10.1098/rsta.2003.1267

THE ROYAL SOCIETY

Pattern formation in large domains

By A. M. Rucklidge

Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, UK

Published online 3 November 2003

Pattern formation is a phenomenon that arises in a wide variety of physical, chemical and biological situations. A great deal of theoretical progress has been made in understanding the universal aspects of pattern formation in terms of amplitudes of the modes that make up the pattern. Much of the theory has sound mathematical justification, but experiments and numerical simulations over the last decade have revealed complex two-dimensional patterns that do not have a satisfactory theoretical explanation. This paper focuses on quasi-patterns, where the appearance of small divisors causes the standard theoretical method to fail, and ends with a discussion of other outstanding problems in the theory of two-dimensional pattern formation in large domains.

Keywords: pattern formation; quasi-patterns; small divisors

1. Introduction

There is a diverse collection of physical, chemical and biological systems that naturally organize themselves into patterns. In these various situations, the same types of qualitative behaviour appear repeatedly, and universal mathematical models have been developed to understand each characteristic situation. These mathematical models of pattern formation provide a unifying viewpoint and have, in turn, stimulated further research in the relevant experimental disciplines. Pattern formation remains a topic of great current interest that spans diverse areas of pure mathematics, applied mathematics and experimental science.

One well-studied example of a pattern-forming instability is the Faraday wave problem of the formation of waves on the surface of a layer of fluid as it is driven by vertical vibrations. This system has been subjected to intensive scrutiny in laboratory experiments and has come to be regarded as an archetypal pattern-forming system. Clear examples of pattern formation occur in a wide range of other systems, including Rayleigh–Bénard convection, liquid crystals in externally imposed electric fields, nonlinear optics, directional solidification, vibrated granular media, chemical reactions, and catalytic oxidation.

Laboratory experiments in pattern formation have continually prompted theoretical developments, and the physical insights that they reveal are essential to a complete understanding of these phenomena. Numerical simulations have also played a central role, and with advances in experimental technique and computing power,

One contribution of 22 to a Triennial Issue 'Mathematics, physics and engineering'.

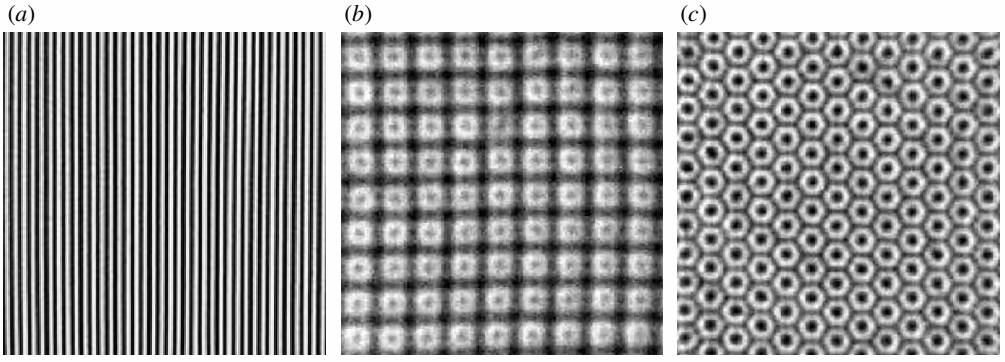

Figure 1. Examples of experimentally observed patterns, viewed from above. (*a*) Stripes (straight rolls) in Rayleigh–Bénard convection. (Reproduced with permission from Cakmur *et al*. (1997). Copyright (1997) American Physical Society.). (*b*), (*c*) Squares and hexagons in a two-frequency forced Faraday wave experiment. (Reproduced with permission from Arbell & Fineberg (2002). Copyright (2002) American Physical Society.) Note the extraordinary degree to which the patterns display spatial periodicity, as well as rotation and reflection symmetry.

attention has turned from smaller to larger domains. Many new types of behaviour have been discovered in recent years, including quasi-patterns and spiral defect chaos. Theoretical understanding of these new types of behaviour is very much lacking, in some cases, apparently, for deep mathematical reasons.

The simplest patterns—stripes, squares and hexagons—have reflection, rotation and translation symmetries; experimentally observed examples of these are shown in figure 1. A comprehensive and very successful theory has been developed to analyse the creation of these patterns from an initial featureless state. This theory, which is based on computing the amplitudes of the various waves (or modes) that make up the pattern, is known as equivariant bifurcation theory, and is expounded in detail in a series of texts (see, for example, Golubitsky & Stewart 2002).

In order to apply rigorous mathematical theories to explain experimental results and other occurrences of pattern formation in the natural world, there are naturally a series of idealizations and approximations that must be made. The first assumption concerns modelling: the experimental configuration is supposed to be describable in terms of some set of equations that predict the future evolution of the system, given its current state. The evolution laws often take the form of partial differential equations (PDEs), particularly when the system under consideration involves a fluid. In many situations, the next idealization is to suppose that in the absence of any driving force, the system will remain featureless, and that if the forcing is turned up, it must reach a critical level before it can overcome any inherent dissipation in the system. If the level of forcing (which is a parameter under the control of the experimentalist) exceeds this critical value, the featureless state will be unstable, and any small disturbances will grow. These cannot grow for ever, and one possible outcome is that the system will settle down to a steady state with some degree of spatial structure: a pattern.

Two further idealizations are often made when computing the mathematical properties of patterns. First, the experimental boundaries are ignored, and so in effect the experiment is supposed to be taking place in a container of infinite size; and second, the observed pattern is supposed to have perfect spatial periodicity. By only con-

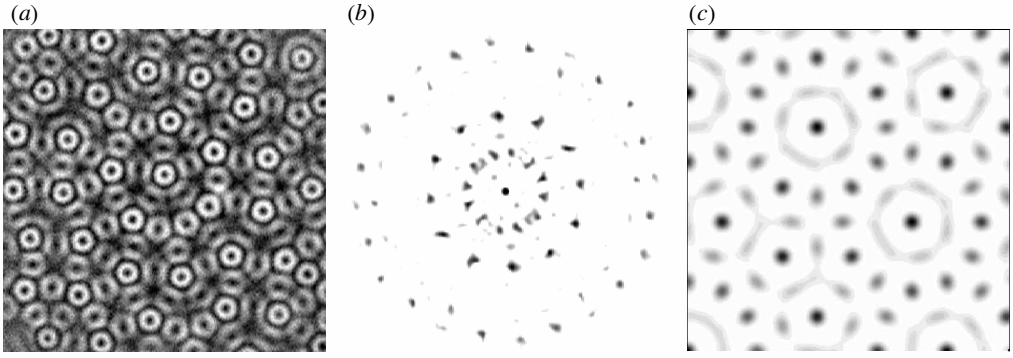

Figure 2. Quasi-patterns: (*a*) 12-fold quasi-pattern observed in a two-frequency forced Faraday wave experiment; (*b*) spatial Fourier transform, showing the 12-fold rotational order in spite of the absence of any translation symmetry. ((*a*) and (*b*) reproduced with permission from Arbell & Fineberg (2002). Copyright (2002) American Physical Society.) (*c*) Synthetic quasi-pattern, constructed from the sum of 12 modes with wavevectors spaced equally around a circle (see equation (2.6)).

sidering patterns that are periodic in space, rigorous theory can be applied to prove the existence of stripe, square and hexagon (and other) solutions of the nonlinear PDEs that model the experimental situation. Given that in some highly controlled experiments the idealization of spatial periodicity appears to hold over dozens of repeats of the pattern (as in figure 1), these assumptions are perfectly reasonable when the objective is to understand the nature of these periodic patterns.

