Effects of turbulence and a patchy environment on the dynamics of plankton populations

Richard Matthew Hillary

B.Sc. (Hons), M.Sc.

A thesis submitted for the Degree of Doctor of Philosophy

Department of Mathematics & Statistics

University of Surrey,

Guildford, Surrey, GU2 7XH.
This thesis is dedicated to the memory of my father, John Brendan
Hillary.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents</td>
<td>i</td>
</tr>
<tr>
<td>List of Figures</td>
<td>iv</td>
</tr>
<tr>
<td>List of Tables</td>
<td>viii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>ix</td>
</tr>
<tr>
<td>Summary</td>
<td>xii</td>
</tr>
<tr>
<td><strong>1 Introduction</strong></td>
<td>1</td>
</tr>
<tr>
<td>1.1 Modelling Planktonic Populations</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Turbulent Advection</td>
<td>8</td>
</tr>
<tr>
<td>1.2.1 Turbulent diffusion</td>
<td>9</td>
</tr>
<tr>
<td>1.3 Alternative models of turbulence</td>
<td>11</td>
</tr>
</tbody>
</table>
1.3.1 Chaotic advection ........................................ 12
1.3.2 Synthetic turbulence .................................. 13
1.4 Patch Dynamics .......................................... 15

2 Chaotic advection and patchiness: An example case using a reduced Lagrangian model for Langmuir circulations 20

2.1 Turnstile lobe dynamics ................................. 25
2.1.1 Lobe dynamics and Lagrangian chaos .............. 26
2.2 Melnikov analysis and particle mixing .................. 27
2.3 Chaotic transport versus diffusion ....................... 35
2.4 Conclusions ............................................... 37

3 Motile, spheroidal organisms in small scale shear flows 40

3.1 Linear Analysis: Equivalent diffusion tensors ........... 44
3.2 Linear & Non-Linear Analysis: Different diffusion tensors 49
3.3 Conclusions ............................................... 58

4 Effects of excitability, inertia and turbulent advection 61

4.1 Turbulent advection of inertial organisms ............... 63
4.1.1 Viscous and inertial effects ........................... 63
4.1.2 Derivation of the long-time equation of motion ...... 65
4.1.3 Aggregation of non-neutrally buoyant organisms ...... 66
4.1.4 Spectral characteristics of inertial particles in turbulence 67
4.1.5 Synthetically turbulent flow .......................... 72
6.3.2 Deterministically related patches .............................................. 157
6.4 Making use of the constructed meta-population time series ............... 162
  6.4.1 Predictive estimation .......................................................... 163
6.5 Applications to model fitting of planktonic dynamics ......................... 173
6.6 Discussion on the Effects of System Noise .................................. 176
6.7 Conclusions .............................................................................. 179

7 Conclusions & Further Work ................................................................ 182

8 Appendix ....................................................................................... 194

  8.1 Ergodic theory of smooth dynamical systems ..................................... 194
     8.1.1 Basic Measure Theory .......................................................... 195
     8.1.2 Liapunov Exponents ............................................................. 197
     8.1.3 Practical computation of Liapunov exponents ......................... 198
     8.1.4 Reorthogonalisation .............................................................. 199
  8.2 Numerical integration methods ....................................................... 200
     8.2.1 ODE integration ................................................................. 200
     8.2.2 PDE integration ................................................................. 201
List of Figures

1.1 Dynamics of an NPZ model with varying the closure rate .... 7

2.1 Steady convection rolls in Langmuir circulations. ............ 22
2.2 Retaining and mixing regions for the reduced Langmuir model . . . 23
2.3 Orbits under the action of the Poincare map of the reduced Langmuir model .................................................. 24
2.4 Sketch of how lobes are formed from the tangling of the stable and unstable manifolds. ........................................... 26
2.5 Prediction of the lobe non-escape area as a function of $\omega$ from the Melnikov analysis. .............................................. 33
2.6 Numerical plot of the non-escape area as a function of $\omega$. ....... 34
2.7 Diffusivity/advection ratio analytically obtained from the Melnikov analysis ............................................................ 36
4.1 Energy spectra for inertial particles in turbulent flow .......................... 69
4.2 Snapshot of the "frozen" turbulent velocity field ................................. 76
4.3 Schematic representation of the excitable nature of the phytoplankton population .......................................................... 79
4.4 Propagation of the phytoplankton excitation wave via diffusion and turbulent advection .............................................. 82
4.5 Bloom formation in the extreme inertial regime .................................. 84
4.6 Snapshots of the phytoplankton population during the oscillatory bloom .......................................................... 85
4.7 Oscillatory bloom in the mean concentrations of phytoplankton and zooplankton .................................................. 86
4.8 Spatiotemporal evolution of the phytoplankton population for the suprathreshold vortex model ........................................ 87
4.9 Oscillatory bloom for the reduced vortex model ................................. 88
4.10 Areas in (\(\lambda_p, \lambda_z\)) space in which we will won't (II/I) see excitation due to inertial separation ........................................ 89
4.11 Complete view of one period of the oscillatory bloom ....................... 92

5.1 Maximal normal Liapunov exponent, \(\lambda_{\perp}^{\text{max}}\), versus the patch coupling, \(\epsilon\) .................................. 105
5.2 Attractors in the \((N_1, N_2)\) plane before and after the blowout bifurcation 106
5.3 \(N_i(t) - N_{i+1}(t)\), for \(i = 2, 3, 4\) and 5 after any transient behaviour has gone. The synchronised solution is unstable throughout the whole lattice ........................................ 115
5.4 A schematic representation of the riddling bifurcation. .............. 120
5.5 Neutral normal hyperbolicity curve in $(\epsilon_1, \epsilon_2)$ space. .......... 124
5.6 Neutral normal hyperbolicity curve in $(d, \epsilon)$ space. ................. 125
5.7 The three Liapunov exponents of $A$ as we vary the closure parameter, $d$. ............................................. 126
5.8 Maximal Liapunov exponent around the chaotic regime, $d = 0.0142$. ... 127
5.9 Phase-synchronised patch dynamics. .................................... 128
5.10 The chaotic Rössler attractor in $(x_1, x_2)$ space. ...................... 132
5.11 Attractors in the $(N_1, N_2)$ plane for the unsynchronised and synchro-
nised motions in the drive response system ............................... 133
5.12 Neutral normal hyperbolicity curve in $(d, \epsilon)$ space for the quadratic
closure model. ....................................................... 137
5.13 Maximal Liapunov exponent around the chaotic regime, $d = 0.5$ for
the quadratic closure model. ........................................ 138
6.1 Reconstruction of the NPZ dynamics using phytoplankton as the mea-
measurement function. .............................................. 145
6.2 Time series of two patch nutrient measurements from a system of
three, diffusively coupled NPZ systems, (6.8). .......................... 147
6.3 Typical phase portraits for coupled patches ............................... 151
6.4 Continuity, inverse continuity and homeomorphism confidence statist-
tics for the four scenarios shown in figure (6.3). ....................... 153
6.5 Homeomorphic confidence statistic for case (ii) as we increase the amount of available data, \( N \) ........................................... 154

6.6 Homeomorphic confidence statistic for case (iv) as we increase the amount of available data, \( N \) ........................................... 155

6.7 Average inter patch error for the sample mean and weighted linear average cases ........................................... 160

6.8 Average inter patch error for the sample mean and weighted linear average cases ........................................... 162

6.9 Original and fitted time series for a chaotic data set. ........................................... 168

6.10 Fitting error between original and fitted time series for a chaotic data set. ........................................... 169

6.11 Phase separation of original and fitted time series for a periodic data set. ........................................... 170

6.12 Propagation of small errors when attempting to predict chaotic time series. ........................................... 171

6.13 The fitting error between the data sets \( D_1 \) and \( D_2 \). ........................................... 176

6.14 The fitting error between the data sets \( D_1 \) and \( D_3 \). ........................................... 177

6.15 The fitting error between the data sets \( D_1 \) and \( D_4 \). ........................................... 178
List of Tables

1.1 Relative dominance of swimming & turbulent diffusion for a range of spatial scales. ........................................ 10

5.1 Default parameter values for the NPZ model. .............................. 101

5.2 Ergodic bursting measure for various lattice points after locally and globally originating blowout bifurcations .................. 116

6.1 Cumulative error when varying the number of data points and the number of basis centres. ......................................... 172
Acknowledgements

Thank you firstly to my supervisor Dr. Martin Bees, without whose help none of the following would be possible and for making my PhD. both profitable and enjoyable at the same time.

I also would like to thank all the staff at the University of Surrey (Mathematics & Statistics). The Department is one where help is easy to come by and thankfully so is a sense of humour. In particular, thanks to Miguel, Alex, Alexandra, David, Tassos, Kyriakos, Samer, Sarker, Leanne, Francis, Debbie, Georg and Stella and all the other postgrads and postdocs in the department over the years.

I am very grateful to the EU for two grants from the Improving Human Potential (IHP) programme that allowed me to visit the Universitat de Barcelona. These two visits proved invaluable and most enjoyable. Many thanks to Prof. Francesc Sagues, Dr. Ramon Reigada, Dr. Jose Maria Sancho and all the people I met in Barcelona who made me feel very welcome.
Many thanks to the EPSRC who supplied the grant for the work contained in this thesis.

Much gratitude and love to my family, Mum, Jude and Oma for all their patience and support. All my love to Clare for all the fun we've had in such a short time.

Finally, much love to my dad, Brendan, who sadly never lived to see this day finally come. I will miss you always.
The main aim of this project has been to investigate how turbulence and a patchy environment can affect the dynamics of plankton. In the first chapter, a brief introduction to the subject, both the methods and models we use and some of their behaviour in relation to observed phenomena, is given. In the second chapter, a reduced model for Langmuir circulations [12, 55] is used as a paradigm for chaotic advection of planktonic species. The patchiness of plankton due to such advection is discussed along with the effects of swimming/turbulent diffusion and how we can sometimes gain an analytical hold on the transport of organisms, using Melnikov analysis [19, 109], in certain flows and we extend the results of previous work [19] in this area.

In Chapter 3, the possibility of pattern formation of swimming, spheroidal organisms in a simple, steady shear flow is investigated. The forms of the diffusion tensors can be more complicated than for isotropic diffusion due to the spheroidal...
shape and behaviour of the organisms [11, 41, 81]. For organisms with identical diffusion tensors it is shown that the shear can have a destabilising effect, due to the non-standard diffusion tensors, contrary to the analysis in [24] for simple Fickian diffusive behaviour in the reacting particles/organisms. For the case of non-identical diffusion tensors, a weakly non-linear extension is derived about the zero wavenumber, no-flow instability. The resulting amplitude equation suggests that zero may not always be the critical wavenumber and that criticality of the bifurcation is highly dependent on the non-linear reaction terms. The work in this chapter generalises the results of Doering & Horsthemke [24] and Spiegel & Zaleski [96] to swimming, spheroidal organisms.

In Chapter 4, the effects of turbulence and any inertial effects (buoyancy or density differences) is considered with regard to the initiation and subsequent propagation of phytoplankton blooms. A stochastic model for synthetic turbulence is employed [62] and the derivation of the steady equation of motion for inertial organisms in homogeneous turbulence is derived, as in Reigada et al. [91], and the different aggregation zones for inertial organisms are revealed.

An analytical expression for the energy spectrum of inertial organisms immersed in steady, isotropic turbulence is derived, using a wavenumber dependent diffusion approximation to the non-linearity in the Navier-Stokes equations first seen in the work by Powell & Okubo [88]. The expression yields results consistent with previous work [53, 54] on 2D turbulent energy spectra. It is seen that any inertial differences between the phytoplankton and zooplankton act to try and separate the two species into different areas of the flow, thus aiding the occurrence of a bloom.
A small sized oscillatory bloom is observed with period of order 1 year suggesting that inertial/density differences [31, 65] between the plankton and the surrounding ocean can influence the formation of blooms.

Using a continuum contact rate expression, based upon the inertial driven “relative” fluxes of the phytoplankton and zooplankton, and other well known spatial statistics the temporal changes in the contact rates and the spatial distributions of the planktonic species are studied. As expected, the bloom increases the contact rate of predator and prey and even when the spatial correlations suggest it would be decreasing due to inertial separation. Our continuum expression detects this as it is a dynamic measure, based on where the plankton are moving to, not where they are.

The effect of a patchy environment is studied first, in Chapter 5, for the situation where we know explicitly the dynamics of the plankton. We take a spatially discrete, coupled oscillator approach to the patchy dynamics of plankton. Such systems show differing degrees of synchronised behaviour [6, 17, 50], depending on the strength of the coupling between the patches. Using a well known NPZ (Nutrient-Phytoplankton-Zooplankton) [26, 27, 102] model for the patchy dynamics, we demonstrate that spatio-temporally varying patch dynamics can arise from a variety of sources, including low levels of coupling and even small levels of system noise.

It is conjectured and numerically demonstrated that, for a general system of diffusively coupled oscillators, if the coupling is low enough to see unsynchronised dynamics in a localised area, we see the emergence of non-synchronous dynamics
in the whole patch system. Large scale forcing has been suggested as a mechanism that can induce the synchronisation of seemingly un-coupled populations [17, 32]. By adapting a test for the detection of so called generalised synchronisation first presented in Abarbanel et al. [1], the non-linear mechanism behind this forced synchrony is revealed, for both regular and chaotic dynamics. Also, the role of chaos is discussed in a general setting for patchy ecosystems. At least for two coupled NPZ systems with different dynamical models, it appears that chaotic dynamics seems to facilitate the formation of smooth, generally synchronised [94] dynamics between the patches more (overall) than regular dynamics. These results are discussed in the context of observed spatially coherent chaotic oscillations such as the celebrated Canadian hare-lynx [17, 49, 52] data set.

The case where the dynamics of the patches is not known and, consequently all we might have is time series data is studied in Chapter 6. Given only measurements of the dynamics of some patchy population, we present a way of trying to deal with the patchy data in a more rigorous framework as mention of the inherently heterogeneous environments being measured [46, 90] is made but is mostly ignored.

A method of first distinguishing independent patch dynamics from deterministically related dynamics is presented based on the algorithm first seen in Pecora et al. [78]. Given this deterministic bond between the patches we then set about creating a meta-population [90] time series representing the collective dynamics of the population. This new time series is constructed so as to try and preserve as much of the individual dynamics as possible. Using a non-linear prediction algorithm [18, 22, 87, 99] ideally suited to possibly short data sets, we suggest this new
time series can be used to improve short term predictions of general trends in the dynamics and also for the purposes of model fitting. Particular attention is given to using algorithms that can be applied to relatively short data sets, often a problem in studying time series data of ecosystems.
Plankton play a major role in the global carbon cycle and are at the bottom of a very long food chain. Also various planktonic species are food for many species of fish larvae such as cod, which constitutes an ever dwindling, yet economically important human food resource. These are just a few of the many reasons why much research centres on the understanding of the processes that involve planktonic communities. For instance blooms of phytoplankton, which consist of a possibly short lived yet often extreme increase in the phytoplankton population, can have devastating effects on the local fish populations (red tides) [57] but also help dictate the year class strength of a particular fish stock, by this we mean the abundance of fish relative to previous years. Fish eggs spawned close to the location of a phytoplankton bloom have a readily available food source when they hatch into larvae [57] and these excellent growth conditions can subsequently influence the number of juvenile fish that recruit into the fish stock. One of the most widely observed phenomena is the “patchiness” of plankton, put more formally that plankton populations vary spatially
as well as temporally [33, 61, 66]. Also, the qualitative nature of this patchiness is very variable, from regular to complex (possibly chaotic) patterns [66, 71] to travelling fronts [16, 33] suggesting that many different processes interact in various ways to produce such a variation of patterns.

There is also a wealth of spatio-temporal scales to consider as well. For a more complete review of how the oceanographic and biological space and time scales overlap, the interested reader is referred to the articles by Folt & Burns [31] and Steele & Henderson [103]. The basic ideas are that, for length scales of around the order of 10m or less and time scales in minutes or less, the dominant influence is the individual behaviour of the plankton themselves. Factors such as swimming, diel vertical migration and food location can cause micro-patchiness [23] at the smallest scales and are more influential than say turbulence or temperature/salinity variations. As we move to length scales of kilometres and above and time scales on the order of days up to possibly years, the physics of the ocean begin to dominate the proceedings. From the measurements of Okubo [73], the turbulent “diffusion” at these scales is much higher than any swimming “diffusion” and so individual effects become much less important.

1.1 Modelling Planktonic Populations

The first step to modelling such processes would normally begin with developing a model of the basic population interactions of the planktonic species. There are individual based models, which concentrate on the behaviour of each plankter with
regards to its local environmental state [86, 111]. These type of models are more suited to the smaller spatio-temporal scales as this is where individual effects are most prominent. While patchiness and turbulence occur at these smaller scales [23], the work in this thesis concentrates more on the larger scales so another methodology is required.

In the late seventies/early eighties, work such as that by Fasham [29] and Steele and Henderson [101] concentrated on trying to model the collective dynamics of the populations (temporally, spatially or both). Such an approach can either be discrete or continuous in time but, for example, the time scale of changes in phytoplankton populations can be of the order of hours while changes in some species of zooplankton are more of the order of days [57]. Discrete time models cannot fully capture, and can even miss, such behaviour so a continuous in time set up is generally deemed more appropriate. These latter types of ODE/PDE models are the ones we will be concentrating on. Such dynamical system type models allow for more general ideas, such as different advective scenarios [2, 16, 71, 92], to be explored and also apply to a more general type of ecosystem model [43] than an individual based model would as mean field type models concentrate on generic population trends, as opposed to representing an individual plankter’s behaviour.

There are a variety of models but most concentrate on representing either the biomass of nutrient, phytoplankton and zooplankton (NPZ) [26, 27, 102] or simply the phytoplankton and zooplankton concentrations (PZ) [60, 66, 85, 107] with the nutrient included in the phytoplankton equation. More trophically complex models do exist and an excellent study of these can be found in [112]. The general structure
of an NPZ model can be represented as follows:

\[
\begin{align*}
\frac{dN}{dt} &= -f_1(N, P) + f_2(N) + f_3(P, Z), \\
\frac{dP}{dt} &= \alpha f_1(N, P) - f_4(P) - f_5(P, Z), \\
\frac{dZ}{dt} &= \beta f_6(P, Z) - f_6(Z),
\end{align*}
\]

where \( N, P \) and \( Z \) represent the biomass of the nutrient, phytoplankton and zooplankton respectively.

The uptake of nutrient by phytoplankton is defined by the function \( f_1 \) while mixing/sinking and excretion regeneration are contained in \( f_2 \) and \( f_3 \), respectively. As an example, [27] incorporated both nutrient saturation and phytoplankton self-shading into the term \( f_1 = aN/(e+n)(b+cP) \). The excretion regeneration terms are proportional to the trophically relevant predation functions \( f_5 \) and \( f_6 \) while sinking and/or mixing is generally assumed to be linearly related to the biomass of the species.

In the phytoplankton equation, \( f_4 \) incorporates both sinking/mixing of phytoplankton to and from the upper mixed layer and both of which are usually assumed to relate linearly to the biomass. The function \( f_5 \) models predation by zooplankton and can been modelled in a variety of forms, from a simple Lotka-Volterra term to Michaelis-Menten [60] to Holling type responses [26, 27, 107]. Finally, \( f_6 \) is the higher predation or “closure” term (\( \alpha, \beta \) are efficiency parameters) which represents zooplankton predation by species such as fish. Again, many types of term are used such as simple linear and quadratic terms [27] to the Holling type III functional response [60]. A good treatment of the different functional forms used in such models can be found in [26, 27, 102].
For PZ models we have a slightly simpler template and

\[ \frac{dP}{dt} = g_1(P) - g_2(P, Z), \]
\[ \frac{dZ}{dt} = \gamma g_2(P, Z) - g_3(Z), \]  

(1.2)

with \( g_1 \) usually being a logistic growth type term. Similar to (1.1), \( g_2 \) represents zooplankton predation of phytoplankton and \( g_3 \) the predation of zooplankton by higher predators and \( \gamma \) is an efficiency parameter. Popular choices for the functional form of \( g_2(p, Z) \) have been the Michaelis-Menten form \([60]\), \( g_2(p, Z) = ZP/(\mu + P) \), or the Holling type III form \([107]\), \( g_2(p, Z) = ZP^2/(\mu^2 + P^2) \). The Michaelis-Menten form is a monotonically increasing function while the Holling type III function gives a sigmoidal predator-prey response curve (with \( \mu \) measuring the prey saturation rate). For the closure term, \( g_3(p, Z) \), well used forms are \( g_3(Z) = dZ, dZ^2, \) and \( dZ^2/(\nu^2 + Z^2) \), with \( \nu \) again measuring the saturation level of zooplankton, with respect to the higher predator, and \( d \) is the rate of predation. For example, the linear and quadratic closure terms model different responses by the higher predators to the amount of prey present but differ in that the quadratic term decreases quicker as the amount of prey gets very small but also increases faster when there are more prey to be found. The linear response is less influenced by the amount of available prey. Terms such as the Holling type-III term model a quadratic response for low-levels of prey which saturates (depending on the value of \( \nu \)) as the prey becomes more abundant. These particular terms have been seen to have a significant effect on the nature of the dynamics of the system for both the pure ODE systems \([60, 107]\) and spatially extended systems \([16, 61, 66, 71]\).

Models of the form shown in (1.2) have displayed a variety of interesting proper-
ties from multistability, limit cycles and activator-inhibitor behaviour [60] to properties such as excitability [107]. Also, NPZ models have been shown to exhibit various equilibria, limit cycles, chaotic oscillations and even complex attractor coexistence [26, 27, 102] for a variety of forms of the function \( f_6 \). Figure (1.1) shows the change in the dynamics for the NPZ model, with linear closure, first seen in [26] when varying the closure rate. We see that ever more complex oscillations can arise as we increase the closure rate \( d \), resulting in chaotic oscillations (arising from the usual period-doubling [25] route).

When combining these types of behaviour with diffusion and turbulent advection one can see the pattern forming potential and we shall address several of these mechanisms in this thesis.

The advantages of using these PZ and NPZ models is that they represent a low-dimensional way of modelling generic aspects of the various species' dynamics. This simplicity will be important later on in the thesis as we investigate the dynamics of many coupled plankton populations and also link a model for the spatially homogeneous phytoplankton-zooplankton dynamics to a stochastic model for turbulent flow. If the models are trophically to complex then the numerical and theoretical analysis of the systems becomes considerably more difficult. Also, with so many parameters and influential processes, it become increasingly difficult to say what process is causing the observed behaviour.

This simplicity is also one of their obvious limitations. In real situations, there are many different types of phytoplankton; from small algal cells \( \mathcal{O}(\mu m) \) up to larger diatoms which are around the mm scale in size [57]. Also, the number of
Figure 1.1: Dynamics of the NPZ model in [26] in the \((P, Z)\) phase-plane under variation of the closure rate. Clockwise from the top left; \(d = 0.14\), \(d = 0.1415\), \(d = 0.14185\) and \(d = 0.142\).

Species of zooplankton is huge and they have a varied array of sizes, shapes and behaviour [31] suggesting that simply grouping them into one species is sometimes an oversimplification. Intra-trophic effects such as cannibalism [46] are known to occur as well as the possibility of intra-trophic competition. Both these topics have been touched upon, from a modelling sense, in the work in [84, 85]. As long as we are mindful of both the applicability and results arising from the use of these low-dimensional plankton models, they represent a relatively simple way of approaching...
1.2 Turbulent Advection

Turbulence is an integral factor in all studies of oceanic processes. However, a complete theory of turbulence has still not been achieved. During the forties, fifties and sixties, the work of Kolmogorov [51], von Karman [108] and Kraichnan [53, 54] began to reveal the complex transfer of energy and small scale structure in turbulent flow. For 3D turbulence, we have the celebrated Kolmogorov $k^{-5/3}$ decay in the energy spectrum for the inertial sub-range (in between the energy generating large scale eddies and the smaller scales where viscosity takes precedence) and Kraichnan's [53] corresponding decay term in $k^{-3}$ for 2D turbulence.

In terms of simulating turbulent flows, the natural starting point would be to turn our attention to the Navier-Stokes equations:

$$
\partial_t U + (U \cdot \nabla) U = -\frac{1}{\rho} \nabla P - gk + \nu \nabla^2 U,
$$

$$
\nabla \cdot U = 0,
$$

(1.3)

where $U$ is the fluid velocity, $P$ is the pressure and the parameters $\rho$, $g$ and $\nu$ represent density, gravity and dynamic viscosity, respectively. We can non-dimensionalise (1.3) via the following re scalings;

$$
U = U U, \quad X = L x, \quad Y = L y, \quad Z = L z, \quad t = \frac{L^2}{U} \quad P = \rho U^2 p,
$$

where $U$ and $L$ are representative velocity and length scales, respectively, of the
flow. We now obtain the following dimensionless form of (1.3):

\[
\partial_t u + (u \cdot \nabla) u = -\nabla p - g k + \frac{1}{\text{Re}} \nabla^2 u, \\
\nabla \cdot u = 0,
\]

(1.4)

with the dimensionless parameters \( g = gL/U^2 \) and the so-called Reynolds number \( \text{Re} = UL/\nu \). As first observed by Reynolds [93] in experiments, as the parameter \( \text{Re} \) is increased, the laminar solutions of (1.4) can become unstable to small perturbations. The resulting flows can be highly irregular (turbulent) and a review of the mechanisms involved in these transitions to turbulence, as well as a review of the spectral properties touched upon here, can be found in [70]. Direct numerical simulation (DNS) of (1.4) is computationally intensive and the existence and uniqueness of solutions, even in simple 3-D geometries, has not yet been established. The rapidly increasing complexity of solutions, as the Reynolds number \( \text{Re} \) is increased [70], also means we are usually restricted to relatively low Reynolds number flows, which in turn limits the scales we can investigate.

### 1.2.1 Turbulent diffusion

Research has also centred on the general dispersive nature of turbulent flows both from a theoretical [54, 88] and experimental [73] viewpoint. In real life computations, the resolution of the computational grid will be finite. Given some fixed grid spacing \( \Delta \), the largest wavenumber we can incorporate is \( k^* = 2\pi/\Delta \). However, with using say the Kraichnan [53] energy spectrum, \( E(k) \),

\[
E(k) \propto k^3 \exp \left( -\frac{k^2}{k_0^2} \right),
\]

(1.5)
Table 1.1: Relative dominance of swimming & turbulent diffusion for a range of spatial scales.

<table>
<thead>
<tr>
<th>Spatial scale</th>
<th>$O(m)$ or less</th>
<th>$O(10m)$ - $O(100m)$</th>
<th>$O(km)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turbulence ($D_E$)</td>
<td>$D_E \leq 5 \times 10^{-4}$</td>
<td>$D_E \in (5 \times 10^{-3}, 10^{-2})$</td>
<td>$D_E \geq 0.5$</td>
</tr>
<tr>
<td>Swimming ($D_S$)</td>
<td>$D_S \in (10^{-4}, 10^{-2})$</td>
<td>$D_S \in (10^{-4}, 10^{-2})$</td>
<td>$D_S \in (10^{-4}, 10^{-2})$</td>
</tr>
</tbody>
</table>

The spectrum may fall away to practically nothing well before the largest wavenumber $k^*$, meaning we are missing the dispersive effects on any particles/organisms below this scale. One way around this is to use the relationship derived by Okubo [73], which uses measured dye dispersal to give an empirical equation for the horizontal "eddy" diffusivity, for a given length scale. It was seen that the eddy diffusivity, $D_E$, was proportional to $L^{1.15}$, where $L$ is the length scale. This length scale dependent diffusivity can then be used to model turbulent dispersion at the relevant/necessary scales and has been utilised in [16, 66, 71, 92] and will be used in chapters 3 & 4 in this thesis. In Table 1.1 we give a very rough idea of the spatial scales where the relevant dispersive behaviours; swimming and turbulent, are dominant, comparable or too small to have an effect. The swimming "diffusivity", $D_S$, was calculated by first using the range of mean swimming velocities, $V_S$, from the work on algal cells by Hill & Häder [40]. Their data suggested that $V_S \in (10^{-2}, 10^{-1})$ms$^{-1}$, and that given some reorientation time $r_r$ of $O(1s)$, the swimming diffusivity can be approximated by the expression $D_S = \frac{V_S^2}{r_r}$, which leads to the range of values given in Table 1.1.
We can see now that, at the smaller scales of metres or less, that individual swimming behaviour is likely to dominate the dispersive nature of the plankton while as we increase the spatial scale, the influence of turbulent dispersion becomes more important until, for scales above a kilometre, the swimming nature of the plankton is virtually irrelevant.

1.3 Alternative models of turbulence

For many years, people have been looking into alternative methods of simulating turbulent flows to qualitatively represent the mixing nature of turbulent flow, while sometimes sacrificing either the non-linearity of the Navier-Stokes [62] equations or some of the statistical properties. Often, there are well defined flow features which dominate the large scale nature of the circulation such as the north Atlantic drift, with more complex (possibly turbulent) motion occurring at smaller spatial/temporal scales. For example, shear flows occur frequently in the ocean and some papers have focused on a simple, imposed shear drift with any smaller scale turbulence effects being modelled by a diffusive term. This situation was looked at in the general form in [24, 96], with regards to shear driven instabilities. With plankton having varying degrees of motility, shape and density [31, 65], it might be possible to see these shear driven instabilities at the smaller spatial scales. This could happen because turbulent diffusion doesn’t dominate, meaning there could be distinct differences in the dispersive behaviour of the interacting plankton. This can lead to the formation of Turing type patterns, especially in the presence of advect-
tion as, in a chemical system model exhibiting activator-inhibitor properties [95], this radically decreased the difference needed between the diffusion coefficients of the reactants to see an instability. The micro-patchiness of planktonic organisms has been investigated with regards to motile behaviour [68, 69] and shear effects on both diffusion of passive tracers [72] and on the diffusion of motile, spheroidal organisms [11, 42]. This work suggests that shear can accentuate even small differences in swimming or shape characteristics. In [16], the breaking of excitation waves by simple shear flows was suggested as a possible route to plankton patchiness so, specifically when considering large scales or the smaller scales where swimming becomes more important than turbulent diffusion, it can be simpler to consider this type of flow dichotomy. By this me mean where the larger scale features are modelled with a simpler advection term and smaller scale effects are modelled via the diffusive term.

1.3.1 Chaotic advection

A simpler method, than direct simulation of the Navier-Stokes equations, for modelling complex flows has been the use of chaotic advection [19, 71, 76] (for a thorough review of the topic the interested reader is referred to the paper by Aref [5]). In [12], advection by a reduced model for Langmuir [55] circulations was considered as a mechanism driving plankton patchiness. Also, in [105] a chaotic model for estuarine flow was employed in a study of the transport of pollutants down rivers. In this thesis, we use the chaotic model for Langmuir circulations from [12] to demonstrate that complex flows can greatly influence the aggregation and contacts rates of
plankton and that, for a certain class of chaotic flows, we can gain some analytical insight into the transport of plankton and how this may relate to the environmental conditions.

When we have some unbounded, incompressible chaotic flow, \( u \) (possibly derived from a Lagrangian viewpoint as in the Langmuir model), then the continuum system

\[
\partial_t c + u \cdot \nabla c = D \nabla^2 c, \tag{1.6}
\]

with periodic boundary conditions is said to exhibit Lagrangian chaos, where \( c \) represents the concentration of the passive scalar and \( D \) is the diffusivity.

### 1.3.2 Synthetic turbulence

In [62], an algorithm for developing "synthetic" turbulence was proposed. This type of stochastic velocity field [54, 70] represents a departure from the non-linear approach of the Navier-Stokes equations to developing a random flow with the relevant statistical properties (spatial correlations and the energy spectrum) being effectively imposed upon the equations of motion. This method basically begins by representing the spatio-temporal evolution of the stream function [70] as a Langevin [36] equation such as the following,

\[
\frac{\partial \eta(r,t)}{\partial t} = \nu \nabla^2 \eta(r,t) + \xi[\lambda^2 \nabla^2] \nabla \cdot \zeta(r,t), \tag{1.7}
\]

where \( \nu \) is the kinematic viscosity, \( \lambda \) is a length scale parameter and \( \zeta(r,t) \) is a Gaussian white noise. The major advantage of this type of method is that, while being easy to implement (relative to the integration of (1.4)), the statistical properties of the flow such as the energy spectrum can be defined through the choice of
the stirring operator, $Q[·]$. This is a kind of reverse approach in that the equation of motion is constructed so as to contain the suitable statistical properties required of the flow. Another advantage is that we can also readily control the relevant length and time scales of the flow using stochastic calculus [62]. This makes it ideal for our purposes as we want a realistic model for turbulence but also one where we can exert strict control over the time and length scales. We will thus be incorporating empirical data into the full reaction-diffusion-advection simulations in chapter 4.

A final point to consider is that both zooplankton and phytoplankton have different viscous and inertial properties [31, 65] to the surrounding oceanic medium. In [63], the equation of motion for a spherical particle in a non-uniform flow field was derived and, using this equation, it was seen in [91, 97] that heavy and light particles/organisms will accumulate in different areas of the flow which could result in non-uniform spatial distributions and other possible effects depending on the type of reaction dynamics considered. Phytoplankton bloom occurrence, patchiness and predator-prey contact rates are all factors likely to be influenced by such processes.

Also, turbulent energy has been shown to be generated, certainly in 2D turbulence which we consider, inside the large scale eddies. With the biased aggregation of inertial organisms being highly dependent on the vorticity [91], there may well be some effect on the energy spectra of inertial organisms immersed in turbulent flow. This idea was touched upon, from the point of view of spectral changes due to reaction dynamics, by Powell & Okubo [88] for passive tracers and is something we wish to explore in this thesis when we consider non-passive inertial organisms.
1.4 Patch Dynamics

The analysis of plankton patchiness is not finished when the patches have formed. In most cases there will be some form of coupling between the patches. This coupling could appear in many guises such as diffusion [66] or cross-predation/higher predation similar to the intra-trophic ideas investigated in [84, 85]. Also, in many cases a (spatial) continuum approach may not be advisable so, to have a broader definition of patchiness, it may be better to consider a spatially discrete set up. As an example, the coupling between patches may be a result of higher predatory effects such as highly mobile fish schools [61]. In this case the patches may be separated by large (relative to the size and motile ability of the patch inhabitants) distances, something not incorporated into a spatially continuous model. Also, in a continuum sense, simple isotropic diffusion does not allow one to vary the directional strength of the patch coupling. A spatially discrete set up also allows for easier analysis, both numerical and analytical, as it requires solving a system of ODEs as opposed to solving a complex, non-linear PDE.

Assuming that we can represent the planktonic populations as continuous-in-time variables, a 1-D patch lattice can be represented as the following:

\[
\dot{S}_i = G_i(S_i) + \sum_j \Gamma_j(S_j, S_i), \quad i \neq j.
\]  

(1.8)

Here, \( S_i \) represents the populations of patch \( i \), the vector valued functions \( G_i(\cdot) \) are the reaction dynamics and the patch coupling is represented by the \( \Gamma_j(\cdot) \). We can further break this into two distinct scenarios:

1. We choose explicit forms for the \( G_i(\cdot) \) and \( \Gamma_j(\cdot) \).
2. We know nothing of the form of the $G_i(\cdot)$ and $I_j(\cdot)$ and all we have is experimental data.

In the first case, we have a system of (possibly many) coupled oscillators. Such systems are known to exhibit various forms of synchronisation [6, 34, 77, 82]. This could be phase synchronisation [83], where all the oscillators lock phases. This particular type of behaviour has been observed in nature, the often quoted Canadian hare-lynx [17, 49, 52] data shows strong phase similarity over a very large spatial sampling range. Transitions from synchronous to non-synchronous dynamics have been seen in a variety of systems such as electrical [8] and biological [17] systems and for a variety of different mechanisms such as low coupling strength, patch parameter differences and low-level system noise. It is conceivable that such a loss of synchronicity in the smaller scale patchy dynamics could result in larger scale spatio-temporal variability and this idea is investigated, for a variety of different scenarios in chapter 5.

Full synchronisation is where the oscillators' dynamics are asymptotically identical but there is a more relaxed form called generalised synchronisation [1, 78, 94] where the oscillators are asymptotically related via some suitable continuous (possibly differentiable) function. This means that the collective and individual dynamics could evolve in a number of different ways, depending on the form and strength of the coupling.

For the second case, all we have is experimental measurements of the populations and in the extreme case we may only have data on one of the patch inhabitants.
This can be overcome thanks to the work of Takens [106], who proved that one could embed the dynamics of a single measurement function of one system variable into a suitably higher dimensional space. Basically it meant that one could reconstruct, under certain conditions, the dynamics of a possibly high-dimensional dynamical system using just one of the measured variables.

Given these time series measurements from various patches it would be helpful to know if the dynamics of the patches are deterministically linked (generally synchronised) or if they evolve in an unrelated manner. In the case of the previously mentioned Canadian hare-lynx data, the populations are in generalised synchronisation via phase synchronisation. Also, Fox et al. [32] observed high levels of synchrony in the recruitment time series of plaice (Pleuronectes platessa L.) around the British Isles. This was detected using a variant on the correlation of the time series but this only tests for some level of linear relationship between the two time series. As in the hare-lynx data, the relationship could be highly non-linear but still deterministic and the correlation approach may yield low results while a strong bond between the two time series may be present. This suggests a more non-linear approach to patchy time series analysis might be advantageous.

