

Modeling—A Primer

Version 2.0

A BioQUEST Collection Text Module by

William C. Wimsatt University of Chicago

Jeffrey C. Schank University of Chicago

User's Manual

A BioQUEST Library VII Online module published by the BioQUEST Curriculum Consortium

The BioQUEST Curriculum Consortium (1986) actively supports educators interested in the reform of undergraduate biology and engages in the collaborative development of curricula. We encourage the use of simulations, databases, and tools to construct learning environments where students are able to engage in activities like those of practicing scientists.

Email: bioquest@beloit.edu Website: <http://bioquest.org>

Editorial Staff

Editor:	John R. Jungck	Beloit College
Managing Editor:	Ethel D. Stanley	Beloit College, BioQUEST Curriculum Consortium
Associate Editors:	Sam Donovan	University of Pittsburgh
	Stephen Everse	University of Vermont
	Marion Fass	Beloit College
	Margaret Waterman	Southeast Missouri State University
	Ethel D. Stanley	Beloit College, BioQUEST Curriculum Consortium
Online Editor:	Amanda Everse	Beloit College, BioQUEST Curriculum Consortium
Editorial Assistant:	Sue Risseuw	Beloit College, BioQUEST Curriculum Consortium

Editorial Board

Ken Brown University of Technology, Sydney, AU Peter Lockhart Massey University, NZ Joyce Cadwallader St Mary of the Woods College Ed Louis The University of Nottingham, UK Eloise Carter Oxford College Claudia Neuhauser University of Minnesota Angelo Collins Knowles Science Teaching Foundation Patti Soderberg Conserve School Terry L. Derting Murray State University Daniel Udovic University of Oregon Roscoe Giles Boston University Rama Viswanathan Beloit College Louis Gross University of Tennessee-Knoxville Linda Weinland Edison College Yaffa Grossman Beloit College Anton Weisstein Truman University Raquel Holmes Boston University Richard Wilson (Emeritus) Rockhurst College Stacey Kiser Lane Community College William Wimsatt University of Chicago

*Copyright © 1993 -2006 by John N. Calley and John R. Jungck
All rights reserved.*

Copyright, Trademark, and License Acknowledgments Portions of the BioQUEST Library are copyrighted by Annenberg/CPB, Apple Computer Inc., Beloit College, Claris Corporation, Microsoft Corporation, and the authors of individually titled modules. All rights reserved. System 6, System 7, System 8, Mac OS 8, Finder, and SimpleText are trademarks of Apple Computer, Incorporated. HyperCard and HyperTalk, MultiFinder, QuickTime, Apple, Mac, Macintosh, Power Macintosh, LaserWriter, ImageWriter, and the Apple logo are registered trademarks of Apple Computer, Incorporated. Claris and HyperCard Player 2.1 are registered trademarks of Claris Corporation. Extend is a trademark of Imagine That, Incorporated. Adobe, Acrobat, and PageMaker are trademarks of Adobe Systems Incorporated. Microsoft, Windows, MS-DOS, and Windows NT are either registered trademarks or trademarks of Microsoft Corporation. Helvetica, Times, and Palatino are registered trademarks of Linotype-Hell. The BioQUEST Library and BioQUEST Curriculum Consortium are trademarks of Beloit College. Each BioQUEST module is a trademark of its respective institutions/authors. All other company and product names are trademarks or registered trademarks of their respective owners. Portions of some modules' software were created using Extender GrafPak™ by Invention Software Corporation. Some modules' software use the BioQUEST Toolkit licensed from Project BioQUEST.

Introductory Note:

Welcome to the Unfinished Land of Model-Building!

The software with this manual is quite stable, and contains mostly errors of omission rather than of commission. The manual itself is another matter. It has continued to grow like topsy, and will ultimately evolve into a book on strategies of model-building, criticism, and evaluation, using the software as a means of delivery, with exercises using it keyed to all of these activities. To our knowledge, there is as yet no text like this at any level, much less at a level which renders it approachable to students without far more math and/or programming experience than are the norm for non-hard-science majors. Although designed originally for a biology course, they have also been used to good advantage in two philosophy of science courses, and are now being tried out as part of the methodological part of the course in an upper division/graduate course on computer modeling in the biological sciences. Those who can should perhaps better write the programs to do these things themselves, my physicist friends tell me, but we don't want to disenfranchise those who can't. Perhaps we can motivate some of them to become those who can.

As it is, there are many things to explore in there pages and programs, but the exercises could be better arranged, and there is no indication of relative difficulty (or even of relative length) of the exercises. We have focused on getting in the relevant materials, rather than on their optimal arrangement, or even presentation. These two simulations presented here evolved from two different exercises used in one form or another for the last 16 years in Wimsatt's Biology Common Core course, BioSci 147: "Genetics in an Evolutionary Perspective," and they have not yet been sufficiently integrated. The index needs a lot of further work--and does not at present include references to many of the appendices. In the meantime, we welcome suggestions for what you found difficult, what you would like to see included, any additional exercises you have added, and the like.

The course in which these were used was a non-majors course which tended to attract a fair number of math, physics, chemistry, and economics majors because it used computer simulations, and a fair number of humanities majors because it had a lot of history. Needless to say, the distribution of interests and special talents was distinctly bimodal! Thus, there are exercises for students who have never had calculus (and didn't like what algebra they did have), and for students who have gone far beyond it. I hope that you have as much fun with this as we have.

We owe an enormous debt of thanks to the students over the years in Wimsatt's BioSci 147 for acting as guinea pigs for and contributors to many of the ideas found in this module--first implemented on HP-25 programmable calculators in 1976, and on Apple Macintoshes since 1985. Special thanks are due to Apple for their conception of a people-friendly machine which solved the problems of bimodal distribution mentioned above, for their friendly support (including a "Paris" prototype, when nothing else was fast enough), and to the 1990 BioSci 147 class, who used something approximating the versions of the programs and labs you see here, but without much of the multiple windows, text processing, and recording support of this version, and who put up with twice-weekly version updates near the end. The 1992 class had to live through the

comparable uncertainty (including suggestions that they not do certain labs until the following week so we had time to implement new things for them to try). Extra special thanks are due to Scott Franklin, whose bifurcation/accumulation program provided the idea testing and inspiration for the N vs r plot mode, to Michael Post, whose elegant "TwigMaker" program deserves to make it into the next version of the Model-Building exercises, and to Matt Ryan and other members of the BioSci 267 class, whose β -testing enthusiasm verged on the diabolical.

Note: the manual is written in Palatino, Times, Geneva, and sometimes Chicago fonts. Make sure that these screen fonts are on your computer if you wish to preserve formatting for reading or printing the manual. If you should decide to print out the manual, or lots of graphics, see the note in the "how to" section on the much greater speed of "Quickdraw" printers as compared to Postscript printers for "Quickdraw" graphics. (The difference in printing time can easily be a factor of 20, and some of the graphics in this manual may take as much as 15-20 minutes to print out on a slow Postscript printer.)

Editor's note: This document is also available in PDF format. In most cases you will probably find it quicker and easier to use the PDF format if you are reading the manual on a computer screen or if you want to print it. VGV 8/13/00

William C. Wimsatt
Jeffrey C. Schank
The University of Chicago
June 12, 1992

Table of Contents

What You Need To Know and General Overview	1
Introduction To Using Models in Biology	3
Model-Building and Exploration with the Minimum-Density Limited Logistic Growth Equation.....	13
Why Study "Simple" Growth Equations?	13
A Basic Strategy of Model-Building and Analysis:	17
Start with Simple Models	17
Complexity.....	18
Human Limitations on Representing Complex Systems.....	18
The Evolution of Models.....	20
Pedagogy.....	20
Major Factors and the Detection of Biases and Errors	21
Combinatorial Complexity.....	21
Taking a Closer Look at These “Simple” Growth Equations.....	23
Let’s start with the difference equation	25
Let’s add more realism and get the Logistic Growth Equation:	27
Let’s add a bit more realism	37
The Importance of Time Lags:	38
A Simple Mathematical Artifact	39
Some Examples of the Use of Various Options in Plotting N vs t Curves:.....	40
Other Kinds of Systematic Investigations with the Logistic Models—More Advanced Explorations of the Capabilities of the Simulator:.....	45
Finding Boundaries and Widths of Attractors:.....	46
Indirect method—Monte Carlo Simulations.....	48
Direct Method for Estimating Boundaries Using Spanning Trajectories.....	49
The Janus Face of Computational Roundoff Errors	50
Chaos in Discrete State Systems	51
A Second Monte Carlo Simulation	63
Phase Plots.....	65
The Analysis of Periodicities.....	73
Temporal Effective Population Size.....	81
Mathematical Artifacts.....	84
Relationships Between r vs N Plots and r vs $N_e(t)$ Plots.....	85
Other Food for Thought:	94
I. Simplifications of the Logistic and MDLS Equations	94
II. Behavior of Dynamical Systems.....	95
III. Other Thought Questions	95
Understanding the Role of Mendelian Inheritance in Evolution by Simulating Blending Inheritance.....	96

The Genotype-Phenotype Distinction.....	101
Investigating the Consequences of Blending Inheritance.....	116
Basic Description of the Model.....	117
Building a Binomial Distribution Model of Blending Inheritance	117
Simulating Blending Inheritance with the Binomial Model	119
Doing the Blending Inheritance Simulation	119
Investigating Blending Inheritance: Exercises.....	117
Further Exercises and Analysis	126
Projects Involving Current Disputes.....	128
Appendix A: How To Use Logistic Growth and Blending Inheritance	170
<p>Appendix A has been removed from this copy of Modeling to reduce the size of the document. It can be found in the Modeling Tools folder inside the Collection Candidates folder titled "Modeling Tools User Manual"</p>	
Appendix B: Functions of False Models as Means to Truer Theories.....	171
Appendix C: Reductionistic Problem-Solving.....	175
Heuristics and their Biases	175
I. Properties of Heuristics.....	175
II. The Origins of Bias in Reductionistic Heuristics.....	175
III. Specific Reductionistic Biases	176
A. Biases of Conceptualization.....	177
B. Biases of Model-Building and Theory Construction.....	178
C. Biases of Observation and Experimental Design.....	178
D. Functional Localization Fallacies.....	179
E. Other Important Biases.....	179
Appendix D: Analyzing the Assumptions of the Logistic and MDLS Models	180
Analysis of Assumptions.....	180
Assumptions of the Logistic and MDLS Models.....	181
Mathematical Assumptions	181
Biological Assumptions	181
Measurement Assumptions and Problems.....	182
Comparative Evaluation of Assumptions--A Look at Other Models	183
Topics for Discussion	183
Appendix E: The Use of Graphics Options To Explore and Enhance Pattern-Detection.....	185
Spurious Patterns Which Are Artifacts of Mode of Presentation	187
Real Differences Which Are Products of How the Data Are Sampled.....	193
Roundoff Errors Unmasked—Quantized Escape Orbits.....	196

Appendix F: Analysis of the Growth Models Using Mathematica	202
References:	207
General Mathematical Background.....	207
References on Visualization.....	207
References on Population Growth Models	208
References on Chaotic Behavior.....	211
References on Model-Building	212
References on Blending Inheritance and Binomial Distribution	216
References on Developmental Constraints in Evolution.....	218
Index.....	219

What You Need To Know and General Overview

A Note on Using This Manual on the Computer:

If you are using this manual on the computer you will derive an important benefit, and will suffer a significant loss. The benefit is that you will see many PICTs produced by the program reproduced as they are drawn by the simulation—thus the simulations are re-run before your eyes—in color if they were drawn in color, and you have a color monitor. The significant loss is a frustratingly slow pace if you try to scroll through, following the text, in places where diagrams are frequent. Not only are the diagrams drawn before your eyes, but they are often redrawn over and over again, a line at a time as you try to scroll them into view. To largely avoid this state of affairs, page through the text a whole page at a time using Word 4.0's print preview feature, until you get where you want to go. This will draw the graphs, but without the constant redrawing. In the single-page mode (rather than in the default double-page mode) you will probably be able to read most of the main text (forget the footnotes!), and you can in any case scroll through rapidly until you find the section you want, using the PICTs as markers for where you want to go. (This mode of navigation obviously becomes more effective the more familiar you are with the manual). GOOD LUCK!

In this module and its accompanying programs (Logistic Growth and Blending Inheritance in the *Modeling Tools* module) you will explore how to use, analyze, and criticize some important and historically influential models in biology. We will focus on simple models in ecology and population genetics, since these disciplines have a long-standing tradition of model-building, and have explored its powers and limitations with the greatest sophistication. This module will allow you to explore mathematically simple models using a variety of modes of visualization provided by the programs Logistic Growth and Blending Inheritance. Mathematically, you should need no more than a background in high school algebra to understand and to do most things.¹ A basic understanding of the concepts of differentiation and integration of the sort presented in many high school introductions to calculus will help, though we will never do more than the finite sum and difference analogues of these operations. You will not need to be able to differentiate or to integrate anything, though you should know that dN/dt is a time derivative and **not** an expression from which the d 's cancel and it reduces to N/t . Knowledge of computer programming is not required, but you should be familiar with the basic operations for running a Macintosh program. You may find it useful to go directly to the Users Manual for the Logistic Growth and Blending Inheritance programs (in the *Modeling Tools* folder) so that you can read how to use

¹Most students should be able to do most or all of the suggested exercises without further mathematical study. Nonetheless, an excellent source for relevant mathematical background for both of the model units in this module (and for most other biological topics), beginning with topics from high school algebra and running up through probability, matrices, calculus, and differential equations is Batschelet, 1979. This book illustrates all topics with biological examples, and contains worked and unworked problems from the current research literature. It is a clear and well motivated book. See particularly his discussion of growth equations, and of the binomial, normal, and Poisson distributions.

these programs and try different simulations and graphical views of the data. Doing this should result in a variety of data patterns you will later want to better understand.

We will show you, with the help of the accompanying programs, that models are very powerful and interesting tools in science, which not only can generate predictions (a commonly emphasized function of modeling) but also (and more importantly) can guide explanation, interpretation, understanding, and discovery in science. When models are used in combination with computers, we can use our most powerful strategy for analyzing the behavior of models—*visualization*. The use and understanding of techniques of visualization is a major focus of this module. We will see that data generated by a model (or biologically real system) can be visualized in multiple ways, and that visualizing data in multiple ways allows us to detect novel patterns in data which may be missed by applying only one or two strategies of analysis. We come to understand the modes of transformation better, and to recognize characteristic “signatures,” “footprints,” or diagnostic patterns whose presence in one mode of visualization indicates what is going on in another. By viewing data in multiple ways we can also detect patterns in the data which are *robust*, that is, by detecting something in a variety of ways we are in a much better position to say that it is a *real* feature of the model or biological system we are studying and not an artifact of our analysis or strategy for manipulating that system.

In the next section, we provide a brief but general introduction to a *philosophy* or *methodology* of using models in science. The introductory section to the two main parts of this module (see next paragraph) expands on themes raised in the Introduction. In the Introduction, we point out not only some of the important uses, functions, and benefits of using models in science, but also biases, errors, and, in general, things to keep cautiously in mind when using models. (These biases can have a variety of sources, from the nature of a mathematical equation, and how it responds to different changes, to how it is applied to a system or implemented in an experimental design, to how the equation is visualized on a computer, to how we perceive, think about, and respond to that visualization.) We also point out that the kinds of models used in science extend beyond mathematical and computer models to experiments as models of natural systems. Thus, the points we make about the uses and abuses of models in science are applicable more broadly to the practice of science in general.

There are two main sections to this module which when used together with the programs Logistic Growth and Blending Inheritance illustrate a number of important uses and functions of models in science. The first section, Model-Building and Exploration with the Minimum-Density Limited Logistic Growth Equation, explores with a variety of graphical modes the behavior of several simple and closely related models of population growth. Among other things, you will be introduced to the concept of *chaos* and how to use the computer to visually analyze the behavior of chaotic systems.

The second section, Understanding the Role of Mendelian Inheritance in Evolution by Simulating Blending Inheritance, explores the consequences of a false but very influential model of inheritance proposed by Charles Darwin and widely accepted in the 19th century. A central point of this section is to show that even literally false models can be very useful for understanding and interpreting models such as those of Mendelian genetics which we take to be correct.

We have also included several appendices which should give further relevant help. Appendix A is a copy of the User's Manual for the programs Logistic Growth and Blending Inheritance. Appendix B provides a list of productive things you can do through the deliberate and knowledgeable use of false models—many of which enormously aid theory construction in science, and many of which can be done in no other way. Appendix C provides a list of “reductionistic” problem-solving heuristics, their properties, biases, and potential errors. It is intended to make you more sensitive to the kinds of errors frequently made when simplifying the system for analysis. Appendix D provides a list of simplifying assumptions made in the logistic and minimum-density logistic growth equations, and provides, in effect, a specific case for analysis in the light of Appendices B and C. Appendix E discusses explicitly how the mode of visual presentation affects the information presented, and analyzes a number of phenomena connected with pixel and computational roundoff errors—focusing on the need to “calibrate” the computer as an instrument. And finally, Appendix F illustrates how to use *Mathematica* to solve and plot the continuous versions of these equations. These appendices should provide material for further thought and analysis both of the equations themselves, and of how accurately they may apply to different real situations. It is hoped that a close analysis of these cases will prime your consideration of the other models you will encounter elsewhere in science.

Introduction To Using Models in Biology

There are two crucial differences between models and theories:² (i) People often call things models when they are not yet very confident about them. When their confidence increases, they promote the “model” to a “theory” or even a “fact.” In 1953 it was appropriate to talk about the Watson-Crick double-helix model of DNA structure, but that seems strange talk nowadays. (ii) More importantly, the term model is often used for conceptual or mathematical tools which have known falsehoods, problematic approximations, or implausible idealizations imbedded in their structure—often to make them simpler to work with.³ Theories use falsehoods too (called “approximations,” “limiting cases,” “idealizations,” and other such fancy language), so perhaps what separates theories from models here is the believed frequency and seriousness of the falsehoods, and, with models, the self-conscious and

²As the term is used here, a *model* is like a *theory* in that it is an articulated, often formalized or mathematical, account of a mechanism or structure which may be used in predicting, explaining, or reasoning about its behavior. Often (but not always) models and theories are mathematical equations or set of equations which are interpreted to apply to a kind of object or system in the real world. The main difference between a model and a theory is our attitude of acceptance. The more we are inclined to accept, for example a set of equations as adequately describing the behavior of a physical system, the more we are inclined to view that set of equations and its interpretation as a theory.

³It is often claimed (particularly by philosophers or logicians) that models are neither true nor false, but instead are kinds of instruments which work well or not in a given context. Quite apart from the philosophical rationales they might give for this, this way of talking seems most plausible when a model is used as a template to look for patterns in data—in which it is often picked up and put down in various places, where it fits better or worse and spoken of as “applying” or not. If the focus is on a particular application, rather than on the model, it seems more natural to speak of truth or falsity. We have chosen to speak of models as (true or) false to emphasize that models may be useful means to the truth even if they fail to fit specific situations very well if their failures can be used to point us in the right direction for their improvement.

deliberate use of the falsehoods in reasoning strategies. As strange as it may seem, reasoning with falsehoods is often an effective road, and sometimes the only road, to building better or truer theories and models.

Mathematical and computer models have a variety of uses and functions in biology, as in other sciences, but the complexity of biological systems (and sometimes our inability to manipulate them effectively in controlled ways) makes the use of computer simulations an especially productive tool. We have structured this module not only so that you can read about the reasons for using and building models, but also so that you will be able to use the accompanying programs to discover several facts about their use:

- The falseness of a model does not, by itself, make it useless—indeed, it often makes it more useful for certain purposes.
- The full analysis and understanding of even very simple models is often no trivial task—such that we are still learning new things about models that have been around for a long time.
- A simple model can often be used to serve quite a variety of different functions in different contexts, which place quite different demands on it, and which it addresses with varying success. One man's noise may be another's data. This is particularly true with our discussion of the logistic growth equation, since it has a number of features which make it useful not only for scientific study of a variety of phenomena (population growth processes, the nature of chaotic phenomena, scale effects in discrete state systems), but also for teaching about the relation of models to one another, and the variety of misleading and artifactual results that can arise if you don't analyze carefully how the phenomena are produced (e.g., the effects of pixel and computational roundoff errors).
- There are some commonly used problem-solving heuristics (“rules of thumb”) and procedures which make it easier to analyze models efficiently and effectively, and their study can help you to solve problems better and faster. (Some of the most effective are illustrated below and include breaking problems into subproblems, looking for limiting results or places where the behavior changes qualitatively, bracketing the case of interest with cases on either side, interpolation and extrapolation, looking at related families of cases, and using models as templates to better detect disparities.)
- The behavior of even very simple and deterministic models can be very complex and even chaotic if they are nonlinear, and, as we are learning, nature is far more often nonlinear than anyone would have supposed only a generation ago.
- Finally, visualization, often by multiple means, is an important tool for the analysis of models, and especially so for the study of chaos.

Some biologists are suspicious of the use of models and computer simulations because they believe that models are too biologically unrealistic to be useful and predictive. They believe that since mathematical or computer models are much simpler than the biological systems they are supposed to represent, they must be false and so cannot be useful in any interesting way. This objection has a parallel, which we will return to below, in objections to the use of laboratory experiments to draw inferences about the behavior of similar organisms or systems in the wild, where the conditions are presumably very different. Both of these objections reflect real and important worries—simulations and laboratory results have not infrequently lead to wrong predictions about nature. Unfortunately, we have no choice. There is no way around the need to make simplifications in models and in the conditions for laboratory experiments. The art is to ask the right questions, to make the right simplifications for those questions, and to recognize (since we often make mistakes, or change simplifications or questions) when our simplifications are getting us into trouble, and how to check up on them. To reject models just because they are drastically simplified representations of nature is to fundamentally misunderstand the nature and functions of models and theory in most areas of science.

Norbert Wiener, the MIT mathematician and father of Cybernetics, once said that the best model of a cat is another cat. But this was intended as a joke. A useful model is one which abstracts away from many of the details and complexities of any given system or organism to see how the behavior of the system depends upon the relatively few (and hopefully most important) factors included in the model. Leaving out detail is the only way a model or theory can apply to more than one case.⁴ There are then several functions and reasons for using models simpler than the systems they represent:

- Including or excluding factors from a model, and then comparing a model's behavior with that of a natural system is one of the central ways for figuring out which of the details of a natural system are important, and which details are largely irrelevant for the purposes of the model. Indeed, in general, theories should leave out irrelevant detail, and, usually, also factors which have smaller effects, are less common, and sometimes factors which are always present and quite important, but are only rarely required as supplements to information more commonly available.⁵
- Many times we know that there is information which is very important for understanding a natural system, but we don't know it specifically or we can't measure it. In these cases, we have to substitute something else and hope that this something else will do. In such cases, it may be possible to construct simpler models—as that something else—and then use these simpler models to help discover new and

⁴Wouldn't the cat have been an even better model if it had been an identical twin? An identical twin raised in the same way? ... and so on.

⁵Lower level reductionistic genetic or molecular detail often but not always has this status for population genetic and ecological problems. More detail is not always better. This point is eloquently made in Levins (1966, 1968, ch. 1) See especially his discussion of "sufficient parameters".

previously undetected features of a biological system (see next section for further discussion).

- Adding biological detail to a model will often increase the model's complexity. But complexity can make a model analytically and computationally intractable (see the next section), so simplifications and approximations must be made if we are going to be able to use a model to better understand biological systems.

A good theory manages to do all of these things without losing explanatory power, and, in so doing, remains applicable to natural systems because of the properties of those systems it does represent. The crucial considerations in making simplifications, idealizations, and approximations in a model are (a) whether the major causal factors and mechanisms are adequately represented and (b) whether we are aware of the biases that result from making these simplifications, idealizations, and approximations. Note that we may be aware of the biases without knowing what the true answer is—for example, if we know that the model presents a lower extreme case, and is therefore a conservative underestimate.

We get benefits not only from using simulations, but also from writing them. The computer is very literal-minded, and has no imagination at all, so the program for a simulation has to be written out in excruciating detail—detail which often teaches us more about exactly what our assumptions mean. If we're not sure exactly which way to do it, we can try alternative ways and see what the results are, and we learn from comparing their results. There is not currently anything in *The BioQUEST Library* that requires you to program your own models, though several applications come close to giving you the freedom to do so. (Of course any programming language can be used to build your own models, if you know how.) Extend, for example, is a model-building program that provides a user-friendly graphical interface and allows you to write small programs in a language that is similar in some respects to the programming language C, but simpler to use. Stella is a less powerful, but still simpler version of the same thing.

Simpler than programming your own model from scratch is hooking up parts or subsystems with already defined properties in special or customizable ways to see what happens. The Odums' Environmental Decision Making (available in the BioQUEST Collection folder on *The BioQUEST Library* CD) module uses Extend, and parts they designed, so by using their simulations you can hook up parts in your own way without having to program them. This, and also some modification of the properties of the parts, can also be done in Biota (available in the BioQUEST Collection folder on The BioQUEST Library CD,) where you are free to define new species, how they interact with other species, how they migrate, and how many there are, and where, at the start.⁶

⁶Once we get something that will run (because we have told the computer all it needs to know) we are indeed often surprised at the results, not only because qualitatively new kinds of things (like chaos) might emerge, but also because we can quickly and painlessly investigate and compare different conditions and assumptions, and will often see patterns and effects which we never would have had the time to generate.

But all this control and freedom to modify the system you are modeling is too much freedom unless you have some idea of what to do with it. With any complex system, you usually can't just set things up at random and expect something interesting to happen. And you probably won't recognize it if it does, unless you already have expectations which are honed from the study of simpler related systems. This is one important reason for starting with simple models (also see next section).⁷ In a way this module is misnamed. This is not a "how to" manual for programs like Stella or Extend. It is closer to a "why to" manual for model-building, together with a "how to" manual for the strategy of model-building and exploratory model evaluation. All of the work you will do here is with models we have already written. They are very flexible models, and you can study them in a variety of ways—including getting them to behave like some other models. But we are teaching something different here than how to hook up different mechanisms to make a complex system. We are trying to teach how *methodologically* you should approach the task of doing it.

The two most important skills of the model-builders' art are *first*, to learn how, when, and where to simplify, and *second*, to learn how to tell when you have oversimplified, and how to diagnose, from the behavior of the model what is causing it. This tends to be learned best by studying specific cases in some detail—a study which is more useful the closer the cases are to the area you want to investigate. An important lesson to learn early is that there is no single best model for all purposes. What models are appropriate usually depend upon what questions you want to ask and how accurately you must know the answer. (Increased accuracy usually requires more complex models.) We may, in different situations, use different subsets of the relevant causal factors so that we don't have to deal with all of their complexities at once, or because different factors assume different importance in different circumstances. Good theory as well as bad models—and everything in between—must simplify. We can use them all—we just need to be able to learn how to tell one from the other.

