A NON-LINEAR FIELD MODEL OF PATTERN FORMATION
WITH APPLICATION TO BIOLOGICAL REGULATION

by

Wendy Anne Maria Brandts

A Thesis submitted in conformity with the requirements
for the Degree of Doctor of Philosophy in the
University of Toronto

©Copyright by Wendy Anne Maria Brandts 1989
ABSTRACT

The study of complex physical systems promises inspiration, new conceptual approaches, and techniques for creating a theoretical foundation of biology. In particular, in the area of pattern formation, questions which are being explored from the physical point of view, such as the dynamics of growth, formation of pattern boundaries, occurrence of phase transitions, changes in symmetry, etc., may also be extended to the biological context. The multi-disciplinary search is ongoing, not only for the solutions to dynamical equations, but even for appropriate basic models that capture the essence of pattern formation mechanisms.

Useful insights into pattern formation problems in biological systems have been gained from the 'polar coordinate model' of positional information (French et al., 1976). On the other hand, the polar coordinate model lacks the formal structure to deal with global reorganization when an overall change in the size of the system occurs. To overcome this limitation and explore the insights offered into pattern formation by this established biological framework, we propose a morphogenetic field model of pattern formation, a physical model (Ginzburg-Landau theory) guided into particular form by biological phenomenology. We apply our model to pattern changes observed experimentally in Tetrahymena doublets, which show, in the conversion of two oral structures back to one, a surprising intermittent state with three oral structures. Our morphogenetic field model allows us to see 'why 3 is between 2 and 1', and to make detailed predictions on the location of pattern elements, novel cell configurations, and dynamical pathways of pattern formation.

We conclude that the morphogenetic field approach has been very fruitful, in offering a 'simple' physical explanation for the enigmatic patterns of Tetrahymena, in suggesting new directions -both experimental and theoretical- for further studies of pattern formation, and in supporting the unification of the formalism, phenomenology and concepts of physical theory with the foundations of theory in biology.
Acknowledgements

I would like to thank my supervisor of the past near decade, Lynn Trainor, who both introduced me to the green land between physics and biology (at a time when the world was beginning to seem to me too broad to be appreciated from a single discipline) and made it possible for me to learn and research there. Course on physical theory in biology → Summer job → Masters → Doctorate. His numerous helpful suggestions, open-minded views, and apparent confidence in the maturity and sensibility of us students has been invaluable.

The legacy of this work also owes great debts to John Totafurno, who made a serious start at biological pattern formation within this physics enclave, brilliant enough to be followed. John’s pioneering efforts have made it possible to say ‘morphogenetic field’ in room 314 without blushing, despite his good sense of humour. Thanks to Prof. Joseph Frankel of Iowa for his enthusiasm towards my project, for providing me with valuable information and suggestions from the model’s first inception, through polishing manuscripts, and on to the ‘coming out’ at the Gordon Conference. And to E. Marlo Nelsen, also in Iowa, for the many hours spent in discussion and demonstrations. To Tim Rogers goes appreciation for giving his talk on a formally related topic at just the right time, and for the following positive discussions, comments and reference sharing. Trevor Hart has been a willing advisor on some mathematical subtleties. And my hat off to Brian Goodwin of U.K. for whispering “Tetrahymena” to us in the first place.

Thanks to Douglas Gingrich for introducing me to CERN libraries and for computing aid, and the Oxford Nuclear Physics Dept., especially Roger Cashmore, for computing facilities at Oxford University, U.K. during the hottest summer of the century. (Credit also goes to F. James and M. Roos, authors of MINUIT, and to all the others who maintain and make available the valuable CERN Computer Centre.
My oral exam rehearsal benefited from the indulgent incisiveness of Marc Dignum, Ed Chahramani, Frank Jones, Dave Robert and Chris Roland. The real oral exams were enlivened by long standing committee members Profs. J. Prentice and D. Rowe (who has now thoughtfully read two of my theses), with guest appearances by Profs. J. Frankel, E. Larsen, A. Jacobs, R. Kapral, and M. Walker. Thank you all for lending an ear and keeping your brains in gear while inventing some wonderfully savory questions.

The general turmoil and irresponsibility of graduate student living were made possible by the contrasting careful toil and responsibility of Tony Kev and Marianne Khurana.

The continuing financial assistance provided by the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged, as well as support from the University of Toronto.

Rooftop celebrations courtesy of Marc and Dave (White Star Extra Dry), Chris and Frank. Not to mention Edith Piaf in Kensington. Zoom Schwartz Perfigliano. Thanks to our “group” (Jeff, Al, Conrad, Dave, Trevor, Frank, Louis, Kathy, Vic, Pat, Jo.in) who have blossomed beyond, –especially Frank, who let me escape being the last by being endlessly helpful to all in need of technical advice– for your vitality and humour and even help. And to Peter Turney, a model friend and scientist, for longstanding scientific and professional advice/encouragement.

Then there’s my family, never to be forgotten; lunches, flowers, with the city sisters Marees and Elisa, who have also attended to my plantation during extended voyages. And the parental units, always ready to celebrate and encourage. Thank you Douglas, for showing such faith in me and saying that one couldn’t write a decent thesis in 3 weeks. The 10th floor myth claims that it was written in 7 days.

Lastly, to the house of Twinings from which inspiration was drawn.
# Contents

1 Introduction  
1.1 Dynamical Models of Pattern Formation  
1.2 Positional Information  
1.3 Morphogenetic Field Concepts  
1.4 Fields in Physics and Biology  
1.5 Pattern Formation in *Tetrahymena*  
1.6 Outline  

2 The Field Model  
2.1 Positional Information of the Field  
2.2 The Energy Functional  
2.3 Discrete Form of E and Numerical Computation  

3 Solutions of the Field Model  
3.1 Form of $\phi(x)$ and $\theta(x)$:SYM, RI Solutions  
3.2 Degeneracy of RI Solutions  
3.3 RI Solution – Dependence on $\beta^2$ and L  
3.4 Energy Diagram and Transitions  
3.5 Convergence of Solutions  

4 Biology of *Tetrahymena*  
4.1 Cell Surface Markers  
4.2 Development  
4.3 Homopolar Doublets  
4.4 Reverse–Intercalation Model  

5 Field Model – Overview of Application  
5.1 Summary of Solutions  
5.2 Energy Diagram and Transitions  

6 Field Model Configurations of *Tetrahymena*  
6.1 Positional Designations  
6.2 Symmetric (SYM) Solutions  
6.2.1 Singlets and Doublets  
6.2.2 CVP Sets  
6.3 Reversed (RI) Solutions  
6.3.1 Singlets  
6.3.2 Triplets  
6.3.3 Dynamical Pathways
7 Summary and Conclusions

A Derivation of Euler–Lagrange Equation

B Derivation of the Dynamical Field Equation

C Elliptic Solutions

D Tanh Interface

E Analytic SYM Solution

F The Rotational Symmetry of SYM

G Random Walk and Relative Balance

H Sensitivity of SC1 Width to $\beta^2$

I Glossary of Abbreviations
   I.1 Experimental Terms .................................. 83
   I.2 Notation from our Model .............................. 84

List of Figures

1 The Symmetric Solution with $W = 1$ ...................... 17
2 The Double Well ........................................... 20
3 Solutions of Gradient Field $\phi(x)$ ...................... 24
4 Gradient Field Saturation ................................ 25
5 $\phi(x)$ Solution from Negative ‘Seed’ .................. 25
6 Solutions of Field $\theta(x)$ .............................. 26
7 RI Field Degeneracy ...................................... 27
8 $\phi(x)$ : RI Field for Range of $\beta^2$ Values ........... 27
9 $\theta(x)$ : RI Field for Range of $\beta^2$ Values ........... 27
10 $\phi(x)$ : RI Field as Circumference is Reduced ........ 28
11 $\theta(x)$ : RI Field as Circumference is Reduced ........ 28
12 Lowest Energy Solutions ................................ 29
13 Energy Diagram .......................................... 29
14 Convergence of $\phi(x)$ Solution with Lattice Size .... 31
15 Convergence of $E$ with Lattice Size .................... 31
16 The Normal Growth of Wildtype Tetrahymena .......... 33
17 Polar Projection of Tetrahymena ......................... 33
18 The Morphogenetic Field ................................ 39

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
LIST OF TABLES

19 Solution Classes: SYM(W=2), SYM(W=1), RI .................. 40
20 Symmetric Solution SYM(W=2) .................................. 45
21 RI Solution L = 3.250\pi .................................................. 45
22 RI Solution L = 3.125\pi .................................................. 45
23 RI Solution L = 3.000\pi .................................................. 45
24 RI Solution L = 2.875\pi .................................................. 45
25 Symmetric Solution SYM(W=1) ...................................... 45
26 Model Distribution of CVP Sets ...................................... 46

List of Tables

1 Dependence of RI Solution Domain on $\beta^2$ ...................... 30
2 Semicell Lengths of Triplet States ................................... 50
3 Probability to Get Triplets For Different L Values ............... 56
4 Dynamical Pathways of Singlets ................................... 61
5 Dynamical Pathways of 1-CVP Set Cells ........................... 62
6 Relative Balance in Random Walk Model ............................ 82
7 Semicell Widths and Imbalance for Different Values of $\beta^2$ .. 82
1 Introduction

1.1 Dynamical Models of Pattern Formation

To create a theoretical foundation for biology, and connect it with physics and the other sciences, it is necessary to develop dynamical models of biological systems. This is especially clear in the area of pattern formation, where scientists from many different disciplines have been concerned with such questions as the dynamics of growth, formation of pattern boundaries, occurrence of phase transitions, changes in symmetry, etc. The search is ongoing, not only for the solutions to dynamical equations, but even for appropriate basic models that capture the essence of pattern formation mechanisms. The challenge of biological systems is particularly great, because the elementary ‘constituents’ of a pattern, as well as the fundamental ‘interactions’ of the processes are often unknown; moreover, the great degree of complexity often found in biological systems makes it difficult to choose the most fruitful and tractable simplifications at any of the many levels of modeling (e.g. molecular, supramolecular, tissue, macro-structure, etc.).

In this work, we develop and demonstrate one possible approach to the creation of a dynamical model of pattern formation. We treat pattern formation at a global or ‘whole system’ level: building on the qualitative biological concepts of positional information and intercalation (e.g. see Wolpert, 1969; French et al., 1976; Nelsen & Frankel, 1986), we develop a mathematical model of a morphogenetic field, expressed as a non-linear vector field, containing the positional information, and driven by global energy minimization. For a conceptually similar treatment of a biological system, supernumerary production in limb regeneration, see (Totafulno & Trainor, 1987). The general approach we take also has parallels in the theory of modulated structures in crystals in condensed matter physics (Cowley et al., 1979); for example, phase transitions to commensurate or incommensurate charge-density-wave states

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
1.2 Positional Information

Useful insights into pattern formation problems in biological systems have been gained from the concept of positional information, which expresses the idea that cells come to 'know' their position with respect to some boundary or origin and then interpret that information (e.g. decide what cell type to differentiate into) according to their genetic program. In particular, the polar coordinate model (French et al, 1976) has helped to provide a qualitative framework for understanding some aspects of regulation, the process by which biological organisms control and maintain their organization (i.e. their spatio-temporal patterns). Intercalation in positional information systems, the insertion or interpolation of new positional values, is typically attributed to the occurrence of a discontinuity in positional values (such as when a piece of tissue is surgically removed and distant cells, not normally neighbours, are brought into proximity) which causes the new positional values to be generated (i.e. intercalated). Two distinct processes involving readjustment of positional values are recognized:

**epimorphosis** Involves growth and localized pattern formation, leaving much of the remaining tissue unaltered. (New growth at the cut surface $\rightarrow$ new positional values inserted locally at the discontinuity.)

**morphallaxis** Entails long range interactions within a group of cells. (No growth $\rightarrow$ rearrangement of positional values at the discontinuity and beyond.)

We shall be concerned with a sort of generalization of the morphallactic concept, wherein rearrangement of patterns (e.g. positional values) may occur over a large part of the system, even when there is no single isolated discontinuity. This can bring about intercalation within a morphallactic as well as an epimorphic system.
1.3 Morphogenetic Field Concepts

Our concept of morphogenetic fields as an ordering influence for pattern formation in biological systems is closely related to positional information. The focus of both these concepts is the information on where parts of a biological structure are in relation to the whole (e.g. where particular features are located on the surface of the organism), and how this information enables highly complex patterns to form in an integrated and coordinated fashion. The morphogenetic field in our proposal is the source of organizational information for the organism, i.e. the configuration of the field determines or is the distribution of pattern features, and therefore it specifies the biological patterns. The means by which the organism translates the field values into growth statements (e.g. molecular assembly) is not specified. We are concerned only with the generative dynamics of the spatial differentiation patterns, whether they be manifested by order parameter fields, chemical waves, visco-elastic strains, electric potentials, morphogen gradients, molecular conformations, intrinsic clock phases, genetic switches, or whatever. Our aim is to create a simple but explicit and quantitative global law for pattern formation, and test whether such a principle is capable of predicting experimentally observed patterns.

The ability of parts to assess their relationship to other parts of the whole suggests that some kind of activity is effected globally\footnote{The activity may be local (i.e. local interactions) while its occurrence globally (i.e. everywhere in the system) leads to overall order, or globally integrated information.}. The logic of this activity (interaction) is given by the dynamics, which determine how the positional information (morphogenetic field) develops in time, e.g. what solutions it may attain. In our model—and indeed in many physical systems—the governing dynamics are equivalent to a requirement that a function, such as the total energy, be minimized. This minimization principle can be compared to the ‘shortest intercalation rule’ of the polar coordinate model (French et al., 1976) which specifies that a minimal
continuous set of positional values be added (intercalated) at a discontinuity. In
our case however, the 'rule' is given an explicit mathematical form, so that not only
can it be compared with physical principles (and perhaps be seen as an analogous
biological principle), but its behavior becomes more specific and may be tested on
a quantitative basis.

To define a field or distribution, an underlying space or manifold is required,
along with some coordinatization to specify locations. We express the fundamental
geometry of our biological system and the coordinatization of its surface by a grid,
defined to be a coordinatized space of positions (e.g. like a grid of longitude and
latitude lines on the surface of a sphere) with a metric or scale to measure distances
in the biological system. This grid may be associated with some structural repeat
distance of the system (e.g. regular rows of hair) or it may have no visible mani-
festation. A distribution of positional information or morphogenetic field values, in
particular, the definition of their 'smoothness' or spacing over the biological system,
deeps on this grid, against which such spatial distributions are to be measured.

Perturbing the system by removing a part of it affects both the morphogenetic
field (positional information) values –by creating a discontinuity– and the underly-
ing grid –which is reduced in total size. In the case of epimorphic growth, further
distortions in the grid may in principle take place, e.g. it may not grow back to
the original size, nor mesh with the pre-existing grid. In practice then, it may be
hard to separate the effects of the discontinuity in field (positional information)
values from the change in the grid (typically, this distinction is not made or can
not even be made in positional information models). Usually only the positional
information value discontinuity aspect is treated (as in the polar coordinate model
of French et al., 1976) because there is no explicit reference to a grid in the defini-
tion of positional values, let alone a comprehensive model of growth incorporating
the underlying grid or its regeneration.
Although we do not propose a model of growth, we explicitly recognize the role of the grid. We make simple assumptions about how it changes during regulation/development; in particular, we assume that it expands or contracts uniformly by the addition or removal, respectively, of 'grid points' at fixed separation distance. We study the effect of altering the grid by considering a morphallactic context: morphogenetic field values over the whole system may rearrange after a perturbation which changes the size of the underlying grid (but does not create a discontinuity). The question we would like to address is 'Can we get intercalation—or intercalation-like behavior—just by growth/shrinkage of the system (the grid) over which the morphogenetic field is defined?'. We note that according to the polar coordinate model for epimorphic regulation (French et al., 1976), while new positional values may be generated as a result of a discontinuity, there is no explicit provision made for intercalation of values if the system size changes without creating a discontinuity. It is this additional consideration that we wish to accommodate with our morphogenetic field model.

1.4 Fields in Physics and Biology

The use of fields that we make here has its formal origins in the role that fields play in physical theory. The mathematical properties of physical fields inspire their adoption to biological problems, not only because there already are well developed formal mathematical structures pertaining to physical fields which make possible quantitative predictions of dynamical behaviors, but because biological fields actually exhibit qualitative similarities to physical fields (e.g. the 'smoothness' property of positional information fields discussed below). Fields in physics are very general objects: they constitute a distribution or function of something (a scalar, vector or other quantity) over space, whose behavior in time is prescribed by dynamical laws, the field equations. Some physical fields are readily conceived of, from a
reductionist standpoint, merely as constructs which summarize the microscopic behavior of identifiable entities (e.g. molecules in chemical concentration fields obeying reaction-diffusion equations); however, other physical fields seem in themselves to be the 'substance' (e.g. electric fields) and are not easily thought of as a collection of more fundamental objects (e.g. quantum fields).

Whatever the status of the field, there are properties which are common to many kinds of fields and are of particular relevance in relation to biological fields. The 'smoothness' properties of fields, such as their continuity and differentiability, are an important aspect both from the physical and the mathematical perspectives. Continuity for fields means that continuous mathematical constructs (e.g. differentiable functions, subject to laws of calculus) can be used to represent them. In addition, smoothness criteria applied to dynamics of a field guide the field so as to maintain that property. Some smoothing behavior is in fact central to all positional information models that account for new patterns by 'filling in' missing positional values, i.e. by the process of 'intercalation' (e.g. French et al., 1976). Smoothness criteria in essence help to shape the dynamical laws of these positional information fields. The great significance of dynamical laws here is that once they are identified, and particular boundary values have been discovered, then an entire field can be interpolated within the boundaries. In electrostatics for example, one can determine a unique electric field everywhere in a region of space, given sufficient information about the field in a restricted region (e.g. given the appropriate Neumann or Dirichelet conditions on the boundary of the region). In reference to the biological context, we view this ability to interpolate the field as analogous to the ability of an organism to regenerate a whole from the information contained in a few of its parts (e.g. Kauffman & Ling, 1981).

The global principle which determines the dynamical laws in physical problems can often be stated in the form of a minimization of some action or energy function.
Requiring a system to evolve so that it minimizes a function of the system configuration often leads to a unique direction of development in the time evolution of the system (e.g. increase in the entropy function associated with irreversible behavior in thermodynamics).

The manner in which dynamical laws can be considered to operate in physical fields is also of interest in the biological extension. Dynamical equations often have alternative mathematical representations, which may in fact be entirely equivalent but entail seemingly different interpretations. For example, one form may be the differential equation, which expresses field values and interactions locally; the other may be the action integral or its analog which expresses the dynamics in terms of global properties. A particular example is found in classical mechanics, where a particle can be seen to move either according to Newton's laws (local, differential equation) or so as to minimize action over any segment of its future path, even its entire future path. The implication for biological fields is that global order might be thought of as due to local influences; thus, using the field concept for biological problems does not attach to it any non-physical 'action at a distance' or mysterious teleological attributes – any more so than is already accepted for physics that is!

1.5 Pattern Formation in *Tetrahymena*

The problem of *Tetrahymena* has not previously found a fully satisfactory solution, despite several ingenious attempts. For example, the inadequacies of two component reaction diffusion systems in dealing with the problem are discussed by Nelsen & Frankel (1986), and a representation of *Tetrahymena* using Laplace fields is constructed by Goodwin (1980), but he does not provide dynamical laws for their behavior. Our model not only provides insights into *Tetrahymena* but, in general, speaks strongly for the potential of field models in other intercalation or pattern formation systems as well.
Tetrahymena doublets are arrangements of two cells, fused side by side. During the regulation of these doublets, which initially have, for example, two complete oral structures, a configuration with three oral structures occasionally appears, before the subsequent conversion to a singlet with one oral structure. Our morphogenetic field model allows us to see ‘why 3 is between 2 and 1’ in a natural way, and to make detailed predictions about the location and orientation of pattern elements (biological markers), such as the oral structures.