Being able to distinguish the two types of behaviour would be advantageous, for a number of reasons. Firstly, when averaging the individual behaviour to obtain a view of the collective system we should perhaps use a different approach for the uncorrelated patch evolution than for the generally synchronised system. The need for a more rigorous approach to create a representative time series leads on to the second reason. There has been recent work with regards to fitting ecological
models to obtained planktonic data [20, 46] to see what type of density dependent processes are possibly at work. Also, model fitting forms the backbone of many stock assessments of fish stocks which are known to reside in sometimes distinct populations inside the management/assessment area [90]. Given some discretely measured data of the population of plankton $s_n = (p_n, z_n)$, we might want to fit this data to some specific model for the plankton populations, e.g. $S_n = (P_n, Z_n)$ and $S_{n+1} = g(S_n)$. If $\theta$ denotes some parameter(s) of the explicit model $g$ then to maximise the goodness of the fit we would ideally want to minimise the error between the data and the model variables. Put more formally, let the (time discrete) residual $r_n = s_n - s_n$. We would like to calculate the parameter vector $\theta$ which minimises the least squares error, $\varepsilon^2(\theta)$, where

$$e^2(\theta) = \sum_{n=1}^{N} r_n^T V^{-1} r_n = \sum_{n=1}^{N} \left( \frac{(P_n - p_n)^2}{\sigma_p^2} + \frac{(Z_n - z_n)^2}{\sigma_z^2} \right), \quad (1.9)$$

where $V$ is the variance [36] matrix which re-scales the error in the event of highly variable data as many minimisation algorithms [89] can be highly affected by badly scaled variables. Even if patches are generally synchronised, sampling data from different ones (effectively random spatial sampling) or just one patch repeatedly (a fixed measurement station perhaps) may yield a dataset that is not completely representative of the meta-patch dynamics. Also, simply taking the sample mean of the patch populations could bias the dynamical behaviour of the patch system. What we would like to consider is a way of systematically differentiating the nature of the patch dynamics and then creating a representative meta-population for the purposes of short-term predictions and model fitting. The method should also be
able to deal with issues such as phase/delay effects, gaps in the data and so on. Using the idea of delay reconstruction [25, 99, 106] of a system's dynamics from a single species measurement and powerful predictive techniques [18, 22, 87, 99] from the non-linear theory of time series analysis we can try and factor these effects into the measurement/prediction/fitting processes.
Chaotic advection and patchiness: An example case using a reduced Lagrangian model for Langmuir circulations

A popular alternative to direct simulation of the Navier-Stokes equations has been the use of chaotic advection and for an excellent review of the subject the reader is referred to the paper by Aref [5]. In [12], advection by a reduced model for Langmuir [55] circulations was considered as a mechanism driving plankton patchiness. Also, in [105], a chaotic model for estuarine flow was employed in a study of the transport of pollutants down rivers. Chaotic fluid flow models have a wide range of uses, particularly in physics and engineering [5, 76] and increasingly in advective models of planktonic dynamics [2, 33, 71]. In this chapter we extend the ideas presented in [12] by performing a Melnikov analysis [19, 109] of the chaotic flow, resulting from a time dependent perturbation to the steady dynamics. This analysis, for a limited range of parameters, can then be used to predict the qualitative behaviour [19] of the chaotic transport of plankton immersed in the flow. We also use the analysis to predict when any motile "diffusive" behaviour is likely to have any noticeable effect and how this ultimately depends on the environmental conditions.
Many of the statistical aspects, such as the energy spectrum, are absent with such a phenomenological approach, compared to a Navier-Stokes based model, but many of the qualitative features of the problem may be captured with a good design for the particular flow model [105]. A simple example would be the following reduced Lagrangian model from [12] for Langmuir circulations, here \((\hat{x}, \hat{y}, \hat{z}) = (u, v, w)\) and

\[
\begin{align*}
    u &= 1 + z - C \cos(\pi z) - B \cos(\pi y), \\
    v &= A \pi \cos(\pi z) \sin(\pi y), \\
    w &= -A \pi \sin(\pi z) \cos(\pi y),
\end{align*}
\]

(2.1)

where \(A\) represents the circulation speed and both \(B\) and \(C\) can be expressed in terms of \(A\), [12]. This reduced model represents the leading order terms in a non-linear analysis of the full equations for Langmuir circulations by Leibovich et al. [58]. While containing none of the temperature and solute terms it serves as a useful approximation to the convection cells resulting from the initial instability in the full set of PDE's defining Langmuir circulations. The instability arises as the wind driven circulation overcomes the stabilising vertical temperature stratification, giving rise to Rayleigh-Bernard type convection rolls in the vertical plane but with a varying drift in the (forward) horizontal direction. Since the flow is independent of \(x\), we can effectively ignore the \(u\) term and we can consider the following Rayleigh-Bernard stream function (Hamiltonian) for the motion in the \((y, z)\) plane:

\[
\Phi = A \sin(\pi z) \sin(\pi y),
\]

(2.2)

such that \(v = \partial \Phi / \partial z\) and \(w = -\partial \Phi / \partial y\). In figure (2.1), some the steady streamlines in the convection rolls in the \((y, z)\) plane are plotted.
Figure 2.1: Steady convection rolls in the \((y, z)\) plane in Langmuir circulations, \(A=0.1\).

To simulate the effect of some periodic wind current in the \(y\)-direction we can perturb the system via the following,

\[
y_{\text{new}} = y_{\text{old}} + \epsilon f(t)(z + 1),
\]

where \(f(t)\) is just the time-periodic function \(f(t) = \cos(\omega t)\). The term in \((z + 1)\) is there to simulate the fact that the influence of the transverse wind current decreases with depth, effectively dying away as we reach the lower boundary. Assuming that the magnitude of the perturbation \(\epsilon\) is small, we may Taylor expand in \(\epsilon\) about \(y_{\text{old}}\) to first order and we obtain

\[
\Phi_\epsilon = A \sin(\pi z) \sin(\pi y) + \epsilon Af(t)\pi(z + 1)\sin(\pi z)\cos(\pi y),
\]

which gives a stream function similar to the one used in [19] to study chaotic transport in the “even” Rayleigh-Bernard instability.

Time dependent perturbations of Hamiltonian systems often display coherent
structures such as the Kolmogorov-Arnold-Moser or KAM [109] tori. In the full 
\((y, z, t(\text{mod}2\pi))\) phase-space, these quasi-periodic structures represent the surviving 
remnants of the orbits plotted in figure (2.1). The celebrated KAM theorem [109] 
asserts the persistence of these perturbed structures, for small perturbations of the 
underlying Hamiltonian system.

These perturbed Hamiltonian systems also exhibit regions of chaotic behaviour 
[109]. To illustrate this, in figure (2.2) we plot the non-escape area of the Poincare 
map of the flow defined by (2.4) (the return time naturally being the period of the 
perturbation oscillation, \(T = 2\pi/\omega\)). Figure (2.3) shows orbits, of varying initial 
conditions, under this Poincare map.

![Figure 2.2: Non-escape areas (black) and mixing regions (grey-scale) of the Poincare map of the flow defined in (2.4). Here \(\omega = 1.2, \varepsilon = 0.1\) and \(A = 0.1\).](image)

The non-escape (black) regions in figure (2.2) are made up of the KAM tori we
see in figure (2.3), while the mixing (grey-scale) regions correspond to the regions where we see so called Arnold diffusion, outside the KAM tori, in figure (2.2).

Figure 2.3: Orbits (under the action of the Poincare map) of initial conditions in both the mixing and retaining regions. Here $\omega = 1.2$, $\varepsilon = 0.1$ and $A = 0.1$. The deformed circles represent the KAM tori with the points being the orbits subject to chaotic advection (Arnold diffusion).

This complex (fractal) boundary between the regions of aggregation (KAM tori) and the regions of high mixing (outside the KAM tori) arises because of the complex interaction of the perturbed stable and unstable manifolds of the saddle points at $(y, z) = (n\pi, 0)$ and $(y, z) = (n\pi, -1)$, and $n \in \mathbb{Z}$.

In the introduction, we introduced the idea of chaotic advection (Lagrangian chaos) and in the following we wish to outline how we can quantify the magnitude of the mixing using Melnikov analysis, [109].
2.1 Turnstile lobe dynamics

After perturbation, the heteroclinic connection between the saddle points at \((y, z) = (n\pi, 0)\) and \((y, z) = (n\pi, -1)\) disappears and the respective stable and unstable manifolds separate. For a critical value of \(\epsilon\), the two manifolds will be tangent to each other and after this critical value they will intersect transversely \([19, 109]\). This gives rise to the creation of lobes as shown in figure (2.4). These “turnstile” lobes dictate the geometry of the non-escape area as an organism may only leave its initial cell via these lobes. This can be seen from figure (2.2) as we can visualise the lobes surrounding the black regions. Since these turnstile lobes dictate inter-cellular flux it would be helpful to calculate the area of the lobes to try and quantify the transport of particles between adjoining Langmuir cells. The previously described lobes move in and out of neighbouring cells, under the action of the flow, and to estimate their area we can use the fact that \(\epsilon\) is small \([109]\). This means that, with relatively weak constraints on the size of \(\epsilon\), we can ensure that the lobes are contained entirely in one region as the perturbation is not very strong, see figure (2.4).

The notation used in figure (2.4) is defined as follows; the lobe \(L_{i,i\pm1}\) denotes the lobe in region \(R_i\) that, after one iteration (of time length \(T = 2\pi/\omega\)) of the Poincare mapping of the flow defined in (2.4), will be completely contained in the adjoining region \(R_{i\pm1}\). By making \(\epsilon\) small, we can guarantee that each lobe will satisfy the conditions \(L_{i,i\pm1} \subset R_i\) and \(\mathcal{F}(L_{i,i\pm1}) \subset R_{i\pm1}\), where \(\mathcal{F}\) denotes the Poincare mapping.
Figure 2.4: Sketch of how lobes are formed from the tangling of the stable and unstable manifolds (see text). The schematic is set in the vertical $(y, z)$ plane.

2.1.1 Lobe dynamics and Lagrangian chaos

The observed chaotic behaviour in such systems arises from the creation of horseshoe \cite{19, 25, 109} type sets as the stable and unstable manifolds begin to accumulate on one another close to the hyperbolic fixed points. Even though these horseshoe sets have zero Lebesgue \cite{28} measure, any orbits getting close to them experience a transient period of stretching and folding by these horseshoes, which causes the observed chaotic motion.

However, if we were considering some unbounded array of these Langmuir cells this would not be chaotic motion in the strictest sense as, even though we have exponential separation of nearby initial conditions, the orbits are unbounded. The usual notion of chaotic motion is that this must happen on some compact domain or manifold \cite{25, 109}. The imposition of periodic boundary conditions at the edges of
the Langmuir cell deals with the requirement for a non-compact domain, and hence we can see so-called Lagrangian chaos for orbits in the mixing regions.

### 2.2 Melnikov analysis and particle mixing

In this section, we shadow the analysis of Camassa & Wiggins [19] with application to the system defined by the stream function in (2.4).

Firstly, we use the fact that we can compute the distance, normal to the unperturbed heteroclinic connection, between the perturbed stable and unstable manifolds, $W^s_\epsilon(s^\pm_\tau)$, $W^u_\epsilon(s^\pm_\tau)$ respectively, of the saddle points, $s^\pm_\tau$, located at $(y, z) = (n\pi, 0)$ and $(y, z) = (n\pi, -1)$, and $n \in \mathbb{Z}$. Note that this method does not involve solving the full set of equations governing the flow, but does utilise the fact that the perturbed system can be split into the sum of two Hamiltonians as follows;

$$\Phi_\epsilon = H_0(y, z) + \epsilon H_1(y, z, t).$$  \hspace{1cm} (2.5)

We will denote the distance between the perturbed manifolds as $d(\tau, \epsilon)$, where $\tau \in \mathbb{R}$ parameterises the unperturbed heteroclinic connection, $q_0(-\tau)$, so

$$q_0(-\tau) = (y(-\tau), z(-\tau)) \in W^s_0(s^\pm_\tau) \cap W^u_0(s^\pm_\tau).$$  \hspace{1cm} (2.6)

Let $q^s_\epsilon(-\tau)$ and $q^u_\epsilon(-\tau)$ denote the perturbed parameterisations of the stable and unstable manifolds, respectively. The actual distance between the stable and unstable manifolds is just $\| q^s_\epsilon(-\tau) - q^u_\epsilon(-\tau) \|$ but to calculate this would require solving the equations of motion, a considerably difficult task. The basic idea of Melnikov's method [19, 109] is to replace this with a signed measure of the distance,
\( d(\tau, \varepsilon), \) between the stable and unstable manifolds\textit{ perpendicular} to the unperturbed heteroclinic connection, \( q_0(\tau), \) so
\[
d(\tau, \varepsilon) = \frac{\mathbf{f}(-\tau) \cdot (\mathbf{q}_u(-\tau) - \mathbf{q}_s(-\tau))}{\| \mathbf{f}(-\tau) \|}, \tag{2.7}
\]
where \( \mathbf{f}(\tau) = (\partial_x H_0(\tau), -\partial_y H_0(\tau)) \) is the unperturbed dynamics. Using the fact that the manifolds vary differentiably with the perturbation \( \varepsilon \) \cite{Wiggins}, it follows that we may Taylor expand the bracket in (2.7) as \( \varepsilon \) is small and
\[
d(\tau, \varepsilon) = \varepsilon \frac{\mathbf{f}(-\tau) \cdot \left( \frac{dq_u(-\tau)}{d\varepsilon} - \frac{dq_s(-\tau)}{d\varepsilon} \right)}{\| \mathbf{f}(-\tau) \|} + O(\varepsilon^2), \tag{2.8}
\]
with the gradients evaluated at \( \varepsilon = 0. \) The variational equation for small perturbations to the original heteroclinic orbit is given by the following,
\[
[H_0, H_1]_\mathcal{P} = \frac{\partial H_0}{\partial y} \frac{\partial H_1}{\partial z} - \frac{\partial H_0}{\partial z} \frac{\partial H_1}{\partial y}. \tag{2.9}
\]
and \( [\cdot, \cdot]_\mathcal{P} \) is the Poisson bracket. If we now integrate this over the whole of time we obtain the Melnikov function \cite{Melnikov} defined as,
\[
M(\tau) = \int_{-\infty}^{\infty} \left[ H_0(y(t), z(t)), H_1(y(t), z(t), t + \tau) \right]_\mathcal{P} dt, \tag{2.10}
\]
and finally from \cite{Arnold, Melnikov} we have that,
\[
d(\tau, \varepsilon) = \varepsilon \frac{M(\tau)}{\| \nabla H_0(y(-\tau), z(-\tau)) \|} + O(\varepsilon^2), \tag{2.11}
\]
where \( \| \cdot \| \) is the Euclidean norm on \( \mathbb{R}^2. \)

For a more complete overview of the definitions and theory used in this chapter the interested reader is referred to the book by Wiggins \cite{Wiggins}. The Melnikov function is an indicator of the distance between the stable and unstable manifolds.
perpendicular to the unperturbed manifold and, to $O(\epsilon^2)$, gives the location of the heteroclinic points. In fact, the intersections of the stable and unstable manifolds, again to $O(\epsilon^2)$, are simply the zeros of $M(\tau)$. This can be visualised by applying the implicit function theorem. If there exists a $\tau^*$ such that $M(\tau^*) = 0$ ($\partial_\tau M(\tau^*) \neq 0$) then the manifolds must intersect transversely within in an $\epsilon$-neighbourhood of $(y(-\tau^*), z(-\tau^*))$. Finally, we can calculate an estimate for the lobe area, $\mu(L)$.

$$
\mu(L) = \epsilon \left| \int_{\tau_j}^{\tau_{j+1}} M(t) \, dt \right| + O(\epsilon^2),
$$

(2.12)

where $\tau_{j+1}$ and $\tau_j$ are two consecutive zeros of $M(\tau)$ and $L$ is any lobe. The Melnikov function and hence the lobe area are independent of the particular heteroclinic connection. We use the following parameterisation from [19] for the unperturbed heteroclinic orbit, $q_0(t - \tau),$

$$
y(t - \tau) = 0, \quad z(t - \tau) = \frac{1}{\pi} \sin^{-1}(\text{sech}(A\pi t)).
$$

(2.13)

We now evaluate $[H_0(y(t), z(t)), H_1(y(t), z(t)), t]_P$ at $\tilde{t} = t + \tau$ and substitute this parameterisation into (2.10) to obtain an expression for $M(\tau),$

$$
M(\tau) = \int_{-\infty}^{\infty} \left( \frac{A^2 \pi^3}{2} \sin(2\sin^{-1}(\text{sech}(A\pi t))) \cos(\omega(t + \tau)) \left( \frac{1}{\pi} \sin^{-1}(\text{sech}(A\pi t)) + 1 \right) 
+ A^2 \pi^2 \cos(\omega(t + \tau)) \text{sech}^2(A\pi t) \right) \, dt.
$$

(2.14)

Firstly, we deal with the awkward term in $\sin(2\sin^{-1}(\text{sech}(A\pi t)))$. Trivially, let $\theta/2 = \sin^{-1}(\text{sech}(A\pi t))$ so $\sin(\theta/2) = \text{sech}(A\pi t)$. What we then need to calculate is the value of $\sin(\theta)$ and we use the half-angle formula,

$$
\sin(\theta) = \frac{2u}{1 + u^2},
$$

(2.15)
where \( u = \tan(\theta/2) \). All we need to calculate \( \tan(\theta/2) \) is \( \cos(\theta/2) \) and by considering the triangle with a hypotenuse of length 1 and an opposite side of length \( \text{sech}(A\pi t) \) (i.e. the angle is \( \theta/2 \)), the adjacent side must be of length \( \tanh(A\pi t) \). This means that \( \cos(\theta/2) = \tanh(A\pi t) \) and, after substituting \( u = \tan(\theta/2) = \sin(\theta/2)/\cos(\theta/2) \) into (2.15) we find that

\[
\sin(\theta) = \tanh(A\pi t) \text{sech}(A\pi t).
\] (2.16)

We then proceed by expanding the term \( \cos(\omega(t + \tau)) \) in terms of sines and cosines. By systematically integrating by parts and using a combination of MAPLE calculations and the fact that an odd trigonometric function integrated between the limits \( \pm \infty \) vanishes, we arrive at the following reduced version of (2.14),

\[
M(t) = A^2 \pi^2 \omega \int_{-\infty}^{\infty} \frac{\sin(\omega t)}{\text{sech}(A\pi t) \cos(\omega t)} \, dt
\]

\[
+ A^4 \pi^2 \int_{-\infty}^{\infty} \frac{\cos(\omega \tau)}{\text{sech}^2(A\pi t) \cos(\omega t)} \, dt.
\] (2.17)

The integrals in (2.17) can now be analytically computed by using the method of residues [104]. To do this, we move the integrals in the complex plane and the first integral in (2.17) is changed as follows,

\[
\int_{-\infty}^{\infty} \frac{\cosh(i\omega t)}{\text{sech}(A\pi t) \cos(\omega t)} \, dt = \int_{-\infty}^{\infty} \frac{\cosh(i\omega t)}{\cosh(A\pi t)} \, dt,
\] (2.18)

and by standard hyperbolic/exponential identities this further reduces to the following two integrals:

\[
\int_{-\infty}^{\infty} \frac{e^{i\omega t}}{e^{A\pi t} + e^{-A\pi t}} \, dt + \int_{-\infty}^{\infty} \frac{e^{-i\omega t}}{e^{A\pi t} + e^{-A\pi t}} \, dt.
\] (2.19)
We integrate around the contour defined by the perimeter of the rectangle defined by its vertices at $R$, $R + i/A$, $-R + i/A$, and $-R$. By Cauchy's residue theorem [104], this contour integral is equal to $2\pi i$ times the sum of the residues of the poles (contained inside the contour) of the integrand. The poles of the integrands in (2.19) are given by the zeros of $\cosh(A\pi t)$ which are $t = i/2A$. By letting $R \to \infty$, the integrals from $\pm R$ to $\pm R + i/A$ vanish. By making the following coordinate changes; $t \to -z + i/A$ and $z = -x$ [104], the integrals in (2.19) become

$$\left(1 + e^{-\omega/A}\right) \left(\int_{-\infty}^{\infty} \frac{e^{i\omega x}}{e^{A\pi x} + e^{-A\pi x}} \, dx + \int_{-\infty}^{\infty} \frac{e^{-i\omega x}}{e^{A\pi x} + e^{-A\pi x}} \, dx\right) = 2\pi i \sum \text{res} \left(\frac{i}{2A}\right) (2.20)$$

where the sum represents the sum of the residues of the two integrands.

The residues [104] of the two integrands are given by $e^{-\omega/2A}/2\pi i A$ and $e^{\omega/2A}/2\pi i A$, respectively. It is the sum of the integrals in the bracket of (2.20) we are interested in, after dividing both sides by $1 + e^{-\omega/A}$ and a little algebra, we finally obtain the following expression for the integral in (2.18).

$$\int_{-\infty}^{\infty} \text{sech}(A\pi t) \cos(\omega t) \, dt = \frac{1}{A} \text{sech}(\omega/2A), \quad (2.21)$$

as was obtained in [19, 109].

As for the second integral in (2.27), by integration by parts we find that

$$\int_{-\infty}^{\infty} \text{sech}^2(A\pi t) \cos(\omega t) \, dt = \omega \int_{-\infty}^{\infty} \tanh(A\pi t) \sin(\omega t) \, dt, \quad (2.22)$$

and using the the hyperbolic identities $\sin(\omega t) = \sinh(i\omega t)/i$ and $\tanh = \sinh / \cosh$

we arrive at a similar complex exponential integral to that seen in (2.19)

$$\frac{1}{2i} \left(\int_{-\infty}^{\infty} \frac{e^{(\pi A + i\omega)t} + e^{(\pi A - i\omega)t}}{e^{A\pi t} + e^{-A\pi t}} \, dt - \int_{-\infty}^{\infty} \frac{e^{(\pi A - i\omega)t} + e^{(i\omega - \pi A)t}}{e^{A\pi t} + e^{-A\pi t}} \, dt\right). \quad (2.23)$$
As before, the poles of the integrands are at $t = i/2A$ and using the same two coordinate changes, we apply Cauchy's residue theorem to obtain the following,

$$\int_{-\infty}^{\infty} \text{sech}^2(A\pi t) \cos(\omega t) \, dt = \frac{1}{A} \text{cosech}(\omega/2A).$$  \hspace{1cm} (2.24)

So we can now define our Melnikov function, using (2.19), (2.21) and (2.23) as

$$M(\tau) = \pi^2 \omega (\sin(\omega \tau) \text{sech}(\omega/2A) + A \cos(\omega \tau) \text{cosech}(\omega/2A)),$$  \hspace{1cm} (2.25)

whose zeros are given by

$$\tau = \frac{1}{\omega} \tan^{-1}(-A \coth(\omega/2A)) + n\pi,$$  \hspace{1cm} (2.26)

and $n \in \mathbb{Z}$. The first expression in (2.25) is, apart from the factor of $\pi^2$ due to the different expression for the stream function from that given in [19], the same as that found for the Melnikov function in the Camassa & Wiggins paper [19] on time-perturbed Rayleigh-Bernard systems. They used a non-depth dependent perturbation and the effect of the use of the depth dependent perturbation appears in the second expression in (2.25).

The zeros of the Melnikov function are as mentioned the location, to order $\varepsilon^2$, of the heteroclinic points. Here we have infinitely many zeroes which suggested that we do indeed have tangling of the manifolds. So, again to $O(\varepsilon^2)$, the area contained in one turnstile lobe is found by integrating the Melnikov function between two consecutive zeros. We note that this estimate to the area of a lobe is well defined as it doesn't depend on the choice of zeros between which you integrate.

There is a closed form expression for the lobe area (computed in MAPLE) but it is too long to show here. Figure (2.5) plots the escape area (due to lobe transport
Figure 2.5: Prediction of the lobe non-escape area, $1-\mu(L)$, as a function of $\omega$ from the Melnikov analysis.

$1 - \mu(L)$ as a function of the forcing frequency $\omega$, which predicts almost no intercellular flux for very small $\omega$, but this flux increases to a maximum around $\omega = 0.15$ and decreases monotonically with increasing forcing frequency. We can only check this for a relatively small range of $\omega$ because as we increase the forcing frequency the role of the secondary intersection points play their part. For the simple set up in figure (2.4), we are considering only the primary intersection points, (PIPs), of the manifolds. However, near the saddle points as the manifolds accumulate on each other, secondary tangencies occur and the creation of secondary intersection points, (SIPs). These in turn form new lobes which, as $\omega$ increases, have an increasing effect on the inter-cellular flux.
In figure (2.6), we numerically plot the non-escape area of the Langmuir cell, as a function of the forcing frequency $\omega$. Since only the lobe formed by the heteroclinic connection around $(y, z) = (0, 1/2)$ will play any real part in the lobe transport (because lobes close to the saddle points quickly form secondary and tertiary tangencies and play little part in the simple transport set up based on PIP’s) we expect some agreement between figures (2.5) and (2.6). This is the case, but only for a small range of values of $\omega$ due to the increasing influence of the secondary/tertiary lobe dynamics. The qualitative form of the graph in figure (2.6) can change with the resolution and definition of when a particle has “escaped” the Langmuir cell. In all cases in this chapter, the computational grid had $500 \times 500$ cells and we defined escape to mean escape after 100 iterations of the Poincare map. Thus we hypothesise...
that, at least for small $\omega$, the differences in the numerical and analytical predictions of the non-escape area are largely due to resolution (both spatial and temporal).

This suggests that such an analysis has a fairly limited scope of applicability as it relies on both small perturbations $\varepsilon < 1$ and low frequency variations in the perturbations $\omega \ll 1$ due to the simplistic assumptions on transport in the flow.

### 2.3 Chaotic transport versus diffusion

So far, we have only considered the transport of passive scalars via the turnstile lobes but when, if at all, must we also consider diffusion? To compare the dominance of particle transport by lobe dynamics or diffusion, [109] considered the time scale of transport across a particular lobe. Intuitively, if we have some particle/organism with diffusivity $D$, then the time taken to diffuse across a lobe is given by $T_D$ and

$$T_D = \frac{d(\tau, \varepsilon)^2}{D} \quad (2.27)$$

where $d(\tau, \varepsilon)$, as defined in (2.11), is effectively the width of the lobe. Now $T$, the time taken for a lobe to move from one region to the adjoining region is the period of oscillation of the forcing oscillation so $T = 2\pi/\omega$.

Now, if lobe transport dominates diffusion, $T_D \gg T$. To investigate if and when this is true we first fix $\tau = \frac{\pi}{4}$, this simply ensures that our reference point for the value of $d(\tau, \varepsilon)$ is non-zero. Now, combining equations (2.11) and (2.25), we have the following expression for the ratio of the diffusive and advective time-scales;

$$\frac{T_D}{T} = \frac{\varepsilon^2 \pi \omega^3}{2DA^2} \left( \sin \left( \frac{\omega\pi}{4} \right) \text{sech}(\omega/2A) + A \cos \left( \frac{\omega\pi}{4} \right) \text{csch}(\omega/2A) \right)^2. \quad (2.28)$$
In figure (2.7), we plot the ratio $T_D/T$ where we set $D = 10^{-4} \text{m}^2\text{s}^{-1}$ (a typical swimming diffusivity for algal cells [40]) and $A = \epsilon = 0.1$. We see that, to begin with, for small $\omega$, the ratio is zero as we have transport primarily by diffusion. With increasing $\omega$, the lobes soon begin to dominate transport reaching a peak at around $\omega = 0.4$. The reason this differs from the maximum lobe area predicted in figure (2.5) is because (2.27) assumes that the lobe is square in shape, when in reality it can form very complicated shapes. At this value of $\omega$, the diffusion plays a lesser role in the mixing of the system. However, as we exceed this resonant value of the forcing frequency, the lobe area reduces and diffusion becomes the dominant force with respect to particle mixing.

![Figure 2.7: Diffusivity/advection ratio analytically obtained from the Melnikov analysis. Here $D = 10^{-4} \text{m}^2\text{s}^{-1}$ and $A = \epsilon = 0.1$.](image)
2.4 Conclusions

There is some compelling field evidence to support the idea of chaos helping to cause the patchiness of plankton observed in oceans and lakes. Franks [33] showed pictures of lake dwelling planktonic populations with a spatial distribution consisting of many deformed filaments of plankton biomass, something widely observed in the passive advection of particles by chaotic motions [5, 76]. Also, such reduced/made-to-measure models like those presented by Bees [12] and Stirling [105] can capture the qualitative features of the relevant system quite well, such as regions of high mixing and regions of aggregation. While direct numerical simulation of the Navier-Stokes equations, even in simple geometries, still remains such an intensive process we will need viable alternatives when studying the motion of organisms in specific situations.

We discussed how some of these models of chaotic advection can give rise to regions of organism aggregation (KAM tori) and regions of strong, complex mixing (regions where we see Arnold diffusion). Both these processes are of interest with regard to the dynamics of advected plankton as these two processes can highly influence the patchiness of the system and the factors such as the phytoplankton-zooplankton contact rates, both are factors seen to influence phytoplankton bloom formation [57, 92].

In some cases, analysis of the transport and mixing of the organisms can be performed but not always. The Melnikov analysis performed in this chapter relies on the magnitude of the perturbation (in this case $\epsilon$) being small. Also, it was observed
that, by comparing the analytical estimates of cell flux due to lobe transport and numerically measured estimates, the accuracy of the analytical estimates was highly dependent on the frequency of the perturbation, $\omega$, to the steady dynamics. There was good agreement between the two estimates for small values of $\omega$ but, as the frequency increased and the subsequent dynamics became ever more complex the analytical estimates rapidly diverged from the numerical ones.

Using a simple estimate for the relative diffusive time scales of lobe transport and simple Fickian “swimming” diffusion first seen in [109], we demonstrated that any motile behaviour may or may not be relevant, depending again on the frequency of the perturbation. Time perturbed Hamiltonian systems of this kind are known to exhibit such resonant behaviour, with regard to the influence of lobe dynamics [109]. This suggests that the comparative importance of any motile effects due to swimming plankton may depend very strongly on the particular environmental conditions, such as the wind frequency considered here.

In Stirling [105], the perturbation $K$ to the steady regime represented the ratio of the magnitude of buoyancy driven and curvature (of the river) driven advection in the vertical plane, which was greater than one. Consequently, any Melnikov analysis is pointless as the accuracy of the expressions for lobe area and the Melnikov function are $O(\varepsilon^3)$. However, because such models are usually 2+1 or 3+1 dimensional, mixing can be measured by calculating the Liapunov exponents [25] of the system. These give the measure of the exponential separation of nearby particles/organisms immersed in the relevant flow. Factors such as low-dimensionality and adaptability mean that, while lacking the rigour and statistical properties of the Navier-Stokes
equations, chaotic advection models can be useful for investigating the turbulent transport of planktonic organisms in specific scenarios.
Motile, spheroidal organisms in small scale
shear flows

Shear flows are a common feature in the ocean and occur at all spatial scales but one question we hope to tackle is whether such flows, coupled with reaction dynamics and swimming behaviour, can cause pattern formation. If the shearing motion can cause instability in a homogeneous medium then it could contribute to the patchiness of oceanic plankton. The average concentration of plankton in the ocean is very dilute, from Bees & Edwards [13] around 10 mg Chl m$^3$, so we can safely assume that its movement has little or no feedback on the flow. This being the case, the general equations for swimming particles, $S_i$, immersed in some incompressible flow, $U(x, y)$ are

$$\partial_t U + (U \cdot \nabla) U = -\frac{1}{\rho} \nabla p - g j + \nu \nabla^2 U,$$

$$\nabla \cdot U = 0,$$  \hspace{1cm} (3.1)

$$\partial_t S_i = \mathcal{R}_i(S) - \nabla \cdot [S_i(U + V_i(p_i)) - D_i \nabla S_i].$$

The flow coefficients and variables are as defined in Section I (1.3), $\langle p_i(x) \rangle$ and $D_i(x)$ represent the mean swimming direction and translational diffusivity tensor.
of species $i$, respectively and $V_{si}$ is the organism's mean swimming speed. The functions $R_i$ denote the reaction terms and $S = (S_1, ..., S_N)$, $N$ being the number of types of interacting organisms.

To calculate the quantities $\langle p \rangle$ and $\mathbf{D}$ (where we drop the index for convenience) we will need to know something of the probability distribution function (pdf), $P(p)$, of the organisms' swimming directions. In Pedley & Kessler [81] this was modelled by the following Fokker-Planck equation,

$$\partial_t P + \nabla \cdot (\dot{p} P) = D_r \nabla^2 P,$$

(3.2)

where $D_r$ represents the rotational diffusivity, which models the intrinsically random part of the organisms' orientation. This description is supported by the experimental evidence of Hill and Hader [40]. The rate of change of the swimming direction, $\dot{p}$, was expressed in [81] as a deterministic torque balance where, in the absence of a gravitational torque,

$$\dot{p} = \frac{1}{2} \Omega \wedge p + e p \cdot \mathbf{E} \cdot (I - pp),$$

(3.3)

and $\Omega = \nabla \wedge \mathbf{U}$ is the fluid vorticity, $e$ is the eccentricity of the organism. The eccentricity is defined in terms of the major $a$ and minor $b$ axial lengths of the ellipse defining the 2D view of the organism and $e = (a^2 - b^2)/(a^2 + b^2)$. Lastly, $\mathbf{E}$ is defined to be the rate-of-strain tensor. From [81],

$$\langle p \rangle = \int_{S^2} p P dx,$$

(3.4)

where $S^2$ is the surface of the unit sphere. If one assumes that there is some constant organism reorientation time, $\tau_r$, in which the organism settles into a new direction,
then the diffusivity tensor can be approximated (at least for small shear rates) by

$$D = V_s^2 \tau_r \langle V(\mathbf{p}) - (\mathbf{p})^2 \rangle,$$

(3.5)

where $V = \langle V^2 \rangle / V_s$ and $\langle V^2 \rangle$ is the mean square swimming speed. This approximation was shown by Hill & Bees [42] to be incorrect for larger shear rates. It was seen that the increasing vorticity effects acted so as to constrain the organism in ever decreasing circles of motion, causing $\|D\| \to 0$ in the limit of increasing shear.

When $\mathbf{p}$ is defined as in (3.3), without the so called gyrotaxis terms, related to gravitational torques, defined in [81] and [11], then $\langle \mathbf{p} \rangle = 0$. This simplifies the full equations in (3.1) and the approximation to the diffusivity tensor. Hence,

$$D = \langle V^2 \rangle \tau_r \langle \mathbf{p} \mathbf{p} \rangle.$$

(3.6)

In general, $D$ will be a symmetric, non-diagonal tensor whose diagonal elements may also differ. If we assume that $\mathbf{U}$ is a simple, steady shear flow, e.g. $\mathbf{U} = \alpha y \mathbf{i}$ then the governing equations simplify dramatically and the elements of $D$ depend only on the strength of the shear, $\alpha$, the eccentricity $e$ and the mean swimming speed, $V_s$. See [11] for a comprehensive review of how the elements of $D$ vary w.r.t. the system parameters $\alpha$, $e$ and $V_s$.

For such swimming behaviour to have any effect we must clarify the spatial scales we are considering. Presuming that zooplankton swimming speeds are comparable with that of motile algal cells with some typical reorientation time, $\tau_r = \mathcal{O}(1)$ s, then the diffusion can be approximated by $D_s \approx V_s^2 \tau_r$, which is $\mathcal{O}(10^{-2})$ m$^2$s$^{-1}$ (based on data for algal cells from [40]). Okubo [73], derived an experimental scaling for the effective diffusion of tracers by turbulence with respect to the relevant length.
scale. Using this observed relationship between $D$ and $L$ ($D = 0.01 L^{1.15}$ in units of centimetres and seconds) means that, for swimming diffusion to be comparable to turbulent diffusion, our length scales are $O(10)$m or less.

Instability of systems advected by shear flows have been investigated before with regard to specific reaction models, [60, 72]. The small-scale patchiness of swimming micro-organisms, due to shear, has also been identified in Mitchell et al. [68, 69]. The linear stability of the general reaction model case was investigated by Doering & Horsthemke [24] for the case where the diffusion tensors of each species were identical. It was found that, in this case, the shear had an overall stabilising effect. In section 3.1, we extend this analysis to systems of swimming, reacting, spheroidal organisms which have equivalent diffusion tensors, but allowing for non-diagonal diffusion terms and variation in the vertical and horizontal diffusion components. Here, it is seen that instability is possible, but this is dependent on the strength of the non-diagonal terms in the diffusion tensor compared to the vertical component of the diffusion tensor. In section 3.2, we extend the work of Spiegel & Zaleski [96], by again considering swimming, spheroidal organisms and perform a weakly non-linear analysis about the no-flow instability and allow for different diffusion tensors for different species. The Turing instability mechanism has been suggested as a possible mechanism for plankton patchiness at large spatial scales [60]. This is unlikely as the Turing instability is driven by having a suitably large difference in the diffusion coefficients of the two reacting species, something which is not possible at the km spatial scale as the dominance if turbulent dispersion causes the diffusion coefficients to be effectively equal. In chemical reactor systems, Satnoianu et al. [95]
demonstrated that advection can radically decrease the disparity in the diffusion coefficients needed to see a Turing type instability. However, we aren’t considering isotropic diffusion tensors and the instability may arise from the more complex diffusive characteristics and we may not require the Turing criteria to be satisfied.

3.1 Linear Analysis: Equivalent diffusion tensors

We initially follow the linear analysis of [24] but because we are allowing the organisms to be spheroidal and to swim, the analysis is slightly more difficult due to both straining and Taylor dispersive effects which cause the diffusion tensors to be anisotropic and non-diagonal in nature.