However, experiments that are carried out in large domains are quite capable of producing patterns that cannot be analysed in this way. A notable example of this is quasi-patterns, which are most readily found in Faraday wave experiments in which a tray of liquid is subjected to vertical vibrations with two commensurate forcing frequencies (Edwards & Fauve 1994). A recent survey of experimental results can be found in Arbell & Fineberg (2002), and one experimental example of a quasi-pattern is shown in figure 2*a*. This pattern is quasi-periodic in any horizontal direction, that is, the amplitude of the pattern (taken along any direction in the plane) can be regarded as a sum of modes with incommensurate spatial frequencies. In general, quasi-patterns exhibit long-range rotational order, most evident in their spatial Fourier transform (figure 2*b*), but they lack spatial periodicity. In this respect, there are obvious similarities with quasi-crystals, which were discovered about a decade earlier (Levine & Steinhardt 1984). Examples of quasi-crystals that are quasi-periodic in one, two or three spatial directions have been found (Janot 1994).

Models of quasi-patterns have been developed by several researchers without the theoretical background required to justify their use. These models are derived using methods that are successful for periodic patterns; however, when the methods are applied to the case of quasi-patterns, a difficulty known as the problem of small divisors arises. This problem appears whenever quasi-periodic behaviour is found in a nonlinear set of differential equations and attempts are made to compute the quasi-periodic solution by a series of approximations, using a method known as perturbation theory. Given an approximate solution, a correction can be found, which, when added to the first approximation, yields a new approximate solution that is (hopefully) closer to a solution of the equations. In many cases, including the case

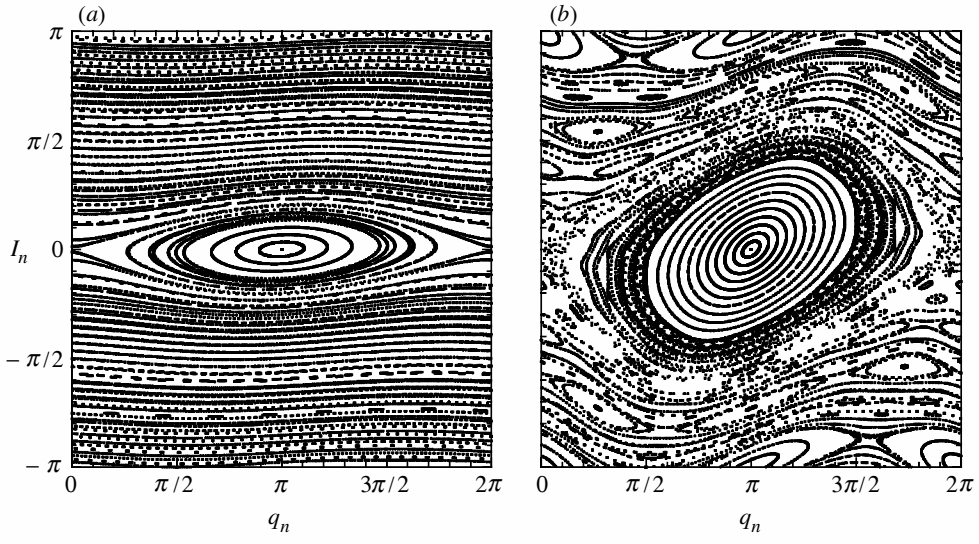

Figure 3. Trajectories in the standard map (1.1): (*a*) $\epsilon = 0.1$; (*b*) $\epsilon = 0.8$. For small ϵ, several features are apparent: there are fixed points (at $(\theta, I) = (0, 0)$ and $(\pi, 0)$), periodic orbits and two types of quasi-periodic orbit—those that have a bounded range of θ (in an island centred on $(\pi, 0)$), and those for which θ increases or decreases monotonically. For larger ϵ, more islands are visible, as well as chaotic dynamics between the islands, and yet some quasi-periodic trajectories persist.

of spatially periodic patterns, it can be proved that this process, if carried to the limit, will indeed converge to a true solution. However, in the case of quasi-periodic behaviour, the corrections turn out not to be uniformly small, owing to the appearance of small numbers in the denominators, and convergence is called into question.

This difficulty was faced first by Poincaré in the context of celestial mechanics in the late nineteenth century. In the absence of any gravitational interaction between planets, each planet in the Solar System orbits the Sun with its own period, and the system as a whole is quasi-periodic in time. Poincaré considered the question of whether or not the Solar System is quasi-periodic given the presence of weak interactions between the planets. Formally, the problem could be solved by perturbation theory, but Poincaré realized that small divisors called convergence of the perturbation series into question.

The small-divisor issue was resolved for this type of problem by Kolmogorov, Arnol'd & Moser (KAM) in the 1950s and 1960s, who showed under what circumstances quasi-periodic behaviour would be found (see, for example, Moser 1973). To take an example, consider the so-called standard map:

$$I_{n+1} = I_n + \epsilon \sin(\theta_n), \qquad \theta_{n+1} = \theta_n + I_{n+1} \mod 2\pi, \tag{1.1}$$

which models a freely rotating pendulum in the absence of gravity, subjected to periodic impulsive forces. The map also models a chain of particles connected by springs and subjected to a sinusoidal potential (Aubry 1983). Here, n plays the role of time (or space in the particle model). When $\epsilon = 0$, all trajectories are of the form $(\theta_n, I_n) = (\theta_0 + nI_0, I_0) \mod 2\pi$, and are periodic with period q if $I_0/2\pi = p/q$ is rational (with p and q integers), and quasi-periodic otherwise. Both periodic and