The biological system that we have chosen to apply our morphogenetic field model to is atypical in two respects. Firstly, it is a single celled system, whereas the traditional application of positional information models is to multicellular organisms. In multicellular organisms, individual cells might be thought of as responding genetically to their local information, and effecting nuclear changes which lead to a particular differentiated state in that cell. In Wolpert’s positional information picture, for example, (Wolpert, 1969, p.15):

“...the positional information of [a] cell ... may be converted into the activation of a specific enzyme, which in turn may lead to activation of a gene coding for a structural protein which is an enzyme which leads to the formation of...”

Hence, which genes get turned on determines what kind of cell develops. Ultimately, however, knowledge of the detailed role played by genes (or other molecules) is inessential in the design of an organizing dynamic – much as the detailed properties of charged particles are not crucial to an understanding of macroscopic electric field behavior, or as the deviation of a real classical body from an ‘ideal point particle’ is irrelevant to the classical laws of motion. From this point of view, it is natural to extend the idea of positional information to systems wherein genetic changes cannot be the sole repository of positional cues, namely to the single cell.
Secondly, while the typical application of positional information emphasizes epimorphic regulation, our chosen system does not. Further, its development is followed by clonal lineages rather than in individuals. In its analogue for the single cell, regulating from generation to generation of cell growth and division, the distinction between epimorphic and morphallactic is quite subtle. Pattern formation in organisms exhibiting longitudinal growth followed by cell division might arguably be considered epimorphic, with the persisting regions of the cell serving as 'fixed differentiated tissue' and the new labile zone of growth as the effective 'intercalated tissue'. On the other hand, to the extent that the new growth does not arise primarily as a regulative response to a perturbation of the system's pattern, it appears as a natural part of cell division, and hence the new pattern forming in the labile zone could instead be regarded as resulting from 'long range interactions within the field' without new growth. We choose the interpretation\(^2\) of morphallactic regulation for *Tetrahymena*, since the patterns seem to change by *in situ* respecification in response to the reduction in system size (i.e. as the cell reduces from a doublet to a singlet size), rather than in response to the creation of a discontinuity with subsequent new localized growth.

In our morphogenetic field model, as the system size is changed, transitions from solutions with 'normal' symmetry to solutions with pattern *reversal* regions take place, suggesting how reverse intercalation phenomena can arise in morphallactic regulation even without the presence of a discontinuity. The effects of positional information in epimorphic regulation, already evidenced in a multiplicity of amputation and transplant experiments, might now also be tested in a morphallactic context, for single celled organisms (such as *Tetrahymena*), and should provide additional insights into pattern formation and positional information fields.

\(^2\)The experimental work on which our model is based (Nelsen & Frankel, 1986; Frankel & Nelsen, 1986) also view the problem as one of morphallaxis.
1.6 Outline

In this work, we develop a mathematical, one free (adjustable) parameter model of biological pattern formation akin to the *qualitative* polar coordinate (French et al., 1976) and reverse-intercalation (Nelsen & Frankel, 1986) models, and to the *quantitative* vector field model of limb regeneration of Tota furno & Trainor (1987). In section 2 we describe the form of the field and the general biologically inspired behaviors sought in its dynamics. An explicit form for the governing energy functional is given (section 2.2), the main feature of the non-linear energy density being a double-well (quartic) in the field gradient. The discrete form of the energy functional required for numerical computation is then outlined (section 2.3). The solutions of our model, in particular the transitions as the system size is changed, between equilibrium solutions of different symmetry (one set corresponding to reverse intercalated (RI) solutions) is discussed in section 3. The effect of varying the model’s free parameter is also addressed here (section 3.3). In addition, the reliable convergence of the solutions is demonstrated (section 3.5).

We then apply our morphogenetic field model to the regulation of Tetrahymena from a doublet state back to a singlet state. First the biology of Tetrahymena is summarized, its cytogeometry and its development, including the experimentally created doublet states (section 4). We then consider how the polar coordinate model (French et al., 1976) can be adapted to account in a qualitative fashion for the general features of the observed regulation (Nelsen & Frankel, 1986). In sections 5 – 6, we turn to the predictions of our model for this system; in particular, we translate the solutions of our model into specific statements about cortical patterns in Tetrahymena. The two types of solutions, ‘symmetric’ and ‘reverse intercalation’, and the transitions between them are discussed, making a detailed comparison with the cell features examined in the Tetrahymena experiments (Nelsen & Frankel, 1986; Frankel & Nelsen, 1986). We conclude with suggestions for further experimental
tests as well as for some refinements of the model (section 7).

The notable successes of our model are in producing states with a reversal in direction of 'winding' of the field, in making quantitative predictions for the patterns of Tetrahymena, including a new classification scheme based on the symmetry of the field solutions, and in discovering the essential role of the 'smoothness' term of the field, an aspect not previously recognized in the biological context of the qualitative polar coordinate model. The ability of our quantitative model to produce the experimental states, while also making new predictions, lends valuable support to the field concept in pattern formation and more generally, to the program of unifying biological understanding with theoretical physics.
2 The Field Model

2.1 Positional Information of the Field

Our morphogenetic field model is intended to apply to a system for which we can define a pattern organized about a circumference (i.e. a 1-dimensional ring), independently of other dimensions. For example, just as one can define the azimuthal positions on the surface of a cylinder independently of the height positions, one could apply our 1-dimensional ring treatment to any organism in which, for example, anterior-posterior and dorsal-ventral (azimuthal) organization is independent (perhaps laid out or fixed at a separate time in development) of proximal-distal (height) organization.

The morphogenetic field is a system of continuous values, spread or distributed around the circumference of an organism. It is important to distinguish the actual locations or positions (in ‘real space’) on the circumference of the organism, from the positional information values or field values (in ‘morphogenetic space’) that specify particular biological markers (pattern elements). The field has a reality separate from the actual spatial positions, independent of any coordinatization used to identify spatial locations; the field is a function of spatial position. A comparison with a familiar physical field serves to illustrate this important distinction. A ‘temperature field’ specifies a distribution of temperatures over some region of space: real space is coordinatized to describe locations, and at any point in real space there is some temperature field value (in ‘temperature space’) measured for example in units of ‘centimeters of mercury’.

Our morphogenetic field must contain sufficient positional information so that every point about a 1-dimensional ring (e.g. the circumference of a cell) has a unique designation. In addition, we require that the field be continuous and periodic around the ring, to ensure that it has the character of a real physical field and
can be interpreted in a biologically sensible way. Discontinuities (such as exist in the polar coordinate model (French et al., 1976) at 0/12, or in the reverse-intercalation model (Nelsen & Frankel, 1986) at 0/10) which do not actually represent real or physical discontinuities in the field are not allowed. The labeling of positions (coordinatization of a space) can have discontinuities (e.g. due to its topology) but the field defined over that space must nevertheless be continuous if there are no underlying physical discontinuities. These three conditions, uniqueness, continuity, and periodicity require the use of at least 2 degrees of freedom in the morphogenetic space. In particular, a minimal description is achieved by using either two truly periodic (around the circumference) scalar functions (since more than one is needed to ensure that every position on the cell has a unique field specification), or perhaps more conveniently by using a vector field (e.g. see Sibatani 1981, p.440; Totafurno & Trainor, 1987). We choose for simplicity a vector field with a constant magnitude, and a direction $\theta(x)$, the angle relative to some arbitrary morphogenetic field axis. $\theta(x)$ specifies the morphogenetic field value at each point $x$ on the circumference.

The starting point of the mapping between our vector field (morphogenetic field) and circumferential biological markers (e.g. cell features) is arbitrary, just as the point where one assigns positional value “1” in a polar coordinate model is arbitrary. Once a feature of the organism is assigned the initial value of the morphogenetic labeling scheme, the subsequent features found in the normal system will be labeled in the manner that they appear sequentially around a normal circumference. For example, 0-1-2-3-4-5-6-7-8/0 in a polar-coordinate model would be analogous to $\uparrow \nearrow \rightarrow \downarrow \rightarrow \searrow$ in a 2-dimensional vector model, and if a 'mouth' is labeled 5 ($\nearrow$) the 'pore' normally found somewhat to its right might be labeled 7 ($\searrow$).

In Fig. 1 we show a morphogenetic field at eight selected points along the circumference in real space. The 'value of the morphogenetic field' is given by the angle
θ relative to the morphogenetic field axis. (To follow our previous illustration, the
'value of the temperature field' is given by the height of mercury relative to the
scale on the thermometer). Though we have drawn the morphogenetic field axes
perpendicular to the spatial z axis, these two axes have no geometrical relationship
between them because they refer to different spaces. (Cf. the orientation of a ther-
mometer relative to space is irrelevant to the temperature reading, since the height
of mercury is defined w.r.t. the 'temperature–space axis' of the thermometer, not
w.r.t. any real–space direction.) That we have depicted the morphogenetic field
axis to have the same orientation relative to real space, at each point in space, is a
mere diagrammatic convenience.

The simplest solution or configuration of the field (Fig. 1) represents one com-
plete set of positional values, with a winding number of \( W = 1 \). Winding number
is defined as the number of full rotations of the morphogenetic field vector as one
goes around the circumference. We call this simple \( W = 1 \) solution the symmetric
solution, SYM(W=1), because the even spacing between vectors over the length
(\( L \)) of the circumference (i.e. their constant 'rate of rotation') implies that this
field configuration has certain symmetry properties. The SYM(W=1)solution will
be seen to represent the normal or wildtype configuration of our biological system.
Other arrangements can be represented by different lengths \( L \), by other winding
numbers \( W \), and by variations in the direction of the vectors (i.e. variation in their
rate of rotation and direction of rotation about the circumference). For comparison,
the biological assignments dorsal, ventral, anterior, and posterior can be thought of
roughly as orthogonal directions \( \uparrow \rightarrow \downarrow \leftarrow \) of the field, though we emphasize that
whereas the D-A-V-P system often implies special and distinguishable axes, our
field allows continuous transitions from dorsal to anterior to ventral to posterior.

Figure 1: The Symmetric Solution with \( W = 1 \). See next page.
Fig. 1  The Symmetric Solution with $W = 1$

The normal, simple configuration of the morphogenetic field, the symmetric solution with winding number $W = 1$, $SYM(W=1)$. The field is a 2-dimensional vector whose direction $\theta$ (the angle relative to some arbitrary morphogenetic field axis) specifies the morphogenetic field value for each circumferential position. The symmetric, winding number $W = 1$, configuration shown here at eight selected points of real space, represents one complete set of positional values, evenly spaced over the circumference.
Having now defined the morphogenetic field (and its relation to positional information), we need to assign to it some dynamical properties, so that we can determine what configurations (and hence what biological patterns) might arise in the cases where the normal field SYM(W=1) is perturbed, esp. when the circumference on which it sits is altered by growth/shrinkage or surgical manipulation (as in our example of *Tetrahymena* regulation of doublets with circumference $2L$, shrinking down to singlets with circumference $L$). The biological basis of our morphogenetic field model derives from characteristics of morphallactic regulation encompassed by the positional information scheme of Nelsen & Frankel (1986). The characteristics that we shall adapt to our morphogenetic field and its dynamics are summarized below (*ib.*, p.66. Note: Nelsen & Frankel employ a system of (scalar) positional values running from 1 to 10/0.):

1. There is continuity everywhere around the cell, as in a clockface. 10 is identical to 0.

2. The positional values can be ordered in either direction around the cell.

3. There exists an optimal spacing of positional values, with upper and lower thresholds of stable distances between adjacent values on both sides of this optimum.

4. When values become too closely spaced or when discontinuities develop, the system tends to regulate to a state of continuity, following a route that involves a minimum number of newly generated positional values.

These properties have correlates in our model as follows:

1. The field is continuous and periodic around the cell circumference.

2. The field can wind in either direction (clockwise or counterclockwise).
3. There is an optimal gradient\(^3\) of winding, i.e. an optimal value of $|\frac{\partial \theta}{\partial x}|$, where $\theta$ is the angle of the morphogenetic field vector, and $x$ is the circumferential position.

In addition, we shall require that changes in the direction of winding are generally not favored.

4. The dynamics are given by minimization of an energy functional; the energy density is a functional of the field. In particular, the energy density will have minima at the values of the optimal field gradient.

This last point 4. expresses a conceptual advancement from the reverse-intercalation and polar coordinate models. The implicit introduction of a grid (section 1.3) allows us to abandon the 'shortest intercalation rule' for a more generalized 'energy minimization rule' since we have now made explicit the distinction between the gradient of the field (the 'spacing of positional values') and the underlying space. In the next section, we develop the energy functional of the field. The term 'energy' as used here is not necessarily the same as the physical energy of the system. Presumably it is, however, a dynamical quantity which the system minimizes.

2.2 The Energy Functional

In the spirit of Ginzburg-Landau theory, we develop an energy functional whose minimization will determine the dynamics of our field. The energy density functional represents the energy (per unit length) contained locally in the morphogenetic field. The total energy of the system for a particular configuration of the field (e.g. a whole cell) is obtained by integrating the energy density around the entire circumference. It is that total energy which will be minimized.

\(^3\)Gradient here refers to how fast the vector field (morphogenetic field) winds, i.e. angle of winding per unit distance taken around the circumference.
The normal circumference of the biological system, which represents its usual or wildtype size, and the optimal gradient are defined consistently in terms of each other: the morphogenetic field having the optimal gradient everywhere, winds exactly once over a distance equal to the normal circumference (this morphogenetic field is a SYM(W=1) configuration, see Fig. 1). In other words, for a single winding of the morphogenetic field vector over a normal circumference $L_0$, the value of the optimal gradient is $\pm \frac{2\pi}{L_0}$, so that the field angle $\theta$ changes by $2\pi$ in the length $L_0$.

We desire the simplest form for an energy functional which is 1) smooth and 2) has minima at the (magnitudes of the) optimal gradient of the field. Our choice of a symmetrical double well (shown in Fig. 2) reflects that both directions of winding

are favored equally, other things being equal. The simplest form for an energy meeting the above criteria is a quartic in the field gradient and so we take for the first term in the energy density:

$$e_1 = \left[ \left( \frac{\partial \theta}{\partial x} \right)^2 - \left( \frac{2\pi}{L_0} \right)^2 \right]$$

which is intrinsically non-negative and has minima at $\pm 2\pi/L_0$. While this term alone is the complete analogue to postulate 3. above of the reverse-intercalation model (Nelsen & Frankel, 1986), we have discovered that an additional feature of the morphogenetic field is explicitly required in a quantitative model, namely that changes in direction of winding from clockwise to counterclockwise or v.v. are not favored (i.e. changes cost energy). This additional requirement discourages 'oscillating' configurations, which have the optimal gradient spacing for the most part but alternate between the positive and negative winding (e.g. in terms of positional information, states like 5 - 6 - 5 - 6 - 6 - 5 are discouraged because each change in winding direction of the morphogenetic field $\uparrow \downarrow \uparrow \downarrow \uparrow \downarrow \uparrow \downarrow$).
**Fig. 2  The Double Well**

The first term of the energy density, \( e_1 \), has the general form of a symmetrical double well, with minima at \( \pm \frac{2\pi}{L_0} \). It is a quartic in the field gradient \( \frac{d\theta}{dx} \).
costs energy). The additional term in the energy density, the manifestation of a ‘smoothness’ requirement, is taken to be

\[ e_2 = \beta^2 \left( \frac{\partial^2 \theta}{\partial x^2} \right)^2, \]  

(2)

where \( \beta^2 \) is a ‘weighting’ of the influence of ‘smoothness’, relative to the importance of attaining the optimal gradient (i.e. of \( |\frac{\partial \theta}{\partial x}| \) being close to \( \frac{2\pi}{L_0} \)). Putting this all together, we have for the total energy \( E \) in the circumferential morphogenetic field:

\[ E = \int_0^L \left\{ \left[ \left( \frac{\partial \theta}{\partial x} \right)^2 - \left( \frac{2\pi}{L_0} \right)^2 \right]^2 + \beta^2 \left( \frac{\partial^2 \theta}{\partial x^2} \right)^2 \right\} dx \]  

(3)

where

\[ L = \text{actual circumference (system size) at some stage of development}, \]

and as before:

\[ L_0 = \text{normal circumference (size) of system}, \]

\[ x = \text{circumferential position} \ (0 \leq x \leq L), \]

\[ \frac{2\pi}{L_0} = \text{optimal gradient} \]

\[ \beta^2 = \text{relative weight of smoothness term}. \]

The closed topology of the circumference means that we will have to satisfy periodic boundary conditions for the field: \( \theta(L) = \theta(0) + 2\pi W, \ W = 0, \pm 1, \pm 2 \ldots \). The winding number \( W \) specifies the number of times the field vector rotates through the complete range \( 2\pi \) of its values in one transit around the circumference (and the sign of \( W \) gives the direction of winding).

In order to simplify the above \( E \) functional, we define the gradient field

\[ \phi(x) \equiv \frac{\partial \theta}{\partial x} \]
(which can be thought of as the rate of change of angle $\theta$ in morphogenetic space, rather than the direction of the field vector itself). If we choose, for simplicity, $L_0 = 2\pi$ for the normal circumference, so that the optimal gradient has the value unity (amounts to choosing appropriate units for length around the circumference), then $E$ expressed as a functional of $\phi(x)$ takes the form:

$$E = \int_0^L \left\{ \left[ \phi^2 - 1 \right]^2 + \beta^2 \left( \frac{\partial \phi}{\partial x} \right)^2 \right\} dx$$  \hspace{1cm} (4)

with $\phi(L) = \phi(0)$, and $\int_0^L \phi dx = 2\pi W$;

$$0 \leq x \leq L.$$

The integral boundary condition (i.e. integral winding condition) arises from the original periodic boundary condition on $\theta(x)$.

The parameters $\beta^2$ and $L$ (and $W$) cannot be scaled away, and represent real freedoms of the system, parameters with which to label independent solutions. It is possible to develop a dynamical field equation for $\phi(x)$ using a variational technique to minimize the energy integral $E$ (cf. time dependent Ginzburg-Landau theory). The resulting differential equation$^4$ is a non-linear diffusion equation (see Appendix B):

$$\frac{\partial \phi}{\partial t} + \frac{\beta^2}{2} \frac{\partial^2 \phi}{\partial x^2} + \phi - \phi^3 - \frac{\lambda}{4} \hspace{1cm} (5)$$

where $\lambda$ = a constant is the Lagrange multiplier for the integral winding condition.

At equilibrium (i.e. when $\frac{\partial \phi}{\partial t} = 0$) it is the second order Euler–Lagrange equation (see Appendix A):

$$\phi^3 - \phi + \frac{\lambda}{4} - \frac{\beta^2}{2} \phi'' = 0. \hspace{1cm} (6)$$

The equilibrium equation also has a first order form

$$(\phi^2 - 1)^2 - \beta^2 \phi^2 + \lambda \phi = c \hspace{1cm} (7)$$

$^4$The differential equations are related to the Cahn-Hilliard equations for phase transitions in binary systems.
or
\[
\phi^4 - 2\phi^2 + \lambda \phi - \beta^2 \phi^2 = c - 1. 
\] (8)

Since this differential equation is non-linear and does not have a simple analytic solution (except in the case \( \lambda = 0, c = 0 \), see Appendix D; it will have solutions in terms of Jacobi elliptic functions, see Appendix C) we choose a numerical approach based directly upon the \( E \) function.

### 2.3 Discrete Form of \( E \) and Numerical Computation

In order to do numerical calculations, particularly to minimize the functional \( E(\phi(x)) \) with respect to the form of \( \phi(x) \), we devised a semi-discrete form of \( E \). The space dimension \( x \) is taken to be a discrete 1-dimensional lattice with \( N \) lattice points. If the lattice spacing is \( h \), and there are \( M \) lattice points in the normal circumference:
\[
dx \rightarrow h = \frac{L_0}{M} = \frac{2\pi}{M}; \quad x \rightarrow (i-1)h, \quad i = 1, 2, \ldots, N + 1. 
\]
The field retains its continuous nature, but is now only evaluated at discrete space points: \( \phi(x) \rightarrow \phi_i \).