By assuming the flow to be a steady linear shear flow, \( \mathbf{U} = \alpha y \mathbf{i} \), and using the fact that \( \langle \mathbf{p} \rangle = 0 \), the full equations in (3.1) can be reduced to the following, spatially two-dimensional, reaction-diffusion-advection equation:

\[
\partial_t S_i = \mathcal{R}_i(S) - \alpha y \partial_x S_i + D_{xx}^{**} \partial_x^2 S_i + D_{yy}^{**} \partial_y^2 S_i + 2D_{ij}^{**} \partial_x \partial_y S_j, \tag{3.7}
\]

where \( S(x, y, t) = (S_1, ..., S_N) \) is the vector of population values, the operator \( \mathcal{R}(\cdot) \) represents the linear and non-linear reaction terms. For the remainder of this chapter, the operators \( \partial_V \) and \( d_V \) will represent the partial and normal derivatives, respectively, with respect to the variable \( V \). We shall assume that each of the diffusion tensors \( \mathbf{D}^{**} \) are diagonal, where each star can be replaced by an \( x \) or a \( y \), and \( x \in (-\infty, \infty) \) and \( y \in [-L, L] \). The boundaries at \( y = -L, L \) are assumed to be
impermeable leading to the following, no flux boundary conditions:

$$\partial_y s_i(x, \pm L, t) = 0. \quad (3.8)$$

We non-dimensionalise (3.7) via the following:

$$x = L \tilde{x}, \ y = L \tilde{y}, \ t = \frac{L^2 \tilde{t}}{D}, \ S = s_r s, \ \mathcal{R}_i(s) = \frac{D s_r}{L^2} \tilde{\mathcal{R}}_i(s) \quad (3.9)$$

where $s_r$ is a reference population. Since we are considering only organisms with identical swimming and shape characteristics, $D_{ij} = D^{**}$, and we define $D = D^{**}$.

After dropping the tildes we obtain the following, dimensionless form of (3.7):

$$\partial_t s_i = \mathcal{R}_i(s) - P_e y \partial_x s_i + \tilde{D}^{xx} \partial^2_x s_i + \tilde{D}^{xy} \partial_y^2 s_i + 2 \tilde{D}^{yx} \partial_x \partial_y s_i, \quad (3.10)$$

where $P_e = \alpha L^2 / D$, $\tilde{D}^{**} = D^{**} / D$ (and so $\tilde{D}^{xx} = 1$). The new boundary conditions follow logically from (3.8),

$$\partial_y s_i(x, \pm 1, t) = 0. \quad (3.11)$$

As is customary, we assume the existence of a homogeneous equilibrium solution, $s^*$, of the nonlinear reaction system, i.e. $\mathcal{R}_i(s^*) = 0$. Trivially, $s_i = s_i^*$ will be a homogeneous equilibrium state for (3.10). We will conduct a linear stability analysis of this solution using suitably small perturbations of this state, $s_i = s_i^* + \epsilon s_i^1$ ($\epsilon \ll 1$), yielding the linearised version of (3.10):

$$\partial_t s_i^1 = \mathcal{J}_{ij} s_j^1 - P_e y \partial_x s_i^1 + \partial^2_x s_i^1 + \tilde{D}^{xy} \partial_y^2 s_i^1 + 2 \tilde{D}^{yx} \partial_x \partial_y s_i^1, \quad (3.12)$$

where $\mathcal{J}_{ij} = \partial \mathcal{R}_j / \partial s_i(s^*)$ and with the usual ansatz

$$s_i^1 = A_i \theta(y) e^{\omega t + ikx}, \quad (3.13)$$
such that the eigenfunctions, $\theta(y)$, satisfy $d_\gamma \theta(\pm 1) = 0$ in accordance with (3.11).

After substituting this ansatz into (3.12) we see that

$$\left( I(-\omega - ik P_y y - k^2 + 2 \tilde{D}^{\gamma\nu} d_y + \tilde{D}^{\nu\gamma} d_y^2) - \mathcal{J} \right) \Lambda \theta = 0, \quad (3.14)$$

where $I$ is the identity matrix. We are interested in the spectrum of $\omega$. To calculate this we first need to obtain the eigenfunctions $\theta(y)$ and eigenvalues $\eta$ which satisfy

$$\tilde{D}^{\gamma\nu}(d_y^2 - \eta) \theta(y) = ik (P_y y - 2 \tilde{D}^{\gamma\nu} d_y) \theta(y). \quad (3.15)$$

Once we have ascertained the spectrum of $\eta$, the dispersion relation in (3.14) requires that

$$\det \left( (\eta - k^2 - \omega) I - \mathcal{J} \right) = 0, \quad (3.16)$$

for non-trivial solutions. If we now let $\tilde{\omega} = \omega + k^2 - \eta$ then (3.16) now becomes

$$\det (\mathcal{J} - \tilde{\omega} I) = 0. \quad (3.17)$$

and the $\tilde{\omega}$ are the eigenvalues of $\mathcal{J}$ and by the assumption that $s^*$ is a stable equilibrium for the spatially homogeneous system, $\text{Re}(\tilde{\omega}) < 0$. So, once the spectrum of $\eta$ is known, we see that $\omega = \tilde{\omega} - k^2 + \eta$. The difference between (3.14) and the analogous equation in [24] is the additional term in $d_y \theta$ occurring due to the more complex nature of the diffusion tensor. Returning to (3.15), to remove the term in $d_y \theta$ we can consider the new eigenfunction, $\varphi(y)$, defined as

$$\varphi(y) = \theta(y) e^{ik \tilde{D}^{\gamma\nu} y}, \quad (3.18)$$

which gives us, after substitution in (3.15), the following new eigenfunction-eigenvalue equation

$$\left( d_y^2 - \mu \right) \varphi(y) = ik \xi y \varphi(y), \quad (3.19)$$
where \( \xi = P_c/D_{uv} \) and the eigenvalue, \( \mu \) is given by

\[
\mu = \eta - \frac{(D_{xy})^2 k^2}{(D_{uv})^2}.
\]  

(3.20)

Combining (3.11) and (3.18) we obtain the boundary conditions for (3.19):

\[
d_y \varphi(\pm 1) = ik \frac{D_{xy}}{D_{uv}} \varphi(\pm 1). \tag{3.21}
\]

If we now multiply (3.19) by \( \bar{\varphi}(y) \), the complex conjugate of \( \varphi(y) \), and integrate over the vertical domain, \( y \in [-1, 1] \), we obtain

\[
\int_{-1}^{1} \frac{1}{\varphi} \varphi^2 \varphi dy - i k \xi \int_{-1}^{1} y |\varphi|^2 dy = \mu \int_{-1}^{1} |\varphi|^2 dy. \tag{3.22}
\]

Integration by parts on the first integral in (3.22) yields

\[
\int_{-1}^{1} \varphi \varphi^2 dy = [\varphi \varphi^2]_{-1}^{1} - \int_{-1}^{1} \varphi^2 \varphi^2 dy, \tag{3.23}
\]

and \( [\varphi \varphi^2]_{-1}^{1} = ik \frac{D_{xy}}{D_{uv}} (|\varphi(1)|^2 - |\varphi(-1)|^2) \), from (3.21), which is purely imaginary.

So the real part of \( \mu \) is given by

\[
\text{Re}(\mu) = - \int_{-1}^{1} |d_y \varphi|^2 dy / \int_{-1}^{1} |\varphi|^2 dy \leq 0, \tag{3.24}
\]

and the imaginary part of \( \mu \) is given by

\[
\text{Im}(\mu) = k \left( \frac{D_{xy}}{D_{uv}} (|\varphi(1)|^2 - |\varphi(-1)|^2) - \xi \int_{-1}^{1} y |\varphi|^2 dy \right) / \int_{-1}^{1} |\varphi|^2 dy, \tag{3.25}
\]

Hence, the real part of \( \mu \) is always non-positive but if \( k \xi \neq 0 \) it must actually be negative because \( \text{Re}(\mu) = 0 \) implies that \( \varphi(y) = c \), for some \( c \in \mathbb{C} \). This cannot be if \( \varphi(y) \) is to satisfy (3.19). Also, \( \text{Im}(\mu) \neq 0 \) if \( k \) and the term in the bracket in
(3.25) is non-zero, suggesting the existence of oscillatory time-scales. Returning to the spectrum of our advective-diffusive eigenvalue, \( \eta \), we can see from (3.20) that
\[
\eta = \mu + \frac{(\hat{D}^{xy})^2 k^2}{(\hat{D}^{yy})^2}
\]  
(3.26)

and upon substituting this expression for \( \eta \) back into (3.17) we obtain the following equation for the eigenvalue \( \omega \):
\[
\omega = \bar{\omega} - k^2 + \mu + \frac{(\hat{D}^{xy})^2 k^2}{(\hat{D}^{yy})^2}.
\]  
(3.27)

By assumption \( \text{Re}(\bar{\omega}) < 0 \) and we have shown that \( \text{Re}(\mu) < 0 \) so any destabilisation depends upon the following term,
\[
\left( \frac{(\hat{D}^{xy})^2}{(\hat{D}^{yy})^2} - 1 \right) k^2.
\]  
(3.28)

If the term in the bracket is positive then there will be some finite wavenumber, \( k_c \), such that for all \( k > k_c \), \( \text{Re}(\omega) > 0 \) and the system will be unstable to perturbations above this wavenumber. The bracket in (3.28) will be positive if
\[
|\hat{D}^{xy}| > |\hat{D}^{yy}|,
\]  
(3.29)
as \( \hat{D}^{**} = D^{**}/D^{xx} \). The modulus bars are necessary as it has been observed that, for certain shear rates, \( D^{xy} < 0 \). How does this relate to the physical constraints on the diffusivity tensor, i.e. \( \det(D) \geq 0 \). This positivity condition requires that
\[
D^{xx} D^{yy} - (D^{xy})^2 > 0,
\]  
(3.30)

and in the final section of the chapter, we discuss the physical possibility of both inequalities (3.29) and (3.30) being satisfied.
Combining (3.28) and (3.27), the critical wavenumber \( k_c \), such that \( \text{Re}(\omega(k_c)) = 0 \), is given by

\[
k_c = \sqrt{\frac{-\text{Re}(\bar{\omega} + \mu)}{\left(\frac{D^{xx}}{D^{yy}}\right)^2 - 1}}
\]

(3.31)

This result is interesting in that it suggests that instability is possible, even for identical diffusion tensors, contrary to the work of Doering & Horsthemke [24]. The key difference is the more complex diffusive behaviour of swimming, spheroidal organisms immersed in even simple shear flows. It is this diffusive behaviour that allows the possibility of an instability.

The paper by Spiegel & Zaleski [96] investigated the no-flow instability (about the zero wavenumber) for a general reaction-diffusion system subject to linear shear and having distinct diffusion coefficients. It was seen that the amplitude equation, resulting from a weakly non-linear analysis, was ill-posed for certain parameter regimes suggesting that zero was not always the critical wavenumber. We extend the analysis in [96] to swimming, spheroidal organisms and also consider a different non-linear reaction term and discuss the implications of this term on the criticality of the instability.

3.2 Linear & Non-Linear Analysis: Different diffusion tensors

In most cases, the diffusion tensors for phytoplankton and zooplankton are likely to be different. For this scenario, the analysis is a little more difficult. To make the
equations more tractable, we follow the analysis in [96] and look at perturbations of the no-flow instability around the zero wavenumber. Returning to (3.10), we have the following non-dimensional equations for variations, $\delta r_i$, from the equilibrium values, $s_i^*$, of the species $s_i$:

$$\partial_t \delta r_i = \mathcal{L}_{ij} \delta r_j + \mathcal{N}_i(r) - P_{ey} \partial_x \delta r_i + D_{ij} \partial_x^2 \delta r_j + D_{ij} \partial_y \delta r_j + 2D_{ij} \partial_x \partial_y \delta r_j, \quad (3.32)$$

and, for simplicity, we assume that there are only two reacting species, the tensors $\mathbf{D}^{**}$ are diagonal but now we set $D = ||\mathbf{D}^{xx}||$ (hence $\mathbf{D}^{**} = \mathbf{D}^{**}/D$) and obviously $r = 0$ is a solution of (3.32). The boundary conditions follow immediately from (3.11) such that

$$\partial_y r(\pm 1, x, t) = 0. \quad (3.33)$$

As for the non-linear term, $\mathcal{N}_i(r)$, we assume that it consists of a sum of polynomial terms like

$$\mathcal{N}_i(r) = \sum_{k_1, k_2 = 0}^{d_2} \sum_{k_1, k_2 = 0}^{d_2} \mathcal{N}_{k_1 k_2} r_{i1}^{d_1 - k_1 + k_2} r_{i2}^{k_1 + k_2}, \quad (3.34)$$

where $d_1 + d_2$ is the degree of $\mathcal{N}$ (and $d_1$ represents the minimum degree of the non-linear terms hence $d_1 \geq 2$). If we start from the no-flow scenario ($P_e = 0$), $\mathbf{D}^{xx} = \mathbf{D}^{yy} = \mathbf{D}$ and $\mathbf{D}^{xy} = 0$ because any movement of the organism is due solely to the (assumed) intrinsically stochastic spatial orientation of the particle, swimming at some constant speed, hence modelled by a simple diffusive term. The more complex diffusive terms in $\mathbf{D}^{xy}$ are caused by the shape characteristics of the organism and occur only when the flow is present. The linearised version of (3.32)
admits solutions of the form,

\[ r_i = B_i \cos(l \pi y) e^{\omega t + i \nu x}, \]  

(3.35)

which satisfy the boundary conditions stated in (3.33). Also, \( l \) is an even integer and \( B \) satisfies the following equation,

\[ (\mathcal{L} - (m^2 + l^2 \pi^2) \bar{D}) - \omega I) B = 0. \]  

(3.36)

Equation (3.36) will have a non-trivial solution if

\[ \omega^2 - \omega \text{Tr}(\mathcal{L} - (m^2 + l^2 \pi^2) \bar{D}) + \text{det}(\mathcal{L} - (m^2 + l^2 \pi^2) \bar{D}) = 0. \]  

(3.37)

We thus have a bifurcation when \( \text{det}(\mathcal{L} - (m^2 + l^2 \pi^2) \bar{D}) = 0 \), as then \( \omega = 0 \), and we can avoid the Hopf bifurcation (for clarity) by stipulating that \( \text{Tr}(\mathcal{L} - (m^2 + l^2 \pi^2) \bar{D}) < 0 \). Later in the analysis, we derive the amplitude equations in a long horizontal length-scale so we consider the case where \( l = 0 \). Then, for small horizontal wave-numbers \( m \), we may consider the case where \( \text{det}(\mathcal{L}) \approx 0 \) (as \( l = 0 \) and \( m \ll 1 \), see equation (3.51)). If we now let

\[ \mathcal{L} = \mathcal{P} + \epsilon^2 \mathcal{Q}, \]  

(3.38)

where \( \epsilon^2 \ll 1 \) and \( \det(\mathcal{P}) = 0 \), then the \( m = 0 \) mode will be marginally stable when both \( \epsilon \) and \( P_e \) are zero so we can expand about the zero wavenumber. We mention that a major reason for expanding about the zero wavenumber is to make the resulting equations more tractable. Without loss of generality we can assume that the matrix \( \mathcal{P} \) can be placed in the form

\[ \mathcal{P} = \begin{pmatrix} \kappa_1 & -\kappa_1 \\ -\kappa_2 & \kappa_2 \end{pmatrix}. \]  

(3.39)
because the diagonal similarity transformation that achieves this leaves the diffusion tensors diagonal [96]. The \( \kappa_i \) are real and to once again avoid the Hopf bifurcation \( \sum \kappa_i < 0 \). To further expand in \( \epsilon \) we introduce the following re-scalings:

\[
\partial_t = \epsilon^2 \partial_T, \quad \partial_x = \epsilon \partial_X, \quad r_i = \epsilon^{2j(i-1)} u_i. \tag{3.40}
\]

The parameter \( d_1 \) is the degree of the smallest non-linear terms. The scalings are used so that we produce the amplitude equation for disturbances in the long horizontal scale, \( X \), and so that the first non-linear terms appear at \( O(\epsilon^3) \).

This now gives us the following version of (3.32):

\[
\tilde{D}_{ij}^{yy} \partial_y^2 u_j + \mathcal{P}_{ij} u_j = \epsilon \left( P_{xy} \partial_X u_i - 2 \tilde{D}_{ij}^{xy} \partial_X \partial_y u_j \right) + \epsilon^2 \left( \partial_T u_i - \tilde{D}_{ij}^{xx} \partial_X u_j - Q_{ij} u_j - \tilde{N}_i(u) \right),
\tag{3.41}
\]

subject again to the boundary conditions in (3.33) and \( \tilde{N}_i(u) \) represents the non-linear terms of order \( d_1 \) only. We now expand the \( u_i \) in \( \epsilon \), so

\[
u = u^0 + \epsilon u^1 + \epsilon^2 u^2 + O(\epsilon^3). \tag{3.42}
\]

To order zero we find that

\[
\tilde{D}_{ij}^{yy} \partial_y^2 u_j^0 + \mathcal{P}_{ij} u_j^0 = 0, \tag{3.43}
\]

which is satisfied for any \( u_j^0 = u_j^0 \) such that

\[
u^0 = A(X,T) \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{3.44}
\]

for some, as yet, arbitrary function \( A(X,T) \). The \( O(\epsilon) \) equations are

\[
\tilde{D}_{ij}^{yy} \partial_y^2 u_j^1 + \mathcal{P}_{ij} u_j^1 = P_{xy} \partial_X u_i^0 - 2 \tilde{D}_{ij}^{xy} \partial_X \partial_y u_j^0. \tag{3.45}
\]
but the last term on the r.h.s. of (3.45) is zero because the \( u_i^0 \) are independent of \( y \).

As in [96], an appropriate solvability condition on (3.41) is that, at each order,

\[
\int_{-1}^{1} v^T F dy = 0, \tag{3.46}
\]

where \( v = (\kappa_2, \kappa_1) \) is the solution of the homogeneous adjoint problem and the vector \( F \) represents the inhomogeneous terms appearing on the r.h.s. at each order. This orthogonality condition guarantees that any secular terms will disappear and our solutions at each order will exist. This is because of the Fredholm alternative.

Firstly, consider the following boundary value problem,

\[
\begin{align*}
A u &= F, \tag{3.47} \\
B u &= 0,
\end{align*}
\]

where \( A \) is a linear operator, \( F \) is the inhomogeneous "forcing" and \( B \) are the boundary conditions, and the homogeneous adjoint system,

\[
\begin{align*}
A^* v &= 0, \tag{3.48} \\
B^* v &= 0,
\end{align*}
\]

and assume we have already solved (3.48) for the vector \( v \). If we let \( \langle \cdot, \cdot \rangle \) denote a suitable inner product then

\[
\langle F, v \rangle = \langle Au, v \rangle = \langle u, A^* v \rangle = \langle u, 0 \rangle = 0. \tag{3.49}
\]

The first step in (3.49) follows from (3.47), the second from the definition of the adjoint operator and the final step follows from (3.48). This demonstrates that if the full inhomogeneous boundary value problem is to have a solution then the inhomogeneous "forcing" \( F \) must be orthogonal to the solution \( v \) of the homogeneous adjoint problem.
At first order, the solvability condition in (3.46) is satisfied because
\[
\int_{-1}^{1} v^T (P_y y \partial_x A(X, T), P_y y \partial_x A(X, T)) \, dy
\]
\[
= \int_{-1}^{1} (\kappa_1 + \kappa_2) P_y y \partial_x A(X, T) \, dy
\]
\[
= \left[ \frac{1}{2} (\kappa_1 + \kappa_2) P_y y^2 \partial_x A(X, T) \right]_{-1}^{1} = 0. \tag{3.50}
\]
Now we can find the solution for \( u^1 \). Writing the \( u^1 \) as
\[
u^1 = P_y \Phi_i(y) \partial_x A(X, T), \tag{3.51}
\]
for some function \( \Phi_i(y) \), the first order equations in (3.45) become
\[
\bar{D}_{11} \phi^2 + \kappa_1 (\Phi_1 - \Phi_2) = y, \tag{3.52}
\]
\[
\bar{D}_{22} \phi^2 + \kappa_2 (\Phi_2 - \Phi_1) = y,
\]
where the \( \Phi_i \) satisfy the boundary condition:
\[
d_y \Phi_i(\pm 1) = 0. \tag{3.53}
\]
The solutions of (3.52), easily obtained using MAPLE and subject to the boundary conditions in (3.53), are
\[
\Phi_1(y) = \frac{1}{6} \left( \frac{\kappa_1 + \kappa_2}{\kappa_1 D_{22} + \kappa_2 D_{11}} y^3 \right) + \nu y + \frac{\eta D_{11} D_{22}}{\sqrt{\zeta}} \tan \left( \frac{\sqrt{\zeta} y}{D_{11} D_{22}} \right),
\]
\[
\Phi_2(y) = \frac{1}{6} \left( \frac{\kappa_1 + \kappa_2}{\kappa_1 D_{22} + \kappa_2 D_{11}} y^3 \right) + \frac{\kappa_1 (\bar{D}_{11} - \bar{D}_{22})}{\kappa_1 D_{22} + \kappa_2 D_{11}} y + \nu y
\]
\[
- \frac{\eta \kappa_2 (\bar{D}_{11})^2}{\sqrt{\zeta}} \tan \left( \frac{\sqrt{\zeta} y}{D_{11} D_{22}} \right). \tag{3.54}
\]
The constant terms $\nu$, $\eta$ and $\zeta$ are defined to be

$$
\nu = \frac{1}{2} \frac{(\kappa_1 + \kappa_2)(\kappa_1 \bar{D}_{11}^{yy} + \kappa_2 \bar{D}_{22}^{yy}) + 2\kappa_1 \bar{D}_{11}^{yy}(\bar{D}_{11}^{yy} - \bar{D}_{22}^{yy})}{(\kappa_1 \bar{D}_{22}^{yy} + \kappa_2 \bar{D}_{11}^{yy})^2},
$$

(3.55)

$$
\eta = \frac{\kappa_1 \bar{D}_{22}^{yy}(\bar{D}_{11}^{yy} - \bar{D}_{22}^{yy})}{(\kappa_1 \bar{D}_{22}^{yy} + \kappa_2 \bar{D}_{11}^{yy})^2}, \quad \zeta = \bar{D}_{11}^{yy} \bar{D}_{22}^{yy} (\kappa_1 \bar{D}_{22}^{yy} + \bar{D}_{22}^{yy} \kappa_1).
$$

So, the equations at $O(\varepsilon^2)$ are

$$
\bar{D}_{ij} \frac{\partial^2 u_j^2}{\partial y^2} + P_e \frac{\partial u_j^2}{\partial y} = P_e \frac{\partial X u_j^1}{\partial y} - 2 \bar{D}_{ij} \frac{\partial X}{\partial y} \frac{\partial u_j^1}{\partial y} + \frac{\partial^2 u_j^2}{\partial y^2} - Q_j u_j^2 - \bar{N}_i(u^0)
$$

(3.56)

If we now apply the solvability condition at $O(\varepsilon^2)$, after a little algebra we have the following amplitude equation for $A(X,T)$:

$$
\partial_T A = D \frac{\partial^2 A}{\partial X^2} + \sigma A + \gamma A^d,
$$

(3.57)

and the constant $\sigma = (\kappa_2(Q_{11} + Q_{12}) + \kappa_1(Q_{22} + Q_{21})) / (\kappa_1 + \kappa_2)$. The constant term $\gamma$ depends upon the $N_{i\kappa_1 \kappa_2}$, $\kappa_1$ and $\kappa_2$ in a similar way to $\sigma$ depending on the $Q_{ij}$, $\kappa_1$ and $\kappa_2$. The effective diffusivity, $D$, again obtained easily using MAPLE to apply the solvability condition at $O(\varepsilon^2)$ using the previously computed solutions for the $\Phi_i(y)$, can be split as follows,

$$
D = \frac{\bar{D}_{11}^{xx} \kappa_2 + \bar{D}_{22}^{xx} \kappa_1}{\kappa_1 + \kappa_2} + P_e^2 \frac{2(\kappa_1 + \kappa_2)}{15(\bar{D}_{11}^{yy} \kappa_2 + \bar{D}_{22}^{yy} \kappa_1)} - P_e \frac{2(\bar{D}_{11}^{yy} \kappa_2 + \bar{D}_{22}^{yy} \kappa_1)}{3(\bar{D}_{11}^{yy} \kappa_2 + \bar{D}_{22}^{yy} \kappa_1)}
$$

$$
+ P_e^2 \frac{4 \kappa_1 (\bar{D}_{11}^{yy} - \bar{D}_{22}^{yy})}{3(\bar{D}_{11}^{yy} \kappa_2 + \bar{D}_{22}^{yy} \kappa_1)^2} \left((\kappa_1 + \kappa_2) \bar{D}_{22}^{yy} - 2 \kappa_1\right)
$$

(3.58)

$$
+ P_e \frac{2(\bar{D}_{11}^{yy} \kappa_2 + \bar{D}_{22}^{yy} \kappa_1)}{3(\bar{D}_{11}^{yy} \kappa_2 + \bar{D}_{22}^{yy} \kappa_1)^2} (\bar{D}_{11}^{yy} (\bar{D}_{11}^{yy} \kappa_2 + \bar{D}_{22}^{yy} \kappa_1) + 2 \bar{D}_{22}^{yy} \kappa_1) + (P e^2 \int_{-1}^{1} \tan(\sqrt{\zeta} y) \text{terms}).
$$
The first term in (3.58) is the effective diffusivity of the reaction-diffusion system (which we assume to be positive) and the second corresponds to Taylor's shear dispersion. The third and fifth terms are where the off-diagonal nature of the diffusion tensor presents itself. These terms in $\tilde{D}_{ij}$ and $\tilde{D}_{ji}$ occur at $O(\text{Pe})$ whereas the terms involving only the vertical diffusion terms, $\tilde{D}^{\text{vw}}_{ii}$ and $\tilde{D}^{\text{ww}}_{ii}$, appear at $O(\text{Pe}^2)$ suggesting that these new terms due to the interaction of the shear and the shape of the organisms have more effect on the stability characteristics of the system with imposed shear. In the limit of no-flow ($\text{Pe} \to 0$), equation (3.5.7) reduces to the reduction of the reaction-diffusion problem and in the limit of the organisms' eccentricity tending to zero (spherical organisms), the amplitude equation is of the same form as that seen in [96].

Even if we assume that $\tilde{D}_{ij}^{\text{vw}} \kappa_2 + \tilde{D}_{ji}^{\text{vw}} \kappa_1 > 0$ and $\tilde{D}^{\text{vw}}_{ii} \kappa_2 + \tilde{D}^{\text{ww}}_{ii} \kappa_1 < 0$, the sign of $D$ is uncertain due to the fact that we are considering the situation where $\tilde{D}_{ij}^{\text{vw}} \neq \tilde{D}_{ji}^{\text{vw}}$. This means that there will be certain cases where the non-linear approximation in (3.57) is ill-posed (due to the negative diffusion coefficient) and, as noted in [96], zero is probably not the critical wavenumber for those cases. If, however, $D > 0$ then we might ask what criteria define the criticality of the instability. If we have that $\sigma > 0$ then small disturbances, $A > 0$, will begin to grow exponentially. The crucial term is the term in $A^{d_1}$, and we assume, for simplicity, that $d_1 = 2$ (i.e. the smallest non-linear term is quadratic). If the coefficient $\gamma < 0$ then we will have saturation of the exponential linear growth and the bifurcation will be supercritical.
The coefficient $\gamma$ can be easily computed to give

$$\gamma = \frac{\kappa_2 \sum_{k_1=0}^{2} N_{1k_10} + \kappa_1 \sum_{k_1=0}^{2} N_{2k_10}}{\kappa_1 + \kappa_2},$$

(3.59)
as only the coefficients of the quadratic non-linear terms appear here. By assumption, $\kappa_1 + \kappa_2 < 0$ so criticality depends upon the sign of the numerator in (3.59). If it is positive then the bifurcation will be supercritical and we can predict that disturbances will develop around the zero wavenumber. The numerator in (3.59) will vary greatly from model to model so no concrete statements on criticality can be made about the zero wavenumber instability in the general reaction model case.

It was mentioned in [96] that, after a linear analysis of their original equation and their amplitude equation, destabilisation was seen at small wave-numbers followed by stabilisation at larger wave-numbers. As in [96], it was seen that the reduced system in (3.57) may well be ill-posed due to the complex diffusive nature of the full system. This effect may well be accentuated by the extra dispersive terms appearing from the interaction of the shear and the organisms spheroidal nature and in these cases, the critical wave-number will probably not be zero. Further to the work in [96], we also discussed the criticality of the shear induced instability when the amplitude equation was well posed. We derived conditions on the linear and non-linear coefficients of the general reaction system which indicate the critical nature of the instability.
3.3 Conclusions

The shape and swimming behaviour of microorganisms can cause the nature of their diffusivity tensors to alter quite radically from the normally employed kind [81], especially when the organisms are subject to advection by even simple velocity fields [11, 41]. In the first section, we demonstrated that a system of swimming, spheroidal organisms, with identical motile and body-shape characteristics, could be driven unstable by imposed shear and that this was a result of the non-diagonal diffusion terms appearing due to the interaction of the shear, the swimming behaviour and the shape characteristics of the organisms themselves. This extends the analysis of [24] to swimming, spheroidal organisms but, contrary to their findings, we see that it is possible for the shear to have a destabilising influence on the system, even for identical diffusive behaviour. In reality, could we observe such an instability? The mechanism involved requires the non-diagonal diffusive term, $\bar{D}^{xy}$, to be greater than the vertical diffusion coefficient, $\bar{D}^{yy}$. The work on gyrotactic spherical particles by Hill & Bees [41] suggested that for small Peclet numbers, this would not be the case. However, when increasing the Peclet number the horizontal and vertical diffusion coefficients seemed to decay faster than the non-diagonal terms. No results are as yet available for spheroidal organisms but the shape characteristics may well interact with the increasing shear (Peclet number) so as to result in a situation where $|\bar{D}^{xy}| > |\bar{D}^{yy}|$ and $\text{det}(\mathbf{D}) \geq 0$, and instability is possible.

In the second section, attention is turned to the case where the diffusion tensors of the interacting organisms are not equal. By considering the equation for the
variations for two species from their averaged equilibrium values, we followed the analysis in [96] and looked at the non-linear behaviour of small disturbances around the zero wavenumber for the no-flow instability. By performing a non-linear analysis about the zero wavenumber, we derived an amplitude equation for the behaviour of small disturbances (in the long horizontal length scale). The equation is similar to that derived in [96] but with additional terms coming from the more complex nature of the swimming terms in the diffusion tensor. As in [96], the equation is not always well-posed suggesting that the zero wavenumber is not the critical wavenumber in these cases. Also, we consider a more realistic form of the general reaction model and derive conditions which determine the criticality of the bifurcation. For supercritical bifurcations, the amplitude equation resembles Fisher's equation which suggests the possible emergence of travelling waves.

For a bifurcation at some non-zero, finite wavenumber the equations become much more complex. Also, the previous analysis dealt with small values of $P_e$, but what is likely to happen for large values of the Peclet number? In [41], using a generalised Taylor dispersion approach it was seen that the situation for large $P_e$ is more complicated than was first thought and that the assumption of equation (3.5) is invalid. As $P_e \to \infty$, the vorticity of the flow constrains the particles to ever decreasing circular paths, allowing no dispersion of the particles. This means that $D_{ij}^* \to 0$ as $P_e \to \infty$ and consequently the diffusion tensors become identical. This analysis was performed with regard to gyrotactic spherical particles but other forces will become important when considering motile spheroidal particles. For example, they are subject to straining forces due to the nature of the shear flow.
For this analysis, we are effectively considering a spheroidal organism which moves in one direction (i.e. along its major axis). The straining field would act by attempting to constrain the organism at some angle $\beta$ to the horizontal (it should be noted that this will not give the organism a preferred swimming direction as the orientation, with respect to the organisms' major axis, has equal probability of being either $\beta$ or $\beta + \pi$ degrees to the horizontal). It has been seen that the cell eccentricity $e$ determines the dispersive behaviour in the limit as $P_e \to \infty$, the vortical effects or the straining effects. If the eccentricity is too small then it is the vorticity induced decay in dispersion that will dominate. In this case, no instabilities are likely to be seen for large Peclet numbers as $D_{ij}^{**} \to 0$. However, if the straining effects become dominant, for sufficiently elongated cells, then, for organisms with differing motile and body shape characteristics, there are likely to be differences still present in the diffusive characteristics (for large Peclet number) suggesting that instability is quite likely.

It was seen by Mitchell et al. [68, 69] that gyrotaxis [81], in the presence of shear, caused small-scale heterogenities of plankton. The added straining influence exerted on spheroidal organisms can also seemingly cause different dispersive behaviour in swimming plankton and small-scale patchiness may well be heavily influenced by physiological and motile responses to the surrounding flow fields.
Effects of excitability, inertia and turbulent advection

All oceanic organisms will experience advection by turbulence but the possible effects on the population dynamics and distribution of these organisms are not well understood. Turbulence can cause inhomogeneous distributions of advected organisms when reaction dynamics are considered, [2]. Moreover it has become clear that we cannot necessarily assume that plankton behave like passive scalars as they have some motile abilities [31] as well as different viscous and inertial properties to the surrounding fluid [31, 65].

It has been observed that very different aggregation zones occur for particles advected by turbulence if their densities are different from the surrounding fluid [91, 97]. For a range of parameters it was seen that light particles are forced into areas of large vorticity while heavier particles are constrained to regions of high strain. If one of the organisms is light/heavy this means we will have some degree of separation of predator and prey. At large spatial scales, phytoplankton blooms have been observed where the population of phytoplankton rapidly rises and then sinks...
back. This is the hallmark of an excitable system and Truscott & Brindley [107] presented a two component PZ model which has the characteristics of an excitable system whose excitability is robust over a realistic parameter range. Possibly many mechanisms such as salinity, temperature and mixing help cause this bloom but we would like to look at the contribution of inertial separation by turbulent advection.

In [16], the breaking of excitation waves by shear flow was suggested as a possible mechanism for plankton patchiness and [71] investigated a chaotically advected system with some driving some influx of nutrient (in this case iron) to the system. In both cases the system needs some perturbation to see excitation (in the latter case it was exactly this addition of iron). The scenario we would like to present generally requires no extraneous perturbation and is driven solely by the interaction of inertial forces and turbulent advection.

In particular, we would like to address the question of whether the separation of phytoplankton and zooplankton is sufficient to seed a bloom. If a large proportion of the phytoplankton are allowed to grow without predation then this could cause a supra-threshold type perturbation indicative of excitable systems. Also, what consequences does this inertial separation and bloom propagation have on the distribution of the organisms and the predator-prey contact rates? Both these questions are important for the understanding of the possible mechanisms of plankton blooms and its consequences on the recruitment of larval fish, a subject of paramount importance to fisheries managers.
4.1 Turbulent advection of inertial organisms

The assumption that zooplankton and phytoplankton can be treated as passive scalars is at best hopeful. They can both, in varying degrees, have different viscous and inertial [31, 98] properties to the surrounding fluid.

4.1.1 Viscous and inertial effects

In [63] the equation of motion was derived for a spherical particle at position $X_n$, with velocity $V_n$, where $n$ is the particle index, in a non-stationary velocity field, $U(X_n)$:

\[ m_p \frac{dV_n}{dt} = m_f \frac{DU(X_n)}{Dt} + 6\pi \alpha \mu (U(X_n) - V_n) \]

\[ -\frac{m_f}{2} \left( \frac{dV_n}{dt} - \frac{dU(X_n)}{dt} \right) - 6\pi \alpha \mu \int_{-\infty}^{t} \frac{d(V_n - U(X_n))/dr}{\sqrt{\nu(t - \tau)}} d\tau. \]

where $m_p$ and $m_f$ are the particle and fluid masses respectively, $\nu$ and $\mu$ are the kinematic and dynamic viscosities, respectively, and $\alpha$ is the radius of the particle.

We bring attention to the differing derivative terms, $D/dt$ and $d/dt$. The first represents the material derivative (following a fluid element) while the second represents the derivative following the particle in the fluid and

\[ \frac{dU(X_n)}{dt} = \frac{\partial U(X_n)}{\partial t} + (V_n \cdot \nabla) U(X_n), \]

as opposed to the more familiar definition of $D/dt$. The first term is the Bernoulli term which is the force from the undisturbed flow, the second term is the Stoke's viscous drag, the third is the added mass term and the final term is the Basset
history force. The equation in (4.1) incorporates the assumption that the particle, associated Reynolds number and fluid gradients around the particle are small [63].

If we further assume that the drag and inertia terms are dominant (and the particle is small) then we may neglect the Basset history term. To further simplify, matters we follow Auton et al. [10] by assuming that the work done displacing a fluid element must be dependent on the forces on the fluid at that point. With this in mind we can exchange the particle derivative for the fluid derivative in the added mass term to obtain the following reduced version of (4.1)

$$\frac{dV_n}{dt} = \alpha (U(X_n) - V_n) + R \frac{DU(X_n)}{Dt},$$  \hspace{1cm} (4.3)

and $\alpha = 6\pi a \mu$ which is just the Stoke's viscous drag parameter from (4.1). If we consider the length and velocity scales of the undisturbed flow to be $l_0$ and $u_0$, respectively, then we if we re-scale time in (4.3), using the flow timescale $l_0/u_0$, we obtain the following,

$$\frac{dV_n}{dt} = A(U(X_n) - V_n) + R \frac{DU(X_n)}{Dt}.$$  \hspace{1cm} (4.4)

where

$$A = \frac{6\pi a \mu l_0}{(m_p + \frac{1}{2} m_f) u_0}, \quad R = \frac{3m_f}{2m_p + m_f},$$  \hspace{1cm} (4.5)

so $A$ becomes a balance of viscous and inertial forces and will henceforth be referred to as the dimensionless inertia parameter. For neutrally buoyant particles, $R = 1$, heavy particles $R < 1$ and for light particles $R > 1$. 