quasi-periodic orbits lie on horizontal lines (invariant curves) in the (θ, I)-plane, but the lines are made up of individual periodic points in the first case, while a quasi-periodic orbit will eventually visit a neighbourhood of each point on the line. When ϵ is perturbed away from zero (see figure 3*a*), the question is which of these families of trajectories will persist as invariant curves of the map? The essential content of the KAM theorem is that, for small enough perturbations, and for almost every irrational value of $I_0/2\pi$, there will be an invariant curve close to the unperturbed invariant curve, and the corresponding quasi-periodic trajectory survives the perturbation. The curves that persist are those that satisfy a Diophantine condition, that is, for which there are constants $K > 0$ and $\delta > 0$ such that $I_0/2\pi$ satisfies

$$\left| p - \frac{I_0}{2\pi} q \right| \geqslant \frac{K}{(|p| + |q|)^{\delta}} \tag{1.2}$$

for every pair of integers p and q, apart from $(0, 0)$. The exponent δ is an indication of the 'irrationality' of $I_0/2\pi$, so, for example, $(\sqrt{5} - 1)/2$ satisfies (1.2) with $\delta = 1$. In general, curves with smaller values of δ persist to larger values of the perturbation ϵ. Invariant curves with rational values of $I_0/2\pi$ are immediately broken up into elliptic and hyperbolic periodic points, with a web of chaotic trajectories near the hyperbolic equilibria (see figure 3*b*).

KAM theory has been applied successfully to a variety of problems in which small divisors arise, for instance quasi-periodicity in the Solar System and in the dynamics of charged particles in tokamak magnetic fields. However, the methods of KAM (based around canonical coordinate transformations) were developed for problems in which quasi-periodicity occurs in only *one* direction (time), whereas quasi-patterns are quasi-periodic in *two* spatial directions. For this reason, KAM theory is not applicable to quasi-patterns, at least not directly, and either the theory must be extended to cover this case, or alternative methods must be developed. In principle, similar issues arise in solid-state quasi-crystals, though the main theoretical approaches for these are developed around aperiodic Penrose tilings of the plane or three-dimensional space, and around projecting higher-dimensional periodic lattices down to three dimensions (Janot 1994), whereas a wave-based approach is more natural for the fluid dynamical quasi-patterns.

The purpose of this paper is to draw attention to some of the theoretical difficulties that are preventing progress in the development of a mathematical understanding of two-dimensional pattern formation in large (or unbounded) domains. The example discussed in most detail is that of quasi-patterns, where some progress has recently been made in coming to terms with the small-divisor problem (Rucklidge & Rucklidge 2003). Their results are outlined here: § 2 introduces a particularly simple pattern-forming PDE (the Swift–Hohenberg equation) and indicates how the small divisors arise. Limits on the magnitude of these small divisors are calculated in § 3, and the perturbation theory for the quasi-pattern solution of the Swift–Hohenberg equation is concluded in § 4, with an indication that the problem of small divisors does indeed cause the perturbation theory to fail. In the last section, these results are related to other outstanding problems in pattern formation associated with large domains, including the intriguing state of spiral defect chaos, and the problem of developing rigorous descriptions of long-wavelength instabilities of two-dimensional patterns. All these are connected to what is perhaps the fundamental open question of pattern formation: why do periodic patterns form at all in large domains?

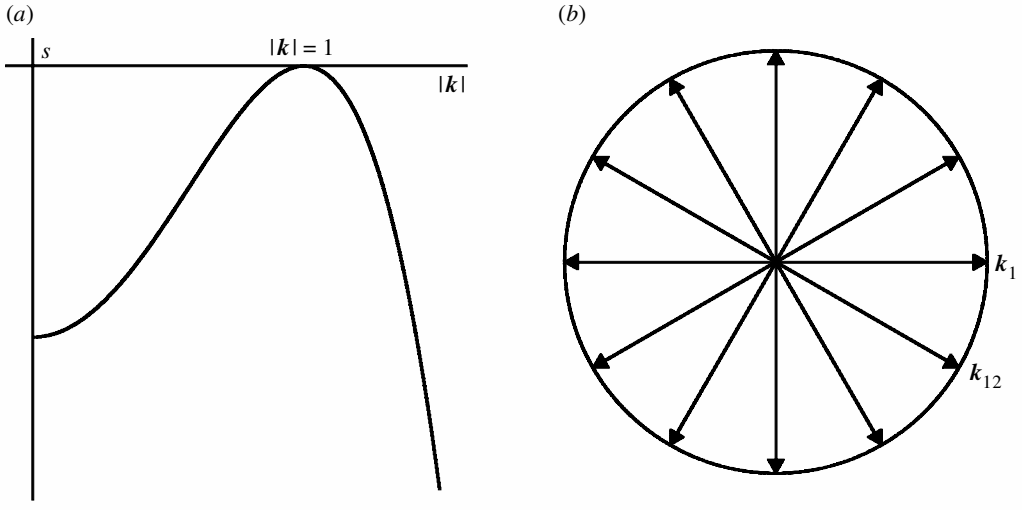

Figure 4. (*a*) Schematic growth (decay) rate s of a mode $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$, as a function of $|\boldsymbol{k}|$ at $\mu = 0$. Modes with $|\boldsymbol{k}| = 1$ are marginally stable. (*b*) Twelve wavevectors on the circle $|\boldsymbol{k}| = 1$. Adding equal amounts of 12 modes with these wavevectors (numbered $\boldsymbol{k}_1$ to $\boldsymbol{k}_{12}$) results in the synthetic pattern in figure 2*c*.

2. Model equations

One of the key mathematical questions concerning quasi-patterns is one of *existence*: do PDEs that model pattern-forming problems have solutions that are quasi-periodic in space, along the lines of the experimentally observed pattern in figure 2*a*? Rather than try to answer this question in the context of a PDE that specifically models the Faraday wave problem, it seems sensible to start with the simplest possible pattern-forming PDE: the Swift–Hohenberg equation (Swift & Hohenberg 1977). In fact, considering the Swift–Hohenberg equation is not such a simplification, since many pattern-forming problems can be cast into this form, or variations (Melbourne 1999). The simplest variant is

$$\frac{\partial U}{\partial t} = \mu U - (1 + \nabla^2)^2 U - U^3. \tag{2.1}$$

The equation is posed on the plane, with $\boldsymbol{x} = (x, y) \in \mathbb{R}^2$, and $U(x, y, t) \in \mathbb{R}$ supposed to be bounded as $(x, y) \to \infty$. The parameter μ represents the force that will drive the pattern formation.

This PDE has a spatially uniform trivial solution $U(x, y, t) = 0$, and the stability of this solution can be investigated by linearizing (2.1). The linearized equation has wavelike solutions: $U = \mathrm{e}^{st}\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$, with growth rate s and wavevector $\boldsymbol{k}$, with the growth rate related to μ and $|\boldsymbol{k}|$ by $s = \mu - (1 - |\boldsymbol{k}|^2)^2$. This relation is plotted in figure 4*a* in the case $\mu = 0$: with this value of μ, all modes are damped (have negative growth rate) apart from those with wavenumber $|\boldsymbol{k}|$ equal to 1. With μ just above zero, modes with $|\boldsymbol{k}|$ close to 1 will grow, until the nonlinear term in (2.1) causes the amplitudes of these modes to saturate at a level related to the value of μ.

