Contents

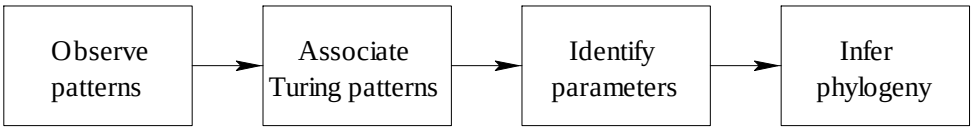
Acknowledgements

I wish to especially thank Prof. Jotun Hein and Dr Andrea Rocco, from the Oxford Centre for Gene Function, for their ongoing supervision and guidance throughout this project. Their comments and suggestions were very much appreciated. I would also like to thank Prof. Nick Trefethen, from the Oxford University Computing Laboratory, for his help with coding in MATLAB to carry out the numerical simulations in this project. Finally, I am grateful to everyone who assisted me throughout this project.

Introduction

The problem of explaining the rich spectrum of different patterns exhibited by animals has a long history. In a controversial paper [1] Turing proposed a Reaction-Diffusion (RD) model to explain at a macroscopic level the process of pattern formation by the occurrence of what he called a diffusion-driven instability. The simplest and most typical Turing system consists of two chemical species, usually referred to as *activator* and *inhibitor*, reacting in such a way that their steady state is stable to small perturbations in the absence of diffusion but, under certain conditions, can be driven unstable in the presence of diffusion. In Turing's original model, he assumed that the chemicals reacted with each other in a linear way and this meant that if the steady state became unstable then the chemical concentrations would grow exponentially. From a biological point of view, this is clearly unrealistic. A number of models have been proposed since then, where the chemicals react with each other in a much more complicated way by means of nonlinear terms. By introducing these nonlinear terms, it is found that when the uniform steady state is driven unstable it is bounded and may evolve to a spatially, non-uniform, steady state, i.e. a spatial pattern. It is supposed that if one of the chemical concentrations goes above or below some threshold value, usually taken to be the steady state value, then the cells differentiate accordingly and hence in this way we may observe a spatial pattern. Even though these proposed models have different chemical motivation and derivation, most of them are capable of generating spatial patterns. In [2] Maini *et al* classify the different ways in which the models are derived: (i) phenomenologically, (ii) to model a hypothetical reaction and (iii) empirically. Several examples include the Gierer-Meinhardt model [3], the Gray-Scott model [4], the Lengyel and Epstein model [5], the Schnakenberg model [6] and the Thomas model [7]. In this project we have focused on the Schnakenberg model, which is an example of (ii); it is based on a cubic autocatalytic process. The main reason we chose this model is because it has the fewest parameters and therefore the analysis will be the simplest. Even so, as will be seen, a wide range of patterns are easily generated.

The next stage of this project will be to introduce some evolutionary process in our RD model by allowing the parameters in the model to evolve. It turns out that this is not too difficult to accomplish if we assume that our parameters evolve by Brownian motion Once we have defined our evolutionary model it should, in principle, be possible to construct a maximum likelihood phylogeny. Here is a schematic illustration of how we intend to accomplish this:



Firstly, we need to observe a set of patterns. Rather than looking at real data, we

will produce the patterns ourselves by simulating our evolutionary model. Then we shall associate these patterns to specific Turing patterns that our model generates. With the analysis we intend to carry out on our RD model, we should then be able to identify the values of the parameters in our model which produce these Turing patterns. By assuming that our parameters (or characters) have evolved by Brownian motion, we can then apply existing methods to calculate a maximum likelihood phylogeny. There is much literature on the estimation of evolutionary trees from quantitative characters (which can be found, for example, in [8], [9] or [10]) and the key assumption of all these methods is that each character evolves independently according to Brownian motion.

In chapter one, we examine the Schnakenberg model in detail, and ascertain what effects the parameters have on pattern selection. Throughout this examination, our analysis will be verified with numerical simulations, which, in general, play a crucial part in our understanding of the pattern formation process.

Once the effect of the parameters has been adequately understood, in the next chapter we seek to develop a stochastic evolutionary model for the evolution of our Turing system that accounts for evolutionary randomness directly at the phenotypic level. This model will only apply to the phenotypic level and will make no statement about the underlying genetic mechanism. Here, we assume that patterns are produced by means of a Turing mechanism. To achieve evolution between the realized phenotypes (e.g. between spotted patterns and striped patterns) we will let our parameters evolve in time by simple Brownian motion; as the parameters vary in this way we will see that transitions are made between different patterns. In fact, for the sake of simplicity, we will only allow one of our parameters to evolve in this way. Thus, this chapter will be devoted to how to let the parameters evolve by Brownian motion in our RD model and to which parameter we should choose to evolve in this way.

With the stochastic dynamics defined on the parameters in this way, this should allow us to run along the branches of the phylogenetic tree. It is to this phylogenetic tree that we turn our attention in the concluding chapter. Focusing on the three-species case, we try to construct a maximum likelihood phylogeny for a given tree topology, supposing that we are given the patterns at the tips of the tree. This can be achieved by firstly inferring the values of our parameters from the patterns at the tips, and secondly, assuming that these parameters have evolved by Brownian motion, by using a restricted maximum likelihood (REML) estimate. By using an REML estimate, we can obtain estimates for the branch lengths of the given tree. The REML approach to this type of problem was first introduced by Felsenstein in [9]. As a final step, we simulate the evolution of one of the parameters in our RD model and produce a three-species phylogeny. We will then take this phylogeny and try to infer the branch lengths of the tree (assuming that we know the tree topology).

Chapter 1
Analysis of the Model

1.1 The Model

We begin by giving a brief derivation of our RD model which is based on the hypothetical Schnakenberg reaction [6]. It is possible to derive the following equations for reaction diffusion mechanisms (see [11], chapter 9, for such a derivation):

$$\frac{\partial \mathbf{x}}{\partial t} = \mathbf{f}\left(\mathbf{x}\right) + D\nabla^2\mathbf{x},$$

where $\mathbf{x}$ is the vector of morphogen concentrations, $\mathbf{f}$ represents the reaction kinetics and D is a diagonal matrix of (positive) diffusion coefficients. Since we are using the Schnakenberg model, we will only be concerned with two morphogen concentrations, $X(\mathbf{r}, t)$ and $Y(\mathbf{r}, t)$, in which case the system becomes:

$$\frac{\partial X}{\partial t} = F(X, Y) + D_A\nabla^2 X,$$

$$\frac{\partial Y}{\partial t} = G(X, Y) + D_B\nabla^2 Y, \tag{1.1}$$

where F and G are the nonlinear kinetics. The Schnakenberg reaction, which is one of the simplest but chemically plausible reactions for two chemical concentrations, is based on the following reactions:

$$X \underset{K_4}{\overset{K_1}{\rightleftharpoons}} A, \quad B \overset{K_2}{\rightarrow} Y, \quad 2X + Y \overset{K_3}{\rightarrow} 3X. \tag{1.2}$$

where X and Y are the two chemical species, A and B are another two chemical species, assumed to be maintained at a constant concentration, and the $K_i's$ are the rates of reactions. If we write $X^* = [X]$, etc., to denote the concentration of these chemicals then, using the Law of Mass Action, this leads to the kinetics (on dropping the asterisks):

$$F(X, Y) = k_1 - k_2X + k_3X^2Y \ ,$$

$$G(X, Y) = k_4 - k_3X^2Y \ , \tag{1.3}$$

where $k_1 = K_4A$, $k_2 = K_1$, $k_3 = K_3$ and $k_4 = K_2B$. Before we go any further, we must nondimensionalise the reaction diffusion equations

with the reaction kinetics above. Let L be a typical length scale and set

$$X = (\frac{k_2}{k_3})^{1/2}u, \quad Y = (\frac{k_2}{k_3})^{1/2}v, \quad t = \frac{L^2}{D_A}t^*, \quad \mathbf{x} = L\mathbf{x}^*,$$
$$a = \frac{k_1}{k_2}(\frac{k_3}{k_2})^{1/2}, \quad b = \frac{k_4}{k_2}(\frac{k_3}{k_2})^{1/2}, \quad d = \frac{D_B}{D_A}, \quad \gamma = \frac{L^2k_2}{D_A}. \tag{1.4}$$

Dropping the asterisks for algebraic convenience, the dimensionless reaction diffusion system is then

$$u_t = \gamma(a - u + u^2v) + \nabla^2 u,$$
$$v_t = \gamma(b - u^2v) + d\nabla^2 v. \tag{1.5}$$

Which we now also write, for future convenience, as

$$u_t = \gamma f(u, v) + \nabla^2 u,$$
$$v_t = \gamma g(u, v) + d\nabla^2 v. \tag{1.6}$$

where $f(u, v) = (a - u + u^2v)$ and $g(u, v) = (b - u^2v)$.

1.2 Linear Stability Analysis

For any steady state to be driven unstable in the presence of diffusion, there are certain conditions which must be satisfied. We now find out what these conditions are; this discussion follows [11], chapter 14, pp. 380-384. To formulate the problem mathematically we need to specify initial conditions and boundary conditions. We take these to be given initial conditions (which will be small perturbations about the steady state) and zero flux boundary conditions. Therefore the mathematical problem is

$$u_t = \gamma f(u, v) + \nabla^2 u, \qquad v_t = \gamma g(u, v) + d\nabla^2 v,$$
$$u(\mathbf{r}, 0), v(\mathbf{r}, 0) \qquad \text{given}, \qquad (\mathbf{n} \cdot \nabla) \begin{pmatrix} u \\ v \end{pmatrix} = 0 \qquad \text{for } \mathbf{r} \text{ and } B, \tag{1.7}$$

where ∂B is the closed boundary of the domain B in which we are solving the system, and $\mathbf{n}$ is the unit outward normal to ∂B. In our case, we are only concerned with two dimensional patterns on animals. Therefore, we will be working in two space dimensions and our two dimensional domain will be a square and have a fixed size for simplicity.

In the absence of diffusion any homogenous (positive) steady state must be linearly stable. It is only in the presence of diffusion that we want this steady state to be driven unstable. With no diffusion terms, we have

$$u_t = \gamma f(u, v), \qquad v_t = \gamma g(u, v). \tag{1.8}$$

It can be derived (again, see [11], chapter 14) that linear stability is guaranteed if

$$f_u + g_v < 0, \qquad f_u g_v - f_v g_u > 0, \tag{1.9}$$

where the partial derivatives of f and g here, and from now on, are taken to be evaluated at the steady state (u_s, v_s).