Including the major causal factors and mechanisms responsible for producing something a biological system is observed as doing (e.g., an organism exhibiting circadian rhythms or a colony of the cellular slime mold *Dictyostelium discoideum* undergoing spiral and concentric waves of cell movement) is complicated by two additional facts during the model-building process. First, biological systems exhibit multiple levels of organization (e.g., organisms are composed of cells, tissues, organs, and organ systems; each of these components are located at different levels of organization), and major causal factors may cross levels of organization (either from higher to lower or from lower to higher levels of organization).⁸ This can present both

⁷How many of you did "chemical experiments" as a kid which consisted of mixing together a bunch of chemicals and seeing what would happen? I always got sort of an inert grayish brown looking sludge that did nothing interesting at all—until I got some hints as to what to mix together, and in what rough proportions. Once you had something that did something, small variations on it sometimes did other interesting things. This last represents a generalizable and rather deep truth about evolution and about problem-solving—that it is a lot easier to get something interesting by modifying a design which is at least moderately successful than by starting from scratch.

⁸For further discussion of levels of organization see Wimsatt (1976). For a detailed example of modelling at several levels of organization simultaneously, see Schank, (1991). Some of the most challenging situations arise when we must consider relations between different theoretical perspectives on the same system which are NOT related as levels of organization—a more common situation as the

opportunities for simplification and sources of bias. Opportunities for simplifications can occur because there are often one or a few upper-level variables which summarize the effects of a large number of lower-level variables.⁹ But this multilevel character of biological systems is also a source of bias because redescription at one level may help to hide the dependency of that variable on conditions at other levels. Thus, we treat fitness as a property of genes or of genotypes in selection models in population genetics, easily forgetting that fitness is a property of whole organisms imbedded in their environments. Fitness attaches to individual genes only conditionally—in the context of their given genetic, somatic, and ecological environments. If any of these changes, the fitness of a gene can change without any internal change in its structure, so fitness is not an intrinsic property of genes.¹⁰

Second, what factors and mechanisms we take to be causally relevant to something a biological system is observed as doing is relative to a *theoretical perspective* of what the system is doing.¹¹ To the extent that a model incorporates the major causal factors and mechanisms of a biological system (where these factors and mechanisms may occur at different levels of organization), as viewed from one or more theoretical perspectives, a model will be viewed as biologically realistic. To the extent that major factors and mechanisms are missing from a model, it will be viewed as unrealistic relative to one or more theoretical perspectives of the system modeled. But, as we will see, the biological realism of a model is only one dimension against which to judge its usefulness.

It is also sometimes stated that a model represents only a mathematical reformulation of what we already know, so it cannot be predictive. In short, on this view, you cannot get any more out of a model than what you put into it. But this view mistakenly supposes that we already know all of the consequences of our axioms and assumptions, and thus cannot be surprised by seeing them derived. That is obviously false, as has become more obvious now that we have computers to extend our reach in exploring the consequences of our models—as you will see below. The models presented here are extremely simple, but using them you will be able to find and analyze phenomena which have never before been seen or reported. And this is true with virtually any open-ended simulation environment.

system becomes more complex. For more on "perspectives", see Wimsatt (1974).

⁹This usually happens when variables are related as upper-level and lower level redescrptions of the same thing, and upper level variables or redescrptions of the lower level can be substituted for lower level variables in a variety of contexts, producing enormous simplifications in the model. Thus fitness—as an upper level variable in evolutionary models—can for some, but not for other purposes, be treated merely as a number reflecting the contributions of viability and reproductive success—lower level variables, without asking about anything about the adaptations which give these benefits. Cf. Levins, 1966.

¹⁰Paradoxically, George C. Williams (Williams, 1966) who makes this point very clearly, is also, with Richard Dawkins, an ardent advocate of the gene selectionist view, which often suggests the most excessive disregard of it.

¹¹For example, an organism can be viewed from multiple theoretical perspectives such as an anatomical structure, a biochemical network, a developing organism, a parent, a system exhibiting circadian rhythms, a part of an evolving population, etc. For further discussion of issues regarding multiple perspectives and the complexities that can arise when neat orderings between levels break down, see Wimsatt (1974).

You will also be made aware of phenomena which you have probably never noticed or have noticed but taken for granted—part, as it were, of the *theory of the instrument*. Computers are instruments, no less than thermometers, gel columns, or spectrophotometers, and until you know how they transform the data at hand, you do not know how they can break down, and you cannot distinguish artifacts in simulation modeling situations from the phenomena you wish to study.

Experimental biologists, like mathematical and computer model builders, are also constrained in the design and execution of laboratory experiments. These constraints are in many ways similar to the constraints on the modeler. Our limitations in manipulating systems are most obvious on the ecological and evolutionary scales, but it is no less true in laboratory biology. The organisms we study often seem to be strange choices (something often made fun of in popular presentations of science), being—at first look—neither interesting nor important to our livelihood. *Drosophila* (the commonly misnamed fruit fly), *Dictyostelium* (the slime mold), and guinea pigs would seem to have no relevance for us, but they have become model organisms and we have developed standardized systems for manipulating them because they are particularly good for answering particular questions which have much broader applicability. They may also often be used to answer other questions for which they are less suitable because they are what we have around, and if we already know something about them, that gives us an advantage.¹² They may be studied in circumstances which are very different from those found in nature. And they may even become significantly changed over generations—either consciously or unconsciously genetically manipulated and selected, in effect, domesticated to the laboratory—so that their behavior, immune reactions, and even their anatomy are no longer representative of the wild species from which they are descended, much less of the other species for which they stand as surrogates in the laboratory. Thus, in various ways, experimental biologists also build false models when they do laboratory experiments.¹³

The reason is that in order to discover causal factors and mechanisms, simplifications and approximations will typically have to be made. Experimental biologists start by developing an *experimental design*. In the simplest case the experimental design will specify treatment and control conditions. In developing an experimental design, an experimenter must consider a number of things including what factors may intervene to confound the experiment. Thus a simulation experiment may involve some efforts to test the random-number generator, or key data-transformation algorithms, or make checks for the significance of round-off errors. This is part of the reason for setting up treatment and control conditions in an experiment—to establish what the major causal factors and mechanisms are, and to isolate the system from other factors which might contaminate the results. Given that natural environments are

¹²Indeed, we will do just that with the models introduced here as well. As a case in point, see the proposals below to study the deterministic growth equation with integral values for N as an example of a discrete state system--something which raises other kinds of questions and which does not use its properties as a growth equation at all.

¹³See Griesemer and Wade (1988) for discussion of how laboratory experiments are a type of model and how biologists use experiments to reason about true causes in nature, and Schank, (1991) for further elaboration of this point.

normally much richer than laboratory environments, the more realistic the environment of an experimental system is allowed to be, the more likely it is that unknown factors may intervene to possibly confound an experiment. This is why so much of biology is conducted in the laboratory, where greater control can be achieved over the biological system being studied. Thus increasing the realism in biological experiments will often require increasing the complexity of the design and analysis of an experiment, and leads to analogous problems of tractability to those found with mathematical and computer models.¹⁴ Therefore, in issues regarding the biological realism of a mathematical or computer model on the one hand, and a laboratory experiment on the other, the model builder and experimental biologist are in the same boat; they will typically be forced to make simplifications and approximations—and in various ways to assess the consequences of their use.

Therefore, the fact that a model or simulation is false, by itself, tells us nothing about its usefulness in science. Knowledge of the ways in which false models can be useful is crucial for understanding the productive use of models in biology and throughout the sciences.¹⁵ In this module, we illustrate with two examples of mathematical modeling, how false models can serve a variety of functions in science despite (and indeed sometimes because of) the fact that they are false. In our first example, we will look at three simple models of population growth called the exponential, logistic, and minimum density-limited logistic growth models. These models are false in the sense that they leave out a number of factors which can affect population growth, and give at best qualitatively accurate descriptions of the action of the factors they do include. They are nevertheless realistic in the sense that they do include major causal factors which influence population growth (birth and death rates summarized in a population growth rate parameter, r ; limitations on the number or density of organisms that an environment can support, summarized as K , the “Carrying Capacity” of the environment for a species; and a lower density or number threshold, C , below which the population goes extinct, for a variety of possible reasons—for sexual species, most obviously, due to the increasing difficulty of finding a mate of the opposite sex!). They include other causally significant factors as well (such as the time lag, Δt , whose significance is sometimes missed when one trusts only to intuition, rather than submitting these intuitions to the rigors of a simulation. These very simple models will illustrate the important points that one can often get much more out of a model than what is put into it, that visualization (and alternative modes of visualization) is an important tool for getting this additional information, and demonstrate a number of the techniques for analyzing and getting information out of models. Their idealizations and simplifications will also be discussed.

The second example we will consider is a model of inheritance which is false in a deeper way than are the simple models of population growth. This model is a simple

¹⁴Similarities in the constraints on laboratory experiments and mathematical (and computer) models are not the only parallels. The experimental design of an experiment share many features in common with a computer program, in fact, there are many similarities between the development of computer software and the development of an experimental design (Schank 1991).

¹⁵In Appendix B we have listed a number of ways in which models can be false and how the ways that models are false can be used to develop truer theories (see Wimsatt 1987 for more discussion on these issues).

binomial model of Darwin's theory of blending inheritance which gets the basic mechanism of inheritance fundamentally wrong, though it operates in the right context, and thus, surprisingly, can get some things right. This example illustrates that models can be false not only by leaving out detail, but also by misrepresenting or oversimplifying what is there. It might seem harder to justify the usefulness of literally false models in science, but it is not—even though they are seldom studied in standard science courses. This is a serious mistake. As we shall see below, even literally false models perform many useful functions (see also Wimsatt, 1987). In particular, we will show that this literally false model of biological inheritance illustrates the importance of Mendelian inheritance for Darwinian evolution, helps to clarify the genotype-phenotype distinction, and is useful for detecting errors in modern group selection models. Furthermore, some results are the same under blending as under Mendelian inheritance—which illustrates that they do not depend in detail upon the mechanisms of Mendelian inheritance. Discovering what something does **not** depend upon is sometimes as important as discovering what it does depend upon. Thus, results of this sort might plausibly be expected to remain true where the subject of the investigation is cultural inheritance and cultural evolution (see Boyd and Richerson, 1985), where the Mendelian laws of inheritance do not apply. This is the sort of stuff of which important discoveries are made.

This module is much more of a general methodology module than any other but the Introduction. How does the main point there—the importance of studying all three of the "3P's"—play in this chapter? What is the role of problem-posing, problem-solving, and persuasion in model-building? Seeing a list of assumptions may help at all stages with the 3P's. It may be helpful at first to recognize that, appearances to the contrary, the 3P's do not refer unambiguously to definite stages of a scientific investigation. Not only is it true, as Jim Stewart and John Jungck observe in their introduction ("Introduction to BioQUEST", found in the BioQUEST Collection folder on *The BioQUEST Library* CD,) that the three processes are often complexly interdigitated, as they must be in any process of generating an answer through iterative refinements. It is also true that different people have different perspectives, or different frameworks for generating problems, so that one person's problem-solving may be someone else's problem-posing, and a third person's persuasion. The solution to a problem of determining the assumptions of a model (not always a trivial problem—see Appendix D), may be an antecedent to posing a problem about how to evaluate the applicability of a model to a specific ecosystem, a list of relevant information for use in designing a simulation, or part of a persuasive argument that a given conclusion was unsound.

Consider another case, one from the design of this module, and the cause of the latest (and still largely unplumbed) additions to the software. This may not mean much to you now, but you will see below! We start by noticing that the behavior of a chaotic system is (by definition) highly sensitive to roundoff errors, which raises questions about the validity and utility of detailed predictions with chaotic systems (*thus relative to persuasion?*) which I was already familiar with (so it became an established phenomenon which could be used for other ends, which...); provided a perfect way of studying the effects of roundoff errors—since any difference makes a difference! (here we have a key to the solution of a problem which had not originally been posed!); led to the development of exercises (on pixel and computational roundoff; problem-solutions for me which will produce problem-posings for you) and program facilities (control over graph coordinates, and the numerical precision and round N options), the

last of which led in turn to an entirely new set of possible investigations of finite automata, using the logistic engine with the round N option and the setting of K to determine the number of states accessible to the automaton, and the setting of r to determine the number and character of cycles it produced. (This was the invention of a new problem-space—or rather the recognition of an old one in an unexpected place!) In the process, roundoff error became transformed from something to be avoided (by following certain practices of numerical hygiene) to something to be studied (as a kind of artifact to be recognized and avoided), to a means of producing a center stage phenomenon (a class of tunable finite-state automata which produce rich complex behavior and are worthy of study both for their own sake, and also for their relation to the (supposedly) continuous case!)

Model-building and exploratory model analysis have this kind of open-ended structure. I have been working in class with the logistic and blending inheritance simulations in these modules since their first realizations in an HP-25 programmable calculator in the spring of 1976. Every once in a while, I think I have discovered everything you can do with them—that is, until I get the next year's student labs, or we put the next new feature in the software. Not only do I get to find out all these new things, but I also get to write about them. Is everything you can learn from a model already contained in its basic assumptions? Not a chance!

Model-Building and Exploration with the Minimum-Density Limited Logistic Growth Equation

Why Study "Simple" Growth Equations?

When studied with the computer, the much maligned simple growth equations, i.e., the exponential (EXP), logistic (LOGS), and minimum-density limited logistic (MDLS), take on a new life. They provide many opportunities for almost unlimited exploration, and the complexity of the behavior of these simple equations makes them natural objects for further study. An important and often ignored feature of models is that they come in related families, and studying how they are related and the comparative analysis of their behavior to determine how it depends upon their assumptions is a major activity of model builders. The exponential, logistic, and MDLS are closely related, and are thus interesting to study for this reason. Another important reason is that the logistic and minimum-density limited logistic equations are nonlinear. Thus, as simple as they are, they are already capable of chaotic and unpredictable behavior of the form that has led to the much publicized "new science" of chaos. Indeed, the analysis of the logistic and other population growth equations was a significant early problem for chaos researchers (see the article by May and the books by Gliesh, Stewart, and Becker and Dorfler in the **References** section). The exercises on model-building and analysis using this simulation are designed to do the following:

- Teach you the behavior of these growth equations, which have a time-honored history as models in ecology and elsewhere (see Kingsland, 1986) and their strengths and limitations as models of population growth.

- Teach you how they are related to each other and to more complex models and, using them as an example, how you can build more complex models out of simpler ones by adding or changing parts of the model.
- Teach you some general and simple features of how to investigate the behavior of dynamical systems, including elementary stability analysis and illustrations of Monte Carlo simulation.
- By employing four distinct modes of representing the data— t vs N plots, $N(t)$ vs $N(t+m)$ or "phase" plots, r vs N or "bifurcation" plots, and r vs $N_{\text{eff}}(t)$ plots—show how different ways of analyzing the data can give you different tools for exploring its patterns, a common feature in science which aids both in discovery and in getting a more robust understanding of the phenomena.
- Introduce you to the study and nature of chaotic phenomena, which are important not only in ecology, but increasingly, for the study of nonlinear, irregular, and apparently random behavior throughout the sciences.

What is Chaos?¹⁶

Chaos or chaotic behavior was in effect first discovered by French mathematician Henri Poincaré while working on the problem of the stability of the solar system, when he showed quite to his surprise that it was stable only under very special conditions—and in general that it was unstable. It was rediscovered and characterized by MIT meteorologist Edward Lorenz in 1960 while doing a computer simulation of a mathematical model of the atmosphere. He had done a long simulation, and wanted to extend it. As a safety precaution (so that he could check that he had restarted it correctly) he entered the values for the parameters corresponding to some time before the end of the last run, and re-ran the simulation from there. To his surprise the values, which initially started out in the same way, began to diverge from those of his preceding run, with no apparent cause for the difference. The explanation he soon arrived at was that he had gotten the values to restart the simulation off the printout from the last run, but the values in the printout had been rounded off from more precise (and therefore, different) values used by the computer. These very small differences appeared to totally change the outcome a sufficient time later. Most deterministic systems have a very small number of points or places where small differences can diverge to make a big difference. (Thus, a small difference in location of a raindrop falling just east or west of the continental divide can lead it to end up in the Gulf of Mexico or the Pacific Ocean, but small differences in location through most of the United States have no ultimate effect). But Lorenz found that for some ranges of some parameters, almost every small difference grew with time into a big difference. This was new, and led to its baptism as the so-called “butterfly effect.” You will be able to study this effect for the logistic and MDLS equations below. It has remained as one of the central definitions of chaotic behavior—sensitivity “almost everywhere” to arbitrarily small changes in initial conditions. Chaotic systems have a large number of other properties, one of the most famous being behavior which is apparently random. This was one of the properties first emphasized by Robert May in 1974, who first discovered it in the logistic growth equation—probably the simplest system to show chaotic behavior, and the system in which we will study it. These properties, and many others, have been shown (by physicist Mitchell Feigenbaum, in 1976) to be very general features of nonlinear differential equations, and therefore to be very broadly found in the physical, biological, and social sciences. (For more on this, see, e.g., Stewart, 1989.)

The study of chaos, in particular, has a number of other advantages for illustrating general techniques of model-building:

- Although the study of chaos is done numerically, and requires amounts of computation which would have been prohibitive before the computer (or at least the programmable calculator!), many of the main lessons of studying chaos point to the virtues of visual representation and qualitative analysis.

¹⁶A Note on format: Special topics of broader interest whose fuller development in the text would disturb the flow of the argument are boxed with a single line. These may include historical remarks, expansions on technical points, and the like. They can normally be ignored in reading the text, although they may help in understanding some points. There are many exercises suggested in the text. When several occur together, they are boxed with a double line.

- Visualization plays a large role in the understanding of chaos, and this simulation has been designed to make qualitative visual classifications as easy as possible by providing lots of control over how the data are represented. In this, the study of chaos has much in common with many other problems which require visualization for the presentation of complex multidimensional data, a process which both stimulates and feeds upon new developments in software for exploratory data analysis (see Wimsatt, 1991).
- Even though chaotic behavior is a case where, paradigmatically, precise values matter, it is also somewhat paradoxically an ideal place to learn that more precision is not always better: sometimes precision is essential (as when constructing a model to be programmed in the computer, or choosing initialization values to run on the computer), but sometimes it is not only unimportant, but actually gets in the way—as in the analysis of some of the data produced by these precise programs. Efforts have been made in the design of these simulations to support qualitative inferences and visual analysis of the patterns wherever possible, while making quantitative data available where necessary.
- The infinite sensitivity of chaotic behavior to the values of the parameters provides a natural way to see that for all of its strengths, the computer, like any other instrument, has its limitations and its biases. No finite machine can use numbers that need to be specified with infinite accuracy, but chaos feeds on any difference, no matter how small, and magnifies it, producing differences in supposedly identical simulation runs.¹⁷ The computer also occasionally throws away differences that we had hoped to keep, making identical simulation runs that should be different. The infinite sensitivity of chaotic systems makes them ideal for illustrating the possible effects of roundoff error. We have included special tools (like the capability to vary the numerical precision used, either totally or selectively, in computations) and exercises to aid in studying these effects.
- More generally, the fallibility and systematic biases of virtually any means of observation, measurement, computation, experimentation, or theorizing are essential lessons to learn about model-building. These biases and errors can be corrected for only if we recognize that they exist. The program Logistic Growth and suggested exercises using it have been designed to study some of these biases and errors.¹⁸
- The infinite complexity of the chaotic regime provides an open-ended microcosm for study—at least down to the resolution afforded by the substantial but still limited accuracy of the computer. Like the highly

¹⁷This is the basis for the so called “butterfly effect.” That is, if the mechanisms interact to produce the weather on Earth are chaotic, then the fluttering of a butterfly in Africa may eventually result in a Tornado in Kansas or a hurricane on the coast of Texas.

¹⁸See Wimsatt, 1980b, 1981a, and appendix C-III for further relevant discussion.

touted Mandelbrot set, the bifurcation diagram provides a map for exploring the dynamical space of a class of equations. Unlike most of the now common programs for exploring the Mandelbrot set, however, this simulation is designed so that you can go beyond the map and analyze the dynamical processes which go to make up the map using the other three modes of analysis.

A Basic Strategy of Model-Building and Analysis: Start with Simple Models

The exponential, logistic, and minimum-density limited logistic growth equations are closely related in that each one builds on the previous one by taking it and adding a simple representation of an additional causal factor to increase the model's realism. This is a frequent strategy of model-building in science: start with a simple model that captures an important causal factor, and then add additional factors or variables to the model in a sequential fashion, in each case comparing its performance to that of the preceding model and to that found in nature.¹⁹ Despite their enormous simplifications, there are usually circumstances where most simple models work relatively well, and the exponential, logistic, and MDLS equations are no exception. Here, features of the behavior of natural systems which are captured by models at a given stage are assumed to be explained by the causal factors included in those models, and features which are explained only after the inclusion of additional factors are held to require those factors for their explanation. In this way, a series of comparisons and contrasts among models can be used to attribute responsibility for various phenomena to various parts or complexes of parts of a complex mechanism. This strategy is a heuristic, and thus is not absolutely foolproof (see Wimsatt, 1985, and Taylor, 1985), but it works very frequently, and is an effective way to generate improved models and explanatory hypotheses.

This strategy of model-building suggests a linear ordering of models from simpler to more complex, which is at the same time a progression from less realistic to more realistic models. This picture is only partially correct—a more accurate picture would be a branching tree of models of increasing detail, specialized to give more accurate and realistic pictures of different systems or of the operation of different factors in a given system which is too complex for realistic representations of all of the factors in it at the same time. But in either case, the natural question to ask is: Why should we bother to build and analyze the simpler models if they are less realistic or

¹⁹This strategy is used not only in developing models, but also in understanding and teaching them, and is found to varying degrees in all of the BioQUEST simulations. Thus, for example, in using GCK, one will almost certainly consider a 2 allele at 1 locus cross with Mendelian dominance before attempting the analysis of a 2 allele at 2 locus system, and that before a 4 locus system involving sex linkage, linkage, and multiple alleles. This is crucial when learning how to do the problems, but there is also a corresponding preference for the simplest possible explanation when considering different hypotheses to explain the behavior of a new unknown system. For engineers, simpler systems are easier to build, for auto mechanics, they are easier to fix, for evolution (with some qualifications) they are easier to evolve, and for scientists, simpler hypotheses are easier to construct, to modify, and to test. For all of this, as Darwin once observed, "Nature is as full of contraptions as it is of contrivance." and the correct solution is not always the simplest one. But simplicity is still a good starting place and working hypothesis.

accurate? Why, if we have the means, not go all of the way to begin with and build the most realistic models we can? There are several reasons for not doing so (and for others, see Appendix B, on the uses of False Models or Wimsatt, 1987), all of which cluster around the concept of *complexity*.

Complexity

Human Limitations on Representing Complex Systems

Miller (1956) found that there is an upper limit of seven (plus or minus two) on the number of bits of information that we can simultaneously remember, and this limitation appears to be independent of content. Although we have developed techniques for extending our short-term memory (e.g., memorizing and chunking several bits of information into a single item, making lists, drawing diagrams), complex systems can quickly tax these techniques. This suggests that there are important constraints on how we represent and think about complex systems. One very powerful and important way we deal with the overwhelming complexity of many systems is to *abstract* away the details and focus on what we view as the major factors. What we abstract away from a system and what we include in an abstraction of a system are relative to how we view a system, the level of organization at which we study it, and, in general, the aims and uses we want to make of an abstraction of a system. These abstractions of systems provide the conceptual basis for such simple models as the Logistic Growth and MDLS equations described here.

But the fact that we tend to abstract away the details of complex systems does not mean that the way we tend to think about them is inadequate for representing their complexity. When it is desirable to understand more about a complex system than is represented at a given level of abstraction, we often propose decompositions of a system into *parts* (and their interactions) and into different *kinds* of components and systems. Likewise, in certain circumstances, we can construct a more adequate representation of a complex system by combining representations of parts and kinds at lower levels of abstraction and organization. By recognizing different levels of organization and levels of abstraction we are often able to keep the task of representing and modeling complex systems manageable. This way of representing complex systems generates two ways of hierarchically viewing a complex system—where levels of abstraction generate *kind of* hierarchies and levels of organization generate *part of* hierarchies—as illustrated in Figure 2-1.²⁰ In a *kind of* hierarchy, an element is a type, which may have more special types which are instances of it, and may itself be an instance of a more general type. For example, a Chevrolet is a kind of car (Chevrolet → car).

²⁰See also Simon, 1981, chapter 7.

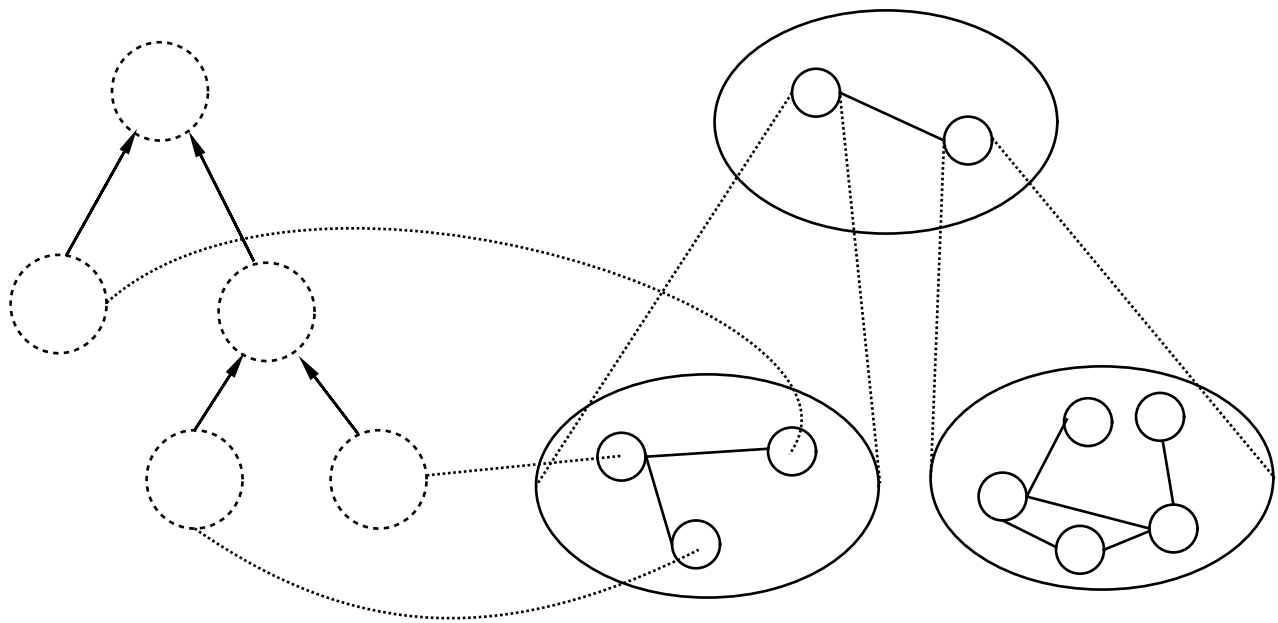
Kind of Hierarchy**Part of Hierarchy**

Figure 2-1. *Kind of* and *Part of* hierarchies.