In many pattern-forming problems, the standard method known as perturbation theory can be used to compute how the amplitude saturates, with the assumption that the parameter μ and the amplitude of the pattern are both very small. This

degree of smallness is explicitly introduced as a small parameter $\epsilon \ll 1$, and U is written in the form

$$U = \epsilon U_1 + \epsilon^3 U_3 + \epsilon^5 U_5 + \cdots . \tag{2.2}$$

The absence of even terms ($\epsilon^2 U_2$) is because of the symmetry $U \to -U$ in equation (2.1). The connection between the small forcing μ and the small parameter ϵ is made explicit by setting $\mu = \epsilon^2$. The expansion (2.2) is inserted into the Swift–Hohenberg equation (2.1) and like powers of ϵ are collected together:

$$0 = \epsilon \mathcal{L}(U_1) + \epsilon^3 (U_1 + \mathcal{L}(U_3) - U_1^3) + \epsilon^5 (U_3 + \mathcal{L}(U_5) - 3U_1^2 U_3) + \cdots ,$$

where, to make the presentation simpler, only steady patterns are considered. The linear differential operator $\mathcal{L}(U)$ is $-(1 + \nabla^2)^2 U$.

In order for this equation to be satisfied for all parameter values, the coefficient of each power of ϵ must separately be zero, and so the equation can be solved formally by considering each power of ϵ in turn. The leading-order equation is

$$\mathcal{L}(U_1) = 0. \tag{2.3}$$

The operator $\mathcal{L}$ acting on a mode $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$ yields $-(1 - |\boldsymbol{k}|^2)^2 \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$, which is zero only when $|\boldsymbol{k}| = 1$, so equation (2.3) has non-trivial solutions that are made up of linear combinations of modes with wavevectors $\boldsymbol{k}$ on the unit circle. Any set of such wavevectors is possible at this level, but a natural choice to make when studying quasi-patterns is

$$U_1(x, y) = \sum_{j=1}^{12} A_j \mathrm{e}^{\mathrm{i}\boldsymbol{k}_j \cdot \boldsymbol{x}},$$

where the 12 vectors $\boldsymbol{k}_1$ to $\boldsymbol{k}_{12}$ are equally spaced around the circle (figure 4*b*). This choice of modes is inspired by the evidence in the Fourier transforms of experimentally observed quasi-patterns (as in figure 2*b*). In order for U to be real, the amplitudes must satisfy $A_{j+6} = \bar{A}_j$. Setting each A_j to the same real value results in a quasi-pattern of the form depicted in figure 2*c*.

At third order in ϵ, the equation to solve is

$$\mathcal{L}(U_3) = -U_1 + U_1^3 = -\sum_{j=1}^{12} A_j \mathrm{e}^{\mathrm{i}\boldsymbol{k}_j \cdot \boldsymbol{x}} + \sum_{j=1}^{12} \sum_{k=1}^{12} \sum_{l=1}^{12} A_j A_k A_l \mathrm{e}^{\mathrm{i}(\boldsymbol{k}_j + \boldsymbol{k}_k + \boldsymbol{k}_l)\cdot \boldsymbol{x}}. \tag{2.4}$$

Notice that U_1^3 contains cubic interactions between the modes in U_1, which take the form of modes with all possible combinations of three of the 12 original wavevectors (allowing repeats). Some combinations (for example, $\boldsymbol{k}_1 + \boldsymbol{k}_1 + \boldsymbol{k}_7 = \boldsymbol{k}_1$) lie on the unit circle, but most ($\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3$) do not. Modes with different wavevectors are orthogonal, so the coefficients of each mode on the left and the right of equation (2.4) must be equal. In particular, the coefficient of modes with wavevectors on the unit circle is zero on the left, since $\mathcal{L}$ acting on such a mode is zero. Setting the coefficient of (for example) $\mathrm{e}^{\mathrm{i}\boldsymbol{k}_1 \cdot \boldsymbol{x}}$ to zero on the right-hand side results in an equation relating the amplitudes of the modes:

$$0 = A_1 - 3(|A_1|^2 + 2|A_2|^2 + 2|A_3|^2 + 2|A_4|^2 + 2|A_5|^2 + 2|A_6|^2)A_1, \tag{2.5}$$

with similar equations resulting from the other modes. One solution of the amplitude equations is for all the amplitudes to be zero (the trivial solution); setting all amplitudes to have the same non-zero modulus results in a quasi-pattern. One particular

solution is $A_1 = \cdots = A_{12} = 1/\sqrt{33}$, and so, in terms of the original variables, the pattern is

$$U(x,y) = \sqrt{\frac{\mu}{33}} \sum_{j=1}^{12} \mathrm{e}^{\mathrm{i}\boldsymbol{k}_j \cdot \boldsymbol{x}} + \cdots . \tag{2.6}$$

This result suggests that the quasi-pattern solution is created when μ increases through zero, with an amplitude proportional to $\sqrt{\mu}$.

This might appear to be the end of the story: the amplitude of the quasi-pattern has been computed as a function of the driving force, and a little more effort leads to an estimate of the stability of the pattern. This kind of calculation has been carried out in a variety of situations, starting either from equations describing the Faraday wave experiment or other experiments, or just using considerations of the symmetry of the quasi-pattern. All these calculations result in amplitude equations similar to (2.5), and all suffer from two severe drawbacks.

The first drawback is that equation (2.5) determines only the amplitudes of the complex numbers A_j, and not their phase. In all, there are six free phases: two of these are fixed by considering resonances that occur at fifth order; two are genuinely free, and are associated with translating (but not changing) the pattern; and two phases (called phason modes) are not determined even by high-order resonances. In this context, the phason modes describe relative translations of two hexagonal sublattices generated by $\boldsymbol{k}_1$, $\boldsymbol{k}_3$, $\boldsymbol{k}_5$ and $\boldsymbol{k}_2$, $\boldsymbol{k}_4$, $\boldsymbol{k}_6$, and may play a role in long-wave instabilities of the quasi-pattern (Echebarria & Riecke 2001). However, as they have a marked effect on the appearance of the pattern, they ought to be determined in a satisfactory theory without long-wave considerations.