Now let us find the steady states of our diffusionless system (1.8) by putting $f(u, v) = 0 = g(u, v)$.

$$f(u,v) = 0 \Rightarrow v = -\frac{a}{u^2} + \frac{1}{u}.$$
$$g(u,v) = 0 \Rightarrow v = \frac{b}{u^2}.$$

Therefore,

$$-\frac{a}{u^2} + \frac{1}{u} = \frac{b}{u^2}.$$

Hence, the unique steady state is:

$$u = u_s = a + b, \qquad v = v_s = \frac{b}{(a+b)^2}. \tag{1.10}$$

Thus, it is clear that the linear stability conditions above will put restrictions on the parameter ranges of a and b. The exact restrictions will be explicitly shown later. We are also only concerned with positive solutions for the steady states and therefore we have the following

$$b > 0, \qquad a + b > 0. \tag{1.11}$$

Firstly, we need to look at the full reaction diffusion system and linearise about the steady state again. This will give us further conditions on the parameters. Set

$$\mathbf{w} = \begin{pmatrix} u - u_s \\ v - v_s \end{pmatrix}. \tag{1.12}$$

Then linearising our system about the steady state which is now $\mathbf{w} = \mathbf{0}$, for $|\mathbf{w}|$ small, we get

$$\mathbf{w}_t = \gamma A\mathbf{w} + D\nabla^2\mathbf{w}, \qquad A = \begin{pmatrix} f_u & f_v \\ g_u & g_v \end{pmatrix}, \qquad D = \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}. \tag{1.13}$$

We can solve this system of equations with zero flux boundary conditions by first defining $\mathbf{W}(\mathbf{r})$ to be the time independent solution to the eigenvalue problem

$$\nabla^2\mathbf{W} + k^2\mathbf{W} = 0, \qquad (\mathbf{n}\cdot\nabla)\mathbf{W} = 0 \qquad \text{for } \mathbf{r} \text{ on } \partial B, \tag{1.14}$$

where k is the eigenvalue. In this context, it is a measure of the wave-like pattern and will therefore be referred to as the *wavenumber*. There is a discrete set of wavenumbers when the domain is finite, which ours is. Let $\mathbf{W}_k(\mathbf{r})$ be the eigenfunction corresponding to the wavenumber k. Because the problem is linear, we can now solve it by writing

$$\mathbf{w}(\mathbf{r}, t) = \sum_k c_k e^{\lambda t}\mathbf{W}_k(\mathbf{r}), \tag{1.15}$$

where the constants c_k can be found by a Fourier expansion of the given initial conditions in terms of $\mathbf{W}_k(\mathbf{r})$. So depending on the sign of the eigenvalue λ, the eigenfunctions will either tend to zero over time, or grow exponentially, implying instability of the steady state. We now find λ and show that it is actually a function of k. To do this we start by substituting (1.15) into (1.13). Therefore, for each k we have, on cancelling the $e^{\lambda t}$ terms and using (1.14), the following:

$$\lambda \mathbf{W}_k = \gamma A \mathbf{W}_k + D\nabla^2 \mathbf{W}_k$$

$$= \gamma A \mathbf{W}_k - Dk^2 \mathbf{W}_k$$

This is now of the form $B\mathbf{x} = \lambda \mathbf{x}$, where $B = \gamma A - Dk^2$ is a matrix, $\mathbf{x} = \mathbf{W}_k$ is an eigenvector and λ is the corresponding eigenvalue. Hence the λ are roots of the characteristic polynomial

$$\left|\gamma A - Dk^2 - \lambda I\right| = 0.$$

If we evaluate this determinant with A and D given from (1.13) we get

$$\lambda^2 + \lambda[k^2(1+d) - \gamma(f_u + g_v)] + h(k^2) = 0, \tag{1.16}$$

$$h(k^2) = dk^4 - \gamma(df_u + g_v)k^2 + \gamma^2 |A|, \tag{1.17}$$

where $|A| = f_u g_v - f_v g_u$. We require $\operatorname{Re} \lambda(k) > 0$ for some $k \neq 0$ for exponential growth in (1.15) and hence for the steady state to now be unstable (note that if $k = 0$, then we would be back to the diffusionless system given by (1.8)). The coefficient of λ in (1.16) is positive since we already have the restriction $f_u + g_v < 0$ from (1.9) for linear stability when there is no diffusion present. Therefore, it is only possible to have $\operatorname{Re} \lambda(k) > 0$ if $h(k^2) < 0$ for some k. Since $d > 0$ and we require $|A| > 0$ from (1.9), the only way that $h(k^2) < 0$ is if the constant coefficient of $h(k^2)$ in (1.17) is negative. Thus, we now have the additional restriction

$$df_u + g_v > 0. \tag{1.18}$$

Because $f_u + g_v < 0$, this means that f_u and g_v must have opposite signs. It also implies that $d \neq 1$, for otherwise it would be impossible to satisfy both inequalities. Remembering that d is the ratio of the two diffusion coefficients, this means that $D_A \neq D_B$ and therefore the two chemicals cannot be diffusing at the same speed.

Inequality (1.18) is a necessary condition for $h(k^2) < 0$, but it is not quite enough. For a sufficient condition we require the discriminant of the quadratic equation given by (1.17) to be positive so that the minimum of $h(k^2)$, for some k, is negative:

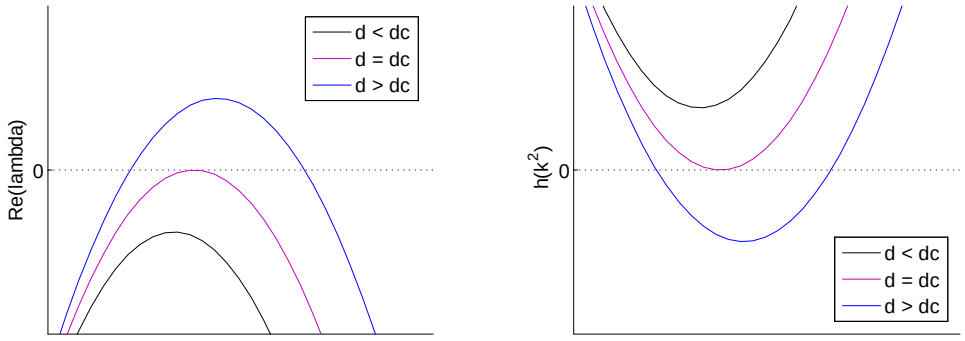
$$[\gamma(df_u + g_v)]^2 - 4\gamma^2 d |A| > 0$$

$$\Rightarrow (df_u + g_v)^2 > 4d |A| \tag{1.19}$$

This is the final condition for diffusion driven instability. Notice that $4d\,|A| > 0$, so the last condition encapsulates condition (1.18) as well. When the minimum of $h(k^2)$ is zero, we have a bifurcation of stability and this happens when the discriminant is zero. The partial derivatives of f and g are functions of the kinetics parameters, and so if they are fixed we can define a critical diffusion coefficient ratio d_c as the appropriate root of

$$\begin{aligned} &[\gamma(d_c f_u + g_v)]^2 - 4\gamma^2 d_c\,|A| = 0 \\ \Rightarrow\ &(d_c f_u + g_v)^2 - 4d_c(f_u g_v - f_v g_u) = 0 \\ \Rightarrow\ &d_c^2 f_u^2 + 2d_c(2f_v g_u - f_u g_v) + g_v^2 = 0. \end{aligned} \tag{1.20}$$

For $d > d_c$ there are values of k for which $h(k^2) < 0$ and hence for which $\mathrm{Re}\,\lambda > 0$. Figure 1.1 illustrates this.



(a) (b)

Figure 1.1(a), (b). (a) Plot of $\mathrm{Re}(\lambda(k^2))$ using the largest of the eigenvalues from (1.16). We see that when the diffusion coefficient ratio d passes beyond the critical value d_c there is a range of values of k for which $Re(\lambda) > 0$ and hence the steady state can be driven unstable. (b) Plot of $h(k^2)$ defined by (1.17). When $d > dc$ there are values of k for which $h(k^2) < 0$.

1.3 An Illustration of Diffusion Driven Instability

Now let us illustrate these diffusion-driven instability conditions and the critical diffusion coefficient with a concrete example using our system (1.5) with the steady state given by

(1.10). We write the conditions, the kinetics and the steady state again here for the sake of reference:

$$f_u + g_v < 0, \qquad f_u g_v - f_v g_u > 0, \qquad (df_u + g_v)^2 > 4d\,|A|\,, \tag{1.21}$$

$$f(u,v) = (a - u + u^2 v), \quad g(u,v) = (b - u^2 v), \tag{1.22}$$

$$u = u_s = a + b, \qquad v = v_s = \frac{b}{(a+b)^2}, \qquad b > 0, \qquad a + b > 0. \tag{1.23}$$

Now,

$$f_u = \frac{b-a}{b+a}, \quad f_v = (a+b)^2, \quad g_u = -\frac{2b}{a+b}, \quad g_v = -(a+b)^2, \tag{1.24}$$

so using (1.21),

$$f_u + g_v < 0 \qquad \Rightarrow \qquad (a+b)^3 > (b-a), \tag{1.25}$$

$$f_u g_v - f_v g_u > 0 \qquad \Rightarrow \qquad (a+b)^2 > 0, \quad \text{automatically satisfied,} \tag{1.26}$$

$$(df_u + g_v)^2 > 4d\,|A| \qquad \Rightarrow \qquad [d(b-a) - (a+b)^3]^2 > 4d(a+b)^4. \tag{1.27}$$

