 2L + O(d(K_a + 2)). \quad (59)$$

Here the line segments $\boldsymbol{p}(0, [-d/2, d/2])$ and $\boldsymbol{p}(L, [-d/2, d/2])$ at the two ends have been taken into account. On the other hand, by equation (57),

$$\int_\Omega v_2 dxdy = \text{area}(C|d) = Ld + O(K_a d^2). \quad (60)$$

In combination, we have proven the following theorem.

Theorem 6. *Let K_a denote the total absolute curvature of C. Suppose the aspect ratio $\varepsilon = d/L \ll 1$. Then*

$$\begin{aligned} d &= \frac{2\int_\Omega v_2 dxdy}{\int_\Omega |Dv_2|}\,(1 + O(\varepsilon(K_a + 2))) \quad \textit{and} \\ L &= \frac{1}{2}\int_\Omega |Dv_2|\,(1 + O(\varepsilon(K_a + 2))), \end{aligned} \quad (61)$$

where the order symbol $O(\bullet)$ is independent of the given curve C and d.

The above analysis has been established only for a single piece of simple curve C. Note that the number 2 inside the order symbol $O(\bullet)$ is caused by the *two* endpoints of C. For a general skeleton curve Γ and its associated skeleton stripe pattern $(\Gamma)_d$ defined in the proceeding section, let v_2 denote the indicator function of $(\Gamma)_d$ in the interior of a reaction-diffusion domain Ω. By applying Theorem 6 to each one of the connected components predicted by Theorem 5, and by the special counting effect (as explained later) of the *functional* total absolute curvature K_a

$$K_a = K_a(d) = \int_\Omega |\kappa||Dv_2| = \int_\Omega \left|\nabla \cdot \left[\frac{\nabla v_2}{|\nabla v_2|}\right]\right| |Dv_2|, \tag{62}$$

we have the following theorem for general skeleton stripe patterns.

Theorem 7. *Let Γ be a compact skeleton curve with total length L, and $(\Gamma)_d$ its associated skeleton stripe pattern with $\varepsilon = d/L \ll 1$. Then*

$$d = \frac{2\int_\Omega v_2 dxdy}{\int_\Omega |Dv_2|}\,(1 + O(\varepsilon K_a)) \text{ and } L = \frac{1}{2}\int_\Omega |Dv_2|\,(1 + O(\varepsilon K_a)). \tag{63}$$

This theorem has been applied to a real stripe pattern generated from the reaction-diffusion equations (see Figure 16). Let A denote the side length of the reaction-diffusion domain. Then our numerical results show that $d = A/30$ and $L = 15A$, both implying that travelling horizontally (or vertically), one shall encounter 15 black-white cycles on average. Notice that the task of estimating the total length L is extremely difficult for human vision.

Furthermore, if the second order curvature measures are to be employed, it is even possible to extract more detailed geometric information about a stripe pattern. This is being explored by the authors in another work. Below we briefly state one of the most interesting results.

First, divide the singular set Γ^* of a skeleton curve Γ into two natural categories:

$$\Gamma_+ = \{\boldsymbol{p} \in \Gamma^* : \boldsymbol{p} \text{ is a } Y\text{-junction}\},$$
$$\Gamma_- = \{\boldsymbol{p} \in \Gamma^* : \boldsymbol{p} \text{ is an end point}\}.$$

Let $z_\pm$ denote their cardinalities separately. Our goal is to develop suitable geometric measures to compute or estimate $z_\pm$.

Similar to the definition of the total absolute curvature in (62), define the functional total *signed* curvature of a skeleton stripes pattern $(\Gamma)_d$ to be

$$K = K(d) = \int_\Omega \nabla \cdot \left[\frac{\nabla v_2}{|\nabla v_2|}\right] |Dv_2|, \tag{64}$$

Figure 16: Estimating the average width d and the total length L of a matured stripe pattern by Theorem 7. Our numerical results show that $d = A/30$ and $L = 15A$, where A is the side length of the reaction-diffusion domain. Both results imply that travelling horizontally (or vertically), one should encounter 15 black-white cycles on average. Notice that estimating the total length L is an extremely challenging task even for human vision

where $v_2(x, y)$ denotes the indicator function of $(\Gamma)_d$ in Ω.

Definition. (Linear Skeleton Curves) A skeleton curve Γ is said to be *linear* if any connected component of its regular set Γ° is a line segment.