Figure 2-1 is an abstract representation of two types of hierarchies. On the left is a *kind of* hierarchy. Arrows indicate the direction of the *kind of* relation between kinds or types. Thus, a dashed circle at the tail of an arrow is a subkind or subtype of the dashed circle at the head of the arrow. The kind of hierarchy in Figure 2-1 could, for example, represent a simple taxonomic hierarchy of living things—or part of one, where the dashed circle at the top of the hierarchy is the class of living things, the two circles below are plants and animals, and below animals are birds and mammals. On the right is a *part of* hierarchy and solid lines connecting smaller circles within a larger circle indicate strong causal interactions between parts (Simon 1981; Wimsatt, 1974). Dashed lines leading down from a small circle to a larger circle indicate a further decomposition of a part at a higher level of organization into several interacting parts. The *part of* hierarchy in Figure 2-1 could represent two interacting ecosystems at the highest level; at the next level down, these interacting ecosystems decompose into interacting populations of living things. The dotted lines linking the *kind of* hierarchy to the *part of* hierarchy indicate what kinds of things the parts are; for example, the dotted lines in Figure 2-1 could indicate that the larger circle on the lower left of the *part of* hierarchy represents three interacting populations, one a plant, another a bird, and the other a mammal population.

The Evolution of Models

Another reason for starting with simple models is the evolutionary nature of model-building in science. Models are usually not orderable in a single dimension any more than species are: just as evolution leads to local differentiation of populations or

species to adapt to local conditions, and a tree-like pattern of descent, so also models are adapted in different ways to apply to different systems, to differentially increase the detail of different parts of the model, to study different questions, and to get answers of different degrees of accuracy and different degrees of generality.²¹ There is no such thing as a single best model for all contexts. Rather, in general, different models are best for different contexts, depending on the aims and conditions of their use. Combinatorial Complexity prevents us from modeling all of the details of real systems and increases in complexity along one dimension are usually bought at the cost of simplifications elsewhere. (See the box on Combinatorial Complexity below.)

Pedagogy

As is to be expected from the previous reasons for focusing on simple models, simpler models are easier to teach, and provide a natural starting point which is easier to understand, and from which one can extend one's analysis and understanding to more complex models. You can't fully understand a model containing the interaction of several factors until you see how it works in the absence of each of those factors—and indeed, ultimately, in the absence of each combination of these factors. It is also a historical fact that more complex models are usually derived and modified from earlier simpler models. This in turn accounts for the practice that in teaching science, one often replicates (at least approximately) the sequence in which the models were developed.

Major Factors and the Detection of Biases and Errors

Even though simple models may leave out detail at other levels of abstraction and organization, they can still be relatively realistic at their focal level of organization or abstraction. This is because even very simple models can capture factors of major importance. The modes of interaction of these factors provide the conceptual basis for understanding what may come later and how a simple model can be improved. That is, the major factors represented in simpler models can be used to organize and classify deviations from the predictions made by these models. Classifications of deviations from these simpler models can then be used to detect biases and error in the

²¹ Though models can't be ordered in a single line from simple to complex, or from less to more realistic, it is not clear that a tree structure will do either. Indeed, the only reason why it seems plausible to construct a tree diagram relating models is through their descent relations, in effect constructing a "phylogeny" of models. But this won't work either, because, unlike the standard (and not always correct) picture given for species, models may claim descent from multiple "parents". The relations among models are thus even more complex than this. Another approach would be to recognize as a dimension every major mode of variation among models. Thus, for population growth models, the number of species involved could be one dimension, the number of age-classes recognized for each species another, etc. Some of these dimensions could be simple binary variables--e.g., deterministic vs. stochastic, except that for many variable models we would have to consider all combinations of assignments to these binary values to the variables. Though some (e.g., Levins, (1968), chapter 1) speak of a "space of all possible models" of a given phenomenon, there is no generally agreed upon way to represent the relations among different models, and different ways are likely to serve for different purposes. Some suggestion of the variety of possible models for population growth can be gotten by looking at appendix D, which classifies the assumptions of the simple growth equations, and some of the ways in which they can be modified to suggest other models.

assumptions they make, which can then be used to correct these errors, and thus to propose new models based on the simpler ones.

Combinatorial Complexity

More complex models are more complex to analyze in several respects:

- With more variables, more information needs to be specified or measured for each case, making simulating or testing them more difficult.
- Models with more variables take longer to compute, and often require much more computer memory to store the results. For the kinds of models discussed here, this is not a problem with current computers, though models seemingly just a little more complex (such as those which introduce multiple species, age structure, and density dependence, combine genetic and ecological parameters, or combine any of these with models which allow for spatial heterogeneity) can produce quite a stretch even for current equipment. (BIOTA is, for example, quite a stretch when more than a few species or more than a few spatial patches are included.)

Combinatorial Explosions

The expansion of possibilities that can arise when a number of independently variable factors must be considered is so rapid that it has a special and graphical name—combinatorial explosion. This rate of increase is easy to underestimate. Cases like this are basically the flip-side of the multiplicative rule for the probability of independent events. Just as such probabilities are multiplied together (and get smaller rapidly, because they are less than one), so the number of distinct outcomes (or cases) for each independent factor is greater than one, and when these are multiplied together, it gets very large rapidly. The probability of any given outcome is very small because it is just one of a very large number of outcomes generated by combinations of independently variable kinds of things or factors. Thus, GCK allows construction of models with up to 6 alleles (which together can produce 21 distinct genotypes at each locus) at up to 5 loci. That's a double combination of genes to make genotypes at different loci— $(21)^5$ to be exact. A full analysis of this model could involve looking at relations in inheritance patterns among any combination of these 4,084,101 possible genotypes—even ignoring the further complications of sex-linkage, linkage, chromosomal groupings, and recombination. (Luckily, we can usually break the problem into parts, and solve for most of these factors one or two at a time, so this enormous potential complexity seldom arises—see also discussions of near decomposability and quasi-independence. Indeed, if it did arise often, science as we know it would probably be impossible.) Spatial heterogeneity in ecological models is even more problematic. BIOTA allows any combination of up to 10 species to be placed in each of a 12x12 array of patches. Even ignoring (1) their initial numbers and classifying each species as just present or absent, (2) the possible choices for how they may interact, and (3) the different possible specifications for modes of migration between and among the 144 patches, this gives $(2^{10})^{144} = 2^{1440}$ which is approximately 10^{432} different cases. For comparison, the astrophysicist Sir Arthur Eddington estimated the number of elementary particles in the universe at a paltry 10^{79} . That was in 1939 so he could have been a little off. But were the universe a million times bigger (it's not clear that it could be without gravitational collapse!) that would only be 10^{85} . There have been approximately 10^{19} seconds since the big bang, and the current fastest estimate for a physical event is the amount of time it would take a photon to traverse the diameter of a hydrogen nucleus—or 10^{-24} seconds. So with 10^{24} of these occurring per second, the total number of events in the universe since the big bang must be no more than 10^{128} . Clearly, we're not even playing in the same ballpark, and it must take some art in exploring our space of possibilities in studying a model. This is where analytical classification and representative sampling (discussed below) play their parts. In any case, this combinatorial explosion is not all bad—it is responsible for our genetic individuality, and makes the world a much more varied and interesting place.

- More variables also means many more distinct cases that must be analyzed if the behavior of the model is to be completely described or understood: if each of v variables can take on s interesting states, the number of possible cases to be considered is s^v . This deceptively simple relationship hides daunting complexity. With 2 binary variables, each capable of being in 2 states, there are just 4 cases to study, but with 4 variables, each capable of 4 states, the number of cases jumps to 256. With 10 variables, each capable of 10 states, this number leaps to 10,000,000,000, and even this is small compared to the range of possibilities allowed with most BioQUEST simulations.

- Related to this is the fact that, as the number of variables increases, the possibilities for complex interactions among many variables increases. There are more such possible interactions, and the manipulations required to analyze each one, involving more variables, is more complex to analyze.

An enormous number of different kinds of cases to study means that they can't all be surveyed, and a premium is placed on either being able to group the cases into a smaller number of types (see kind of hierarchies above), such that one of each can be studied, as either a representative case, or the "best" or "worst" case (depending what you are looking for) of that group, or being able to sample from the large number of cases in a statistically representative fashion. (The latter is the strategy of Monte Carlo simulation.) Both techniques—**analytic classification** of a large number of cases into a smaller number of types, and designing a **representative sampling** of cases for a Monte Carlo simulation—are aided by studying related problems in simpler models in order to decide what features are relevant to the analytical classification or should be taken as representative.

Taking a Closer Look at These "Simple" Growth Equations

The simple growth equations in turn give the basis for further extensions to more complex models as we would expect from the discussion in the last section. The exponential and logistic growth equations provide the basis for generalizations to cases where organisms live to reproduce more than once, and have changing probabilities of having young and dying as a function of age.²² The logistic growth equation also provides the basis for models of multispecies interactions—the so-called Lotka-Volterra equations for species interacting by competition, predation, and symbiosis. They may also be used as the basis for models involving spatially heterogeneous environments, where different numbers of organisms of different types are found in and can migrate between different neighboring cells. All of these represent extensions which are used in the BIOTA simulations.²³ Even when things get more complicated with age structure and multispecies interactions the logistic and MDLS still often provide rough qualitative predictions of the growth of single species, and in any case, revealing comparisons for evaluating the magnitude of the effects of leaving the more complex details out. Often we can derive conditions from this comparison which tell us when the simpler model is adequate for predictions in the more complex case, and sometimes can tell us something about the effects of embedding simpler systems in a larger and more complex system. (See, e.g., Exercise 6 below about the destabilizing effects of time lags in multispecies systems.)

²²These are so called "age-structure models", which are themselves the simplest in a series of models necessary for reasonable predictions of population change in complicated cases like that of human demography and many animal species. See, e.g., Crow and Kimura, 1970, chapter 1, or Roughgarden, 1979, chapter 18.

²³In this case, each species is assumed to grow in a fashion suggested by the logistic, to which are added additional algebraic terms for the numbers of other species present, and how much they help or hurt its growth. The community change as a whole is given by the simultaneous solution of these interspecific growth equations for each of the species in the community.

We won't provide a detailed analytical and biological discussion of these simple growth equations since more extensive discussions are widely available elsewhere.²⁴ See the discussion of the assumptions of these models, and their analysis in Appendix D. Here we show how using these simple equations exemplifies model-building strategies, and biases; how their idealizations and ways they are false do not necessarily hinder, and may even promote, their usefulness; how the computer has revolutionized the study of mathematical models of natural phenomena, and how these models and the computers used to study them, like any other instruments, need calibration, thoughtful use, and the exercise of judgment. With this, even simple models such as these can be sensitive exploratory probes for sampling the rich complexities of nature.

The exponential, logistic, and MDLS equations will all be presented and used in a common form—as so-called finite difference equations. In many cases, using this form turns what would otherwise be a calculus problem (the solution of differential equations) into a closely related algebra problem.²⁵ Instead of integrating the differential equation to get a solution for population size as a function of time, the algebraic right-hand side of the differential equation is taken as specifying a growth rate at a given time t , as a function of population size $N(t)$, and other relevant variables, which holds exactly for a finite time interval, Δt . The population growth in that time interval, ΔN , is taken to be the product of the growth rate (dN/dt) and this time interval, Δt . The change in N , ΔN , is added to the preceding population size to produce a new population size, $N(t+\Delta t)$, which is used in the calculation for the next time interval. Not only are the equations easier to present and to solve in this way, but it naturally suggests a loop in a computer program, which is just how they are solved. (This kind of solution is called a numerical solution.) Indeed, barring the use of a symbolic manipulation program (like Mathematica) which gives exact, symbolic, or analytic solutions, it is the only way that they can be solved on a computer.

With difference equations (also called “finite difference equations”), we can approximate the solutions we would need with calculus if we were (like most physical scientists), interested in the continuous case—and in fact, with the computer, the study of difference equations is a natural way of learning differential calculus. To physical scientists, the difference equation serves as an approximation to the differential equation they really want to study. Perhaps most surprisingly, in population genetics and ecology, the situation is often or usually reversed. Because generations take a finite amount of time, and there is often a lag between changes in population size and when this change feeds back to affect the growth rate of the population, the difference equation is often the more realistic one, and the differential equation is used, if at all, as an approximation! (This turns out to be a crucial issue for the production of chaotic behavior, and one reason why it may be particularly important in ecology, because

²⁴See, in order of increasing sophistication, Wilson and Bossert, 1971, Crow and Kimura, 1970, and Roughgarden, 1979). Roughgarden is particularly good in providing many examples of how these and other simpler models are used in, presupposed by, or motivate more complex ones.

²⁵In particular, the difference equation will often have exactly the same solution as the differential equation. The difference equation may not be stable if or when the differential equation is, but the converse is always true: if the difference equation is stable, the differential equation must be. This can be readily seen using this program by choosing values of Δt closer and closer to 0, since in the simulation of a difference equation, the limit of this transformation as $\Delta t \rightarrow 0$ yields the corresponding differential equation.

difference equations are less stable and generate chaos more readily than their corresponding differential equations.)

Logos—Logistic:

Historically, the logistic equation (logistic, from the Greek word, *Logos*, meaning law) for population growth was first studied as a continuous differential equation because its first major exponent, geneticist Raymond B. Pearl, saw it as another extension of a more general law of growth. This law was first expressed in differential equation form for the growth of a chemical species in a chemical reaction as the reaction went to saturation. Pearl extended this equation to cover the growth of individual organisms in development, and the growth of populations. Lotka and Volterra were also first motivated to construct their multi-species growth equations via extension from related chemical and physical theories. For more on this, see Sharon, B. Kingsland, *Modelling Nature: The Origins of Theoretical Ecology*, Chicago: The University of Chicago Press, 1986.

A standard (though not the only) way of presenting a difference equation is as a pair of equations, the first specifying a rate of change of a variable with respect to time, dN/dt , and the second specifying how this rate will affect the change of the variable over a finite period of time. This second equation basically involves a linear extrapolation of the derivative or slope, dN/dt , for a finite time, Δt , and is the same in all of the equations we will consider.

Let's start with the difference equation version of exponential growth

$$\frac{dN}{dt} = rN \quad (1a)$$

and

$$N(t+\Delta t) = N(t) + \frac{dN}{dt} \Delta t \quad (1b)$$

where:

N = population size (number of individuals) or density (number of individuals per unit area).

r = intrinsic growth rate, normally evaluated under low density conditions, or best circumstances, if these are not at low densities (as with the MDLS equation).

$N(t)$, $N(t+\Delta t)$ = size or density at time t , and at time $t+\Delta t$.

Δt = a standardized time interval, either in "physical" time (a second, a minute, a year), or organism-relative time (a generation, or fractions or multiples thereof: a generation for *E. coli* is about 20 minutes, for *Drosophila*, about 14 days, for man (depending on how it is defined, and the age of reproduction in different

populations) 14-23 years, for a century plant (and some species of bamboo), roughly 100 years.)

The rate equation (1a) says that the rate of change (or growth) of a population is proportional to its size, N . The constant of proportionality, r , is called the intrinsic growth rate. The parameter r is the difference between birth and death rates, $r = b - d$. So if $b = d$ (a population is at equilibrium in size—as many are being born as are dying), then r is zero and the population is neither growing nor declining. If $b > d$, r is positive and the population is growing. If $b < d$, r is negative, and the population is declining.²⁶

When Darwin said that all species, by nature, increase at a geometric rate, he had essentially this exponential growth equation (or the corresponding algebraic geometric growth equation) in mind, and his talk of increase implies that he believed that for all natural organisms (at least under ideal conditions) $r > 0$, *or else they would no longer be here*. This same exponential equation, however, also applies to other situations where r is negative, such as the decline in the amount of a radioactive substance as it decays with time, or to the change in water level in a cylindrical bucket with a hole in the bottom. There are also clearly cases in biology (such as with local or global extinctions of a population or species) where r is negative. This equation is not directly studied in this simulation, but it occurs as a component of the logistic and MDLS equations, and at very low densities (a small proportion of the Carrying Capacity), logistic growth is approximately the same as exponential growth.

The second equation, (1b), is the equation used to get from the net growth rate (dN/dt , given by equation 1a) to the change in population size expected some time (Δt) later, using a method of *linear extrapolation*.

Let's add a little more realism and get the Logistic Growth Equation:

$$\frac{dN}{dt} = rN \left[\frac{K-N}{K} \right] \quad (2a)$$

(2a)

$$\text{and, as before: } N(t+\Delta t) = N(t) + \frac{dN}{dt} \Delta t \quad (1b/2b)$$

(1b)/(2b)

(in the future, (1b) will not be repeated but understood)

where:

N , r , etc. as before, except that r is assumed to be positive.

²⁶In these models, b and d do not enter directly as parameters, except through r . In more realistic models (including those for age-structured populations), b and d are used directly, so that the effects of changing them in different ways can be studied. See appendix D.

K = Carrying Capacity (equilibrium point for low values of $r\Delta t$.)

Note that equation (2b) is exactly the same as equation (1b). This is because it is here treated as a general method (one of *linear extrapolation*) for getting a value of a curve at a later point given its value (N) and slope (dN/dt) at an earlier point, and we will not change it in these models. The model changes because of changes in the equation for the growth rate—equation (2a). The second equation is a general way (and probably the simplest way) of going from a differential equation to a corresponding difference equation.)

With the logistic growth equation, in equation (2a) the equation for exponential growth (equation 1a) is multiplied by an additional factor, $[(K-N)/K]$, to reflect the density dependence of growth rates. At higher densities, the net effective growth rate usually decreases. (Indeed, at a high enough density, it **must** decrease—this was the insight Darwin got from Malthus.) This may be for a variety of reasons, either separately, or taken together—limitations of food, nest sites, increased costs of or time spent in foraging or socializing, intraspecific competition, or other negative interactions. The net decrease in population growth rate may be due to increases in death rates, decreases in birth rates, or both. In this model as in the MDLS below, r is kept constant as a coefficient for the species, indicating what it can do in the best of circumstances, and the effects of density through changes in birth or death rates are built into this new factor. This algebraic factor in this model represents perhaps the simplest possible case: a linear effect of density on growth, declining from 1 at $N = 0$ through 0 at $N = K$ to become proportionately more negative as N gets increasingly larger than K . Let us now point out some things about the behavior of this equation—or of any species which grows in this fashion (items (2), (4), and (5) below may be validated via natural extensions of exercise 1 in the box on the following page):

- (1) At very low densities, as $N \rightarrow 0$, $[(K-N)/K] \rightarrow 1$, so the logistic behaves more like the exponential. Indeed, when $[(K-N)/K] = 1$, this equation reduces to the exponential, and this limiting behavior is very revealing. (See graph 2 below, though if you try the values in the paragraph below the box, the effect will be even more pronounced).
- (2) With small positive r , for N between 0 and K ($0 < N < K$), N increases. (Try the same run, now from $N = 1$ for 50 generations.) It will approach $N = K$, slowing down as it gets closer. If r is small, we don't have to try this for larger starting populations between 0 and K because it already goes through a fairly dense sample of population sizes between its starting value and K , so we already have an idea of what it does for those values. This means that we can take any trajectory starting with N between 0 and K as representative for the others meeting that condition. This is an example of simplification of the number of cases considered through *analytic categorization*, as described above.

Functions of “Limiting Value” Arguments:

This kind of argument—looking for limiting values under which one equation behaves exactly or very much like another—is an extremely important way of gaining understanding of both equations, and is commonly practiced throughout the mathematical sciences. ***It is one of the most important problem-solving heuristics in mathematical modeling.*** It sometimes helps you to solve the more complex equation, but even where it does not, it gives understanding because it tells you what conditions on the more complex (and presumably more realistic) model yield behavior like the simpler one. In doing so, it tells you when you can get away with using the simpler one, and what you have to do to do it. When the equations are interpreted causally, this kind of limiting behavior allows us to tease apart the role of different causal factors in producing the results, and to see how the new causal factors of the more complex model interact with the existing ones of the simpler model. Limiting behavior also plays an important justificatory or testing role: when new, more detailed and complex models are first advanced, it is often a condition of accepting the new model that it produce behavior like older, simpler (but well established) models. In effect, when you are exploring new territory, it is better to start from a well known base camp. (A special case of this arises in testing a new computer model, where it is a standard technique to make sure that it works numerically in cases where the answer is already known. Indeed, restarting his run in the middle (through which Lorenz discovered the sensitivity of his equations to the exact initial conditions—see the Chaos box) was a special application of this technique: Lorenz was using it as a check to make sure that he had reentered the values correctly, for which the answer was, “yes” (approximately) and “no” (exactly).) For some further functions of this kind of procedure, see the descriptions of the functions of “successional reduction” in Wimsatt, 1976, and for discussion of more examples, see Wimsatt, 1987. (A list of functions of false models and examples is found in Appendix B.)

- (3) Determinism entails that (a) if we have exactly the same conditions again, exactly the same thing will happen. This, which could be called the ***principle of determinism***, is true for all of the cases we will consider. It does NOT entail the stronger claim, which could be called the ***principle of smooth determinism***, that (b) if similar conditions obtain, similar things will happen—though we often tend to believe it anyway. And it usually does happen, with very rare exceptions, e.g., the raindrop at the continental divide.²⁷ We are basically assuming the truth of (b) in making this analytic classification, and we got away with it this time. But with chaotic behavior, and this is the important difference, (b) is almost never true.

²⁷ Many discussions by philosophers of causal determinism focus on what is above called the ***principle of determinism***, arguing that our notion of physical law requires commitment to the idea that, in exactly similar circumstances, exactly the same thing will happen. Anomalously, they do not discuss what is here called the ***principle of smooth determinism***, which is probably even more important for our concept of a physical law. Roughly, in circumstances where we are ready to recognize the existence of a law, it must be the case that there are substantial ranges of conditions within which the principle of smooth determinism holds. One of the anomalous features of chaotic behavior is that under conditions where it is realized, this principle breaks down almost everywhere.

Exercise 1: The Exponential and Linear Limits

You can investigate this effect (and everything else discussed here) with the Logistic Growth program found in the *Modeling Tools* module. Logistic growth is a special case of the minimum-density limited logistic growth equation attained by setting $C = 0$, so do that first. (see Appendix A for instructions on how to use this program.) Now start with a very small value of N , a pretty small value of r , and run the population for just a few generations t , such that $N(t)$ is a small fraction of K . (Thus, try $N = 1$, $K = 100$, $C = 0$, $\Delta t = 1$, and $r = 0.1$ for $G = 10$ generations, with the bottom radio button set for 1 run.) The discrete exponential would predict a population size at the end of generation 10 to be $(1.1)^{10} = 2.593742$. What do you see? How close is it? With strictly geometric increase, if you multiply the starting value by a constant value (say 10) the ending value should be multiplied by the same amount. Is it? Try it and see. The fact that it isn't (and also the fact that you didn't get exactly 2.593742 after 10 generations in the first run) is due to the second density dependent multiplier. The fact that the deviations from the exponential get proportionately larger in the second run is due to the higher densities, and the fact that $[(K-N)/K]$ is no longer close to 1. Consider now what happens as r gets very small. (This suggests a still simpler *linear* approximation, in which for $rG \ll N$, ΔN is approximately constant.) What is the difference? Under what conditions can a geometric rate of increase be approximated by a linear relation?

- (4) We can also see from the equation that at $N = K$, $[(K-N)/K] = 0$, so dN/dt is also 0 and there should be no change in the population. Thus $N = K$ is an equilibrium.
- (5) For $N > K$, $[(K-N)/K]$ is < 0 , so we would normally assume that the population would decline, until $N = K$.
- (2) through (5) together imply that $N=K$ is a stable equilibrium, which means that if the system is perturbed or changed slightly from that value, it returns. But, $N=K$ is not *always* a stable equilibrium, as we will see below in Exercise 3.

Exercise 2: The Classical Vision of the Logistic

These last few ideas can be tested directly. Set $K = 100$, $N = 25$, the increment in r at 0.1, and the bottom radio button for 10 runs. Let generations run from 0 to 25. (*If you don't understand these instructions, see Appendix A.*) You will now see that $N = 100$ is an equilibrium for 10 different values of r , $.1 \leq r \leq 1.0$. Now set $N = 50$, and the top radio button for 10 runs at an increment in N of 10. This will show you that $N = K$ is a stable equilibrium for values of N from 50 to 140, since from all of these values, N approaches K as time goes on. (If you want to, change K to verify this.) Thus, at low densities, growth rates are roughly exponential. As densities increase, density-dependent effects slow down growth, and the population slowly approaches its maximum density at $N = K$. If, by some chance, the population goes to too high a value of N (as could happen through a rapid influx of migrants (increasing N) or rapid decreases in carrying capacity, K —as happens for many species every winter), the population declines, approaching $N=K$ from above. This is the basic, classically understood behavior of the logistic growth equation—before chaos, and before Robert May's 1974 paper in *Science*.

Exercise 3: Period Doubling and Chaos

With $\Delta t = 1$, investigate different values of r ($0 \leq r \leq 3$) to see what values are stable. (We speak of values of r as being stable when $N=K$ is a stable equilibrium for those values.) To do this, it is useful to pick a starting value of N close to $N=K$, and see whether from that distance (technically both above and below $N=K$, but in this case, it doesn't matter too much) the N -values get closer to $N=K$ or further away. What happens for $r > 2$? When you understand these conditions, try varying Δt to see how changes in that affect the stability of r . Can you find a simple condition or equation relating r and Δt that determines when $N=K$ is a stable value? Increment r slowly above that value. What happens? Can you find 2-point stable oscillations? Over what range? (A stable oscillation, like a stable equilibrium, is a set of points such that small perturbations from it in any direction tend to return to it. To test for the stability of a 2-point oscillation, try oscillations which are slightly smaller and slightly larger to see if they both converge to the oscillation in question.) Can you find 4-point stable oscillations? 8-point? How long do you think this sort of thing can continue? Can you find out? Where do things start to get unpredictable? And what do we mean by predictable and unpredictable anyway?

HINT: the behavior of the logistic (we'll look at the MDLS later) for $r > 3$ is not interesting dynamically because the population always goes too high and then crashes immediately to 0 or below, though the program converts negative population sizes (which don't make too much biological sense) to zero. This dynamical behavior may not be very interesting (in effect it stops the behavior) but biologically this is very interesting since it means that the population has gone extinct. Those biologists who thought that chaotic behavior in the model said something about real populations in nature (which was not everybody, and very few people initially) thought this was very interesting. Many others, though, thought that the mathematics didn't have too much to do with reality. Arguments for the importance of chaos have grown stronger since. See Wimsatt 1980a and Shaffer 1984 for discussions of the reality of chaos in nature.