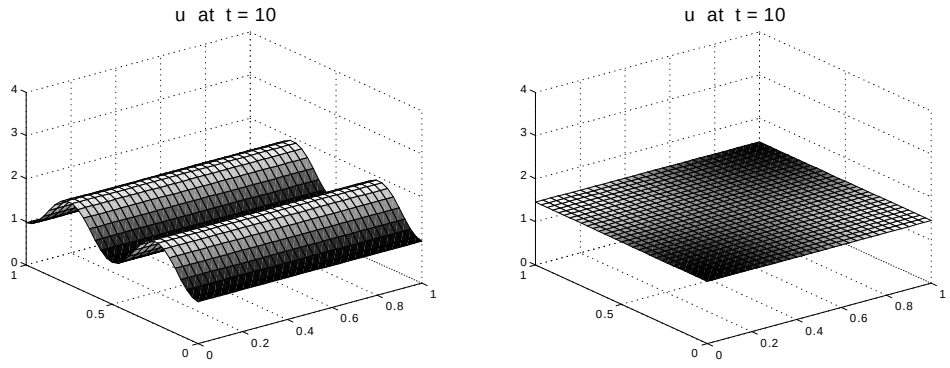
Remember that f_u and g_v must have opposite signs and since g_v is always negative f_u must be positive and hence

$$b > a. \tag{1.28}$$

If we choose kinetics parameters as follows,

$$a = 0.05, \qquad b = 1.4,$$

so that the inequalities (1.23), (1.25) and (1.28) are all satisfied, it only remains to find the appropriate d so that inequality (1.27) is also satisfied. This can be done via (1.20), which gives us the value for the critical diffusion coefficient $d_c = 13.85$. For $d > 13.85$, $\mathrm{Re}\,\lambda > 0$ and inequality (1.27) is indeed satisfied as one can easily varify. Hence if $d > 13.85$, then our values for a, b and d lie inside the so-called Turing space, the domain in (a, b, d) parameter space within which we can expect to observe diffusion-driven instability (as long as γ is large enough, see section 1.4), as the figures below indicate. The figures below are from numerical simulations of the full nonlinear RD system (1.7) with our kinetics on a square domain. The equations were solved on a grid with 30×30 sites by a simple Euler method. A very small time step was used to ensure numerical stability and the calculation time was taken so that the patterns were stable. The calculations were achieved by using MATLAB and the actual code which we used can be found in the Appendix. For further methods on solving partial differential equations numerically, you may refer to [12]. All of our numerical simulations were carried out in this way.



(a) (b)

Figure 1.2(a), (b). Numerical simulations of the full nonlinear reaction diffusion system given by (1.7) with kinetics (1.22) and with the initial conditions as small perturbations from the steady state. Both simulations have parameter values $a = 0.05$, $b = 1.4$ and $\gamma = 400$. (a) d is chosen as $d = d_c + 1 = 14.85$ so that $d > d_c$. We can see that a spatial pattern forms. (b) d is now chosen as $d = 12.85$ so that it is below the critical diffusion value $d_c = 13.85$. Now we see that u settles back down to the spatially homogenous steady state.

Finding a and b so that inequalities (1.23), (1.25) and (1.28) hold is not difficult. For example, any $b > 1$, $b > a > 0$, will work, since, for this particular choice, (1.23) and (1.28) are clearly satisfied and $(a + b)^3 > b > (b - a)$ so (1.25) is satisfied too. There are of course other choices that can be made. (For different kinetics, it may not be possible to satisfy (1.23), (1.25) and (1.28) so easily). However, different choices will result in drastically different critical diffusion values,

$$a = 0.01, \qquad b = 1, \qquad \Rightarrow \qquad d_c = 6.2$$
$$a = 0.5, \qquad b = 1.5, \qquad \Rightarrow \qquad d_c = 79.2$$
$$a = 1, \qquad b = 10, \qquad \Rightarrow \qquad d_c = 996.8$$

and it may, for example, be unreasonable to have the two chemicals diffusing at such different speeds. It is possible to determine the parameter space where spatial patterns can be generated for our kinetics analytically, and this has been done (for example, see [11]). However, the algebra is messy, and in general this is only possible for simple kinetics, such as ours.

1.4 The Range of Unstable Modes

There remains one parameter that has not yet been discussed, which plays a crucial role in the pattern formation process: γ. To see how γ comes into the picture, we must return to the wavenumbers k and the function $h(k^2) = dk^4 - \gamma(df_u + g_v)k^2 + \gamma^2 |A|$, given by (1.17). When d is greater than d_c, we know from earlier that $h(k^2) < 0$ and hence we can determine the range of unstable wavenumbers $k_1^2 < k^2 < k_2^2$ for which $h < 0$ by calculating the roots k_1^2 and k_2^2 of $h = 0$. Explicitly,

$$k_1^2 = \gamma\{(df_u + g_v) - [(df_u + g_v)^2 - 4d\,|A|]^{\frac{1}{2}}\} < k^2$$

$$< \gamma\{(df_u + g_v) + [(df_u + g_v)^2 - 4d\,|A|]^{\frac{1}{2}}\} = k_2^2. \tag{1.29}$$

Therefore

$$k_1^2 = \gamma A(a, b, d) < k^2 < \gamma B(a, b, d) = k_2^2, \tag{1.30}$$

since the partial derivatives are functions only of the kinetics parameters a and b. The effect of γ is quite clear now: if we increase or decrease γ, we increase or decrease respectively the range of possible unstable wavenumbers. Since there is a discrete set of wavenumbers in a finite domain, (1.30) implies that if a,b, and d are fixed (and lie inside the Turing space) and γ is sufficiently small, then there may be no wavenumbers k which lie in this range and hence no eigenfunctions will be driven unstable. This means that even if the parameters a,b, and d lie inside the Turing space, it is possible that all the eigenfunctions in (1.15) will tend to zero exponentially and the steady state will remain stable in the presence of diffusion. On the other hand, when γ is large enough, several wavenumbers will lie inside the range given by (1.30) and hence, for large t, (1.15) implies that

$$\mathbf{w}(\mathbf{r}, t) = \sum_{k_1}^{k_2} c_k \exp[\lambda(k^2)t]\mathbf{W}_k(\mathbf{r}), \tag{1.31}$$

since all the other modes tend to zero exponentially. Clearly there will be a fastest growing mode in this summation for a wavenumber k_m where $\mathrm{Re}\,\lambda(k_m^2)$ takes the largest possible value. This wavenumber can be located by first finding the value of k where $\mathrm{Re}\,\lambda(k^2)$ attains its maximum (or equivalently, where $h(k^2)$ attains its minimum), and then finding the wavenumber k_m closest to this value of k. Although these eigenfunctions, or spatial patterns, in (1.31) are linearly unstable and grow exponentially with time, what actually happens is that these eigenfunctions are eventually bounded by the nonlinear terms in the reaction diffusion system found in the kinetics, given by the functions f and g. Hence a steady state spatially non-uniform solution (i.e. a spatial pattern) emerges, as indicated in figure 1.2. We expect the fastest growing mode to dominate and hence give the realized spatial pattern but this is not always the case as we shall see below.

We now return to our model and find out what the wavenumbers actually are. We choose a square domain defined by $0 < x < 1$, $0 < y < 1$ and let ∂B denote the square boundary.

So (1.14) is now

$$\nabla^2 \mathbf{W} + k^2 \mathbf{W} = 0, \qquad (\mathbf{n} \cdot \nabla)\mathbf{W} = 0 \qquad \text{for } (\mathbf{x}, \mathbf{y}) \text{ on } \partial B, \tag{1.32}$$

and therefore the eigenfunctions are

$$\mathbf{W}(x, y) = \mathbf{C}_{n,m} \cos n\pi x \cos m\pi y, \quad k^2 = \pi^2(n^2 + m^2), \tag{1.33}$$

where n and m are integers. Now, if $n = 0$ or $m = 0$, then we can expect to see a striped pattern emerge since there will only be spatial variation in one direction. However, if $n \neq 0$ and $m \neq 0$, then we can expect a spotted pattern to emerge. The problem arises when γ is sufficiently large so that both striped and spotted modes lie within the range of unstable wavenumbers. It is not necessarily the fastest growing mode that dominates, and there is a battle between the mode selection of stripes versus spots which, in fact, depends on which parameter values you use from the Turing space. The underlying reasons for this mode selection process is a complicated one and much research has been done in this field. For a complete understanding of this process, a full nonlinear analysis is required. The papers by Ermentrout [13] and by Nagorcka and Mooney [14] indicate that the crucial factor in the pattern selection process is the type of nonlinearity present. Moreover, it has been shown that cubic terms favour stripes while quadratic ones favour spots, once the reaction diffusion equations have been transformed via (1.12) so that the uniform steady state is $(0, 0)$.

1.5 Spots vs Stripes in our Model

In [15] there is an extremely useful figure (figure 1.3), created by numerical simulations, showing the parameter space of our kinetics for fixed d ($d = 20$) with the regions where different patterns are exhibited. γ was chosen large enough so that many modes can be driven unstable. Spotted patterns emerge in a large area of the Turing space, as indicated by the filled circles. In the upper part of the Turing space, striped patterns may be observed, as indicated by open circles. (In the very narrow region between the upper boundary of the Turing space and the striped area, reversed spotted patterns emerge, as indicated by the squares combined with dots, but we will not concern ourselves with this region).

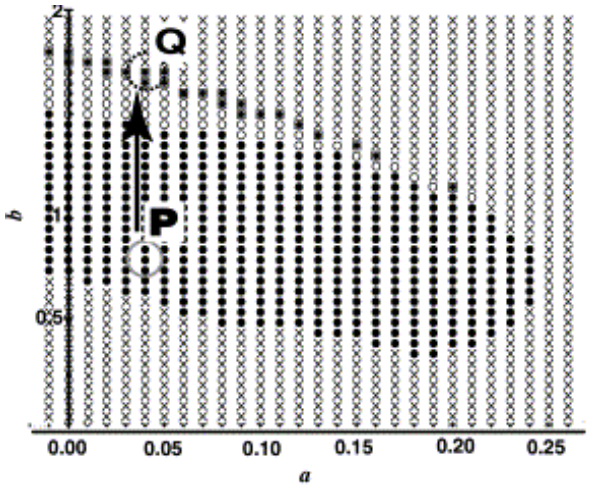
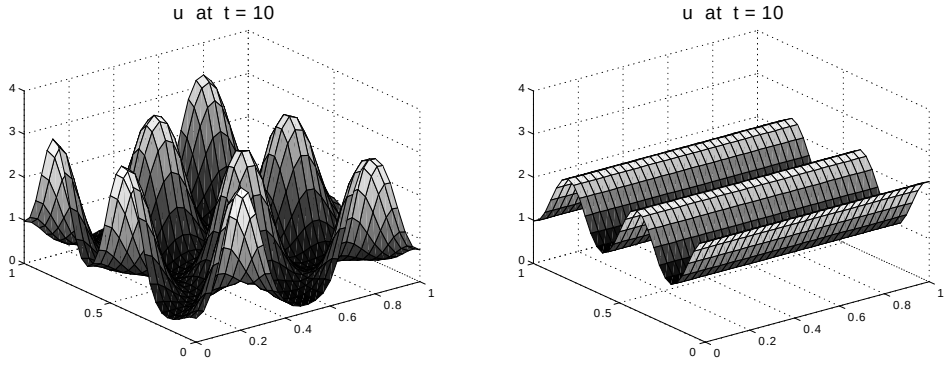


Figure 1.3. Turing space indicating the regions where spots and stripes are formed in our model for $d = 20$ and $\gamma = 10,000$. For an explanation of the different regions refer to the text. This figure is taken from [15].