The second drawback becomes apparent only when an attempt is made to compute higher-order corrections to the pattern. Returning to equation (2.4), all modes with wavevectors on the unit circle have already been taken into account by solving (2.5). The remaining modes all have wavevectors off the unit circle ($|\boldsymbol{k}| \neq 1$), and so the linear operator $\mathcal{L}$ can be inverted to find U_3:

$$U_3 = -\sum_{|\boldsymbol{k}_j+\boldsymbol{k}_k+\boldsymbol{k}_l|\neq 1} \frac{A_j A_k A_l}{(1-|\boldsymbol{k}_j+\boldsymbol{k}_k+\boldsymbol{k}_l|^2)^2} \mathrm{e}^{\mathrm{i}(\boldsymbol{k}_j+\boldsymbol{k}_k+\boldsymbol{k}_l)\cdot\boldsymbol{x}},$$

since the operator $\mathcal{L}^{-1}$ acting on a mode $\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}$ yields $-\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}}/(1-|\boldsymbol{k}|^2)^2$, defined as long as $|\boldsymbol{k}| \neq 1$.

However, if $|\boldsymbol{k}|$ is close to one, $\mathcal{L}^{-1}(\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}})$ can be arbitrarily large. This does not pose difficulties for computing U_3, but continuing the calculation to higher-order results in combinations of vectors that can come arbitrarily close to the unit circle. Specifically, U_3 involves sums of three of the original 12 vectors, and U_N will involve integer combinations of up to N of the 12 vectors $\boldsymbol{k}_1$ to $\boldsymbol{k}_{12}$. If the original choice of vectors had been two, four or six, in an attempt to describe striped, square or hexagonal patterns, the integer combinations of vectors arising at high order would not have come close to the unit circle, instead forming a lattice. Choosing 12 evenly spaced vectors (or any other number) leads to integer combinations of vectors that come arbitrarily close to the unit circle. Small divisors arise when the operator $\mathcal{L}$ is inverted, which raises doubts as to whether or not the power series (2.2) for U will converge.

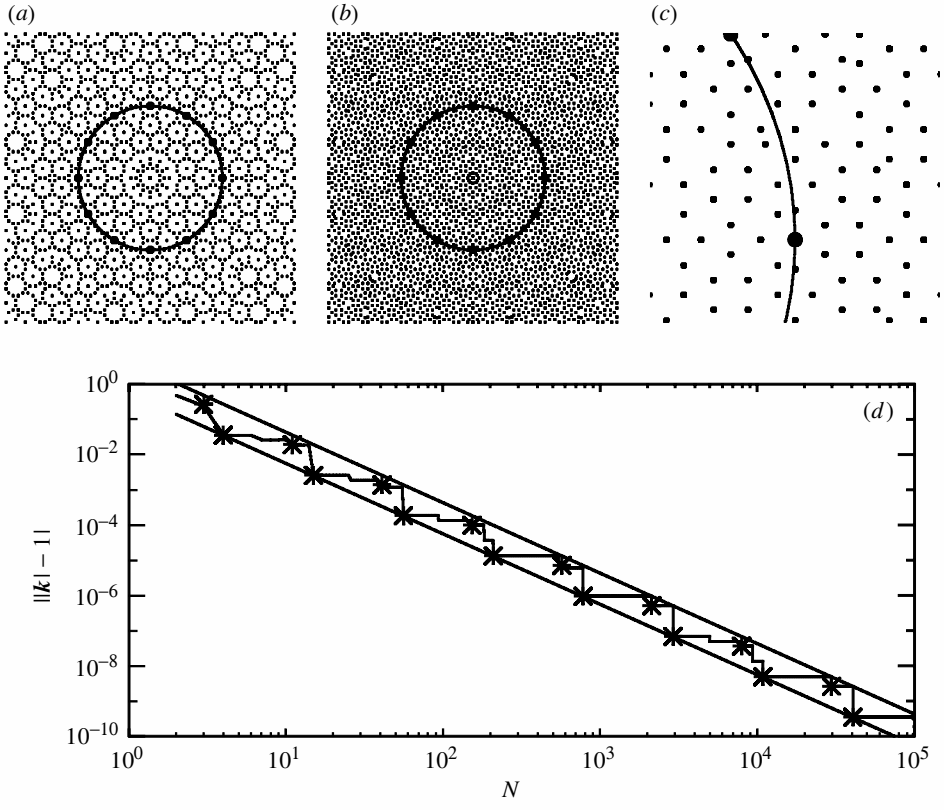

Figure 5. Positions of combinations of up to N of the original 12 vectors on the unit circle, with (*a*) $N = 11$, (*b*) $N = 15$; (*c*) detail of (*b*). The circle indicates the unit circle, $|\boldsymbol{k}| = 1$, the large dots are the original 12 wavevectors, and the small dots are integer combinations of these. Note how the density of points increases with N, and the proximity of points to the unit circle decreases with N. (*d*) Smallest non-zero distances from the unit circle $||\boldsymbol{k}_m| - 1|$ as a function of the total number of modes $|\boldsymbol{m}| = N$. Stars mark distances calculated from equation (3.2), and straight lines indicate the scaling N^{-2}. (After Rucklidge & Rucklidge (2003).)

3. Small divisors

Does the smallness of the small divisors arising from inverting $\mathcal{L}$ cause the sum (2.2) for $U(x, y, t)$ to diverge? To answer this question, the first stage is to derive a Diophantine-like condition for integer combinations of up to N of the 12 original vectors on the unit circle (such combinations arising at order N in the power series for U). It turns out that, for a given N, the smallest non-zero distance from the unit circle of a combination of N vectors is bounded above and below by a constant times N^{-2}.

An explanation of how this is derived begins with figure 5*a–c*, illustrating the locations of combinations of up to $N = 11$ and 15 wavevectors. Note how the density of points increases with N, and how the minimum distance between points and the unit circle goes down with N. Figure 5*d* shows results for the smallest non-zero distance from the unit circle as a function of the total number of vectors. The solid lines in figure 5*d* confirm numerically that the scaling for the distance to the unit

Table 1. *Continued-fraction approximations to* $r = \sqrt{3}$, *as a function of the order* l *of the truncation*

	l	0	1	2	3	4	5	6	7	8	9	10
$r = \sqrt{3}$	$\frac{p_l}{q_l}$	$\frac{1}{1}$	$\frac{2}{1}$	$\frac{5}{3}$	$\frac{7}{4}$	$\frac{19}{11}$	$\frac{26}{15}$	$\frac{71}{41}$	$\frac{97}{56}$	$\frac{265}{153}$	$\frac{362}{209}$	$\frac{989}{571}$

circle is of order N^{-2}, and the stars represent explicit combinations of wavevectors close to the unit circle, which were found as follows.