The \( E \) functional in discrete space takes the form:
\[
E = \sum_{i=1}^{N} \left\{ \left[ \phi_i^2 - 1 \right]^2 + \beta^2 \left( \frac{\phi_i - \phi_{i-1}}{h} \right)^2 \right\} h
\] (9)

with \( \phi_0 = \phi_N \), and \( \sum_{i=1}^{N} \phi_i h = 2\pi W \);

\[
i = 1, 2, \ldots, N. 
\]

\( E \) was minimized with the \( \phi_i \) as variable parameters, using the function minimization program MINUIT (long, double precision version, algorithms SIMPLEX followed by MIGRAD) from the CERN library. MINUIT requires the input of an initial set of minimization parameters, the \( \phi_i \). We used a 'seed' of -1.0 values among a 'background' of +1.0 values to get rapid convergence (a few thousand iterations).
The periodic boundary conditions (period $L$) were employed in both 'fixed ends' and 'free ends' schemes. Fixing the ends ($\phi(0) = \phi(L) = \text{constant}$) has negligible effect on the solution energy when the number of lattice points $N$ is sufficiently large (i.e. fixed and free end solutions become identical when $M \geq 16$, see section 3.5). We used a lattice of $M = 32$ ($N_{\text{max}} = 64$), which is large enough that neither the exact value of $M$ itself, nor the initial conditions, nor the fixed vs. free ends conditions, affect the solution (see section 3.5). Values of $\beta^2$ between 0.0 and 1.0 (usually 0.2) were used for most of this work.

3 Solutions of the Field Model

3.1 Form of $\phi(x)$ and $\theta(x)$: SYM, RI Solutions

From the minimization of $E(\phi(x))$, or $E(\phi_i)$, we find two different symmetry classes of solutions for $\phi(x)$, or $\phi_i$ (see Fig. 3). The simpler form of solutions, the symmetric

![Figure 3: Solutions of Gradient Field $\phi(x)$. See next page.](image-url)

\textit{ric solutions} $\text{SYM}(W=2)$ and $\text{SYM}(W=1)$, and their energies, can be calculated analytically (see Appendix E) and were verified numerically. We obtained the second class of solutions, the \textit{reverse intercalation solutions} (RI), by numerical minimization as described in section 2.3. Fig. 3 illustrates the two symmetry classes of solutions $\phi(x)$, at a system size (circumference $L$) and value of $\beta^2$ for which the RI solution has lower energy than the SYM solutions. The constant functions $\phi(x) = \pm 1$ appear to be \textit{saturation levels} for the $\phi(x)$ solutions (i.e. values which effectively bound $\phi(x)$ from above and below): the $\phi(x)$ solutions can reach these levels but do not exceed them. Only the upper saturation level $\phi = +1$ is actually reached for this solution, but for other values of $\beta^2$ and $L$, the lower level $\phi = -1$ is reached as well. In Fig. 4 for example, it is apparent that as $\beta^2$ is reduced, greater curvature is allowed in the boundary between $\phi > 0$ and $\phi < 0$, and the reversal
Fig. 3 Solutions of Gradient Field $\phi(x)$

$\phi(x)$ for $L = 3\pi$; $\beta^2 = 0.2$. Examples of solutions belonging to the two symmetry classes, SYM and RI, of gradient field $\phi(x)$ resulting from the minimization of $E(\phi(x))$, are shown. There appear to be saturation levels (i.e. upper and lower bounds) at $\phi = \pm 1$. The upper level $\phi = +1$ is reached for this RI solution but the lower level $\phi = -1$ is only reached by solutions with other values of $\beta^2$ and $L$. The form of $\phi(x)$ reflects its affinity for $|\phi| = 1$ values, as well as for smooth boundary regions. The optimal value of $\phi(x)$, $\phi = +1$, is shown as the horizontal line $\phi_{optimal}$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
region approaches closer to the lower saturation level $\phi = -1$. By $\beta^2 = 0.1$, it appears that this lower level has been reached and $\phi(x)$ flattens out along it (i.e. the domain of $\phi(x) = -1$ grows). The basic form of $\phi(x)$ reflects the tendency of the solution to seek values of $|\phi| = 1$, with the boundary regions between $\phi = +1$ and $\phi = -1$ being as smooth as possible.

The symmetry between winding directions implies that in addition to the RI solution shown, which has $\phi > 0$ for the most part and reaches the upper saturation level, and the SYM solutions which are everywhere positive, there will be corresponding solutions with $\phi < 0$ for the most part, for both types of solution, RI (which will then reach only its lower saturation level) and SYM. These complementary solutions are easily obtained and are indeed identical (see Fig. 5) to a reflection of $\phi(x)$ (shown in Fig. 3) through the x-axis. Because of this complete symmetry between positive and negative (clockwise and counterclockwise) winding directions, we shall only present results for the positive solutions (i.e. solutions with $\phi > 0$ for the most part).

Because we are assuming that there are no ‘special’ points about the circumference of the biological system (cf. any point on the circumference could be taken as the ‘starting point’ for assigning morphogenetic field values), any arbitrary point on the circumference may be taken to be the location where $\phi = 0$; hence, there is in fact a whole set of solutions of $\phi(x)$, each with a different location of $\phi = 0$ (i.e. each with different phases labeled by values of $x_0$ as defined in $\phi(x_0) = 0$). We have chosen $\phi(0) = 0$ for convenience, which puts one of the RI boundaries at $x = 0$; however, the RI boundaries could in principle occur anywhere around the circumference.

Figure 4: Gradient Field Saturation. See next page.

Figure 5: $\phi(x)$ Solution from Negative ‘Seed’. See next page.
Fig. 4  Gradient Field Saturation

$\phi(x)$ for $L = 3.5\pi$; $\beta^2 = 0.1, 0.2, 0.3, 0.4, 0.5$. As $\beta^2$ is reduced, greater curvature appears in the boundary between $\phi > 0$ and $\phi < 0$. The 'tongue' of the reversal region approaches $\phi = -1$ but does not exceed it. These solutions are not the global minimum energy solutions; at this $L$ value, SYM(W=2) has lower $E$ than RI does.
Fig. 5  $\phi(x)$ Solution From Negative ‘Seed’

$\phi(x)$ for $L = 3\pi; \beta^2 = 0.2$. $\phi(x)$ solutions obtained with a predominantly ‘negative seed’ in the initial conditions. Corresponding to Fig. 3 which has $\phi > 0$ for the most part, these are solutions with $\phi < 0$ for the most part. These complementary solutions are identical to a reflection of $\phi(x)$ of Fig. 3 through the $x$–axis.
We obtained the field $\theta(x)$ from the gradient field $\phi(x)$ by numerical integration, using constant of integration $\theta(0) = 0$. The arbitrariness of the assignment $\theta(0) = 0$ is related to the arbitrary starting point of labeling around the circumference of the system. Fig. 6 shows $\theta(x)$ for the same parameter values as $\phi(x)$ in Fig. 3.

Figure 6: Solutions of Field $\theta(x)$. See next page.

It is apparent that at this circumference, the SYM(W=2) solution has a gradient of $\theta(x)$ greater than the optimal gradient $|\phi| = 1$; in other words, SYM(W=2) is steeper (field angle winds 'faster') than the optimal slope (the optimal slope would be parallel or perpendicular to the line $\theta(x) = x$). In similar terminology, the SYM(W=1) solution is everywhere flatter than the optimal slope, whereas the RI solution has the optimal slope for most of its domain—except close to the smoothed region of transition between winding directions.

A fundamentally important difference between the SYM and RI solutions $\theta(x)$ lies in their different symmetry: SYM has a rotational symmetry that RI does not. (A formal statement of the different rotational symmetry of SYM and RI can be found in Appendix F.) As $L$ decreases, the system undergoes a transition from SYM(W=2) to RI and the symmetry of SYM is broken: all points on the circumference are no longer equivalent due to the appearance of an RI region. This spontaneous symmetry breaking will be seen to have significant implications for the "crowding of positional values" in the application to Tetrahymena.

### 3.2 Degeneracy of RI Solutions

In arriving at specific solutions $\theta(x)$ we have assigned the values of $\theta(0) = 0$ and $\phi(0) = 0$, which correspond to specific choices of the phases of the periodic functions $\theta(x)$ and $\phi(x)$, respectively. In Fig. 7 we show a subset of solutions $\theta(x)$ with the same $\theta$-phase, $\theta(0)$, but different $\phi$-phases, $\phi(0)$. All these solutions have the
Fig. 6  Solutions of Field $\theta(x)$

$\theta(x)$ for $L = 3\pi; \beta^2 = 0.2$. The field $\theta(x)$ is shown for the same parameter values as the gradient field $\phi(x)$ in Fig. 3. At this circumference ($L = 3\pi$), the SYM(W=2) solution is 'steeper' than the 'optimal slope', i.e. it has a gradient of $\theta(x)$ greater than the optimal gradient (the optimal gradient $\frac{d\phi}{dx} = 1$ would appear parallel or perpendicular to the line $\theta = x$). The SYM(W=1) solution is 'flatter' than the 'optimal slope', i.e. it has a gradient of $\theta(x)$ less than the optimal gradient. The RI solution has the optimal slope (gradient) for most of its domain, except near the region of transition between winding directions (i.e. near $x = 2.5\pi$ and $3\pi$).
same energy (they are degenerate) and are related by simple linear transformations. Fortunately, it turns out that 'without a loss of generality', we can restrict our attention to the particular subset of solutions \( \theta(x) \) with \( \theta(0) = 0 \) and \( \phi(0) = 0 \). We will see how all the biological configurations (i.e. the spatial arrangements of biological elements making up the biological pattern) predictable by our model, can be derived from this one particular subset of solutions.

### 3.3 RI Solution – Dependence on \( \beta^2 \) and L

We now examine the behavior of RI solutions as the model's parameters are varied. Our model has two\(^5\) independent parameters, \( L \) and \( \beta^2 \). In our interpretation, the parameter \( \beta^2 \) is free to be adjusted to fit the experimental situation. On the other hand, \( L \) serves as a kind of time marker or dynamical variable (in a more complete model explicitly incorporating growth/shrinkage, the size of the system \( L \) would become a dynamical variable \( L(t) \)) and is therefore not a fitting parameter per se.

In Fig. 8 and Fig. 9 we show RI solutions differing only in the value of \( \beta^2 \). \( L \) was chosen so that we would get a large range of \( \beta^2 \) for which RI was actually the lowest energy solution\(^6\). Increasing \( \beta^2 \) reduces the curvature in the region of transition between \( \phi > 0 \) and \( \phi < 0 \), and forces the reversal region away (above) the lower saturation level at \( \phi = -1 \). The main effect on \( \theta(x) \) of increasing \( \beta^2 \) is to reduce the curvature at the boundary between regions of opposite winding direction,

---

\(^5\)Changes in scale of \( z, E, \) and \( \phi(x) \) do not further reduce the number of parameters.

\(^6\)The values of \( L \) which have RI as the lowest energy solutions, for several values of \( \beta^2 \), can be determined from Fig. 13.

---

Figure 7: RI Field Degeneracy. See next page.

Figure 8: \( \phi(x) \): RI Field for Range of \( \beta^2 \) Values. See next page.

Figure 9: \( \theta(x) \): RI Field for Range of \( \beta^2 \) Values. See next page.
Fig. 7  RI Field Degeneracy

$\theta(x)$ for $L = 3.25\pi; \beta^2 = 0.2$. A subset of solutions $\theta(x)$ with the same $\theta$-phase, $\theta(0)$, but different $\phi$-phases, $\phi(0)$, is shown. The particular form of solution ($\theta(0) = \phi(0) = 0$) used in this work and for biological interpretation is emphasized.
Fig. 8 \( \phi(x) \): RI Field for Range of \( \beta^2 \) Values

\( \phi(x) \) for \( L = 3.125\pi; \beta^2 = 0.1, 0.2, 0.3, 0.4 \). Same parameters as Fig. 9. RI solutions differing only in the value of \( \beta^2 \) are shown. Values of \( \beta^2 \) are in the range for which the RI solution has lower energy than the SYM solution. The main effect of increasing \( \beta^2 \) is to reduce the curvature of the boundary between \( \phi > 0 \) and \( \phi < 0 \), forcing the minimum of \( \phi(x) \) upwards.
Fig. 9 \( \theta(x) \): RI Field for Range of \( \beta^2 \) Values

\( \theta(x) \) for \( L = 3.125\pi \); \( \beta^2 = 0.1, 0.2, 0.3, 0.4 \). Same parameters as Fig. 8. RI solutions differing only in the value of \( \beta^2 \) are shown. Values of \( \beta^2 \) are in the range for which the RI solution has lower energy than the SYM solution. For \( L = 3.125\pi \) this includes \( 0 \leq \beta^2 \leq 1/2 \). The main effect of increasing \( \beta^2 \) is to reduce the curvature at the boundary between regions of opposite winding direction, increasing the \( x \)-domain of the nearly linear segment of \( \theta(x) \) and lowering the peak in the RI region.
increasing the x-domain of the nearly linear part of the solution and lowering the peak (and hence the range of \( \theta \)-values reached) in the RI region. An increase in \( \beta^2 \) also increases the energy of the RI solution, so that while each of the four curves shown (Fig. 8 and Fig. 9) is the minimum energy solution for its particular \( \beta^2 \) value, the solutions increase in energy with \( \beta^2 \) (the SYM solutions are independent of \( \beta^2 \) since they have zero curvature in \( \theta(x) \)).

In Fig. 10 and Fig. 11 the behavior of RI solutions at fixed \( \beta^2 \), but decreasing

Figure 10: \( \phi(x) \) : RI Field as Circumference is Reduced. See next page.

Figure 11: \( \theta(x) \) : RI Field as Circumference is Reduced. See next page.

circumference \( L \), is illustrated. With the chosen boundary conditions \( \theta(0) = \phi(0) = 0 \) each curve has a virtually identical segment from \( x = 0 \) to somewhat above \( 2\pi \). In \( \theta(x) \) these segments are linear and are essentially the same as a SYM(W=1) solution, except for the flattening at the RI boundaries which makes them extend beyond the normal (circumference) domain of \( 0 - 2\pi \). The reduction of \( L \) essentially affects only the RI region, causing the length (x-domain) of the reversal region and the extent of field values reached (Fig. 10: \( \phi \)-range or depth of minimum; Fig. 11: \( \theta \)-range or the height of the peak) to decrease.
Fig. 10  \( \phi(x) \): RI Field as Circumference Is Reduced
\( \phi(x) \) for \( L = 3.25\pi, 3.125\pi, 3\pi, 2.875\pi; \beta^2 = 0.2 \). Same parameters as Fig. 11. RI solutions at fixed \( \beta^2 \) but decreasing circumference \( L \) are shown. Reducing \( L \) affects only the RI region, reducing the length of the RI region and the range of \( \phi < 0 \) reached.
Fig. 11  \( \theta(x) \): RI Field as Circumference Is Reduced

\( \theta(x) \) for \( L = 3.25\pi, 3.125\pi, 3\pi, 2.875\pi; \beta^2 = 0.2 \). Same parameters as Fig. 10. RI solutions at fixed \( \beta^2 = 0.2 \), but decreasing circumference, \( L \), are shown. The boundary conditions \( \theta(0) = \phi(0) = 0 \) were chosen so that the curves would have a nearly linear part (extending from 0 to just above \( 2\pi \)) similar to a SYM(W=1) solution. The reduction of \( L \) affects only the RI region, causing the length of the region, and the range of field values attained to decrease.
3.4 Energy Diagram and Transitions

A set of minimum energy solutions (for $\beta^2 = 0.2$) over the whole range $^7$ of $L$ between $4\pi$ and $2\pi$ is summarized in Fig. 12. The progression: $\text{SYM}(W=2) \rightarrow \text{RI} (W=1) \rightarrow \text{SYM}(W=1)$, can also be followed on the Energy Diagram (Fig. 13), which shows the energy $E$ of a given solution as a function of $L$.

We follow the $L$ dependence of the solutions, starting with large $L$. At a relatively large circumference ($L \approx 4\pi$) the lowest energy solution is $\text{SYM}(W=2)$. As the circumference decreases, a point of transition is eventually reached where, depending on the value of $\beta^2$ assumed in the model, $\text{SYM}(W=2)$ no longer has a lower energy than RI. At this point ($L \approx 3.3\pi$, for $\beta^2 = 0.2$) a major shift in the symmetry of the solutions occurs, with the winding number $W$ decreasing from 2 to 1. There is a well defined range of $L$, the RI solution domain, over which RI will continue to be the lowest energy solution. This domain depends on $\beta^2$: as can be seen in Fig. 13, there is a different curve for each $\beta^2$ (Table 1 summarizes the domains). For values greater than $\beta^2 \approx 0.5$ the domain decreases to nothing; that is, when $\beta^2 \geq 0.5$ we would not expect to see the RI solution appear for any circumference size. As $L$ decreases below the domain of RI solutions ($L \approx 2.8\pi$, for $\beta^2 = 0.2$), a second transition point is reached, and the system now switches from RI to $\text{SYM}(W=1)$. This second transition does not involve a change in $W$ however, and the field undergoes a less drastic alteration than at the first transition.

---

$^7$This is the range of $L$ that will be of interest in our application to Tetrahymena. The two basic transitions described below could be found repeated at other particular (larger and smaller) values of $L$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Fig. 12  Lowest Energy Solutions

$\theta(x)$ for $L = 4\pi, 3.75\pi, 3.5\pi, 3.25\pi, 3\pi, 2.75\pi, 2.5\pi, 2.25\pi, 2\pi; \beta^2 = 0.2$. The set of minimum energy solutions for $\beta^2 = 0.2$, over the range of circumferences from $L = 4\pi$ to $L = 2\pi$ (at intervals of four lattice points, $N = 4$) is shown. The progression from $\text{SYM}(W=2) \longrightarrow \text{RI} \longrightarrow \text{SYM}(W=1)$ solutions is evident with decreasing $L$. The transition from $L = 3.00\pi$ to $L = 2.75\pi$, going from RI to $\text{SYM}(W=1)$, is apparently less drastic than the transition between $L = 3.50\pi$ and $L = 3.25\pi$ which requires a change in the winding number, $W$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Fig. 13 Energy Diagram

Energy vs. $L$ (circumference): SYM($W=1$), SYM($W=2$), and RI ($\beta^2 = 0.1, 0.2, 0.3, 0.4, 0.5$). The energy of SYM and RI solutions, as a function of $L$, is plotted. The energy of lowest-energy solutions, in the progression SYM($W=2$) $\rightarrow$ RI $\rightarrow$ SYM($W=1$) for decreasing $L$ (shown in Fig. 12), is found by following along the SYM($W=2$) curve until the transition at $L \approx 3.3\pi$ (for $\beta^2 = 0.2$); lowest energy solutions are subsequently obtained along the RI energy curve until the second transition at $L \approx 2.8\pi$ where the SYM($W=1$) curve acquires the lowest energy. The value of $\beta^2$ alters the energy of only the RI solution, but consequently also affects the range of $L$ over which RI are the lowest energy solutions.
Table 1: Dependence of RI Solution Domain on $\beta^2$. The relative RI solution domain is defined as the range of circumference lengths $L$ for which the RI solutions are the lowest energy solutions, divided by the length of the total range of $L$, $4\pi$.

The curves on the energy diagram (Fig. 13) are not symmetric about the half way point of the circumference $L = 3\pi$ (since the energy density 'well' is not symmetric about $|\phi| = 1$). This means that the range of $L$ for which SYM(W=1) has minimum energy is greater than for SYM(W=2) and there should be a predominance of SYM(W=1) solutions over SYM(W=2), as is evidenced in Fig. 12.