A better view of what is happening can be achieved by doing a different kind of graph—resulting in what is sometimes known as a "bifurcation plot" or "fig tree," a play on the name of one of its inventors, physicist Mitchell Feigenbaum.²⁸ This graph has come, perhaps more than any other, to be emblematic of chaotic behavior. It is an unusual sort of graph in that many y -values (rather than just one) are plotted for each x -value. In this graph, it is as if a separate N vs t graph were plotted for each r -value (or more exactly, at the r -value corresponding to the location of each pixel along the x -axis). But here the t -values are no longer plotted explicitly. Rather, the N -value is plotted for a series of successive values of t all at the r -value for that run. (These graphs take a long time to construct. The graph below, Figure 2-2, is 400 pixels wide, and plots the population size between generations 200 and 300 for each of these 400 runs. Since it also has to calculate the first 200 generations to know where it is to start plotting at generation 200, the computer is doing $400 \times 300 = 120,000$ iterations of the MDLS equation, so you should expect to wait for a while—though you can watch while it makes the graph in the standard version of this program.)²⁹

²⁸Actually, this mode of plot was first used (in 1976) by ecologist (and ex-physicist!) Robert May, but he did not extend his calculations systematically beyond the onset of chaos, and thus did not see the whole "tree".

²⁹We have now added an option so that the bifurcation plots can be done in color, which allows seeing some interesting additional phenomena, and which will be discussed further below. One consequence of this color option (activated by checking the color PICT box in the dialog box for r vs. N) is that all of the computation must be done before the graph is drawn, so you get to watch the "beach ball" cursor for quite a while as it does it.)

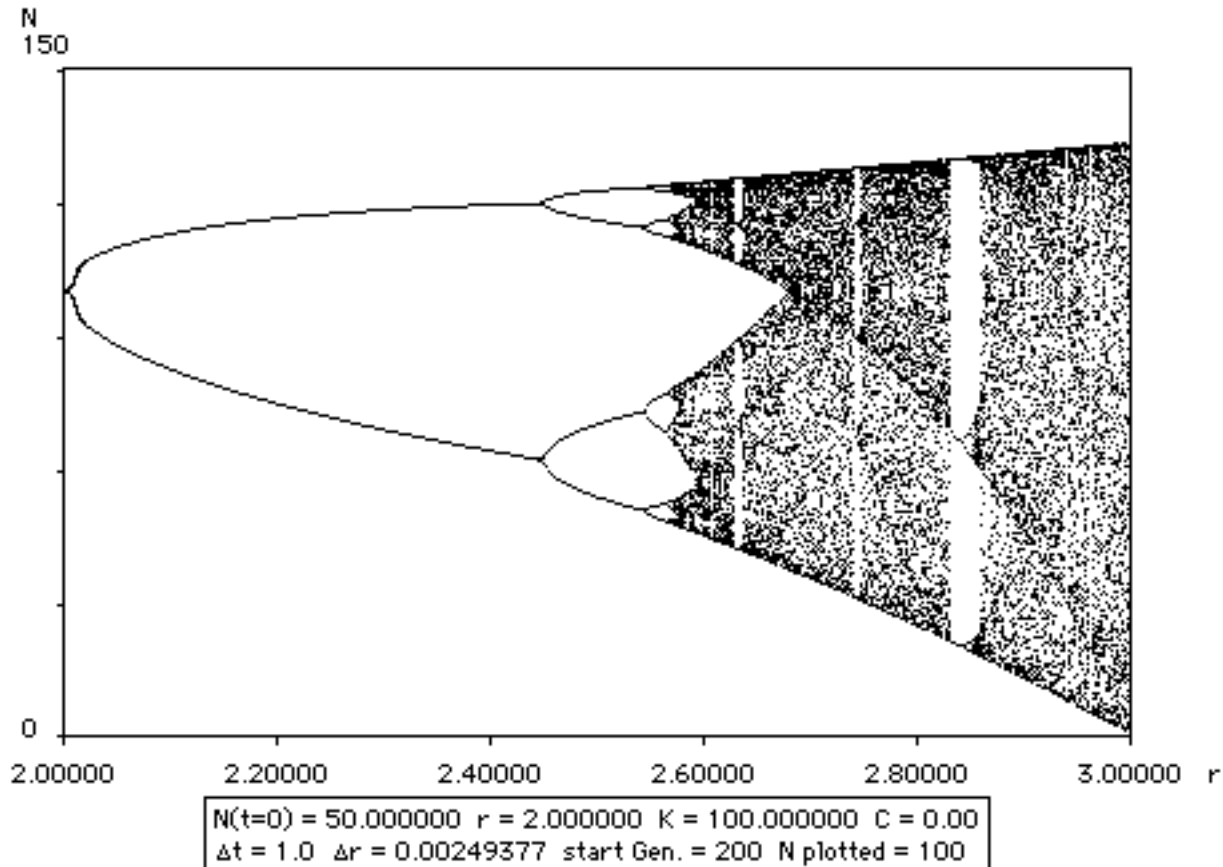


Figure 2-2: Bifurcation diagram showing the behavior that the system settles down to after 200 generations. (In this program, the first 200 generations are calculated, to give the system a chance to settle down if it is going to do so quickly, and then the next 100 generations are plotted.)

If you now want to see an overview of what is happening, with the stable and unstable cycles, look at Figure 2-2—which is called a bifurcation diagram. To make this kind of diagram, a population is started out from a given value of N and of r , and run for a given number of generations (called “start Gen” in the label for Figure 2-2), after which another determined number of points are plotted (“ N plotted”), all for that same value of r . The value of r is then incremented by a small amount, and the process is then repeated. The values of r are laid out along the x -axis, and the values of N along the y -axis, producing a special kind of N vs r plot. In general, if the population is at equilibrium for the value of r and the points plotted, they will all be plotted at the same place. If the population has achieved a stable 2-point oscillation, all of the points plotted will map back and forth into one or the other of them for that value of r . If the population is in a stable n -point oscillation, and at least n points are plotted, n points will be shown above that value of r , and so on. If the population is not at equilibrium, the points will move around. See if you can discover how they move around if (a) you are approaching equilibrium, (b) you are approaching a stable short-period cycle (say, 2, 4, or 8), or (c) you are moving chaotically.

In the r vs N plot in the program, you can adjust the number of generations calculated before plotting and the number of generations plotted. Smaller values for

the first number give you a sense of what the system is doing while it is settling down. If you are looking at this (called **transient behavior**, because it goes away after a while, unlike **equilibrium behavior**, which persists) it is useful also to use a fairly small value for the second number, so you don't get confused by a large number of points. You can experiment with these. What transient behavior you get, and how long it lasts, is a function both of the starting population size, N_0 , and of r and C . (Transient behavior is at least as interesting as the equilibrium behavior sought in the above "bifurcation plot," but seldom investigated—set the first number to 0 and the second to 5 for $0 < r < 3$. Can you relate this behavior to the $N(t)$ vs. $N(t+m)$ "phase plots" discussed below?) (Or see the following plot for generations 1-5 for $2.5 < r < 3$ produced in this way.) Smaller values for the second reduce the density of points in the chaotic region. Smaller values for either make the program run faster. See further descriptions in Appendix A.

These diagrams are usually plotted for enough generations that most of the things that are going to settle down into a stable cycle will have done so, and those that aren't going to make a band of points above that value of r . If we now look from left to right along the x -axis to compare what happens for different values of r , we see that a single line branches into two (at around $r = 2.0$), that these branch again, producing 4 branches (at around $r = 2.44$), that these branch again, producing 8 (at around 2.54), etc. These branchings or "bifurcations" are the onset of stable 2-, 4-, and 8-point cycles as r is increased through those values. In general, stable periods will be represented by a series of roughly parallel bands or "branches" for those r -values. If you want to see the curves settling down to their stable cycles, try a smaller value for "**start Gen**" (for more speed, you might try 25 for both **start Gen** and **N plotted**), and you will see the "trunks" thicken before the branching happens. (If they don't, try changing N_0 a little, and/or decreasing **start Gen** even more.)

This bifurcation plot predicts the first bifurcation to a 2-point cycle at $r = 2$, the second (to a 4 point cycle) at $r = 2.44$, etc. In particular, have a look at the ballooning thicket-like mess above about 2.57, and the bands of clarity or "windows" at various r values above (the largest readily visible being between about 2.83 to 2.85, although other bands appear elsewhere under greater magnification). These bands represent regions where there are stable cycles. They are bridged by twigs at different heights—one "twig" for each value in the cycle. Feigenbaum (1980) has shown that this diagram in fact predicts transitions in all nonlinear equations whose behavior gets more complex and ultimately goes chaotic via period doubling bifurcations.³⁰ See the Stewart (1989) book for a clear and deep exposition, though in fact, this particular case is so well known that virtually any book on chaos will do.³¹ With appropriate rescaling

³⁰A period-doubling bifurcation is a process through which, for increasing values of r , a stable period switches over to another stable period which is twice as long. Thus $2 \rightarrow 4$, $3 \rightarrow 6$, and $17 \rightarrow 34$ are all period doubling bifurcations.

³¹Physicists do the bifurcation diagram for a closely related but different equation (a kind of normalized or rescaled logistic), which they call the logistic. This equation is the general normalized quadratic form ($dN/dt = \lambda x(1-x)$) in which it was originally investigated for these purposes by Robert May. It behaves in a basically identical manner as ours does for the logistic (i.e., with $C = 0$) except that as our r goes from 0 to 3, the parameter along the x -axis there, λ , goes from 1 to 4, and that as our vertical scale goes from 0 to $4/3(K)$, that vertical scale goes (proportionately) from 0 to 1. With those caveats, you can read what is happening in most books on Chaos (e.g., Stewart, 1989) from this diagram.

for C/N , it works for the minimum-density limited logistic too, and for a lot of other growth equations.

You may have wondered how we got these numbers so accurately. You can find them using appropriate techniques by exploring the logistic, but you may also want to explore the bifurcation diagram using the “ N vs r ” option at the bottom of the **Simulation** menu. The bifurcation diagram furnishes a map for where (what values of r and N) interesting things are happening, and with this program you can blow up portions of it to see more clearly. To do so, select “ N vs. r ” from the **Simulation** menu. The minimum and maximum values for N and r specified there are default values. If you select minimum and maximum values that are closer together, you will get a blowup of that region. Thus plotting r from 2.8 to 2.9 produces a magnification of that region 30 times as great as doing a plot from 0 to 3. Use the Record coordinates option at the bottom of the **Simulation** menu, and click the mouse on locations that interest you to record their coordinates in the Run Record. A lot of the details of what is happening require the other views provided by our N vs t , and “ $N(t)$ vs $N(t+m)$ ” plots, however. These three, and another plot type, “ $Ne(t)$ vs r ,” to be discussed later, confirm and complement one another. Some things are best studied using one view mode, and some are best studied using another. Like most interesting things, this behavior is worth several different views!

As mentioned above, one can also get different views of what is going on by taking different time slices of the r vs N plot. The next figure, Figure 2-3, takes the first 5 generations by setting the initialization generations at 0, and the plot generations at 5, and then the next 5 generations by resetting the initialization generations at 5. With this setup, you can take different windows (of whatever width you choose; determined by the number of generations plotted) at whatever time slice you want; determined by the number of initialization generations. This raises a lot of questions considered in the box below. You might want to try some of them out. (If you are reading this on a computer with a color monitor, you will see that Figures 2-3a and 2-3b are drawn in color. Color PICTs can be drawn and saved with the r vs N mode by checking the Color PICT option in the dialog box, and then either copying the graphs to the scrapbook or using the Save PICT option in the **File** menu. These files can be quite large (roughly 8 times larger than black and white bitmaps), and it is quite easy to make a color PICT in this mode which is too big for you to save. If so, decrease the number of generations plotted, or decrease the graph width using the Graph coordinates option. The last thing to notice about these color PICTs in the r vs N mode is that color codes for generation number. The same seven colors are used in a cyclical repeating order in all graphs: The first seven generations will be colored red, rose, yellow, green, light blue, dark blue, and black, followed by a repeat. Thus, generations 1, 8, 15, 22, 29, ... , $1+7n$, ... , will be colored red; 2, 9, 16, ... , $2+7n$, will be colored rose, etc. Thus, red appears in trajectory 1 (the constantly climbing shallow sloping line near the top of the top graph), and in trajectory 8 (the undulating bottom curve starting at the leftmost edge of the bottom graph). You might try producing something like Figure 2-2 with the color PICT option on. Do you see any places where well-defined branches (in the non-chaotic regions or bands) switch colors in the middle? If not, try another value for N_0 . Can you explain what is happening here? Try looking at the transient behavior of earlier generations. Can you give a general explanation for what is happening here?

Can you get a generalizeable explanation for the occurrence of phase shifts with changes in N_0 or r ?

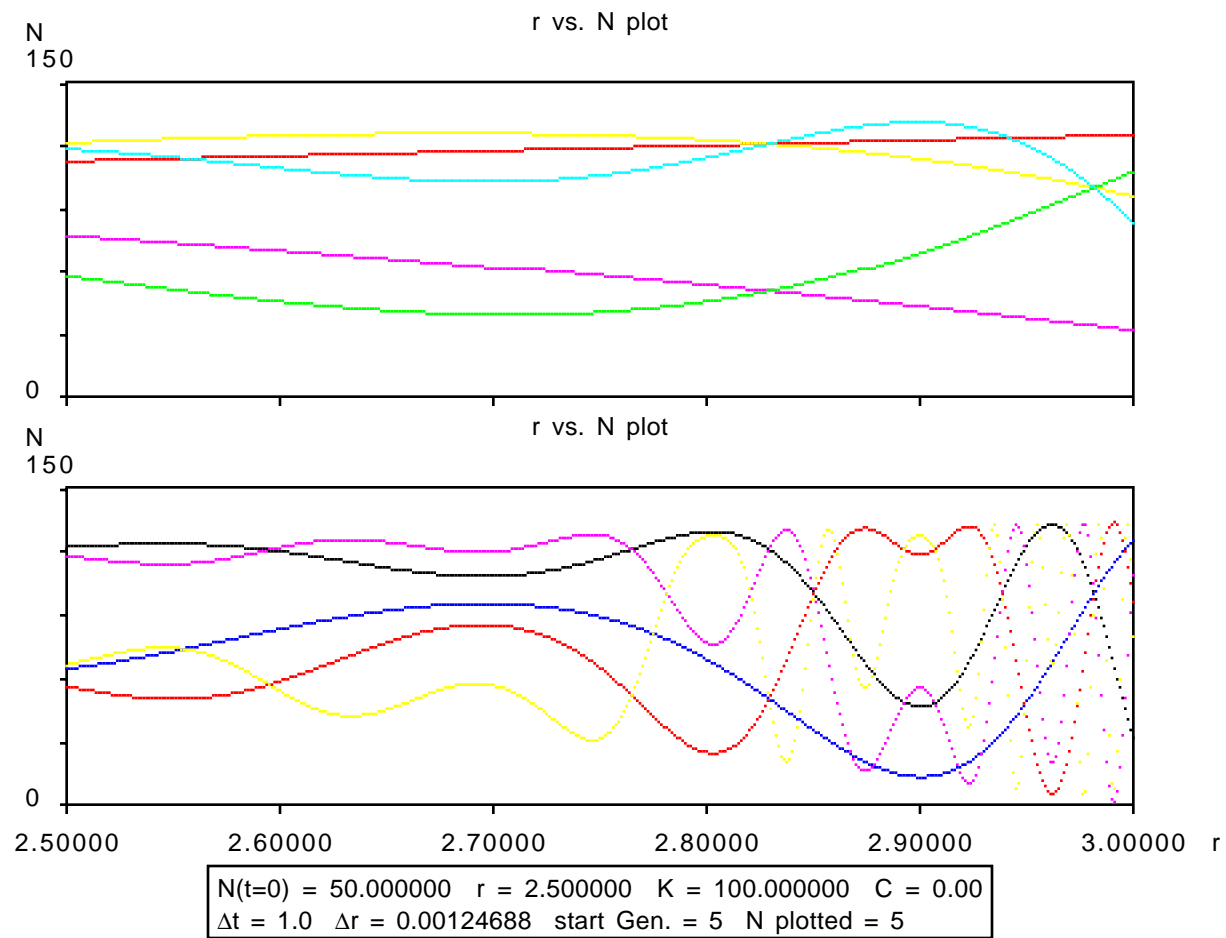


Figure 2-3. r vs N plot for $2.5 < r < 3$, and $0 < N < 150$, with start generations = 0 (top) and = 5 (bottom) and plot generations = 5. The starting value of N is also important in determining what the individual curves look like.

Exercise 4: Tasting Time Slices of Bifurcation Plots

Look at the curves for N as a function of r in successive generations in Figures 2-3a and 3b, or experiment with making some yourself. What does it mean when a curve goes through a local maximum or minimum in this kind of plot? What is happening the generation before and the generation after? Can you explain why higher generations get curvier?³² What (if anything) does it mean when trajectories cross? Are there any differences in the cases where two curves cross and where multiple curves cross? Consider the generations in which they cross. You could (in the program, by clicking with the cursor) find out the locations (the N and r values) where interesting things are happening (intersections—and which generational curves are intersecting?), maxima or minima) and see what happens at or around these values using the N vs r or $N(t)$ vs $N(t+m)$ plots. Can you relate this to the $N(t)$ vs $N(t+m)$ "phase" plots described below? Save a plot with well defined local maxima and minima (i.e., one with low generation numbers). Now change the starting population size (N_0). What happens? Why?

It is now time to introduce the third and most general of our growth equations, the minimum-density limited logistic, or MDLS for short.

Let's add a bit more realism and get the minimum density limited logistic:

$$\frac{dN}{dt} = rN \left[\frac{K-N}{K} \right] \left[\frac{N-C}{N} \right]$$

where:

N , r , K as before.

C = minimum density threshold (e.g., reflecting the probability of finding a mate, or lower thresholds for density or numbers where cooperative inter-actions are important, as for warmth, cooperative hunting, or cooperative defense against predators.)

We've done the analysis relating the logistic to the exponential. Now it's your turn to do the analysis relating the MDLS to the logistic. If you're not sure what happens, explore it with the simulation. The following exercises (in Box 5) should give you a lot of points for comparison:

³²Be careful here—these curves are not trajectories, but functions relating how the population size varies through the first 5 generations as a function of changes in an underlying process. To put it another way, the growth equation is actually changing (because the value of r is changing) as you move along the x-axis. Nonetheless, the different generations are represented here—by different curves. (The situation is thus just the opposite of that in the N vs. t plots, where the different curves represent different values of r or of N , and the distance along the x axis or along the curve the time).

Exercises 5: The MDLS and Its Relation to the Logistic

1. Set $C > 0$, but a relatively small fraction of K (i.e., if $K = 100$, let C be 20.). For some fairly small value of r (say, about 0.1 to 0.3) click on the top radio button, and set N close to but $< C$. Choose an increment in N such that 5 runs will start N at values bracketing C from below and above. (Choosing $N = 18$ and $\Delta N = 1$ will do for a start if $C = 20$.) Explore lower and higher values if you like.
2. How does the behavior differ from the logistic? Evaluate their behavior when N is $> C$, but close to it, for how fast N approaches K , and for the stability of the curve near $N = K$ for different given r values. Can you explain why the MDLS is stabler?
3. Can you predict when (i.e., for what r values) the MDLS will bifurcate (relative to the logistic) for different r values?
4. Try varying C to see what happens. Describe qualitatively the behavior of the MDLS, by analyzing $[(N-C)/N]$ for different values of N , and show that when C is set = 0, the MDLS reduces to the logistic as a special case. Given this, under what conditions would you expect the behavior of the two equations to be strongly similar? What conditions would you look for to test which of the two equations best fit the behavior of a population? Where would you expect their agreement to break down? Is this the same question?
5. Try some of these exercises using N vs t plots, and then compare the results for that value of C over a range of r -values using an N vs r plot. How are they similar and how do they differ? Try N vs r plots for different values of C . What regularities do you find? Is there a relation between K , C , and the r -value for which the lower branch of the "fig tree" in an N vs r plot intersects the line $N = C$? For the altitude (the N -value) of the upper branch at this r -value?
6. What does all of this mean biologically?

The Importance of Time Lags:

The time lags represented in these equations by Δt are very important to understanding the dynamics that they predict. Δt , remember, is the time lag between when something happens (e.g., an increase in population size) and when that change has a causal effect on further changes; it is thus a response lag or reaction time. We do things on a number of different time scales, and a failure to adjust or respond on the right time scale can cause problems. We watch our balance as we walk, and as we spend. We are fortunate indeed that our bank does not require us to respond to our checking account balance in fractions of a second—and we don't (and indeed cannot) walk so slowly that adjustments to the local terrain will do as long as they are in by the end of the month. Richard Levins (1968) tells the story of a species which has two generations per year, one each in winter and in summer, and always makes for the next generation the phenotype that was best for this one. This creature would always be "out of synch"—better if it could either predict what environment its offspring would occupy, or reproduce much faster, while this generation's phenotype was still good. We began accumulating significant amounts of CO_2 in the atmosphere soon after

1800—about 7 or 8 generations ago. We first detected that rise in the late 1800s, and have only recently come to appreciate its significance in the “greenhouse effect.” One thing everyone agrees upon is that we would have a much easier time adjusting now if we could have come to an agreement on its significance much earlier. The next group of exercises show indeed that our sense of the importance of time lags, and responding with appropriate speed, is reflected in the equations modeled here:

Exercise 6: Time Lags in MDLS

1. Time lags turn out to be very important for analyzing and understanding the dynamic behavior of these equations. To see this, try everything you have tried before for a Δt of 1 for different values of Δt , say of 2, or 1/2, 5 or 1/5, etc. (Notice the correlated constraints given in the dialog box on Δt and the number of generations. The main thing to remember is that you can't plot more than 400 points in a trajectory.)
2. Why should changing Δt make a difference? Why do you think that the behavior is so sensitive to relatively small changes in Δt ? How do these changes affect behavior?
3. Keeping K and C constant, can you guess an equation relating r and Δt to predict period doublings in the logistic relative to when they occur when $\Delta t = 1$?
4. Can you guess an equation relating r , Δt , K , and C to predict corresponding period doublings (or “bifurcations” as they are called) for different values of these four quantities?
5. Does the kind of dependency on Δt you find make biological sense? Why should small Δt 's be stabilizing, and big Δt 's be dangerous? Suppose someone says to you, “Chaos can't be important for this species because its growth rate is too small,” but you know that it interrelates with a variety of other species which also interact with each other in such a way that causal effects could propagate for a long way (and a long time) in the ecological web before coming back. On what time scale is its growth rate small? Is this species safe from chaos? (It should be possible to further study this multispecies case using BIOTA.)

A Simple Mathematical Artifact, or How Strange Effects Can Creep in!

Here is a mathematical artifact of the kind we always have to look out for in mathematical modeling. We said above that negative r values are meaningful for the exponential growth function, where a negative value produces exponential decay rather than exponential growth. Suppose we try a negative r value in the logistic and MDLS. What happens (a) when $N < C$, (b) when $C < N < K$, and (c) (**especially!**) when $N > K$? What happens to the stability or instability of behavior at $N=K$, and at $N=C$, i.e., what happens to N for small fluctuations above or below each of these points? Do all of these make biological sense? (Actually none of them do, but case (c) above seems most strikingly paradoxical to most people. We know that some people thrive under stress, but this is too much! So you shouldn't trust everything you read in a mathematical equation.)

What went wrong? Simply that when rationalizing the use of the additional multiplicative factors, $[(K-N)/K]$ and $[(N-C)/N]$, to turn the exponential into the logistic and minimum density limited logistic, people had just assumed that the growth rate

was positive (remembering Darwin and Malthus' comments on geometric rates of increase perhaps). Given this, it would be natural to assume that the exponential that worked just fine by itself, and the two additional factors which worked fine for positive r 's, would work fine together under all of the conditions when the exponential did. But they don't. (Figure out for yourself what goes wrong for the logistic model with negative values for r .) This is another maxim of model-building (and computer programming, and experimental design too): parts that work well in isolation need to be checked out to see how they work together under the important variations in conditions in which they are likely to be used.

These discussions have illustrated another important strategy of problem-solving, namely, **problem decomposition**, made possible by the peculiar relation of the exponential, logistic, and MDLS equations. Their behavior was discussed in terms of the behavior of 3 modular terms, Nr , $[(K-N)/K]$, and $[(N-C)/N]$. We first discussed the growth equation for the first module or term, then discussed the behavior of the second term or module, and how its addition affected the behavior of the first, and then encouraged you to investigate the behavior of the third module, and how its addition to the first two modules affected the net behavior of the three together. This strategy, which could be called "divide and conquer" is a very powerful and effective problem-solving strategy which saves an enormous amount of time and effort. It was first effectively elaborated by Herbert A. Simon in a series of essays reprinted in his (1981). Problems for which the divide and conquer strategy works (or systems which can be analyzed in this way) are called "nearly decomposable," and this general approach to problem-solving characterizes "reductionistic research strategies" (see Wimsatt, 1980b).

Some Examples of the Use of Various Options in Plotting N vs t Curves:

Figures 2-4 through 2-6 illustrate the use of the MDLS program for simulating the logistic, i.e., with C set = 0. These figures show:

- The effects of increasing r (each of the next three figures illustrates the behavior of a set of curves which differ in r).
- The dynamical effects of changing Δt (Figure 2-5 is produced by the same conditions as Figure 2-4, except that Δt is .5 rather than 1, so the curves are sampled twice as often and, dynamically, halving Δt has approximately the same effect as halving r)

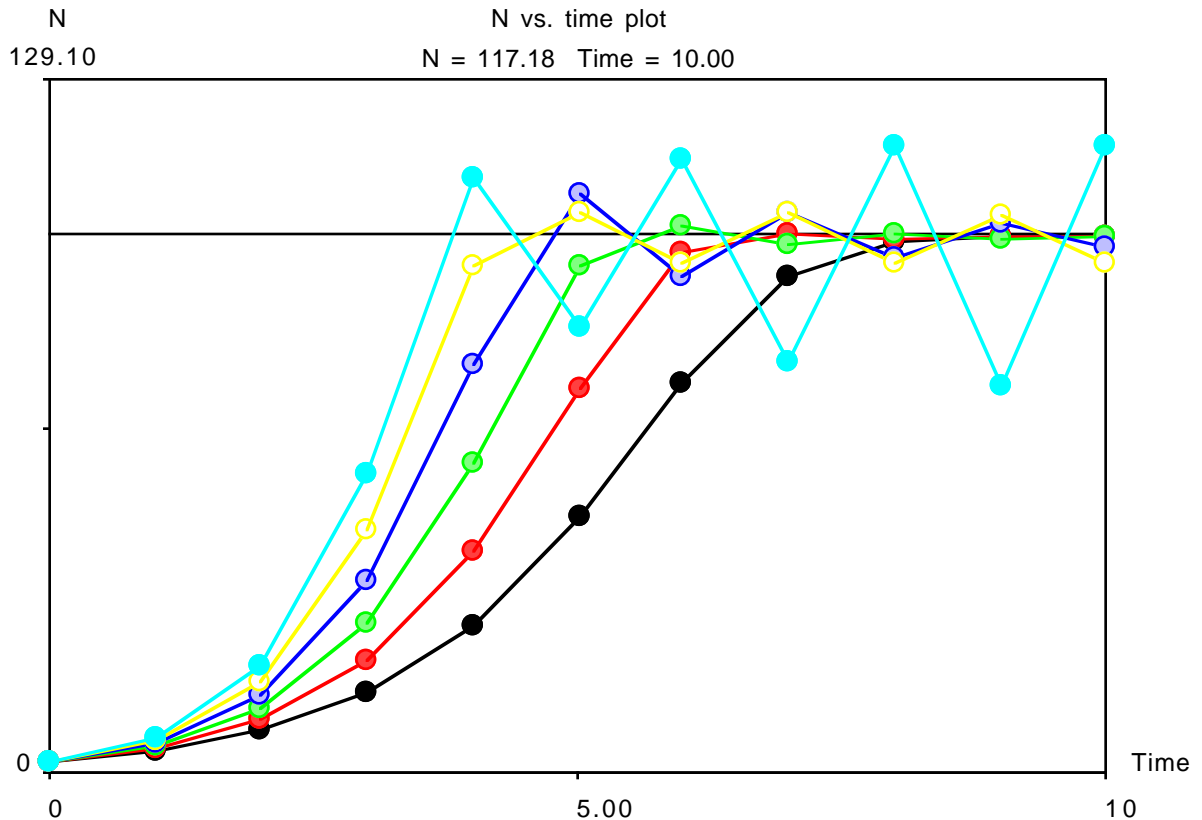


Figure 2-4. Starting parameters: $N = 2$, $r = 1.2$, $\Delta r = .2$, 6 runs; $\Delta t = 1$; $C = 0$; $K = 100$.