64
4.1.2 Derivation of the long-time equation of motion

Later we shall move to an Eulerian frame so we need to establish the existence of an Eulerian velocity field. To obtain a simplified, long time expression for the corrected velocity field we follow the approach of [64, 91]. Henceforth, we refer to \( V(r) \) as the effective velocity field at position \( r \) associated with \( V_n \) for all \( n \). The analysis begins by reforming (4.4) as a non-linear integro-differential equation as follows:

\[
X(t) = X(0) + \frac{1}{A} V(0)(1 - e^{-At}) + \int_0^t U(X(s))(1 - e^{-A(t-s)}) ds + \frac{R}{A} \int_0^t (U(X(s)) \cdot \nabla)U(X(s))(1 - e^{-A(t-s)}) ds,
\]

(4.6)

and integration by parts yields

\[
X(t) = X(0) + \frac{1}{A} V(0) + \int_0^t U(X(s)) ds - \frac{1}{A} U(X(t)) + e^{-At} \left( U(X(0)) - V(X(0)) + \frac{R}{A}(U(X(0)) \cdot \nabla)U(X(0)) \right) + \frac{R}{A} \int_0^t (U(X(s)) \cdot \nabla)U(X(s)) ds - \frac{R}{A^2} (U(X(t)) \cdot \nabla)U(X(t)) + O(\frac{1}{A^2}).
\]

(4.7)

If we now assume \( A \) to be suitably large, we can neglect the effect of the exponential term in (4.7). This gives us a long time equation of motion, obtained by differentiating (4.7) after the exponential terms have decayed, which yields

\[
\frac{dX(t)}{dt} = U(X) + \frac{R}{A} (U(X) \cdot \nabla)U(X) - \frac{1}{A} \frac{dU(X)}{dt} + O(A^{-2}),
\]

(4.8)

and naturally \( \frac{dX}{dt} = V \). If we now apply (4.3) we have that

\[
\frac{dU(X)}{dt} = \frac{\partial U(X)}{\partial t} + (U(X) \cdot \nabla)U(X) + O(A^{-1}),
\]

(4.9)
and, because the ambient flow is stationary, we have the following expression [91] for the effective velocity field, \( \mathbf{V}(r) \),

\[
\mathbf{V}(r) = \mathbf{U}(r) + \frac{R-1}{A} (\mathbf{U}(r) \cdot \nabla) \mathbf{U}(r) + \mathcal{O}(A^{-2}).
\] (4.10)

4.1.3 Aggregation of non-neutrally buoyant organisms

Aggregation of particles will occur in regions of negative divergence, \( \nabla \cdot \mathbf{V} < 0 \), and the divergence of (4.10) can be written in terms of the local strain, \( S = \sqrt{S_{||}^2 + S_{\perp}^2} \) \( (S_{||} = \partial_y u + \partial_x v, \ S_{\perp} = \partial_x u - \partial_y v) \) and the local rotation \( (2 \times \text{vorticity}), \Omega = \partial_y u - \partial_x v, \) of the original turbulent flow, \( \mathbf{U}(r) = (u, v) \), as follows. Firstly, we note that

\[
(U \cdot \nabla) \mathbf{U} = (\partial_x u + \partial_y v)^2 - 2\partial_x u \partial_y v + 2\partial_y u \partial_x v + (u \partial_x + v \partial_y)(\partial_x u + \partial_y v)
\]

\[
= 2(\partial_y u \partial_x v - \partial_x u \partial_y v),
\] (4.11)

by application of the incompressibility condition, \( \partial_x u + \partial_y v = 0 \). From the above definitions and after removing the terms that cancel each other out,

\[
S^2 - |\Omega|^2 = (\partial_x u)^2 + (\partial_y v)^2 + 4\partial_y u \partial_x v - 2\partial_x u \partial_y v,
\] (4.12)

\[
= (\partial_x u + \partial_y v)^2 + 4(\partial_y u \partial_x v - \partial_x u \partial_y v),
\]

and by applying the incompressibility condition again in conjunction with (4.11), we obtain the following expression for the divergence of the inertial velocity field,

\[
\nabla \cdot \mathbf{V} = \frac{R-1}{2A} (S^2 - |\Omega|^2),
\] (4.13)

which leads to two different possibilities for aggregation. If the organism is heavy \( (R < 1) \) then accumulation occurs in regions where \( S^2 > |\Omega|^2 \), regions of high strain.
and low vorticity. Conversely, if an organism is lighter than the surrounding medium \((R > 1)\) accumulation occurs in regions where \(S^2 < |\Omega|^2\), thus in areas with high vorticity and low strain.

### 4.1.4 Spectral characteristics of inertial particles in turbulence

In this section, we discuss the role that inertia plays in determining the energy spectrum of particles immersed in turbulent flows. This situation has been considered by Fung et al. [35], for neutrally buoyant particles in unsteady flows and it was seen that the energy spectra of the particle's velocity field and the fluid it was immersed in differed due to the drag forces. For passive advection, the wavenumber-dependent energy spectrum of the particle is just that of the ambient flow, \(E(k)\). Given the long term equation of motion in (4.10), we can gain some analytical insight into how the spectral properties of the inertial particles might vary from those of the fluid they are immersed in. Below, we will take the Fourier transform of equation (4.10) but, following Powell & Okubo [88], replace the complicated non-linearity by utilising a wavenumber-dependent diffusivity, \(D(k)\), defined to be

\[
k^2 D(k) \tilde{U}_i(k) = \int dk' k'_j \tilde{U}_j(k-k') \tilde{U}_i(k').
\]

The term on the right hand side of the above equation is the appropriate Fourier component of the non-linear term in the Navier-Stokes equations [70] and the Fourier transform of \(U(r)\), \(\tilde{U}(k)\), is defined in the usual way,

\[
\tilde{U}(k) = \int dr U(r)e^{ikr}.
\]
So, in Fourier space and with the diffusive simplification, equation (4.10) is defined to be

\[ \tilde{\mathbf{v}}(\mathbf{k}) = \bar{\mathbf{v}}(\mathbf{k}) + \frac{R-1}{A} k^2 D(\mathbf{k}) \tilde{u}(\mathbf{k}). \]  

(4.16)

We let \( F(k) \) be the energy spectrum of the corrected velocity, \( \mathbf{\tilde{V}}(\mathbf{k}) \), and following [88] one can now suitably integrate over all wave-numbers \( k \) in Fourier space to obtain an expression for the corrected energy spectrum,

\[ F(k) = E(k) \left( 1 + \frac{k^2 D(k)(R-1)}{A} \right), \]  

(4.17)

and the \( k \) dependence disappears as we have integrated over the whole Fourier space which yields the "averaged" diffusion coefficient, \( D(k) \). Hence, similar to (4.10), the energy spectrum consists of both that of the original flow and a correction term, whose magnitude depends on both the inertial and viscous properties of the particle.

Okubo [73], based on empirical observations of dye dispersal in the ocean, derived an approximate relationship between the length-scale and the horizontal 2D diffusivity due to turbulent motion. For a given length scale, \( l \), the diffusivity, \( D(l) \), was given by

\[ D(l) = 0.01 l^{1.15} \text{cm}^2 \text{s}^{-1}. \]  

(4.18)

Now since \( l \propto k^{-1} \) we can approximately say that, from the above relationship,

\[ D(k) \propto k^{-\frac{2}{3}}, \]  

(4.19)

which, as noted by Okubo [73], is slightly different to the analytical \( D(k) \propto k^{-\frac{2}{3}} \) expression [70]. Substituting this relationship into (4.17) yields the following novel
result,

\[ F(k) \approx E(k) \left(1 + \varepsilon k^\frac{\beta}{2}\right), \]  

(4.20)

where

\[ \varepsilon = \frac{\beta(R - 1)}{A}, \]  

(4.21)

and \( \beta \) is the constant of proportionality from (4.19). Due to the fact that by assumption, \( A \) is large, and that \( \beta \) is \( \mathcal{O}(10^{-1}) \) [73] the correction coefficient \( \varepsilon \) will be small, as was the correction value in (4.10). Figure (4.1) plots the normal, \( E(k) \), and corrected, \( F(k) \), Kraichnan [53] energy spectra for both light and heavy inertial regimes.

Figure 4.1: Normal Kraichnan energy spectrum (solid line) and the corrected spectrum (dashed line) for light (left; \( A = 10, R = 2 \)) particles and heavy (right; \( A = 10, R = 0 \)) particles.

What implications does this have for the energy spectra of inertial particles?
Firstly, we briefly explain what the energy spectrum means in terms of physical quantities. Standard results from turbulence work [70] state that the characteristic mean squared velocity of the flow, \( u_0^2 \), (which determines the appropriate characteristic velocity scale) is given by the expression
\[
\frac{1}{u_0^2} = \frac{\langle U_i(r)U_i(r) \rangle}{\langle U_i\rangle^2} = \int_0^{\infty} dk E(k),
\]
(4.22)

As we have seen, for lighter particles \( R > 1 \) so the correction term will be positive suggesting that, for less dense particles, \( F(k) \geq E(k) \) for all \( k \), and from (4.22) this means the lighter particles will have a higher mean square kinetic energy than neutrally buoyant particles would have. Does this agree with other theoretical work on 2D turbulence? Kraichnan’s [53] pioneering work on 2D turbulence recognised that the energy was contained in the large-scale eddies. Furthermore, that the enstrophy (in turbulent flow this is the mean squared vorticity [88]), and its related dissipation rate, \( \theta(k) \), characterised the turbulence in this regime. Looking at figure (4.1), we see that the difference in the passive and corrected spectra is greatest around the characteristic wavenumber (in this case \( k_0 = 1 \)) in the Kraichnan spectrum. This wavenumber corresponds to the length-scale of the largest eddies in the flow, the eddies where the turbulent kinetic energy is generated. As was demonstrated in section 4.1.3, lighter particles will accumulate in the energy containing eddies hence they are more likely to have a larger energy spectrum at these corresponding wave-numbers than passive particles simply following the flow.

For heavy particles, the correction coefficient will be negative and consequently,
the largest wavenumber we can consider (after which \( F(k) < 0 \)) is

\[
k = k_{\text{max}} = \varepsilon^{-\frac{3}{5}} \gg 1,
\]

so for heavier particles \((R < 1)\) we have that \( F(k) < E(k) \), for \( k \in [0, k_{\text{max}}] \). This also makes intuitive sense as heavier particles are constrained to the regions outside the eddies, where the turbulent energy is generated. Also, as before this difference in the two spectra is greatest at the large eddy scales.

The expression in (4.20) should be readily testable, especially with the type of model for turbulent flow we consider in the following section as we can define the exact energy spectrum of the flow. We can consider the following general form for equation (4.20),

\[
F(k) = E(k) \left(1 + \varepsilon k^n \right).
\]

Since \( F(k) \) can be numerically calculated and both \( E(k) \) and \( \varepsilon \) are known we may compute the suitable value of the index \( n \). If \( n \approx 5/6 \) then this should be fairly convincing evidence that the expression in (4.20) is fairly accurate. The diffusive approximation made in (4.14) was found in [88] to give the familiar \( k^{-\frac{4}{3}} \) and \( k^{-3} \) power laws for Kolmogorov's inertial sub-range and Kraichnan's two-dimensional turbulence via dimensional analysis. It is thus a valid approximation, except for large wavenumbers in the heavy particle regime as it predicts negative energy at suitably large wavenumbers. To sum up, the analysis suggests that the energy in the corrected particle velocity fields for lighter/heavier particles will be higher/lower than that of the ambient fluid due to the aggregation caused by the inertial bias.
4.1.5 Synthetically turbulent flow

We work with a statistically homogeneous, isotropic and stationary two-dimensional velocity field (but can be extended to three dimensions) which represents a synthetic, or kinematic, turbulent flow first introduced in [62]. This flow corresponds to the velocity of the elements of the fluid in which the particles, \( U \), are immersed, although in the presence of inertia the velocity of these particles, \( V \), is different from the velocity of the fluid at the particle position.

The starting point of our two-dimensional simulations is a Langevin equation for the stream function, \( \eta(r, t) \), first employed by Marti et al. [62],

\[
\partial_t \eta(r, t) = \nu \nabla^2 \eta(r, t) + Q[\lambda^2 \nabla^2] \nabla \cdot \zeta(r, t),
\]

(4.25)

where \( \nu \) stands for the kinematic viscosity, \( Q[\lambda^2 \nabla^2] \) denotes an operator which controls the spatial correlations, with \( \lambda \) being the typical length scale of the flow and \( r \) and \( t \) represent the location and time variables. Furthermore, \( \zeta(r, t) \) is a Gaussian white-noise field with zero mean value and whose covariance is given by

\[
\langle \zeta^i(r, t) \zeta^j(r', t') \rangle = 2\epsilon_0 \delta(t - t') \delta(r - r') \delta^{ij}
\]

(4.26)

where \( \epsilon_0 \) is the parameter that determines the intensity of the noise and the \( \delta(\cdot) \) and \( \delta^{ij} \) are the Kronecker-delta functions. The Gaussian process is used to simulate an intrinsically random stirring force, statistically governed by the stirring operator, \( Q[\cdot] \). The Langevin equation can be formally integrated to get the temporal evolution of the stream function. The incompressible two-dimensional velocity field \( U \) is now
obtained in the usual manner and
\[
U(r, t) = \left( -\frac{\partial \eta(r, t)}{\partial y}, \frac{\partial \eta(r, t)}{\partial x} \right).
\] (4.27)

The two-point velocity correlation is defined to be
\[
C_{ij}(r - r', t - t') = \langle U_i(r, t) U_j(r', t') \rangle,
\] (4.28)
and the radial correlation function is defined as
\[
R(r, s) = \frac{1}{2} [C_{xx}(r, s) + C_{yy}(r, s)],
\] (4.29)
where \( r = |r - r'| \) and \( s = |t - t'| \). The principal kinematic properties of the flow are:

1. The mean square velocity (intensity), \( u_0^2 = R(0, 0) = \int_0^\infty E(k) \, dk \), where \( E(k) \) is the energy spectrum of the flow.

2. The characteristic time scale, \( t_0 = \frac{1}{u_0^2} \int_0^\infty ds R(0, s) \).

3. The characteristic length scale, \( l_0 = \frac{1}{u_0^2} \int_0^\infty dr R(r, 0) \).

The parameter \( u_0 \) is fairly self explanatory, the parameter \( l_0 \) determines the size of the large scale vortices and the time-scale \( t_0 \) determines the temporal lifespan of these eddies. For example, with the “frozen” flow \( t_0 = \infty \) and the flow remains fixed for all time. These parameters can, in turn, be computed from the original parameters, \( \mu \), \( \nu_0 \) and \( \lambda \) [62]. In this work, only the Kraichnan [54] spectrum is considered as it is a generic feature of two-dimensional turbulence. Kraichnan’s spectrum displays a well defined peak at a wavenumber \( k = k_0 \). Other spectra such as the Karman-Obhukov
spectrum were developed to look at Kolmogorov turbulence [51, 70] and subsequently
display the long 

"-5/3" tail in the inertial subrange. Such spectra can be similarly
applied but care must be taken when accounting for planktonic dispersion at smaller
spatial scales. The Kraichnan spectrum has the spectral property that

\[ E(k) \propto k^3 \exp \left( -\frac{k^2}{k_0^2} \right), \]  

where \( k_0 \) is the characteristic wavenumber. This means that the stirring operator \( Q \)
should be defined such that

\[ Q[\lambda^2 \nabla^2] = \exp \left( \frac{\lambda^2 \nabla^2}{2} \right). \]

The Fourier transform of (4.25) is

\[ \tilde{\delta}(\eta(k, t)) = -\mu k^2 \eta(k, t) - iQ(-\lambda^2 k^2)k^i \tilde{\zeta}^i(k, t), \]

where \( k^2 = ||k||^2 \), \( k^i \) is simply the \( j \)th component of the vector \( k \) in Fourier space,
and summing over indices is understood. To gain an insight into the movement
of energy in the flow, we can introduce the structure function \( S(k, t) \) [62] which is
defined in terms of the equal time auto-correlation function, \( \langle \eta(\cdot) \eta(\cdot) \rangle \), such that

\[ 4\pi^2 \delta(k + k')S(k, t) = \langle \eta(k, t) \eta(k', t) \rangle. \]

From [62], by combining this with the expression in (4.32), we obtain

\[ \partial_t S(k, t) = -2\nu k^2 S(k, t) + 2\epsilon_0 \nu k^2 Q^2(-\lambda^2 k^2) \]

and the energy spectrum, \( E(k, t) \), satisfies \( E(k, t) = k^3 S(k, t)/(4\pi) \). Combining this
relationship with the above expression tells us that

\[ \partial_t E(k, t) = -2\nu k^2 E(k, t) + 2k^2 W(k) \]
where

\[ W(k) = \frac{\nu \epsilon}{4\pi} k^3 Q^2(-\lambda^2 k^2). \]  

In the steady-state, we see that \( E(k) = \nu^{-1} W(k) \). For more of the details on the synthetic flow and its properties, the interested reader is referred to the paper by Marti et al. [62]. It is now clear to see from (4.35) that the stream function has a balance between energy input, from the stirring operator \( Q \), and energy dissipation from the viscous term (at time and length scales dictated by \( \nu \)). In Fourier space, we see that (4.25) corresponds to building up the field from its independent Fourier modes. Using this type of model for turbulence implies a change from an intrinsic randomness (associated with the complex behaviour resulting from the non-linearity of the Navier-Stokes equations) into a system of independent Fourier modes coupled to an external noise with prescribed statistical characteristics. The algorithm for the numerical integration of (4.25) in discrete Fourier space may be found in the Appendix.

A detailed presentation of the algorithm just introduced can be found in [62]. Although the general method of flow generation allows us to work with a non-stationary field, we are not going to make use of this possibility here. For the sake of simplicity, and in order to avoid the existence of a great number of temporal scales, we take a stationary flow, so that, we fix \( t_0 = \infty \) in our flow simulations.

The original equation derived by Maxey & Riley [63] has certain requirements for its safe application. The main condition is that \( (a^2 u_0)/(l_0 \nu) \ll 1 \). Since we have chosen that \( a \leq 10^{-2} \text{m}, \ u_0 = 0.1 \text{m/s}^{-1}, \ l_0 = 5 \times 10^4 \text{m} \) and \( \nu = 10^{-6} \text{m}^2 \text{s}^{-1} \) we can
Figure 4.2: *Snapshot of the "frozen" turbulent velocity field. The direction of the lines represent the direction of the flow at that point and the length of the lines is the magnitude of the flow there.*

safely use the Maxey & Riley equation. In line with a realistic range of values for $\alpha$ from [65, 98] we henceforth consider $\alpha = 0.0257s^{-1}$, this corresponds to zooplankton of a radius of around 1 cm and implies that $A \gg 1$, justifying the preceding analysis.

We use a steady flow because any temporal changes in the undisturbed flow, \( \mathbf{U} \), will be followed by an exponential decay in \( \mathbf{U}(\mathbf{X}) - \mathbf{V} \) at a rate $\propto A$, because of the first term in the equation for the time evolution of the corrected velocity in (4.4). Since $A \gg 1$, this means that any effects due to structural changes in the flow will happen on a time scale much faster than any changes in the reaction dynamics, which supports the use of a stationary flow representatively.
In figure (4.2) we show a flow realisation used in the simulations. This flow has been generated using a square lattice of 128 x 128 points, and \( l_0 \) corresponds to a tenth part of the linear size of and also roughly the size of the large eddies in the flow.

### 4.2 Incorporating reaction kinetics and diffusion

To include the effects of planktonic interactions and diffusion/motility we now work in an Eulerian frame with the reaction-diffusion-advection equation:

\[
\frac{\partial P}{\partial t} = -\nabla \cdot (V^P P - D_P \nabla P) + R_P(P, Z),
\]

\[
\frac{\partial Z}{\partial t} = -\nabla \cdot (V^Z Z - D_Z \nabla Z) + R_Z(P, Z),
\]

where \( V^i \) represents the velocity, with inertial and viscous corrections, of species \( i \) from advection by the turbulent flow at length-scales above the grid size. The diffusion terms \( D_i \) represent the dispersion of plankton beneath the scales that the flow cannot simulate due to the rapid decay in the Kraichnan spectrum. Truscott & Brindley [107] suggested an excitable model for phytoplankton-zooplankton interactions and

\[
R_P(P, Z) = rP \left( 1 - \frac{P}{K} \right) - \frac{\gamma Z P^2}{P^2 + \kappa^2},
\]

\[
R_Z(P, Z) = \frac{e \gamma Z P^2}{P^2 + \kappa^2} - \delta Z,
\]

where \( P \) and \( Z \) are a measure of the biomass density of the phytoplankton and zooplankton, respectively.
In this model, \( r \) and \( K \) represent the growth rate and carrying capacity of the phytoplankton, in the absence of zooplankton, respectively. The parameters \( \gamma \) and \( \kappa \) are the zooplankton grazing rate and the half-saturation constant of the grazing response while finally \( e \) is the grazing efficiency and \( \delta \) the zooplankton mortality rate.

The nullclines for the above model are:
\[
\dot{P} = 0 \Rightarrow Z = \frac{r(1 - P/K)(P^2 + \kappa^2)}{P},
\]
\[
\dot{Z} = 0 \Rightarrow P = \frac{\delta \kappa^2}{1 - \delta},
\]
which, in the phase plane, typically look like those sketched in figure (4.3).

We non-dimensionalise (4.37) by rescaling such that
\[
P = Kp, \quad Z = Kz, \quad x = l_0 \bar{x}, \quad y = l_0 \bar{y}, \quad t = \frac{t l_0}{u_0}, \quad V^i = u_0 \bar{v}^i,
\]
and the \( r \) employed here is different to that seen in chapter 2. After dropping the tildes this gives us the non-dimensional version of (4.37),
\[
\frac{\partial p}{\partial \tau} = -\nabla \cdot (v^p p - d_p \nabla p) + \beta p(1 - p) - \frac{\gamma p^2}{p^2 + \chi^2},
\]
\[
\frac{\partial z}{\partial \tau} = -\nabla \cdot (v^z z - d_z \nabla z) + e \left( \frac{\gamma z^2}{p^2 + \chi^2} - \sigma z \right),
\]
where the dimensionless parameters are defined to be
\[
\beta = \frac{r l_0}{u_0}, \quad \epsilon = e, \quad \chi = \frac{\kappa}{K}, \quad \sigma = \frac{\delta l_0}{u_0 e}, \quad \tilde{\gamma} = \frac{\gamma l_0}{u_0}, \quad d_i = \frac{D_i}{u_0 l_0}.
\]

4.2.1 Parameter values for the full simulation model

We chose \( L \) as the length scale \( \approx 50 \text{km} \) (and \( l_0 \) is related to \( L \) in the sense that \( l_0 \) determines the scale of the eddies and in this case \( L \approx 10 l_0 \)) and the characteristic
Figure 4.3: Sketch of the nullclines of the excitable system in (4.38). Sub-threshold perturbations (A) return quickly to the equilibrium, E. Supra-threshold perturbations however (from B) go on a long excursion from E but still returns to E (as E is linearly stable).

velocity is to be $u_0 = 0.1 \text{ms}^{-1}$, similar to that used in both [2] and [71]. Following Truscott & Brindley [107] we set $\beta = 0.43$, $\chi = 0.053$, $\sigma = 0.34$ and $\epsilon = 0.05$ (so the zooplankton is the slow/recovery variable). The parameter $\gamma$ is set equal to one so advective and reactive effects are comparable. Following Marti et al. [62], the system is spatially discretised using a square lattice of $128 \times 128$ cells, with a grid spacing of 0.5, and the time step used in the numerical scheme is $\Delta t = 0.001$.

Therefore, the non-dimensional length of the region is 64. Kraichnan’s spectrum
has a maximum occurring at $k = 0.75$ and drops to a very small value at $k = 2$
(this is not true of the Karman-Obhukov spectrum which has relatively high energy
for large $k$). Hence, the length-scale associated with the highest energy is equal to
$2\pi/0.75 = 8.4$, or about a tenth of the length of the simulation region. The length-
scale associated with a wavenumber of 2 (residual energy is given by $2\pi/2 = \pi$, or
about one 20th of the simulation region). If the simulation region is 50 km then
the dimensional length-scale associated with low energy is $L/20 = 2.5 \times 10^5$ cm.
This is the length-scale for which we wish to implement a turbulent diffusivity into
the reaction-diffusion-advection system. Hence, the diffusivity $D = 0.01 \times \ell^{1.16} =
1.6 \times 10^4 \text{cm}^2 \text{s}^{-1} = 1.6 \text{m}^2 \text{s}^{-1}$ (according to the empirical relationship extracted by
Okubo [73] from experimental data).

Therefore, the synthetic turbulence accounts for large scale flow, centred about a
wavenumber $k_0$. The energy spectrum drops off sufficiently quickly for larger wave-
numbers such that the turbulent mixing can be modelled by simple diffusion at small
length-scales. We neglect any swimming diffusion for the following reason. If we
assume an upper bound on the typical swimming speed, $V_s = \mathcal{O}(10^{-1}) \text{m}^2 \text{s}^{-1}$, and
some typical reorientation time, $\tau_r = \mathcal{O}(1)$ s, then the diffusion can be approximated
by $D_s \approx V_s^2 \tau_r$ which is $\mathcal{O}(10^{-2}) \text{m}^2 \text{s}^{-1}$ (based on data for algal cells from [40]).

The integration of (4.40) was achieved by using a two-step Lax-Wendroff scheme
[89] which gave good numerical convergence for the parameter values employed. For
a review of the Lax-Wendroff scheme and the second order Runge-Kutta scheme [89]
used to integrate the reaction dynamics, see the Appendix.
4.3 Simulations for various advective and inertial regimes

Before we proceed with the simulation results we first state the nature of the scenarios we will consider.

1. **CASE I**: Simple diffusive propagation of the excitation wave from a suitable perturbation.

2. **CASE II**: Spread of a bloom under turbulent advection with no inertial or viscous corrections.

3. **CASE III**: Bloom initiation from an homogeneous initial distribution in an extreme inertial scenario.

4. **CASE IV**: Bloom initiation from an homogeneous initial distribution in the case of neutrally buoyant phytoplankton and slightly heavy zooplankton.

In figure (4.4), the phytoplankton population is shown after a supra-threshold perturbation. The phytoplankton excitation wave propagates via diffusion (CASE I) and turbulent advection (CASE II). In both cases the excitation spreads and eventually dies out with the phytoplankton population returning to the quiescent equilibrium level.

At \( t = 0 \), \( p \) and \( z \) are equal to their equilibrium values \( p^* = 0.03827 \) and \( p^* = 0.04603 \), respectively, except for a small central zone where a value of \( p \) large enough to cause excitation is added. These figures are qualitatively in agreement with
Figure 4.4: Propagation of the phytoplankton excitation wave via diffusion (middle) and turbulent advection (right) from a supra-threshold perturbation (left). The axes represent the \((x, y)\) plane.

[71] in that bloom propagation occurs in a non-uniform manner (for the turbulent flow aided bloom propagation). We do not investigate how the intensity of the turbulence affects bloom propagation (in line with the distinctions between local and global excitations defined in [71]) as we are more interested in the inertial initiation of blooms. What is evident is that the system tends towards the homogeneous equilibrium state, \((p^*, z^*)\), for large times.

For the simulation of inertial organisms (Cases III and IV) we begin with a spatially homogeneous distribution at the unexcited, equilibrium values \(p^*\) and \(z^*\). The inertial separation under turbulent advection acts as the perturbation to the equilibrium distribution and, analogous to the threshold nature of excitable systems, we find that there exist critical values of the Bernoulli parameters which separated regimes where no bloom occurred from regions where we saw localised excitation.

82
and subsequent bloom propagation. In figure (4.5) we see the evolution of both populations in the extreme inertial regime, $R_p = 1.33$ and $R_z = 0.667$ (CASE III). As the phytoplankton accumulate in the eddies the population exceeds the threshold value and excitation begins inside the eddies. Diffusion causes the excitation to move out from the eddies almost covering the whole domain but the flow eventually forces this wave back into the vortices. All the time the excitation is also decaying and these three processes seem to reach an equilibrium state leaving a steady, heterogeneous distribution of plankton, with phytoplankton at the excited state inside the eddies and below it outside. For suitably small Bernoulli parameters the separation is not strong enough to induce excitation. Hence, we reach a steady, quasi-homogeneous distribution around the non-excited equilibrium values.

One of the most interesting cases occurs for the regime where we have neutrally buoyant phytoplankton ($R_p = 1$) and slightly heavy zooplankton ($R_z = 0.9$). Here we see the emergence of an oscillatory bloom in the average phytoplankton population with a period of around 1 year which is interesting when we consider the annual nature of the observed blooms (although seasonal forcing is likely to have a role). In figure (4.7) we see the nature of the oscillatory bloom in the mean population values of phytoplankton and zooplankton and in figure (4.6) we show the evolution of the whole phytoplankton population. The initial excitation forms now due to the absence of zooplankton in the eddies and, as before, the bloom propagates via diffusion throughout the domain. The main difference now is that the expanding phytoplankton wave is not constrained to the vortices so the bloom decays and returns to the quiescent equilibrium before the process begins again. What is more
Figure 4.5: Formation of steady, non-homogeneous phytoplankton (top; a.) and zooplankton (bottom; b.) distributions in the extreme light/heavy inertial regime (CASE III).

fascinating is that this regime is the most physically realistic as phytoplankton are usually neutrally buoyant but zooplankton slightly denser than the surrounding ocean [31, 98].

4.3.1 Reduced model of the vortex dynamics

To elucidate the phenomena observed in the previous section we suggest a more simplistic model of the dynamics in the vortical regions as it would appear that the attraction/expulsion for light/heavy organisms is the driver behind the blooms. In constructing a simplified model, we make the assumption that firstly we need effectively consider only the radial aspect, $r$, of the vortex and hence use a 1D
Figure 4.6: Snapshots of the phytoplankton (top) and zooplankton (bottom) populations during the oscillatory bloom, (CASE IV).

model where the vortex is a domain of length $L_v$. Secondly we assume that the rate of expulsion/attraction is proportional to the distance from the vortex core and the size of the Bernoulli parameter, $R_i$. We model the vortex dynamics of the phytoplankton and zooplankton as follows

$$\frac{\partial p}{\partial \tau} = -\partial_r \left( -\lambda_p r p - d_p \partial_r p \right) + \beta p (1 - p) - \frac{zp^2}{p^2 + \chi^2},$$

$$\frac{\partial z}{\partial \tau} = -\partial_r \left( \lambda_z (L_v - r) z - d_z \partial_r z \right) + \epsilon \left( \frac{zp^2}{p^2 + \chi^2} - \sigma z \right),$$

(4.42)

where $r \in [0, L_v]$ and the $\lambda_i \geq 0$ represent the Bernoulli parameters (in the sense that $\lambda_i \propto |1 - R_i|$) of the organisms. Also all the parameters are as in the previous section apart from the $\lambda_i$, $i = p, z$, which are assumed to be small. The boundary conditions for the PDE in (4.42) were assumed to be no-flux. For various combinations of
values of these $\lambda_i$ we see much of the qualitative behaviour of the full model. For most values of $\lambda_p$ and $\lambda_z$ we see that the system begins to show accumulation (if one of the $\lambda_i > 0$). Depending on the size of $\lambda_i$ we get either a steady sub-excitation distribution or we see the excitation wave begin to propagate and then settle to steady distribution but with the vortex centre in a permanently excited state. Figure (4.8) shows the evolution of the spatial distribution of phytoplankton for $\lambda_p = \lambda_z = 0.1$. We see how the excitation wave begins to propagate then recedes, settling into a steady yet excited (at the core) distribution.

In figure (4.9) we see that the reduced vortex model also admits an oscillating bloom solution. As in the full system, this only occurs for neutrally buoyant phytoplankton, $\lambda_p = 0$, and slightly dense zooplankton. We found that too large a value of
$\lambda_z$ led to the formation of a steady, excited distribution and too small led to a quasi-homogeneous distribution of phytoplankton. We also add that the size of the spatial domain was also a factor. If the size of the eddy was too small we saw no oscillation (due to the no-flux boundary condition). The PDE in (4.42) was integrated using the operator splitting [89] method (Lax-Wendroff advection step, weighted Crank-Nicholson for diffusion and second order Runge-Kutta for the reaction dynamics) and a review of the numerical scheme can be seen in the Appendix.

![Diagram](image)

Figure 4.8: Spatiotemporal evolution of the phytoplankton population for the suprathreshold vortex model, $\lambda_p = \lambda_z = 0.1$.

In figure (4.10), we plot the curve separating $(\lambda_p, \lambda_z)$ space into areas where the inertial/advective separation is too weak (region I) to cause excitation and areas
Figure 4.9: Oscillatory bloom for the reduced vortex model. Here, $\lambda_p = 0$ and $\lambda_z = 0.075$.

where the separation is sufficiently strong enough (region II). As can be seen clearly, this bifurcation curve of inertial induced excitation is symmetric and linear, due to the simple linear nature of the advection terms in the reduced model in (4.42). It does however demonstrate the observed requirement of sufficient separation strength before any excitation is seen. The curve was obtained by first fixing one of the $\lambda_i$ and increasing the other from zero. If at any point in the computational grid the phytoplankton population exceeded the value needed to cause a bloom [107], this was then a point on the curve in figure (4.10) separating the two regions in $(\lambda_p, \lambda_z)$ space.
Figure 4.10: Areas in $(\lambda_p, \lambda_z)$ space in which we will/won’t (II/I) see excitation due to inertial separation.

The mechanism for these oscillations is analogous to the full system. A lack of zooplankton at the origin (vortex core) leads to a rapid increase in phytoplankton which seeds an excitation that propagates outwards due to diffusion. However, the pulse is diluted due to the form of the advection term and dies out, thus returning to the beginning of the cycle. One point we make is that there seemed to be some dependence on the radius of the eddy as, at least for smaller eddies $L_v = 12$ (computational units only), the oscillatory bloom disappeared. Our next area of interest is how these processes affect the aggregation, segregation and predator-prey contact rates of the phytoplankton and the zooplankton.
4.4 Aggregation, segregation & contact rates

In this section, we investigate the effects of bloom formation on the statistical properties of the organisms' spatial distributions and how it could affect the grazing rate of zooplankton. We use four different statistics which we will now introduce.

To measure the magnitude of aggregation of each species we have the following two possibilities. The first we call $\Pi_1$ and is defined as

$$\Pi_i(t) = \frac{\langle C_i(r, t) \rangle^2}{\langle C_i^2(r, t) \rangle}$$

where $\langle \cdot \rangle$ denotes a spatial average and $C_i$ is the population density of species $i$. If the distribution is homogeneous then $\Pi_i = 1$, while if all the organisms are all at one point then $\Pi_i = 1/N^2$ (where $N^2$ is the number of mesh points in the computational grid).

To measure the degree of segregation of predator and prey in the system we use a population correlation function, $C_{ij}$, defined as follows

$$C_{ij} = \frac{\langle C_i(r, t) C_j(r + r', t) \rangle}{\langle C_i(r, t) \rangle \langle C_j(r + r', t) \rangle} \quad i \neq j, \ r' = 0,$$

and the indices denote the different species in the system. If species $i$ and $j$ have identical spatial distributions then $C_{ij} = 1$ and if they are located in completely separate areas of the spatial domain then $C_{ij} = 0$.

The final statistic we wish to use is a variation on the idea of the predator prey contact rate. It uses the difference in the relative fluxes of the phytoplankton and zooplankton which gives a measure of the rate at which the two organisms “meet”. In one spatial dimension, the relative flux of phytoplankton past a zooplankter, at
a given spatial location, is
\[
\Gamma_p(t) = \left| \frac{J_z}{z} - \frac{J_p}{p} \right| p,
\]
(4.45)
\[
= \frac{|p J_z - z J_p|}{z},
\]
as both \(p\) and \(z\) are non-negative and assumed non-zero. To now obtain an average contact rate (per unit time only) we compute
\[
E[\Gamma_p(t)] = \frac{\int_L \Gamma_p(t) z dx}{\int_L z dx \int_L p dx},
\]
(4.46)
where \(L\) is the length of the domain. Such an contact rate should be symmetric, in the sense that it should be invariant when we swap \(z\) for \(p\) and vice versa. We can see this by first noting that the relative flux of zooplankton past a phytoplankter, at some point in the domain, is given by
\[
\Gamma_z(t) = \frac{|p J_z - z J_p|}{p},
\]
(4.47)
analogously to (4.45). Similarly the mean contact rate is given by
\[
E[\Gamma_z(t)] = \frac{\int_L \Gamma_z(t) p dx}{\int_L z dx \int_L p dx},
\]
(4.48)
and so we find that
\[
E[\Gamma_z(t)] = E[\Gamma_p(t)] = \bar{\Gamma}(t) = \frac{\int_L |p J_z - z J_p| dx}{\int_L z dx \int_L p dx},
\]
(4.49)
as required.

Generalising this idea to 2-d we have an analogous expression for the averaged contact rate, \(\bar{\Gamma}(t)\),
\[
\bar{\Gamma}(t) = 2\pi \int_D \frac{\| p J_z - z J_p \| dA}{\int_D z dA \int_D p dA},
\]
(4.50)
where $\mathcal{R}$ is the zooplankter's perceptive radius and $D$ is the spatial domain being considered. From [31], we used $\mathcal{R} = 6$ cm in line with the idea of the perceptive radius being around 2 to 4 body lengths.

In figure (4.11) we observe how the statistical measures discussed vary over the period of one bloom oscillation.

![Graph](image)

Figure 4.11: Complete view of one period of the oscillatory bloom.

With this figure, and with the help of the corresponding snapshots in figure (4.6), we can describe in detail the behaviour of the planktonic populations and their spatial distributions. From (a) to (b) in figure (4.11), the phytoplankton abundance slowly increases but mostly spreads (via diffusion) inside the vortices. During this period $\Pi_p \approx 1$ since there is no localisation, only diffusion inside the
vortices. During this period, zooplankton numbers decrease due to the lack of prey, especially inside the eddies because of the inertial drift trying to exclude them from these areas.

The phytoplankton bloom begins at point (b). From (b) to (c), the lack of zooplankton allows the phytoplankton to increase, although it is always confined to the vortices. Hence, $\Pi_p$ decreases to a low value, a signature of localisation. At the same time, the zooplankton starts to grow again due to the presence of larger amounts of phytoplankton. This is also reflected in the sudden growth of the contact rate, $\Gamma(t)$, just before the phytoplankton reaches its maximum just after (c).