Theorem 8. *Let Γ be a $C^{(m)}$ $(m > 2)$ compact skeleton curve inside a reaction-diffusion domain Ω. Then*

$$z_+ - z_- = \frac{1}{\pi} \lim_{d \to 0} K(d)\,, \tag{65}$$

$$z_+ + z_- \leq \frac{1}{\pi} \lim_{d \to 0} K_a(d)\,, \tag{66}$$

and the second equality holds if and only if Γ is a linear skeleton curve.

In particular, as in spline theory, define

$$x^{\pm} = \frac{|x| \pm x}{2}, \qquad \text{for any real scalar } x \in R,$$

and the two half-sided total curvatures,

$$K_{\pm} = K_{\pm}(d) = \int_{\Omega} \kappa^{\pm} |Dv_2| \ \text{ with } \ \kappa = \nabla \cdot \left[\frac{\nabla v_2}{|\nabla v_2|} \right].$$

Then Theorem 8 leads to the simple formulae:

$$z_{\pm} \leq \frac{1}{\pi} \lim_{d \to 0} K_{\pm}(d), \tag{67}$$

and the two equalities hold when Γ is a *linear* skeleton curve.

Theorem 8 and formulae in (67) allow to effectively estimate the singular points. They depend on the crucial assumption of convexity in the definition of Y-junctions.

5. Conclusion

In the current paper, statistical as well as geometric theories and frameworks have been developed to identify, classify, and characterize generic patterns arising from reaction-diffusion systems.

In the vast reaction-diffusion literature that permeates biology, chemistry, and computer simulations, the current paper presents the *first* most systematic work in the efforts of combining pattern-theoretic analysis with pattern formation research. Many novel ideas and methodologies have thus been developed inevitably at the risk of being completely wrong.

The present work once again proves, that the numerous contributions by the community of modern image, vision, and pattern analysis, are *not only* applicable in the important fields of computer vision, artificial intelligence, medical imaging, astronomic data analysis, or natural language processing, *but also* universally powerful for various feature and pattern mining tasks in physical, chemical, and biological simulations.

Acknowledgments

We are profoundly grateful to Professor Hans Othmer for his inspirational teaching on biochemical pattern formation and mathematical biology, without whose generous support and encouragement this project would have been absolutely impossible.

This project is partially supported by USA-NSF's Program of Applied Mathematics under grant number DMS-0202565.

References

[1] L. Alvarez, F. Guichard, P.-L. Lions, J.-M. Morel, Axioms and fundamental equations of image processing, *Arch. Rational Mech. Anal.*, **123** (1993), 199-257.

[2] G. Bellettini, G. Dal Maso, M. Paolini, Semicontinuity and relaxation properties of a curvature depending functional in 2D, *Ann. Scuola Norm. Sup. Pisa Cl. Sci., 4*, **20** (1993), 247-297.

[3] V. Castets, E. Dulos, J. Boissonde, P. De Kepper, Experimental evidence of a sustained standing Turing-type non-equilibrium chemical pattern, *Phys. Rev. Lett.*, **64** (1990), 2953-2956.

[4] A. Chambolle, P.L. Lions, Image recovery via Total Variational minimization and related problems, *Numer. Math.*, **76** (1997), 167-188.

[5] T.F. Chan, S.-H. Kang, J. Shen, Euler's elastica and curvature based inpainting, *SIAM J. Appl. Math.*, **63**, No. 2 (2002), 564-592.

[6] T.F. Chan, S. Osher, J. Shen, The digital TV filter and nonlinear denoising, *IEEE Trans. Image Process.*, **10**, No. 2 (2001), 231-241.

[7] T.F. Chan, J. Shen, Variational restoration of non-flat image features, models and algorithms, *SIAM J. Appl. Math.*, **61**, No. 4 (2000), 1338-1361.

[8] T.F. Chan, J. Shen, Mathematical models for local nontexture inpaintings, *SIAM J. Appl. Math.*, **62**, No. 3 (2001), 1019-1043.

[9] T.F. Chan, J. Shen, On the role of the BV image model in image restoration, *Amer. Math. Soc. Contemporary Mathematics,* volume on *Recent Advances in Scientific Computing, Partial Differential Equations,* Editors*, S. Y. Cheng, C.-W. Shu, and T. Tang*, **330** (2003), 25-41.

[10] T.F. Chan, J. Shen, Vaiational image inpainting, *Comm. Pure Applied Math.,* in press (2004).