The vectors $\boldsymbol{k}_1, \boldsymbol{k}_2, \ldots, \boldsymbol{k}_{12}$ are labelled anticlockwise around the circle starting with $\boldsymbol{k}_1 = (1, 0)$, with $\boldsymbol{k}_{j+6} = -\boldsymbol{k}_j$ (figure 4*b*). Integer combinations of N of these vectors can be written as

$$\boldsymbol{k}_m = \sum_{j=1}^{12} m_j \boldsymbol{k}_j, \quad \text{with } |\boldsymbol{m}| = \sum_j |m_j| = N.$$

Including equal and opposite vectors, $\boldsymbol{k}_j$ and $\boldsymbol{k}_{j+6}$ will only increase N without coming any closer to the unit circle, so only $m_1, \ldots, m_6$ are considered, but these are allowed to be negative. With this restriction, the squared length of a vector $\boldsymbol{k}_m$ is

$$\begin{aligned}|\boldsymbol{k}_m|^2 = {} & m_1^2 + m_2^2 + m_3^2 + m_4^2 + m_5^2 + m_6^2 \\ & + m_1m_3 + m_2m_4 + m_3m_5 + m_4m_6 - m_5m_1 - m_6m_2 \\ & + \sqrt{3}(m_1m_2 + m_2m_3 + m_3m_4 + m_4m_5 + m_5m_6 - m_6m_1).\end{aligned}$$

This is of the form $|\boldsymbol{k}_m|^2 = 1 + p - rq$, where $r = \sqrt{3}$ is irrational and p and q are integers. If $p - rq$ is close to zero (that is, if r is well approximated by the rational p/q), then $|\boldsymbol{k}_m|^2$ can come close to 1 (but can only be exactly 1 if $p = q = 0$).

It is clear that the theory of continued-fraction approximations of irrationals will be useful here. The continued-fraction expression for $r = \sqrt{3}$ is

$$r = \sqrt{3} = 1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2 + \cdots}}}}.$$

That the value of this continued fraction is $\sqrt{3}$ can readily be shown by solving

$$r = 1 + \cfrac{1}{1 + \cfrac{1}{2 + (r - 1)}} = \frac{3 + 2r}{2 + r}.$$

This equation can be rearranged to give $r^2 - 3 = 0$, so $r = \sqrt{3}$ is the positive root. Since this irrational satisfies a quadratic equation with integer coefficients, $\sqrt{3}$ is called a quadratic irrational.

If the fraction is truncated after l terms, the successive fractions p_l/q_l that approximate $r = \sqrt{3}$ are given in table 1. The theory of continued fractions for quadratic

irrationals (Hardy & Wright 1960) shows that

$$\frac{K_1}{q_l^2} < \left|\frac{p_l}{q_l} - r\right| < \frac{K_2}{q_l^2} \quad \text{and} \quad \left|\frac{p_l}{q_l} - r\right| < \left|\frac{p}{q} - r\right|, \tag{3.1}$$

where K_1, K_2 are constants, and q is any integer satisfying $0 < q < q_l$. These inequalities mean that the truncated continued-fraction expansions p_l/q_l approximate r well, but not too well, as l becomes large, and that if p_l/q_l is the truncation of the continued-fraction approximation of an irrational r, no other fraction with a smaller denominator comes closer to r.

Apart from those vectors $\boldsymbol{k}_m$ that fall exactly on the unit circle (which would have $p = q = 0$), the relations in (3.1) can be used to show that $|\boldsymbol{k}_m|^2$ can approach 1 no faster than order N^{-2}:

$$||\boldsymbol{k}_m|^2 - 1| \geqslant \frac{K}{N^2},$$

where $|\boldsymbol{m}| = N$ and K is a constant—this lower limit is shown as a straight line in figure 5*d* (see Rucklidge & Rucklidge (2003) for more details).

The order N^{-2} rate of approach is indeed achieved by special combinations of vectors, which were found after a prolonged examination of the distances plotted in figure 5*d*. Choosing

$$\boldsymbol{k}_m = p_l\boldsymbol{k}_4 + (q_l - 1)\boldsymbol{k}_9 + (q_l + 1)\boldsymbol{k}_{11} = (1, p_l - \sqrt{3}q_l), \tag{3.2}$$

with $|\boldsymbol{m}| = N = p_l + 2q_l$ and $|\boldsymbol{k}_m|^2 - 1 = (p_l - \sqrt{3}q_l)^2$. As N (or, equivalently, l or q_l) increases, p_l and q_l are related by $p_l \sim \sqrt{3}q_l + \mathcal{O}(1/q_l)$, so $q_l = \mathcal{O}(N)$, and $|\boldsymbol{k}_m|^2 - 1 = \mathcal{O}(N^{-2})$. These particular choices of $\boldsymbol{k}_m$ are plotted on the graphs in figure 5*d* as stars. A little numerology suggests that $p_l + 2q_l = q_{l+2}$, so the values of N at which there is sudden drop in $|\boldsymbol{k}_m|^2 - 1$ are $N = 3, 4, 11, 15, \ldots$.

In summary, given an integer N, the vector $\boldsymbol{k}_m$ with $|\boldsymbol{m}| = N$ that comes closest to the unit circle (without being on the unit circle) satisfies

$$\frac{K}{N^2} \leqslant ||\boldsymbol{k}_m|^2 - 1| \leqslant \frac{K'}{N^2},$$

for constants K and K', for 12 equally spaced original vectors. The numerical evidence in figure 5*d* suggests values $K = 0.56$ and $K' = 4.34$.

4. The question of convergence

The results of the previous two sections imply that when $\boldsymbol{k}_m$ is close to the unit circle, $\mathcal{L}^{-1}(\mathrm{e}^{\mathrm{i}\boldsymbol{k}_m\cdot\boldsymbol{x}})$ can be as large as a constant times $N^4\mathrm{e}^{\mathrm{i}\boldsymbol{k}_m\cdot\boldsymbol{x}}$, with $N = |\boldsymbol{m}|$. This is so large that it clearly could lead to divergence of the power series (2.2) for U, particularly when nonlinear interactions of these large contributions are taken into account. This problem of small divisors is not just a feature of the particular Swift–Hohenberg equation (2.1) used for illustration here, but arises in any calculation of the properties of quasi-patterns based on perturbation theory.