To illustrate this the following figures are, once again, of numerical simulations using parameter values where we can expect to observe these different patterns (based on the above figure), with γ sufficiently large so that several modes lie within the range of unstable wavenumbers. Again, the calculation time was taken such that the patterns were stable.



(a) (b)

Figure 1.4 (a), (b). Simulations carried out with parameter values $a = 0.05$, $d = 20$ and $\gamma = 600$.Depending on the value of b we may observe spotted patterns or striped patterns; as predicted by figure 1.3. (a) $b = 1$, so that we lie in the spotted region of the Turing space given by figure 3.1. (b) $b = 1.6$, so that we now lie in the striped region.

1.6 Mode Selection - Further Analysis

As shown above, while it is reasonably straight forward to produce striped and spotted patterns from our model by choosing the parameters appropriately, we still do not know how many stripes or spots we can expect to observe. To understand this, we must return to the range of unstable wavenumbers given by (1.29) and (1.30), and the wavenumbers themselves given by (1.33). The following table lists the wavenumbers of the first several unstable modes in ascending order using $K^2_{n,m} = \pi^2(n^2 + m^2)$ - of course n and m can be interchanged.

$K^2_{1,0}$	$n = 1, m = 0$	9.87
$K^2_{1,1}$	$n = 1, m = 1$	19.74
$K^2_{2,0}$	$n = 2, m = 0$	39.48
$K^2_{2,1}$	$n = 2, m = 1$	49.35
$K^2_{2,2}$	$n = 2, m = 2$	78.96
$K^2_{3,0}$	$n = 3, m = 0$	88.83
$K^2_{3,1}$	$n = 3, m = 1$	98.70
$K^2_{3,2}$	$n = 3, m = 2$	128.30
$K^2_{4,0}$	$n = 4, m = 0$	157.91
$K^2_{4,1}$	$n = 4, m = 1$	167.78
$K^2_{3,3}$	$n = 3, m = 3$	177.65
$K^2_{4,2}$	$n = 4, m = 2$	197.39
$K^2_{4,3}$	$n = 4, m = 3$	246.74
$K^2_{5,0}$	$n = 5, m = 0$	246.74

Table 1.1

Clearly (1.30) implies that it is not just γ that affects the range of unstable wavenumbers; we could affect the range by changing a, b and d as well. Recall that for fixed a and b, we can find a critical diffusion coefficient d_c via (1.20) such that if $d > d_c$, Turing instability occurs. Thus we can choose $d = d_c + \varepsilon$, where $0 < \varepsilon << 1$ so that $h(k^2) < 0$ for a very small range of wavenumbers (see figure 1.1) and therefore isolate a single mode. Then, by using γ, we can position this range of wavenumbers to choose the mode so desired, and thus force a particular pattern to arise; even forcing a striped pattern to emerge in the region of the Turing space where typically spotted patterns emerge, and vice versa. For example, $a = 0.05$, $b = 1.4$, gives $d_c = 13.85$. Choosing $d = 13.86$ and $\gamma = 230$ gives the range $86.84 < k^2 < 92.41$, thus isolating the mode given by $K^2_{3,0} = 88.83$, whereas $\gamma = 250$ gives the range $94.39 < k^2 < 100.45$, which isolates the next mode given by $K^2_{3,1} = 98.70$. However, choosing d in such a way does have some drawbacks. Firstly, a problem arises when trying to produce the patterns given by $K_{4,3}$ and $K_{5,0}$, which have the same wavenumber but induce different patterns: the former gives spots while the latter gives stripes. In these cases it would be difficult to forecast which mode would be selected, and we would have to return to the choice of a and b, look at the Turing space and find out where any bifurcations occur. Secondly, the resulting patterns have a very low amplitude, since d is so close to d_c, and it would be preferable to produce patterns which are more pronounced.

Now, for fixed d, we know (e.g. from figure 1.3) that the values of a and b are restricted (for diffusion driven instability to occur), and we can only change them by a relatively small amount. It can be seen numerically that the allowed changes do not affect the range much; for a given γ it may be possible to change the range enough to switch between two close by modes, but that is all. Therefore, to really change the range of unstable wavenumbers by a significant amount, γ must be used.

If we fix a and b, choose (and fix) an appropriate $d > d_c$, we may then use γ to select the specific mode we want to dominate in the pattern regime. With $a = 0.05$, $b = 1.4$ and $d = 13.86$, we run into the problems mentioned above plus the additional fact that many values of γ result in a range of wavenumbers such that no mode lies within that range and hence no mode is driven unstable. For example, $\gamma = 70$ gives the range $26.43 < k^2 < 28.12$, which is a long way from the two closest modes $K_{1,1}$ and $K_{2,0}$. To reduce the number of values of γ for which no mode is selected, it is simply a matter of choosing d appropriately: we find that, for $a = 0.05$ and $b = 1.4$, choosing $d = d_c + 1 = 14.85$ has the desired effect for the first few modes. This also avoids using a value of d so close to d_c, and hence the patterns will indeed be more pronounced. In fact, in what follows, we will only look at the first four modes for simplicity and thus we need not concern ourselves with the problem of mode selection in the case where different patterns have the same wavenumbers, such as with modes $K_{4,3}$ and $K_{5,0}$.

There still remains one more issue with this choice of a, b and d. We find that for $97 < \gamma < 143$ (using only integer values for γ), mode $K_{2,0}$ is no longer isolated and mode $K_{2,1}$ becomes available as well. A similar thing happens to mode $K_{2,1}$ when $\gamma = 154$; mode $K_{2,2}$ also becomes available. If we assume that the fastest mode will dominate (see below), in both cases, then we can find the range of values of γ for which we can expect to observe each pattern using (1.30),

$$k_1^2 = \gamma\{(df_u + g_v) - [(df_u + g_v)^2 - 4d\,|A|]^{\frac{1}{2}}\} < k^2$$

$$< \gamma\{(df_u + g_v) + [(df_u + g_v)^2 - 4d\,|A|]^{\frac{1}{2}}\} = k_2^2.$$

Now $\lambda(k^2)$ attains its maximum half way between k_1^2 and k_2^2, i.e. at $(k_1^2 + k_2^2)/2$, and hence whichever mode is closest to the mid-point of k_1^2 and k_2^2 will be the fastest growing mode, and will dominate. From numerical simulations, it turns out that this assumption that the fastest mode will dominate is actually correct for these simple modes. It appears that the interaction caused by the nonlinearities is much more complex with the higher modes than when only the simpler modes are linearly unstable. As we are only dealing with the first four modes, we need not worry about bifurcation diagrams such as figure 1.3; as long as a and b lie inside the Turing space, for a fixed d, we can predict which patterns will be produced as γ changes. The table below shows the ranges of γ (again, using only integer values), where we can expect to see the different patterns obtained from the first four modes, as observed from numerical simulations and as predicted by linear theory, assuming that the fastest mode dominates. Figure 1.4 shows the patterns (obtained from numerical simulations) corresponding to these first four modes. It should be mentioned that the patterns are not unique in the sense that they may have a different orientation or polarity. The exact orientation and polarity of the patterns depends on the initial conditions.

Mode Selected	Predicted by Linear Theory	From Numerical Simulations
no mode	$1 < \gamma < 19$	$1 < \gamma < 21$
$K_{1,0}$	$20 < \gamma < 35$	$22 < \gamma < 35$
no mode	$36 < \gamma < 38$	$36 < \gamma < 41$
$K_{1,1}$	$39 < \gamma < 71$	$42 < \gamma < 70$
no mode	$72 < \gamma < 76$	$71 < \gamma < 77$
$K_{2,0}$	$77 < \gamma < 112$	$78 < \gamma < 117$
$K_{2,1}$	$113 < \gamma < 162$	$118 < \gamma < 166$