Since $\beta^2$ determines both transition points simultaneously, we could in principle choose its value with reference to experimental data, i.e. by fitting to the location ($L$ value) of one of the transitions (giving at the same time a prediction of where the other one falls). In the preceeding analysis, we generally used the value $\beta^2 = 0.2$, to make our illustration of the model definite. This chosen value of $\beta^2$ seems to be a good 'middle' value, within the realm where RI solutions are permitted, and with a broad enough range of $L$ that the progression through changing $L$ can readily be followed.

In summary, near the minimum and maximum circumferences considered, i.e. near $L = 2\pi$ or $L = 4\pi$, the lowest energy solution will be SYM(W=1), SYM(W=2) respectively. At intermediate circumference sizes, the RI solution will have the lowest energy. Both the RI solution domain (Table 1) and the location of the two transition points are determined by $\beta^2$.

The meaning of $\beta^2$, as will be recalled, is that it measures the importance of the
second ("smoothing") term in the energy functional, relative to the first ("optimal spacing") term. The fact that $\beta^2$ has to be small to get RI behavior in this model, in particular that $\beta^2 < 1$, indicates that the influence of this second term is indeed less than that of the first term. Hence we can understand why the polar coordinate type models have some limited success, even though they effectively represent only the first term.

### 3.5 Convergence of Solutions

To test the effect of discretization in this model, several different lattices were used for the space ($x$) variable. In particular, the numerical minimization was executed for several different basic lattice sizes. Recall that the normal circumference has $M$ lattice points, and that the actual circumference $N$ may vary between $M$ (normal singlet) and $2M$ (normal doublet). In Fig. 14 four solutions $\phi(x)$, $\beta^2 = 0.2$, $L = 3.00\pi$ with fixed endpoints $\phi(0) = \phi(L) = 0$ are shown, each calculated on a different basic lattice ($M = 8, 16, 24, 32; N = 12, 24, 32, 46$). All three solutions $M \geq 16$ appear to fit the same curve. For $M = 8$ the solution is overly sensitive to initial conditions and convergence criteria; in particular, for the conditions used to obtain these solutions, the $M = 8$ solution did not give a reliable curve (e.g. here $\phi(x)$ is below the upper saturation level, a slight change in the initial 'seed' causes $\phi(x)$ to go visibly above the upper saturation level). For $M > 16$ the lattice used (size $M$) does not significantly affect the form of the solution. In Fig. 15 the energies of these solutions ($L = 3.00\pi$, $\beta^2 = 0.2$) are plotted, along with the corresponding $E$ for solutions that were not required to satisfy $\phi(0) = \phi(L) = 0$ (though of course
Fig. 14  Convergence of $\phi(x)$ Solution with Lattice Size

$\phi(x)$ for $M = 8 \, *, 16 \, \square, 24 \, \Delta, 32 \, \circ; \, N = 12, 24, 36, 48; \, \beta^2 = 0.2$. For lattice sizes above $M = 8$ the $\phi(x)$ solutions with fixed endpoints appear to converge to one curve. For $M \leq 8$ the convergence is not stable, i.e. the form of $\phi$ depends sensitively on initial conditions and on convergence criteria.
Fig. 15  Convergence of $E$ with Lattice Size

KEY: $\beta^2 = 0.2 \square$, $\beta^2 = 0.4 \bigcirc$; 'Fixed ends': solid markers, 'Free ends': open markers.

The energy $E$ as a function of the basic lattice size $M$. The energy for a particular solution $\phi(x)$ appears not to depend significantly on the lattice size $M$ when $M \geq 16$. $E$ increases with $M$ monotonically; for small $M$ the effect of the Free end condition is pronounced.
\( \phi(0) = \phi(L) \) must still be satisfied; in addition, a similar set of solutions with \( \beta^2 = 0.4 \) is shown. The \( E \) of each solution converges monotonically with increasing lattice size. Only at the smallest lattice size shown \( (M = 8) \) is the difference between fixed \( \phi(0) = \phi(L) = 0 \) and free \( \phi(0) = \phi(L) = \text{anything} \) ends apparent. We chose \( M = 32 \) as our basic lattice size for this work, since it has a safe margin of convergence and is not excessively demanding of the numerical computations (which must minimize with \( 2M - 1 = 63 \) variable parameters).
4 Biology of Tetrahymena

Tetrahymena is a ciliated protozoan (Protozoans are both a single cell and a whole animal at the same time, two thrills in one, the hydrogen atom of biology.). Its clonal growth can be schematized as the elongation and sequential segmentation of a cylinder (Frankel et al., 1987). This kind of growth and division implies that any pattern that can be propagated longitudinally can be inherited ("cortical inheritance", the cortex is the cell surface) and hence this system allows us to address the question: is inheritance determined by the cell nucleus, or does the cortex (cell surface) as a whole control a self perpetuating pattern, relying on nuclear genes only for the synthesis of the necessary basic materials? The experiments we are concerned with here follow the sequences of patterns that appear many cell generations after a mechanical perturbation. Regulation tends towards the initial pattern which was 'in harmony' with the genotype. The 'inverse' experiments, involving a genetic perturbation with subsequent regulation towards patterns more compatible with the new genotype, have also been done (Frankel et al, 1987) and would make an interesting complementary application of our model, potentially leading to further insights into the interplay between genetics and dynamics in pattern formation.

4.1 Cell Surface Markers

The normal (wildtype) configuration of Tetrahymena has one oral apparatus (OA) —making it by definition a singlet— as well as several other distinguishing features (see Fig. 16 and Fig. 17). The general shape of the cell is roughly ellipsoidal, with

Figure 16: The Normal Growth of Wildtype Tetrahymena. See next page.

Figure 17: Polar Projection of Tetrahymena. See next page.

a narrowed anterior end. There are 18 to 21 longitudinal rows of cilia, the cil-
Fig. 16 The Normal Growth of Wildtype *Tetrahymena*

The cell grows longitudinally and then segments laterally. Cell markers shown are: CR, ciliary rows; RP, LP, right and left postoral ciliary rows, respectively, extending from oral area; OA, oral apparatus with UM, undulating membrane whose asymmetric location helps define the 'handedness' of the OA; CVP, contractile vacuole pores, which appear to the cell's right of the OA; OP oral primordium; FZ predivision fission zone.
Fig. 17  Polar Projection of *Tetrahymena*

The cell surface in a polar view (projection), to illustrate the circumferential relation of cell markers. Cell markers and distances shown are: oral apparatuses pOA1, pOA2, sOA; oral meridians pOM1, pOM2, sOM; semicells sc1, sc2; CVP midpoints CM1, CM2; CM locations d1, d2; sOA location (sector) sr. p- denotes primary, s- secondary. -1 denotes smaller semicell, -2 denotes larger semicell.
...which meet in a regular ring at the anterior (A) end and terminate at irregular positions at the posterior (P) end. If the system of pattern specification information (e.g. positional information or the morphogenetic field) covers the same region as the ciliary rows, then the surface will have the topology of a cylinder (i.e. with open ends, it can be 'unrolled' into a 2-dimensional sheet), even though the cell itself is closed and hence topologically spherical. Either way, an equatorial circumference (closed ring) of morphogenetic field values can be defined, representing the circumferential organization which is maintained over most of the A–P range, and is normally quite stable. The oral apparatus (OA) at the anterior (A) end is a compound structure: three membranelles and an undulating membrane (UM). The asymmetry of the OA makes it possible to define a ‘handedness’, which in the wild type has the UM to the cell’s right of the membranelles and is called a right-handed OA. The two ciliary rows at the longitude of the OA and terminating just posterior to it are labeled as the right postoral (RP) and left postoral (LP) rows. To the right of the RP, located near the posterior end of one, two, or three adjacent ciliary rows are the contractile vacuole pores (CVP's). A group of CVP's on adjacent ciliary rows is referred to as a CVP set. The location of the CVP's (d) (Fig. 17), as measured by the number of ciliary rows between the RP and the midpoint of the CVP set (CM), is approximately proportional to the cell circumference (c) (Nanney 1966a, 1967, cited in Frankel et al., 1987). Many of the useful definitions of this section (as well as other experimental jargon and terms pertaining to our model) can be found for the convenience of the reader with finite recall, in the Glossary (Appendix I).

4.2 Development

The cell surface development of Tetrahymena during clonal growth (Fig. 16) begins with the appearance of new basal bodies, which organize into an oral primordium
(OP) to the left of the midregion of the RP (a normal OP or OA is designated as primary, denoted pOP and pOA). A lateral break in the ciliary rows, the equatorial fission zone (FZ), appears anterior to the OP when it is nearly completed. This fission zone divides the ciliary rows into anterior and posterior segments. Just anterior to the break in the ciliary rows, new CVP’s form, usually along the rows established by the old posterior CVP’s. In the final stage, cleavage occurs along the fission zone, and two new individuals are formed, the anterior product containing the original OA and new CVP’s, the posterior product having a new OA and the original CVP’s.

4.3 Homopolar Doublets

Homopolar doublets are arrangements of two cells, joined side by side with anterior-posterior (A-P) axes parallel, obtained by various manipulations of cells or cell pairs (Nelsen & Frankel, 1986). They begin with approximately double the singlet number of ciliary rows, and with two OA’s and two CVP sets (Fig. 17 shows a polar view of the cell features and circumferential distances defined between them. This is the usual representation used by Frankel and Nelsen). A semicell of a homopolar doublet is defined as a ‘half’ of a cell, including a primary oral meridian (pOM: located at the right postoral (RP) ciliary row of the pOA) and that portion of the cell to its right, up to but not including the next oral meridian (OM).

The subcloning experiments begin with balanced doublets (i.e. the OA’s are on opposite sides of the cell; the semicells are equal in size), which then regulate in many generations (20 - 200) back to singlets. The reversion to singlets involves a loss of ciliary rows and the disappearance of one of the OA’s, and of one of the CVP sets. For doublets that remain ‘balanced’, when the number of rows is reduced to less than about 30 (i.e. halfway between the initial doublet and final singlet number) one of the OM’s is lost. Usually the loss of rows entails a shift in the relative positions of
the OM's so that the cell becomes 'unbalanced'; however, imbalance does not seem to arise solely from differential reduction of rows in the two semicells but rather by "slippage of the OP", as indicated by the correlation between a reduction of rows in one semicell with an increase in the other (Nelsen & Frankel, 1986, p.57). During the reduction of cells with two OA's to ones with a single OA, a transient state with three OA's may appear! These additional oral structures, the secondary OA's or OP's (sOA, sOP respectively), have a highly variable location but are found only in the narrow (smaller) semicell, and are always of abnormal shape – often resembling a left-handed version of the normal right-handed OA. Occasionally, oral areas that can be interpreted as the outcome of a fusion between a neighboring sOP and pOP are seen. In addition, throughout the development various other configurations of doublets and singlets appear (e.g. balanced doublets; unbalanced doublets; 1-CVP-set doublets; partial doublets with a normal and an abnormal OA and one or two CVP sets; 2-CVP-set singlets; singlets). These will be discussed in the context of our morphogenetic field model (section 6).

4.4 Reverse–Intercalation Model

In order to organize and explain the results of the doublet experiments, Nelsen & Frankel (1986) propose a modified version of the 'shortest–intercalation rule' of the polar coordinate model which states that (French et al., 1976, pg. 970):

When normally nonadjacent positional values in ... the circular ... sequence
[of positional values] are confronted ... growth occurs at the junction until cells with all intermediate positional values have been intercalated.
... continuity of the circular sequence means that there are two possible sets of intermediate values between any nonadjacent positional values ...
A critical stipulation of the model is that when cells with nonadjacent positional values in the circular sequence are brought into contact,
intercalation is always by the shorter of the two possible routes.

In their "reverse-intercalation model" points around the cell circumference are assumed to carry particular positional values (say, 1 to 10/0) which specify what structures will form there. This system of positional values is separate from the ciliary rows, although the surface structures develop next to ciliary rows and both can presumably respond to the cell size (however, ciliary rows do not undergo reversals of asymmetry). The reverse-intercalation model can be summarized in five postulates (Nelsen & Frankel, 1986, p.66):

1. Typically, there is continuity everywhere around the cell, as in a clock face. Ten is identical to zero.

2. The positional values can be ordered in either direction around the cell.

3. There exists an optimal spacing of positional values, with upper and lower thresholds of stable distances between adjacent values on both sides of this optimum.

4. When values become too closely spaced or when discontinuities develop, the system tends to regulate to a state of continuity, following a route that involves a minimum number of newly generated positional values.

5. The positional coordinates may be propagated longitudinally as the ciliate undergoes cellular and clonal growth.

The fifth postulate is a consequence of the way that ciliates grow, namely the elongation and division cycles allow longitudinal values to be preserved whereas latitudinal values must change (the cycle of A-P values must double). We are primarily concerned with longitudinal positional information here, the dimension corresponding to a circumferential ring, since this is the dimension which exhibits
unusual variations in the doublet experiments. The complementary problem of how the latitude values are duplicated in division does not concern us here, because the outcome of A-P pattern formation is essentially unaltered from the normal process in these experiments (e.g. prior to each division, doublets produce two new OA's at the same A-P position that singlets produce one new OA prior to division). Additional evidence that the circumferential dimension of positional information can be separated from the A-P dimension is provided by the mutant Janus, a mutation which selectively alters the circumferential pattern without affecting A-P polarity (Frankel et al., 1987).

The reverse-intercalation model is an ingenious scheme to account for the doublet to singlet transitional forms in a qualitative way, given the additional ad hoc aid of “non-uniform crowding of positional values” during shrinkage. While not an a priori unreasonable assumption in itself, it would be challenging to discover a scheme which could lead to observed reversed (RI) states even from uniform crowding. Moreover, it would be helpful to understand why ‘stretching’ did not also lead to intercalated states, and to account for the absence of ‘oscillatory’ solutions. In addition, and perhaps more significantly, one would like to have a testable quantitative model, especially given the fine detailed numerical/statistical data available for Tetrahymena doublets and mutants. In this work, we developed such a model, a quantitative morphogenetic field model of pattern formation in systems of changing size; we will apply the results (solutions) of our model to a detailed analysis of the case of Tetrahymena doublets regulating to singlets.

5 Field Model – Overview of Application

5.1 Summary of Solutions

We shall briefly review the results of our model so far, in terms appropriate to the analysis of Tetrahymena. In section 2 we constructed a morphogenetic field model
of pattern formation, reflecting the basic behaviour expressed qualitatively by the reverse-intercalation model. We then examined solutions to our model (section 3). The most important and striking solution, the reverse intercalation (RI) solution, with symmetry different from the normal singlet and balanced doublet symmetric (SYM) solutions, was obtained even with the assumption of uniform (homogeneous) shrinkage of the cell about the entire circumference. This is a valuable result because it shows that, although uneven shrinkage around the cell circumference might actually occur, it is not an essential prerequisite for the creation of RI states. In other words, an ad hoc non-symmetrical shrinkage does not have to be put into the model in order to generate the observed reversals. This is an especial advantage as there is little information about likely asymmetry of shrinkage in the cell. As we shall now see, the basic patterning of the doublet, singlet and reversal states is very well represented without including any non-uniformity a priori.

Our morphogenetic field specifies the circumferential morphogenetic (positional) information of a biological system (see Fig. 18), in the present case the organiza-

Figure 18: The Morphogenetic Field. See next page.

tion of positional longitudes of a Tetrahymena cell. This field contains sufficient information so that every point about the circumference of the cell has a unique designation. In addition, the field is continuous and periodic (after one transit around the circumference, the field again takes on its initial value). Associated with each possible configuration of the field is an energy, determined by the energy functional in the model. The configurations that specify the actual patterns found in the biological system will be the solutions of lowest energy.

These solutions can be shown on a graph of $\theta(x)$ vs. $x$ (see Fig. 6 and Fig. 19) where $\theta(x)$ is the morphogenetic field angle (the angle of the morphogenetic field vector, in morphogenetic space, relative to some fixed but arbitrary morphogenetic
Fig. 18 The Morphogenetic Field

The field is continuous, but is shown here only at discrete points about the circumference. The field values change with circumferential positions $x$; they do not change with anterior-posterior (A–P) location. The field angle $\theta(x)$ is the angle between the field vector at point $x$ and some fixed but arbitrary morphogenetic axis (shown as a dashed line).
axis) and $x$ is the circumferential position on the cell. Two main classes of (minimum energy) solutions were found, the symmetric (SYM) and the reverse intercalation (RI) solutions. The SYM solutions are the simplest solutions, the field angle winds at a constant rate $\frac{d\theta}{dx}$ (see Fig. 19). The number of complete turns the field vector makes in a transit around the cell circumference is called the winding number $(W)$. Our symmetric solutions with winding number $W=2$, denoted SYM($W=2$), represent doublet cells. The symmetric solutions with winding number $W=1$, denoted SYM($W=1$), represent singlet cells; in particular, the normal Tetrahymena is a SYM($W=1$) singlet.

The second class of solutions, the RI solutions, are more complex than the SYM solutions. The RI configurations contain a region of field which winds in the opposite sense to the rest of the field, e.g. the field may wind one full rotation in the clockwise sense, then wind some more clockwise, followed by a change in winding direction and a portion of counter-clockwise winding (see Fig. 19). Higher order (and higher energy) RI solutions could in principle occur corresponding to more complex arrangements with regions of positive winding interdigitated with regions of negative winding.

Which of the previous solutions has the lowest energy depends on the model's parameter $\beta^2$ (defined in section 2.2), which is fixed here at $\beta^2 = 0.2$, and on the system size or circumference $L$ which decreases as the cell shrinks. The system size or circumference is given by $L$: a normal singlet is defined to have a circumference $L = 2\pi$ and a doublet therefore has an initial circumference of $L = 4\pi$.

There are three domains of $L$, reached successively in time, corresponding to circumstances where each one of the three solutions SYM($W=2$), RI, SYM($W=1$) in turn has the lowest energy. These solutions of the model can most usefully be

Figure 19: Solution Classes: SYM($W=2$), SYM($W=1$), RI. See next page.
Fig. 19 Solution Classes: SYM(W=2), SYM(W=1), RI

There are three classes of solutions, SYM(W=2), SYM(W=1), RI. Which of these types will give the minimum energy solution depends on the circumference L. Three representative morphogenetic field configurations and their corresponding representations \( \theta(x) \) vs. \( x \), where \( \theta(x) \) is the morphogenetic field angle, are shown.
thought of as undergoing transitions. as \( L \) is decreased, from \( \text{SYM}(W=2) \) to \( \text{RI} \), and then from \( \text{RI} \) to \( \text{SYM}(W=1) \).

5.2 Energy Diagram and Transitions

The progression: \textit{doublet} \( \text{SYM}(W=2) \rightarrow \text{doublet or singlet} \) \( \text{RI} (W=1) \rightarrow \text{singlet} \) \( \text{SYM}(W=1) \), can be followed on the Energy Diagram (Fig. 13), which shows the energy \( E \) of a given solution as a function of \( L \). (The set of minimum energy solutions \( \theta(x) \) (for \( \beta^2 = 0.2 \)) over the whole range of \( L \), from 'initial' doublet to normal singlet, is shown in Fig. 12).