- Figures 2-4 and 2-5 also show a number of features (discussed in Appendix A on the use of the programs) which allow a lot of freedom in representing the data, and help in its analysis:

- autoscaling**: the maximum value is always scaled at 10% higher than the maximum point plotted in the graph. Aside from producing well-scaled graphs, this also has the benefit that the value on the y -axis (129.10 for Figure 2-4 vs 110.05 for Figure 2-5) always accurately indicates the difference in these maxima. (To get the maxima, multiply these values by 10/11.)

- using **grids** (in Figures 2-5 and 2-6). Both dimensions of the grids can be manually set in ways to facilitate measurement, pattern detection, and/or detection of periodicities.

- change of **graph coordinates** for minimizing **pixel roundoff** (noticeable in the height difference in Figures 2-4 vs 2-5)

- changing **point size** (4 units in Figures 2-4, 2-6; 3 units in Figure 2-5)

- the use of the **point grabber** cursor to get the coordinates of the first point in Figure 2-4, and the highest point in the first (displayed under the graph titles in both cases). Note effects of pixel roundoff.

- with the Record coordinates option activated (at the bottom of the **Simulation** menu) clicking on the mouse button not only grabs the

point coordinates to the top of the graph, but prints them in the Run Record Window. Thus a number of points may be recorded (with the **graph parameters**, which are automatically sent to the run record each time a new data set is created or a new kind of graph is drawn from the data), and comments on these points may be added either interstitially, or at the end of the list, creating an annotated lab record of the runs done using the program. (More details on this are found in the text at the start of the Run Record Window.)

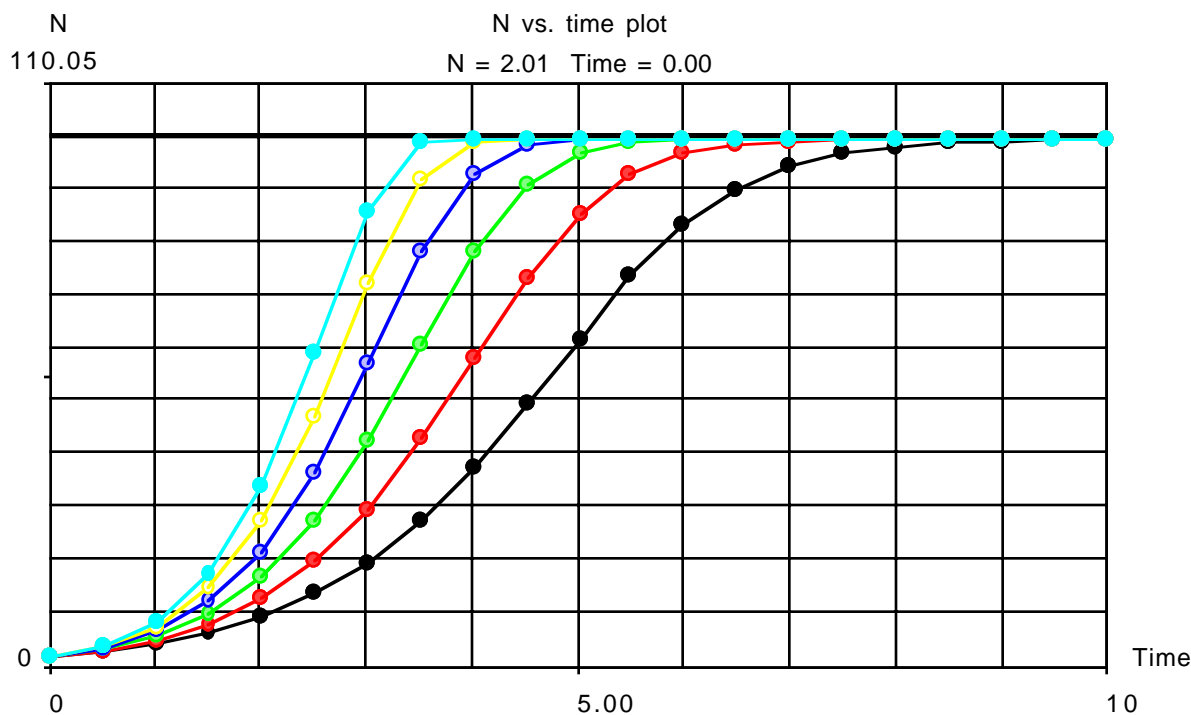


Figure 2-5. Same parameters as before except $\Delta t = 1/2$, and the Grids option is used.³³

- Figure 2-6 illustrates the use of the "get N" utility on the first graph to pick a starting value for N in Figure 2-6. The N-value chosen is the stable 2-point amplitude of one of the curves with an intermediate r -value, illustrating how curves with still lower values of r decay in amplitude from that point while curves with larger values of r amplify. The **point grabber** is particularly useful in stability analysis, since you can get the coordinates (in particular the N -value) of points or regions where things appear to be converging or diverging.

Several things other than the parameters of the model have been changed in Figure 2-5:

³³ Note that the curves grow faster (like interest that is compounded more frequently). Note also that none of them overshoot—they are stabler than the curves of the first graph because the population reflects the effects of its increased size after $1/2$ generation rather than after a whole generation.

- Most obviously, a **grid** has been added, and it has been customized: the **autoscaling** default grid would have the 10th division at the highest point plotted, and one division from there to the top, at 110.05. (From this reading at the upper right of the graph, we can get the highest point plotted as $(10/11) \times (110.05) = 100.045$ —a very slight overshoot or pixel-rounding error. You can check on this by opening the **Grids** menu. The autoscaling default menu will have a spacing in that box which is $1/10$ of the maximum value plotted to several decimal places. In fact, the **Grid** menu doubles in this way as a place to read out the **maximum value plotted** in any graph.³⁴ The grid here has been changed to have a horizontal line every increment of 10 individuals in population size, putting the value $N = K$ at the 10th division. (This same change has also been made in Figure 2-6 below.)
- Similar calculations on the top coordinate of Figure 2-4 yield a highest point plotted of 117.36. As noted above, this point was measured by the point grabber at 117.18. The error here is due to “pixel roundoff.” See the box on the next page for a discussion of pixel roundoff and the use of the graph coordinates settings.

³⁴This works both for the logistic and the blending inheritance simulations, and will come in handy in several different contexts later.

Pixel roundoff and graph coordinate settings:

A “pixel” is the minimum sized data point or block of visual space under control of the computer on the screen. Lines and other images appear a little jagged because they are not continuous, but made up of pixels. The point grabber has to get the location of a pixel wherever you put the cursor when you click the mouse. That is in the center of the plotted point—if, that is, *you* have centered the cursor accurately over the point! But, placing the cursor as accurately as possible does not eliminate the possibility of error. Pixels have a definite size. So when the point is plotted by the program as it constructs the graph, it is plotted as centered at the location of the closest pixel, but this is almost never exactly the same as the calculated value. So there is an error—in this case of $.18/117.36 = .15\%$. Most pixel roundoff errors fall in about this range. (If you try to replicate this run, you won’t get exactly the same results, unless you use the same monitor—a standard 640 x 480 pixel Mac II monitor in this case—since the default graph size is automatically scaled to the pixel dimensions of the monitor, and the pixel errors are a function of the pixel dimensions. Pixel errors are smaller on larger screen monitors, but the ratio in linear pixel dimensions between the smallest and largest Macintosh screens is only about 2:1, so using a small screen is not a severe disadvantage on that score.) Pixel roundoff can occur in the location of everything in the graph, and in fact generates the double width of the line at $N = K$.

It can be minimized by (1) making the vertical dimensions of graphs = $Nk + 1$ where k is the largest integer such that $Nk + 1$ is less than or equal to the difference between the top and bottom coordinates given when you boot the program or first open the graph coordinates dialog box. N for these purposes should be the population size listed in the upper left hand corner rounded off to the nearest integer. (As explained there, those coordinates are the largest allowable for the screen size of the computer you are using.) When you do this, the vertical scale is k pixels per individual. With these settings, it is usually also easier to make K some simple round number like 100—except of course when you are investigating the effects of changing K . The corresponding corrections in the horizontal direction occur when using, instead of $Nk + 1$, the largest integer h , such that $(G\Delta t)h + 1$ is \leq the difference between left and right coordinates given, where G is the number of generations. Pixel roundoff is probably most easily detected in the horizontal dimension when plotting several runs for a large number of generations (100 or more) with the line option activated and where stable cycles are occurring. In this case, a kind of Moire effect or local imperfections in the pattern are usually seen which can be eliminated by changing the horizontal graph dimensions as suggested above. (See Appendix E below for examples.) In fact, it was our desire to be able to eliminate this phenomenon (which renders the analysis of periodicities more difficult) which originally motivated this option. Pixel error is one of the two kinds of computer-induced error you can see using this simulation—the other being computational roundoff error.

You can **calibrate pixel width** in the x and y directions for given values of the graph coordinates and ranges of graphed values by computing them, but it is actually easier to do so using the **record coordinates** option. Simply pick a point in the graph where a horizontal and a vertical straight line intersect. Using the cursor crossbars, click right at their intersection, and one pixel above, below, left, and right. This in effect gives you two measurements for cursor width in each direction to check against each other. (If you err by one pixel, one of the two measurements in that direction should be twice the other.) Since many changes in the variables, and all changes in the graph coordinates, will change the increments in x and y signified by a pixel, and the size of a pixel will be a good indicator of probable error in picking points off the screen, you may wish to do this often.

Exercise 7: Values of r , Overshoots, and Oscillations

Try to change r and Δr to get the same stability or overshoot effects in Graph 2-5 as in Graph 2-4. (You won't be able to simultaneously match the growth rates.) It is useful in doing this to find a reference point to compare in the two simulations. A very useful one is to find out for what value of r in each case the oscillations have the same amplitude. Do this for several different amplitude pairs.

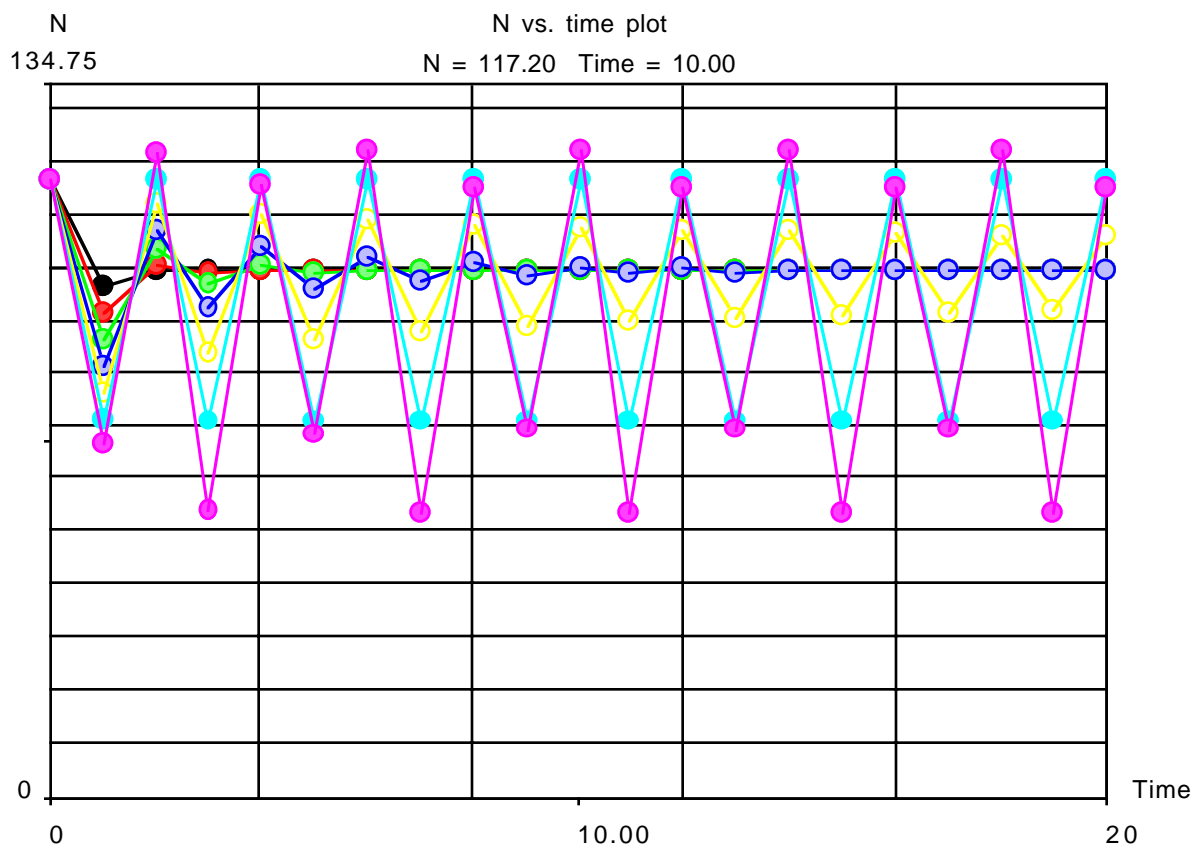


Figure 2-6. Graph illustrating that equations having different r -values have different equilibria. All 7 runs start at the same N -value, separating in the first generation due to their different r 's. The curves with the four lowest r -values decay rapidly to a single-point equilibrium; the fifth decays slowly; the sixth remains at its initial amplitude in a stable 2-point cycle, and the seventh amplifies to a larger amplitude 4-point cycle. Note that the grid has been set horizontally to a period of 4 generations to emphasize the cycle. This can be very useful as we will see below.

Other Kinds of Systematic Investigations with the Logistic Models—More Advanced Explorations of the Capabilities of the Simulator:

Other graphs illustrate other possible kinds of systematic investigations. The number of possibilities given the self-similar variety of the bifurcation map is virtually limitless. These are but specific examples of kinds of investigations you can do in many different places. They have been chosen in part for their biological interest, for their usefulness as a technique when investigating dynamical systems, or sometimes both. This presentation is a mixture of teaching by illustration and suggested things you can do to carry the investigations further. There are often other and sometimes even better ways of doing these things. These ways have been chosen for their intuitive character, but we encourage you to try your own.

Finding Boundaries and Widths of Attractors:

A domain of attraction or "attractor" is a range of values of a variable or set of variables within which a system converges to an equilibrium or recurrent behavior. It can be thought of by analogy with a drainage basin—a geographical region within which water drains into a more compact point, region, or trajectory. Boundaries between attractors are like ridges—water to one side of the boundary goes one way, into one basin, water on the other side of the ridge goes the other way, into another basin—like the continental divide. So-called "adaptive topographies" or "fitness surfaces" are constructs for describing the evolutionary dynamics of populations of organisms which are quite like this, in that populations evolve toward the peaks in their surfaces (regions of greater fitness) just as water flows down. If you know the analogs to contours in such a surface, you can predict the direction of change for that system. The number of attractors for a system is the number of persistent types of behavior or behavioral outcomes it can exhibit. (This ignores the question of the diversity and length of transients.) For more on domains of attraction, see W. R. Ashby's *Introduction to Cybernetics* (1956). Knowledge of the attractors of a system and the location of their boundaries is extremely important for understanding their behavior (see also Kauffman, 1992).

Simple attractors are one-dimensional objects—ranges of population sizes, which may be thought of as marked off along segments of the y -axis in an N vs t plot.³⁵ There are two different ways for estimating the size of attractors for this system. The most obvious way is to find the location of each of the relevant boundaries and add up the lengths of the relevant line segments. Thus, if this is a region where the system tends to fall into a three-point cycle, you want to know the distance around each point within which the system is attracted to that point. Since the system moves among the

³⁵ In fact, one of the characteristics of chaotic or fractal systems is that their attractors are not always so easy to characterize. In fact, when a system is in a region of chaotic behavior, the domains of its attractors are inextricably intertwined almost everywhere. Compare the case of the discrete-state logistic for $K = 20$ discussed below with this one here, and ask what happens to a system that behaves like the discrete state logistic as the number of states and number of cycles increases without bound.

three points in order, the total size of the attractor is the sum of distances of attraction around each of the three points.

A sometimes much simpler way is to randomly pick points along the line and ask whether they are in the domain of the attractor or not, determining that by which direction trajectories move from there. If we can reliably score whether or not they are in, then the relative size of the domain of the attractor should be the number of "hits" divided by the total number of tries. (This procedure is a paradigmatic example of a Monte Carlo simulation. It is particularly interesting because it extends to cases which involve an arbitrary number of dimensions. Enclose the n -dimensional region of interest in a box (an n -dimensional one, that is) of known size. Then an estimate of the proportion of that box occupied by that region is obtained by counting the proportion of tries within the box that also fall in the region. The more tries you make, the better the estimate should get. Of course, this assumes that your random number generator is truly random. You can also turn this exercise around: if you already *know* the proportion of the box occupied by the region, or can find it out by another means, you can use that to *check* how good your random number generator is. That's what we will do below.

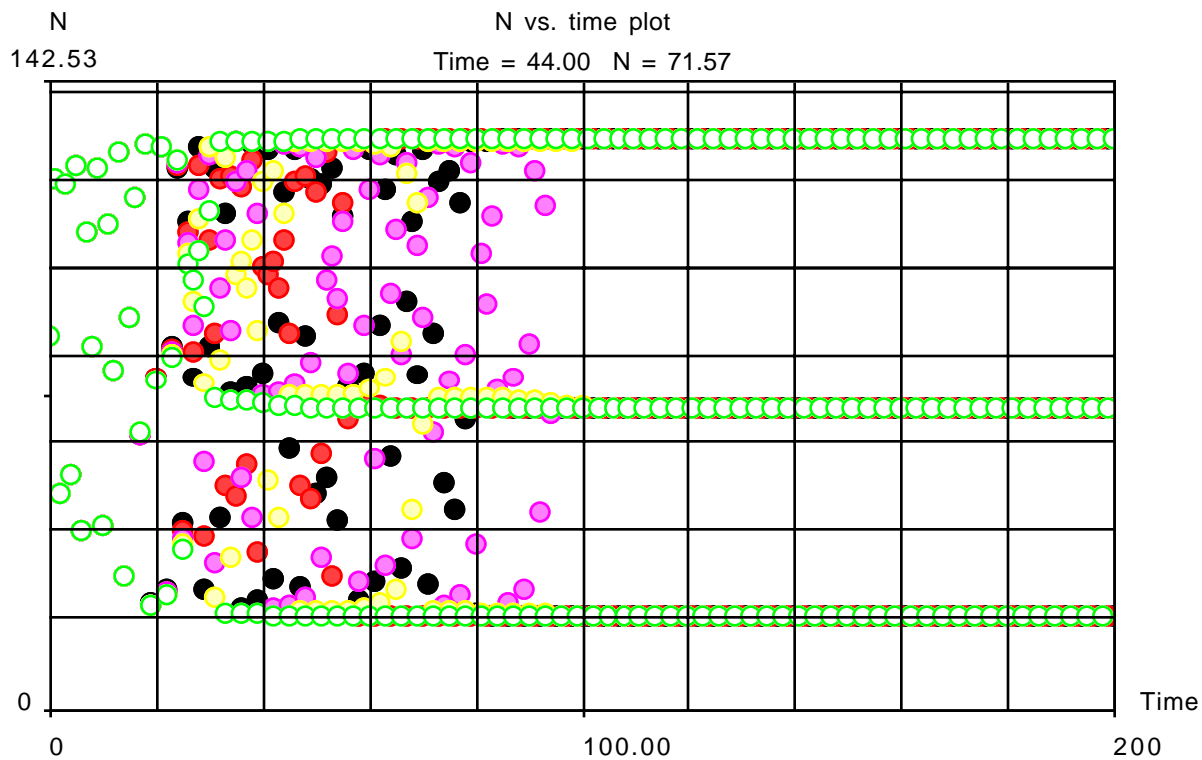


Figure 3-1. N vs t plot for the logistic ($C = 0$), with $K = 100$, $\Delta t = 1$; $r = 2.83$, initial $N = 85$, $\Delta N = +.00001$, for 5 trajectories for 200 generations. Points are size 4, in pattern mode. Vertical and horizontal grid divisions of 20.

We will look at a series of 5 trajectories run at $C = 0$, $r = 2.83$ (near the beginning of the period 3 window visible in the " N vs r plot" in Figures 2-2, 3-2, or 3-6), the top radio button set to run increments in N , starting at $N = 85$, with a positive increment of .00001 (so we get 85.00000, 85.00001, ... , 85.00004.) This plot illustrates nicely the **infinite sensitivity to initial conditions** characteristic of the chaotic regime, since you

can see from the graph that there is no noticeable difference between the trajectories for about the first 17 generations (since the white circles of the last trajectory plotted totally cover the earlier different patterned circles), after which they diverge rapidly so that within about 10 generations, they are essentially randomly scattered between the bottom and top extreme values. They then drop apparently randomly into the middle (and largest) hole of the 3-point attractors, where they converge rapidly to a stable 3-point cycle. Here we see the infinite sensitivity twice over—through the initial divergence around generation 17, and through their essentially random order of capture by the 3-point attractor. (Check for yourself that their order of capture is not a simple function of their starting values.)

Indirect method—Monte Carlo Simulations and a Test of the Randomness Assumption

The last trajectory to drop into the attractor in Figure 3-1 is the middle (medium gray) one, at generation 93. (If you are using the program, check that the rightmost point not in one of the 3 lines is it by clicking in the middle of it with the point grabber to identify its generation) See if you can identify where the other trajectories fall in, by finding their rightmost points not in one of the 3 lines, and adding 1. Now if trajectories in the chaotic regime are really random (and you will have to decide when (or if) this condition is satisfied), then the ratio of the number of points falling in to the total number of trials (i.e., for each trajectory, starting from when you think they are randomly distributed, to when each trajectory drops in) should be equal to the ratio of the total width of the attractor to the total range of points explored. How wide do you estimate that the 3-point attractor is at $r = 2.83$? (It needn't be the same width at other r -values, so this indicates other projects you can do of the same form. Or the members of a class could each do this exercise for a different r value constituting a series spanning the window and plot the appearance, growth, and decline of the attractor.)³⁶ This kind of probabilistic estimate of a parameter is called a Monte Carlo simulation, and is very important in the analysis of dynamical systems. They are only as good as their randomness assumptions however, and are limited in accuracy by sample size.

Direct Method for Estimating Boundaries Using Spanning Trajectories and the Cheap Way out for an n-Point Attractor

We can check this estimate, since there is another way of estimating it more directly. I will illustrate with the top boundary of the middle (and largest) region of the attractor. With the point grabber get the coordinates of the start of the lowest trajectory that diverges (light gray, $N = 71.27$ in generation 44), as well as the highest trajectory that converges (same, $N = 70.36$, a few generations later.) Run a series in N

³⁶ Some of these questions, such as where the three point attractor appears, and when (and how) it bifurcates are probably better explored with the N vs. r plot mode and "blowing up" the r -scale. In general, when given a problem specified in terms of a plot mode, it is worth asking whether you can learn anything about the phenomenon by using another plot mode to supplement or check your analysis. And even if another plot mode is not essential, you get a much richer understanding of the phenomena by seeing how they look in different plot modes. This is a generalizeable lesson: ***look at everything you can get in as many different ways as you can.***

with the increments chosen to span this interval. You don't have to match it exactly—thus 10 trajectories from 71.27 down to 70.37 will probably capture it, and have the advantage of a nice increment of 0.1. Trajectories at one end will fall in, and those at the other end will fall out. Pick the two successive trajectories in the series which split in this way, and use them as endpoints for the next series. In this way, you should be able to improve your estimate of the boundary by a factor of i in each generation, where i is the number of iterations you do. Determine top and bottom boundaries for the middle and the 2 extreme cycles. (There's also a "cheap way out": if you got the top and bottom boundaries for the bottom region accurately, you should be able to run each of these points for 2 more generations (one point and 1 generation at a time) to get the other boundaries accurately using the autoscale feature of the grids option to determine the maximum value plotted, since values on an edge for one region of the attractor should be very close to an edge for the other two. Although it should be falling slowly in or out, the rate of divergence can be estimated in the same way by comparing its value at the top of the first and second complete cycles.

Exercise 8: How Good Is Chaos as a Random Number Generator?

Now calculate the ratio to the total range as before and compare with your other results. If there are significant differences between your estimates, can you explain why this might be so in terms of the randomness assumption and what you have found out about the behavior of trajectories? This exercise is in effect a check on how good a random number generator chaotic processes are. Note that if you wish to check an assumption used in generating an answer, you have to find a way of getting that answer without using that assumption, and compare results. If they agree, the assumption is probably OK. (Not necessarily though; this is a good heuristic, but not foolproof—see Wimsatt, 1981). If they disagree, you know that something is wrong in one or the other approaches, and you have to try other things to figure out what. In this case, there is not much else that could be wrong, so the randomness assumption is suspect.