Table 1.2

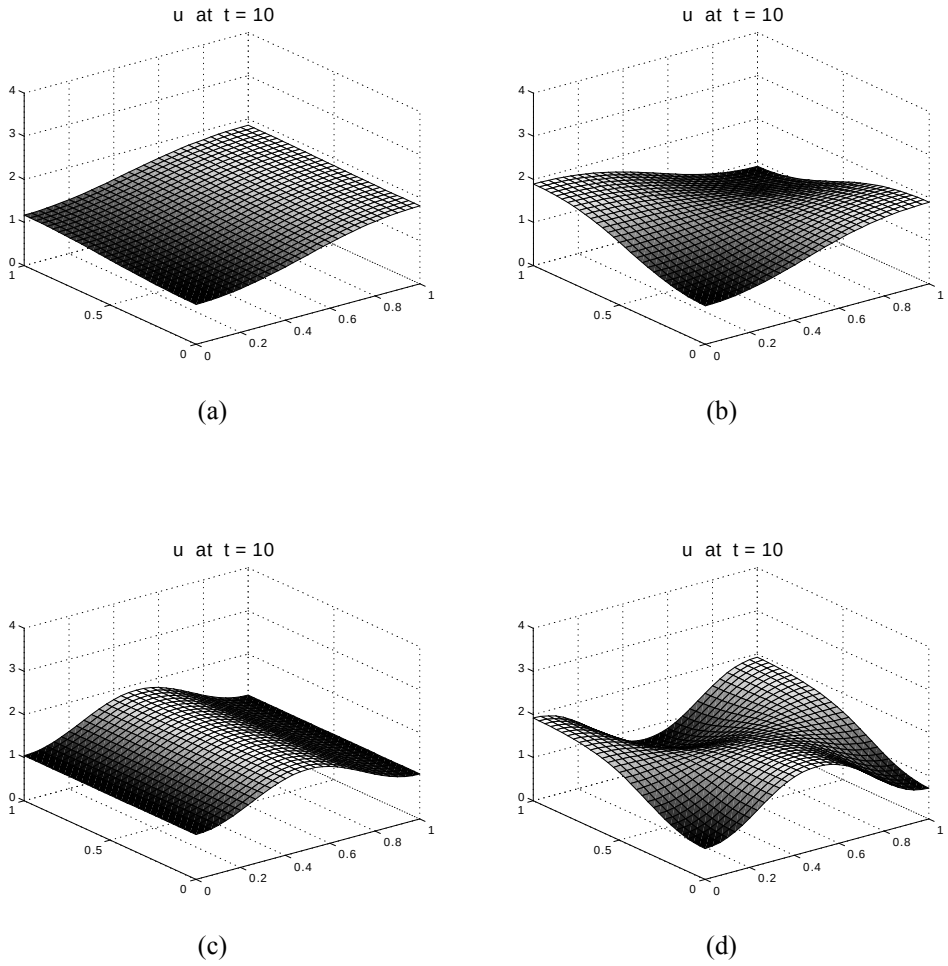


Figure 1.4 (a)-(d). The patterns corresponding to the first four modes. They all have parameter values $a = 0.05$, $b = 1.4$ and $d = 14.85$.The values of γ are chosen so that they lie within the appropriate ranges given by table 1.2. (a) The first mode, $K_{1,0}$.(b) The second mode, $K_{1,1}$.(c) The third mode, $K_{2,0}$.(d) The fourth mode, $K_{2,1}$.

Chapter 2
Stochastic Dynamics

2.1 Understanding the Model

Now that we understand the effects that the parameters have on the pattern selection process, we turn our attention to developing a stochastic evolutionary model of our Turing model that accounts for evolutionary randomness. Here we make the assumption that patterns are created by means of a Turing mechanism - our Turing mechanism. Morphogenesis is a complex dynamic process and there are, of course, many levels of the patterning process, each with its own dynamics. We do not attempt to model the dynamics of the underlying random genetic mutations; our dynamics will only apply to the phenotypic level. Put most simplistically, random mutations will result in a change in our parameters which, in turn, can result in transitions between different phenotypes, i.e. different patterns, as has been seen in the previous chapter. Therefore, we seek to model the dynamics at the phenotypic level by the evolution of our parameters in our Turing system, to account for the transitions between different patterns that can, for example, be observed in the real world. We accomplish this by allowing the parameters to evolve by a stochastic process in the way described in the following section. In fact, to keep things as straightforward as possible in chapter three, we wish to choose only one of the parameters to evolve in this way.

2.2 How Do We Implement Stochasticity?

As an illustration, let's consider fixing d ($d = 20$) and γ (with γ large) so that the transition between spots and stripes can be made quite easily by letting a and b explore the Turing space (see figure 1.3). Or, easier still, by fixing a as well and just allowing b to vary. If we fix $a = 0.05$, $d = 20$ and $\gamma = 600$, numerical simulations show that $b = 1.41$ produces a spotted pattern while $b = 1.42$ produces a striped pattern and hence there is a bifurcation at around $b = 1.41$. In dimensional terms we have from (1.4)

$$X = (\frac{k_2}{k_3})^{1/2}u, \quad Y = (\frac{k_2}{k_3})^{1/2}v, \quad t = \frac{L^2}{D_A}t^*, \quad x = Lx^*,$$

$$a = \frac{k_1}{k_2}(\frac{k_3}{k_2})^{1/2}, \quad b = \frac{k_4}{k_2}(\frac{k_3}{k_2})^{1/2}, \quad d = \frac{D_B}{D_A}, \quad \gamma = \frac{L^2k_2}{D_A}, \tag{2.1}$$

and hence a variation in the rate of reaction k_4 will result in a variation of the dimensionless parameter b, whilst all other parameters remain unchanged. By using this feature we can develop a stochastic model of the evolution of our RD model which accounts for evolutionary randomness and allows transitions across the Turing space, from spotted regions to striped ones. For the sake of simplicity, suppose that we keep all the parameters fixed except b, and

let b evolve in time by simple Brownian motion $b = b(T) = b_0 + x(T)$ where b_0 is the initial value of b at $T = 0$ and $x(T)$ is a random variable, normally distributed, with zero mean and delta correlation in time, (i.e. we assume that the noise is white). The time scale on which b evolves is much greater than the time scale on which the pattern evolves: we assume that the generated pattern is set out in the early stages of embryonic development, whereas it is only down the generations of a particular species that different patterns may be induced by changes in b. Therefore, it is logical to let the changes in b be discrete and write $b_{T+1} = b_T + x$ where b_0 and x are as above, and so the successive values b_T represent values of b at successive generations. Then for each value b_T, we evaluate the RD system and find out what pattern is induced at that generation. In this case, however, we already know when a spotted or striped pattern will arise as b changes, since we know that there is a bifurcation at $b \approx 1.41$. Therefore, by suitably choosing b_0 and the standard deviation of x, we can generate a sequence of values b_T such that we will obtain a spotted pattern for several generations, followed by a striped pattern for several generations, etc. This also includes the possibility that b_T lies outside the Turing space for several generations and hence no patterns arise. The following table gives a list of values b_T (to 2 decimal places) generated by numerical methods (by MATLAB) using $b_0 = 1.50$ and x with mean zero and standard deviation 0.05, demonstrating how it is possible to make a transition between different patterns. Of course it is possible, for example, that if $b_0 = 1$ and x has a very low standard deviation, then it would take a long time for b to pass beyond the bifurcation point $b \approx 1.41$ and hence it would take a very long time for a transition to be made from spots to stripes.

$b_0 = 1.50$	
$b_1 = 1.49$	$b_{11} = 1.45$
$b_2 = 1.49$	$b_{12} = 1.40$
$b_3 = 1.51$	$b_{13} = 1.41$
$b_4 = 1.48$	$b_{14} = 1.37$
$b_5 = 1.47$	$b_{15} = 1.40$
$b_6 = 1.45$	$b_{16} = 1.43$
$b_7 = 1.45$	$b_{17} = 1.42$
$b_8 = 1.52$	$b_{18} = 1.32$
$b_9 = 1.49$	$b_{19} = 1.32$
$b_{10} = 1.43$	$b_{20} = 1.40$

Table 2.1

2.3 Choosing which Parameter to Evolve

We can see that by fixing d and γ, as above, and by either fixing a and letting b vary or by fixing b and letting a vary, a transition can easily be made from spots to stripes. However, as has been noted in section 1.6, we do not know how many spots or stripes would be present

if we chose to do this, and this information is actually essential to our problem in chapter three. As is probably clear from our analysis in section 1.6, it is γ that we choose to evolve. We can see from (2.1) that a change in γ results from a change in L, k_2 or D_A, but that only a change in L will leave the other parameters unchanged. Therefore, it makes sense to interpret a change in γ as a change in the domain size L.

Now if we fix $a = 0.05$, $b = 1.4$ and $d = d_c + 1 = 14.85$, as we did earlier in section 1.6, then we know precisely which values of γ give rise to the patterns given in figure 1.4 by using table 1.2. It will be clear in section 3.1 that this is exactly what we need to know. It should be noted that although this choice of a, b and d is more or less arbitrary, since this model is based on a hypothetical reaction, we may well choose the parameter values in order to simplify matters. Thus, we will assume that a, b and d are fixed as above, and that only γ evolves by Brownian motion in the way discussed earlier. By choosing γ_0 and the standard deviation of x appropriately, as γ evolves over time we can expect different patterns to be observed corresponding to the first four modes, shown in figure 1.4.

Chapter 3
Constructing a Maximum Likelihood Phylogeny

3.1 The Problem

By letting γ evolve in our RD model according to a Brownian motion process, we can expect a phylogenetic tree to be formed where different patterns are produced at the tips corresponding to different values of γ. In this chapter, we will only focus on the simple patterns given by figure 1.4. Now, supposing that we are just given these patterns (from figure 1.4) with no knowledge of how they were produced, we would like to construct a maximum likelihood phylogeny for these patterns. This can be done by using the the following strategy. Given the patterns, we assume that they were generated from our Turing mechanism. Furthermore, we assume that it is γ that has evolved in a random way (and hence is responsible for the resultant phylogeny), by the process of Brownian motion, whilst all other parameters remain fixed. The next step is to identify these patterns with our Turing patterns by referring to figure 1.4. We will then be able to infer the values of γ from which the patterns are formed by using table 1.2. Once we have inferred the values of γ at the tips of the tree, we can construct a maximum likelihood phylogeny by using the REML approach, which is based on γ evolving by Brownian motion. It is to this inference problem that we devote the remaining sections.

3.2 Maximum Likelihoods and the REML Approach

Now, let us consider the evolution of γ by Brownian motion. Each displacement of γ is independent with mean zero and variance σ^2, and the variance of each individual displacement is the same regardless of what the value of γ is. The net displacement is the sum of the individual steps, and the total variance is the variance of this sum. This, in turn, is equal to the sum of the individual variances since the individual steps are independent. Therefore, if σ^2 is the variance that accumulates per unit time, then the variance of the net displacement after time t will be $\sigma^2 t$. If the total displacement of γ along a branch is the sum of a large number of independent quantities then this means that the total displacement is, in fact, normally distributed. Therefore, we know that the total displacement along a branch of length t is taken from a normal distribution with mean zero and variance $\sigma^2 t$, and furthermore, since the small steps along different branches are all independent, the displacements in different branches of a tree are independent.

By assuming that the character γ evolves by Brownian motion, we can calculate the likelihood for any given tree quite easily since the changes in each branch are independent and normally distributed. We will focus on the three-species case and use the restricted

maximum likelihood (REML) as explained by Felsenstein in [16] for a three-species model with multiple characters. By using the REML approach we can find maximum likelihood estimates for the branch lengths but not for the root node or any interior nodes. We now go through the details of the method for the simple case where we only have one character, namely γ. Figure 3.1 shows a three-species evolutionary tree with the root node (0), and the interior node - or branch point - (4), labelled as well. The v_i next to each branch of the tree is the amount of variance (i.e. $\sigma^2 t$) expected to accumulate during evolution along that part of the tree (we refer to the $v_i's$ as branch lengths).