We follow the \( L \) dependence of the solutions, starting with large \( L \). The doublet starts out with a relatively large circumference \( (L \simeq 4\pi) \), for which the lowest energy solution is that of the \( \text{SYM}(W=2) \) doublet. As the circumference decreases with shrinkage and loss of ciliary rows, a point of transition is reached where (depending on the value of \( \beta^2 \) assumed in the model) \( \text{SYM}(W=2) \) no longer has a lower energy than \( \text{RI} \). At this point \( (L \simeq 3.3\pi) \), a major shift in the symmetry of the solutions occurs, with the winding number \( W \) decreasing from 2 to 1. There is a well defined range of \( L \), the \textit{RI solution domain}, over which \( \text{RI} \) will continue to be the lowest energy solution. As \( L \) decreases below this domain of \( \text{RI} \) solutions, a second transition point is reached \( (L \simeq 2.8\pi) \) where the system switches from \( \text{RI} \) to \( \text{SYM}(W=1) \). This second transition does not involve a change in \( W \) however, and we will see that the field indeed appears to undergo a less drastic alteration than at the first transition. Since a cell expressing the \( \text{RI} \) state can be a doublet or a singlet (i.e. it can have 1 or 2 oral apparatuses), the transition from doublet \( \rightarrow \) singlet does not necessarily coincide exactly with the transition from \( \text{SYM} \) to \( \text{RI} \). The underlying dynamical transition, or 'phase transition', is the \( \text{SYM} - \text{RI} \) transition, whereas the doublet–singlet transition is merely a partial manifestation of it. Note however that the doublet–singlet transition must take place somewhere within the
RI solution domain, and that all states before (i.e. larger $L$ than) the first transition are both SYM and doublet, while all states after (i.e. smaller $L$ than) the second transition are both SYM and singlet.

The curves on the energy diagram are not symmetric about the half way point of $L$ (at $L = 3\pi$). This means that the range of $L$ for which SYM($W=1$) has minimum energy is greater than for SYM($W=2$). There should thus be a favoring of "large singlets" over "small doublets" (assuming size changes linearly with time or generation number), a prediction which remains to be tested.

Since $\beta^2$ determines both transition points (Fig. 13), we could in principle choose its value with reference to experimental data, i.e. by fitting to the location (i.e. $L$ value) of one of the transitions, the theory should then predict where the other one falls. Likewise, we could determine the maximum expected size of a (SYM) singlet, given the minimum size of a (SYM) doublet. This might in practice prove to be a difficult project, because studies are done on clonal populations, with intrinsically variable numbers of ciliary rows, rather than by following individuals through all stages, and there seems to be no exact correlation between the number of ciliary rows and the stage of development.

Different clones—or even different individuals within a clone—might have different values of $\beta^2$ (e.g. it might be genetically determined). Or $\beta^2$ might vary in time or even with location on the circumference. In our analysis however, we have chosen the fixed value $\beta^2 = 0.2$, which gives appropriate transition points (circumference values) for Tetrahymena regulation, as well as can be inferred from the experimental data.

If $\beta^2$ is chosen to be greater than about 0.5, then RI solutions will always have higher energy than either of the SYM solutions. In that case, we would get a direct transition from SYM($W=2$) doublets to SYM($W=1$) singlets, at about halfway between the doublet and singlet size. Such transitions are observed for balanced
doublets (Frankel & Nelsen. 1986). a particularly significant observation since the
SYM(W=2) solutions are also balanced configurations (i.e. both semicells are identi-
tical). We note in addition, that direct transitions from doublet to singlet without a
transitory reversal region or third OA can also arise even when the RI intermediary
solution is attained (in this case, the RI region would not include the OA).

In summary, when the circumference (L) is near the normal values for a singlet or
doublet, the lowest energy solution will be SYM(W=1), or SYM(W=2) respectively;
otherwise, the RI solution will have the lowest energy. Both the exact RI solution
domain and the location of the two transition points are determined by $\beta^2$, which
we have taken to be $\beta^2 = 0.2$ for *Tetrahymena*.

The meaning of $\beta^2$ in our model is that it measures the relative importance of
a term which does not have a direct counterpart in the reverse-intercalation model
(Nelsen & Frankel, 1986), but is necessary for there to be appropriate solutions in
our quantitative formulation. The fact that $\beta^2 < 1$ indicates that the influence
of this term is less than that of the term originating from polar coordinate model
concepts, and hence we can understand why the reverse-intercalation model was
qualitatively quite effective for *Tetrahymena*.

6 Field Model Configurations of *Tetrahymena*

6.1 Positional Designations

The morphogenetic field values $\theta$ must be translated into designations of cell fea-
tures. To accomplish this, the normal sequence of biological features around the
cell circumference is associated with a sequence of field values. The initial point of
this association or mapping, i.e. where the 0 ($= 2\pi$) of the field falls, is arbitrary.
For example, any one of the field values may be designated as the value giving an
OA. This arbitrariness constitutes a degeneracy or a symmetry in the model and
means that there are no 'special' positions around the cell circumference. The ulti-
mate result of this degeneracy is that the RI region may occur anywhere about the circumference.

The effect of the degeneracy of the $\theta(x)$ functions (Fig. 7) will be included here by allowing the OA to take on, in turn, any positional value, while restricting consideration to only one of the degenerate $\theta(x)$ diagrams. This is entirely equivalent to choosing a single mapping between field values and cell features, and including all the degenerate $\theta(x)$ diagrams. In the following Figures (graphs of $\theta(x)$ vs. $x$), we will denote the morphogenetic field value $\theta$ of an OA by $R$ or $L$, depending on whether the field values around it are winding clockwise or counterclockwise, respectively. The idea is that some particular $\theta$ value specifies an OA, and the 'handedness of the field' in its neighbourhood (i.e. the local sense of rotation of the field) determines the handedness of the OA. When the field value of an OA appears on a reversal boundary, $L/R$ will be used, denoting a fusion of a left- and a right-handed OA.

The morphogenetic field values for the CVP sets are determined from three items of experimental data (Nelsen & Frankel, 1986):

1. A singlet has 18–21 ciliary rows.

2. The CVP's are along 1, 2, or 3 adjacent rows.

3. In a typical singlet, the average relative distance between pOA and the midpoint of the CVP set (CM) is $0.227 \pm 0.035$.

From 1. and 2. above we get that the maximum relative proportion of circumference (relative to a circumference of 1.0) over which a CVP set is normally formed is about 0.17 (i.e. 3/18). In terms of a system of morphogenetic field values running from $\theta = 0.0 \rightarrow 2\pi$, this is a range in $\theta$ of $(0.17 \times 2\pi = ) 1.1$. Since our $\theta(x)$ graphs are marked in grid intervals of $\pi/8$, the maximum range in $\theta$ for the formation of a CVP set will span $\approx 2.7$, or nearly 3, grid intervals of $\theta$ on the $\theta(x)$ graphs. The
average midpoint of the CVP set. CM, in a normal singlet (3. above) allows us to calculate the approximate separation in morphogenetic field value between pOA and the CM: $2\pi \times 0.227 = 1.4$, or 3 to 4 grid lines. For illustrative purposes, we shall mark CM locations on the $\theta(x)$ graphs by a heavy line extending from the $x$-positions corresponding to 3, to the $x$-positions corresponding to 4, grid intervals above (i.e. ‘to the right’ for clockwise winding) the field value of the OA. The full extent of CVP formation (up to th-ee adjacent ciliary rows) will be indicated by a dashed line extending out from the solid line of CM location. A representative set of minimum energy configurations $\theta(x)$ vs. $x$ for decreasing $L$ values is shown in Figs. 20–25.

Figure 20: Symmetric Solution SYM(W=2).

Figure 21: RI Solution $L = 3.250\pi$.

Figure 22: RI Solution $L = 3.125\pi$.

Figure 23: RI Solution $L = 3.000\pi$.

Figure 24: RI Solution $L = 2.875\pi$.

Figure 25: Symmetric Solution SYM(W=1).

6.2 Symmetric (SYM) Solutions

6.2.1 Singlets and Doublets

The experimentally recognized division of cells into two main classes, the ‘doublets’ and the ‘singlets’, (specifying how many oral apparatuses a cell has) is not the natural grouping for our model. Our model has two classes of solutions, SYM and RI, which reflect differences in their underlying symmetry. Both SYM and RI solutions can be manifested as doublets (SYM(W=2) doublets, or RI triplets not completely expressed), or as singlets. Conversely, doublets may be represented by
Symmetric and Reverse Intercalation Solutions

An illustrative sample of the minimum energy solutions at several values of $L$ for $\beta^2 = 0.2$. Each cell configuration is read along a horizontal line. The handedness of OA's are determined by the direction of winding of the field at the field value for OA: right handed (normal) OA's are denoted $R$, left handed ones $L$. Number 1, or 2 affixed indicates the semicell, when sc's are unequal size. The domain of CVP formation is indicated by solid ("most likely location") and dashed lines.

![Diagram of Symmetric Solution SYM(W=2)](image)

Fig. 20 Symmetric Solution SYM(W=2)

$L = 3.50\pi$. The SYM solutions are linear. All SYM(W=2) solutions correspond to doublets with 2 CVP sets.
Fig. 21  RI Solution $L = 3.250\pi$

Reading from the top down (i.e. beginning with the configuration along the horizontal line $\theta = 15 \cdot 2\pi/16$) the cell configurations are: (2x) singlet with 3 CVP sets; (1x) singlet with 2 CVP sets, one set broad; (1x) singlet with 2 CVP sets, one set narrow; (8x) singlet with 1 CVP set; (3x) triplet with 1 CVP set; (1x) fusing triplet with 2 CVP sets.
Fig. 22  RI Solution $L = 3.125\pi$

Reading from the top down (i.e. beginning with the configuration along the horizontal line $\theta = 15 \cdot 2\pi/16$) the cell configurations are: (1x) singlet with 2 CVP sets, one set broad; (1x) singlet with 3 CVP sets; (1x) singlet with 2 CVP sets, one set broad; (1x) singlet with 2 CVP sets, one set narrow; (9x) singlet with 1 CVP set; (2x) triplet with 1 CVP set; (1x) fusing triplet with 1 CVP set.
Fig. 23  RI Solution $L = 3.000\pi$

Reading from the top down (i.e. beginning with the configuration along the horizontal line $\theta = 15:2\pi/16$) the cell configurations are: (1x) singlet with 2 CVP sets; (1x) singlet with 3 CVP sets; (1x) singlet with 2 CVP sets, one set broad; (1x) singlet with 2 CVP sets, one set narrow; (9x) singlet with 1 CVP set; (2x) triplet with 1 CVP set; (1x) fusing triplet with 1 CVP set.
Fig. 24  RI Solution $L = 2.875\pi$

Reading from the top down (i.e. beginning with the configuration along the horizontal line $\theta = 15 \cdot 2\pi/16$) the cell configurations are: (1x) singlet with 1 CVP set, CVP set atypically far to the right; (1x) singlet with 2 CVP sets, one set broad; (1x) singlet with 2 CVP sets; (1x) singlet with 2 CVP sets, one set narrow; (10x) singlet with 1 CVP set; (1x) triplet with 1 CVP set; (1x) fusing triplet with 1 CVP set.
Fig. 25  Symmetric Solution \( \text{SYM}(W=1) \)

\( L = 2.75\pi \). The SYM solutions are linear. All \( \text{SYM}(W=1) \) solutions correspond to singlets with 1 CVP set.
SYM(W=2) or by RI solutions, and singlets may be represented by SYM(W=1) or by RI solutions. The SYM solutions give the simplest ('most symmetric') forms for doublets and singlets in our model.

Doublets (SYM(W=2)) and singlets (SYM(W=1) as well as RI — see section 6.3) should both be quite plentiful, assuming the experiments are able to support the appropriate range of cell sizes.

6.2.2 CVP Sets

The linear nature of the SYM solutions causes the location of the CVP midpoint, CM, to be linearly proportional to the size of the cell (the cell size is the total of the two semicells sc1, sc2, with sc1 being the narrower 'half cell' between two OA's, and sc2 being the wider 'half'); hence the relative value of CM in the SYM doublet will be the same as for the SYM singlet. This linear behavior accords with the experimental observation that "the mean distance of CM2 from pOM2 is linearly proportional to sc2" and that the average relative CM for doublets (0.185 ± 0.046) was not significantly different from the value for singlets (0.227 ± 0.035). The distribution of CVP sets for SYM(W=2) doublets is shown in Fig. 26 and agrees well with

Figure 26: Model Distribution of CVP Sets. See next page.

experiment.

6.3 Reversed (RI) Solutions

6.3.1 Singlets

In our model, RI solutions give rise to singlets when the RI takes place over a range of morphogenetic field values that includes the OA. Most of the RI solutions (≈ 80%) are singlets, the percentage of singlets increasing as the circumference (L)
Fig. 26 Model Distribution of CVP Sets

The distribution of CVP sets for 2-CVP-set doublets and 2- or 3-CVP set singlets for our model is shown. A. The distribution of CVP sets in 2-CVP-set doublets (all are SYM(W=2) solutions). The CM2 distribution is significantly wider than the CM1 distribution. The doublet distributions are narrower and higher than the 2-CVP-set singlet distributions. B. The distribution of CVP sets in 2-CVP and 3-CVP set singlets (all are RI solutions). The greatest spread of locations is found for CM2 of 2-CVP-set singlets. Doublets and 2-CVP-set singlets show approximately the same location of CM1, and the CM2 peak of 2-CVP-set singlets falls about halfway between the doublet peaks. All of the above features agree with the experimental distributions.
decreases, with 1-CVP-set states predominating. We also predict several new singlet states: some are subtle variations of typical singlets, others are more striking.

The most striking new singlet configuration is the 3-CVP-set state. In some instances, 2 of these 3 CVP sets are sufficiently close that they may appear as 2-CVP-set cells (depends on the intrinsic range of field values which correspond to CVP formation); however, there are also cases where all 3 sets are distinct. These 3-CVP-set configurations are analogous to triplet states (3-OA states) but do not manifest a reversal of internal symmetry of CVP's due to their intrinsically symmetric nature. That is, even if the CVP were formed in conformance to a global field of reversed winding direction, it's lack of intrinsic 'handedness' would presumably make it indistinguishable from a normal CVP.

Related to the 3-CVP-set singlets are the 2-CVP-set singlets wherein one of the CVP sets is unusually broad. These broad regions are CVP analogs of fused OA states. In our illustrative sample (from Figs. 21–24), all but one of the 2-CVP-set singlets contain one CVP set at approximately the typical singlet location. The additional CVP set(s) are further to the right of the OA (CM2/CM1 = 2.7 ± 0.1), though left of the location expected for CM2 of a doublet (for SYM(W=2), CM2/CM1 = 5.5 ± 0.1), which indicates that 2-CVP-set RI singlets are not just like doublets without a second OA, but are singlets with merging CVP sets (see Fig. 26). Taken together, we would therefore expect the inclusive CM (including all CVP sets) of 2-CVP-set singlets to be shifted to the right, and more the larger the cell is (greater L). The one singlet without a CVP set at the typical location occurs for small L, and the single CVP set is located distinctly further to the right.

Our distribution of CVP sets (Fig. 26) in 2-CVP-set doublets (which are all SYM(W=2)) and in multi-CVP-set singlets (RI solutions) can be compared to

---

8 Although not reported in the original doublet study of right-handed cells (Nelsen & Frankel, 1986) 3-CVP-set cells have now been observed in experiments on left-handed cells (Nelsen & Frankel, 1989).
the experimental distribution of CVP sets found for "typical doublets" and "2-CVP-set singlets" (A. and B. of Fig.12 in Frankel & Nelsen, 1986). We expect a fairly close correspondence, since our SYM(W=2) doublets would be observed as "typical doublets", and our 2-CVP-set singlet RI solutions would be observed experimentally as 2-CVP-set singlets; however, the correspondence is not exact because we also include 3-CVP-set singlets, thereby adding entries at the high end of "Distance from pOM to CM". Moreover, the experimental distribution of CVP's includes doublets and singlets of 'all sizes' (all cells with 28 or more ciliary rows), indistinguishably mixed together, whereas we do not include any overlap in sizes of doublets \( (L = 3.875\pi - 3.50\pi) \) and singlets \( (L = 3.25\pi - 2.875\pi) \) (we include only solutions which have the lowest energy for that size).

Our theoretical distribution of CVP's in doublets and 2-CVP-set singlets (Fig. 26) is qualitatively similar to the experimental distribution quoted above. Looking at "A. Doublets" first, we note that the CM2 distribution is wider than the CM1 distribution, both in our theoretical examples and in the experimental sample. Again in both theory and experiment, doublets and 2-CVP-set singlets show approximately the same location of CM1, and the CM2 peak of 2-CVP-set singlets falls about halfway between the doublet peaks. The greatest spread of locations is found for CM2 of 2-CVP-set singlets, both theoretically and experimentally, reflecting the dynamic range of 2-CVP-set singlet sizes during regulation. In overall appearance, the doublet distributions are 'sharper': they are both narrower and have significantly higher peaks than the 2-CVP-set singlets.

Since the RI region can occur anywhere in the morphogenetic field, we might expect the likelihood of getting a multiple-CVP-set singlet to be the same as a triplet OA; however, the case of the CVP sets is complicated by the fact that CVP's can form over a broader range of field values than can the OA's. The result of the broad range for CVP's is that the proportion of RI cells which can be expected to have
more than 1 CVP set is approximately fixed for all (except the smallest) sizes \( L \) of the RI domain, and is equal to \( \approx 25\% \). The details of whether these are 2-CVP sets or 3-CVP sets, and how broadened a CVP set may be depends specifically on the length of the circumference \( (L) \) and on the exact morphogenetic field value range of CVP specification.

6.3.2 Triplets

Appearance of Third OA \((L)\) in Smaller Semicell \((sc1)\)

A triplet state has three OA's: two primary right-handed ones \( R_1 \) (pOA1), and \( R_2 \) (pOA2), and a left-handed or secondary one \( L \) (sOA). (Note: experimental triplets are called "doublets with sOA" by Nelsen & Frankel.) Though much rarer than singlet states, particularly for small cell circumferences (small \( L \)), a variety of triplet configurations is obtained (see Table 2 for examples).

It is clear from Table 2 (and the \( \theta(x) \) graphs in Figs. 21 – 24) that the semicell in which the reversed or left-handed oral apparatus appears is always the narrower semicell \((sc1)\). For RI solutions, the difference between semicell sizes \( sc1 \) and \( sc2 \) is unmistakable, and the presence of a \( L \) (sOA) in a semicell is perfectly correlated with that semicell being distinctly smaller.

Correlation of Position of \( L \) (sOA) and Width of Smaller Semicell \((sc1)\)

According to our model, the \( L \) (sOA) may appear anywhere within the smaller semicell \( sc1 \), with equal likelihood to be near the \( R \) (pOA) to its left or to its right. The theoretical\(^9\) mean relative position of \( L \) (sOA) in \( sc1 \) (i.e. \( sr/sc1 \)) is therefore \( 0.5 \pm 0.2 \), but ranges from 0.0 to 1.0. This is consistent with the experimental finding that although the variability was great, it seems that all positions between \( pOM1 \) and \( pOM2 \) could have the sOA.

\(^9\)The approximate average and standard deviation is calculated from measurements at half-grid-divisions of \( \theta(x) \) graph and excludes fusion \( L/R \) events.
<table>
<thead>
<tr>
<th>$\theta$ of OA</th>
<th>sc2</th>
<th>sc1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>units of $2\pi$</td>
<td>sc2/c</td>
</tr>
<tr>
<td>$L = 3.25\pi$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 · $2\pi/16$</td>
<td>1.03</td>
<td>0.63</td>
</tr>
<tr>
<td>1 · $2\pi/16$</td>
<td>1.00</td>
<td>0.62</td>
</tr>
<tr>
<td>2 · $2\pi/16$</td>
<td>1.00</td>
<td>0.62</td>
</tr>
<tr>
<td>3 · $2\pi/16$</td>
<td>1.01</td>
<td>0.62</td>
</tr>
<tr>
<td>Average:</td>
<td>1.01</td>
<td>0.62</td>
</tr>
</tbody>
</table>

| $L = 3.125\pi$ |     |     |
| 0 · $2\pi/16$ | 1.03 | 0.66 | 0.52 | 0.33 |
| 1 · $2\pi/16$ | 1.00 | 0.64 | 0.56 | 0.36 |
| 2 · $2\pi/16$ | 1.00 | 0.64 | 0.56 | 0.36 |
| Average:       | 1.01 | 0.65 | 0.55 | 0.35 |

| $L = 3.00\pi$ |     |     |
| 0 · $2\pi/16$ | 1.03 | 0.69 | 0.45 | 0.30 |
| 1 · $2\pi/16$ | 1.01 | 0.67 | 0.49 | 0.33 |
| 2 · $2\pi/16$ | 1.01 | 0.67 | 0.49 | 0.33 |
| Average:       | 1.02 | 0.68 | 0.48 | 0.32 |

| $L = 2.875\pi$ |     |     |
| 0 · $2\pi/16$ | 1.05 | 0.73 | 0.38 | 0.26 |
| 1 · $2\pi/16$ | 1.02 | 0.71 | 0.42 | 0.29 |
| Average:       | 1.04 | 0.72 | 0.40 | 0.28 |

Table 2: Semicell Lengths of Triplet States. (Note: Since $sc1 + sc2 = c$, then in units of $2\pi$ a "full sized" normal doublet has $c=2$ and a normal singlet has $c=1$.)
The actual experimental value for the mean relative position of the sOA was 0.383, which seems to indicate a slight asymmetry of sOA positioning within sc1 (sOM has a tendency to be closer to pOM1 than to pOM2). However, as this value lies well within one standard deviation of our theoretical value, and considering the experimental instability of sOA expression, it is difficult to say whether this is a real asymmetry or just a statistical fluctuation away from 0.5.