The Janus Face of Computational Roundoff Errors Seeing Similarity Where There Is None (The Loss of Significant Differences) and Creating Differences Where There Are None (The Loss of Significant Similarities)

Exercise 9

One other thing to try with this plot before you leave it: do 10 runs starting at $N = 85$, with an increment of .000001. What do you see? Can you find the value between .00001 and .000001 for which this change occurs? (As accurately as possible!) Try (for another exercise with the opposite result) comparing trajectories for which you increment up to a value vs specify that value directly. (For example, compare the behavior of the 5th trajectory in the series, $r = 2.81$, $\Delta r = 0.25$, with the supposedly identical trajectory for $r = 2.91$.) Do you get the same result? (Compare a point not in the attractor as far to the right as possible using the point grabber.) This should alert you to a second sort of roundoff error: Just as **pixel roundoff** can cause errors, so can **computational roundoffs**. Are these numbers direct estimates of the smallest difference the computer can handle? Why or why not? (You may want to look at the discussion of the last case in Appendix E.) This will be discussed further below.

This exercise and others like it can also be performed, and the effects of roundoff error studied more systematically, by using the Numerical Precision and the Round N options available under most plot modes.

If checked, the Numerical Precision option (available under the Initialize and Plot... selection from the **Simulation** menu) allows setting the number of decimal places used in ALL calculations involving any of the model or simulation variables (N , C , K , r , Δt , ΔN , or Δr) to 0 through 6.³⁷ These variables are rounded to this number of places after each computation through the use of special operators for +, -, *, and ÷, which automatically perform this rounding to the set number of decimal places whenever one of these operations is used, in all intermediate results, as well as at each generation. [Number of significant figures can be further varied by changing value of K , e.g., to 1, 10, 100, 1000.] Precision without rounding is between 8 and 9 places total, or for default values ($K = 100$), between 6 and 7 to the right of the decimal place. These options can be used for studying the effects of roundoff error and different degrees of precision in numerical calculations. This option is currently available for the N vs t and the $N(t)$ vs $N(t+m)$ plot modes only.

The "Round N " option, if checked, rounds the population size, N , to the nearest integer after each generation is calculated; this integer value is used in all calculations involving N for the next generation. (This option is accessed as above for the N vs t and $N(t)$ vs $N(t+m)$ plot modes, but it can also be chosen for the r vs N plot mode. According to some criteria, the use of this option ought to make for a more realistic

³⁷In cases where it is inconvenient to have this rounding act on ΔN and Δr , it may be preferable to do only 1 run at a time with this option activated, since these act only on the 2nd and later runs.

model, because organisms come only in whole numbers, but you should decide this for yourself. If you decide that this model is less realistic in its behavior (try it particularly for small values of K), can you say how its behavior is less realistic? Can you explain why the model might behave less realistically even though it models the real-world situation more realistically? Can you think of a further change in the model—not necessarily one you can do here—that would improve the situation? (Look at Appendix D on assumptions of the logistic, or to the next section below for ideas if you don't know what to do with this one.)

Chaos in Discrete State Systems, Large and Small: How Large Does a System Have To Be to Behave as if It Were Continuous? And Some Interesting Questions about the Behavior of Smaller Finite-State Automata

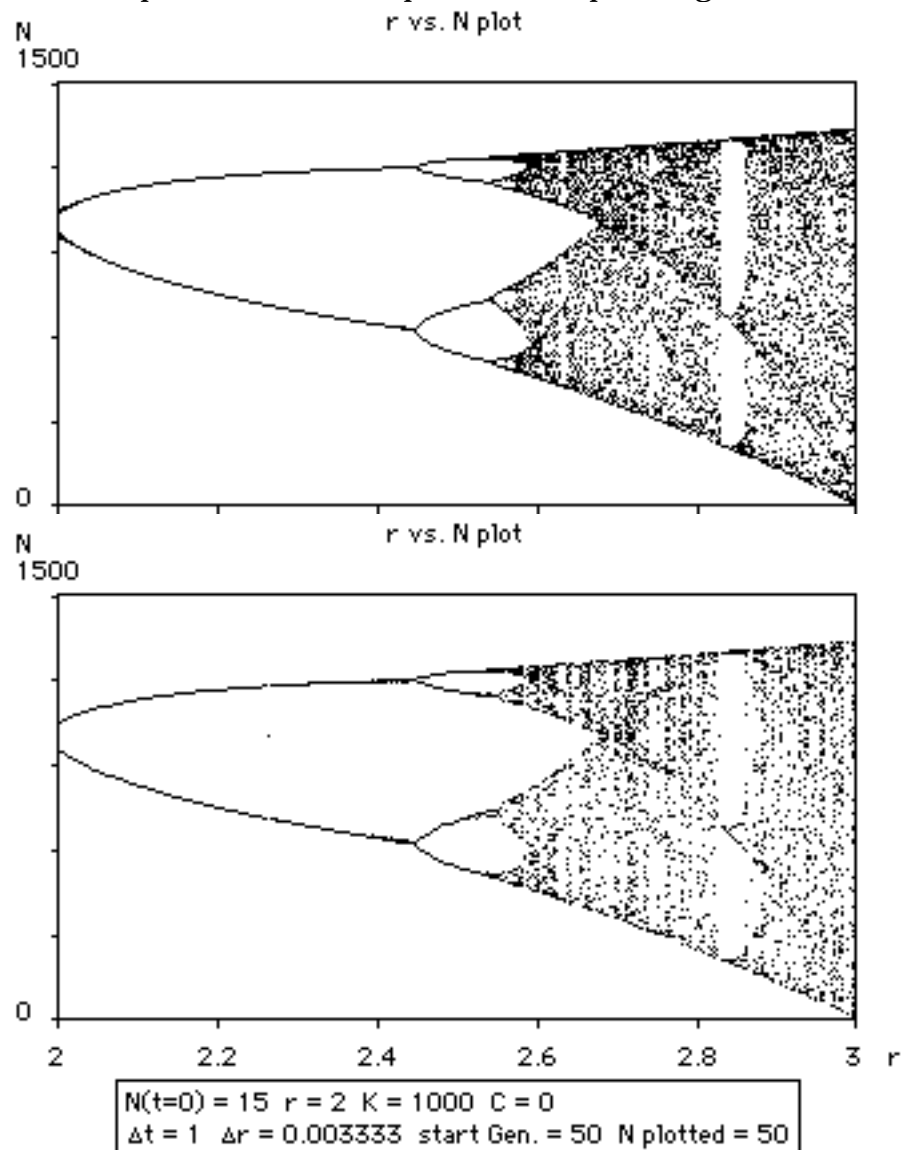
The Round N option can be used for a whole other set of exercises—to investigate the cyclic and chaotic behavior of finite-state automata (see, e.g., Kauffman, 1969, 1991, 1992). This is a topic that has attracted increasing attention among chaos theorists, and we can use Round N to investigate it. It is, curiously, a topic that appears not to have been studied for the logistic map by physicists—probably because they use a normalized version of the equation for dN/dt parameterized in terms of x rather than N , $[dx/dt = Kx(1-x)]$ in which K does not appear explicitly, and, in the context where they study it, it would never occur to them to suppose it would be interesting to ask what would happen if one limited N (their x) to integral values. In fact, since they let x range from 0 to 1 as we let N range from 0 to $4/3$, the idea of limiting this variable to integral values would probably seem incoherent. (This is a nice example of how different formulations of the same model can have quite different consequences for how it is extended!)

Viewed as an abstract model, the logistic or MDLS with the Round N option activated can be viewed as a deterministic discrete-state automaton with allowable states ranging over all of the integers in between 0 and what $(4/3)K$ rounds to, inclusive, i.e., including 0 and the largest integer less than or equal to $(4/3)K + .5$. What this means is that varying K controls the number of states the system is capable of assuming.³⁸ Thus, if $K = 10, 100, 1000, \dots$, then $N_{\max} = 13, 133, 1333, \dots$, and you can construct a system capable of the desired number of states, say Σ states, by setting $K = (3/4)\Sigma$.³⁹ Secondly, with r , we have a "control parameter" through which we can modulate the dynamical stability of the system—changing its behavior from what in

³⁸Actually, the system can start out in states larger than $4/3 K$, but for $r > 0$, will decrease until it is \leq that quantity, and not venture outside of that range thereafter.

³⁹This ignores the complications induced by roundoff error at the upper end of the range. Sometimes the maximum allowable of N will be greater than $(4/3)K$ and sometimes less, depending on the exact setting of K , but it will always be true that $(4/3)K - .5 \leq N_{\max} \leq (4/3)K + .5$.

the continuous case passes for stable equilibration up through various

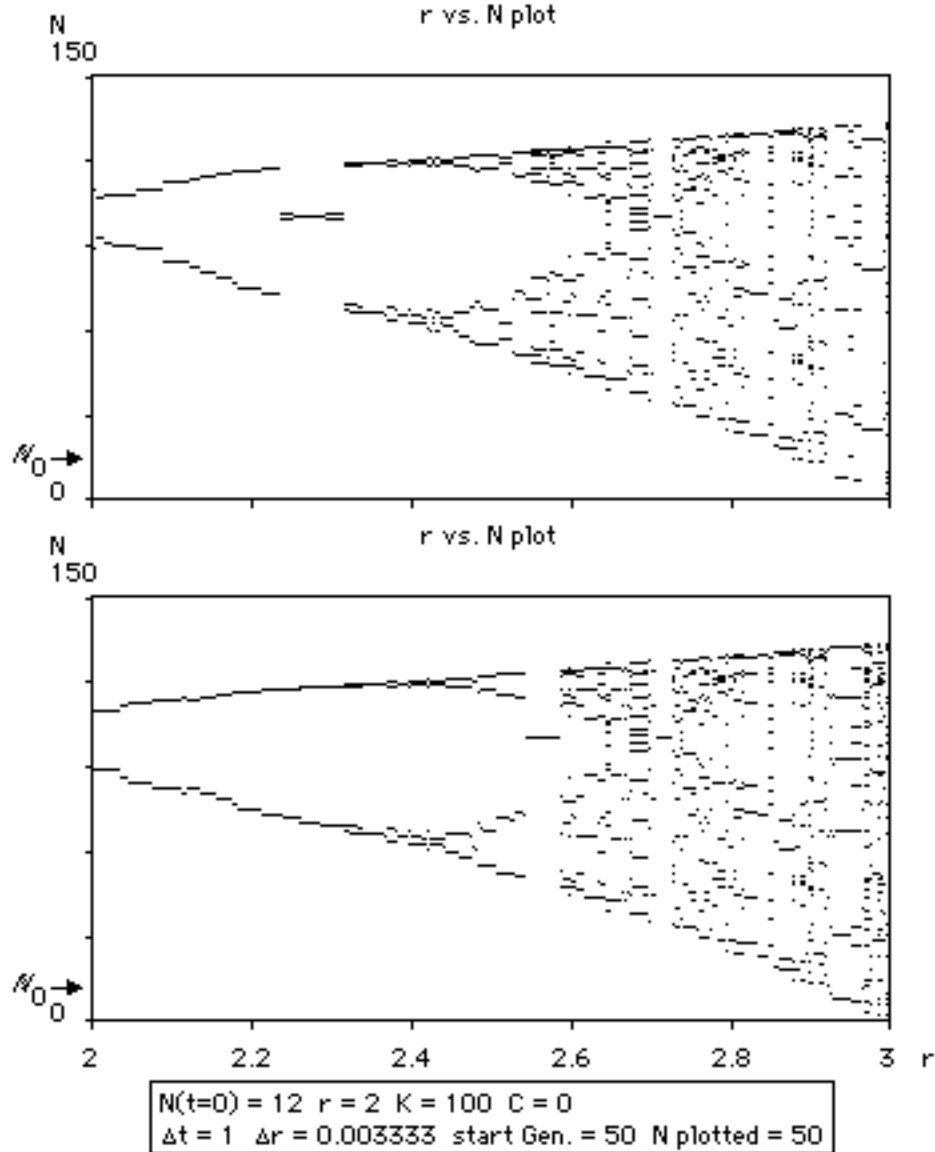


Figures 3-2 and 3-3. Bifurcation plots of systems with different numbers of allowable states.

periodicities into the chaotic regime. Thus we have the ability to study the effect of the number of allowable states on the dynamical behavior of the system.⁴⁰ In the first two figures of this series (above), we compare a standard bifurcation plot with one in which $K = 1000$, so that there are about 1333 allowable states. There are striking differences in these plots. At the vertical pixel resolution of this plot (170 pixels), there are nearly 8 allowable states per pixel—much finer resolution than is visible in the picture. Thus, you might think that it shouldn't matter whether there are thousands or billions of allowable states, but surprisingly, it clearly does. Round N in effect makes every

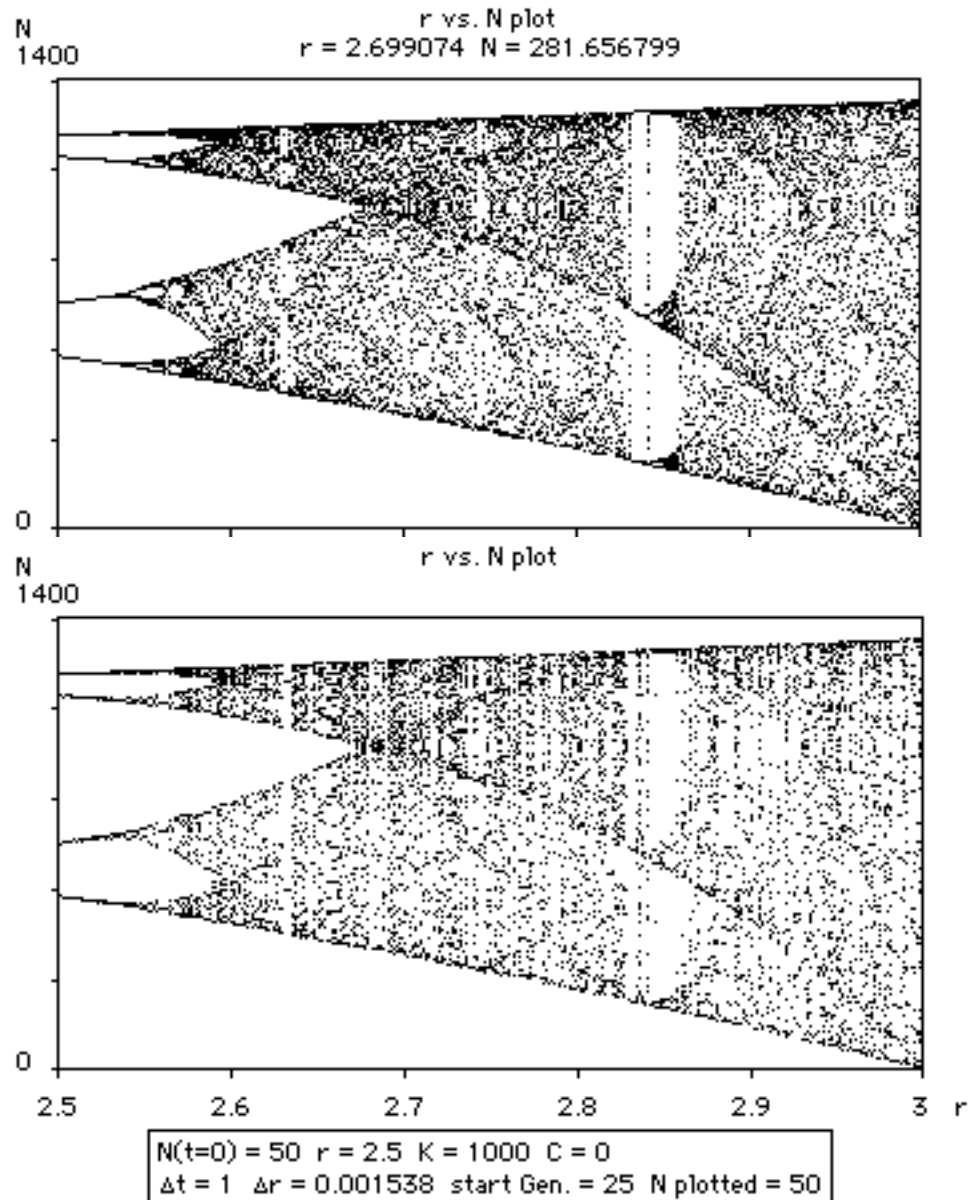
⁴⁰ Of course, there's something funny going on here, because in a digital computer, all states are discrete, and the simulations we describe as continuous (i.e., with "Round N " and Numerical Precision" turned off) are (in this software) actually discrete state systems with about 10^8 - 10^9 allowable states.

integer a quantized attractor, which has manifest effects, from the lack of equilibrating spread near $r = 2.0$, to the single point at $N = K$ for about $r = 2.26$, to the many "bands" and generally thinner appearance in the "bush" of the tree. We'll return to this later.



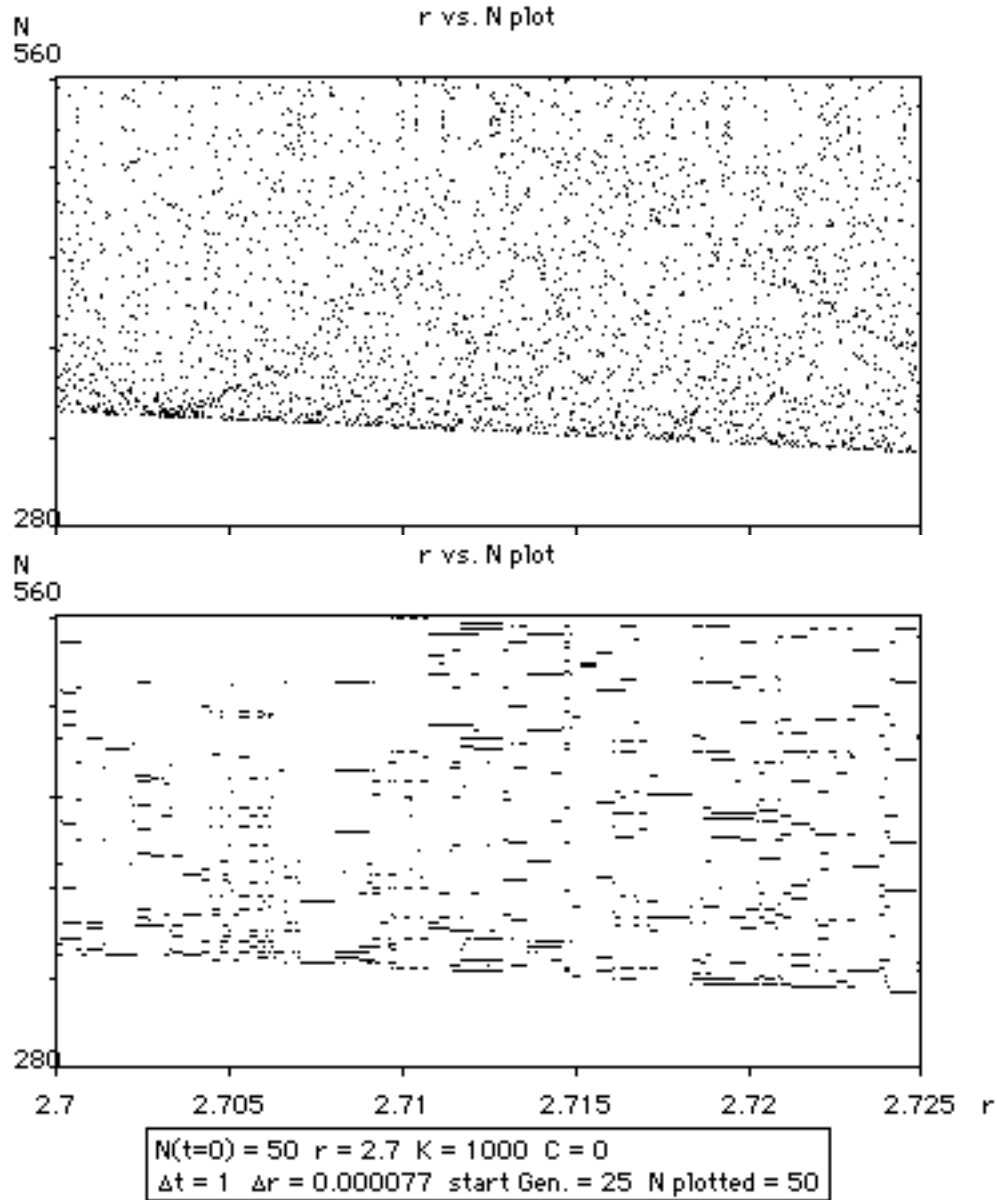
Figures 3-4 and 3-5. Dependence on starting value of systems with small numbers of states.

In the two pictures immediately above, K is decreased to 100. The thinning continues, but these graphs also show another interesting phenomenon—the obvious dependence of many details of the pattern on N_0 , 15 at the top and 12 at the bottom. This kind of dependence is actually there, but not obvious in the earlier figures. Note the appearance of many obvious periodic orbits which are stable across ranges of r appearing as bands along the x-axis. We'll return to this, after a side trip to look more at the reduced number of points in the "bush" of the tree for $K = 1000$. Figures 3-6 and 3-7 rehearse this phenomenon for slightly different settings (in part demonstrating that it is robust), arranged so as to blow up the region from $r = 2.5$ to 3.0. Again, we compare the appearance with round N turned off and turned on:



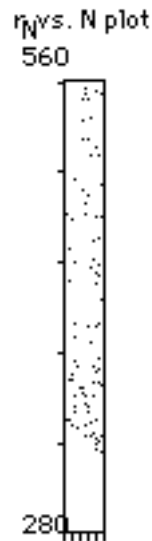
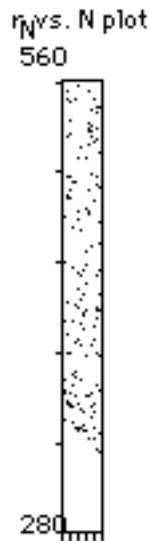
Figures 3-6 and 3-7. Like figures 3-2 and 3-3 but for somewhat different values.

As an aside, all of these paired comparisons are perhaps most effectively made by making them small enough so as to be totally visible in the .i.scrapbook,; pasting them in as neighboring pictures and then flipping back and forth between them. This serves very effectively to heighten and localize the differences. This makes maximally effective heuristic use of the template matching capabilities of our visual system. Suppose now that we wanted to "blow up" a region in the 2 preceding plots to count and compare the number of "dots" as a way of quantifying the effects that Round N had in reducing the number of trajectories. (This would be doing it via a kind of sampling, rather than analytically.) Consider the region, $280 \leq N \leq 560$, and $2.700 \leq r \leq 2.725$, a small square in the original plots. Should we just blow them up to the original size and then count, as in the 2 plots below?



Figures 3-8 and 3-9. Blowups of corresponding regions in Figures 3-6 and 3-7.

There is something clearly wrong in these "blowups," as is indicated by the fact that there are clearly many more points here than in the original regions. Can you explain why? (Compare the size of Δr in the two sets of plots, and recall how the r vs N plots are constructed, or see the discussion of r vs N or r vs N_{eff} plots in Appendix E.) The right way to preserve the point count would be as below—keeping the horizontal scale the same, and cutting out a strip of the right width and location. (This can be reset with the Graph Coordinates option, by making the graph $1/20$ th as wide—just 16 pixels to be exact. It messes up the legends, but that's not important in this context.) The vertical blowup (of 5x) doesn't affect the point count (if it covers the same range of N -values), and separates the points better to simplify counting. (Its only observable effect is the steeper lower boundary of the region.)



Figures 3-10 and 3-11. The "correct" blowups of Figures 3-6 and 3-7.

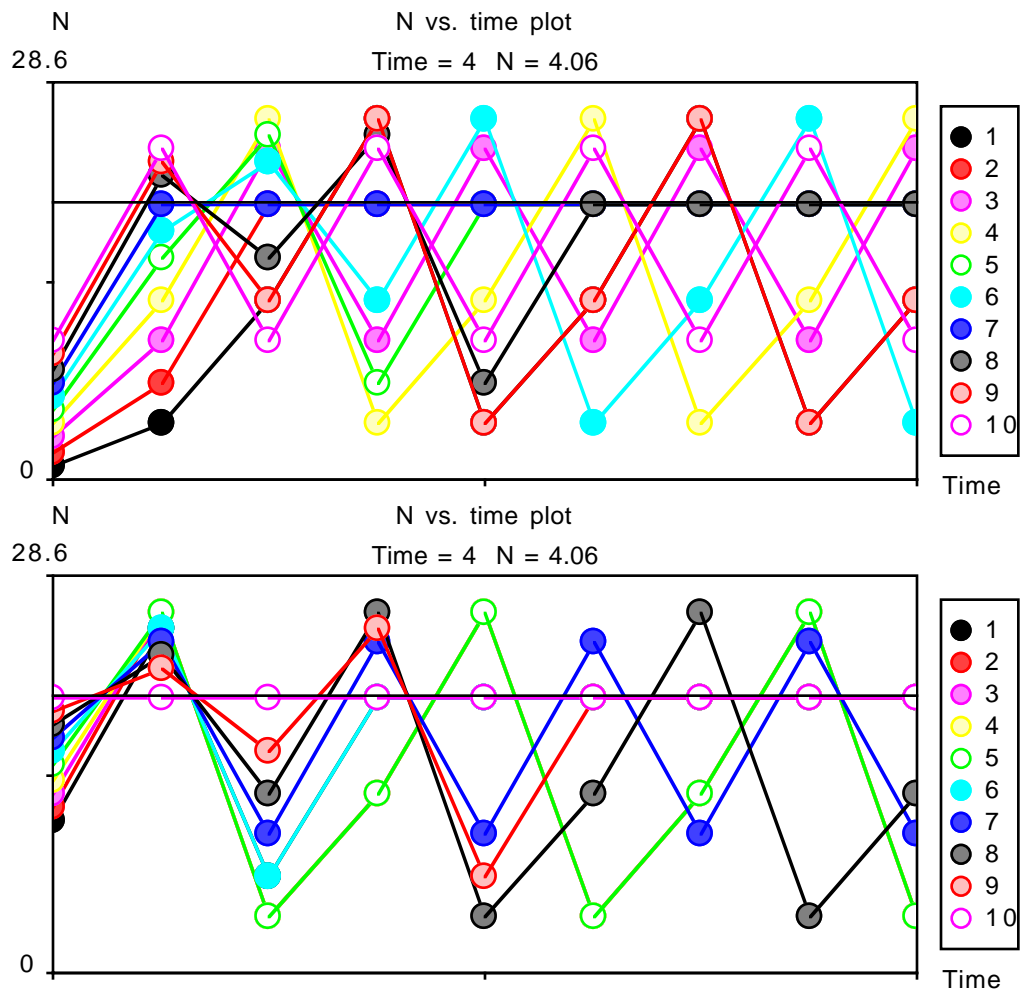
The counts in these two cases (Continuous at left, Round N at right) are 80 and 135 points, respectively. This was facilitated by printing out the two plots (separately, before they were inserted side-by-side (in a "table") in this text) blown up to 300% on a LaserWriter, which makes them each full page height, and makes the counting a breeze. (This can be done from within the program, using the Print setup... command in the **File** menu.)