This failure of convergence can be illustrated dramatically in the particular Swift–Hohenberg example by carrying out the perturbation theory calculation to high order

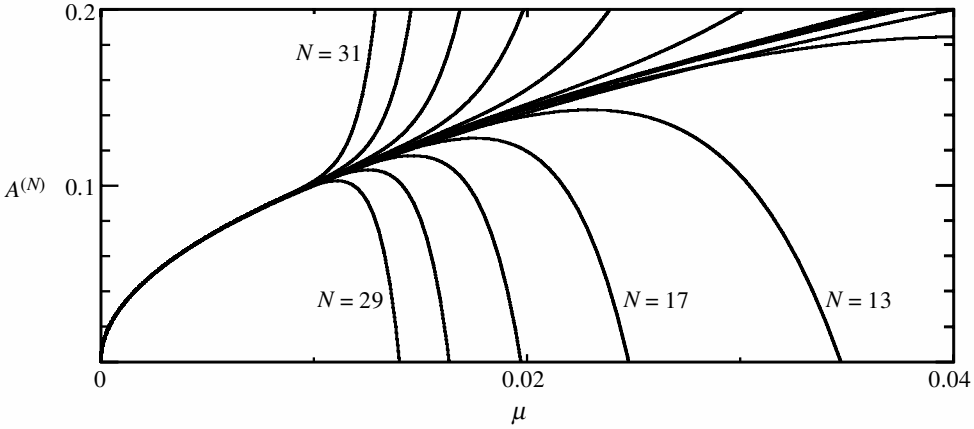

Figure 6. Amplitude $A^{(N)}$ as a function of μ, for different levels of truncation $N = 1, \ldots, 31$. Increasing the order of truncation leads to graphs of $A^{(N)}$ that diverge for μ closer and closer to zero as N becomes larger. The amplitude has been scaled to remove a factor of $1/\sqrt{33}$. (From Rucklidge & Rucklidge (2003).)

(33rd order in this case). If the series (2.2) is truncated to include powers of ϵ up to and including $N + 2$, the resulting expression for $U^{(N)}$ is of the form

$$U^{(N)} = A^{(N)} \sum_{j=1}^{12} \mathrm{e}^{\mathrm{i}\boldsymbol{k}_j \cdot \boldsymbol{x}} + \text{other modes},$$

so $A^{(1)} = \sqrt{\mu/33}$, from (2.6). The amplitude $A^{(N)}$ of the basic quasi-pattern is shown as a function of μ in figure 6, for $N = 1, \ldots, 31$. In this calculation, only modes with wavenumbers up to $\sqrt{5}$ were kept, to keep the total number of modes within manageable limits. Even so, there were more than 15 000 modes generated at the highest order—without this truncation, there would have been almost 2 000 000. Since the modes that were dropped from the calculation were the most heavily damped, their contribution to the total amplitude was quite small (of the order of 1%), and restricting the number of modes in this way had no effect on how close combinations of wavevectors could get to the unit circle.

It is clear in figure 6 that, at each level of truncation N, the graph of $A^{(N)}$ against μ diverges at a value of μ that decreases as N becomes larger. The value of μ at which the sum up to order N diverges is related to the smallest distance from the unit circle achieved by combinations of N of the 12 original wavevectors. Since this distance goes to zero as N increases, the sum $A^{(N)}$ will continue to diverge closer and closer to $\mu = 0$. In contrast, the equivalent calculation for spatially periodic patterns has a non-zero radius of convergence (Rucklidge & Rucklidge 2003).

5. Discussion and speculation

The main conclusion of the calculation is that even if perturbation theory does generate a convergent series approximation to the quasi-pattern for small enough μ, the series certainly diverges if the parameter μ is bigger than about 0.01. It might be possible that the series does converge for smaller μ, though there is a strong argument that this is not the case. However, even if the series does diverge for all

non-zero μ, a low-order truncation may still give a useful asymptotic approximation of the quasi-pattern, assuming that the equations do have a quasi-pattern solution. It is on this basis that other researchers have proceeded.

There are two related issues at stake. First, existence: do pattern-forming PDEs (like the two-dimensional Swift–Hohenberg equation) have quasi-pattern solutions? A more general formulation of this question, using the Swift–Hohenberg equation as an example, becomes apparent by setting $\mu = \epsilon^2$ in (2.1), scaling U by ϵ and seeking a steady solution. The resulting equation can be written as

$$\mathcal{L}(U) = \epsilon^2(-U + U^3),$$

which incidentally demonstrates that this is not a singularly perturbed problem. When $\epsilon = 0$, *any* linear combination of waves with wavevectors on the unit circle solves this equation. The question is: which of these solutions persist to small but positive ϵ? Current theory can so far only answer this question for those solutions that are spatially periodic. The limits on the rate of approach of wavevectors to the unit circle will play a central role in an eventual existence theory for quasi-patterns.

The second issue is, given the small-divisor problem, are there methods that yield useful approximations to quasi-pattern solutions? Standard perturbation theory does not converge sufficiently rapidly (or slowly) to provide an answer unequivocally one way or the other. However, if quasi-pattern solutions exist, then the series ought to provide an asymptotic approximation to those solutions. Nonetheless, this approach will be left with difficulties, such as the undetermined phason modes, and so should not be regarded as a reliable way of computing properties of quasi-patterns.

What is needed is a method that converges more rapidly. Each order in the standard theory gains a factor of ϵ^2 as well as large factors from any small divisors that arise. There are other methods, developed for proofs of KAM theory, that converge more rapidly, and these may be required for a rigorous treatment of quasi-patterns as well. The difference between the KAM situation and that of quasi-patterns is that in the KAM case, the solutions of interest are quasi-periodic in only one dimension (time), while in the second, quasi-patterns are quasi-periodic in two space directions.

The problems that confront a proper mathematical theory of quasi-patterns are related to the difficulties that have arisen in other aspects of pattern formation: long-wave modulation of two-dimensional patterns, and the coexistence of spirals and spiral defect chaos with straight rolls. At the heart of this is the question of why, given that there is always a *range* of excited wavenumbers, spatially periodic patterns, characterized by a *single* wavenumber, are often observed in spatially extended (effectively infinite) domains.

This last question has been answered rigorously in the context of one-dimensional patterns (stripes) in an infinite domain. Just above the onset of pattern formation, a range of wavenumbers close to $|\boldsymbol{k}| = 1$ are linearly unstable. Periodic patterns with wavenumbers within a certain interval, which can be computed explicitly, are stable to all perturbations, and all small-amplitude initial conditions will eventually settle down to one of these stable periodic patterns (Melbourne 1999). A periodic pattern with wavenumber outside this interval is unstable to long-wave modulations that have the effect of altering the wavenumber to bring it into the stable interval. The equation that governs this process is known as the Ginzburg–Landau equation.