In addition to the ratio sr/sc1 (which by itself is not sufficient to discover possible correlations between sc1 width and distance of sOA to either pOM), we examine the (minimum) distance of the sOA to its closest pOM and the corresponding sc1 width for our model. We find that for a given cell size \( L \), when the sOA is relatively close to a (either) pOM, sc1 is slightly smaller than when the sOA appears near the middle of sc1. Though this correlation holds for all the cell sizes examined (i.e. \( L/\pi = 3.25, 3.125, 3.00, 2.875 \); in Figs. 21 - 24), the maximum difference in sc1 widths for a given cell size \( L \) is only around 3%. Translating into numbers of ciliary rows, this comes to \( \leq 1 \) ciliary row (for a doublet of 30 rows) and hence would be difficult to detect in the experiments, especially as this 3% is buried beneath the larger variation in sc1 width with cell size \( L \).

Position of \( L \) (sOA) and Width of the Wider Semicell (sc2)

The experiments did not discover any relationship between the average position of sOA and the width of sc2; nor does our model indicate any strong correlation. The only link potentially being that sc2 is slightly narrower (\( \approx 2\% \)) when the sOA was very near to a \( R \) (pOA). As this represents a spacing effect of even less than 1 ciliary row, it would not have been detected by the experimental cell measuring scheme (i.e. the counting of ciliary rows).

Widths of Semicells (sc1 and sc2)

Detailed widths of sc1 and sc2 are given in Table 2 for four \( L \) values. The width of sc2 is approximately constant. The difference in total circumference of the
doublets is due primarily or exclusively to sc1, which decreases as L decreases both in the absolute and relative sense. That is, both the absolute width of sc1 and the relative width sc1/c, also called the relative balance, decrease with decreasing L. The relative balance (sc1/c) decreases from 0.38 to 0.28 (where 0.5 would correspond to perfect balance) over the domain of RI solutions. This corresponds to only about a 10% change in difference of semicell sizes (i.e. from approximately a 60–40 split to a 70–30 split).

Data on number of ciliary rows of doublets and relative balance (Table 1, Nelsen & Frankel, 1986) shows great variability, leading the authors to conclude that there is “no clear association between average total number of ciliary rows and mean imbalance”. Nevertheless, for the clone (subclone No.4) which clearly has both the lowest average and lowest range of values of ciliary rows, the relative balance (sc1/c) also clearly takes on the lowest value: 0.34 ± 0.07. This value not only lies in the range of our lowest imbalance (0.28) but agrees with our finding that the smallest L (the smallest circumference or least number of ciliary rows) shows the smallest relative balance.

The matter of how cells become unbalanced and the correlation between balance and total ciliary row number was further addressed in another experiment on a convenient clone (Frankel & Nelsen, 1986). It is noted that doublets which start out balanced “tend to become increasingly unbalanced while simultaneously undergoing loss of ciliary rows in successive cell generations”. It appears that although rows are indeed being lost, the primary cause of imbalance is “cortical slippage” of oral meridians, wherein the new OA develops along a ciliary row other than the OM of the previous generation (rarely more than one row over). The question of whether there is any differential loss of ciliary rows in the two semicells leads to an examination of the correlation of imbalance with total number of ciliary rows. If there is differential loss in sc1, a positive correlation should appear between relative
balance and total number of rows. The experimental correlation found was 0.37, leading Frankel & Nelsen to conclude that although imbalance does exhibit a tendency to increase with decreasing number of rows, the contribution of differential loss of ciliary rows to imbalance of doublets is probably small relative to that of slippage.

We note that our differential reduction in sc1, noted above, leads to a positive correlation between relative balance and total ciliary row number, in agreement with the experimental assessment just discussed. However, we will now examine how our model differs in the interpretation of this correlation. To test the random slippage scheme, it is desirable to discover whether a random process could in principle lead to imbalanced semicells. For this purpose, Frankel & Nelsen invoked an Absorbing Markov Chain model of random slippage (states get 'absorbed' when the imbalance is so great that a transition to a singlet state occurs). Given certain assumptions, they concluded that it is at least plausible from the point of view of the Markov Chain model that random slippage could by itself have resulted in a decrease in semicell balance. The concept of random slippage has, however, a serious detraction, namely the circular logic of the scheme, as pointed out by the authors themselves. That is, the "compression of positional values through random slippages...[implies that] the positional value itself appears to be dependent upon a structure whose location it controls".

In our model, slippage does apparently occur, i.e. new OA's appear at different longitudes than old OA's (as evidenced by our semicell width data), but it is not random. Instead, slippage is determined by the energy function controlling the system of morphogenetic field values: the semicell widths change in such a way as to decrease the energy of the system. Hence our model provides insight into (the direction of) slippage (viz. the observed negative correlation of relative balance and total number of rows). In fact, we can go further than this and actually determine
how much of the correlation arises from overall loss of rows (which presumably does occur at random, i.e. a row has the same probability of being lost whether it's in sc1 or in sc2) and how much is due to the (energetically favored) shift of the boundary between sc1 and sc2 (the slippage).

We model our problem of the overall random loss of rows in two semicells, with each row having the same probability of being lost, in analogy to a one dimensional random walk of fixed stepsize (see Appendix G). This model tells us what semicell widths (number of rows) would be expected, on average, from a purely random decrease, starting with balanced semicells (each with $M$ rows) and ending with the total number of rows $N < 2M$.

The expected average relative balance if cell shrinkage involved only random loss of ciliary rows, $0.461 \pm 0.005$ for our model with $M = 32$, $0.447 \pm 0.006$ for the experimental minimum number of ciliary rows in a normal singlet $M = 18$, is clearly larger (i.e. semicells more nearly equal) than the relative balance $0.33 \pm 0.04$ actually obtained by solutions of our field model. The balance calculated in the random loss scheme decreases (by 3% or 4% between $L = 3.250\pi$ and $2.875\pi$) with cell circumference, but not enough to account for the corresponding decrease in the model (30%). This comparison to a random loss scheme indicates that the semicell boundary does shift more than might easily be accounted for by a random loss of rows. Hence, the correlation between relative balance and total ciliary row numbers turns out to be primarily a result of the energy directed semicell boundary shifting.

A further indication that the boundary between the semicells shifts comes from the experimental observation that “the average width of sc2 of cells in unbalanced subclones was greater than one half of the largest number of ciliary rows found in any doublet”. This observation corresponds in our model to the result that the “singlet portion” (the approximately linear section covering $\theta = 0$ to $2\pi$ in the $\theta(x)$ graphs of Figs. 21 - 24) of the RI doublets extends in circumference beyond
2π (though not by much) – that is, for all RI solutions, it is wider than the normal singlet.

The general agreement exhibited between our model and the experiments regarding correlation in the widths sc1 and sc2 strongly suggests that the cell shrinkage ("crowding of positional values") probably occurs fairly uniformly, as was assumed in our field model. For it is especially in this kind of measurement – the precise spacing of markers about the circumference – that disagreement between the experiments and our model would be anticipated, were cell shrinkage actually highly nonuniform.

Moreover, when we compare the actual value of the average relative balance for RI solutions found in the model (0.33 ± 0.04), to the average relative balance of experimental doublets with pOM1 and pOM2 (0.33), the spacing predictions of the model are even more convincingly confirmed. With regard to this measurement, it is significant to note that since sc1 typically has ≈ 10 rows and sc2 ≈ 20 rows, then these doublets have an average of ≈ 30 rows, and hence are likely to be RI doublets rather than SYM(W=2) doublets10. Accordingly, the relative balance of SYM(W=2) doublets (i.e. 0.5) would not be expected to enter into this relative balance average.

The supposition that these doublets were RI doublets (i.e. "triplets" not expressing the latent sOA) is consistent with the experimental observation that "[certain] geometric factors did not differ significantly in doublets with sOA or without sOA" and the experimental indications (via studies of pre-division cells) that sOA expression is not stable.

Frequencies of Occurrence

The probability of obtaining a triplet vs. a singlet is proportional to the height in \( \theta(x) \) of the ‘RI bump’ on the \( \theta(x) \) field graphs (Figs. 21 – 24), since it is only when

---

10See example on page 57 for conversion of number of ciliary rows to \( L \) values.
the morphogenetic field value of the OA lies within this range (in other words, when the RI occurs near the OA field value) that triplets arise. The probabilities for getting triplets are shown for different $L$ values in Table 3.

Hence, for a population of cells undergoing RI at random locations about the circumference, we would expect a maximum of $\approx 20\%$ triplets to result. In a mixed sample with cell sizes from doublet to singlet, this figure should at first glance be reduced by a factor of the relative domain of RI solutions (fraction of all solutions from $L = 4\pi$, to $L = 2\pi$, that are RI) since not all cells will be in an RI configuration; however, there are likely more RI solutions in the experimental samples than might be guessed from a simple reading of the domain of RI solutions (Energy Diagram Fig. 13 and Table 1) for the following reason. The average (with very small variance) number of ciliary rows in a singlet, 18–21, should imply 36–42 rows in a complete doublet. But in fact, the mean number of ciliary rows seen in doublets is only in the high 20’s and low 30’s and in singlets within the same clone, around 25. This might mean that the experiments do not see $L = 4\pi$ or neighboring SYM(W=2) doublets, but begin at somewhat smaller $L$, though still within the domain of the SYM(W=2) solutions. Similarly, the experiments might not be pursuing singlets right down to the neighbourhood of the $L = 2\pi$ singlet size. The restricted domain of observation wrt. $L$ in the experiments, relative to the scope of the model, means that the effective domain of RI solutions might in practice (i.e. experimentally) seem relatively larger than the transition points in
the model would suggest. The effect this has on the (experimental) statistics of cell configurations is to weight disproportionately the contribution of RI states, relative to the SYM states (which exist only at the extreme $L$ values).

For example, if we take 20 ciliary rows for the normal singlet, 40 rows for the normal doublet (and translate between row numbers and $L$ values with the factor $2\pi/20$), then the transition points of $\text{SYM}(W=2) \rightarrow \text{RI} \rightarrow \text{SYM}(W=1)$ would occur at roughly 33 and 28 ciliary rows, respectively. By taking the experimental value of "mean number of ciliary rows" to represent our observed experimental range of sizes, we would conclude that our samples only cover a small domain of solutions, the largest cell being already in the RI region and the smallest only half way down the SYM($W=1$) domain. In this case then, all the doublets observed would be RI doublets, and more than half of the singlets would be SYM($W=1$) singlets.

The statistics on singlet/doublet/triplet formation should be compared with the experimental statistics. In general, the percentage of cells that were doublets was low (0%, 3%, 4%, 7%, 16% for different clones, Nelsen & Frankel, 1986) and the proportion of doublets with an sOA (i.e. triplets) was near or less than 20%. Note that in the model, all RI doublets contain an sOA and hence could be expressed as triplets whereas no SYM($W=2$) doublet will show an sOA. Hence we can conclude either that most doublets observed were SYM($W=2$) doublets, or contrariwise that they were RI doublets (as the previous example might suggest) but that developmental expression of sOA's is low\textsuperscript{11}.

**CVP Set Formation**

The predictions regarding CVP set formation in RI triplets in the model fit extremely well with the experimental findings for doublets/triplets (recall that triplets are called "doublets with sOA" in the experiments). Again we consider our RI triplets to correspond to (some subset of) experimental doublets, i.e. we hypo-

\textsuperscript{11}There is ample experimental evidence that expression of sOA's is much less stable than expression of pOA's.
esize that some experimental doublets are RI triplets with the sOA not expressed (and other experimental doublets are SYM(W=2) doublets). Experimental triplets are, of course, represented by fully expressed RI triplets. The three main theoretical predictions regarding CVP formation (in RI solutions) are:

1. Most triplets have 1 CVP set.

2. The CVP set is usually located in the wider semicell, sc2, at a fixed relative distance. (The fixed CM in sc2 results from the linear and nearly linear forms of $\theta(x)$ in sc2, for both SYM(W=2) and RI, respectively. See Figs. 20–25 and 21–24.)

3. The rare exception in which there is a CVP set in each semicell occurs when sc1 is especially large (Fig. 21). Our one RI solution ($L = 3.25\pi$, $\theta$ of OA $\approx$ 0.0) with a CVP in sc1, to the left of sOA, has a relative sc1 width$^{12}$ of 0.37. No smaller sc1 in the model contains a CVP. (Larger semicells exist, at $\beta^2$ = 0.2, but only for the SYM(W=2) doublet solutions and these indeed generally have 2 CVP sets.)

The corresponding experimental findings are virtually identical.

(1) Only a few doublets had a CVP in sc1 (in addition to the CVP set in sc2). CVP's were even rarer in sc's of doublets that showed sOA than in doublets without sOA, in accordance with our assumption that while experimental doublets are composed of SYM(W=2) (having therefore 2 CVP sets) and not fully expressed RI triplets (mostly 1 CVP), sOA doublets are all RI triplets. If there was also an sOA, the CVP appeared to its left (mean relative position of CVP1 in sc1 with sOA: 0.25).

$^{12}$An estimate of the significance of this model number can be obtained by considering the sensitivity of predicted semicell widths sc1 as the one free model parameter $\beta^2$ is changed. A calculation (see Appendix H) of the mean sc1 width and its standard deviation, for $L = 3.00\pi$, obtained by averaging over the range of $\beta^2$ values consistent with the Energy diagram transitions (i.e. $\beta^2 = 0.1$ - 0.4), gives $0.30 \pm 0.03$, or a standard deviation of 10%.
(2) Essentially all doublets were seen to have CVP's in the wider semicell sc2; the mean distance of CM2 from pOM2 was found to be linearly proportional to sc2, with no influence from sc1, nor from sOA in sc1 (i.e. the relative midpoint of CVP2 does not vary much). The actual value of the relative CVP2 set in sc2 ($\approx 0.19 \pm 0.05$), though not exactly the same as the average relative midpoint of CVP in singlets ($\approx 0.23 \pm 0.04$) is easily within one standard deviation, and agrees with a fixed CM as is predicted by the model\textsuperscript{13}.

(3) In doublets (without sOA), CVP's were rarely found in sc of less than 11 rows. Translating this number of rows into our measure gives $\frac{(11 \text{ row sc1})}{(30 \text{ row average total})} = 0.37$ as the relative minimum sc1 width for expected CVP formation, in exact agreement with the case found in our model. The extra CVP set in this case appears about at the midpoint of sc1, due to the symmetry of the RI solution (the 'bump' is parabolic). The experimental value ($\approx 0.3 \pm 0.1$) seems to suggest a slight asymmetry. This might be an indication that intercalation is actually not completely symmetric wrt. both winding directions\textsuperscript{14}.

The finding that “d1 was not proportional to sc1” indicates that sc1 is not contained in a linear region of $\theta(x)$ (cf. discussion on linear behaviour for SYM solutions). Nor is it in the model. This may be an additional indication of a directional asymmetry in the intercalation process\textsuperscript{15}.

6.3.3 Dynamical Pathways

RI Transitions

Biological time is incorporated in our model by the changing circumference $L$, i.e.

\textsuperscript{13}The model predicts an approximately fixed CM, but the actual value of CM was put into the model to translate from $\theta(x)$ to cell configurations.

\textsuperscript{14}See also recent evidence of “directional intercalation” in (Nielsen & Frankel, 1989).

\textsuperscript{15}Asymmetric intercalation would give rise to an asymmetric RI region in the $\theta(x)$ function. It might be incorporated into our model by the addition of a small asymmetric term in the energy functional.
we assume that the circumference \( L \) decreases as regulation/development proceeds. For any \( L \), there is a whole set of equivalent minimum energy solutions, differing only by a linear transformation, basically a rotation of the circumference (see degenerate solutions Fig. 7). Because we have considered only these equilibrium solutions (strictly, at fixed \( L \), we consider time independent field configurations) rather than the total sequential dynamical development of our system through all sizes \( L \), we cannot automatically connect an individual solution for one \( L \) value to a unique 'later time' solution at a different circumference \( L \); however, given the reasonable assumption that the complete dynamics would express some kind of 'inertia' to changes in the pattern, combined with our energy minimization principle, we can make interesting hypotheses as to the dynamical relationship between individual patterns (solutions) of different \( L \). In particular, we make the simple assumption that the field will change as little as possible in going from a lowest energy configuration at one value of \( L \) to a lowest energy configuration at another value of \( L \). This can be achieved by the almost linear portion (the segment corresponding to a normal singlet configured semicell) of the field remaining essentially unchanged as the circumference decreases through the entire RI domain. (Fig. 11 shows these fields of different \( L \) suitably matched up.) We will now follow the consequences of this hypothesized dynamical correspondence for several instances, over the RI solution domain (values of \( L \) which give RI as lowest energy solutions).

In Table 4 we have particular cell types listed as they go from \( L = 3.250\pi \) to \( 2.875\pi \). Each row entry corresponds to one sequence of development, taken in sequence from the four graphs of the RI field (Figs. 21 – 24. Follow one horizontal grid line along from graph to graph, i.e. from \( L \) value to \( L \) value).

The first example singlet, the one with morphogenetic field value of OA equal to \( 15 \cdot 2\pi/16 \) (topmost grid line, \( \theta(x) = 15 \cdot 2\pi/16 \), begins as a 3-CVP singlet (at \( L = 3.25\pi \), Fig. 21). Then, as it shrinks, the two CVP sets closest to the OA (R2)
merge \((L = 3.125\pi, \text{ Fig. 22})\) and it becomes a 2-CVP-set cell with a broad CVP domain. Subsequently, the broad CVP domain shrinks \((L = 3.00\pi, \text{ Fig. 23})\) until it is lost, leaving a 1-CVP-set singlet with an unusually large (i.e. far to the right of the OA) CM \((L = 2.875\pi, \text{ Fig. 24})\).

The second example singlet (second from top grid line, \(\theta(x) = 14 \cdot 2\pi/16\), on the graphs in Figs. 21 – 24) also begins as a 3-CVP-set cell, with greater spacing between the two sets closest to the OA, which subsequently \((L = 3.00\pi, \text{ Fig. 23; or } L = 2.875\pi, \text{ Fig. 24})\) merge into one broad CVP set, leaving a 2-CVP-set singlet. In the third example (third from top grid line, \(\theta(x) = 13 \cdot 2\pi/16\)), we begin with a 2-CVP-set singlet, the set farther from the OA broadened, and eventually shrinking to a normal width, ending with a 2-CVP-set singlet at \(L = 2.875\pi\) (Fig. 24). The last example (fourth from top grid line, \(\theta(x) = 12 \cdot 2\pi/16\)), initially is a 2-CVP-set singlet (Fig. 21), but the set farther from the OA is very narrow, likely containing only one CVP row. As \(L\) is reduced, this narrow set simply approaches the other set (moves left) and is still a 2-CVP-set cell at \(L = 2.875\pi\) (Fig. 24). The general behavior of the CVP sets in these examples is then, a gradual merging or shrinking, with movement towards the left and ending in a state with either fewer, or more compact, CVP sets.