From (c) to (d), the larger phytoplankton population aids the regeneration of the zooplankton, especially around the straining regions of the flow around the eddies, see figure(4.6), which is where the phytoplankton had accumulated previously. For this reason, $\Pi_z$ decreases, a signal of zooplankton localisation. The zooplankton continues to grow, even while the phytoplankton abundance decreases (due to the fast/slow nature of the phytoplankton/zooplankton dynamics). It is observed how, at this stage, zooplankton penetrates (via diffusion) inside the eddies, see figure(4.6), counteracting the inertial drift. After point (d), in order to complete the oscillation, the zooplankton population begins to decay due to the lack of prey, and the phytoplankton travels below the equilibrium concentration before finally returning to this state at (a). During this phase, both localisation functions increase and the contact rate decreases due to the reduction of the number of both species.
4.5 Conclusions

In this chapter, we have demonstrated how inertial separation of phytoplankton and zooplankton could be a driver/contributing driver (coupled with excitable reaction dynamics from [107]) to generating phytoplankton blooms. Undoubtedly, other major drivers for the observed blooms could be factors such as temperature, irradiance and nutrient availability (see the excellent review paper by Legendre [57]). However, we suggest that the inertial characteristics of both phytoplankton and zooplankton can certainly aid in the occurrence of and in some regimes cause the bloom when we incorporate turbulent flow.

In section 4.1.4, we developed an analytical expression for the change in the energy spectrum of an inertial particle immersed in steady, homogeneous turbulence. Kraichnan's [53] work on 2D turbulence concluded that the energy in the system was contained in the large scale eddies. It is known [97, 91] that light particles aggregate inside these eddies while heavier ones prefer to be outside the eddies in the straining regions. Intuitively, we might expect a higher/lower value of $u_0^2$ for lighter/heavier particles as lighter particles are forced into the energy containing eddies, while the heavier particles are excluded from these regions. This was what was indeed suggested by the analysis.

In section 4.2, we brought together the excitable reaction dynamics and the turbulent flow, with inertial corrections. Not only did we observe that inertial factors can cause patchiness of the plankton but can also help cause a bloom, for realistic inertial regimes. The resulting oscillatory bloom had a period of approximately one
year. This situation differs from other related work on advection enhanced blooms [71, 16] as it requires no initial perturbation and is self-sustaining.

Using a simplified vortex model in one spatial dimension, we back up our findings that the emergence of a bloom is dependent on the strength of inertial separation/advection. In fact, there appears to be some critical value of the advection/inertial parameter which separates an excited regime from a quiescent one (see figure (4.10)) as well as the oscillatory bloom, supporting our explanation of what we observed in our numerical investigations of the full system.

In the final section we investigated how bloom formation affects the statistical properties of the spatial distribution of the organisms and the contact rates. In all but the oscillatory case we saw that the aggregation statistics $\Pi_i$ would go through some transient phase and then reach an equilibrium value. This value would either be excited ($\Pi_i < 1$) corresponding to a heterogeneous distribution or unexcited ($\Pi_i \approx 1$) corresponding to a near homogeneous distribution. This is not surprising considering the dependency of excitation on the strength of separation/advection.

When considering contact rates we developed a type of continuum contact rate based on the relative flux of the two species. Furthermore, the separation of the two organisms can be either good or bad for the zooplankton population. For the extreme inertial case seen in figure (4.5), the contact rate decreased, as we would expect because predator and prey reside in different areas. However, for the oscillatory case, the bloom can propagate without restraint which means more phytoplankton move into zooplankton populated areas thus increasing the contact rate. In particular, it is clear that the spatially averaged grazing rate varies significantly in response to
physical and population dynamics cues and it is unlikely that the averaged dynamics can be properly represented by a classical system of coupled ODE's.

For phytoplankton populations subject to excitable dynamics and turbulent advection with inertial effects, the mean contact rate does depend on the extent of the patchiness in contrast to other approaches, such as the probabilistically derived results of Pitchford & Brindley [86].

What we have demonstrated is that, for zooplankton at the larger end of the size spectrum (up to 1 cm), inertial effects can help initiate a phytoplankton bloom. Realistically, the mechanism behind the observed blooms is likely to contain contributions from a variety of different sources. We hypothesise that inertia is unlikely to be the sole contributor but is almost certainly a contributor to the bloom formation process as a whole.
Plankton Patch Dynamics: The coupled oscillator approach

The observation of the patchiness of oceanic plankton populations is a well documented phenomena [33]. Many driving mechanisms for patchiness have been suggested, from large scale turbulent advection [2] to small scale individual responses such as predator avoidance and buoyancy [31]. Regardless of the mechanism, more often than not these "patches" of plankton are not independent as many forms of coupling can exist between nearby patches (for instance diffusive coupling or the effects of higher predation). In the first part of this chapter, we would like to demonstrate, when considering the plankton to inhabit a patchy environment, that spatiotemporally varying dynamics can arise from a number of different sources. It has been observed that time series of larval plaice populations around the British isles display high levels of synchronicity [32], even when the habitats are far apart and may or may not be linked via the population dynamics of the patches. In the second section, we show how external forcing can force the synchronisation of patches, even when they aren't explicitly coupled and exert no effect on the forcing.
5.1 Bidirectionally Coupled Patches

In the following section, we consider the case where the patches are coupled together explicitly. A simple starting point would be a reaction-diffusion equation such as:

\[ \partial_t s = G(s) + D \nabla^2 s, \]  

(5.1)

where the vector \( s \) represents the interacting species, the reaction dynamics are given by \( G \) and \( D \) is the diffusion (which, depending on the spatial scale, may denote turbulent advection or swimming behaviour). In [66] a “patchy” version of (5.1) was considered whereby each patch is diffusively coupled but has spatial variations in the reaction system \( G \). Here, we propose a spatially (1D) discretised paradigm for patch dynamics. Coupled lattice maps generally have the form

\[ s_i^{n+1} = (1 - 2\epsilon)G(s_i^n) + \epsilon(G(s_{i+1}^n) + G(s_{i-1}^n)) \]

(5.2)

where \( i \) denotes the lattice point and \( n \) is the discrete time variable. However, plankton populations are best represented as continuous time variables due to the effect of overlapping generations [46], so it is preferable to use the following model

\[ \dot{s}_i = G(s_i) + \epsilon_{i+1}(s_{i+1} - s_i) + \epsilon_{i-1}(s_{i-1} - s_i), \]

(5.3)

whereby we have continuous reaction dynamics. For the situation we consider here we employ the following “no-flux” type boundary conditions, \( \epsilon_0 = \epsilon_{n+1} = 0 \). Here, we shall mostly consider the symmetric coupling case \( \epsilon_{i-1} = \epsilon_{i+1} = \epsilon, \forall i \) but do assess the non-symmetric case.

To represent the reaction dynamics, \( G \), we use a well known three component nutrient-phytoplankton-zooplankton (NPZ) model, so that \( s = (N, P, Z) \). The NPZ
model that we will use was constructed in [102] but investigated in detail in [26] and [27]. It takes the form

\[
\frac{dN}{dt} = -\frac{Na}{(e + N)(b + cP)} P + \gamma P + \frac{\lambda P^2}{\mu^2 + P^2} Z + \gamma dZ + k (N_0 - N),
\]

\[
\frac{dP}{dt} = \frac{Na}{(e + N)(b + cP)} P - \gamma P - \frac{\lambda P^2}{\mu^2 + P^2} Z - (s + k) P, \tag{5.4}
\]

\[
\frac{dZ}{dt} = \frac{\alpha \lambda P^2}{\mu^2 + P^2} Z - dZ.
\]

Here \(N, P\) and \(Z\) are the biomass of the nutrients, phytoplankton and zooplankton respectively. \(a\) is a measure of the maximum growth rate of \(P\), \(b\) represents light attenuation by water and \(c\) specific light attenuation by the phytoplankton themselves. The higher predation is denoted by \(d\) and \(e\) is the half-saturation constant due to the uptake of nutrient by the phytoplankton. Phytoplankton are lost from the system by two mechanisms, sinking of \(P\) given by \(s\) and the cross-thermocline exchange rate (with deep water devoid of phytoplankton) denoted by \(k\). \(N_0\) represents the nutrient level below the mixed layer and \(r\) the phytoplankton respiration rate. Here, \(\alpha\) and \(\beta\) describe zooplankton growth efficiency and excretion. Finally, \(\gamma, \lambda\) and \(\mu\) denote the rates of recycled higher predation, zooplankton grazing and the zooplankton grazing half-saturation coefficient respectively. See [102] for more details. Typical parameter values and units of the above quantities are presented in Table 5.1. The nature of the higher predatory response is a somewhat contentious subject. The model as above employs a linear functional response, but it has been suggested that a quadratic or Holling type III form may be more accurate. We are more concerned with the collective dynamics and use this model purely for the
sake of example. The dynamics of the uncoupled system are well documented [27] from equilibria to stable limit cycles to chaos under variation of the closure (higher predation) rate \( d \). Unless explicitly stated we shall consider cases where the individual patch dynamics are chaotic as these cases are the most interesting in terms of possible routes to desynchronised patch dynamics. In the next two sections we show that, in our spatially discrete system, the transition to non-synchronous collective dynamics can correspond to one of two possible bifurcations and later, we look at what we can expect to see when we have variations in the individual dynamics.

5.1.1 Patch Synchronisation

As mentioned, our main interest is to see under what conditions the individual patch dynamics cease to be synchronous giving rise to spatial (as well as temporal) irregularity throughout our patch lattice. Much work in recent years has considered the general behaviour and synchronisation of coupled oscillators [6, 34, 77, 82]. By synchronisation, we mean that the asymptotic dynamics of all the individual patches are identical and are constrained to some manifold, which we henceforth refer to as \( M_S \), defined by

\[
M_S = \{ N_i^t, P_i^t, Z_i^t \mid N_i^t = N_i^t, P_i^t = P_i^t, Z_i^t = Z_i^t \} \quad \forall i \neq j, \forall t.
\]

The boundary of synchronous and non-synchronous behaviour corresponds to a symmetry breaking bifurcation such that the synchronous attractor \( A \in M_S \) loses stability transverse to \( M_S \). This “blowout” bifurcation can be calculated using a variant on the concept of Liapunov exponents [25] (see Appendix for a more thorough re-
Table 5.1: Default parameter values for the NPZ model in (5.4).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Symbol</th>
<th>Default Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phytoplankton growth rate</td>
<td>$a$</td>
<td>$0.2 \text{ (m/day)}^{-1}$</td>
</tr>
<tr>
<td>Light attenuation by water</td>
<td>$b$</td>
<td>$0.2 \text{ m}^{-1}$</td>
</tr>
<tr>
<td>Light attenuation by phytoplankton</td>
<td>$c$</td>
<td>$0.4 \text{ m}^2(\text{g C})^{-1}$</td>
</tr>
<tr>
<td>Higher predation of zooplankton</td>
<td>$d$</td>
<td>$0.142 \text{ g C m}^{-3}\text{day}^{-1}$</td>
</tr>
<tr>
<td>Nutrient half-saturation constant</td>
<td>$e$</td>
<td>$0.03 \text{ g C m}^{-3}$</td>
</tr>
<tr>
<td>Cross-thermocline exchange rate</td>
<td>$k$</td>
<td>$0.05 \text{ day}^{-1}$</td>
</tr>
<tr>
<td>Phytoplankton respiration</td>
<td>$r$</td>
<td>$0.15 \text{ day}^{-1}$</td>
</tr>
<tr>
<td>Phytoplankton sinking</td>
<td>$s$</td>
<td>$0.04 \text{ day}^{-1}$</td>
</tr>
<tr>
<td>Lower mixed level nutrient concentration</td>
<td>$N_0$</td>
<td>$1 \text{ g C m}^{-3}$</td>
</tr>
<tr>
<td>Zooplankton growth efficiency</td>
<td>$\alpha$</td>
<td>$0.25$</td>
</tr>
<tr>
<td>Zooplankton excretion fraction</td>
<td>$\beta$</td>
<td>$0.33$</td>
</tr>
<tr>
<td>Regeneration of zooplankton excretion</td>
<td>$\gamma$</td>
<td>$0.5$</td>
</tr>
<tr>
<td>Zooplankton grazing rate</td>
<td>$\lambda$</td>
<td>$0.6 \text{ day}^{-1}$</td>
</tr>
<tr>
<td>Zooplankton half-saturation constant</td>
<td>$\mu$</td>
<td>$0.035 \text{ g C m}^{-3}$</td>
</tr>
<tr>
<td>Patch to patch flux</td>
<td>$\epsilon_i$</td>
<td>Bifurcation parameter</td>
</tr>
</tbody>
</table>
view of the concepts used in the following section). The Liapunov exponent of the base point \( x \in A \) in the direction \( u \in T_xM_S \) is given by,
\[
\lambda(x, u) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \log \| DG^t(u) \| \, dt,
\]
where \( DG^t(\cdot) \) represents the Jacobian of \( G(\cdot) \) at time \( t \) and \( T_xM_S \) is the tangent space of \( M_S \) at the point \( x \). The normal Liapunov exponent, \( \lambda_\perp(x, v) \), is defined to be,
\[
\lambda_\perp(x, v) = \lim_{T \to \infty} \frac{1}{T} \int_0^T \log \| \Pi_{(T_xM_S)\perp} \circ DG^t(v) \| \, dt,
\]
and \( \Pi_V \) denotes an orthogonal projection onto the vector space \( V \). If we assume that \( A \) supports some natural, ergodic invariant measure [25], \( \mu \), then the time averages defined in (5.5) and (5.6) will be equal to the space averages
\[
\lambda = \int_A \log \| DG(u) \| \, d\mu(x),
\]
and
\[
\lambda_\perp = \int_A \log \| \Pi_{(T_xM_S)\perp} \circ DG(v) \| \, d\mu(x),
\]
and consequently converge to finite set of constant values. Firstly, we should define what we mean when we say attractor and we work with the following two definitions

1. \( A \) is an asymptotically stable attractor if it is stable in the sense of Liapunov and the basin of attraction of \( A, \beta(A) \), contains a neighbourhood of \( A \).

2. \( A \) is a Milnor attractor if \( \beta(A) \) has non-zero Lebesgue measure and there exists no proper compact subset, \( A' \subset A \) such that \( \beta(A')=\beta(A) \) up to a set of zero measure (\( A \) is indecomposable).
where the basin of attraction, $\beta(A)$, is just the set of points whose orbits' forward time limit is the invariant set $A$. By Liapunov stable we mean that, for every neighbourhood $U$ of $A$, we can find another neighbourhood, $V$, such that $G^t(V) \subset U$, $\forall t \in \mathbb{R}$. This is the fairly standard topological definition but we will see that we have instances where $\beta(A)$ contains no neighbourhood of $A$ but is still large in a measure-theoretic sense. This is why we include Milnor's [67] weaker definition.

These normal exponents measure the contraction/expansion of perturbation vectors transverse to $M_S$. If now we define $\lambda_{\perp}^{\text{max}}$ to be the largest normal exponent then the local stability of the synchronous attractor $A$ is determined by the sign of $\lambda_{\perp}^{\text{max}}$. If it is negative then small perturbations will die out exponentially but if it is positive then disturbances initially grow $\sim e^{\lambda_{\perp}^{\text{max}} t}$ until the non-linear terms kick in (and $A$, while still an attractor in $M_S$, has a basin of attraction with zero Lebesgue measure in the full phase space). Parameters such as the diffusive coupling $\epsilon$ are called normal parameters as they don't affect the dynamics normal to $M_S$. This asserts the continuity of the $\lambda_{\perp}$ w.r.t. normal parameters allowing the defining of a clear bifurcation point. For normal parameters categorised the scenario into two types of behaviour:

1. **Hysteretic blowout bifurcation** For some specific value, $\nu_c$, of some arbitrary, normal coupling parameter $\nu$ the chaotic set $A$ will be marginally stable. Beyond the critical value of $\nu$ the attractor is transversely unstable but trajectories with initial conditions close to $M_S$ will experience a chaotic transient orbit which closely resembles the invariant set $A$ but will eventually move off
to the other attractor/attractors away from $M_S$. This is also referred to as the *sub-critical* case as the attractors created after the blowout bifurcation are contained in $M_S$.

2. **Non-hysteretic blowout bifurcation**

After we have passed through $\nu_c$ there will be a phenomenon called *on-off intermittency*. This is where the orbits close to $M_S$ will spend a long time very near $M_S$ but every once in a while they will burst away from the attractor in the manifold. The time spent in these bursts is typically short compared to the time spent near the manifold. The non-hysteretic case is called *supercritical* as the family of attractors emerging from a blowout exist in the whole of phase space but are said to be *stuck* [8] to the invariant manifold, $M_S$.

In figure (5.1) we plot the maximal normal Liapunov exponent, $\lambda_{\perp}^\text{max}$, which we calculated for the two patch, symmetric coupling case, $\epsilon_1 = \epsilon_2 = \epsilon$. We see that the synchronous state loses transverse stability below $\epsilon_c = 0.002$ (3 d.p.) as $\lambda_{\perp}^\text{max}$ passes through zero. In figure (5.2) we show the attractors in $(N_1, N_2)$ space for $\epsilon_1 = \epsilon_2 = 0.003$ (just above $\epsilon_c$) and for $\epsilon_1 = \epsilon_2 = 0.001$ (just below $\epsilon_c$) to illustrate the difference in the system before and after the blowout bifurcation (this is an example of on-off intermittency).

This shows that if the coupling is sufficiently low, any small differences in the initial conditions of each patch will initially grow exponentially (until the non-linear terms kick in). The calculations were made for the simplest case of two interacting patches but can we extend the idea to arbitrary numbers of coupled patches?
Figure 5.1: A plot of the maximal normal Lyapunov exponent, $\lambda^{\text{max}}$, versus the patch coupling, $\epsilon$. We see the first blowout bifurcation occurs at around $\epsilon_c=0.002$. Just below this value of the coupling, the synchronous state $A$ will cease to be an attractor for the whole system.

5.1.2 Dynamics of coupled oscillator systems

We consider a fairly specific form for our plankton patch lattice system in (5.3) as this type of diffusive coupling suits the system we are considering best but what of the dynamics of a more generic coupled oscillator system. Firstly, let us move to a more more general framework than that in (5.3), so

$$\dot{S} = G(S) + \mathcal{E}_L \otimes H(S), \quad (5.7)$$

where $S = (s_1, s_2, \ldots, s_n)^T$ represents the species present (the $s_i$ are $m$-dimensional vectors and $i = 1, \ldots, n$ denotes the lattice point). The reaction dynamics are governed by the function $G(S) = (G(s_1), G(s_2), \ldots, G(s_n))^T$. The matrix $\mathcal{E}_L$ denotes
Figure 5.2: *Attractors in the $(N_1, N_2)$ plane before, $\epsilon_1 = \epsilon_2 = 0.003$, (left) and after, $\epsilon_1 = \epsilon_2 = 0.001$, (right) the blowout bifurcation. We can clearly see that after the blowout bifurcation, the symmetry of the system is broken and $N_1$ and $N_2$ evolve in a non-synchronous manner.*

the lattice coupling matrix and $H$ is an as yet arbitrary (assumed $C^1$) function encompassing both the nature of the coupling (linear/non-linear) and which of the patch species will be coupled to similar/different species in the other patches in the lattice. Working under the assumption that the synchronisation manifold $(s_1 = s_2 = \cdots = s_n)$ is an invariant manifold constrains the sum of all the elements in a particular row of $E_L$ to sum to zero, so $\sum_j (E_L)_{ij} = 0$.

Systems such as those in (5.7) have been seen to exhibit globally synchronous behaviour (Heagy *et al.* [38]; Pecora and Carroll [79]; Pecora [80]) for various types of local and global coupling matrices and cluster synchronisation (Belykh *et al.* [14]); for certain non-local coupling regimes oscillators synchronise together in groups but there is no synchronisation between the groups. In this work we consider only
coupling matrices which allow only the possibility of a globally synchronous state, both for relevance and breadth of the applications.

5.1.3 General stability analysis of coupled oscillator arrays

Given the global synchronous solution \( s^*(t) = s_1 = s_2 = \cdots = s_n \), the variational equation governing small perturbations \( \xi_i = s_i - s^*(t) \) of the \( i^{th} \) lattice point is given by

\[
\dot{\xi} = (I_n \otimes DG + \mathcal{E}_L \otimes DH) \xi,
\]

where \( \xi = (\xi_1, \ldots, \xi_n)^T \), \( I_n \) is the \( n \times n \) identity matrix, \( \otimes \) is the direct (tensor) product. The tensor product of two square matrices \( A (n \times n) \) and \( B (m \times m) \) is an \( mn \times mn \) block matrix obtained by replacing \( A_{ij} \) by \( A_{ij}B \). Here, \( DG \) and \( DH \) are the Jacobian matrices of \( G \) and \( H \), respectively, evaluated at the synchronous solution.

As a commonly occurring example, the case of nearest-neighbour symmetric diffusive coupling (with no flux boundary conditions) was one of several considered in Pecora and Carroll [79] and Pecora [80]. For this scenario, the coupling function \( H \) is just the \( m \)-dimensional identity matrix (so only similar species variables are coupled in the lattice) and \( \mathcal{E}_L = \epsilon \mathcal{E} \), where \( \epsilon \) is a positive scalar and the coupling matrix is defined by

\[
\mathcal{E} = \begin{pmatrix}
-1 & 1 & 0 & \cdots & 0 \\
1 & -2 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 1 & -1
\end{pmatrix}. \tag{5.9}
\]
The lattice Jacobian, the bracket in (5.8), would then take the following form [80],

\[
\begin{pmatrix}
DG - \epsilon I_m & \epsilon I_m & 0 & \cdots & 0 \\
\epsilon I_m & DG - 2\epsilon I_m & \epsilon I_m & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \epsilon I_m & DG - \epsilon I_m \\
\end{pmatrix}
\]  

(5.10)

To isolate variations transverse to the synchronisation manifold from those inside the synchronisation manifold, discrete Fourier transforms were employed so as to diagonalise the coupling matrix \(\mathcal{E}_L\) and leave the matrix in (5.10) in block diagonal form. This change of coordinates yields the following Jacobian matrix,

\[
\begin{pmatrix}
DG & 0 & 0 & \cdots & 0 \\
0 & DG + \epsilon I_m \gamma_1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & 0 & DG + \epsilon I_m \gamma_{n-1} \\
\end{pmatrix}
\]

(5.11)

where \(\gamma_k = -4\sin^2(\pi k/2n), \text{ for } k = 1, \ldots, n - 1,\) is an eigenvalue of the matrix in (5.9). So, the stability of the spatial nodes in the lattice system are governed by the following individual block variational equations,

\[
\dot{\xi}_k = (DG + \epsilon I_m \gamma_k) \xi_k
\]

(5.12)

with each block corresponding to a discrete spatial Fourier mode. As can be clearly seen from (5.11), the mode \(k = 0\) corresponds to the variational equation for motion inside the synchronisation manifold and the Liapunov exponents of this mode correspond exactly to those of a single oscillator’s dynamics. Orthogonality of the
Fourier modes means that all the other non-zero modes govern the dynamics of small perturbations transverse to the synchronisation manifold. Hence, if all the Liapunov exponents of these modes are negative, the global synchronised solution will be stable.

The dependence of the eigenvalues $\gamma_k$ on the number of oscillators, $n$, means that the scalar coupling values for which we see stable synchronisation will also depend on the number of oscillators present. Indeed, in [79, 80], not only was this effect analytically predicted and numerically observed but that, for non-identity forms of the coupling function $H$ that there could even be a limit on the number of oscillators for which a stable synchronous solution could be realised. The interested reader is referred to the work in [38, 79, 80] for a thorough analytical and numerical review of the dynamics of this example and other types of coupled oscillator systems with constant coupling matrices.

From a biological/ecological viewpoint, the idea that the lattice coupling matrix is constant throughout the lattice seems somewhat simplistic and hopeful. There are likely to be local variations in the coupling rates between lattice points, variations in migration strategies/mixing rates for example. Given the structure we consider in (5.3), this would lead us to the following, non-symmetric form for the coupling
matrix,

\[
E_L = \begin{pmatrix}
-\epsilon_2 & \epsilon_2 & 0 & \cdots & 0 \\
\epsilon_1 & -(\epsilon_1 + \epsilon_3) & \epsilon_3 & \cdots & 0 \\
0 & \epsilon_2 & -(\epsilon_2 + \epsilon_4) & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & \epsilon_{n-1} & -\epsilon_{n-1}
\end{pmatrix},
\]  \hspace{1cm} (5.13)

and again \( \sum_j (E_l)_{ij} = 0 \), so the synchronisation manifold is invariant under the action of the lattice flow (5.7).

In the example system and references given previously, the constant, symmetric nature of the lattice coupling matrix allows for the use of discrete Fourier transforms to block diagonalise the coupling matrix which then allows one to analyse the stability of the resulting spatial Fourier modes inside the synchronisation manifold, \( k = 0 \), and transverse to the synchronisation manifold, \( k \neq 0 \).

The case where we have a coupling matrix such as the one defined in (5.13) does not appear to have been analysed in a similar manner in the literature, presumably because the non-constant, non-symmetric nature of the coupling matrix does not permit the spatially modal decomposition of the lattice Jacobian. In the previous system, the scalar coupling \( \epsilon \) needed to see stable synchronisation depended only on \( DG \), the form of the lattice coupling matrix, the coupling function's Jacobian matrix \( DH \) and the number of coupled oscillators, \( n \). However, with the added possibility of local coupling variations, it would seem likely that any loss of stability of the synchronous state could also be dependent on a particular region of the lattice as well.
To see what possible effects such local coupling variations might have, let us consider the variational equation for the vector variable \( \zeta = (\zeta_1, \ldots, \zeta_{n-1})^T \), where \( \zeta_i = s_i - s_{i+1} \), with the Jacobian matrix \( DG \) evaluated at the synchronous solution \( (\zeta = 0) \),

\[
\dot{\zeta} = (I_{n-1} \otimes DG + E_L^1 \otimes I_n) \zeta,
\]

and the \((n-1) \times (n-1)\) matrix \( E_L^1 \) is given by

\[
E_L^1 = \begin{pmatrix}
-(\epsilon_1 + \epsilon_2) & \epsilon_3 & 0 & \cdots & 0 \\
\epsilon_1 & -(\epsilon_2 + \epsilon_3) & \epsilon_4 & \cdots & 0 \\
0 & \epsilon_2 & -(\epsilon_3 + \epsilon_4) & \cdots & 0 \\
& & & \ddots & \ddots \\
0 & \cdots & 0 & \epsilon_{n-2} & -(\epsilon_{n-1} + \epsilon_n)
\end{pmatrix}. \tag{5.15}
\]

The equation in (5.14) is the variational equation for small perturbations transverse to the synchronisation manifold. From the structure of \( E_L^1 \), we can see that, barring the "boundary" lattice points \( i = 1, n - 1 \) and \( n \), the coupling term \( \epsilon_i \) directly affects only the dynamics of the variables \( \zeta_{i-1}, \zeta_i \) and \( \zeta_{i+1} \). Let us consider the following decomposition of the full lattice phase-space, \( S \),

\[
S = S_1 \oplus S_2 \oplus \cdots \oplus S_n, \tag{5.16}
\]

and \( s_i \in S_i, \forall i \). The variables \( \zeta_{i-1}, \zeta_i \) and \( \zeta_{i+1} \) govern the fate of small perturbations of the synchronisation manifold in the space \( S_i \), defined by

\[
S_i = \oplus_{j=i-1}^{i+2} S_j. \tag{5.17}
\]
The simplest scenario that one could envisage is where, to begin with, $\epsilon_i = \epsilon, \forall i$. We now assume that there is some critical value of the scalar coupling, $\epsilon = \epsilon_c$ (depending again on $DG, DH$, the lattice coupling matrix and $n$), below which the synchronous state is unstable. If we have the situation where $\epsilon > \epsilon_c$, but $|\epsilon - \epsilon_c| \ll 1$, then what happens to the system if we vary just one of the lattice point coupling parameters, $\epsilon_i$?

As was seen, varying only this $\epsilon_i$ affects transverse perturbations of the synchronisation manifold in the localised space $S_i$. If indeed $\epsilon > \epsilon_c$, but $|\epsilon - \epsilon_c| \ll 1$, there should be a threshold value of $\epsilon_i$ for which small perturbations to the synchronisation manifold in $S_i$ didn't die out and in fact grew, leading to the existence of one, positive normal Liapunov exponent. However, this seemingly locally originating blowout bifurcation must in fact manifest itself as a loss of stability of the globally synchronised state. We can analytically demonstrate why this must be the case but the analysis requires that we first assume that there is some value of $\epsilon_i > 0$ below which the maximal normal Liapunov exponent corresponding to the fate of small perturbations to the synchronisation manifold in $S_i$, $\lambda_i^{\text{max}}(i)$, is positive.

By the multiplicative ergodic theorem of Oseledec (1971), we can decompose the space orthogonal to the synchronisation manifold, $TM^\perp$, in the following manner. There exist linear sub-spaces $F_1 \supset \cdots \supset F_k$, such that $TM^\perp = F_1 \oplus \cdots \oplus F_k$ (where $k = m(n-1)$), and

$$\lambda_i^j = \lim_{T \to \infty} \frac{1}{T} \int_0^T \log \| \Pi_{TM^\perp} \circ DG^j(v) \| \, dt, \quad \forall v \in F_j \setminus F_{j+1}, \quad (5.18)$$

where $\lambda_1^j > \cdots > \lambda_k^j$, $\| \cdot \|$ is the Euclidean norm on $R^{m(n-1)}$ and the $\lambda_i^j$ are
the Liapunov exponents normal to $M_S$. Next, we define the linear time evolution operator, $\Lambda(t)$, for $\zeta(t)$, which satisfies the following set of equations,

\[
\begin{align*}
\zeta(t) &= \Lambda(t)\zeta(0), \\
\dot{\Lambda}(t) &= J^+\Lambda(t), \\
\Lambda(0) &= I_{m(n-1)},
\end{align*}
\]

(5.19)

where $J^+$ is the bracket on the r.h.s. of (5.14). The normal Liapunov exponents can now be defined (Eckmann & Ruelle [25], Appendix) to be the logarithms of the eigenvalues of the following limiting matrix,

\[
\lim_{t \to \infty} (\Lambda^T(t)\Lambda(t))^{\frac{1}{d_t}}.
\]

(5.20)

Let $W_1, ..., W_k$ be the eigenspaces of the eigenvalues $\alpha_1, ..., \alpha_k$ of the limiting matrix defined in (5.20). Now,

\[
\begin{align*}
F_k &= W_k \\
F_{k-1} &= F_k \oplus W_{k-1}, \\
& \vdots \\
F_1 &= F_2 \oplus W_1.
\end{align*}
\]

(5.21)

Let us now consider some generic transverse perturbation vector $w \in TM^1$, where

\[
w = w^1 + w^2 + \cdots + w^k
\]

(5.22)

and $w^j \in W_j$. Since the matrix $\Lambda^T(t)\Lambda(t)$ is symmetric, the eigenvectors are orthogonal. Hence,

\[
\|w\|^2 = \sum_{j=1}^{k} \|w^j\|^2.
\]

(5.23)
and furthermore, with \( w^T \Lambda(t) \Lambda(t) w = \parallel \Lambda(t) w \parallel^2 \), we see that

\[
\parallel \Lambda(t) w \parallel^2 = \sum_{j=1}^{k} \parallel \Lambda(t) w^j \parallel^2.
\]  

(5.24)

Using the fact that \( \lambda_j^1 = \log(\alpha_j) \),

\[
\parallel \Lambda(t) w \parallel^2 \approx e^{2\lambda_1^1 t} \parallel w^1 \parallel^2 + \cdots + e^{2\lambda_k^1 t} \parallel w^k \parallel^2.
\]  

(5.25)

and if we factor out the term in \( e^{2\lambda_1^1 t} \) we have that

\[
\parallel \Lambda(t) w \parallel^2 \approx e^{2\lambda_1^1 t} \left( \parallel w^1 \parallel^2 + e^{2(2\lambda_1^1 - \lambda_1^1) t} \parallel w^2 \parallel^2 + \cdots + e^{2(2\lambda_k^1 - \lambda_1^1) t} \parallel w^k \parallel^2 \right).
\]  

(5.26)

Now, as \( t \to \infty \), \( e^{2(2\lambda_1^1 - \lambda_1^1) t} \to 0 \) because \( \lambda_1^1 > \cdots > \lambda_k^1 \). This means that the long term behaviour of the perturbation \( w \) is dominated by the first term in (5.25). However, we have one positive normal exponent and by definition this must be \( \lambda_1^1 = \lambda_{1\text{max}}(i) \). Hence, any generic perturbation will be exponentially expanded and we have a global blowout.

To illustrate this effect numerically, a lattice of eight diffusively coupled NPZ systems was considered. Numerical investigation showed that, for the nearest-neighbour diffusive coupling with a single scalar, \( \epsilon \), to represent coupling strength, the eight patch system exhibited globally synchronous behaviour for coupling strength above \( \epsilon \approx 0.0075 \). In line with the theoretical scenario discussed previously, we reduced the value of one of the coupling parameters (in this case \( \epsilon_4 \)) when the system is close to the global loss of transverse stability. In figure (5.3), we plot the temporal difference in the nutrient variables for adjacent patches, \( N_i - N_{i+1} \), for \( i = 2, 3, 4 \) and 5 so that we look at the dynamics transverse to the synchronisation manifold in the lattice points closest to the region where we have decreased the coupling. For
\( c_4 < 0.001 \), the globally synchronous state loses stability, giving rise to the dynamics seen in figure (5.3).

Figure 5.3: \( N_i(t) - N_{i+1}(t) \), for \( i = 2, 3, 4 \) and 5 after any transient behaviour has gone. The synchronised solution is unstable throughout the whole lattice.

As can be seen from figure (5.3), the magnitude of the bursts from synchronicity are greatest in the two regions exactly adjacent to \( i = 4 \), where we decreased the coupling. In fact, we can quantify this bursting by computing the following average,

\[
\langle \zeta_i \rangle = \lim_{T \to \infty} \frac{1}{T} \int_0^T \| \zeta_i(t) \| \, dt = \lim_{T \to \infty} \frac{1}{T} \int_0^T \| s_i(t) - s_{i+1}(t) \| \, dt, \tag{5.27}
\]

which (under the assumption that the attractor is ergodic) should converge (almost everywhere) to a constant value for each \( i \) by Birkhoff's ergodic theorem (Eckmann &
Table 5.2: Bursting measure defined in (5.27) for the case of a blowout from altering only one coupling parameter (row 1) and where all the coupling parameters are equal but below the synchronisation threshold (row 2).

<table>
<thead>
<tr>
<th>$i$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \zeta_i \rangle (\epsilon_i = 0.001)$</td>
<td>0.0045</td>
<td>0.012</td>
<td>0.013</td>
<td>0.0044</td>
</tr>
<tr>
<td>$\langle \zeta_i \rangle (\epsilon_i = 0.001, \forall i)$</td>
<td>0.054</td>
<td>0.055</td>
<td>0.054</td>
<td>0.056</td>
</tr>
</tbody>
</table>

Ruelle [25]). Non-zero values of the $\langle \zeta_i \rangle$ are indicative of non-synchronous dynamics while for synchronised systems, $\langle \zeta_i \rangle$ will converge to zero as $T \to \infty$. Table 5.2 shows the numerical results for the values of the quantity defined in (5.27) for the situation where we lower only one value of the coupling (first row and $\epsilon_4 = 0.001$) and when all the values of the coupling are the same but below the synchronisation threshold (second row and $\epsilon_i = 0.001, \forall i$).

As can be seen from table 5.2, for the first case we see that $\langle \zeta_i \rangle$ is different as we move away from the lattice point for which we decreased the coupling yet it displays a symmetrical decrease. This is in contrast with the second scenario where the corresponding lattice values of $\langle \zeta_i \rangle$ are almost identical. We point out that this similarity is not supported at the lattice boundary points for the no-flux coupling scenario as the coupling changes at these points but numerical investigations showed that, for periodic boundary conditions, the quantities $\langle \zeta_i \rangle$ were all converging to some value independent of $i$. This is because of the shift-invariant [38] nature of the coupling.
We mention that convergence of the bursting measures \( \langle \zeta_i \rangle \) was quite slow, the results given were for \( 10^6 \) iterations, these differed little from results obtained at 750,000 iterations but did differ somewhat from results obtained at 500,000 iterations. We hypothesise that the reason for this is that the time average in (5.27) must be long enough to smooth out the sporadic bursting followed by longer periods of near synchronous behaviour typical of such systems after a blowout bifurcation has occurred.

These numerical simulations give some weight to the idea that, if we allow for local coupling variations, a global blowout can arise from a more localised event. As seen in table 5.2, the asynchronous bursting is strongest around the region where the coupling parameter was decreased. While computing the point at which the maximal normal Liapunov exponent becomes positive gives us the parameters for which synchronisation becomes unstable, the simple ergodic average bursting quantity defined in (5.27) allows us to (numerically at least) investigate something of the local lattice dynamics after this blowout event, if only in terms of the severity of the asynchronous behaviour.

This situation obviously considered varying only one of the coupling parameters while keeping the rest constant and just above the synchronisation threshold. For more complicated scenarios where more than one parameter is varied, we can expect a more complicated interplay of local stabilising/destabilising influences. In general, when allowing for local coupling variations, one is almost certain to observe such locally originating blowout bifurcations, for a variety of coupling parameter combinations but we also hypothesise that, as was seen in the scalar coupling systems
studied in [38, 79, 80], these parameter values for which we see a blowout bifurcation will also alter with the type of lattice coupling matrix and number of oscillators present.

5.1.4 Further Pathways to Irregular Patch Dynamics

So we see that low enough coupling strength can allow the onset of spatiotemporally heterogeneous patch dynamics but this is not the only possible route in our patch lattice system. For one, we have not yet addressed the situation where there are differences in the underlying reaction dynamics of each patch and there is also the fact that we must consider what effect some low levels of system noise might have.