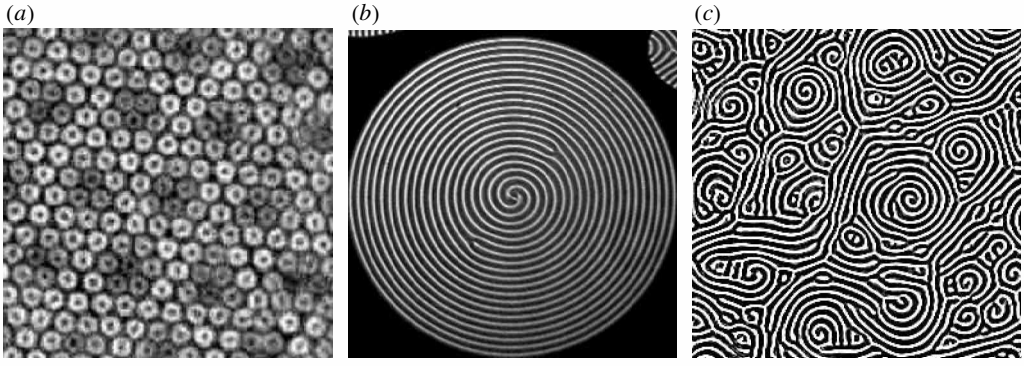

Figure 7. Other experimentally observed examples of pattern formation. (*a*) Spatially modulated hexagons in a two-frequency forced Faraday wave experiment. (Reproduced with permission from Arbell & Fineberg (2002). Copyright (2002) American Physical Society.) (*b*) Giant two-armed spiral in Rayleigh–Bénard convection. (Reproduced with permission from Plapp *et al*. (1998). Copyright (1998) American Physical Society.) (*c*) Spiral defect chaos in Rayleigh–Bénard convection. (Reproduced with permission from Morris *et al*. (1993). Copyright (1993) American Physical Society.)

The difficulty in an infinite two-dimensional domain is that the linear stability problem is highly degenerate: patterns can form with any orientation, since all modes close to the circle $|\boldsymbol{k}| = 1$ are linearly unstable. This leads to difficulties when trying to justify the use of two-dimensional versions of the Ginzburg–Landau equation, which only allow for a selection of these unstable modes. The modes that are not included in this approach could nonetheless be excited by nonlinear interactions between those that are included. It may be that the Ginzburg–Landau approach does provide a qualitatively correct description of numerical and laboratory experiments, but there is as yet no satisfactory justification for this. As a result, long wave modulations of two-dimensional patterns (as in, for example, figure 7*a*) remain beyond reach. In particular, long-range changes in the orientation of a pattern are not captured by any current theory.

It seems unlikely, though perhaps possible, that a theory of two-dimensional patterns based on amplitudes of individual modes will be able to satisfy the two very different requirements of mathematical rigour and of having sufficient flexibility to allow patterns of different orientations, mixtures of (for example) squares and hexagons, quasi-patterns, defects, and so on. Allowing for all these possibilities would mean that amplitudes of all modes within a band around the circle $|\boldsymbol{k}| = 1$ would have to be included, and this would make the convergence problems discussed above much worse. There are patterns that do have modes with wavevectors of all orientations: giant spirals, for example (see figure 7*b*). Even more puzzling and striking is spiral defect chaos (figure 7*c*), which has been observed in convection experiments with low-viscosity fluids. This pattern is made up of fragments of rolls and spirals, of size intermediate between the domain size and the scale of the pattern. This disordered state can occur close to the onset of convection and at the same parameters, the straight roll state is also stable. It is a challenge to see how amplitude-based models might be able to explain this phenomenon.

Of course, solving the PDEs for the particular problem at hand always remains a possibility, but an approach based on amplitude equations is still the most promis-

ing way of understanding pattern formation as a universal, problem-independent, phenomenon.

I am grateful to many people who have helped shape these ideas, in one way or another, over a period of several years. This research is supported by the Engineering and Physical Sciences Research Council.

References

Arbell, H. & Fineberg, J. 2002 Pattern formation in two-frequency forced parametric waves. *Phys. Rev.* E **65**, 036224.

Aubry, S. 1983 The twist map, the extended Frenkel–Kontorova model and the devil's staircase. *Physica* D **7**, 240–258.

Cakmur, R. V., Egolf, D. A., Plapp, B. B. & Bodenschatz, E. 1997 Bistability and competition of spatiotemporally chaotic and fixed point attractors in Rayleigh–Bénard convection. *Phys. Rev. Lett.* **79**, 1853–1856.

Echebarria, B. & Riecke, H. 2001 Sideband instabilities and defects of quasipatterns. *Physica* D **158**, 45–68.

Edwards, W. S. & Fauve, S. 1994 Patterns and quasi-patterns in the Faraday experiment. *J. Fluid Mech.* **278**, 123–148.

Golubitsky, M. & Stewart, I. 2002 *The symmetry perspective: from equilibrium to chaos in phase space and physical space*. Basel: Birkhäuser.

Hardy, G. H. & Wright, E. M. 1960 *An introduction to the theory of numbers*, 4th edn. Oxford: Clarendon.

Janot, C. 1994 *Quasicrystals: a primer*, 2nd edn. Oxford: Clarendon.

Levine, D. & Steinhardt, P. J. 1984 Quasicrystals: a new class of ordered structures. *Phys. Rev. Lett.* **53**, 2477–2480.

Melbourne, I. 1999 Steady-state bifurcation with Euclidean symmetry. *Trans. Am. Math. Soc.* **351**, 1575–1603.

Morris, S. W., Bodenschatz, E., Cannell, D. S. & Ahlers, G. 1993 Spiral defect chaos in large aspect ratio Rayleigh–Bénard convection. *Phys. Rev. Lett.* **71**, 2026–2029.

Moser, J. 1973 *Stable and random motions in dynamical systems*. Princeton University Press.

Plapp, B. B., Egolf, D. A., Bodenschatz, E. & Pesch, W. 1998 Dynamics and selection of giant spirals in Rayleigh–Bénard convection. *Phys. Rev. Lett.* **81**, 5334–5337.

Rucklidge, A. M. & Rucklidge, W. J. 2003 Convergence properties of the 8, 10 and 12 mode representations of quasipatterns. *Physica* D **178**, 62–82.

Swift, J. & Hohenberg, P. C. 1977 Hydrodynamic fluctuations at the convective instability. *Phys. Rev.* A **15**, 319–328.

AUTHOR PROFILE

A. M. Rucklidge

Alastair Rucklidge was born in Oxford, England, and grew up in Toronto, Canada, graduating from the University of Toronto with first class honours in Engineering Science in 1986. After earning an SM in Physics at the Massachusetts Institute of Technology, and a spell at Hatch Associates, a consulting engineering firm based in Toronto, he obtained his PhD in Applied Mathematics from the University of Cambridge in 1992. He held a research fellowship from Peterhouse, Cambridge, the Sir Norman Lockyer Fellowship from the Royal Astronomical Society, and an EPSRC Advanced Research Fellowship from 1998–2003. Now aged 38, he has been a lecturer in Applied Mathematics at the University of Leeds since 2000. His scientific interests include pattern formation, nonlinear dynamics and astrophysical fluid dynamics, and he is a keen outdoor photographer.