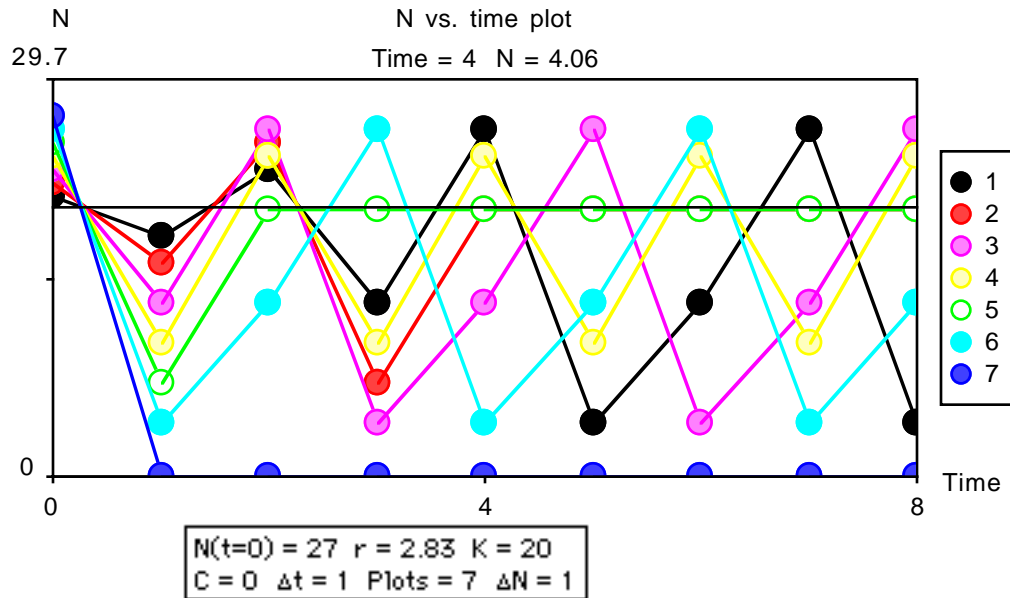
Before we throw away the "wrong solution" graphs preceding these two "count strips," notice another significant difference between the Continuous and Round N graphs with the 20-fold blowup in r -scale. Whereas the points in the Continuous graph almost all remain points in the blowup, the points in the Round N graph are almost all bands of varying lengths stretched out along the r -axis. That is, they represent stable "pivot points"—and usually stable stacks of bands or strata signifying stable orbits, sequences of relatively short period which remain invariant over a range of r -values. Why do you think that is? You might look some more at this, and also see how stable these bands are over different initial values (recalling the sensitivity to N_0 reflected in the two $K = 100$ graphs.)

This can be done qualitatively for these kinds of graphs by making sequences of shots in the scrapbook for successive integral values of N_0 and switching back and forth between them or playing them in sequence as movies. With systems which don't have so many states (i.e., for smaller values of K), the analysis can be done exhaustively and analytically using the N vs t or $N(t)$ vs $N(t+m)$ plot modes and the methods given below. Further methods for finding and investigating periodic behavior with these two representational modes are discussed later in the chapter.

In the graphs immediately below, K is set still smaller, to 20. The maximum N -value attained is 26, and for $N = 27$ and higher, all populations crash to 0, and stay there. Thus we have a finite state automaton capable of 28 states, and whose behavior is exhaustively captured in the next three graphs, as it is started at all integers from 1 to

27. (The behavior for state 0 is captured by trajectory 27.) In these graphs, the Lines option is used to aid in the visualization of the state transitions over time. Working from the simulation, rather than the static pictures, you have several other options available to help you disentangle the various networks of state transitions. Thus you can specify a longer time lag between plotted points with the Redraw option, so you can slow down the plots to see more clearly how they are drawn. (Try 3 as a starter.) Or you can plot fewer trajectories at a time—indeed, only 1 if you prefer. You can also delete or add starting or ending trajectories to a series, to better understand patterns or transitions in behavior from neighboring starting points. (You could also, of course do a complementary series of graphs varying r from the same starting point.)





Figures 3-12 through 3-14. All starting values for the $K=20$ finite state logistic.

In interpreting these patterns, note that:

(1) the behavior of the system is completely deterministic, so that if different trajectories ever enter the same state, they must behave identically thereafter.

(2) Because of the finite number of states allowed, all trajectories must sooner or later enter a cycle, with the maximum length of the trajectory (including transient states + cyclic states) or longest cycle equal to the number of allowable states in the system. The last graph (above) in fact shows all of the possible modes of behavior of this system for this set of parameters, though not all of the ways of getting there. It shows stable equilibria (or cycles of length 1) at $N = 0$ and at $N = 20$. (This shouldn't surprise you, from what you already know about the logistic.) It also shows a cycle of length 2, oscillating between $N = 10$ and $N = 24$. And finally, it shows a period 3 cycle among states 4, 13, and 26. If you look at the two preceding graphs, you will see the stable equilibrium at K (i.e., at 20), and the period 2 and 3 cycles in each.

Some other things to look for:

(3) A reduction in the number of dots visible from one time period to the next indicates a merging of trajectories—look for lines converging at that point.

(4) There are a variety of special patterns that are quite revealing. The compound graph of 10 trajectories on the page below (Figure 3-15) was made under the same conditions as Figure 3-12 in the preceding series, but for 25 rather than 8 generations, and without the lines. Both of these changes were made to ease the perception of repetitive patterns. After things settle down, we see two lines of alternating gray and white, and three lines each having a period 3 repeat of black, light gray, light gray, phase shifted 1 generation to the right for each step from bottom to top. These last are

obviously composites of multiple phase-shifted trajectories—trajectories which are at different points in the same cycle.

(5) Other clues for tracking trajectories come from intuitions about the logistic (using the r -value, how far N is below or above K , and the values of K , C , and Δt) to estimate the next location.

The repeat patterns and the deterministic assumption are clearly jointly useful in decoding the graph. Note that there are two dark gray (or blue and red, if viewed in color) spots at $N = 7$ in generations 0 and 1—and the first dark gray (blue) appears at $N = K$ in generation 1. So the other dark gray (red) must be captured in generation 2. And, similarly, the white (green) and medium gray (gray) at $N = 7$ in generations 4 and 5 must be captured there in generations 5 and 6. This convergence of 4 trajectories to 1 explains 3 out of 4 of the "lost" trajectories. The merging of a black (from 4) and a light gray (or red) (from 23) in generation 1 to the same trajectory (at 13 in generation 2)—also visible in Figure 3-12—explains the last "lost" trajectory.⁴¹

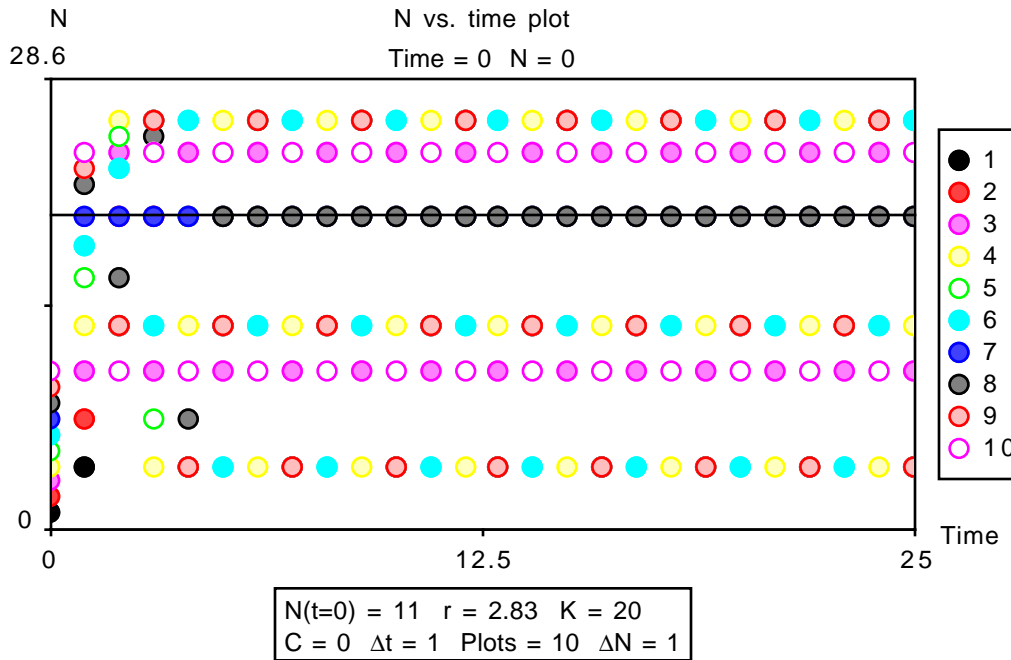


Figure 3-15. Like Figure 3-12, except without lines and for 25 generations to see repeat patterns.

⁴¹ It is important in decoding these graphs not to confuse the order in which the trajectories are drawn by the program (indicated in the legend at the right of the graph) with the order (or generation number) in which the trajectories reach their equilibrium state or cycle. Thus the dark gray (blue) spot starting at 7 in generation 0 goes directly to 20, and is visible there until covered by the medium gray (gray) spot in generation 5, even though it is followed only 1 generation behind by the dark gray (strawberry red) spot, which goes to 7 in generation 1 and 20 in generation 2. This spot is never visible after it gets to 20 however, because the blue spot (trajectory 7) is drawn after the red one (trajectory 2) and thus covers it in every case where they have the same value in the same generation, and thus occupy the same location in the graph.

On the next page, we have listed all of the state transitions you can read off these graphs for this system, for these parameter values of r , K , C , and Δt . This list is accompanied by flow charts or directed graphs organizing these state transitions into the 4 attractors and their domains, specifying under all conditions how this system goes to equilibrium, and what (and what kind of) equilibrium it goes to. Note that the attractor domains are "all mixed up": the first 4 states, 0, 1, 2, 3, end up in 4 different attractors, and this continues. The attractors for the states in order, from 0 through 27, are: 4, 1, 2, 3, 1, 2, 1, 2, 2, 1, 3, 2, 1, 1, 1, 1, 2, 3, 1, 2, 2, 1, 2, 1, 3, 2, 1, 4. There is no simple rule for the order of these attractors, but there is a growing theory of how different classes of random or chaotic finite-state systems should behave.

Flow chart between states:
(R) = already included

0->0
1->4->13->26->4
2->7->20->20
3->10->24->10
4(R)
5->16->25->7
6->18->23->13
7(R)
8->22->16
9->23
10(R)
11->25
12->26
13(R)
14->26
15->26
16(R)
17->24
18(R)
19->22
20(R)
21->18
22(R)
23(R)
24(R)
25(R)
26(R)
27->0

State Transition Diagrams for the Discrete-State Logistic Growth Equation:

$K = 20, r = 2.83$

Only integral values of population size allowed. Attractors indicated by heavy arrows.

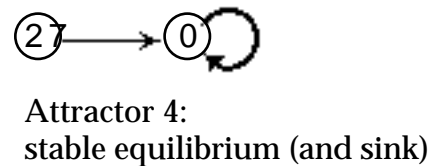
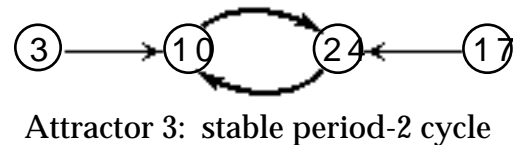
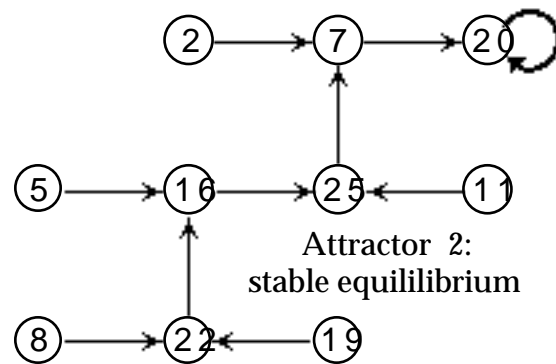
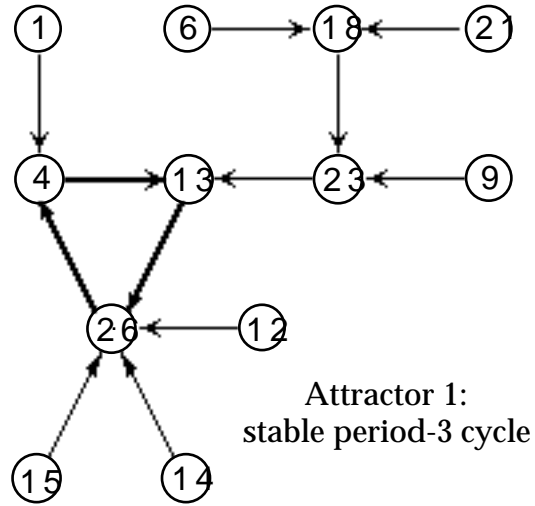


Figure 3-16. Sample state-transition diagrams, attractors, and their domains for a "logistic automaton." The notation "R" in the list indicates that the value in question already appears in the list, and so the transitions from that point on are already given. Which things are assigned R thus depend upon the order of entries in the list, and say nothing about the dynamics of the system. When cycles or equilibria first appear, they are followed until a state is repeated.

Several things are interesting about this class of automata, which we might call "logistic automata":

- (1) they are quite different in structure from other simple discrete state automata, most of which are composed of a number of elements which individually have states and which interact with one another in various ways to change one another's states. (Examples of such automata of widespread current interest include connectionist networks used to model neural and cognitive activity (see Bechtel, 1991), and Kauffman's (1969, 1985, 1991, 1992) models of gene control networks as networks of Boolean automata.) Logistic automata have no parts, at least not under this description.
- (2) Chaotic behavior has been defined for and claimed for all of these other discrete-state systems, which are "tunable" to produce or to avoid chaos—two features which logistic automata also possess.
- (3) Many of the relevant properties of such systems are related to their behavior near the transition to chaos—something which can also be studied for "logistic automata."
- (4) A number of claims have been made about whole classes of these other kinds of networks which appear to be quite general for things in that class, but possibly quite special for that class, though no one knows how special. In particular, some of these claims also have natural interpretations for logistic automata, but no one has investigated to see whether they are true for logistic automata. These automata thus provide an interesting class of systems for comparative investigation of some of these behaviors to see whether any, some, or all of them are more general than supposed, and what we can learn about chaos and about these classes of systems from this.

We won't go into them here, but Kauffman in particular (1969, 1985, 1991, 1992) has made claims about: (1) how mean number of cycles and cycle lengths depend upon the number of states, (2) how the identity and domain of attractors vary with small changes in the system or in its starting state, (3) how many other attractors on the average are reachable via small modifications in the system or in its starting state from given attractors, and (4) how the behavior of the system is to be characterized as it nears the transition to chaos. All of these and more can be investigated for logistic automata, and some of them will require some interesting conceptual work in addition to decide how to define what to change to test which conjectures, and how to measure and evaluate the effects of such changes. This too is a frequent need in state-of-the-art model-building. There are enough original research projects in these topics for a raft of classes, and quite probably some interesting and eminently publishable results. Have fun!

A final thought question: Given what you have seen here, can chaos be defined in the same way for discrete-state systems as for continuous systems? Why not? How would you do it? (See Kauffman, 1991, for his answer.)

A Second Monte Carlo Simulation—Estimating the Dependence of Extinction Rates upon r :

One of the earliest things noted for chaotic behavior in ecological systems (see, e.g., May, 1974), is that the wild swings in population sizes produced by chaotic dynamics might lead to extinction (either directly, through "population crashes" all the

way to zero, or indirectly, through major decreases in populations that were then knocked out by other effects. Our emphasis here is on modeling extinction of local populations rather than whole species,⁴² and supposes the simplest model of extinction easily studied with this model. We suppose that a population goes extinct as soon as (or very soon after) its chaotic fluctuations in size bring it to less than its critical density, C . Figure 3-17 depicts a simulation in which the r value is just high enough to cause populations occasionally to bounce down to a population size slightly below C .

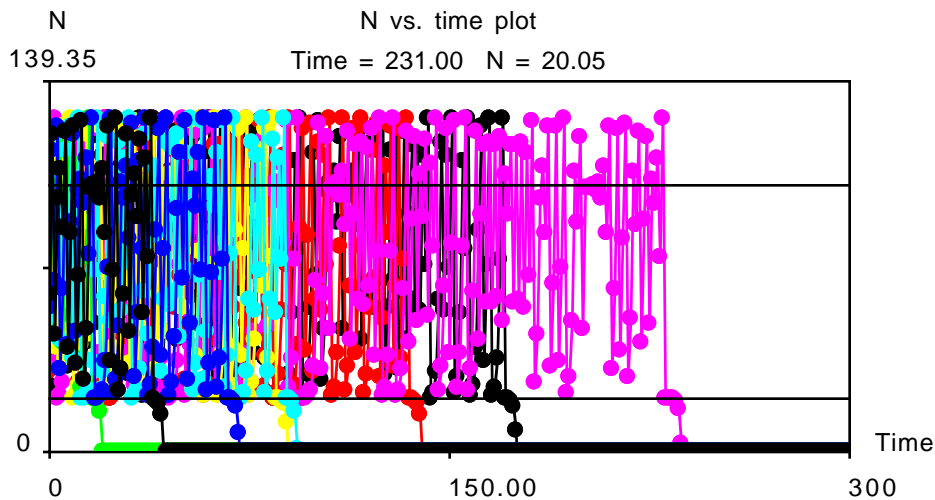


Figure 3-17. Simulation of extinction rates for the MDLS with $K = 100$ and $C = 20$.

Figure 3-17 is set for 8 trajectories with N 's of 25, 30, ..., 60, and for an r value just high enough (3.751) that occasionally populations bounce just below $C = 20$, after which they rapidly go extinct. If you are viewing this on a color monitor, or do the run using patterns, you can identify which population is going out where. The coordinates above the graph are the location at which the last population first goes below C —in generation 231, and at a N value of 20.05. (There is obviously a pixel roundoff error here, since 20.05 is $> C$, not less!⁴³

⁴² For more on statistical modelling of local extinctions, see MacArthur and Wilson, 1967. See Raup, 1991 for an up to date discussion of the causes of extinction at the species level, which discusses direct vs. indirect causes of extinction. Raup's fascinating discussion however focusses on large-scale extinction events, and the hypothesis that these are caused by collisions with asteroids or comets. This is an area which has seen significant mathematical modelling, though the models are quite different in focus (and in levels of analysis) from those discussed here.

⁴³ The choice of $C > 0$ is not arbitrary here. In principle, the same thing could be done in a program by testing for when $N \leq 0$, rather than for when $N \leq C$, but if $C = 0$, then the program default (which plots $N = 0$ from the time that it first is ≤ 0) causes successive trajectories to overwrite each other (like they do in this case at the bottom) and it is not easy to tell when a population first goes extinct. By setting $C > 0$, and making r only slightly greater than necessary to make $N < C$, each population goes out by first hitting just below C somewhere and then decaying. This gives a characteristic "tail" for each population (distinguishable with the patterns or in color) which can be evaluated with the point grabber. This in effect is a matter of "tuning" the simulation, so that the necessary information is readily observable—exactly the same kind of process necessary in the laboratory to get the most (and least ambiguously) from the instrumentation and the experimental design.

You can do several different things with this kind of simulation: You could do a large number of runs (no more than 10 at a time please!) from different starting points—e.g., 100 runs, from 25 to 74.5 by increments of .5, score each for when it goes extinct (using the point grabber) and plot a distribution of how many populations survive for different periods of time. These could be graphed. You could also test to see whether extinction is an exponential decay process by seeing whether the time it takes for half of the populations to go extinct (the “half-life”) is roughly constant for different numbers of surviving populations for a given value of r . (This would also be indicated by a linear graph of the logarithm of the number of populations remaining vs time.) You can also see how this distribution changes for slightly larger r values (3.752, 3.753, etc.), and estimate how the half-life depends upon r . Or you could change to a different C -value, find the point at which populations first begin to drop through it and go extinct, and do comparable studies for that value. (This should actually be a more sensitive method than that given above in exercise box 6 (comparison of bifurcation points) for determining the dependence of behavior on C or C/K .) Can you explain why? Why don't you test this conjecture?

You can also use an independent estimate to see how flat the distribution of points between N_{min} and N_{max} is in this region. You may want to try to estimate the slope of N_{min} at $r = 3.75$ by looking at its value at some earlier point (say, 3.70), relative to that at 3.75 using the r vs N plot and looking at a narrow band of N -values (say $19 \geq N \geq 24$). (To do this, you may have to plot a larger than normal number of points to get a reasonable density. Try setting initialization = 0, and plotted points = 200.) With this estimate, you can try to extrapolate the value of N_{min} out for the values of r you did for your extinction runs, in order to see how far it should go below C for the range of r -values you are investigating. Now try calculating $(C - N_{min}) / (N_{max} - N_{min})$, which is the proportion this overshoot is of the whole range. If the distribution of points is flat over this whole range (meeting one criterion of randomness), then this should give the expected proportion of trials producing an N -value in the interval $[N_{min}, C]$, where any value of $N > C$ constitutes a trial. (If you're curious, Roughgarden (1979, p. 339, Figure 18.9) gives an illustration of the expected distribution for the case $r = 3$, and $C = 0$.)

Exercise 10

Try this one in phase plot format (explained below after the next section), for different r values (3.751, 3.752, 3.753,...), and for different phase lags, m : ($t+5$, $t+10$, ...) What is the relation between the information given in corresponding t vs N and $N(t)$ vs $N(t+m)$ plots?

Phase Plots: Learning Why Bifurcations and Chaos Happens, and Detecting and Analyzing Periodicities

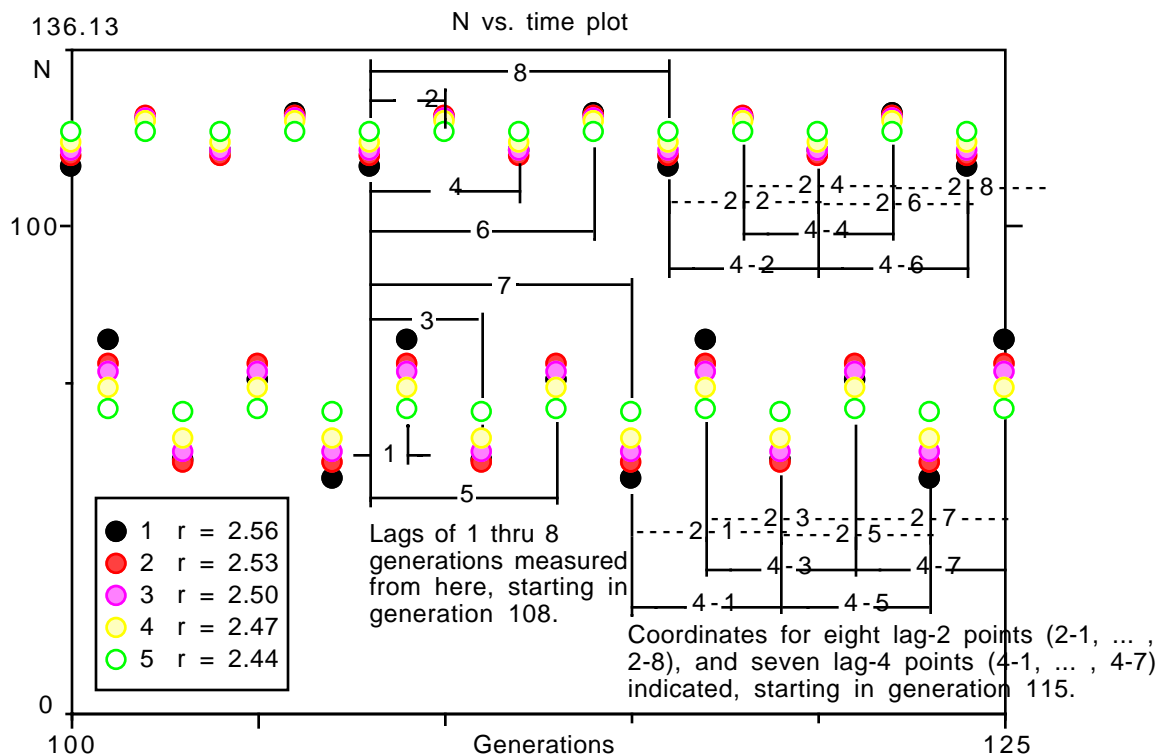
Now we are going to try a different way of presenting the data—a so-called “phase plot” which plots the value of a variable at one time (as the x -coordinate) vs its value some specified time later (as the y -coordinate), for all different times for which the

program has data.⁴⁴ (The phase plot, $N(t)$ vs $N(t+m)$, is done here with data generated using the same format as the N vs t plots—and if you have already generated an N vs t plot, you can use that data—or conversely.) The phase plot is an interesting and revealing way of plotting data, has a large number of different uses, and is becoming far more common for good reason. We will demonstrate a variety of uses here, but there are many more. Several of its uses or features are as follows:

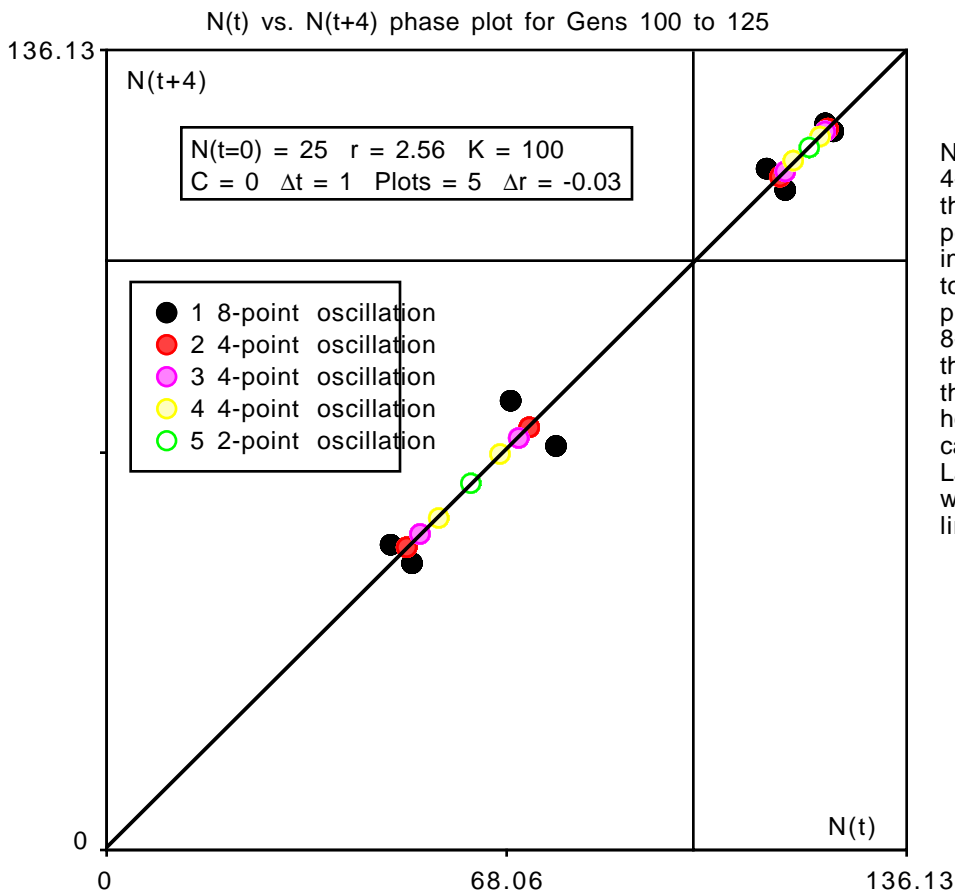
- The 45° or "stability line"⁴⁵ (where $x = y$) has a special significance: points on this line have the same value at t and at $t+m$. Thus, if $m = 1$, the point is not changing over time—it is at equilibrium. If $m > 1$, a point on the line indicates a periodicity, either of length m , or of an integral divisor of m . Any other points on this line (for the same trajectory) will also be part of that periodic orbit or cycle, and conversely. The number of such points should be m , or an integral divisor of it, if m happens to be some multiple of the period. *A cycle of period j should have all of its points falling on the 45° line when m is set equal to j , or to any integral multiple of j .* (Points on a cycle won't fall on the line unless m matches its period in this way.)