A similar analysis can be done on triplet states, in which we now follow the
evolution of the OA's. See Table 5. In our first example (third from bottom grid

<table>
<thead>
<tr>
<th>$\theta$ of OA</th>
<th>$L = 3.250\pi$</th>
<th>$L = 3.125\pi$</th>
<th>$L = 3.000\pi$</th>
<th>$L = 2.875\pi$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3.0 \cdot 2\pi/16$</td>
<td>triplet → singlet → singlet → singlet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$2.4 \cdot 2\pi/16$</td>
<td>triplet → triplet → triplet → singlet</td>
<td></td>
<td></td>
<td>(fusing)</td>
</tr>
<tr>
<td>$2.0 \cdot 2\pi/16$</td>
<td>triplet → triplet → triplet → singlet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$1.0 \cdot 2\pi/16$</td>
<td>triplet → triplet → triplet → triplet</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$0.0 \cdot 2\pi/16$</td>
<td>triplet* → triplet → triplet → triplet</td>
<td>(fusing)</td>
<td>(fusing)</td>
<td>(fusing)</td>
</tr>
</tbody>
</table>

Table 5: Dynamical Pathways of 1-CVP-Set Cells. * This triplet has 2 CVP sets.

line, $\theta(x)=3 \cdot 2\pi/16$), at $L = 3.25\pi$ (Fig. 21) there is a triplet with sOA ($L$) closer to the pOA ($R1$) on its left. In the next step ($L = 3.125\pi$, Fig. 22), the sOA ($L$) and its nearby pOA ($R1$) have vanished, leaving only a singlet. In the second example (second from bottom grid line, $\theta(x)=2 \cdot 2\pi/16$), the sOA ($L$) also approaches the pOA ($R1$) on its left, but this time the singlet does not appear until $L = 2.875\pi$ (Fig. 24). The third example (first from the bottom grid line, $\theta(x)=1 \cdot 2\pi/16$) shows an sOA ($L$) closer to the pOA on the right (i.e. $L$ is closer to $R2$ than to $R1$). Subsequent cell shrinkage does not bring the sOA much closer to a pOA and the cell remains a triplet even at $L = 2.875\pi$ (Fig. 24) but it does end up closer to the other pOA (i.e. $L$ is now closer to $R2$ than to $R1$). In this case, the triplet state has remained over the entire RI domain and a singlet will not be formed until the transition from RI → SYM(W=1). In the last example explicitly marked on the graphs (bottom line, $\theta(x)=0$) we see a fused $L$ (sOA) and $R$ (pOA). In addition, there is a second CVP set, but only when the narrow semicell (sc1) is at its widest.

Development in this fused triplet consists mainly of the abnormal OA ($L/R$) moving in closer to the normal OA ($R$)$^{16}$, and the rapid loss of the second CVP set. By examining intermediate $\theta(x)=constant$ lines$^{17}$ not already drawn on the graphs,

$^{16}$Movement to the left—as in above examples— or right is possible: it is specific to the representation.

$^{17}$Or by interpolating more $L$ value curves.
such as (Table 5) $\theta(x) = 2.4 \cdot 2\pi/16$, we can readily see that for the cases where a $L$ (sOA) and $R$ (pOA) approach, and a triplet converts eventually to a RI singlet, there will be a brief point in this gradual approach that represents a fusion of $L$ and $R$.

In conclusion, the usual fate for triplets is approach and fusion of the $L$ with a $R$ (either one); however, there are a few exceptional scenarios, namely the one in which a fused $L/R$ is maintained throughout the whole RI domain, and the one in which the $L$ is nearly equidistant from either $R$ and remains so, not undergoing fusion. It should also be apparent that the typical fused $L/R$ state is usually a short-lived state (with the one exception noted above), which may help to account for its relative rarity in the experimental samples.

**Biventral Doublets and Two-RI Events**

The appearance in experiments (Nelsen & Frankel, 1986) of rare "biventral doublet" configurations suggests that another regulatory pathway, in addition to the one just described, is operative. A biventral doublet has one pOA and one sOA, with 2 CVP sets between them, one to the right of the pOA, the other to the left of the sOA. According to the reverse-intercalation model, conversion of typical doublet to biventral doublet requires two RI events, thereby reducing the winding number from 2 to 0, and resulting in a bilaterally symmetrical organization of the cell. The reverse-intercalation model cannot account for the instability of the biventral doublet configurations, which seem to revert to 2-CVP-set singlet configurations:

> positional values can be lost by incremental intercalation across the two longitudes of reversal ... [which] will always preserve the biventrically symmetrical organization.

In our model, such biventals with two RI boundaries would generally have higher energies (each reversal boundary accumulates an additional energy contribution in
the second term of the energy functional, the $\beta^2$ term). If biventrals (two RI events) did appear in our model, they would be expected for larger $L$, where there is more 'space to fit' the RI regions. This accords with the experimental observation that biventrals (two RI's) are approximately the size of doublets whereas 2-CVP-set singlets (one RI) are nearer to the typical size of singlets.

\footnote{In contrast, according to the reverse-intercalation model, which does not include an effect analogous to the $\beta^2$ term, even a cell with many alternating regions of positional winding directions could be stable.}

\footnote{Though our model could have two-RI solutions with higher energies, we have not yet classified their forms and energies.}
7 Summary and Conclusions

In this work we have probed into a possible mechanism for pattern formation, particularly relevant to biological systems. We have seen the translation and extension of an important biological concept, positional information or pattern specification into a rigorous and explicit physical formalism: a vector field (e.g. order parameter field) governed by a non-linear Ginzburg–Landau type energy density functional.

The patterns were represented by a distribution or field, the *morphogenetic field*, defined over the surface where pattern formation was observed to occur. We conceived this field to be conceptually similar to physical fields, and defined it in the simplest way that would allow it 1) to represent uniquely all circumferential positions around a 1-dim ring, and 2) to have the continuous character of physical fields. Generating dynamics for the field were derived from an energy density based on local field interactions. We chose the simplest possible interaction scheme that would reflect 1) the basic biological tendency for an even (homogeneous) spacing of pattern features (i.e. regulation towards an 'optimal spacing of positional values', which translated as 'optimal gradient of the vector field'), and 2) the basic tendency of physical fields to be 'smooth'. This behavior was modeled by an energy density functional in the form of a double well (quartic in the field gradient) with minima at the values of the optimal field gradients, including also a higher derivative term, representing the 'inertia' of the field to changing winding direction (This higher derivative term is essential in order for there to be the appropriate RI regions over a range of circumference lengths. The non-linear term was essential in order for there to be $\phi \neq constant$ equilibrium solutions at all relevant circumference lengths.). Our model had one free parameter ($\beta^2$) to reflect the relative weight of these two factors.

The model was solved by finding the minimum energy configurations of the field,
under periodic boundary conditions (the field must regain its initial value after one
transit around the circumference). Two main classes of solutions were discovered,
with transitions occurring between them as the system size is changed: the symmet-
ric (SYM) and the reverse intercalation (RI) solutions. The generation of solutions
corresponding to configurations with a region of RI, which may appear anywhere
about the circumference, is our most significant result. The one free parameter
of the model, $\beta^2$, determines the circumference lengths at which the transitions
between the two classes of solutions, $\text{SYM}(W=2) \rightarrow \text{RI} \rightarrow \text{SYM}(W=1)$ occur.

In the application to pattern formation in *Tetrahymena*, the symmetric solu-
tions (SYM) corresponded to the ‘normal’ states of *Tetrahymena*: 1) the normal
*singlet*, with one complete set of pattern features, and 2) the fusion of two singlets
into a *doublet*, being twice as large and having two complete and similarly aligned
sets of features. Doublets are observed experimentally to reduce in size and event-
tually regulate back to a singlet pattern, during which process they occasionally
express an intermediary *triplet* state with some pattern features appearing in trip-
licate, but with reversed symmetry or handedness. At system sizes intermediate
between doublet and singlet, our model undergoes transitions to the dramatically
different reverse intercalation (RI) configurations. The existence of this domain of
RI solutions is crucial for understanding the intermediate triplets (and other con-
figurations). Our quantitative model allows not only detailed comparison with the
observed *Tetrahymena* patterns but suggests dynamical ‘pathways’ for the regula-
tion of individuals. In particular, our model not only accounts for the circumferen-
tial location of the cell markers (oral apparatuses and the CVP sets) but predicts
and gives a rational for a multiplicity of regulatory pathways: doublets transform-
ing directly to singlets, doublets passing through triplet configurations and then
to singlets, rapidly changing triplets, triplets that appear ‘suddenly’ (like a phase
transition) and then remain essentially unaltered until their eventual conversion to
singlet, cells developing extra CVP sets, merging CVP sets, disappearance of CVP sets, etc..

One of the more profound insights offered by our model refers to the relationship between observed biological states and underlying field configurations (reflecting the presumptive physical processes). Similar observed states may arise from fundamentally different field solutions, and conversely, one type of field solution can give rise to very different biological patterns. For example, the observed 1-CVP-set singlets can result from either a symmetric solution or a RI solution; conversely, the RI field solution can give rise to triplets, 1-CVP-set singlets, and singlets with 3 CVP sets. The lesson to be learned here is that ‘taxonomy’ based only on the superficial expressed patterns may not provide sufficiently complete insight into the fundamental underlying processes or the dynamical relationships between patterns. In particular, the fundamental classes of patterns in this problem may turn out not to be ‘singlets’ and ‘doublets’, nor ‘balanced’ and ‘unbalanced’ doublets, but something more closely related to the symmetry of the field solutions (SYM(W=1) and SYM(W=2), and RI)\(^{20}\). Our model suggests new relationships between patterns, with a simple underlying logic but a vast diversity of observables.

On the one hand, many further experimental tests of this model suggest themselves. More confirmation of the details of our predicted patterns (such as 3-CVP-set singlets) could be undertaken. Furthermore, continued analysis of the statistical features of the patterns in doublets and their derivatives, but grouped according to the basic classifications of our model (the symmetry classes SYM and RI), would be valuable. Particular predictions such as the location of the transition points between SYM and RI, as related to the size of doublets and singlets could be addressed. In addition, two general directions for experimental work are suggested

\(^{20}\)cf. In physics we have numerous examples of this difference between ‘fundamental’ classes and ‘observable’ classes, e.g. in quantum mechanics the observable states of a system may be linear combinations of the eigenstates, such as elementary particles having observable states which are “mixtures” of the basic quarks.
by our model. One involves discovering or creating additional cell markers, so that more of the circumferential patterns can be charted or interpolated. The other suggested direction, likely requiring new experimental approaches and methods, is to follow individual cells, to ascertain the dynamical connection between successive cell configurations (generations), such as the abrupt change in positions of cell features through the two transitions.

On the other hand, much experimental work already done, both on *Tetrahymena* and other developmental systems, could be analysed by our model. For example, the *Tetrahymena* mutant *janus* form and its doublet relatives (Frankel & Nelsen, 1987) would be a very interesting application, or the regulation of left-handed cells and their doublets could be studied. Recent experiments (Nelsen & Frankel, 1989) on left-handed cells have indicated that there might asymmetry between winding directions (i.e. "directional intercalation"), a feature which might be incorporated into our model (e.g. by including a cubic term in the energy functional, to break the symmetry of our field interaction).

We find that our morphogenetic field model has been very fruitful, in offering a 'simple' physical explanation for the enigmatic patterns of *Tetrahymena*, in suggesting new directions - both experimental and theoretical - for further studies of pattern formation in this system. In addition, our very broad interpretation of positional information, viz. as a general means of distinguishing locations in an ordered system, suggests its widespread applicability, not only to pattern formation problems of direct biological relevance but also to patterns in purely physical systems. The difference between pattern specification or positional information in biological vs. physical systems is primarily in its expression\(^{21}\), e.g. the physical or chemical complexity of the fields and the substances in which the patterns are

\(^{21}\)Pattern specification information is often interpreted differently in biological systems than in physical systems, e.g., it is seen in the context of a surviving organism as a 'means' to attaining an 'end'.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
manifested. Biological systems have more complex parts in which to show the effects of positional ordering (e.g. the genetic switching of each cell in Wolpert's (1969) positional information scheme) but the underlying general principles of local activity leading to global ordering may be identical, whether they be, for example, magnetic particles interacting electromagnetically in a ferromagnet, or cells in tissue interacting by series of complex physiological means.

The proof of this argument lies in developing explicit models of generalized morphogenetic information for pattern formation problems, and testing them in individual systems, both physical and biological, such as we have done for Tetrahymena. The success of our program argues strongly for the value of developing general high-level models of biological processes, in addition to the detailed molecular approaches typically undertaken. While the molecular approaches might be imagined to supply a mechanism to support positional information and morphogenetic fields, they generally do not by themselves have the power (the tractability) at such a microscopic level to make the necessary high-level predictions for relatively large-scale patterns. We view the success of our model as greatly supportive to the ambition of bringing about a unification of the formalism, phenomenology and concepts of physical theory with the foundations of theory in biology.
A Derivation of Euler–Lagrange Equation

Many equations of physics can be stated either in terms of a differential equation or as an equivalent variational principle. Classical mechanics provides an important example. Hamilton’s variation principle

\[ \delta \int_{t_1}^{t_2} L dt = 0 \]  \hspace{1cm} (10)

is equivalent to Lagrange’s equations of motion

\[ \frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \]  \hspace{1cm} (11)

where the Lagrangian \( L \) is a function of the generalized coordinates \( q_i \) and velocities \( \dot{q}_i \) and the time. Lagrange’s equations are equivalent to Newton’s laws of motion.

The calculus of variations (see e.g. Corben & Stehle, 1974; Gelfand & Fomin, 1963; Weinstock, 1974) tells us how to convert between these two statements of physical laws. We shall consider the basic problem of the calculus of variations as it applies to our model, i.e. we shall determine a differential equation with the appropriate boundary conditions that is equivalent to our original energy functional representation.

Consider a function \( F(y, y', x) \) where \( y' \equiv \partial y/\partial x \), and the integral

\[ E = \int_a^b F(y, y', x) dx = E[y(x)]. \]  \hspace{1cm} (12)

We want to find the function \( y(x) \) which minimizes the functional \( E[y(x)] \) (in general, which makes \( E[y(x)] \) stationary). That is, we want to find a \( y(x) \) such that if we replace \( y(x) \) by a ‘nearby’ function, the integral \( E \) is unchanged to first order (the variation in \( E \) must vanish). We consider a ‘nearby’ function which takes on the same values at the endpoints \( a \) and \( b \) as \( F(y, y', x) \), and has integral \( E_2 \). It can then be shown that in order for the difference \( \Delta E \equiv E_2 - E \) (i.e. the variation in \( E \)) to vanish for arbitrary changes in \( y(x) \), we require

\[ \frac{\partial F}{\partial y} - \frac{d}{dx} \frac{\partial F}{\partial y'} = 0. \]  \hspace{1cm} (13)
This is the Euler–Lagrange equation.

When the Lagrangian is independent of a coordinate, there will be an integral of the motion, i.e. a conservation law. In the case where \( F \) does not depend explicitly on \( x \) (i.e. \( \partial F/\partial x = 0 \)), then \( x \) is an 'ignorable' or 'cyclic' coordinate and the Euler–Lagrange equations can be reduced to a first order equation

\[
F - y' \frac{\partial F}{\partial y'} = \text{constant.}
\] (14)

The more general variational problem, that of minimizing a functional

\[
E[y(x)] = \int_a^b F(y, y', x)dx
\] (15)

subject to the constraint that another functional

\[
J[y(x)] = \int_a^b G(y, y', x)dx = J_0
\] (16)

be held constant, can be treated by using Lagrange multipliers. The problem of finding \( y(x) \) which minimizes \( E[y(x)] \) under variations \( \delta y(x) \) that are restricted by \( J[y(x)] = J_0 = \text{constant} \), is exactly equivalent to the problem of finding \( y(x) \) which minimizes

\[
E[y(x)] + \lambda J[y(x)]
\] (17)

under these restricted variations, with \( \lambda \) being any constant (the Lagrange multiplier). The function \( y(x, \lambda) \) which minimizes \( E + \lambda J \) for arbitrary \( \delta y \) is given by the Euler–Lagrange equation for \( E + \lambda J \):

\[
\frac{\partial(F + \lambda G)}{\partial y} - \frac{d}{dx} \frac{\partial(F + \lambda G)}{\partial y'} = 0.
\] (18)

If a value of \( \lambda \) can be found so that the solution \( y(x, \lambda) \) satisfies \( J = J_0 \), then this \( y(x) \) also minimizes \( E + \lambda J \) under the restricted variations, and is the solution to the constrained problem.

Turning now to the specific case of our model, we note that since we are seeking solutions of the difference field \( \phi(x) \) (corresponding to \( y(x) \) above) which have
both positive and negative regions (such is the meaning of a reverse intercalation solution), there must be a point \(x_0\) where \(\phi(x_0) = 0\). We choose our coordinate origins so that \(\phi(0) = 0\). The periodicity required of \(\phi(x)\) implies also that \(\phi(L) = \phi(0) = 0\), where \(L\) is the circumference (length) of the system, so that we can consider the endpoints as fixed in our variational problem. Our periodic condition on the field \(\theta(x)\) translates to an integral condition on the difference field \(\phi(x)\), \(\int_0^L \phi(x) dx = 2\pi W\), where \(W\) is the winding number, and will be taken as the constraint condition for the minimization of our energy functional

\[
E = \int_0^L \left\{ (\phi^2 - 1)^2 + \beta^2 \left( \frac{\partial \phi}{\partial x} \right)^2 \right\} dx. \tag{19}
\]

(Without the constraint, minimizing \(E[\phi(x)]\) would give the trivial solution \(\phi(x) = \pm 1\) everywhere, \(E = 0\)). We now calculate the terms in our Euler–Lagrange equation.

\[
F \equiv (\phi^2 - 1)^2 + \beta^2 \phi'^2 = \phi^4 - 2\phi^2 + 1 + \beta^2 \phi'^2
\]

\[
G = \phi \quad \tag{20}
\]

Hence the second order Euler–Lagrange equation is

\[
\phi^3 - \phi + \frac{\lambda}{4} - \frac{\beta^2}{2} \phi'' = 0. \tag{22}
\]

The first order Euler–Lagrange equation is

\[
(\phi^2 - 1)^2 - \beta^2 \phi'^2 + \lambda \phi = c \tag{23}
\]

or

\[
\phi^4 - 2\phi^2 + \lambda \phi - \beta^2 \phi'^2 = c - 1. \tag{24}
\]

(The first order equation can also be obtained directly from the second order equation by integrating with the integrating factor \(\phi'\)).
The meaning of $c$ (it is a 'constant of the motion') can be seen by considering the endpoint $x = 0$ where $\phi(0) = 0$, which gives from the Euler–Lagrange equation,

$$c - 1 = -\beta^2 \left( \frac{\partial \phi}{\partial x} \right)^2 \bigg|_{x=0}. \quad (25)$$

That is, the constant $c$ is related to the slope of the solution $\phi(x)$ at the point $x = 0$, and in general does not vanish.

The meaning of $\lambda$ can be seen by considering the endpoint $x = 0$ in the second order Euler–Lagrange equation,

$$\lambda = 2\beta^2 \frac{\partial^2 \phi}{\partial x^2} \bigg|_{x=0}. \quad (26)$$

The multiplier $\lambda$ can also be considered simply as an amplitude for $\phi(x)$ since $\lambda$ is determined, in principle, from the condition

$$\lambda \int_0^L \phi(x) dx = 2\pi W. \quad (27)$$

### B Derivation of the Dynamical Field Equation

We now show how a dynamical field equation with explicit time dependence can be derived (e.g. Gelfand & Fomin, 1963) and that its equilibrium form is the same as the Euler–Lagrange equation for this system. (The Euler–Lagrange equation (Appendix A) gives the minimum energy ($E$) solution, $\phi(x)$, for a given $\beta^2$, $L$. Time does not appear explicitly in that equation because it is the equilibrium solution.)