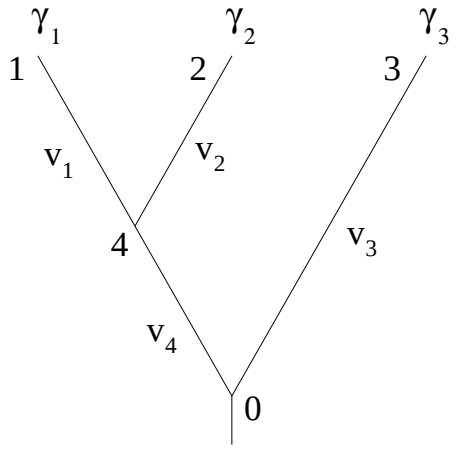


Figure 3.1. A three-species tree that has resulted from the evolution of the character γ. For an explanation of the notation, refer to the text.

The first thing we need to examine is the joint distribution of γ_1, γ_2, γ_3 and γ_4 given γ_0. Now the values of γ_1, γ_2 and γ_3 are

$$\gamma_1 = (\gamma_1 - \gamma_4) + (\gamma_4 - \gamma_0) + \gamma_0,$$
$$\gamma_2 = (\gamma_2 - \gamma_4) + (\gamma_4 - \gamma_0) + \gamma_0,$$
$$\gamma_3 = (\gamma_3 - \gamma_0) + \gamma_0.$$

So γ_1 is the sum of three terms, the first two of which are normally distributed with mean zero and variances v_1 and v_4. Therefore, γ_1 is normally distributed with mean γ_0 and variance $v_1 + v_4$. Similarly, γ_2 is normally distributed with mean γ_0 and variance $v_2 + v_4$, and γ_3 is simply normally distributed with mean γ_0 and variance v_3. Now γ_1 and γ_2 have a covariance due to their common term $(\gamma_4 - \gamma_0)$, which will be the variance of this common term, namely v_4. The covariance of γ_1 with γ_3 is clearly zero, as is the covariance of γ_2 with γ_3. Thus, the three observed values for γ_1, γ_2 and γ_3 are expected to

be multivariate normally distributed with means zero and with variances and covariances as mentioned above - i.e. the covariance matrix is

$$\begin{bmatrix} v_1 + v_4 & v_4 & 0 \\ v_4 & v_2 + v_4 & 0 \\ 0 & 0 & v_3 \end{bmatrix}.$$

Now, to calculate likelihoods, the REML approach says that we should look at the differences between the $\gamma_i's$ and by doing this, we find that the calculation is greatly simplified. The difference $\gamma_1 - \gamma_2$ has mean zero and variance given by

$$\begin{aligned} Var[\gamma_1 - \gamma_2] &= Var[\gamma_1] + Var[\gamma_2] - 2Cov[\gamma_1, \gamma_2] \\ &= (v_1 + v_4) + (v_2 + v_4) - 2v_4 \\ \\ &= v_1 + v_2 \end{aligned}$$

To simplify the calculation of the likelihood, our next strategy is to compute a weighted average of γ in the two tips such that this average is independent of $\gamma_1 - \gamma_2$. By doing this, we can effectively split the tree into two independent trees, since γ_3 is independent of $\gamma_1 - \gamma_2$ as well. We can then calculate the likelihoods of the two trees independently and multiply them together to obtain the likelihood of our original tree. Suppose that the weighted average is $f\gamma_1 + (1-f)\gamma_2$. We want the covariance of this average with $\gamma_1 - \gamma_2$ to be zero:

$$\begin{aligned} & Cov[f\gamma_1 + (1-f)\gamma_2,\ \gamma_1 - \gamma_2] \\ \\ &= fCov[\gamma_1,\ \gamma_1] - (1-f)Cov[\gamma_2,\ \gamma_2] + (1-2f)Cov[\gamma_1,\ \gamma_2] \\ \\ &= fVar[\gamma_1] - (1-f)Var[\gamma_2] + (1-2f)Cov[\gamma_1,\ \gamma_2] \\ \\ &= f(v_1 + v_4) - (1-f)(v_2 + v_4) + (1-2f)v_4 \\ \\ &= fv_1 - (1-f)v_2 \end{aligned}$$

If this is to be equal to zero, then we require that

$$\begin{aligned} & f(v_1 + v_2) - v_2 = 0 \\ & \Rightarrow f = v_2/(v_1 + v_2) \end{aligned}$$

Hence, the weighted average of γ is

$$\gamma'_4 = \frac{v_2\gamma_1 + v_1\gamma_2}{v_1 + v_2} \tag{3.1}$$

where we treat γ'_4 as the value of γ at the interior node 4 in the sense that it is inferred from its descendants. This value γ'_4 has zero covariance with $\gamma_1 - \gamma_2$ and also has zero covariance with γ_3, since

$$Cov[\gamma'_4,\ \gamma_3] = Cov[f\gamma_1 + (1-f)\gamma_2,\ \gamma_3]$$

$$= fCov[\gamma_1, \gamma_3] + (1-f)Cov[\gamma_2, \gamma_3]$$

$$= 0.$$

Hence, the three values ($\gamma_1 - \gamma_2$, γ'_4, γ_3) are independently normally distributed with means (0, γ_0, γ_0). For an REML estimate, we can ignore the values γ'_4 and γ_3 and use only their difference $\gamma'_4 - \gamma_3$, which is normally distributed with mean zero and variance equal to $v_3 + Var[\gamma'_4]$. We may calculate $Var[\gamma'_4]$ using (3.1) and it turns out that

$$Var[\gamma'_4] = v'_4 = v_4 + v_1v_2/(v_1 + v_2). \tag{3.2}$$

Hence $\gamma'_4 - \gamma_3$ has variance $v_3 + v_4 + v_1v_2/(v_1 + v_2)$.

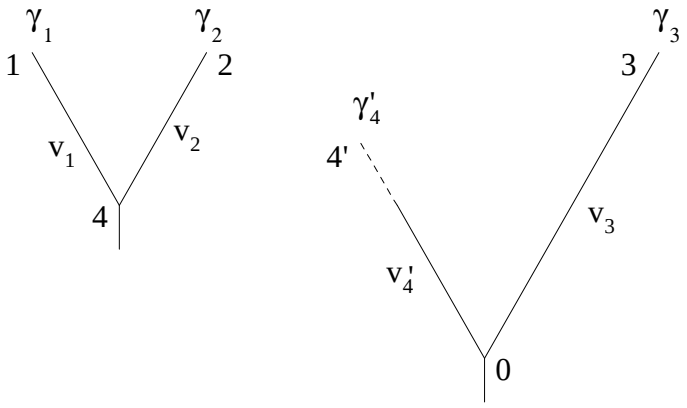


Figure 3.2. Using the pruning method we can effectively split our three-species tree in figure 3.1 into two independent trees. These two trees, when taken together, have the same restricted likelihood as the tree in figure 3.1. γ'_4 and v'_4 are defined in the text.

This process whereby nodes 1 and 2 are effectively replaced by the fictional value γ_4' is known as "pruning" the tree and is an idea that was first introduced by Felsenstein in [9]. Thus, when making an REML estimate, the likelihood of the tree in figure 3.1 is exactly the same as the product of the likelihoods of the two trees in figure 3.2. This method can quite easily be extended to a situation where many species are involved (see [16]). Now, the likelihoods of the two trees in figure 3.2 are

$$L_1 = \frac{1}{[2\pi(v_1+v_2)]^{1/2}} \exp[-\frac{(\gamma_1-\gamma_2)^2}{2(v_1+v_2)}], \tag{3.3}$$

and

$$L_2 = \frac{1}{[2(v_3+v_4+v_1v_2/(v_1+v_2))]^{1/2}} \exp[-\frac{(\gamma_4'-\gamma_3)^2}{2(v_3+v_4+v_1v_2/(v_1+v_2))}], \tag{3.4}$$

since $\gamma_1-\gamma_2$ and $\gamma_4'-\gamma_3$ are both normally distributed, both with mean zero and variance v_1+v_2 and $v_3+v_4+v_1v_2/(v_1+v_2)$ respectively. Therefore, the likelihood $L=L_1L_2$ only depends on v_3 and v_4 through their sum $v_3^*=v_3+v_4$. This means that we cannot find a maximum likelihood for each of these branch lengths individually, but only for their sum. It also means that as far as REML estimation is concerned, the tree in figure 3.3 is exactly the same as the tree in figure 3.1, in that they both have the same restricted likelihood. Hence, we may assume that we are in the case of figure 3.3, and find maximum likelihood estimates for $\hat{v}_1$, $\hat{v}_2$ and $\hat{v}_3$ (dropping the asterisk on v_3).

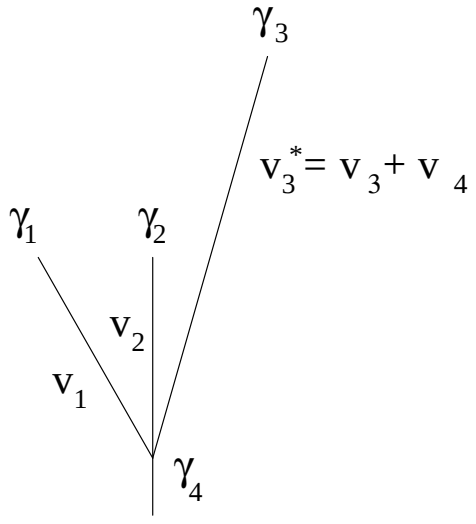


Figure 3.3. This tree has exactly the same restricted likelihood as the tree in figure 3.1.