5.1.5 Riddled basins of attraction

Riddled basins of attraction were first investigated in [3]. A basin of attraction is said to be locally riddled if there exists \( \delta > 0 \) such that, for arbitrary \( x \in \beta(A) \) and any \( \varepsilon > 0 \), the \( \varepsilon \) ball \( B_\varepsilon(x) \) contains a positive measure set of points whose orbits exceed a distance \( \delta \) from \( A \). A basin of attraction is globally riddled if, for arbitrary \( x \in \beta(A) \), \( B_\varepsilon(x) \) intersects \( \beta(A) \) and the basin of some other attractor with positive measure. This means that any open set in \( \beta(A) \) will have a non-zero fraction that will either move a specified distance away from synchronicity (local riddling) or converge to another, non-synchronous attractor (global riddling).

If the basin of attraction of \( A, \beta(A) \), is either locally or globally riddled then the synchronous state will not be stable to low levels of noise. Riddled basins form when one of the saddle cycles, from the usual cascade to chaos, embedded in \( A \) becomes
a repelling cycle. This causes the creation of infinitely many repelling 'tongues' foliating off of $A$. A schematic representation of this is given in figure (5.4). For the case of local riddling, it is not too difficult to see that any amount of noise will, with probability one, push all orbits into one of these repelling regions. The orbit will then move some specified distance from synchronicity. This phenomenon has been called attractor bubbling [6]. For the case of global riddling we have a more extreme version of attractor bubbling as the orbit is continually moved into the basins of attraction of the non-synchronous/synchronous attractors. We can envisage a type of back and forth behaviour associated with this attraction/repulsion from both attractors. Whether the basin is locally or globally riddled, noise driven intermittency (attractor bubbling) will give us similar patch dynamics to the post blowout scenario described previously.

A final consequence of riddling is that our notion of an attracting set must "weaken" and we can only use the notion of $A$ being an attracting set in the sense of Milnor [67]. If $A$ is locally riddled then $A$ cannot be asymptotically stable. This is because we cannot have Liapunov stability as we may define some neighbourhood, $U_\delta$, of $A$ such that,

$$U_\delta = \bigcup_{x \in A} B_\delta(x), \quad (5.28)$$

If we now take $\delta$ to be the $\delta$ from the definition of local riddling, then any neighbourhood, $V \subset U_\delta$ (such that $A \subset V$), contains a subset of positive Lebesgue measure, $\tilde{V}$ (and some $\tau \in \mathbb{R}$), such that

$$\inf_{x \in \tilde{V}} \| G^\tau(x) - G^\tau(z) \| = \delta, \quad (5.29)$$
Figure 5.4: A schematic representation of the riddling bifurcation. Unstable tongues form at a cusp at P and every pre-image of P (only two are shown but there are infinitely many embedded in A). Orbits inside the tongues are repelled from A but outside they are attracted as \( \lambda_{\text{max}}^1 < 0 \).

which tells us that \( G^T(\tilde{V}) \cap U_3 = \emptyset \) and, hence, \( G^T(V) \not\subseteq U_3 \). Note that the condition is not \( \forall T > \tau \) because of the eventual re-injection \([6]\) of such orbits back towards the synchronisation manifold. Hence, A is no longer an asymptotically stable attractor as \( V \) was an arbitrary subset of \( U_3 \). The same argument applies to the case of global riddling but the second condition for asymptotic stability is also violated as \( \beta(A) \) contains no open sets, hence no neighbourhood of A.

The transition from locally riddled to globally riddled basins \([48, 59]\) is now a fairly well understood process. The first pathway from locally to globally riddled basins involves a boundary crisis, where the basin of attraction \( \beta(A) \) of the synchronous state \( A \) collides with the absorption area, \( A(A) \). The absorption area is
defined to be the smallest open set such that $A \subset A$ and $G^t(A) \subset A$, i.e. the smallest open set which contains $A$ and is forward-time invariant. Assuming that initial conditions outside $\beta(A)$ diverge to infinity, if the basin of $A$ is locally riddled then $A \subset \beta(A)$. For some critical value of the coupling, $c_r$ say, the boundary of the basin of attraction, $\bar{\beta}(A)$, collides with the boundary of the absorption area, $\bar{A}$. After this boundary crisis, the basin of attraction $\beta(A)$ is riddled with points that diverge to infinity. The second transition from locally to globally riddled basins requires the birth of a new attractor inside one of the repelling tongues seen in figure (5.4). For a more complete explanation and examples of the two bifurcations involved in the transition from locally to globally riddled basins of attraction, the reader is referred to the work in [48, 59]

To see if the basin of attraction of $A$ could be riddled, for certain values of $c_1$ and $c_2$, we can look at the eigenvalues of the saddle point embedded in $A$. When all the real parts of the eigenvalues of this fixed point become positive, then the basin of $A$ would be riddled. We found one real eigenvalue $\lambda_s = -0.048442$ which is independent of both $c_1$ and $c_2$ so changing these two parameters cannot cause the riddling of $\beta(A)$ by the saddle fixed point. Since the most unstable saddle point in $A$ cannot become a repeller it is very unlikely that the periodic orbits can either. Riddling has been observed in discrete population models with skew product structure, [3], and local riddling seems a generic phenomenon in coupled oscillators so it's absence here does not preclude it's presence in a different system. Intuitively, global riddling is not a likely phenomenon in such coupled systems due to the strong non-linear restraining mechanisms of most continuous population models
(thus removing the possibility of the boundary crisis mechanism described) and the uni-directional nature of the coupling (that is we consider the case where $\epsilon_i > 0$, $\forall i$).

5.1.6 Generalised synchronisation of patches

The final scenario we consider is when we have small fluctuations in the parameters of the reaction dynamics governing each patch. This is a more general case of (5.7):

$$\dot{s}_i = G_i(s_i) + \epsilon_i(s_{i+1} - s_i) + \epsilon_i(s_{i-1} - s_i), \quad (5.30)$$

where we allow for the fact that the dynamics governing each patch will not always be the same. This type of patch parameter variation has been examined in the continuum sense by [66] and [61]. Variation in the underlying parameters of plankton dynamics has been hypothesised to be a driver of phytoplankton blooms such as red tides and it is also feasible that there are some variations in parameters over large spatial scales. What does this imply for our coupled patch lattice model? In general, there are three classes of behaviour for systems with detuned parameters and each depends on the size of the parameter mismatch and the strength of the coupling. Exact synchronisation of the systems is no longer possible, but a less rigid definition of synchronisation has been recently developed, [94] and [1], in which the dynamics of individual patches are related by some continuous (possibly smooth) mapping and are thus said to be in generalised synchronisation (GS). Many terrestrial ecological systems have shown quite high levels of generalised synchronisation [17, 49, 52] and the well known Canadian Hare-Lynx data set exhibits spatially
coherent chaotic oscillatory behaviour (phase synchronisation in this case). In this section we demonstrate how to calculate regions of parameter space for which we going to see this generalised synchronisation of the patches.

Hence, the three possibilities for the dynamics of the patch lattice are:

1. For small parameter mismatch and sufficiently strong coupling, there exists a diffeomorphism mapping the dynamics of one patch to another (A is known as normally hyperbolic, [110]).

2. For coupling strength below a certain value, differentiability (and possibly other properties) is lost but there is still a continuous relation between patches, [94, 78].

3. For the chaotic case, as we increase the parameter mismatch (or equally decrease the coupling), this deterministic relation is lost and the patches evolve in an unrelated manner.

The first condition, normal hyperbolicity of the attractor $A$, can be numerically established again using Liapunov exponents. The definition of normal hyperbolicity [110] requires that vectors transverse to $T_x M_S$ experience contraction stronger than vectors inside $T_x M_S$. If this condition is satisfied then, for small parameter mismatches, the subsequent invariant manifold will be diffeomorphic to $M_S$. In terms of Liapunov exponents, this means that we require that for all $x \in A$, $v \in T_x M_S$ and $u \in T_x M_S$,

$$\lambda_{\perp}^{\max}(x, v) < \lambda_{A}^{\min}(x, u),$$

(5.31)
where $\lambda_{\text{max}}$ is the maximal normal Lyapunov exponent and $\lambda_{\text{min}}$ is the smallest Lyapunov exponent of $A$. However, as noted in [50], this is a necessary condition, but not sufficient, as we cannot calculate the minimal Lyapunov exponent for all the unstable cycles embedded in $A$. Whether this set of zero measure can generically effect the normal hyperbolicity of $A$ is still an open question.

Figure 5.5: A plot of the neutral normal hyperbolicity curve in $(\epsilon_1, \epsilon_2)$ space. Above the curve the GS will be a diffeomorphism (up to a set of zero measure) but below it will not be.

For the NPZ system in the chaotic regime, we found that $\lambda_{\text{min}} = -0.096$. Figure (5.5) shows a neutral normal hyperbolicity curve, in $(\epsilon_1, \epsilon_2)$ space, for the case of two coupled NPZ systems. We add that the calculation of this curve did not employ a type of modified Newton method to find the zeros of the function $G(\epsilon_1, \epsilon_2) = \lambda_{\text{max}}(\epsilon_1, \epsilon_2) - \lambda_{\text{min}}(\epsilon_1, \epsilon_2)$ due to computational constraints. Instead we employed an ad hoc algorithm, which we found captured as much of the curve as possible but
unfortunately couldn't completely retain the symmetry of the curve.

To look at how this property varies with the dynamics of the original patches figure (5.6) shows the same curve in \((d, \epsilon)\) space (symmetric coupling so \(\epsilon_1 = \epsilon_2\)). Surprisingly we see that quite strong coupling is required for \(A\) to be normally hyperbolic, except within a neighbourhood of the chaotic regime around \(d = 0.142\).

![Figure 5.6: Neutral normal hyperbolicity curve in \((d, \epsilon)\) space. Note that strong coupling is required for normal hyperbolicity except around the chaotic regime.](image)

This effect can be understood by consideration of figures (5.7) and (5.8), where we plot the Liapunov exponents of \(A\) with regard to the closure rate \(d\). We find that there is always a relatively large negative exponent except around the chaotic regime. Very negative values of \(\lambda_A^{\text{min}}\) means we require strong coupling so that \(\lambda_A^{\text{max}}\) satisfies (5.31).

In figure (5.9), we see the nature of the generalised synchronisation when the at-
tractor $A$ is normally hyperbolic. The time series of nutrient for two, non-adjacent patches exhibit phase synchronisation. This situation has been observed in the Canadian hare-lynx data [17, 49], even though the individual patches evolve chaotically their phases were almost identical over large spatial scales. If we decreased either the coupling strength or the size of the stochastic parameter perturbation, this phase-synchronised behaviour would be lost.

5.2 Unidirectionally Coupled Patches

In the first section, the coupling between patches was in both directions and full synchronisation (of identical systems) was possible but its stability depended on the strength of the coupling. Another class of coupled systems exists where the coupling
Figure 5.8: Plot of the maximal Lyapunov exponent close up round the chaotic region at $d \approx 0.142$.

goes in one direction only and consequently one system will be independent of the other, while the other system feels the influence of the first system. For two variables these are called skew product systems and in a simple mapping format they have the general structure,

$$x_{n+1} = f(x_n), \quad y_{n+1} = g(x_n, y_n).$$

for the one dimensional variables, $x_n$ and $y_n$. The biological basis for looking at systems of this type is where the plankton populations reside in reservoirs or lakes, or are separated by large distances. In these cases, coupling is unlikely to be from direct interaction of the planktonic species but from large scale environmental factors such as temperature. As previously mentioned, high degrees of synchrony in the time series of larval plaice in the Irish and North Seas have been observed in [32] and it is unknown which factor dominates, migration or temperature. The forc-
Figure 5.9: Phase-synchronisation in the nutrient patch lattice dynamics, $N_i$ and $i = 1, 2$. Here, $\epsilon = 0.075$ and there is a small, stochastic variation of $O(10^{-3})$ in the closure rate $d$.

The synchronisation of synchrony via large-scale environmental effects in linear stochastic systems has been dubbed the Moran effect [44]. In the following section we aim to give a mathematical reason behind how this phenomenon may be occurring in non-linear ecosystem models with external forcing.

### 5.2.1 External coupling & Patch Synchronisation

In line with the general set up of section 4.2, the dynamics of the patch inhabitants is denoted by $G$ but we now consider some external forcing field, $E$, such that $E$ exerts an effect on the patch inhabitants $s$ but is itself unaffected by them. This could be anything from temperature to salinity but we consider it only in a general
sense. This means we have the following so-called drive-response system:

\[ \dot{E} = F(E), \]
\[ \dot{s} = G(s, \Theta, E), \]

(5.32)

and the coupling from the (autonomous) drive system, \( E \), to the response system, \( s \), is characterised by the parameter vector \( \Theta \) (under the assumption that if \( \Theta = 0 \) both systems evolve independently).

In general, systems of this kind cannot fully synchronise (\( \dot{E}(t) = \dot{s}(t) \)) but they can generally synchronise in a slightly weaker sense than defined in the section 4.1.3. In [1] the generalised synchronisation of chaotic drive-response systems, such as the one in (5.32), was defined to be associated with the existence of some transformation \( \Phi: E \rightarrow S \) taking trajectories on the attractor in the drive space \( E \) to the attractor in response space \( S \), i.e. \( s(t) = \Phi(E(t)) \). It is assumed that this is also independent of the initial conditions of the response system, \( s(0) \), as long as they are inside the basin of attraction of the generally synchronised attractor. The existence of \( \Phi \) is also only required after transients have decayed and in [1] the following properties of \( \Phi \) were laid down:

1. \( \Phi \) has no explicit time dependence.
2. Points in \( E \) space are mapped to \textit{points} in \( S \) space. The function \( \Phi \) need not be injective and consequently not invertible but has a finite number of branches (with some rule for moving from branch to branch).
3. On each branch, \( \Phi \) will be locally continuous.
This notion of GS is weaker than in the bidirectionally coupled case but still allows for the prediction of response variables when only the drive variables are known. The main result from [1] was that by constructing an auxiliary system,

\[ \dot{r} = G(r, \Theta, E), \]  

which is identical to the original response system, we can detect the presence of GS. Statistical tests for the confidence of the existence of such a deterministic transformation \( \phi \) exist, [78, 94], but this so called auxiliary systems approach defined in [1] is a method that is easily implemented, both numerically and experimentally. The premise begins with the assumption that if all the response systems \( (s_1, ..., s_N) \) are being driven by the same driving signal, \( E \), then they will inhabit the same attractor, so long as their respective initial conditions lie in the same basin of attraction. It is possible that \( s_i(t) = s_j(t) \), \( i \neq j \), but quite unlikely if \( s_i(0) \neq s_j(0) \), especially in the chaotic case due to exponential separation of nearby orbits. However, if the response systems are synchronised to the drive system (via \( \phi \)) then it is natural that the solution \( s_i(t) = s_j(t) \) exists.

It transpires that the stability of the complex, generally synchronised oscillations, \( s_i(t) = \phi(E(t)) \) is equivalent to the stability of the comparatively simple regime where \( s_i(t) = s_j(t) \). To see this equivalence, first consider the linearised equations for the evolution of \( \xi_i(t) = s_i(t) - \phi(E(t)) \) and \( \xi_j(t) = s_j(t) - \phi(E(t)) \) and naturally \( i \neq j \). We have that

\[ \dot{\xi}_i = DG^t(\phi(E), \Theta, E) \cdot \xi_i, \]
\[ \dot{\xi}_j = DG^t(\phi(E), \Theta, E) \cdot \xi_j, \]  

(5.34)
and \( DG^i = \partial_x G(X, \Theta, E(t)) \) is the Jacobian matrix of \( G \), evaluated for the generally synchronised motions. The linear evolution of \( \xi_i(t) - \xi_j(t) = \mathbf{s}_i(t) - \mathbf{s}_j(t) \) is identical to that of \( \xi_i(t) \) and \( \xi_j(t) \), meaning they have the same Jacobian matrix, \( DG^i(\cdot, \Theta, E) \) as in (5.34). This now means that the linear stability of the generally synchronised motions \( \xi_i(t) = \mathbf{s}_i(t) - \phi(E(t)) \) in \( E \ominus S_i \) is both a necessary and a sufficient condition for the linear stability of the synchronised motions \( \mathbf{s}_i(t) = \mathbf{s}_j(t) \), in \( E \ominus S_1 \ominus \cdots \ominus S_N \). Consequently, if the external forces act strongly on the patches, the dynamics of the individual patches will become synchronised.

The auxiliary system approach was developed in [1] for the detection of GS in these uni-directionally coupled systems. Here, we shall begin from the opposite direction and show that any type of large scale forcing of planktonic populations (with such a drive-response structure) can force the dynamics of uncoupled patches to synchronise, if the “forcing” (coupling) is strong enough. Indeed this could well be the type of process at work where we see synchronicity in seemingly uncoupled larval plaice populations [32] or in the Canadian hare-lynx data [49]. The auxiliary systems approach involves only two response systems but here we have generalised the idea in a logical manner to an arbitrary number of response systems and turned the argument around to show that population synchrony can be a result of the environmental driving system being generally synchronised with the response systems.

To demonstrate, numerically, how this process occurs we shall consider the response (and naturally the auxiliary) system to be defined by the NPZ system in (5.3). For the driving system, we consider the scenario where \( E \) is chaotic but the ideas apply to any kind of forcing dynamics. The Rössler system has long been used
as a dynamical system that exhibits chaotic oscillations. Here, it may have no direct physical relevance but it is autonomous (both easier to integrate and fits the criteria of the drive system) and is comparatively simple:

\[
\begin{align*}
\dot{x}_1 &= -(x_2 + x_3), \\
\dot{x}_2 &= x_1 + 0.2x_2, \\
\dot{x}_3 &= 0.2 + x_3(x_1 - \mu),
\end{align*}
\]

(5.35)

where we use \(\mu = 5.7\). It is conceivable that such a system may represent mixing, temperature, salinity or other physical forcing that experiences no effect from the planktonic populations. Figure (5.10) shows a projection of the Rössler attractor in \((x_1, x_2)\) space.

![Chaotic Rössler attractor](image)

**Figure 5.10:** The chaotic Rössler attractor in \((x_1, x_2)\) space.

Our drive system is defined by \(E = T(x)\) (where \(x = (x_1, x_2, x_3)\) and \(T\) is a linear translation with rescaling so that, for orbits on the attractor, \(E_i \in (0, 1)\)). Again, for simplicity, we consider the case where only one of the drive variables is coupled to only one of the response variables. We only couple the variable \(E_1\) to the
equation for the nutrient evolution from (5.3) and thus

$$\dot{N} = \dot{s}_1 = G_1(s) + \theta (E_1 - s_1),$$

and $\theta > 0$. The corresponding equation for the auxiliary system is identical and in figure (5.11) we plot the attractors in the $(N_1, N_2)$ plane for the synchronised and unsynchronised cases. For some particular value of the coupling, the robustness of the generalised synchronisation is lost and the response and auxiliary systems move from a synchronised to an unsynchronised regime for some $\theta \in (0.005, 0.01)$.

![Figure 5.11: Attractors in the $(N_1, N_2)$ plane for the unsynchronised (left, $\theta = 0.001$) and the synchronised (right, $\theta = 0.01$) motions. For both cases, $N_1(0) \neq N_2(0)$](image)

These ideas are not restricted to the case where $E$ and $s$ exhibit chaotic dynamics. We have shown that, even in the extreme case of both drive and response systems being chaotic, indirectly coupled patches may synchronise if the external forcing is sufficiently strong enough. As before, we could ask the question of what happens
when the response and auxiliary system have slight parameter differences. The answer lies again in the idea that the manifold in question must be normally hyperbolic. The robustness of the manifold $s(t) = \phi(E(t))$, thanks to (5.34), is equivalent to the robustness of the much simpler manifold, $s(t) = r(t)$. If this simpler manifold is normally hyperbolic, such that condition (5.31) holds, then the dynamics of the response and auxiliary system(s) will be generally synchronised themselves, but in a much stronger manner than that implied by the aforementioned definition of the transformation $\phi$.

5.3 Conclusions

By viewing interacting plankton patches as a coupled lattice type model we have demonstrated several different routes by which we see transition from synchronous (spatially homogeneous) to non-synchronous (spatiotemporally varying) dynamical regimes. This non-continuum approach means we can classify some of these transitions in terms of a bifurcation (blowout and riddling bifurcation) by using a two patch system as an example and demonstrate (analytically and numerically) how these results would generalise to $n$ coupled systems. Also, we show what effect some variation in patch system parameters may have. We classify the types of generalised synchronisation of patches and again how we can compute areas of parameter space in which each definition applies.

We considered only normal coupling parameters here but we could expect that most actual reaction parameters would not be normal. A term of some debate in
most models of plankton dynamics is the higher predation or closure term. We employ a simple linear form but, from our perspective, we could consider a coupled form of the higher predation function. Predatory fish can presumably travel from patch to patch significantly more effectively than zooplankton and phytoplankton. One idea we considered is to assume that if a fish saw more prey in one patch it would prefer to concentrate on that patch. Thus we considered the following type of higher predation function

\[
\frac{d z_i^2}{\chi z_i + (1 - \chi) z_j},
\]

where \( i \neq j \) and \( \chi \in [0, 1] \). Thus a predator will choose which patch to feed on based upon zooplankton numbers in each patch. Here \( \chi \) is a type of biased predatory term as if \( \chi < \frac{1}{2} \), the predator prefers \( z_j \) and vice versa for \( \chi > \frac{1}{2} \). Also, it is easily seen that \( \chi \) is a normal parameter and in \( M_S \) the closure term reduces to the original linear form. Using \( d \) as our patch bifurcation parameter is work for the future and it will be interesting to see the importance of higher predation on patch dynamics. Similar transitions are observed when we vary non-normal parameters and a good review of the phenomena observed when varying non-normal parameters can be found in [9].

Another open question concerning the existence of chaotic plankton dynamics. Due to the huge amount of effort required to collect a complete and reliable data set of spatio-temporal plankton distributions, it is difficult to decide between stochastic and possibly chaotic effects. However, field data of the dynamics of diatom communities [100], showed the presence of chaos. Indeed it has been suggested that chaos is
a good natural state for populations, [4], in that chaotic fluctuations allow the persistence of coexisting species even in unfavourable conditions. We hypothesise that a coupled analogue of this phenomena may have been observed here. The strongest form of generalised synchronisation requires that the synchronised attractor be normally hyperbolic, (5.31), which in turn dictates the strength of the coupling that is required. For the linear closure model and the (structurally and parametrically) different quadratic closure model, the most negative Liapunov exponent was large, except around the chaotic regime. This means that, around the chaotic regime, the coupling required for the attractor to be normally hyperbolic is much weaker. This behaviour can not be verified for a general chaotic population model but is likely as the individual dynamics are less locked into their respective dynamics in the chaotic regime than in, say, a periodic regime. The phenomenon of generalised synchronisation has been seen in ecological models [17, 49, 52] and data (Canadian Hare-Lynx system) that also exhibit chaotic behaviour. Perhaps this is because it is easier to generally synchronise when in a chaotic regime as our numerical experiments have suggested.

Edwards & Bees [27] also considered a quadratic $dZ^2$ closure term and found the existence of a chaotic solution, for different parameter values than those described in the linear closure model. To see if this type of generally synchronising behaviour could be a generic trait, we also calculated the values of the Liapunov exponents, for a range of values of $d$, as in figure (5.7). As we see in figure (5.12), there exists a region around the chaotic regime close to $d = 0.5$, see figure (5.13), where the strength of coupling needed to satisfy the normal hyperbolicity condition, bar two
or three isolated discontinuities, is consistently less than anywhere else. This is precisely the effect which allowed the patches to generally synchronise that much easier in the linear closure model. The other patch parameter values for the quadratic closure model can be found in [27].

Figure 5.12: Neutral normal hyperbolicity curve in \((d, \epsilon)\) space for the quadratic closure model. Note that stronger coupling is required for normal hyperbolicity except around the chaotic regime.

The coupling of patches may not always be so apparent. Certainly for such large scale effects as temperature, external forces can act on patches with no direct coupling whatsoever and some evidence [32, 44] exists for this type of forcing having the ability to cause population synchrony. By adapting a test from [1] for the detection of generalised synchronisation of so called drive-response systems, we demonstrated that if the external forcing exerts a strong enough effect on the two patches then they will synchronise. We also mention how this concept generalises to arbitrarily
Figure 5.13: Plot of the maximal Liapunov exponent close up around the chaotic regime, $d \approx 0.5$ for the quadratic closure model.

...many patches and what may happen if the response and auxiliary system(s) have slight parameter differences.

...From macro to micro scales, many ecosystems exhibit a patchy structure and factors such as migration, temperature variation mean that each population patch may be coupled to several of the others (directly or indirectly). Although the work in this chapter is with regard to planktonic dynamics, the non-linear mechanisms and phenomena involved are applicable to terrestrial models as well. Treating these interacting populations as a system of coupled oscillators gives more scope for both incorporating the coupling and numerical study so we suggest that patchy dynamics are better studied in a non-continuum sense.
In the previous chapter, we saw that the dynamics of patches of interacting plankton populations can be linked in a variety of different ways and many methods for dealing with such systems are available. These methods are ideal for when the dynamical systems governing the patches are explicitly known. However, what can be achieved in an experimental situation when only raw data are available? The answer is potentially quite a lot using established ideas from the analysis of non-linear time series. Here we present implementable methods to first statistically measure if several patches of plankton are related randomly or deterministically (generally synchronised) and then create a representative patch system such that the error between the averaged system’s dynamics and the individual patch dynamics is minimised. Furthermore, we suggest that, even for chaotic data, we can circumvent some of the problems of sampling populations over large length and time scales.

The work is organised as follows, in section 6.1 we briefly explain how to reconstruct individual patch dynamics from experimental measurements. The statistical
method of detecting generalised synchronisation of time series, first presented by Pecora et al. [78], is briefly explained in section 6.2 along with illustrative examples of the technique in practice. In section 6.3 we explain our method of constructing a meta-patch model from the individual patch time series, which is constructed in such a way as to keep as much of the individual behaviour as possible. In section 6.4.1, we outline a fairly standard predictive estimator from non-linear time series theory and the pros and cons of using such a method in an experimental setup. Model fitting of experimental data is an area of current interest [20, 46] in planktonic dynamics, specifically detecting various types of density dependence/independence in planktonic populations. The analysis of possibly patchy data is not restricted to just planktonic systems. In fisheries management, large areas are treated as a single stock, but in reality they are often modelled [90] as possibly interacting substocks, which are then combined into a single dynamical model for use in risk assessment/management decisions. In section 6.5, using various constructed “patchy” data sets as examples, we show how our method of constructing a representative meta population time series can help to reduce the incidence of error likely to appear when using less rigorous methods of sampling. Some degree of noise (both process and observational) is always going to be a factor and in section 6.6 we address some of the existing methods of noise reduction and their good and bad points.
6.1 Delay embedding and state space reconstruction

Many theoretical models exist for planktonic populations [112] and they encompass a broad range of complexity and type. Recently, by using experimental time series data, many attempts have been made to actually fit the data to specific types of population models. However, this is not a new field; [37] investigated field data for evidence of Lotka-Volterra type behaviour and in [30] Feller was looking for evidence of logistic type growth in biological systems. In [46], several data sets were analysed for evidence of predator dependent prey dynamics as opposed predator independent constructions, such as Michaelis-Menten or Holling type responses. Other similar work on model fitting can be found in [20]. These approaches first assume some underlying system equations and consequently fit the data to these equations. However, this is always a tricky procedure as we can never truly be sure of the accuracy of the original model. For instance, the model may be chosen with some fixed parameter(s), but seasonally varying parameters, such as temperature, means that the form of the solutions of the original model may change quite drastically with time.

With this in mind we present a method of investigating possibly many interacting populations with no knowledge of the underlying equations governing the population and no restrictions on the number of species present in the patch populations (apart from that there are finitely many of course!). The aims are to pursue a sensible construction of a meta-population system representing interacting populations and to construct a prediction of future population trends, all from individual time series.
data from a selection of patches. Also, we would like to show how this meta-model might be employed for the purposes of model fitting. Given a situation where the only information we have on a dynamical system is a measurement of one system quantity (variable) it might seem quite natural to assume that there is little we can say about the dynamics as a whole. However, in [106] Takens proved that if we reconstruct the dynamics, via a technique called delay reconstruction, from this one measurement then the reconstruction will be an embedding of the underlying dynamics in some suitably higher-dimensional space. This tells us that the dynamics in the reconstructed space are qualitatively the same as (or smoothly conjugate) to the original dynamics. So, to understand the original system we can gain much insight by studying the reconstructed system. However, some information is lost; we don’t know the specific relationship between the reconstructed and original systems. This restricts us to making only qualitative statements about the dynamics of the original system, based on the observation of the reconstructed system.

We assume that we have some specific system, \( r(t) \), we wish to study and that an unknown process determines the time evolution of the system, \( f : r(t_0) \rightarrow r(t) \). If \( \varphi(r(t)) \) is some measurement function (e.g. nutrient concentration, phytoplankton biomass) then the time evolution of \( r(t) \) is now given by the time series \( \varphi(r(t)) \). In practice this measurement will be discrete and we assume for simplicity’s sake that the interval of sampling is constant, i.e. our measurement becomes \( \varphi(r(i\tau)) \) for some suitable fixed \( \tau \) and \( i \in \mathbb{N} \). The choice of the size of the delay is obviously important and highly case dependent. Too small a delay and the dynamics would seem almost at equilibrium, even if the original dynamics are oscillatory. Too large
and the system would seem much more random (if the underlying dynamics are chaotic). Lastly, an extreme (but perhaps very unlikely) case would be choosing a time delay that is an integer harmonic of the frequency of oscillation and, hence, the reconstructed dynamics would consist of a fixed point.

The above method is used to discretely sample the system to generate a new dynamical system as follows. Given $N$ measurements $\varphi_1, \ldots, \varphi_N$, we construct the following $d_e$-dimensional dynamical system (and we obviously require that $N > d_e$),

$$s_i = (\varphi_i, \varphi_{i+1}, \varphi_{i+2}, \ldots, \varphi_{i+d_e-1}),$$

where $i$ is a discrete time variable and $d_e$ is the embedding dimension.

Now Takens Embedding Theorem states that if $d_e \geq 2m + 1$ (where $m$ is the original dimension of the system) then the dynamics, in $\mathbb{R}^d$, are given by $F : s_i \rightarrow s_{i+1}$ and are equivalent to the original (discretised) dynamics $f : r(i\tau) \rightarrow r((i+1)\tau)$ under some (unknown) smooth, invertible transformation $\Pi$, i.e. $f = \Pi^{-1} \circ F \circ \Pi$. Hence, we can qualitatively reproduce the dynamics of a patch just by periodically measuring the concentration of one of the patch organisms. To show how this is achieved we use the NPZ system of the previous chapter as an example (we already know the dynamics but it serves well as an illustrative example). The dynamics are governed
by the following coupled ODEs:

\[
\frac{dN}{dt} = -\frac{Na}{(e+N)(b+cP)}P + rP + \frac{\lambda P^2}{\mu^2 + P^2}Z + \gamma dZ + k (N_0 - N),
\]

\[
\frac{dP}{dt} = \frac{Na}{(e+N)(b+cP)}P - rP - \frac{\lambda P^2}{\mu^2 + P^2}Z - (s+k)P, \tag{6.2}
\]

\[
\frac{dZ}{dt} = \frac{\alpha \lambda P^2}{\mu^2 + P^2}Z - dZ.
\]

This system displays dynamics from equilibria to limit cycles to chaos, under the variation of the closure rate, \( d \). A detailed description of the equations in (6.2) can be found in [27]. By using \( \tau = 1 \) day as our delay we can unfold the original \((N, P, Z)\) dynamics from the 3-dimensional state space into a 7-dimensional state space using only the phytoplankton concentration, \( P \), as follows,

\[
\Pi : (N_t, P_t, Z_t) \rightarrow (P_{t\tau}, P_{(t+1)\tau}, P_{(t+2)\tau}, P_{(t+3)\tau}, P_{(t+4)\tau}, P_{(t+5)\tau}, P_{(t+6)\tau}). \tag{6.3}
\]

Figure (6.1) shows the original dynamics and the reconstructed dynamics for \( d=0.1415 \) and \( d=0.142 \).

So if we can measure the concentration of a particular patch inhabitant (e.g. chlorophyll concentration) then we can reproduce the dynamics of the patch. However, the aim of this work is to see how we can use these ideas to simplify interacting patches. One factor that could prove to be a problem is that if we are sampling over very large spatial scales, we cannot sample all patches at the same time so there will undoubtedly be some intrinsic gaps in the data. Also, the situation we are investigating is where we don’t know if the patches evolve randomly or are synchronised in some way. If they are dynamically coupled then we cannot simply sample each one
Figure 6.1: Reconstructed dynamics using $P$ as the measurement function. In both figures we plot $(P_{tr}, P_{(i+5)r})$ (main), $(P_t, Z_t)$ (inset) for $d=0.1415$ (left) and $d=0.142$ (right).

and move onto the next hoping to bring all the data together at the end regardless of the delay. This would not show the effect that patches have on each other, as we would be comparing the patch dynamics at fundamentally different times.

So, given our patch data we would like to see what, if any, similarities exist between the individual patch dynamics and, if there are similarities, how can we use this fact to simplify the system.

### 6.2 Experimental verification of generalised synchronisation

Given many time series measurements such as those in figure (6.2), we would like to see if the patches evolve randomly or are linked (generally synchronised) in some
deterministic manner. There are many properties that we might want to investigate such as the existence of a continuous function from one patch to another, invertibility (and injectivity) and also differentiability of this generalised synchronisation. Several competing methods have been developed for statistically testing for such properties as continuity and differentiability. We present the method developed by Pecora et al. \[78\] which corresponds to measuring the confidence in the existence of the required properties for generalised synchronisation. Other approaches, such as Rulkov et al. \[94\], used a variation on the idea of false nearest-neighbours to see if any functional relationship exists between the time series. We suggest the method developed in \[78\] primarily because it is simple to apply, very versatile and as we shall see, yields good results even for relatively short data sets. By this we mean that you can decide the "level" of generalised synchronisation you wish to test for, from basic continuity through to diffeomorphic behaviour (differentiable with differentiable inverse). Furthermore, the results consist of confidence probabilities for the various properties desired and are thusly easily understandable and interpretable.

Here, we shall test for the existence of a homeomorphism (continuous, invertible and injective) between the individual patches, as with experimental data it seems somewhat optimistic to test for differentiability. To describe the method for the statistical test of continuity we first go back to the well known \(\varepsilon - \delta\) definition of continuity. The function \(f\) is continuous at a point \(x_0\), in some metric space \((X,d)\) (and \(d\) is the standard Euclidean metric in \(\mathbb{R}^n\)), if \(\forall \varepsilon > 0, \exists \delta > 0\) such that \(d(x,x_0) < \delta \Rightarrow d(f(x),f(x_0)) < \varepsilon\). In rougher terms, for any given neighbourhood
of $f(x_0)$, we can always find a corresponding neighbourhood of $x_0$ such that all the points in this neighbourhood belong to the neighbourhood of $f(x_0)$. This guarantees that neighbouring points are mapped to neighbouring points.

The algorithm itself starts by choosing some values of $\epsilon$ and $\delta$ and finding all the points within a distance $\delta$ of $x_0$. Now we check that all these points are within $\epsilon$ of $f(x_0)$, if not we decrease $\delta$ by some suitable amount and begin the process again. We repeat this process until all the images of points in the $\delta$ set are inside the $\epsilon$ set or $\delta$ is so small that no points can be found in the $\delta$ set. If such a $\delta$ exists, then the function fulfils the standard definition of continuity described above, at least at the resolution $\epsilon$.

Secondly, a null hypothesis is developed to be used in the test for continuity. If we have made $N$ measurements to generate our time series then let $M_\delta$ be the number
of points in the δ set and $M_\varepsilon$ be the number of points in the ε set (typically $M_\varepsilon > M_\delta$ as points outside the δ set may be mapped in the ε set). Naturally, the points $x_0$ and $y_0$ are excluded and do not influence the statistical test. The null hypothesis is that if the two time series are related in a random manner (independently distributed in the image space with respect to points in $(X, d)$), then the probability $p$ of one of the $M_\delta$ points mapping into the ε set is well estimated as $M_\varepsilon/N$. So, consequently, the probability that all the $M_\delta$ points will be mapped into the ε set is $p^{M_\delta} = (M_\varepsilon/N)^{M_\delta}$.

If this probability is suitably small then the null hypothesis can be rejected and the two time series are not randomly related. It must be emphasised that a rejection of the null hypothesis is not proof of the alternate hypothesis, the existence of a continuous function. It merely gives us confidence that all the sample points in the $M_\delta$ set being mapped to the ε set is not due to a random coincidence.

How can we express this confidence numerically? The situation above describes a probability with a binomial distribution, $P(m; M_\delta, p)$ (the probability of finding $m$ points out of $M_\delta$ in the ε set given that the probability of finding one is $p$). The event of interest is in the tail of this distribution ($m = M_\delta$) hence the probability of this happening is $p^{M_\delta}$. The likelihood of the event can be defined as $p^{M_\delta}/p_{\text{sup}}$ where $p_{\text{sup}}$ is the maximum probability in the binomial distribution usually occurring at a certain $m < M_\delta$. This says that not only must $p^{M_\delta}$ be small, it must be small in comparison with the most likely event.

The confidence can now be expressed via the continuity statistic [78]

$$C(\varepsilon, i) = 1 - \frac{p^{M_\delta}}{p_{\text{sup}}}, \quad C(\varepsilon, i) \in [0, 1],$$

(6.4)
where \( i \) indexes the point at which continuity is being tested at \((x_i = x_0)\) and we make clear the explicit dependence on \( \varepsilon \). If \( C \approx 1 \) then we are confident of continuity at \( x_i \) but if \( C \approx 0 \) we are not confident of continuity because we have no statistical basis on which we can reject the null hypothesis.