We begin with the energy functional and the constraint:

$$E = \int_0^L \left\{ (\phi^2 - 1)^2 + \beta^2 \left( \frac{\partial \phi}{\partial x} \right)^2 \right\} dx \quad (28)$$

$$J = \int_0^L \phi dx = 2\pi W. \quad (29)$$
We can ensure a constrained minimization by incorporating the constraint directly into the functional to be minimized via a Lagrange multiplier:

\[
H = E + \lambda J
\]

\[
= \int_0^L \left\{ (\phi^2 - 1)^2 + \beta^2 \left( \frac{\partial \phi}{\partial x} \right)^2 + \lambda \phi \right\} dx
\]

\[
= \int_0^L h(\phi, \phi_x, x) dx.
\]

The dynamics will arise from the requirement that changes in the field \( \phi(x, t) \) with time must decrease the energy \( H \). The change \( \delta H \) in the energy functional, resulting from some change \( \delta \phi \) in the field, can be expressed symbolically in terms of the functional derivative \( \frac{\delta H}{\delta \phi} \) of the functional \( H \):

\[
\delta H = \int \frac{\delta H}{\delta \phi} \delta \phi \, dx.
\]

Time enters the equations explicitly as we now associate a change in the field \( \delta \phi \) (at \( x \)) with a change in the time \( \delta t \):

\[
\delta \phi = \frac{\partial \phi}{\partial t} \delta t.
\]

So the change in \( H \) takes the form:

\[
\delta H = \int \frac{\delta H}{\delta \phi} \frac{\partial \phi}{\partial t} \delta t \, dx.
\]

If we choose the field equations

\[
\frac{\partial \phi}{\partial t} = -\frac{\delta H}{\delta \phi},
\]

we can ensure that the energy will decrease with time:

\[
\delta H = -\int_0^L \left( \frac{\delta H}{\delta \phi} \right)^2 \delta t \, dx < 0 \text{ for } \delta t > 0.
\]

Expressing the functional derivative \( \frac{\delta H}{\delta \phi} \) explicitly,

\[
\frac{\delta H}{\delta \phi} = \frac{\partial h}{\partial \phi} - \frac{\partial}{\partial \phi} \left( \frac{\partial h}{\partial \phi_x} \right)
\]

\[
= 4 \left( \phi^3 - \phi + \frac{\lambda}{4} - \frac{\beta^2}{2} \frac{\partial^2 \phi}{\partial x^2} \right),
\]
gives the field equation
\[
-\frac{1}{4} \frac{\partial \phi}{\partial t} = \phi^3 - \phi + \frac{\lambda}{4} - \frac{\beta^2}{2} \frac{\partial^2 \phi}{\partial x^2}.
\]  
(40)

At equilibrium, $\partial\phi/\partial t = 0$ and we get the same equilibrium field equations as before:
\[
0 = \phi^3 - \phi + \frac{\lambda}{4} - \frac{\beta^2}{2} \frac{\partial^2 \phi}{\partial x^2}.
\]  
(41)

Multiplying through by $\frac{\partial \phi}{\partial x}$ and integrating gives the first order form.

C Elliptic Solutions

Starting with the first order Euler-Lagrange equation for $\phi(x)$, we seek an analytic solution.

\[
c - 1 = \phi^4 - 2\phi^2 + \lambda\phi - \beta^2 \left(\frac{d\phi}{dx}\right)^2
\]  
(42)

\[
\beta \frac{d\phi}{dx} = \pm(\phi^4 - 2\phi^2 + \lambda\phi + 1 - c)^{1/2}
\]  
(43)

\[
dx = \pm\beta \frac{d\phi}{(\phi^4 - 2\phi^2 + \lambda\phi + 1 - c)^{1/2}}
\]  
(44)

\[
x = \pm\beta \int_{\phi(0)}^{\phi(x)} \frac{d\phi}{(\phi^4 - 2\phi^2 + \lambda\phi + 1 - c)^{1/2}} \quad \text{(since $\phi(0) = 0$)}
\]  
(45)

In order to carry out this integration ($x$ is an Elliptic Integral of the First Kind), and express $\phi(x)$ explicitly, we need to know the values of $\lambda$ and $c$; however, to determine $\lambda$, we need to know $\phi(x, \lambda)$ to solve the constraint equation. A similar problem has been undertaken (the one-dimensional equilibrium solutions of the unconstrained Cahn–Hilliard equation, Novick–Cohen & Segel, 1984) using numerical and analytical approximations to incorporate the constraint (composition conservation). Because of the complexity of this approach, which can only provide approximate (numerical) solutions anyhow, we devised an entirely different method, based directly on the energy functional. A notable success of our method is that not only
is it simpler to execute (relying as it does on computational subroutines available in
standard scientific libraries) but our numerical solutions \( \phi(x) \) are at least as precise
as the corresponding curves obtained by this other method (cf. Novick–Cohen &
Segel, 1984, Fig. 3). (For references on elliptic functions, see Bowman, 1953; Byrd
1954; or a reference text on mathematical functions for physicists.).

D  Tanh Interface

From numerical solutions (the RI solutions), we find that \( \phi(x) \) has a maximum near
or at the upper saturation level (‘saturation’ implies that further change in \( \phi \) does
not reduce the energy \( E \)), i.e. \( \phi_{\text{max}} \simeq 1.0 \). There is also a minimum of \( \phi(x) \) near
the lower saturation level, i.e. \( \phi_{\text{min}} \simeq -1.0 \). At an extremum (min or max) the
derivative of \( \phi(x) \) vanishes and the first order differential equation (Euler–Lagrange
equation) for \( \phi(x) \) gives:

\[
(\phi_{\text{max}}^2 - 1)^2 + \lambda \phi_{\text{max}} = c .
\]

Hence

\[
\lambda \approx c .
\]

Considering now the first order differential equation near \( \phi(0) = 0 \):

\[
c = 1 - \beta^2 \left( \frac{d\phi}{dx} \right)_{x=0}^2 \Rightarrow c \leq 1 .
\]

We can gain more information about the value of \( c \) by using an approximate discrete
form for the derivative,

\[
\frac{d\phi}{dx} \simeq \frac{\phi_i - \phi_{i-1}}{x_i - x_{i-1}} ,
\]

in the above expression for \( c \), and determining several \( c \) values (and hence also \( \lambda \))
from the numerical solutions for \( \phi_i \). It seems that \( c \) is small (\( c \leq 0.1 \)). Hence it
is reasonable to consider the analytic solution for \( \phi(x) \), \( c = \lambda = 0 \), as a sort of
asymptotic approximation for \( \phi(x) \). (See Byrd 1954, or Bowman 1953, for elliptic
functions and integrals). Beginning with the expression for $x$ from section C

$$x = \pm \beta \int_0^{\phi(x)} \frac{d\phi}{(\phi^4 - 2\phi^2 + \lambda \phi + 1 - c)^{1/2}}$$ (50)

$$\pm \frac{x}{\beta} \approx \int_0^{\phi(x)} \frac{d\phi}{(\phi^4 - 2\phi^2 + 1)^{1/2}}$$ (51)

$$= \int_0^{\phi(x)} \frac{d\phi}{(\phi^2 - 1)^{1/2}(\phi^2 - 1)^{1/2}}$$ (52)

$$= \int_0^{\phi(x)} \frac{d\phi}{(1 - \phi^2)^{1/2}(1 - \phi^2)^{1/2}}$$ (53)

$$= F(\sin^{-1} \phi(x), 1) \text{ (Elliptic Integrals of First Kind, mod } k^2 = 1)$$ (54)

$$= \text{sn}^{-1} (\phi(x), 1) \text{ (Jacobi Elliptic Function)}$$ (55)

$$= \tanh^{-1} \phi$$ (56)

$$\phi(x) = \tanh \left( \pm \frac{x}{\beta} \right).$$ (58)

While this $\phi(x)$ does not meet the periodic boundary condition, we might nevertheless expect it to be approximately indicative of the shape of $\phi(x)$ at the $\phi > 0$, $\phi < 0$ interface (the reversal region).

E Analytic SYM Solution

To find the solutions $\phi(x) = \text{constant}$ (i.e. $d\phi/dx = d^2\phi/dx^2 = 0$) we need to work directly from the energy functional, since the variational method is not meaningful for integrands without derivatives.

$$E = \int_0^L \left\{ (\phi^2 - 1)^2 + \beta^2 \left( \frac{d\phi}{dx} \right)^2 \right\} dx \Rightarrow$$ (59)

$$E = \int_0^L (\phi^2 - 1)^2 dx \text{ for } \frac{d\phi}{dx} = 0.$$ (60)

Note also that the parameter $\beta^2$ plays no part in these SYM solutions. The constraint equation is now sufficient to give the solution $\phi(x)$:

$$\int_0^L \phi(x) \, dx = \phi \int_0^L \, dx = 2\pi W \Rightarrow$$ (61)
\[ \phi = \frac{2\pi W}{L}, \quad W = 1, 2. \]  

(62)

The energy is found analytically from the \( E \) functional:

\[ E = \int_0^L \left( \left( \frac{2\pi W}{L} \right)^2 - 1 \right)^2 dx = \left( \left( \frac{2\pi W}{L} \right)^2 - 1 \right)^2 L. \]  

(63)

\( E(L) \) for both \( \text{SYM}(W=1) \) and \( \text{SYM}(W=2) \) is plotted in Fig. 13.

F  The Rotational Symmetry of SYM

The rotational symmetry of the SYM solution (which gives it that name) consists of a product of rotations, one in real space, one in morphogenetic space, which leaves the solution invariant.

A transformation \( \hat{r}_x \) on the space variable \( x \) (denoting position on the circumference) is performed, followed by a transformation \( \hat{R}_\theta \) on the morphogenetic field. We define a rotational invariance to exist if there is a combination of \( \hat{r}_x \) and \( \hat{R}_\theta \) for all \( x \), that leaves \( \theta(x) \) invariant. That is,

\[ \hat{R}_\theta \hat{r}_x \theta(x) = \theta(x) \quad \text{for some } \hat{R}_\theta, \hat{r}_x. \]  

(64)

Now, the field \( \theta(x) \) is defined as an integral of the gradient field \( \phi(x) \):

\[ \theta(x) = \theta(0) + \int_0^x \phi(x') dx'. \]  

(65)

\[ \theta(x + a) = \theta(0) + \int_0^{x+a} \phi(x') dx'. \]  

(66)

For the SYM solution: \( \phi(x') \equiv \phi = \text{constant} \); hence,

\[ \theta(x) = \theta(0) + \int_0^x \phi dx' \]  

(67)

\[ = \theta(0) + \phi \int_0^x dx' \]  

(68)

\[ = \theta(0) + \phi \cdot x \]  

(69)

and hence

\[ \theta(x + a) = \theta(0) + \phi \cdot x + \phi \cdot a. \]  

(70)
We define the action of the rotation operators \( \hat{R}_\theta, \hat{r}_x \): 

\[
\hat{R}_\theta \theta(x) \equiv \theta(x) + \mathcal{R} = \theta(0) + \int_0^x \phi(x')dx' + \mathcal{R}.
\]

\[
\hat{r}_x \theta(x) \equiv \theta(x + x_r) = \theta(0) + \int_0^{x + x_r} \phi(x')dx' + \mathcal{R}.
\]

\[
\hat{R}_\theta \hat{r}_x \theta(x) = \hat{r}_x \theta(x) + \mathcal{R} = \theta(0) + \int_0^{x + x_r} \phi(x')dx' + \mathcal{R}.
\]

for SYM \( \Rightarrow \)

\[
\hat{R}_\theta \hat{r}_x \theta(x) = \theta(0) + \phi \cdot x + \phi \cdot x_r + \mathcal{R} = \theta(x) + \phi \cdot x_r + \mathcal{R}.
\]

We can choose \( \phi \cdot x_r + \mathcal{R} = 0 \) so that

\[
\mathcal{R} = -\phi \cdot x_r \quad \text{no} \ x-\text{dependence},
\]

giving

\[
\hat{R}_\theta \hat{r}_x \theta(x) = \theta(x)
\]

for all \( x \), for SYM solution.

For a general solution \( \theta(x) \) (not the SYM solution),

\[
\hat{R}_\theta \hat{r}_x \theta(x) = \theta(0) + \int_0^{x + x_r} \phi(x')dx' + \mathcal{R}
\]

\[
= \theta(0) + \int_0^z \phi(x')dx' + \int_z^{x_r} \phi(x')dx' + \mathcal{R}
\]

\[
= \theta(x) + \int_x^{x_r} \phi(x')dx' + \mathcal{R}.
\]

The integral term of this last expression has \( x \)-dependence, which means that in order for \( \hat{R}_\theta \hat{r}_x \theta(x) = \theta(x) \Rightarrow \mathcal{R} = \mathcal{R}(x) \) we would need an independent field rotation \( \mathcal{R} \) at each point in space \( x \).
Hence we have shown that the SYM solutions have a rotational symmetry that, in general, other solutions (such as RI) do not.

G  Random Walk and Relative Balance

Make a random 1-dim walk of \( n \) steps, with equal probability to move right or left. Denote by \( n_+, n_- \) the number of steps taken to the right and left, respectively, so that

\[
    n = n_+ + n_-.  \tag{84}
\]

The net distance moved from the starting point, \( D \), when the steps are of unit length is:

\[
    D = n_+ - n_.  \tag{85}
\]

The expected root mean square (rms) distance travelled after \( n \) steps is:

\[
    D_{\text{rms}} = \sqrt{\langle D^2 \rangle}  \tag{86}
\]

\[
    = \sqrt{n}.  \tag{87}
\]

Hence, on average:

\[
    n_+ = \frac{(n + D)}{2}  \tag{89}
\]

\[
    = \frac{n}{2} + \frac{\sqrt{n}}{2}  \tag{90}
\]

\[
    n_- = \frac{(n - D)}{2}  \tag{91}
\]

\[
    = \frac{n}{2} - \frac{\sqrt{n}}{2}.  \tag{92}
\]
We model shrinkage and the loss of ciliary rows as independent, random loss of rows from semicells sc1 and sc2. We identify $n_+$ as the number of rows lost from sc1, and $n_-$ as the number of rows lost from sc2.

Hence we find, beginning with initial size or number of rows $N = 2M$, and reducing to $N = FM$ (where $F = L/2\pi$) after a total loss of $n$ rows ($n = 2M - FM = M(2 - F)$), that there will be on average $\#sc1$ rows left in the semicell sc1:

$$\#sc1 = M - n_+$$

$$= \frac{1}{2}MF - \frac{1}{2}\sqrt{M(2 - F)}.$$  \hfill (94)

$$= \frac{1}{2}MF - \frac{1}{2}\sqrt{M(2 - F)}.$$  \hfill (95)

This gives a relative balance:

$$\#sc1/c = \frac{\frac{1}{2}MF - \frac{1}{2}\sqrt{M(2 - F)}}{MF}.$$  \hfill (97)

$$= \frac{1}{2} - \frac{1}{2}\sqrt{\frac{2 - F}{F}}\frac{1}{\sqrt{M}}.$$  \hfill (98)

Note that the relative balance expected for random loss of rows depends on the initial number of rows $M$ in a semicell. We give (Table 6) the relative balance for both $M = 32$ (the discretization used in our model) and $M = 18$ (corresponding to the experimental minimum number of ciliary rows in a normal singlet). For larger $M$, the relative balance approaches even closer to 0.5, i.e. the semicells become more nearly equal.

The relative balance found in our field model is less than either of the relative balance values arrived at by random loss above; in addition, the relative balance decreases with circumference faster than would be suggested by the random loss.
<table>
<thead>
<tr>
<th>$L$</th>
<th>M=32</th>
<th>M=18</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.250 π</td>
<td>0.467</td>
<td>0.456</td>
</tr>
<tr>
<td>3.125 π</td>
<td>0.463</td>
<td>0.450</td>
</tr>
<tr>
<td>3.000 π</td>
<td>0.458</td>
<td>0.444</td>
</tr>
<tr>
<td>2.875 π</td>
<td>0.454</td>
<td>0.439</td>
</tr>
<tr>
<td>average</td>
<td>0.461 ±0.005</td>
<td>0.447 ±0.006</td>
</tr>
<tr>
<td>% change</td>
<td>3%</td>
<td>4%</td>
</tr>
</tbody>
</table>

Table 6: Relative Balance Expected in a Random Walk Model of Ciliary Row Loss.

model. We therefore conclude that random loss is insufficient to account for either the imbalance observed in our field model, or in the experiments.

H Sensitivity of SC1 Width to $\beta^2$

The sensitivity of sc1 width to a change in the model parameter, $\beta^2$, is used as a measure of the significance of the relative balance in the model (see Table 7). Measurements (sc1) were made from $\theta(x)$ (Field Angle) vs. $x$ (Circumferential Position) graphs for several values of $\beta^2$. The average value of imbalance is $\text{sc1}/c = 0.30 \pm 0.03$ (a standard deviation of about 10%).

<table>
<thead>
<tr>
<th>$\beta^2$</th>
<th>sc1</th>
<th>c</th>
<th>sc1/c</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>3.0</td>
<td>12.0</td>
<td>0.25</td>
</tr>
<tr>
<td>0.2</td>
<td>3.5</td>
<td>12.0</td>
<td>0.29</td>
</tr>
<tr>
<td>0.3</td>
<td>3.8</td>
<td>12.0</td>
<td>0.32</td>
</tr>
<tr>
<td>0.4</td>
<td>3.9</td>
<td>12.0</td>
<td>0.33</td>
</tr>
</tbody>
</table>

Table 7: Semicell Widths and Imbalance for Different Values of $\beta^2$. For this table and calculation, $L = 3.00\pi$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
I Glossary of Abbreviations

I.1 Experimental Terms

OA  Oral Apparatus.

OP  Oral Primordium.

OM  Oral Meridian. Located at the right postoral ciliary row of the OA.

pOA Primary OA with normal internal organization.

sOA Secondary OA with reversed or indeterminate orientation.

p  Prefix for primary: pOA, pOP, pOM.

s  Prefix for secondary: sOA, sOP, sOM.

sc  Semicell. The region between two OM.

sc1 Narrower semicell.

sc2 Wider semicell.

CVP  Contractile vacuole pore (set).

CM  CVP midpoint: average of the positions of the CVP (not weighted by the number of CVP’s that are formed along a single row).

CM1 CM in sc1.

CM2 CM in sc2.

pOM1 OM at the left margin of sc1.

pOM2 OM at the left margin of sc2.

sr  Sector between pOM1 and sOM.
Circumference expressed in number of ciliary rows.

Number of ciliary row intervals between pOM and CM.

sc1/c Relative balance.

I.2 Notation from our Model

$\beta^2$ Parameter in energy functional: relative weight of the second term.

$L$ The circumference of the system. Varies between $4\pi$ (doublet) and $2\pi$ (singlet).

$N$ The discrete measure of $L$. Number of lattice points in the system. Varies between $2M$ (doublet) and $M$ (singlet).

$M$ Number of lattice points in a normal singlet. We used $M = 32$ for this model.

$F$ Relative circumference of system, $F = L/2\pi$. Doublet has $F = 2$, and singlet $F = 1$.

$R$ A right handed OA, or pOA. Also used to denote location of OA in the morphogenetic field graphs; neighboring values wind in positive sense.

$R_1$ OA in sc1.

$R_2$ OA in sc2.

$L$ A left handed OA, or sOA. Also used to denote location of OA in the morphogenetic field graphs; neighboring values wind in negative sense.

domain Refers to the $x$-axis of circumferential positions, $N$ values.

range Refers to the $\theta$-axis in $\theta(x)$ graphs (morphogenetic field values).
References


Goodwin, B.C. (1980). “Pattern Formation and its Regeneration in the Proto-


York.