3.3 Finding the Exact Maximum Likelihood Estimates for the Three-Species Case

By applying the pruning process to the tree of figure 3.3 we get the likelihood $L = L_1 L_2$ where L_1 and L_2 are defined by (3.3) and (3.4) with $v_3^* = v_3 + v_4$. On dropping the asterisk on v_3 and using (3.1) we get

$$L = \frac{1}{[2\pi(v_1+v_2)]^{1/2}} \exp[-\frac{(\gamma_1-\gamma_2)^2}{2(v_1+v_2)}]$$

$$\cdot \frac{1}{[2\pi(v_3+v_1v_2/(v_1+v_2))]^{1/2}} \exp[-\frac{[\frac{v_2\gamma_1+v_1\gamma_2}{v_1+v_2}-\gamma_3]^2}{2[v_3+v_1v_2/(v_1+v_2)]}]. \tag{3.5}$$

Which has log-likelihood

$$\ln L = -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(v_1+v_2) - \frac{1}{2}\frac{(\gamma_1-\gamma_2)^2}{(v_1+v_2)}$$

$$-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln[v_3+v_1v_2/(v_1+v_2)] - \frac{1}{2}\frac{[\frac{v_2\gamma_1+v_1\gamma_2}{v_1+v_2}-\gamma_3]^2}{[v_3+v_1v_2/(v_1+v_2)]}.$$

This expression can be simplified by

$$\ln(v_1+v_2) + \ln[v_3+v_1v_2/(v_1+v_2)] = \ln(v_1v_2+v_1v_3+v_2v_3),$$

and by

$$\frac{(\gamma_1-\gamma_2)^2}{(v_1+v_2)} + \frac{[\frac{v_2\gamma_1+v_1\gamma_2}{v_1+v_2}-\gamma_3]^2}{[v_3+v_1v_2/(v_1+v_2)]}$$

$$= \frac{v_3(\gamma_1-\gamma_2)^2 + [v_1v_2/(v_1+v_2)](\gamma_1-\gamma_2)^2 + (v_1+v_2)[\frac{v_2\gamma_1+v_1\gamma_2}{v_1+v_2}-\gamma_3]^2}{v_1v_2+v_1v_3+v_2v_3}.$$

Now, we can continue to simplify the numerator as follows:

$$v_3(\gamma_1-\gamma_2)^2 + [v_1v_2/(v_1+v_2)](\gamma_1-\gamma_2)^2 + (v_1+v_2)[\frac{v_2\gamma_1+v_1\gamma_2}{v_1+v_2}-\gamma_3]^2$$

$$= v_3(\gamma_1-\gamma_2)^2+\frac{v_1v_2}{v_1+v_2}(\gamma_1-\gamma_2)^2+\frac{(v_2\gamma_1+v_1\gamma_2)^2}{v_1+v_2}-2\gamma_3(v_2\gamma_1+v_1\gamma_2)+(v_1+v_2)\lambda_3^2$$

$$= v_3(\gamma_1-\gamma_2)^2+\frac{v_1v_2}{v_1+v_2}(\gamma_1^2-\gamma_2^2)+\frac{v_2^2\gamma_1^2+v_1^2\gamma_2^2}{v_1+v_2}-2v_2\gamma_1\gamma_3-2v_1\gamma_2\gamma_3+v_1\lambda_3^2+v_2\lambda_3^2$$

$$= v_3(\gamma_1-\gamma_2)^2+\frac{v_2\gamma_1^2(v_1+v_2)}{v_1+v_2}+\frac{v_1\gamma_2^2(v_1+v_2)}{v_1+v_2}-2v_2\gamma_1\gamma_3-2v_1\gamma_2\gamma_3+v_1\lambda_3^2+v_2\lambda_3^2$$

$$= v_3(\gamma_1-\gamma_2)^2+v_2(\gamma_1-\gamma_3)^2+v_1(\gamma_2-\gamma_3)^2$$

Using these simplifications we see that the log-likelihood is

$$\ln L = -\ln(2\pi)-\frac{1}{2}\ln(v_1v_2+v_1v_3+v_2v_3)$$

$$-\frac{v_3(\gamma_1-\gamma_2)^2+v_2(\gamma_1-\gamma_3)^2+v_1(\gamma_2-\gamma_3)^2}{2(v_1v_2+v_1v_3+v_2v_3)}. \tag{3.6}$$

For convenience we write this as

$$\ln L = -\ln(2\pi)-\frac{1}{2}\ln T-\frac{S}{2T} \tag{3.7}$$

where

$$S = v_3(\gamma_1-\gamma_2)^2+v_2(\gamma_1-\gamma_3)^2+v_1(\gamma_2-\gamma_3)^2,$$

$$T = v_1v_2+v_1v_3+v_2v_3.$$

Now to maximize the likelihood we can differentiate this log-likelihood with respect to the branch lengths and put all these derivatives equal to zero:

$$\frac{\partial \ln L}{\partial v_1} = -\frac{(v_2+v_3)}{2T}-\frac{(\gamma_2-\gamma_3)^2}{2T}+\frac{S(v_2+v_3)}{2T^2}=0 \tag{3.8}$$

$$\frac{\partial \ln L}{\partial v_2} = -\frac{(v_1+v_3)}{2T}-\frac{(\gamma_1-\gamma_3)^2}{2T}+\frac{S(v_1+v_3)}{2T^2}=0 \tag{3.9}$$

$$\frac{\partial \ln L}{\partial v_3} = -\frac{(v_1+v_2)}{2T}-\frac{(\gamma_1-\gamma_2)^2}{2T}+\frac{S(v_1+v_2)}{2T^2}=0 \tag{3.10}$$

If we multiply the first equation by $2T/(v_2+v_3)$ we get

$$-1-\frac{(\gamma_2-\gamma_3)^2}{(v_2+v_3)}+\frac{S}{T}=0 \tag{3.11}$$

Similarly, for the other two equations we get

$$-1 - \frac{(\gamma_1 - \gamma_3)^2}{(v_1 + v_3)} + \frac{S}{T} = 0 \tag{3.12}$$

$$-1 - \frac{(\gamma_1 - \gamma_2)^2}{(v_1 + v_2)} + \frac{S}{T} = 0 \tag{3.13}$$

Also if we calculate the sum $v_1(3.8) + v_2(3.9) + v_3(3.10)$, we find that

$$-1 - \frac{S}{2T} + \frac{S}{T} = 0$$
$$\Rightarrow \frac{S}{T} = 2,$$

and hence (3.11)-(3.13) can be written as

$$v_2 + v_3 = (\gamma_2 - \gamma_3)^2 \tag{3.14}$$

$$v_1 + v_3 = (\gamma_1 - \gamma_3)^2 \tag{3.15}$$

$$v_1 + v_2 = (\gamma_1 - \gamma_2)^2. \tag{3.16}$$

From which we can now easily obtain the following maximum likelihood estimates $\hat{v}_1$, $\hat{v}_2$ and $\hat{v}_3$:

$$\hat{v}_1 = \frac{1}{2}[(\gamma_1 - \gamma_2)^2 + (\gamma_1 - \gamma_3)^2 - (\gamma_2 - \gamma_3)^2]$$
$$\hat{v}_2 = \frac{1}{2}[(\gamma_1 - \gamma_2)^2 + (\gamma_2 - \gamma_3)^2 - (\gamma_1 - \gamma_3)^2]$$
$$\hat{v}_3 = \frac{1}{2}[(\gamma_1 - \gamma_3)^2 + (\gamma_2 - \gamma_3)^2 - (\gamma_1 - \gamma_2)^2],$$

which can be written slightly more simply as

$$\hat{v}_1 = (\gamma_1 - \gamma_2)(\gamma_1 - \gamma_3) \tag{3.17}$$

$$\hat{v}_2 = (\gamma_2 - \gamma_1)(\gamma_2 - \gamma_3) \tag{3.18}$$

$$\hat{v}_3 = (\gamma_1 - \gamma_3)(\gamma_2 - \gamma_3). \tag{3.19}$$

3.4 Applying an REML Estimate to Our Model for a Three-Species Tree

3.4.1 Simulating the Phylogeny

In this concluding section, we create a three-species tree - using sensible values for γ_0 and the $v_i's$ - which has the same topology as the tree in figure 3.1 by letting γ evolve by Brownian motion. Referring to table 1.2 (using the ranges from the numerical simulations), we then note which patterns will be produced at the tips. We then work backwards, and try to infer the $v_i's$ which maximize the likelihood of that particular tree by using an REML estimate. Of course, this means that we cannot estimate v_3 and v_4 individually but only their sum.

First of all, let us take $\gamma_0 = 118$, i.e. the value of γ where there is a bifurcation between modes K_3 and K_4. Now we want to choose the $v_i's$ such that it is highly probable that γ will still lie within one of the ranges given in table 2.3 after it had evolved. By using computer software such as MATLAB, we find that taking v_3 and v_4 as 30^2 and 40^2 are sensible enough choices. So we simulate the evolution of γ_0 along the first two branches with $\gamma \sim N(0, 30^2)$ (where $\gamma \sim N(\mu, \sigma^2)$ means that γ is normally distributed with mean μ and variance σ^2), and $\gamma \sim N(0, 40^2)$ and obtain the two values 130 and 77 for γ. It makes sense to assign $\gamma_3 = 130$ and $\gamma_4 = 77$, on figure 3.1, since we expect that a branching will be more likely to occur at $\gamma_4 = 77$, given the ranges where the different patterns occur in table 2.3. Now we let $\gamma_4 = 77$ evolve along the next two branches with suitable choices for v_1 and v_2. Simulating this evolution with $\gamma \sim N(0, 5^2)$ and $\gamma \sim N(0, 15^2)$, we obtain the values 81 and 62 for γ, and assign $\gamma_1 = 62$ and $\gamma_2 = 81$ just so that γ_1, γ_2 and γ_3 are in ascending order with $\gamma_1 = 62$, $\gamma_2 = 81$ and $\gamma_3 = 130$. Therefore, at the tips we observe the patterns given by $\gamma_1 = 62$, $\gamma_2 = 81$ and $\gamma_3 = 130$ which are, respectively, the patterns obtained from modes K_2, K_3 and K_4 (refer to figure 2.1 for these patterns).

3.4.2 Inferring the Phylogeny

By assuming that we know the tree topology and have been given these patterns at the tips, we can infer the branch lengths via an REML estimate as follows. Given one of the patterns from figure 1.4, we know, from table 1.2, that there is a distinct range of possible values of γ that could give rise to that pattern. We can, therefore, infer the values of γ by using table 1.2. However, to use an REML estimate, we must work with discrete values. We could take any value of γ from the possible range of values, and any value we take would be perfectly reasonable. Since we need to make a choice, we will simply assume that γ takes the value at the mid-point of the particular range of values that are possible.