This is clearly a test for continuity at a single point, so to gain a statistic for the whole time series we randomly choose \( M_p \) points in the time series and average \( C \) over these points as follows:

\[
C(\varepsilon) = \frac{1}{M_p} \sum_{i=1}^{M_p} C(\varepsilon, i), \quad C(\varepsilon) \in [0, 1].
\] (6.5)

The subsequent statistics are highly dependent on the choice of \( \varepsilon \). If \( \varepsilon \) is of the order of the size of the attractor in reconstruction space then we would expect that \( C(\varepsilon) \approx 1 \) even if the two original time series were randomly related. If \( \varepsilon \) is too small on the other hand, noise in the system or computational resolution will almost certainly give us results where \( C(\varepsilon) \approx 0 \). They are also dependent on the number of data points available as the sum in (6.5) will invariably improve with \( M_p \), this is something we shall address later in this section.

There is no set method for choosing the value of \( \varepsilon \) but we suggest that one can use general information from the time series to provide a sensible estimate for the range of \( \varepsilon \). A logical upper bound for \( \varepsilon \) would be the magnitude of the largest variation in the time series data, \( \varepsilon = \sup \{ d(\varphi_m, \varphi_n) \mid m, n = 1, ..., N \} \). Any value of \( \varepsilon \) larger than this will give \( C(\varepsilon) = 1 \), even if the two time series evolve in an uncorrelated manner. The situation for a minimum bound on \( \varepsilon \) is a little more complicated. For one, we would like to be sure that we are testing local effects in the time series,
but choosing \( \varepsilon \) too small may incorporate effects that we have no real interest in investigating, such as small levels of system noise. A rule of thumb we would suggest is to consider values of \( \varepsilon \) down to a distance an order of magnitude smaller than \( \varepsilon \).

The test for inverse continuity of the generalised synchronisation follows again from the standard \( \varepsilon - \delta \) definition, i.e. the function \( f \) has continuous inverse at a point \( y_0 \), in some metric space \((Y,d)\), if \( \forall \delta > 0, \exists \varepsilon > 0 \text{ such that } d(y,y_0) < \varepsilon \Rightarrow d(f^{-1}(y), f^{-1}(y_0)) < \delta \). Using the method described for the continuity test our statistic for inverse continuity at \( y_0 \) is simply

\[
I(\varepsilon, i) = 1 - \frac{P_{\sup}^{M_{\varepsilon}}}{P_{\sup}}, \quad I(\varepsilon, i) \in [0, 1],
\]

where now \( p = M_0/N \) (note that \( p_{\sup}, M_{\delta}, \text{ and } M_{\varepsilon} \) will not be the same as those for the continuity test). As before we average this point statistic over randomly chosen points in the time series to get the average statistic

\[
I(\varepsilon) = \frac{1}{M_p} \sum_{i=1}^{M_p} I(\varepsilon, i), \quad I(\varepsilon) \in [0, 1],
\]

and all the comments and suggestions on the continuity statistic apply (with the obvious switch for \( \varepsilon \) and \( \delta \)). The final statistic we can construct is a confidence statistic for whether the function \( f \) is a homeomorphism (onto and injective with continuous inverse). This can be expressed by the compound statistic, \( C(\varepsilon)I(\varepsilon) \). If this product is close to unity we can be confident that the function \( f \) is indeed a homeomorphism.

To show how the statistics can be applied and how to interpret them we look at time series taken from the following coupled array of NPZ systems

\[
s_j = G_j(s_j) + \epsilon_{j+1}(s_{j+1} - s_j) + \epsilon_{j-1}(s_{j-1} - s_j),
\]

(6.8)
where the $G_j$ are the same as in the single NPZ system in (6.2) and $j = 1, ..., M$ denotes the patch number. This serves as a good paradigm system for many coupled patches and, using the phytoplankton concentration of each patch as the measurement for the time series construction, we will apply the continuity/inverse continuity algorithm for different parameter scenarios.

Figure 6.3: Typical phase portraits for adjacent patches, $(P_j, P_{j+1})$, where (i) $\epsilon = 0.2$ and no parameter mismatch (top left), (ii) $\epsilon = 0$ (top right), (iii) $\epsilon = 0.005$ and small variation in closure rate $d$ (bottom left) and (iv) $\epsilon = 0.2$ and large variation in closure rate $d$ (bottom right). For these simulations we set $M=3$.

Figure (6.3) shows the phase portrait of the time series of adjacent patches. Details of the fourth-order Runge-Kutta [89] procedure employed in the integration of (6.8) can be found in the Appendix. We use data from (6.8) to generate $M$ time series $\varphi_i^j$ where $i = 1, ..., N$ and $j = 1, ..., M$ for different parameter scenarios. The four situations shown in figure (6.3) represent four likely scenarios for coupled...
systems:

- Identical patches which are strongly coupled. The dynamics are synchronised and the generalised synchronisation relation $\Phi : \varphi_i \to \varphi_i$ is just the identity map.

- Completely uncoupled patches evolving in a random fashion.

- Patches where the coupling is fairly close to the synchronisation threshold $\epsilon = 0.001$ [42] and there is a small, randomly generated mismatch, $\Delta \approx O(10^{-4})$, in the closure rate $d$ in each patch.

- Patches which are highly coupled but there is a relatively large parameter mismatch, $\Delta \approx O(10^{-2})$, in the closure rate $d$ in each patch.

In figure (6.4) we plot the continuity statistics for the four scenarios shown in figure (6.3). We see that for the synchronous case (i) that at all resolution scales $\epsilon$ we have high confidence in the existence of the generalised synchronisation (as it is just the identity map). For the completely asynchronous case (ii) we see that as we decrease the resolution scale $\epsilon$, the confidence rapidly tends to zero suggesting that the patches evolve in a random manner. Case (iii) represents an interesting situation as the dynamics are intermittent and consequently the system spends long periods close to the diagonal, punctuated by bursts from synchronicity. This can be seen from figure (6.4) as we get a gradual decay to zero as we increase resolution, for example we have that at $\epsilon = 0.07$, $C(\epsilon) \approx 0.6$ suggesting that the patches spend around 60% of their time in sync (at this resolution).
In case (iv) we see that the statistics are very good \((C(\varepsilon)I(\varepsilon) \approx 1)\) down to a resolution of around \(\varepsilon = 0.025\). Above this level the strong coupling guarantees generalised synchronisation but below the parameter mismatch effects are evident in the rapid decay to zero of the statistics. All calculations were made using 2000 constructed data points. The value of \(\delta_{\min}\) for the calculation of \(C(\varepsilon)\) was set to the same value as \(\varepsilon_{\min}\) used for the calculation of \(I(\varepsilon)\) and \(\delta_{\min} = \varepsilon_{\min} = 0.001\).
In reality, large and reliable data sets are not always the norm in many biological situations. Measuring over large spatio-temporal scales brings with it very real constraints such as resource availability and financial limitations. This applies to the modelling and measurement of not just planktonic ecosystems but other oceanic and terrestrial systems as well. To assess the sensitivity of the test to the amount of data present, we look at the most extreme case (ii) (where statistically there is no relationship between patches). In figure (6.5) we see that, for a relatively shorter time series, the statistic $C(c)I(c)$ differs from that featured in figure (6.4) ($c$ is fixed at 0.05) but, as we increase the length of the time series the statistic $C(c)I(c)$ rapidly approaches that seen in figure (6.4).

![Figure 6.5: Homeomorphic confidence statistic for case (ii) as we increase the amount of available data, $N$ ($c = 0.05$ throughout).](image)

In figure (6.6) we show a similar plot to figure (6.5) but for case (ii) (strong generalised synchronisation). For a resolution of $c = 0.025$, around where the statistics
in figure (6.4) undergo a marked change, the homeomorphism confidence statistic $C(\varepsilon)I(\varepsilon)$ is plotted against increasing sample size,

![Graph showing $C(\varepsilon)I(\varepsilon)$ against $N$]

Figure 6.6: Homeomorphic confidence statistic for case (iv) as we increase the amount of available data, $N$ ($\varepsilon = 0.025$ throughout).

Even around the resolution ($\varepsilon = 0.025$) where the confidence in generalised synchronisation drops off fairly quickly, the statistics show little difference as we increase the amount of available data. All types of confidence/likelihood tests are prone to inaccuracies for short time series, due to limited data. However, the method employed here seems to converge to the expected result quite quickly with increasing data availability, even in the most extreme cases.
6.3 Creating a meta-population model

In many experimental scenarios [90] it is often advantageous to construct a kind of meta-model which describes the general population trends. How can we use the statistical test described in the previous section to decide which type of averaging procedure is best? We can separate the problem into two main areas as follows.

6.3.1 Randomly related patches

If, after applying the confidence test, we find that the dynamics of the individual patches are related in a random manner, we could average the dynamics by modelling it as a linear stochastic process using, for example, an autoregressive moving average (ARMA) [21] model:

$$\varphi_i = \alpha_0 + \sum_{k=1}^{R} \sum_{j=1}^{M} \alpha_k^j \varphi_{i-k}^j + \sum_{k=1}^{R} \beta_k \zeta_{i-k},$$

where the $\zeta_i$ are uncorrelated random variables, $R$ is the length of the regression and the $\alpha_k^j$ and $\beta_k$ are the model parameters. This is a slight variation on a fairly standard stochastic approach to time series modelling. The main difference is that both individual patch terms and their history are included in the first sum while the usual stochastic history term appears in the second sum. The advantages of such a formulation is that it can certainly capture the apparent complexity of a problem and many routines exist in statistical packages to implement the model. The down side is that because it is inherently stochastic, any predictive power is small in comparison to a deterministic non-linear approach. However, we are really interested in the scenario when the patches are not randomly related so we go no
further with this model.

6.3.2 Deterministically related patches

If we have found that, for a suitable range of our resolution parameter \( \varepsilon \), that both \( C(\varepsilon) \approx 1 \) and \( D(\varepsilon) \approx 1 \) then we can adapt various methods from dynamical systems aspects of time series analysis. The statistics tell us that we can be confident that there exists some homeomorphism between individual patches, \( \Phi(j, j') : \varphi_i^j \rightarrow \varphi_i^{j'} \), where \( j' \) is some arbitrary patch such that \( j \neq j' \). Consequently, we can take any suitable combination of the \( \Phi(j, j') \) and gain a homeomorphism from any patch to another. So we now consider the construction of a representative time series as follows:

\[
\tilde{\varphi}_i = T(\varphi_i),
\]

(6.10)

where \( \varphi_i = (\varphi_1^i, ..., \varphi_M^i) \) and \( T(\cdot) \) is an operator to be established. The beauty of working in delay reconstruction coordinates is that any intrinsic phase differences in the individual patch dynamics \( \varphi_i^j \) can be factored into equation (6.10) right at the start when we reconstruct the dynamics from the original measurements.

Several methods exist for determining a suitable operator \( T(\cdot) \) from an empirical time series. The proper orthogonal decomposition, POD [15], considers (6.10) as a modal decomposition,

\[
\tilde{\varphi}_i = \sum_{j=1}^{M} \lambda_j \mathcal{U}^j(\varphi_i^j),
\]

(6.11)

such that \( (\lambda_j, \mathcal{U}^j) \) and \( M \) represent an eigenvalue/eigenfunction pair and the number of degrees of freedom to be solved for, respectively.
This approach works well when dealing with turbulent data when there is no \textit{a priori} knowledge of the number of degrees of freedom of the system. In most practical applications, we know explicitly the number of patches, $M$. Also, the patch systems are generally synchronised hence, we hypothesise that a slightly simpler approach could be employed, one drawback of the POD is that it is computationally expensive.

The second method consists of reworking of (6.11) where we choose the form of the function, $U^j$, depending on the situation, leaving us a system to solve for the $\lambda_j$.

How do we go about choosing the functional forms of $U^j$? A popular choice in time series is to use \textit{radial basis functions}, [18, 22, 87, 99]. From the individual patch data we choose the \textit{radial basis centres}, $r_1, \ldots, r_R$, and fix $B(x) = \sqrt{x^2 + c}$, for some suitable constant $c$. The averaged time series is now given by

$$\bar{\phi}_i = \sum_{j=1}^{R} \lambda_j B(|| \phi_i - r_j ||),$$

(6.12)

which can work well for global, non-linear approximations as we obtain the non-linear behaviour from good choices of basis centres and the constant $c$, whereas the system is linear in the $\lambda_j$. The parameter $R$ determines how many radial basis centres are used and is problem specific.

The third possibility we envisage is a \textit{weighted linear average (WLA)},

$$\bar{\phi}_i = \sum_{j=1}^{M} \lambda_j \phi_i^j,$$

(6.13)

which is conceptually the simplest of the three possibilities we have mentioned here but is perhaps more accurate than the simplest notion of $\bar{\phi}_i$ being simply the sample
patch mean (SPM) of the $\varphi_i^j$, i.e.

$$\bar{\varphi}_i = \frac{1}{M} \sum_{j=1}^{M} \varphi_i^j,$$

(6.14)

We wish to preserve as much of the dynamics of the individual patches in the new averaged patch time series as possible so we choose to fit using the least squares error

$$e_k^2 = \sum_{i=1}^{N} (\varphi_i - \varphi_i^k)^2,$$

(6.15)

for $k = 1, ..., M$ and where $N$ is the sample length of the data. The error for the whole fit though is given by

$$e^2 = \sum_{k=1}^{M} e_k^2 = \Xi(\lambda),$$

(6.16)

where $\lambda = (\lambda_1, ..., \lambda_M)$. $\Xi(\lambda)$ is quadratic in the $\lambda_j$ and many routines exist for the minimisation of such forms (solving $\partial_{\lambda_j} \Xi(\lambda) = 0$ under the condition that the Hessian matrix is positive definite). Throughout this chapter, we used the Broyden-Fletcher-Goldfarb-Shanno minimisation algorithm [89].

These methods are based on the assumption of a deterministic relation between the individual patch time series. If we apply our rule of thumb for the resolution parameter this gives us a confidence bound of roughly $\epsilon \geq 0.05$ (where $\epsilon$ is the resolution of the continuity/inverse continuity statistic). If the determinism statistics are good above this threshold then we can be confident in applying the ideas from non-linear time series analysis. If not we would be advised to use some kind of stochastic method such as the modified ARMA model (see section 6.3.1). The statistics for cases (i) and (iv) from figure (6.3) are very good, $C(\epsilon)I(\epsilon) \approx 1$, the
statistics for case (iii) are not as good but suggest prolonged periods of synchronicity interrupted by the “random” bursting indicative of systems close to synchronisation.

The statistics for case (ii) rapidly converge to zero as we decrease $\epsilon$, suggesting that the two time series are randomly related (as we would expect of uncoupled chaotic trajectories with differing initial conditions).

![Figure 6.7: Average error, $e_i$, for the synchronous case (iii) at time $i$ between $\phi_i$ and the $\varphi_i\hat{,}$ when using the sample mean (top) and the weighted linear average approach (bottom). The total sample error for the sample mean was $e^2 = 18.764$ and for the weighted linear average $e^2 = 16.26$. Here the number of data points used to obtain the fit $N = 500$ and the number of patches was $M = 6$.](image)

Figure (6.7) plots the inter-patch error, when comparing the sample mean method, (6.14), with the weighted linear average method, (6.13), for case (iii), the syn-
chronous bursting case where we are reasonably confident of some level of generalised synchronisation. It can be seen that the weighted linear average method performs better than simply taking the sample patch mean, though the conceptual similarity of the two approaches is evident as the bursting behaviour occurs in the same places. The total error for the sample patch mean (SPM) was \( e^2 = 18.76 \) while the weighted linear average (WLA) method gave a total error \( e^2 = 16.26 \). Figure (6.8) shows the similar error comparison for a system of 6 patches in phase synchronisation, case (iv). In this case, the total sample error for the sample mean was \( e^2 = 17.89 \) and for the weighted linear average \( e^2 = 16.72 \). In both cases, the WLA method performed better (though not hugely) than the SPM method.

The radial basis function approach was implemented but, for the scenario of near-synchronous patch dynamics that we are considering, was very poor in terms of accuracy and computational time. To gain a level of accuracy appreciably close to those achieved with the WLA and SPM methods, much more computational effort was required. We hypothesise that this is because the radial basis function approach works well for approximating very non-linear systems but is not as suited to approximating the weaker non-linear behaviour seen in the near synchronous systems. On the whole, the WLA method captured the 'local' behaviour better than the SPM method but disappointingly did not radically decrease the inter-patch error as was hoped. The advantages of radial basis functions will become apparent in the next section when we wish to make short term predictions of the patch dynamics, given our meta-population data set.
Figure 6.8: Average error, $e_i$, for the synchronous case (iv) at time $i$ between $\bar{\varphi}_i$ and the $\varphi^i_j$ when using the sample mean (dashed line) and the weighted linear average approach (full line) for a patch system exhibiting phase synchronisation. The total sample error for the sample mean was $e^2 = 17.89$ and for the weighted linear average $e^2 = 16.72$. Here the number of data points used to obtain the fit $N = 500$ and the number of patches was $M = 6$.

6.4 Making use of the constructed meta-population time series

In the previous sections, it was demonstrated how one can first elucidate the link, if any, in the dynamics of interacting patches and then construct an average population time series which preserves much of the individual patch dynamics. We would like to show that one can then use this meta-population time series to both predict short term general trends in the patch system and also improve the quality of any fitting of the data to a specific model. Fisheries scientists have long been trying to predict
trends in fish stocks [90] so as to offer better advice on the allowable catches of these stocks. Planktonic populations have a direct effect on the state of many fish stocks [57] so some way of predicting trends in planktonic abundance may well be advantageous to the economically/ecologically important analysis of fish stocks.

Furthermore, the analysis of many multi-species ecosystems has concentrated on fitting a particular model to the time series measurements of that ecosystem [20, 30, 37, 46]. Particular mention is made by Jost and Arditi [46] that the lake from which they sampled the planktonic populations had several linked populations, but that this was not factored into the treatment of the data with regard to model fitting. We suggest that constructing the meta-population before performing either predictive or model fitting procedures will increase the applicability of the results when dealing with a patchy ecosystem.

6.4.1 Predictive estimation

The main question we want to address in this section, with regard to predicting behaviour from possibly short data sets, is can we use our meta-population time series \( \varphi_i \) for predicting behaviour in the future dynamics of the patches? This question can be addressed, assuming underlying deterministic dynamics, using a consequence of Taken's embedding theorem. Given our meta-time series, \( \varphi_i \), Taken's embedding theorem says that we will have some function \( F \) such that

\[
\varphi_{i+1} = F(\varphi_i, \varphi_{i-1}, \ldots, \varphi_{i-d+1})
\]  

(6.17)
where \( d_e \) is the dimension of the embedding. This asserts the predictability of the time series. However, since \( \mathcal{F} \) is not necessarily injective, it will not be invertible.

What we need to do is find a good approximation to the function \( \mathcal{F} \). This will allow us to predict the future behaviour of the time series because the next point of the time series is determined by the \( d_e \) previous ones. We now return to the notion of approximating \( \mathcal{F} \) by a linear combination of radial basis functions which were briefly discussed in section 6.3. The following ideas have appeared in various works [18, 22, 87] but the algorithm we use was presented in this form by Stark [99]. We look for some approximation \( \tilde{\mathcal{F}} \) to \( \mathcal{F} \) and, in a similar fashion to (6.12), we have

\[
\tilde{\mathcal{F}}(\varphi_i) = \sum_{j=1}^{R} \lambda_j B(|| \varphi_i - r_j ||),
\]

(6.18)

where again \( r_1, ..., r_R \) are the radial basis centres, \( \varphi_i = (\varphi_{i}, \varphi_{i-1}, ..., \varphi_{i-d_e+1}) \), \( B(x) = \sqrt{x^2 + c} \) for some suitable constant c. As before, we choose the \( \lambda_j \) to minimise the least squares error

\[
e^2 = \sum_{i=d_e}^{N} (\tilde{\mathcal{F}}(\varphi_i) - \mathcal{F}(\varphi_i))^2,
\]

(6.19)

where \( N \) is the number of data points in the time series and we begin with \( i = d_e \) as delay coordinates require the last \( d_e \) points in the time series to construct the next point in the reconstructed phase-space. From (6.17), \( \mathcal{F}(\varphi_i) = \varphi_{i+1} \) so combining (6.18) and (6.19) we obtain

\[
e^2 = \sum_{i=d_e}^{N} \left( \sum_{j=1}^{R} \lambda_j B(|| \varphi_i - r_j ||) - \varphi_{i+1} \right)^2,
\]

(6.20)

and upon defining the matrix \( A_{ij} = B(|| \varphi_i - r_j ||) \) and the vector \( \mathbf{v} \), where \( v_i = \varphi_{i+1} \),
we have

\[ e^2 = \sum_{i=1}^{N} \left( \sum_{j=1}^{R} A_{ij} \lambda_j - v_i \right)^2 \]

(6.21)

\[ = \| A \lambda - v \|^2 . \]

So minimising \( e^2 \) boils down to finding the \( \lambda_j \) which minimise the Euclidian norm of the vector \( A \lambda - v \). This is a standard linear least squares problem and the most efficient method is generally considered [18] to be the orthogonal decomposition of \( A \). Consider the following reformulation of \( A \) and \( v \)

\[ A = QR, \]

\[ v = Q \bar{r} + \bar{q}, \]

(6.22)

where if \( Q = [q_1, ..., q_m] \) then \( q_i^T q_j = 0 \) if \( i \neq j \) and \( q_i^T \bar{q} = 0 \) for \( i = 1, ..., m \). The matrix \( R \) is upper triangular with unit diagonal elements so \( R_{ij} = 0 \) for \( i > j \) and \( R_{ii} = 1 \) for all \( i \). Now using the above formulation we see that

\[ A \lambda - v = QR \lambda - Q \bar{r} - \bar{q}, \]

(6.23)

and consequently

\[ \| A \lambda - v \|^2 = \| \bar{q} \|^2 + \| Q(R \lambda - \bar{r}) \|^2 - 2 \bar{q}^T Q(R \lambda - \bar{r}). \]

(6.24)

However, we know that \( \bar{q} \) is orthogonal to the columns of \( Q \) hence \( \bar{q}^T Q = 0 \) and finally we have that

\[ \| A \lambda - v \|^2 = \| \bar{q} \|^2 + \| Q(R \lambda - \bar{r}) \|^2 . \]

(6.25)
So, the \( \lambda_i \) required to minimise (6.25) can be obtained as a solution of \( Q(R\lambda - \mathbf{r}) \).

Since \( R \) is upper triangular we can find the \( \lambda_j \) very easily by back substitution so

\[
\lambda_i = \bar{r}_i - \sum_{j=i+1}^{R} R_{ij} \lambda_j \tag{6.26}
\]

inductively for \( j = R - 1, \ldots, 1 \). Many methods exist for computing the vectors \( \mathbf{q}_i \) such as Givens or Householder transformations but here we use the modified Gram-Schmidt algorithm [56] which is numerically better conditioned. The basic algorithm defines a set of auxiliary vectors \( \mathbf{q}_j^i \) for \( j \geq i \) by the notation

\[
\mathbf{q}_j^{i+1} = \mathbf{q}_j^i - R_{ij} \mathbf{q}_i,
\]

where \( \mathbf{q}_i = \mathbf{q}_i^i \) and \( \mathbf{q}_j^1 = \mathbf{a}_j \), where \( \mathbf{A} = \{ \mathbf{a}_1, \ldots, \mathbf{a}_R \} \), so inductively,

\[
R_{ij} = \frac{\mathbf{q}_j^i \mathbf{q}_i^T}{\mathbf{q}_i^T \mathbf{q}_i},
\]

with the convention that \( R_{ij} = 0 \) if \( \mathbf{q}_i = 0 \). We similarly define the vectors \( \bar{\mathbf{q}}^i \) for \( i = 1, \ldots, R + 1 \) with \( \bar{\mathbf{q}}^1 = \mathbf{v} \) and

\[
\bar{\mathbf{q}}^{i+1} = \bar{\mathbf{q}}^i - \bar{r}_i \mathbf{q}_i,
\]

for \( i = 1, \ldots, R \) with \( \bar{\mathbf{q}}^{R+1} = \bar{\mathbf{q}} \) and

\[
\bar{r}_i = \frac{\bar{\mathbf{q}}^{i+1} \mathbf{q}_i}{\mathbf{q}_i^T \mathbf{q}_i},
\]

Now we simply set \( \lambda_R = \bar{r}_R \) and use (6.26) to compute \( \lambda_{R-1}, \ldots, \lambda_1 \) via back-substitution. The subsequent vector \( \lambda \) gives us the optimal approximation \( \tilde{\mathbf{F}} \) of \( \mathbf{F} \), subject to least squares fitting.

This is a global batch method but since it is extremely unlikely that we will have any real time data, there seems little need in applying a recursive form of the above
procedure. However, a good description of such a method can be found in [99].

Again we have the problem of how to choose the radial basis centres. The usual choice is from data in the time series itself. By this we mean that, for some suitable indexing function \( \eta(j) \), we require that \( r_j = \Phi_{\eta(j)} \), where we choose \( \eta(j) \) such that it gives us a suitable spacing between the basis centres. We can now express the elements of \( A \) as

\[
A_{ij} = B(w_{\eta(j)}^i), \tag{6.31}
\]

where

\[
w_{\eta(j)}^i = \left( \sum_{k=0}^{d-1} (\Phi_{\eta(j)-k} - \Phi_{\eta(j)-k})^2 \right)^{\frac{1}{2}}. \tag{6.32}
\]

In figure (6.9) we plot a chaotic time series, \( \bar{\varphi}_i \) (the one used to produce the right hand side of figure (6.1)), and the approximation \( \psi_i = \bar{F}(\bar{\varphi}_{i-d_e}, \ldots, \bar{\varphi}_{i-1}) \). The first 1000 points were used as the data sample and the first 150 points were used to define the radial basis centres with \( B(x) = \sqrt{x^2 + 2} \) as this value for the constant \( c \) gave the best fit (the embedding dimension will be \( d_e = 7 \) throughout, in line with Taken's \( d_e \geq 2m + 1 \) condition). The choice of the constant \( c \) is problem specific and highly related to the amplitude of the systems oscillations. For example, the prediction of a time series obtained from the chaotic Henon system in Stark [99] used \( c = 10 \), where points \((x_n, y_n)\) in the Henon attractor are contained in the region \((\pm 1.5, \pm 0.4)\). In our reconstructed NPZ system, time series orbits are in the interval \((0, 0.4)\), prompting our smaller choice of the constant \( c \).

The two time series are virtually indistinguishable, figure (6.10) shows the error at each iteration. The average error is around \( 10^{-4} \) for both the first 1000 iterates.
Figure 6.9: Top figure shows the true time series $\tilde{\phi}$, while the bottom shows the approximation $\psi = \tilde{F}(\tilde{\phi})$. The time series was constructed by sampling the phytoplankton population of a single NPZ model with $d = 0.142$.

(in sample) and the second 1000 iterates (out of sample) suggesting that over-fitting has not occurred. By this we mean that the approximation works just as well for the sample used in its construction as it does for the next sample of points in the time series. Later we pay attention to the effect of varying the length of the time series, $N$, and the number of basis functions, $R$.

So now we can construct an accurate estimation of the underlying dynamics, can we then predict forward in time? The answer is yes but the question of just how far forward we can predict depends on the nature of the system itself.

If it is periodic/quasi-periodic then a good approximation $\mathcal{F}^j$ will stay close
Figure 6.10: Fitting error $e_i = \varphi_i - \psi_i$ for the two time series in figure (6.9).

to $\mathcal{F}$ for a while before small errors gradually cause the approximation and the actual time series to become out of phase (see figure (6.11) for the case where the closure rate $d = 0.14$). For the case of periodic systems this shouldn't cause concern because, provided the original time series contains a full cycle, we can simply paste many cycles together. The situation is more difficult if the underlying dynamics are chaotic though. One of the hallmarks of a chaotic system is that small errors (i.e. ones resulting from the approximation) expand exponentially with time [25] until they are of the order of the size of the system in question.

Put more formally, if the system is indeed chaotic (with some associated maximal Lyapunov exponent [25], $\kappa > 0$) then all errors of size $\delta \ll 1$ will grow to an order of $\delta e^{\kappa n}$ after $n$ iterations. Let $\rho$ be some measure of the order of magnitude of the system, such as the mean population, then a simple calculation reveals that if

$$n \approx \frac{\ln (\rho/\delta)}{\kappa}$$

(6.33)

then the error will be of the order magnitude of the system and prediction becomes
Figure 6.11: Gradual phase separation of the approximation (dashed line) and the true time series (solid line) for the periodic case. The time series was constructed by sampling the phytoplankton population of a single NPZ model with $d = 0.14$.

futile ($n$ is often referred to as the Lyapunov time).

Figure (6.12) shows how such a separation occurs in a chaotic time series. For the first 1000 iterates we plot $e_i = \hat{\varphi}_i - \psi_i$ (where $\psi_{i+1} = \hat{F}(\varphi_i)$). From $i = 1000$ we define $\hat{\psi}_i$ as $\hat{\psi}_i = \hat{F}(\psi_{i-d}, ..., \psi_{i-1})$ and plot $\hat{e}_i = \hat{\varphi}_i - \hat{\psi}_i$. Clearly, after around 100 iterations the error in prediction is already of the order of the amplitude of the original time series, see figure (6.9), so predictions beyond that point would be pointless. So, for chaotic time series, short term predictions are all we can achieve but, as mentioned in [47], a short term prediction of a chaotic time series will be more accurate than a short term prediction of an inherently stochastic system (such as the ARMA system). This is because the uncertainty in the chaotic system develops in time, whereas in a stochastic system there is some level of uncertainty right from the start.

With regard to randomly related patches in section 6.3.1, we briefly mentioned
Figure 6.12: *Propagation of small errors when attempting to predict chaotic time series.* At \( i = 1000 \) we begin to predict future points in the time series generated again by sampling the phytoplankton population of a single NPZ model with \( d = 0.142 \).

The ARMA model which requires a historical data set to make an approximation to the dynamics. The more data that are available, the better the fit. The same is true when using methods such as radial basis function approximation but the other factor here is that the accuracy is also linked to the number of radial basis centres that are defined. The more centres we have, the more accurate the fit [18, 22, 99].

In Table 6.1, we demonstrate how the cumulative error of prediction \( e = \sum_{i=1}^{T} |e_i| \) \((e_i = \varphi_i - \psi_i \text{ and } T = 1000)\) varies with both sample length \( N \) and the number of basis centres specified, \( R \). The time series used in the simulations came from a single NPZ system in the chaotic regime. As is to be expected, the case with the longest sample length and the most basis centres is the most accurate. What general trends are observable? Firstly, the dependence on the sample length is non-linear as we see a marked decrease in the accuracy for successive equal decrements in the
Table 6.1: Cumulative squared error when varying the number of data points and
the number of basis centres.

<table>
<thead>
<tr>
<th></th>
<th>N=500</th>
<th>N=250</th>
<th>N=125</th>
<th>N=50</th>
</tr>
</thead>
<tbody>
<tr>
<td>R=100</td>
<td>0.0561</td>
<td>0.0563</td>
<td>0.0942</td>
<td>0.3803</td>
</tr>
<tr>
<td>R=75</td>
<td>0.0737</td>
<td>0.0754</td>
<td>0.1031</td>
<td>0.4701</td>
</tr>
<tr>
<td>R=50</td>
<td>0.1955</td>
<td>0.2153</td>
<td>0.2918</td>
<td>0.9072</td>
</tr>
<tr>
<td>R=25</td>
<td>0.217</td>
<td>0.2361</td>
<td>0.3014</td>
<td>1.1036</td>
</tr>
</tbody>
</table>

sample length, whatever the number of basis centres. Also, the relationship between
accuracy and the number of basis centres seems non-linear as well. There is a small
decrease in accuracy when decreasing the number of basis centres from 100 to 75
and from 50 to 25 but a large one when decreasing $R$ from 75 to 50. This is seen for
all lengths of the data set $N$ suggesting the possibility of a sigmoidal relationship
between $R$ and $e$.

In general, the radial basis approach is quite suited to dealing with possibly short
data sets as this lack of data can, to some degree, be balanced out by specifying
more basis functions. This cannot be said for the ARMA model as it utilises only
the available data. Also, more sophisticated techniques are available [22, 87, 99]
for choosing the radial basis centres whereby initially many are selected and then a
model selection algorithm deselects “unimportant” basis centres.
6.5 Applications to model fitting of planktonic dynamics

The fitting of specific population models to time series data has been an area of interest for decades, from detecting simple forms of population dynamics such as Lotka-Volterra [37] and logistic-type behaviour [30] to sex and age-structured, spatially resolved population models [90]. Such exercises can tell us much about the way the different species interact, for instance whether population density-dependence of various kinds is present [20, 46]. The gathering of accurate data over large spatio-temporal scales is not an easy task for many terrestrial/oceanic ecosystems which will inherently affect the accuracy of any subsequent fit made using such data. As mentioned in [46], inherent spatial (and temporal) heterogeneities mean that you can be sampling data from regions with differing dynamics.

Here, we shall demonstrate that the approach of first classifying the interacting dynamics of patches, then creating a meta-population time series accordingly, can increase the accuracy of fitting a specific model to a patchy ecosystem. By fitting the chosen model to this averaged data we might expect to get a better fit than simply using data sampled from just one patch, arithmetically averaged data or data randomly collected from the various patches, a problem touched upon in [46].

Given discrete measured data of the population of plankton $s_i = (n_i, p_i, z_i)$, for $i = 1, ..., N$, we want to fit this data to some specific model for the plankton populations, e.g. $S_i = (N_i, P_i, Z_i)$ and $S_{i+1} = G(S_i)$. If $\theta$ denotes the vector of parameters of the explicit model $G$, then to maximise the goodness of the fit
we must minimise the error between the data and the model variables. Put more formally, let the residual \( r_i = S_i - s_i \). We then calculate the \( \theta \) that minimise the least squares error, \( e^2(\theta) \), where

\[
e^2(\theta) = \sum_{i=1}^{N} r_i^T \mathbf{V}^{-1} r_i = \sum_{i=1}^{N} \left( \frac{(N_i - n_i)^2}{\sigma_n^2} + \frac{(P_i - p_i)^2}{\sigma_p^2} + \frac{(Z_i - z_i)^2}{\sigma_z^2} \right),
\]

where \( \mathbf{V} \) is the diagonal variance matrix [36] and this rescaling helps to avoid badly scaled errors for highly variable data as many descent type minimisation algorithms are sensitive to badly scaled variables [89]. This is because they require the calculation of the first and second derivatives of the function being minimised. If small parameter changes illicit large changes in the value of the function, the algorithm can move beyond the desired minimising solution. Conversely, for data with very low variability, this rescaling means that the algorithm doesn't locate a false numerical minima due to the "flat" nature of the function being minimised.

As an example case, we consider the data from a system of three coupled NPZ systems (6.2), coupled as shown in (6.8). We introduce a small random variation of \( \mathcal{O}(10^{-3}) \) in the parameter \( d \) and set the base value of \( d = 0.14 \) (a simple periodic solution). The subsequent patch dynamics are generally synchronised but, for the purposes of illustration, we consider the following four data sets to which we want to fit the single NPZ model:

1. The data set \( \mathcal{D}_1 \), constructed using the weighted linear averaging procedure defined in section 6.3.

2. The data set \( \mathcal{D}_2 \), constructed by taking data from each individual patch (three samples of equal intervals). Here we wish to mimic the effect of attempting to
"randomise" the location of sampling.

3. The data set $\mathcal{D}_3$, constructed by measuring the population of just one of the patches, ignoring the rest. This situation is similar to having a fixed measuring station.

4. The data set $\mathcal{D}_4$, constructed by taking the mean (over all patches) population of each patch species.

For the sake of argument, we now attempt to establish the value of the zooplankton grazing efficiency, $\alpha$ from (6.2), which gives the best fit (w.r.t. a least squares error) between the measured dynamics and the NPZ model. To do this we apply the same minimisation routine used in section 6.3 to the least squares error defined in (6.34) to compute this optimising value of $\alpha$. The actual value for $\alpha$ is $\alpha = 0.25$ and we use this as our initial guess. In figure (6.13) we plot the error $e_i = N_i - n_i$ after fitting the data sets $\mathcal{D}_1$ and $\mathcal{D}_2$.

The cumulative error for the fitting of $\mathcal{D}_1$, $e = \sum_{i=1}^{N} |e_i|$, was $e = 1.783$. The corresponding error for $\mathcal{D}_2$ was $e = 2.605$. In figure (6.14), we plot the error $e_i = N_i - n_i$ after fitting the data sets $\mathcal{D}_1$ and $\mathcal{D}_3$. The cumulative error for $\mathcal{D}_3$ was $e = 2.116$.

The final data set we considered, $\mathcal{D}_4$, was that obtained by simply taking the sample mean of the patches at each time-step. Figure (6.15) shows the error for $\mathcal{D}_1$ and $\mathcal{D}_4$. The cumulative error for $\mathcal{D}_4$ was $e = 3.126$.

In all cases, at least for the various constructed data sets considered here, fitting the model to the meta-population time series increased the accuracy of the fit with
Figure 6.13: The fitting error between the data sets $D_1$ (full line) and $D_2$ (dashed line). The optimal values of $\alpha$ were 0.250377 ($D_1$) and 0.250408 ($D_2$) respect to the other approaches considered above.

6.6 Discussion on the Effects of System Noise

One effect we have not yet discussed is that of noise. The main purpose of the methods we have presented is ultimately to use them on real experimental data, where noise will always be a factor. This work has been dealing with purely deterministic “clean” data as an example but in a real world situation the measured signal may be formed from a combination of noisy and deterministic parts, $s_i = \varphi_i + \zeta_i$, where $\varphi_i$ is our deterministic signal and $\zeta_i$ represents some unknown stochastic process. This stochastic element $\zeta_i$ generally will be a combination of both observation errors and process errors (errors from measurement and the inherent stochastic component
of the system). When looking at fitting different types of predator-prey models to measured data, [46] looked at differing fitting processes for different noise types with mixed results. The methods we have described will work best on a deterministic process. Hence, we would require the separation of the measured signal $s_i$ into its two constituents, $\varphi_i$ and $\zeta_i$.