[11] T.F. Chan, J. Shen, *Image Analysis and Processing*, Variational PDE, wavelets, and stochastic methods, SIAM Publisher, Philadelphia (2005).

[12] T.F. Chan, J. Shen, L. Vese, Variational PDE models in image processing, *Notices Amer. Math. Soc.*, **50** (2003), 14-26.

[13] T.F. Chan, L.A. Vese, Active contours without edges, *IEEE Trans. Image Process.*, **10**, No. 2 (2001), 266-277.

[14] A. Cohen, W. Dahmen, I. Daubechies, R. DeVore, Harmonic analysis of the space BV, *Revista Matematica Iberoamericana,* submitted (2002).

[15] T.M. Cover, J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Inc., New York (1991).

[16] I. Daubechies, *Ten lectures on wavelets*, SIAM, Philadelphia (1992).

[17] D.C. Dobson, C.R. Vogel, Convergence of an iterative method for total variation denoising, *SIAM J. Numer. Anal.*, **34**, No. 5 (1997), 1779-1791.

[18] D.L. Donoho, I.M. Johnstone, Ideal spacial adaption by wavelet shrinkage, *Biometrika*, **81** (1994), 425-455.

[19] R.O. Duda, P.E. Hart, D.G. Stork, *Pattern Classification*, John Wiley & Sons, Inc., New York (2001).

[20] S. Esedoglu, J. Shen, Digital inpainting based on the Mumford-Shah-Euler image model, *European J. Appl. Math.*, **13** (2002), 353-370.

[21] L.C. Evans, J. Spruck, Motion of level sets by mean curvature, *J. Diff. Geom.*, **33**, No. 3 (1991), 635-681.

[22] W.H. Fleming, R. Rishel, An integral formula for total gradient variation, *Arch. Math.*, **11** (1960), 218-222.

[23] S. Geman, D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Machine Intell.*, **6**, 721-741 (1984).

[24] W. Gibbs, *Elementary Principles of Statistical Mechanics*, Yale University Press (1902).

[25] E. De Giorgi, Frontiere orientate di misura minima, *Sem. Mat. Scuola Norm. Sup. Pisa* (1960-1961).

[26] E. Giusti, *Minimal Surfaces and Functions of Bounded Variation*, Birkhäuser, Boston (1984).

[27] U. Grenander, *Lectures in Pattern Theory. I. II. and III*, Springer (1976-1981).

[28] W. Heisenberg, Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik, *Zeitschrift für Physik*, **43** (1927), 172-198.

[29] G. Kanizsa, *Organization in Vision*, Praeger, New York (1979).

[30] L. Lengyel, I.R. Epstein, A chemical approach to designing Turing patterns in reaction-diffusion systems, *Proc. Nat. Acad. Sci. USA*, **89** (1992), 3977-3979.

[31] L. Lengyel, G. Rabai, I.R. Epstein, Experimental and modeling study of oscillations in chlorine dioxide-iodinacid reaction, *J. Amer. Chem. Soc.*, **112** (1990), 9104-9110.

[32] K. Lin, S.G. Schirmer, E.T. Camacho Wirkus, Chemical pattern formation in reaction-diffusion systems, Technical report, MSRI Summer program on Dynamics of Low Dimensional Continua, U.C. Berkeley (1999).

[33] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press (1998).

[34] D. Marr, E. Hildreth, Theory of edge detection, *Proc. Royal Soc. London, B*, **207** (1980), 187-217.

[35] S. Masnou, J.-M. Morel, Level-lines based disocclusion, *Proceedings of 5th IEEE Int'l Conf. on Image Process., Chicago*, **3** (1998), 259-263.

[36] Y. Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, volume **22** of *University Lecture Series*, AMS, Providence (2001).

[37] J.W. Milnor, *Topology from the Differentiable Viewpoint*, Princeton Univ. Press, revised edition (1997).

[38] D. Mumford, Elastica and computer vision, In C. L. Bajaj, editor, *Algebraic Geometry and its Applications*, pages 491-506. Springer-Verlag, New York (1994).

[39] D. Mumford, *Geometry Driven Diffusion in Computer Vision*, chapter "The Bayesian rationale for energy functionals", pages 141-153, Kluwer Academic (1994).

[40] D. Mumford, *Pattern Theory*, The mathematics of perception, *Int'l Congress Mathematicians (ICM)*, III, Beijing (2002).

[41] D. Mumford, B. Gidas, Stochastic models for generic images, *Quaterly of Applied Mathematics*, **59** (2001), 85-111.