If we apply this to our patterns given by modes K_2, K_3 and K_4, we obtain $\gamma_1 = 56$, $\gamma_2 = 98$ and $\gamma_3 = 142$. Using (3.17)-(3.19) we have the following maximum likelihood estimates

$$\hat{v}_1 = 3612$$
$$\hat{v}_2 = -1848$$
$$\hat{v}_3^* = 3784,$$

where $\hat{v}_3^* = \hat{v}_3 + \hat{v}_4$. Clearly, negative values are not desirable, as they make no biological sense, and therefore we impose the constraint that each $\hat{v}_i \geq 0$. Putting $\hat{v}_2 = 0$, we now return to equations (3.17)-(3.19) to get the maximum likelihood estimates for $\hat{v}_1$ and $\hat{v}_3^*$. If $\hat{v}_2 = 0$, then (3.18) implies that

$$\gamma_2^2 - \gamma_1\gamma_2 - \gamma_2\gamma_3 + \gamma_1\gamma_3 = 0$$
$$\Rightarrow \gamma_1\gamma_3 = \gamma_1\gamma_2 + \gamma_2\gamma_3 - \gamma_2^2.$$

Putting this into equations (3.17) and (3.19), we get

$$\hat{v}_1 = (\gamma_2 - \gamma_1)^2$$
$$\hat{v}_3^* = (\gamma_3 - \gamma_2)^2.$$

Hence, for our values of γ we obtain the estimates

$$\hat{v}_1 = 42^2$$
$$\hat{v}_2 = 0$$
$$\hat{v}_3^* = 44^2.$$

It is clear that these estimates are quite different to the actual values. Of course one of the main reasons for this is because we have assumed that given a pattern, there is only one γ which corresponds to it, whereas in reality we know that there is a wide range of values that are permissible. However, another concern is the estimate $\hat{v}_2 = 0$. Having $\hat{v}_2 = 0$ essentially means that to maximize the likelihood of our tree, we must take γ_2 as the phenotype of node 4 on figure 3.1, which implies that the phenotype γ_1 is a direct descendent of the phenotype γ_2. This is not what we were looking for, but in actual fact, this was inevitable since we were only using one character to infer the branch lengths. This becomes apparent if we look back at equations (3.17)-(3.19). Suppose that the γ_i are all distinct and that, without loss of generality, $(\gamma_1 - \gamma_2) > 0$. To make sure that $\hat{v}_1 \geq 0$, (3.17) forces $(\gamma_1 - \gamma_3) > 0$. (3.19) in turn forces $(\gamma_2 - \gamma_3) > 0$ to keep $\hat{v}_3^* \geq 0$, but now we can see from (3.18) that we cannot maintain $\hat{v}_2 \geq 0$ since we already have the conditions $(\gamma_1 - \gamma_2) > 0$ and $(\gamma_2 - \gamma_3) > 0$. Hence, one of the $\hat{v}_i's$ will always be negative. However, if multiple characters are used to infer the phylogeny, then this is not necessarily the case. When p characters are used, say, then it can be derived (see [16]) that

$$\hat{v}_1 = \sum_j (x_{1j} - x_{2j})(x_{1j} - x_{3j})/p \tag{3.20}$$

$$\hat{v}_2 = \sum_j (x_{2j} - x_{1j})(x_{2j} - x_{3j})/p \tag{3.21}$$

$$\hat{v}_3 = \sum_j (x_{3j} - x_{1j})(x_{3j} - x_{2j})/p, \tag{3.22}$$

where x_{ij} is the value of character j in species i. Clearly the more characters that are used, the more reliable our maximum likelihood estimates will be. For our particular model this means letting our other parameters a, b and d evolve by Brownian motion, too. We would also have to try and construct a table similar to that of table 2.3, which involves all the parameters and not just γ. Although this may be difficult - we have seen the subtle effects of changing the parameters in the earlier chapters - in principle it should be possible, at least by using numerical methods if not analytical ones.

Concluding Remarks

We have shown that it is actually possible to make a maximum likelihood inference of phylogenies (for the three-species case), the phylogenies of which represent the evolution of RD patterns - via the evolution of the parameters by Brownian motion - by using the REML approach. Although starting off with a simple model is extremely revealing, in order for this inference procedure to have any real biological meaning, we should use reaction kinetics which are empirically based instead. In this way, we could find out the actual ranges of values for the kinetics parameters and the ratio of diffusion coefficients, rather than just choosing values more or less arbitrarily to make the analysis as simple as possible. We should then analyse this more realistic model in the same way we analysed our model, discussed above. In doing do, we should not restrict ourselves to the patterns produced only by the simple modes, but also consider the richer patterns produced by the more complex modes. Next, we should use this more realistic model and allow for several of the parameters to evolve by Brownian motion in such a way that a phylogeny consisting of these richer patterns can be realised. It should then be possible to look at real data, which generally consists of these richer patterns (rather than simulated data used in this study), and apply an REML estimate to this real data in the same way described in this report.

When applying an REML estimate to this data, we know that the estimate will be much more accurate when several characters are utilized. Therefore, in the evolutionary RD model, we should allow as many of the paramaters to evolve as possible so that when it comes to the inference problem, we can treat each of these parameters as a character in the REML estimate. Finally, it is important to remember that the REML estimate that we have employed requires us to know the tree topology in advance, which we may or may not be able to infer from other data. It also assumes that the character states are discrete, and uses discrete values to obtain the maximum likelihood estimates. However, we have seen that the character states are actually continuous and therefore other methods may be more suitable to obtain the estimates being sought. (For further discussion on making maximum likelihood inferences of phylogenies when continuous characters evolve by Brownian motion see [17]). In principle, then, it should be possible to obtain a reasonably accurate maximum likelihood phylogeny which refers to real data by using the methods described in this project, and by implementing the suggestions outlined above.

Appendix A
MATLAB Code

To solve the Reaction-Diffusion Equations, given by (1.7), with reaction kinetics given by (1.22), numerically, we used MATLAB. A simple Euler method was implemented by using the following code.

```
function []=RDzeroflux(N,a,b,d,gamma)

% defining the mesh

N = 30;
h = 1/N;                        % step size in x and y
x = h*(0:N);                    % x coordinates of grid
y = h*(0:N);                    % y coordinates of grid
[xx,yy] = meshgrid(x,y);        % 2D x and y mesh coordinates
dt = .01*h^2;                   % time step - usually small

% parameter values

a = .05;
b = 1.4;
d = 14.85;
gamma = 400;

% Initial data at t=0:

us = (a+b);                     % u steady state
vs = b/us^2;                    % v steady state
u = (a+b)*ones(size(xx));       % u steady state
v = (b./u.^2);                  % v steady state
u = u + .1*randn(size(xx));     % add small perturbations about the steady state
v = v + .1*randn(size(xx));     % add small perturbations about the steady state

% Time-stepping:

t = 0;
tmax = 10;
nsteps = round(tmax/dt);        % number of time steps
```

```
for n = 1:nsteps                    % main time-stepping loop

t = t+dt;
uE = u(:,[2:N+1 N]);
uW = u(:,[2 1:N]);
uN = u([2 1:N],:);
uS = u([2:N+1 N],:);
vE = v(:,[2:N+1 N]);
vW = v(:,[2 1:N]);
vN = v([2 1:N],:);
vS = v([2:N+1 N],:);

% finite difference formula

u2v = u.^2.*v;
u = u + gamma*dt*(a - u + u2v) + dt*(uE+uW+uN+uS-4*u)/h^2;
v = v + gamma*dt*(b - u2v) + d*dt*(vE+vW+vN+vS-4*v)/h^2;

end

% plot solution:

colormap('gray')
surf(x,y,u)
title(['u at t = ' num2str(t)],'fontsize',16)
zlim([0 4])
shg
```

Bibliography

[1] A.M. Turing, *The chemical basis of morphogenesis*, Phil. Trans. R. Soc. Lond. B 237, 37-72, 1952.

[2] P.K. Maini, K.J. Painter and H.N.P Chau, *Spatial pattern formation in chemical and biological systems*, Faraday Trans. 93, 3601-3610, 1997.

[3] A. Gierer, H. Meinhardt, *A theory of biological pattern formation*, Kybernetik 12, 30-39, 1972.

[4] P. Gray, S.K. Scott, *Autocatalytic reactions in the isothermal, continuous stirred tank reactor: oscillations and instabilities in the system* $A + 2B \rightarrow 3B$, $B \rightarrow C$, Chem. Engrg. Sci. 39, 1087-1097, 1984.

[5] I. Lengyel, I.R. Epstein, *Modelling of Turing structures in the chlorite-iodine-malonic acid-starch reaction system*, Science 251, 650-652, 1991.

[6] J. Schnakenberg, *Simple chemical reaction systems with limit cycle behaviour*, J. Theor. Biol. 81, 389-400, 1979.

[7] D. Thomas, *Artificial enzyme membranes, transport, memory and oscillatory phenomena*, in: D. Thomas, J.-P. Kernevez (Eds.), *Analysis and Control of Immobilized Enzyme Systems*, Springer, Berlin, pp. 115-150, 1975.

[8] A.W.F. Edwards, L.L. Cavalli-Sforza, *Reconstruction of evolutionary trees*, in: V.H. Heywood, J. McNeill (Eds), *Phenetic and Phylogenetic Classification*, Systematics Association Publ., pp. 67-76, 1964.

[9] J. Felsenstein, *Maximum-likelihood estimation of evolutionary trees from continuous characters*, Amer. J. Hum. Genet. 25, 471-492, 1973.

[10] L.L. Cavalli-Sforza, A. Piazza, *Analysis of evolution: evolutionary rates, independence, and treeness*, Theor. Pop. Biol. 8, 127-165, 1975.

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

[12] K.W. Morton, D.F. Mayers, *Numerical Solution of Partial Differential Equations*, C.U.P., Cambridge, 1994.

[13] B. Ermentrout, *Stripes or spots? Nonlinear effects in bifurcation of reaction-diffusion equations on the square*, Proc. R. Soc. Lond. A 434, 413-417, 1991.

[14] B.N. Nagorcka, J.R. Mooney, *From stripes to spots: prepatterns which can be produced in the skin by a reaction-diffusion system*, IMA J. Math. Appl. Med. Biol. 9, 249-269, 1992.

[15] H. Shoji, Y. Iwasaa and S. Kondo, *Stripes, spots, or reversed spots in two-dimensional Turing systems*, J. Theor. Biol. 224, 339-350, 2003.

[16] J. Felsenstein, *Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates*, Evolution 35, 1229-1242, 1981.

[17] J. Felsenstein, *Inferring Phylogenies*, Sinauer Associates Inc., Massachusetts, 2004.