Signal separation has been a hotly researched topic in signal processing for many years. One of the classical methods of noise reduction is that of Kalman filtering [36]. This procedure is normally applied to observed data with deterministic (linear) and stochastic (Gaussian) elements. It filters the noisy observation by utilising both autocorrelation in the noisy signal, and the variance-covariance matrix of the predicted deterministic and the observed noisy signal. This method was originally designed for use on linear systems but the generalised Kalman filter [36] uses the
Figure 6.15: The fitting error between the data sets $D_1$ (full line) and $D_4$ (dashed line). The optimal values of $\alpha$ were 0.250377 ($D_1$) and 0.250419 ($D_4$).

Linearisation of a non-linear operator in the same filtering algorithm. This type of approach can be unsuitable for non-linear systems, primarily because the linearisation estimate could be inaccurate, but especially for chaotic systems as, no matter how accurate the linear approximation, the linearised dynamics will exponentially expand small differences, on the average [25].

Perhaps the most popular method, in the case of a (non-linear) deterministic/stochastic mixed signal, is to use the power spectrum [25] of the observed signal. As an example we could consider the signal $s_i$ as some periodically varying deterministic signal $\varphi_i$ plus some Gaussian variate $\zeta_i$. If $s_i$ has power spectrum $E(\omega)$, for some frequency $\omega$, then we would expect the spectrum to have Dirac $\delta$'s at $\omega^*$ which is the frequency of oscillation of $\varphi_i$ (also at the harmonics $2\omega^*, 3\omega^*, \ldots$ etc.). Hence the peaked part of $E(\omega)$, due to $\varphi_i$, can be separated from the broadband
part due to the $\zeta$. Spectral methods fall down when we have a linear combination of chaotic and stochastic signals because both may have a broadband power spectrum and as such cannot be easily separated. There are many types of non-linear noise reduction algorithms which aim to separate chaotic and stochastic signals and a good review of the better ones can be found in Kantz & Schreiber [47]. Also, Kantz & Schreiber developed the time series package TISEAN, available free on-line, with inbuilt non-linear noise reduction algorithms and indeed other types of estimation packages which can deal with systems in delay coordinates.

### 6.7 Conclusions

We have presented a method for classifying interacting patch populations and predicting future trends using methods from deterministic time series analysis. Firstly we check that the patch dynamics are deterministically linked (generally synchronised) using a statistical test developed by Pecora et al. in [78]. We showed that the method is useful even with a relatively short time series, often a problem in physical/biological problems.

In section 6.3, we suggested some methods of constructing a meta-patch times series, if the patches are generally synchronised. While discussing various possibilities, we employed a weighted linear average method as, for the patch systems close to synchrony considered here, it gives agreeable results for very little computational expense and is easy to implement. It was seen that it also performed better than the popular method of taking the sample patch mean as the representative bulk
average population. However, we should point out that this improvement seen in the weighted linear average approach was not huge by any means.

In section 6.4.1 we addressed the issue of how to use the meta-time series to predict short term trends in the patch systems' dynamics. For this problem, we outlined a non-linear prediction algorithm using radial basis functions, first applied with regard to neural networks [18] but later adapted to helping predict chaotic time series [22, 87, 99]. Again, the method is relatively simple to apply and computationally inexpensive as the minimisation routine involves a modified Gram-Schmidt [56] method as opposed to more intensive Newton type descent algorithms. Also, it can deal with possibly short data sets, making it ideal for use in a biological context when plentiful data may be scarce.

In section (6.5), using a constructed example, we showed that the idea of creating such a representative patch time series could improve the quality of the fitting of real data to some specific reaction model. In both [20] and [46], mention was made that data sampling sometimes ignores any spatial or temporal heterogeneities and that this could naturally have a detrimental effect on the quality of model fits made with such data. The work presented here has tried to find ways of circumventing these problems, at least for the case where the interacting patches are not randomly related. Also, we have attempted to outline methods that can be applied to situations where there is relatively little data available or where there are gaps in the data.

The work contained in this chapter is an extension of the work in chapter 5, in that even though the dynamics of the system are unknown, we can still classify the
interactions and make better short term predictions for the dynamics of the patches, based on (possibly small amounts of) population data. Also, as in chapter 5, the ideas presented here are applicable to any type of patchy ecosystem.
Conclusions & Further Work

In the work contained in this thesis, the prime motivations have been to investigate how both the observed patchy nature of many planktonic populations and, their subsequent transport via the ocean, affect the dynamics of the plankton.

In chapter 2, a reduced, time-perturbed model for Langmuir [55] circulations was used to look at the possible role of chaotic advection in the generation of patchiness. Also, models of chaotic advection have been shown to be both simpler to implement than a Navier-Stokes based approach [2, 5, 71] and useful for application to specific systems or geometries [12, 33, 76, 105]. It was shown that certain types of chaotic models can exhibit regions of aggregation and regions of high mixing, which can cause variation in both the patchiness of the system and the transport component of the predator-prey contact rates. In the example presented, it is even possible to analytically investigate the mixing occurring in the system and use these expressions to predict, for what relevant environmental/biological scales, swimming (diffusion) or advection will dominate. This type of analysis is not possible for all models
of chaotic advection, but other methods from dynamical systems theory could be applied to study factors such as mixing and contact rates. Future work could be done by investigating the chaotic advection of planktonic communities in a more general framework.

In Chapter 3, the possible interaction between physical/swimming characteristics and simple flow fields, with regard to possible pattern formation, was investigated. A general non-linear reaction-advection-diffusion equation was considered, with advection being modelled as a simple linear shear flow. Swimming behaviour and shape characteristics can significantly alter the diffusive behaviour of motile organisms, even under the advection of simple flow fields [11, 42, 81]. Coupled with non-linear reaction terms this has long been known to give rise to spatio-temporal pattern formation. This type of system has been investigated before, using specific reaction models exhibiting behaviour such as activator-inhibitor [60] properties and excitability [16].

For the general reaction model case, a linear stability analysis revealed that if the swimming and shape characteristics are the same (equivalent diffusion tensors) then, due to the fact that the diffusion tensors may be quite complicated, the shear can exert a destabilising influence on the system. This result is contrary to that of Doering & Horsthemke [24], who considered simple Fickian diffusion. The non-diagonal terms in the diffusion tensor are the mechanism behind the possible instability and the conditions under which such a bifurcation could be observed were discussed in relation to recent work on the diffusive behaviour of swimming organisms in shear flows, [11, 41].
To investigate what might happen when we allow the organisms' characteristics to differ, a weakly non-linear analysis was performed for a two species system with different diffusion tensors. For simplicity, the analysis was performed about the no-flow, zero wavenumber instability. Similar to the results found in the paper by Spiegel & Zaleski [96], the resulting amplitude equation was not necessarily well-posed, for certain diffusive regimes, suggesting that zero is not the critical wavenumber in these cases. Considering a more general non-linear term than that in [96], the criticality of the instability was defined in terms of the non-linear coefficients appearing in the amplitude equation. The main drawbacks of the analysis was the need to investigate the no-flow, zero wavenumber instability, primarily because the non-zero wavenumber equations become too complex to deal with. This assumption of an instability for small shear rates is probably unrealistic as the differences required in the diffusion coefficients is usually too high, or the time-scale needed to see the instability develop too long. For stronger flow fields, the differences in the diffusive behaviour needed to see instability can decrease quite considerably [95]. Further work could concentrate on using a specific reaction model to see how these biological factors such as swimming speed and shape can aid in planktonic pattern generation, even for simple flow fields.

The work on chapters 2 & 3 was primarily to demonstrate that the advection, from simple shear flows to chaotic flow fields, of plankton and the various characteristics of the plankton, primarily motility and shape, can aid (both individually and in interaction with each other) the patchiness of planktonic communities.

Phytoplankton blooms [57] consist of localised, and/or short lived dramatic in-
creases in the population of phytoplankton and can be beneficial to shaping the year-class strength of a particular fish stock but also poisonous (Red tides) to others [57]. In chapter 4, an attempt was made to incorporate the ideas discussed in chapters 2 & 3; namely advection by complex (possibly turbulent) flows and the interaction of biological factors such as motile behaviour and density variations with the imposed flow. Specifically, we investigated how these complex factors interact and whether they can influence the formation of phytoplankton blooms. Truscott & Brindley [107] proposed a two species Phytoplankton-Zooplankton (PZ) model, which exhibited excitable phytoplankton dynamics, as a simple model for phytoplankton blooms. By incorporating this with a model for synthetic turbulence first presented by Marti et al. [62], the effects of planktonic density variations were considered. It has been observed [91, 97] that inertial particles tend to have preferred regions of aggregation when immersed in turbulent flow and planktonic organisms are known to exhibit varying degrees of density differences [31, 65], with regard to the surrounding ocean. After performing a variety of simulations, for various inertial regimes, it was observed that, for a physically realistic inertial regime, preferred aggregation zones (due to density characteristics) initiated a small scale bloom in the phytoplankton population. This was when we began the simulations with the planktonic populations at their equilibrium levels so, even though inertia alone could not initiate a full scale bloom from a homogeneous state, it means that the perturbing effects of processes such as nutrient up-welling [57, 71], known to help cause blooms, can be aided by the inertial characteristics of the plankton themselves.

A continuum expression for the phytoplankton-zooplankton contact rate, based
on the relative fluxes of the predator and prey, was derived to see how the inertial bloom altered the predator-prey contact rates. Conceptually, inertial separation should lower the contact rate, as predator and prey are being pulled apart to some degree. However, the separation initiates the bloom, which subsequently spreads into areas of higher predator densities, thus increasing the contact rate.

The energy spectrum of turbulent flow has been studied quite extensively [53, 54, 70] in both two and three spatial dimensions. From Kolmogorov’s [51] $k^{-5/3}$ inertial sub-range power law to Kraichnan’s [53] $k^{-3}$ power law for 2D turbulence, the mechanisms of energy decay and formation are well studied. To see how these properties altered for inertial particles (as opposed to passive scalars), we used a wavenumber dependent diffusivity, as used in Powell & Okubo [88], as an approximation to the non-linear term seen in the Navier-Stokes equations. Coupled with the experimentally derived horizontal diffusivity-lengthscale relationship from Okubo [73], it was shown that the energy spectrum changed for light particles (those forced into the eddies) so as to increase the mean squared velocity (intensity) of the corrected inertial flow. The opposite occurred for heavy particles (those forced into the straining regions). This tallies quite well with the theoretical work on 2D turbulence [54] which showed that the turbulent energy is created inside the eddies, characterised by the enstrophy (mean square vorticity). Since lighter particles/organisms aggregate in regions of high vorticity [91], it is logical to expect that they would have a higher intensity than those aggregating in regions of low vorticity. Numerical plots of the corrected energy spectra also showed the inertial influence to be most important around the length scale of the large scale energy generating eddies.
Possible areas for further study of the ideas presented in chapter 4 are a spectral treatment of the simulated blooms. Powell & Okubo [88] examined how simple population dynamics could alter the energy spectra of organisms immersed in turbulent flow. It was seen that the energy spectra altered differently for 2D and 3D turbulence, even for simple population dynamics. The flow implemented here is ideal for numerical study of the bloom spectra as the flow’s energy spectrum is something defined by the user, so deviations due to populations dynamical events could be directly studied as the underlying energy spectrum of the original flow is known. Also, some numerical investigation of how the turbulent spectrum alters with inertial corrections would be useful, to see if the analytical predictions in section 4.1.4 are correct.

The model for bloom dynamics is quite specific in that it is a continuous time model and the bloom is subject to “top down” control by the zooplankton. By “top down” we mean that the formation of the bloom is controlled by the trophic group above the phytoplankton in the food chain. Logically, the other classification is termed “bottom up” as it is the trophic group below the species in question that controls the specific phenomenon. Such “top down” bloom control is not always the case as poisonous blooms such as red-tides [57] are subject to “bottom up” control by the nutrient population as they are not grazed upon by zooplankton. Huppert et al. [45] presented a model for phytoplankton blooms, based upon control by the nutrient population. Also, nutrient up-welling is a well documented [57] driver of blooms, based on field observations and simulations, so further work could be in modelling the influence of up-welling of nutrient and perhaps the competing effects
of top down/bottom up control by zooplankton and nutrient, respectively, in three component models of planktonic blooms.

In chapter 5, the patchy dynamics of planktonic organisms was treated as a coupled lattice of plankton patches. We spatially discretised the problem and treat the resulting system as a lattice of coupled oscillators. This allowed more scope in the parameters that can be varied and also in terms of both theoretical and numerical analysis as much recent work has concentrated on the generic properties and analysis of coupled oscillators [1, 3, 6, 50]. Such systems exhibit various forms of synchronised [34, 77, 82] behaviour and in chapter 5 it was shown that spatio-temporally varying planktonic dynamics could result from suitably low coupling or even small levels of system noise. The first pathway to unsynchronised patch dynamics was firstly demonstrated numerically using a simple two-patch system. Then, for a 1D lattice of arbitrary length, it was conjectured and numerically demonstrated that, for suitably low coupling in just one patch, the whole system can exhibit unsynchronised motion. The idea of spatial variations in the plankton dynamical models has been investigated, in a continuum sense, by Malchow et al. [61] and Medvinsky et al. [66] where fish school motion was modelled as variations in the higher predation (closure) term in the zooplankton dynamics. These papers demonstrated that such variations could cause the appearance of spatio-temporally varying (even chaotic) dynamics for a system perturbed from equilibrium.

Spatial variations in the dynamics governing each patch is to be expected. Again using invariant manifold theory [110], it was shown how we can compute the strength of coupling needed to give rise to a smooth relationship between the various patch
dynamics, at least for small differences in the dynamics. The computational evidence suggested a curious role for chaos in such systems. For two NPZ models with different functional forms of the closure term, it was seen that the coupling needed to guarantee smooth inter-patch dynamics was lower, on the whole, around the chaotic regime than anywhere else. Previously, it has been suggested by Allen et al. [4] that chaos aids the persistence of coexistence dynamics, even in relatively unfavourable conditions, due to the amplification in local population “noise”. We suggest that a similar process may be at work here and that the chaotic dynamics help stop the individual patch dynamics locking on to their particular isolated patch dynamics. This could consequently allow the patch dynamics to settle into a stronger form of synchronous behaviour for lower coupling strength.

Spatially coherent oscillatory behaviour has been seen in terrestrial [17] and oceanic [32] ecosystems where the coupling may not arise from direct inter-patch population interaction but could be a result of external environmental influences such as temperature or precipitation. By reversing an idea on detecting generalised synchronisation in so-called drive-response systems by Abarbanel et al. [1], we demonstrated the non-linear mechanism behind the forced synchrony of patchy populations by external forcing factors and gave a brief numerical example of the process using a rescaled chaotic Rössler system to drive a two patch array.

The treatment of patchy dynamics as a system of coupled oscillators is possibly simplistic but has the advantages of being applicable to a wide range of scenarios and is easier to analyse than a similar continuum system. Also, the ideas and mechanisms presented here apply not only to planktonic dynamics but to any patchy ecosystem;
terrestrial or oceanic.

In the final chapter, the aim was to extend the ideas and concepts of chapter 5 to the situation where all that is available is measured data of possibly one of the patch inhabitants and the dynamics governing the system are unknown. It was shown how the dynamics of the patch system can be reconstructed, using delay coordinates [47, 106], from just one of the patch inhabitants. Then, we presented a relatively simple algorithm, developed by Pecora et al. [78], to detect if any deterministic relationship exists between the patches. Particular attention was given to the applicability of such an algorithm for short, possibly incomplete data sets, often a problem in many real world situations.

If such a relationship was detected, a method was then presented for constructing a meta-population, representing the whole patch, which attempts to retain much of the original dynamics. Using examples of expected types of "patchy" data sets (arising from factors such as fixed measurement stations, random spatial sampling etc.), it was shown that the relatively simple linear weighted average reduced the average patch error more than the arithmetic mean of the patch populations, a more popular method. We also attempted to use a non-linear estimation procedure, involving the use of radial basis functions [18, 22, 87, 99], but these were found to offer only comparable results to the weighted linear approach but were computationally much more intensive. Arguably, for systems very close to synchronisation there is no real difference between the two linear methods. However, for systems in phase synchronisation, there can be large differences in the amplitudes of the patch dynamics and this is the type of effect the weighted linear average is better at capturing because
it assigns specific weights to each patch population, as opposed to assigning each patch the same weighting.

Given this representative patch time series, we considered how to use this (possibly short) meta-population time series for making short term predictions in the planktonic dynamical trends. Since we are generally assuming underlying deterministic (possibly chaotic) dynamics and a deterministic relationship between patches, a non-linear prediction method is employed that was developed in various forms in [18, 22, 87, 99]. The predictive capabilities of the algorithm were displayed, even for chaotic time series data. Also, attention was paid to how the algorithm performs for possibly short or incomplete data sets. It was shown that, unlike other predictive methods such as ARMA models [21, 47], small data sets could be, to a certain degree, compensated for. This was because the non-linear predictor algorithm we employed uses a specified number of basis functions, which can be increased to increase the accuracy of the prediction if the number of available data points is low. This property also means that we can use the predictor to fill in any gaps in the time series, under certain specified conditions.

The second possible use for such a time series is for fitting a particular model for the population dynamics to the observed data. This has been a field of interest in planktonic dynamics recently [20, 46], with regard to identifying density dependence (non-linearities) in the planktonic dynamics. It also forms the backbone of most analyses and stock assessments of fish populations [90]. The patchy structure of the underlying population under consideration is not always accounted for, as mentioned in Jost & Arditi [46]. The construction of the averaged time series goes some way
to incorporating the patchiness and we demonstrated, for four constructed “patchy”
data sets, how the algorithm can improve the quality of the fit. The four data sets
we considered were:

- Data sampled from just one patch, to mimic the effect of a fixed measurement
station.
- Data sampled at equal intervals from the different patches, akin to randomly
sampling from different sites.
- Data constructed by simply arithmetically averaging the patch populations.
- Data constructed from generally synchronised patches using the linear weighted
average method described in chapter 6.

The data set constructed from applying our weighted linear pathc average per-
formed best in terms of the goodness of the fit between the data and the NPZ model
it was fitted to. Again, the gain in accuracy versus the extra work is a payoff that
the analyst must decide is worthwhile. What we hope is that possible effects of some
of the problems that can arise from sampling patchy ecosystems have been outlined
and how, for certain cases, they can be surmounted.

Patchy dynamics will not always be deterministic in nature and consequently the
methods presented here will not be applicable. An increasing amount of research
is being done into using more statistical methods such as maximum likelihood esti-
mators [39] and Bayesian decision theory [90] for highly stochastic populations with
possibly a patchy/migratory structure.
The work in this thesis has been to investigate the complex interaction of the physical and biological influences on the dynamics of planktonic systems. However, the treatment of patchy dynamics is of vital importance to many ecosystems and it is hoped that some of the work presented here will also aid the understanding of these wider issues.

The work from chapter 4 on excitable inertial plankton in turbulence [92] has been published in the Proceedings of the Royal Society of London Series B: Biological sciences under the title *Plankton blooms induced by turbulent flows*. Two papers are currently in preparation; one for submission to Physical Review E [43] called *Plankton lattices and the chaotic suppression of patchiness* and the other [42] entitled *Chaos & synchrony in patchy ecosystems*, for submission to the Bulletin of Mathematical Biology, based on the work from chapter 5.
In this Appendix we give a more detailed explanation of the technical terms, definitions and the numerical methods that were used in this thesis. The first section is all about the theory of smooth dynamical systems and comprises of a condensed review of some of the topics in [25, 28, 109, 110].

8.1 Ergodic theory of smooth dynamical systems

The analysis of smooth dynamical systems, especially chaotic ones [25], has moved into a more probabilistic field than the more classical dynamical systems theory. In this section we briefly discuss basic measure theory and how this leads to the notion of ergodic measures and the probabilistic analysis of complex dynamical systems.
8.1.1 Basic Measure Theory

We call $\mu$ a *measure* on $\mathbb{R}^n$ if it assigns a non-negative value to every measurable (Borel) set in $\mathbb{R}^n$ and:

1. $\mu(\emptyset) = 0$,

2. $\mu(A) \leq \mu(B)$ if $A \subseteq B$.

3. If $A_1, \ldots, A_n$ is a countable (or finite) sequence of measurable sets then

   $$\mu \left( \bigcup_{i=1}^{\infty} A_i \right) \leq \sum_{i=1}^{\infty} \mu(A_i),$$

or if $A_i \cap A_j = \emptyset$ for all $i \neq j$ then

   $$\mu \left( \bigcup_{i=1}^{\infty} A_i \right) = \sum_{i=1}^{\infty} \mu(A_i).$$

We call $\mu(A)$ the *measure* of the set $A$. For example, the $n$-dimensional Lebesgue measure is defined as follows. If we define $S = \{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid x_i \in [r_i, s_i]\}$ as a 'coordinate parallelepiped' in $\mathbb{R}^n$, the $n$-dimensional volume of $S$ is given by

$$\text{vol}^n(S) = (r_1 - s_1)(r_2 - s_2) \cdots (r_n - s_n).$$

We define the *$n$-dimensional Lebesgue measure*, $\mathcal{L}^n$, of a set $S$ to be,

$$\mathcal{L}^n(S) = \inf \left\{ \sum_{i=1}^{\infty} \text{vol}^n(S_i) \mid S \subseteq \bigcup_{i=1}^{\infty} S_i \right\},$$

where the $S_i$ cover the set $S$. Lebesgue measure is intuitively linked to the concepts of area ($n=2$) and volume ($n=3$).

A probability measure $\rho$, defined on some space $\Omega$, is a measure such that

1. $\rho(\Omega)=1$, 

195
2. \( p(A) \leq 1 \) if \( A \subseteq \Omega \).

Any measure \( \mu \), defined on some set \( M \) can be transformed into a probability measure, \( \rho \) on \( M \), where

\[
\rho(A) = \frac{\mu(A)}{\mu(M)},
\]

for any measurable \( A \subseteq M \).

A probability measure \( \rho \) is said to be invariant, under some transformation \( G \), if

\[
\rho(A) = \rho(G^{-t}(A)) \quad \forall t \in \mathbb{R}^+,
\]

and \( A \subseteq \mathbb{R}^n \).

For dynamical systems which model physical systems/processes, we would like to concentrate on physical measures (ones which we assume represent some experimental time average). The main requirement for the numerical approximations of the properties of such measures is that the measure is invariant, (8.1), under time evolution.

When moving to ergodic theory, we consider the long time behaviour of the system; we ignore transients, and concentrate on attracting structures. If we have some attracting set \( A \), of the dynamical system governed by \( G^t \) we have the following definition. If \( A \) supports some probability measure \( \rho \), then \( \rho \) is said to be ergodic if \( \rho(A) = 1 \). This equates to the notion of \( A \) being indecomposable. The orbits of points in the attractor (up to a set of zero measure) explore the whole attractor.

The ergodic theorem states that, for every continuous function \( \varphi \),

\[
\lim_{T \to \infty} \frac{1}{T} \int_0^T \varphi(x(t)) \, dt = \int \varphi(x) \, d\rho(x),
\]
for almost all initial conditions, \( x(0) \).

### 8.1.2 Liapunov Exponents

The eigenvalues of the Jacobian matrix determine the nature of a fixed point, as do the Floquet exponents of a limit cycle. Once we have some attractor, how can we assess the long term behaviour of initial conditions on this attractor? If we Taylor expand to first order about some arbitrary initial state, \( x(0) \), we see that

\[
G(x(t) + \delta) - G(x(t)) = \delta x(t) = DG(x(t))\delta,
\]

where \( DG_{ij} = \partial G_i / \partial x_j \). Assuming that \( \delta x(t) = \delta e^{\lambda t} \) (for some \( \lambda \)), we have that

\[
\lambda = \lim_{T \to \infty} \frac{1}{T} \int_0^T \log \| DG^t(x(0))\delta x(0) \| \, dt.
\]

The constant \( \lambda \) is called the Liapunov exponent of the point \( x(0) \) in the direction \( \delta x(0) \). Furthermore, the ergodic theorem tells us that the time average defined in (8.1) is equal almost everywhere (a.e.) to the space average and

\[
\lambda = \int \log \| DG(x) \delta x \| \, d\rho(x).
\]

The multiplicative ergodic theorem of Oseledec [74] goes even further. If we have some compact invariant set \( M \) and there are real numbers \( \lambda_1 > \lambda_2 > \ldots > \lambda_n \) and a splitting of the tangent space at \( x \in M, T_x M = E_1 \oplus E_2 \oplus \cdots \oplus E_n \) (\( E_1 \supset E_2 \supset \cdots \supset E_n \)) such that

\[
\lambda_j = \lim_{T \to \infty} \frac{1}{T} \int_0^T \log \| DG^t(u) \| \, dt = \lambda_j, \quad \forall u \in E_j \setminus E_{j+1},
\]

then \( x \) is said to be a regular point and, furthermore, the set of regular points of \( G \), \( \mathcal{R}(G) \), satisfies \( \rho(\mathcal{R}(G)) = 1 \). This holds for all \( G \)-invariant probability measures \( \rho \).
supported in $M$. The $\lambda_1, ..., \lambda_n$ are the Liapunov exponents and the theorem tells us that the time average above converges to a finite, constant set of values (up to a set of zero measure). In simpler terms, we can compute the time average of one initial condition and this will represent the behaviour averaged over the whole attractor.

A different definition of the Liapunov exponent is given by first noting, for continuous systems, that the linearised dynamics for small perturbations, $\delta x(t)$, are governed by the following equations:

$$\delta x(t) = A_0 \delta x(0),$$
$$\frac{dA_t}{dt} = D G^t A_t.$$

If we define the eigenvalues of the following limiting matrix

$$\lim_{t \to \infty} (A^T_t A_t)^{1/2},$$

to be called $\alpha_j$ then the Liapunov exponents, $\lambda_j$, are defined by the relationship $\lambda_j = \log(\alpha_j)$. The $\alpha_j$ will be real and positive because $A^T_t A_t$ is a positive definite matrix.

### 8.1.3 Practical computation of Liapunov exponents

Given the various definitions for Liapunov exponents, what is the best way to go about actually calculating them? In practice, calculating the eigenvalues of the limiting matrix defined in (8.2) is very difficult so an alternative method(s) would be useful. One of the most well used ideas is to analyse the rates of growth/reduction in phase space volume. First, we assume we have the existence of $n$ orthonormal vectors $v^1_0, ..., v^n_0$ (with phase space volume $\prod_{i} \|v^i_0\|$). Under the linearised dynamics, at
some time $t > 0$,

$$v_i^j = A_t v_i^0.$$

Now, the component of $v_j^2$, orthogonal to $v_i^1$, grows like $e^{\lambda_2 t}$. In general, the component of $v_j^i$, orthogonal to $v_i^j$, $\forall j = i - 1, ..., 1$, grows like $e^{\lambda_i t}$. So, with this orthonormal basis, we now have that

$$\lambda_i = \lim_{t \to \infty} \frac{1}{t} \log \| v_i^t \|.$$

However, this can lead to numerical overflow if $\lambda_i > 0$ or the distance between $v_i^t$ and $v_i^{t+1}$ becomes so small it cannot be computed accurately.

### 8.1.4 Reorthogonalisation

To circumvent both these practical problems with the theoretical definitions, the linearity of the problem can be exploited. Given some $s, t > 0$

$$v_{s+t}^i = A_{s+t} v_i^0 = A_t (A_s v_i^0) = A_s (A_t v_i^0).$$

If we now fix $s$ to be constant (usually problem specific), we denote by $w_{t+s}^i$ the orthonormal component of $A_s v_i^j$, w.r.t. $A_s v_i^j$, $\forall j = i - 1, ..., 1$ (these can be calculated via a simple Gram-Schmidt procedure). In essence, we go forward for time $s$ and then reorthogonalise and renormalise to avoid the numerical complications mentioned. This now yields

$$\lambda_i = \lim_{t \to \infty} \frac{1}{t_s} \log \| w_{t+s}^i \|,$$

which is a readily implementable numerical procedure for calculating Liapunov exponents.
8.2 Numerical integration methods

In the thesis we numerically integrate a variety of ODE's and PDE's (both deterministic and stochastic) and in this section we give a brief overview of the numerical schemes employed.

8.2.1 ODE integration

For the integration of the various ODE's used in this thesis, two types of Runge-Kutta [89] schemes were employed. The first we will discuss is the second order Runge-Kutta, which is accurate to $O((\Delta t)^3)$ and $\Delta t$ is the discrete time step used.

Given the continuous dynamics,

$$\dot{x} = G(x),$$

in $\mathbb{R}^n$, we define the two vectors $m_1$ and $m_2$ as follows,

$$m_1 = \Delta t \times G(x(t)),$$

$$m_2 = \Delta t \times G(x(t) + m_1).$$

and the update formula from $x(t)$ to $x(t + \Delta t)$ is

$$x(t + \Delta t) = x(t) + m_2.$$

This is the method used for the integration of the excitable phytoplankton dynamics in chapter 4 as the spatial integration schemes were second order accurate in space and time. Hence, the temporal integration scheme was chosen to be of the same accuracy.
The second method we use is the fourth order Runge-Kutta scheme, accurate to $O((\Delta t)^5)$. Assuming that we have the same governing dynamics as above, we define the following four vectors $m_1$, $m_2$, $m_3$ and $m_4$,

$$m_1 = \Delta t \times G(x(t)),$$
$$m_2 = \Delta t \times G\left(x(t) + \frac{1}{2}m_1\right),$$
$$m_3 = \Delta t \times G\left(x(t) + \frac{1}{2}m_2\right),$$
$$m_4 = \Delta t \times G(x(t) + m_3),$$

and the updating formula for $x(t)$ to $x(t + \Delta t)$ is

$$x(t + \Delta t) = x(t) + \frac{1}{6}(m_1 + 2m_2 + 2m_3 + m_4).$$

This is the numerical scheme we use for the integration of the ODE's in chapters 2, 5 & 6 as for a lot of the work in both chapters 2 & 5 we require accurate integration schemes.

8.2.2 PDE integration

The integration of partial differential equations is a little bit more complex a task than integrating ordinary differential equations as we generally have both spatial and temporal derivative terms to account for. Methods usually employed are spectral, pseudo-spectral, finite element and finite difference methods [89]. For the integration of the PDE's in chapter 4, we used only spectral and finite difference methods. The first PDE that we had to integrate was the linear stochastic partial differential equation (SPDE) in (4.25), which describes the temporal evolution of the stream function which defines the synthetically turbulent flow used in this chapter.
By moving into discrete Fourier space, the SPDE from (4.25) becomes a linear un-coupled ordinary differential equations. The discrete Fourier space is defined as in [62] where \( k = (k_x, k_y) = 2\pi N \Delta x(\mu, \nu) \) (where \( N^2 \) is the number of mesh points on the computational grid, \( \Delta x \) is the fixed grid spacing and we henceforth use Greek indices to represent the point in Fourier space).

The subsequent linear nature of the resulting sets of stochastic ODE's means that the exact integration formula [62] for \( \eta_{\mu\nu}(t) \) to \( \eta_{\mu\nu}(t + \Delta t) \) is given by

\[
\eta_{\mu\nu}(t + \Delta t) = \exp(\nu \mathcal{L}_{\mu\nu} \Delta t) \eta_{\mu\nu}(t) + \beta_{\mu\nu}(t) + \gamma_{\mu\nu}(t),
\]

in terms of the new random variables, \( \beta_{\mu\nu}(t) \) and \( \gamma_{\mu\nu}(t) \). These new variables are defined according to

\[
\beta_{\mu\nu}(t) = Q_{\mu\nu} \mathcal{D}_{\mu\nu}^w \int_t^{t+\Delta t} ds \zeta_{\mu\nu}^w(s) \exp((t + \Delta t - s)\nu \mathcal{L}_{\mu\nu}),
\]

\[
\gamma_{\mu\nu}(t) = Q_{\mu\nu} \mathcal{D}_{\mu\nu}^\eta \int_t^{t+\Delta t} ds \zeta_{\mu\nu}^\eta(s) \exp((t + \Delta t - s)\nu \mathcal{L}_{\mu\nu}),
\]

where \( Q_{\mu\nu} \) represents the discrete Fourier transform of the stirring operator, \( Q[\lambda^2 \nabla^2] \), from equation (4.25). In Fourier space, the Laplacian, \( \mathcal{L}_{\mu\nu} \), and divergence, \( \mathcal{D}_{\mu\nu} \), operators are translated to

\[
\nabla^2 \rightarrow \mathcal{L}_{\mu\nu} = \frac{2}{(\Delta x)^2} \left( \cos \left( \frac{2\pi \mu}{N} \right) + \cos \left( \frac{2\pi \nu}{N} \right) - 2 \right),
\]

\[
\nabla \cdot \rightarrow \mathcal{D}_{\mu\nu} = \frac{1}{\Delta x} \left( \exp \left( \frac{2\pi \mu}{N} \right) - 1, \exp \left( \frac{2\pi \nu}{N} \right) - 1 \right).
\]

This gives the details of the integration algorithm but the full details of how the correlations and energy spectrum may be expressed in discrete Fourier space may
be found in the paper by Marti et al. [62].

Next we give a brief review of how we integrate the PDE's seen in equations (4.40) and (4.42) in chapter 4. These represent the reaction-diffusion-advection equations for the full inertial turbulence model and the reduced vortex model, respectively. To integrate these PDE's we first note that they occur in a general form

$$\frac{\partial u}{\partial t} = G(u) - \nabla \cdot (F(u)),$$

where the operators $G(\cdot)$ represents the spatially independent reaction dynamics and $F(\cdot)$ represents the spatial flux. If we assume that $T$ represents the finite difference scheme for the integration of the above PDE, then we can rewrite this finite difference scheme as $T = T_1 + T_2$, where these two new differencing schemes are the schemes for the integration of the operators $G(\cdot)$ and $F(\cdot)$, respectively. This method of operator “splitting” allows us to sequentially apply the differencing schemes to the PDE and the full scheme will be stable [89] if the individual schemes it comprises of are stable.

Given the fixed spatial and temporal grid spacings $\Delta x$ and $\Delta t$, the two-step Lax-Wendroff [89] scheme, in 2 spatial dimensions is given by the following algorithm. Firstly, we use the simple Lax method, this assumes that the system is flux conservative and that the forward centred discrete time derivative can be written as $\partial_t u \approx (u_{i,j}^{n+1} - u_{i,j}^n)/\Delta t$. Here the index $n$ represents the discrete time variable and the indices $i$ and $j$ the location on the 2D computational grid. The Lax scheme [89]
gives us the following initial step,
\[
\begin{align*}
    u_{i+1/2,j+1/2}^{n+1/2} &= u_{i,j}^n - \frac{\Delta t}{2\Delta x} \left( F_{i+1,j+1}^n - F_{i,j}^n \right), \\
    u_{i+1/2,j-1/2}^{n+1/2} &= u_{i,j}^n - \frac{\Delta t}{2\Delta x} \left( F_{i+1,j-1}^n - F_{i,j}^n \right), \\
    u_{i-1/2,j+1/2}^{n+1/2} &= u_{i,j}^n - \frac{\Delta t}{2\Delta x} \left( F_{i-1,j+1}^n - F_{i,j}^n \right), \\
    u_{i-1/2,j-1/2}^{n+1/2} &= u_{i,j}^n - \frac{\Delta t}{2\Delta x} \left( F_{i-1,j-1}^n - F_{i,j}^n \right).
\end{align*}
\]

The Lax-Wendroff scheme now utilises these intermediate quantities to calculate the fluxes \( F_{i+1/2,j+1/2}^{n+1/2}, F_{i+1/2,j-1/2}^{n+1/2}, F_{i-1/2,j+1/2}^{n+1/2} \) and \( F_{i-1/2,j-1/2}^{n+1/2} \) and the final step from \( u_{i,j}^n \) to \( u_{i,j}^{n+1} \) is given by
\[
    u_{i,j}^{n+1} = u_{i,j}^n - \frac{\Delta t}{\Delta x} \left( F_{i+1/2,j+1/2}^{n+1/2} - F_{i-1/2,j+1/2}^{n+1/2} + F_{i-1/2,j-1/2}^{n+1/2} - F_{i+1/2,j-1/2}^{n+1/2} \right).
\]

The two-step Lax-Wendroff scheme is second order accurate in both space and time. The first step avoids the problem of mesh-drifting [89] as the full integration step utilises the grid points either side of the grid location \((i, j)\) as opposed to some methods, such as the simpler Lax method which uses only those grid points to the left or the right of the grid point \((i, j)\). The spatially 1D version of this algorithm was also used for the spatial integration of the reduced vortex model in (4.42). For this scheme, we split the spatial integration into two-steps, one using the 1D Lax-Wendroff scheme for the advection step and the weighted Crank-Nicholson scheme [89] for the diffusion term. The 1D Lax-Wendroff follows from the 2D scheme explained above but involves specifying only two intermediate points to go one full time-step forward.

The weighted Crank-Nicholson scheme uses the following semi implicit-explicit
finite difference scheme for the standard diffusion equation,

\[ u_i^{n+1} = u_i^n - \frac{D \Delta t}{2(\Delta x)^2} \left( \theta (u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}) + (1 - \theta) (u_{i+1}^n - 2u_i^n + u_{i-1}^n) \right), \]

where the optimal weighting factor is \( \theta = 0.55 \) \cite{89} and \( D \) is the diffusivity. Solving the above equation reduces to solving a tri-diagonal linear equation for the \( u_i^{n+1} \) from the known variables \( u_i^n \) and the difference scheme is stable for all grid spacings and time steps.
Bibliography


[93] O. Reynolds, *An experimental investigation of the circumstances which determine whether the motion of water shall be direct or sinuous, and of the law of resistance in in parallel channels*, Phil. Trans. R. Soc. A. 174 935 (1883).