[42] D. Mumford, J. Shah, Optimal approximations by piecewise smooth functions and associated variational problems, *Comm. Pure Applied. Math.*, **42** (1989), 577-685.

[43] J.D. Murray, *Mathematical Biology*, Springer-Verlag, New York (1993).

[44] M. Nitzberg, D. Mumford, T. Shiota, *Filtering, Segmentation, and Depth*, Lecture Notes in Comp. Sci., **662**, Springer-Verlag, Berlin (1993).

[45] A.V. Oppenheim, R.W. Schafer, *Discrete-Time Signal Processing*, Prentice Hall Inc., New Jersey (1989).

[46] H.G. Othmer, *Mathematical Models of Biological Pattern Formation*, Lecture notes and book to appear (2003).

[47] H.G. Othmer, L.E. Scriven, On the eigenvalues of the matrix pencil $A+\mu B$, *Z. Angew. Math. Phys.*, **24** (1973), 135-139.

[48] J.E. Pearson, Complex patterns in a simple system, *Science*, **261** (1993), 189-192.

[49] P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Machine Intell.*, **12** (1990), 629-639.

[50] T. Poggio, S. Smale, The mathematics of learning, dealing with data, *Notices Amer. Math. Soc.*, **50**, No. 5 (2003), 537-544.

[51] I. Prigogine, R. Lefever, Symmetry breaking instabilities in dissipative systems, *J. Chem. Phys.*, **48** (1968), 1695-1700.

[52] G.-C. Rota, J. Shen, On the combinatorics of cumulants, *J. Comb. Theory (A)*, **91**, No. 1 (2000), 283-304.

[53] L. Rudin, S. Osher, Total variation based image restoration with free local constraints, *Proc. 1st IEEE ICIP*, **1** (1994), 31-35.

[54] L. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica D*, **60** (1992), 259-268.

[55] J. Shah, Elastica with hinges, *J. Visual Comm. Image Rep.*, **13** (2002), 36-43.

[56] C.E. Shannon, A mathematical theory of communication, *The Bell System Technical Journal*, **27**, 379-423 (1948), 623-656.

[57] J. Shen, Geometric inpainting and applications, *Proc. SPIE*, **4792** (2002), 102-113.

[58] J. Shen, On the foundations of vision modeling I. Weber's law and Weberized TV restoration, *Physica D*, **175** (2003), 241-251.

[59] J. Shen, Bayesian video dejittering by BV image model, *SIAM J. Appl. Math.*, **64**, No. 5, 1691-1708 (2004).

[60] J. Shen, G. Strang, Asymptotic analysis of Daubechies polynomials, *Proc. Amer. Math. Soc.*, **124** (1996), 3819-3833.

[61] J. Shen, G. Strang, Asymptotics of Daubechies filters, scaling functions and wavelets, *Appl. Comput. Harmon. Anal.*, **5** (1998), 312-331.

[62] G. Strang, T. Nguyen, *Wavelets and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA (1996).

[63] D.S. Thompson, J.S. Nair, S.S.D. Venkata, R.K. Machiraju, M. Jiang, G. Craciun, Physics-based feature mining for large data exploration, *Computing Sci. Eng.*, **4**, No. 4 (2002), 22-30.

[64] A. M. Turing, The chemical basis of morphogenesis, *Philos. Trans. R. Acad. Sci., B*, **237** (1952), 37-72.

[65] G. Turk, Generating textures on arbitrary surfaces using reaction-diffusion, *Computer Graphics, SIGGRAPH '91*, **25** (1991), 289-298.

[66] C. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia (2002).

[67] J.D. Watson, F.H.C. Crick, A structure for deoxyribose nucleic acid, *Nature*, **171** (1953), 737-738.

[68] R. Williams, Xmorphia, Web Publishing at Center for Advanced Computing Research, Caltech, Pasadena, California. http:www.cacr.caltech.edu roy.

[69] A. Witkin, M. Kass, Reaction-diffusion textures, *Computer Graphics, SIGGRAPH '91*, **25**, No. 3 (1991).

[70] S.C. Zhu, D. Mumford, Prior learning and Gibbs reaction-diffusion, *IEEE Trans. Pattern Anal. Machine Intell.*, **19**, No. 11 (1997), 1236-1250.

[71] S.C. Zhu, Y.N. Wu, D. Mumford, Minimax entropy principle and its applications to texture modeling, *Neural Computation*, **9** (1997), 1627-1660.