⁴⁴Thus, if the time lag, m , is 10 time units (or generations), and the simulation is a run for 100 generations, the first pair plotted is $N(1)$, vs. $N(11)$, and the last is $N(90)$ vs. $N(100)$. This also explains why, with a longer lag, fewer points will be plotted from the same data.

⁴⁵ Actually, to get the line $y = x$ (or $N(t+m) = N(t)$) to be at 45°, the horizontal and vertical dimensions of the graph must be adjusted using **graph coordinates** in the menu so that they are the same, as they were for all graphs here.



N vs. t (top) and Lag-4 Phase Plots for the same data, showing how pairs of N-values with different are taken as the x and y coordinates for the phase plot. Compare the appearance of 2, 4, and 8-point in the N vs. t and the Lag-4 phase plots.



Note that all points in the 2 and 4-point oscillations both map to the stability line--into 2 and 4 points on that line, respectively indicating that their variation is totally captured by the Lag-4 phase plot. The 8 points of the 8-point oscillation do not fall on this line, and the magnitudes of their deviation from it indicate how far they are from being captured by a 4-point cycle. In Lag-8 phase plot, these points would also fall on the stability line.

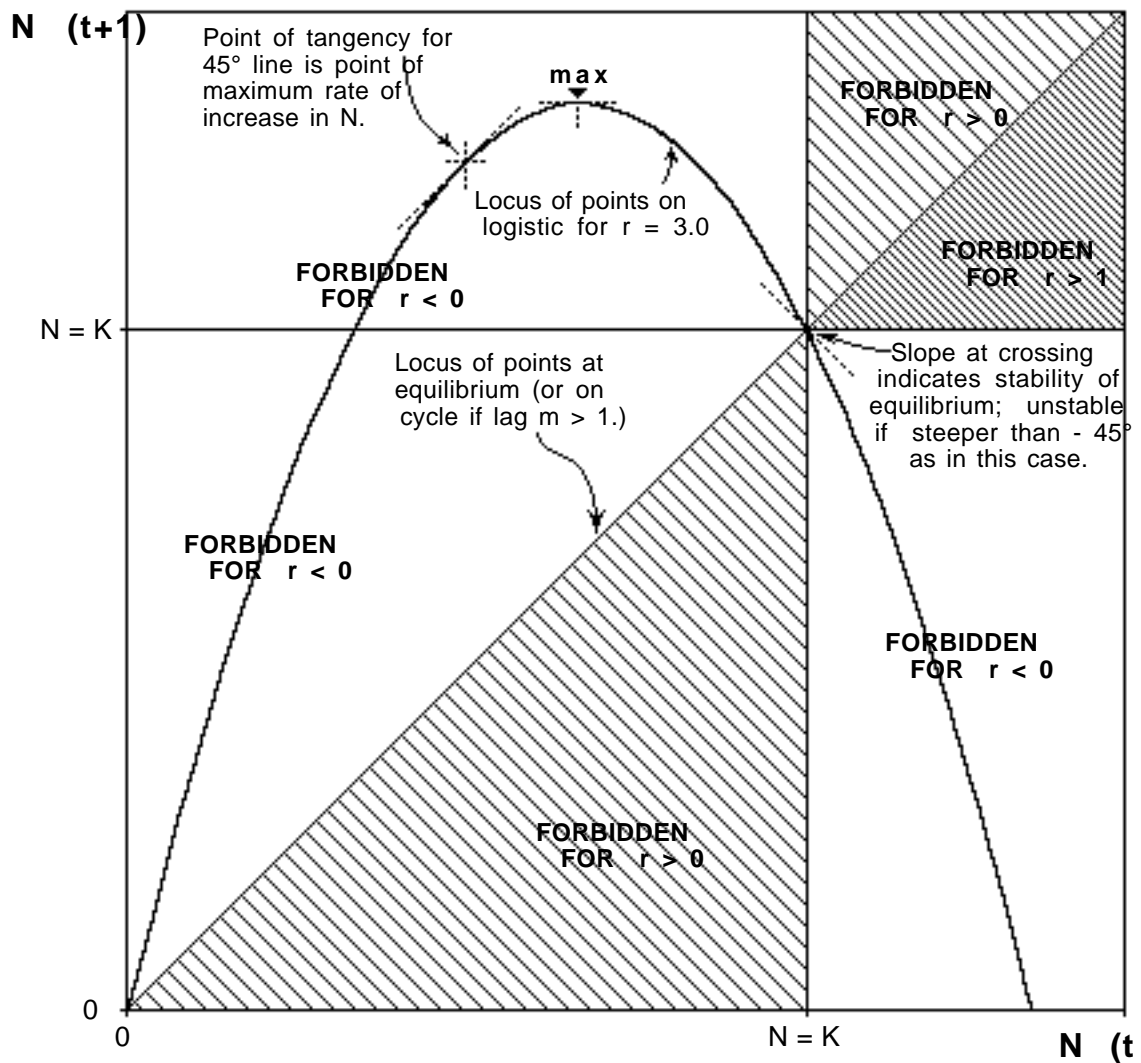
Thus phase plots may be used (by systematically varying m) to look for exact or approximate periodicities in the trajectories of the points.⁴⁶ Since phase plots are probably unfamiliar to you, let's look at some data which are graphed both in a traditional N vs t format (Figure 3-18a), and again, as a phase plot with a lag of 4 ($N(t)$ vs $N(t+4)$) (Figure 3-18b). The values used were picked to give a range of trajectories with different periods—all powers of 2—with cycle times ranging from 2 to 8. To simplify matters, we plotted the population sizes in generations 100 through 125, starting late enough so that all of the trajectories had time to settle into stable periods, so all you see is periodic behavior throughout the whole of both graphs.

Notice that with the phase plot, you lose information on which generation you are in, but gain information on how closely the various trajectories fit a regular periodicity specified by the lag, m . In this case, with a lag of 4, we see how well the various trajectories fit a period 4 cycle. As you can see, all of the period 2 and period 4 curves do so perfectly. The period 2 plot (represented by two white circles on the line), does so degenerately—it runs through its period 2 recurrence twice every period 4 cycle. Two points from each of the three period 4 plots spread outward along the stability line (as r increases) from each of the period 2 points. Each of these makes its period 4 cycle exactly once. The period 8 cycle also spreads outward, in this case orthogonal to the line, indicating that its 8 points no longer fit the 4-point periodicity, but the *deviation* of these 8 points from the stability line shows how large the deviation is from a 4-point cycle at various points of the cycle. There are two period 8 points associated with each point of the 4 point cycle, because the period 8 trajectory deviates from the period 4 cycle differently in its first half than it does the second time around in its second half. (Each "line" of the 4-point branch in an N vs r plot has bifurcated into two branches in the 8-point cycle, and one of each of these pairs of branches is visited every 4 generations—alternate ones of the pair in alternate period 4 passes, just as the period 4 cycle arose by binary bifurcation of the period 2 cycle earlier.) If you redo this run with a lag of 8, you will see the 8 black dots all located on the stability line, spreading outward in pairs on both sides of each of the last period 4 spots. We'll return to this below (see Figures 3-23 through 3-27).

Phase plots are extremely popular because they are very instructive about the dynamics of a system. Because we are dealing with a system whose next value is totally determined by its current value (according to the settings of the various parameters), it is particularly revealing to look at the phase plot for a lag of 1.⁴⁷ Figure 3-19 illustrates, for further discussion, a number of relations expected for phase plots, for the logistic equation, for the special case of a lag of one generation:

⁴⁶Time series analysis in the time domain basically involves looking at a sequence of points using phase plots of all different time lags, and finding that value of m such that the deviation of the points from the line $x = y$ is a minimum according to some measure (least mean squares, etc.). In some circumstances (e.g. periodic exponential growth or decay), the x - or y - coordinates may be transformed before fitting the line, or one may look for fits to a line of different slope. If there are several values of m having the same deviation, the smallest is taken as the period, and the larger should be integral multiples of it.

⁴⁷Setting Δt to less than 1 (or more than 1) does not give you a lag of less than (or more than) 1. In effect, a lag of 1 means a lag of 1 unit of Δt . Except in special cases, however, it is normally supposed that $\Delta t = 1$ generation.



Phase plot relations for the logistic growth equation with a time lag of one generation

Figure 3-19. Qualitative relationships and interesting loci of points for the lag 1 phase plot.

- Since dN/dt (or $\Delta N/\Delta t$, for finite differences) is positive for the logistic if $r > 0$, and $N < K$, then $N(t+1)$ will be greater than $N(t)$, and points cannot under these conditions occur in certain regions of this graph. Think first about what it means for a point to be above or below the diagonal line, and in each of the four quadrants of the graph. Why are points forbidden above the line in the upper right-hand corner? Why are they

forbidden below the line in that corner (in the densely shaded region), if $r > 1$? Can you explain why they are and are not permitted (and under what conditions) in each of the regions of the graph?

- The slope of the curve plotted at the point(s) where it crosses the 45° line (where $x = y$) determines the stability of points falling on the curve for points of the "cycle" (if there is one) m units apart. These plots are also central to explaining why bifurcations and chaotic behavior occur. In particular, if the slope is negative and steeper than -45° (indicating a slope whose magnitude is greater than -1), this indicates that small deviations from the stability line will be followed in the next generation by larger deviations on the other side of it, indicating an unstable or growing oscillation. (Convince yourself that if a point falls right on the -45° line, given the meanings of the coordinates, it will go an equal distance on the opposite side of the equilibrium value in the next generation.) For the same reason, negative slopes shallower than -45° (i.e., with values between -1 and 0) indicate decaying oscillations. (To test your understanding, consider what positive slopes of greater or less than 45° mean.) You can check these conclusions by generating a curve that has the desired properties (of increasing or decaying oscillations and the like) in the N vs t plot mode, and then selecting the phase plot option, to see how it looks here. A book on chaos (such as Stewart, 1989) explains this material in greater detail, and Boyce and DiPrima (1991, pp. 104-107) give a clear exposition of a graphical means using the stability line and the parabola (a so-called "stairstep diagram") to show the stability or instability of a trajectory.

- For values of $r > 1$ (and if $\Delta t = 1$ and $C = 0$), the points for a given curve will fall on a parabola which has a maximum value in the upper left-hand quadrant, which gives the value of $N(t)$ for which the value of $N(t+1)$ will be at its highest value—indicated in Figure 3-19 by **max**. You can determine these values (in an active simulation plot) by clicking that point (or any other points of interest) with the mouse. (Parallel interpretations work here for phase plots with longer time lags, in which case there are also multiple maxima and interesting intermediate minima.) What is the significance in Figure 3-19 of the curve intersecting the x -axis? What would you expect to happen in the next generation for N -values greater than that point of intersection?

- For points above the stability line, there is an increase in population size from generations t to $t+1$. If a line is projected vertically from any point down to the x -axis, the distance that falls below the stability line is subtracted from the total height, and these new y -values are plotted, we are now plotting $N(t+1) - N(t)$ on the y -axis vs $N(t)$ on the x -axis, which is ΔN as a function of N . The stability line now becomes mapped to the x -axis, and the curve falls above or below the x -axis depending upon whether for that value of N , population size will increase or decrease in the next generation. This latter kind of plot (or the related dN/dt as a function of N) is commonly used in population genetics for analyzing the stability of an equilibrium.

- Consider lines parallel to and above the stability line. What would it mean if a point fell on a line parallel to and Y units above it (distance above being measured in a vertical direction, parallel to the y -axis)? Consider now the line parallel to the stability line, and tangent to the parabola giving the locus of points for that r -value of the logistic. What is the significance of its x -value? Its y -value? How big is its ΔN compared to the values of ΔN for other values of N ? If Y stood for yield, how would you find the

maximum yield? If you were the owner of a fishery, having a pond with a carrying capacity of 1000, and an r of 2.7 per year, and you wanted to remove as many fish as possible each year, indefinitely, how many could you take out? How many should you leave in? How many would you lose (we suppose to starvation, though this is not thoroughly realistic here) if you, through

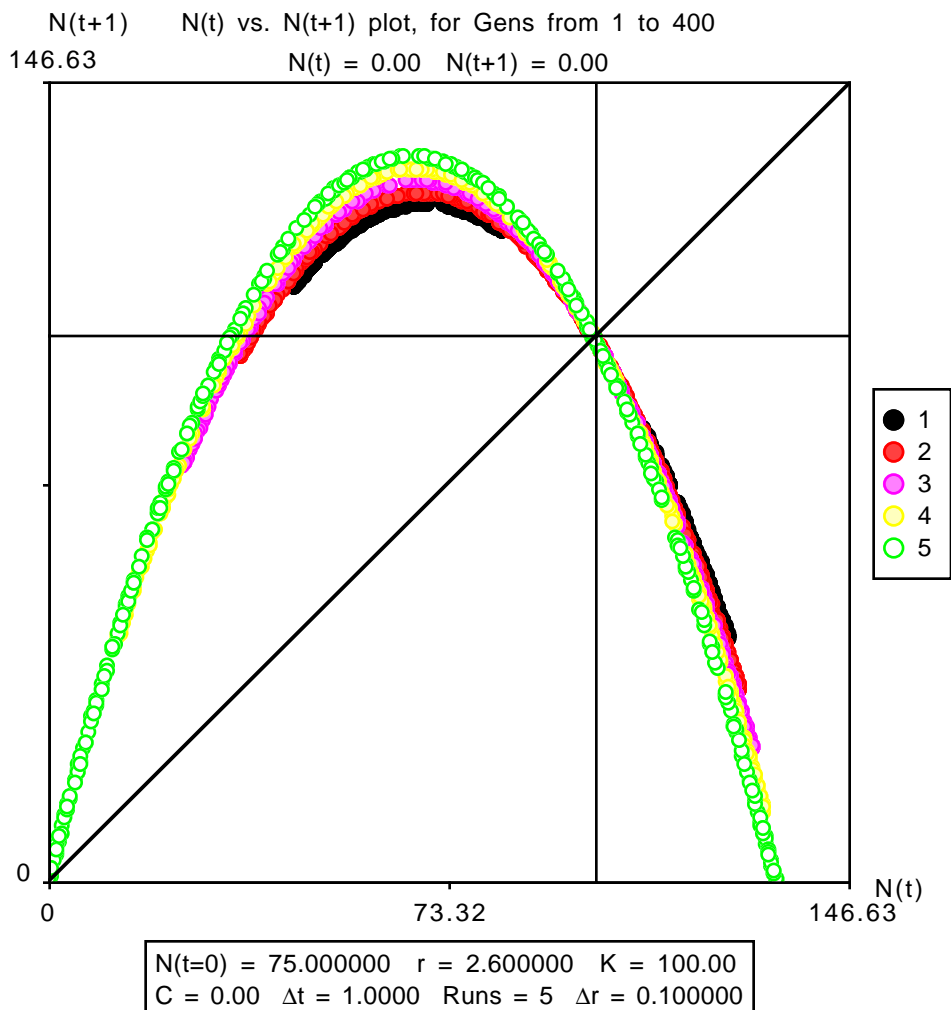


Figure 3-20. Lag-1 phase plot for 5 r -values; all other variables the same.

oversight, didn't take your "catch" out 1 year? How many would you lose in 3 years (defined as the difference between what you would get with optimal removal, vs with the optimal policy for long range recovery from your mistake.) Assume you don't have to eat in the meantime. How would this change if you have to eat, say, 100 fish per year (removable once a year, at harvest)? How many years will it take you to recover? How would having to eat the fish affect your loss due to the mistake? How do your answers differ if r is 2.6? 2.8? 2.9? (All of these problems can be solved with this program, running N vs t or lag-1 phase plots, and recording points from the graphs using the "record coordinates" option.)

The next two plots illustrate the effects of changing two kinds of crucial variables on the appearance of phase plots. In the first (Figure 3-20, above), we are dealing with a lag-1 plot which shows the effects of increasing values of r on the shape of curves in the chaotic regime. (It shows 5 trajectories for r values of 2.6, 2.7, ..., 3.0, printed with dots ranging from black (2.6) to white (3.0).) What does it mean that the parabolas bend increasingly sharply as r increases? What does it mean that the parabola for $r = 2.6$ does not reach to the "ground," while that for 3.0 does? Using the program, determine how the maximum value of N attained changes as a function of increasing r . Determine what value of N in the preceding generation produces this value. Where would you look to see which value was produced in the next generation by the maximum N values for each curve?

In the next figure, Figure 3-21, more trajectories are plotted (10 this time), and the lag is set at 2, rather than 1. This time, $C = 20$, rather than 0. Finally, the trajectories are run in reverse order, from larger r -values (at $r = 3.75$) down to smaller ones ($r = 2.68$) by increments of $\Delta r = -0.12$. (A negative increment for Δr was chosen so that the lower values would be the last ones plotted, and would thus appear "on top," so that they are not hidden as in the preceding graph.) Why might we have chosen to use higher r values in this plot? (If you're not sure, try this plot with the r -values from Figure 3-20, see what it looks like, and try to explain why. Or go back to the discussion of the MDLS in the preceding chapter.) What are the similarities and differences between this graph and the preceding (Figure 3-20)? What are the extra lines doing in this plot? What do the local minima in the middle of the x -axis imply for the different r -value curves? Why are these more spread out than were the maxima in the middle of the preceding graph? Why are the curves steeper near the left and the right side of the graph? Why are there two maxima for some of the curves? What does it mean that some of the trajectories are represented by spots rather than by curves? What does it mean to be represented by a curve rather than a spot or series of spots? What effect or effects has changing the value of C from 0 to 20 had? What changes would occur for a phase plot of this data set if the lag were set to 3? Or 4? Or 5? Or 20? How many maxima would there be? What would the curves look like? Try to guess, and then try it out with the program.

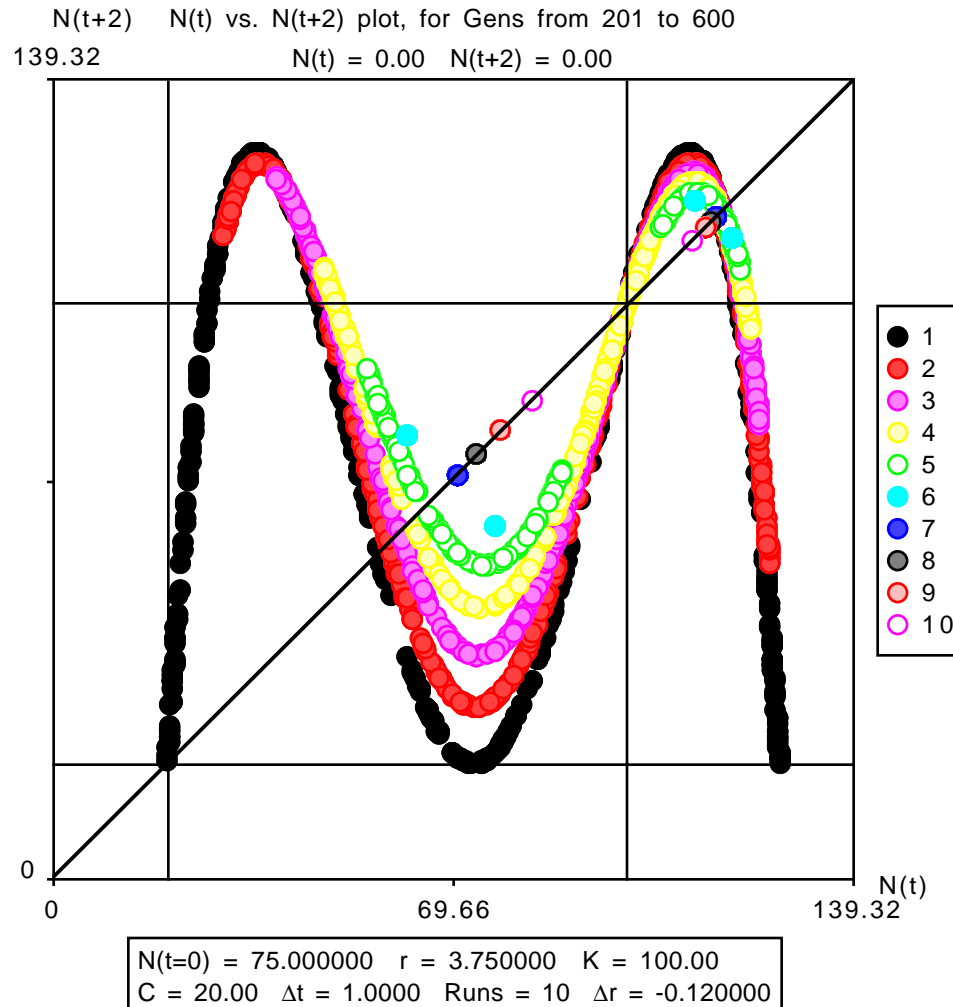


Figure 3-21. Lag-2 phase plot for 10 trajectories from $r = 3.75$ down to 2.67 with $C = 20$.

Exercise 11

Try phase plots for this and other trajectories for different values of m . What happens to the points in a curve as m increases? Can you see how a filled rectangle is produced for trajectories in the chaotic regime for large values of m ? What does the scattering of points in the rectangle mean for the ease of predicting its behavior with that time lag? What does it mean for a point to be in each of the four quadrants for a given value of m ? Can you explain what it means for some quadrants to be empty? Try a phase plot of $N(t)$ vs $N(t+1)$ for 10 trajectories in the chaotic regime starting from the same N -value with gradually increasing r -values, like those of Figure 3-20. Look at the same points for longer time lags. What happens?

The Analysis of Periodicities Using Grids and Template Matching in N vs t Plots and, Alternatively, Using Phase Plots

It is possible to use the graphic capabilities of the program to analyze periodicities generated using different values of r . The grid coordinates at the top of Figure 3-22 (measured from generation 0 at the left end of the graph) locate a set of points (right after the osculating pair four balls from the left, and just below the $N = 75$ crossbar) which fail to match a corresponding set 64 generations earlier, thus demonstrating a period of at least 128. (They are more spread out in the vertical dimension, and in color, show a black annulus that the earlier set lacks.) To see this for yourself, try three different starting N values at $C = 0$ and $r = 2.56995$. But this isn't the only, or even the easiest, way of seeing that it is periodic.

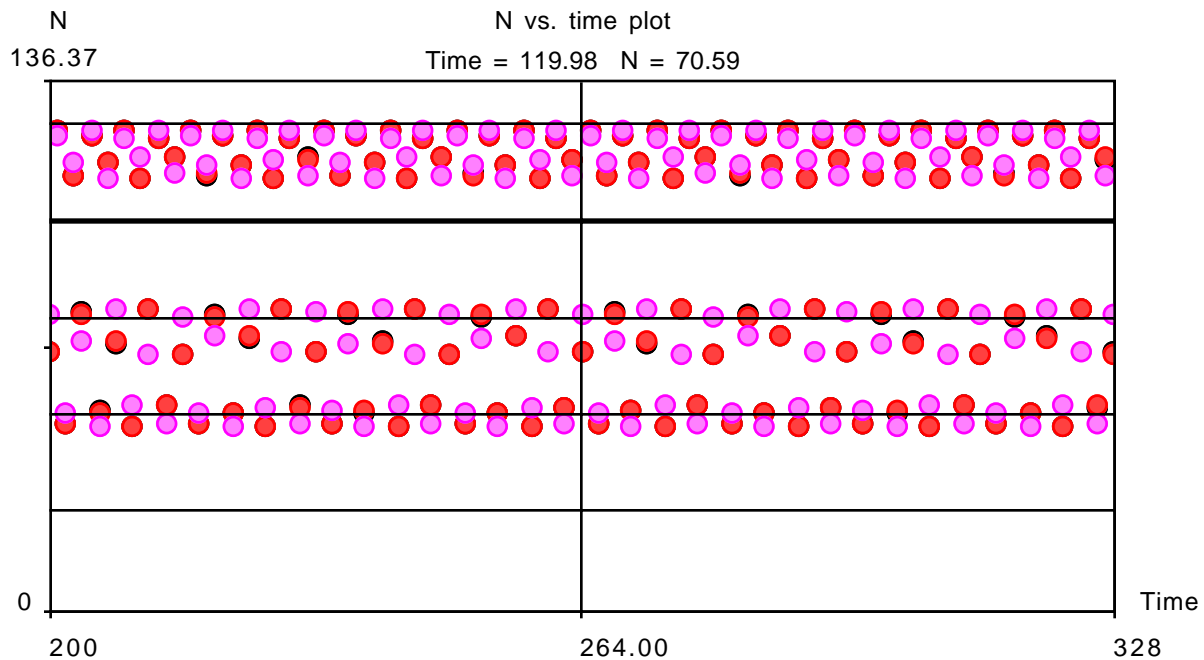


Figure 3-22. Three trajectories: $N_0 = 30, 40, 50$; $r = 2.56995$; start Gen = 200, stop Gen = 328; grid spacing: vertical = 25, horizontal = 64.

Another way to detect periodicities is to use the $N(t)$ vs $N(t+m)$ menu option in Logistic Growth. The following is a series of $N(t)$ vs $N(t+m)$ plots for $m = 1, 2, 4, 8$, and 64 for the same data which produced the time plot in Figure 3-22.⁴⁸ Notice the successively better curve fits for higher order periods in this series, indicating the hierarchical organization of periodicities in the 2^n series, with higher frequencies of lower amplitudes (in the ratio lambda, $\lambda \approx 2.504$, in successive periods—recognizable as the ratio of spreads of successive bifurcations⁴⁹ in the N vs r plots.).

⁴⁸You don't need to recalculate them. Just click on the $N(t)$ vs. $N(t+m)$ option, and the data will be regraphed in the new mode. It makes better sense for such graphs to make them square so that N at 2 different times is graphed at the same scale. You can do this with the graph coordinates option.

⁴⁹ λ and δ are two universal scaling numbers for binary bifurcations discovered by Feigenbaum. See Hofstadter (1981) or Stewart (1989) for further discussion.

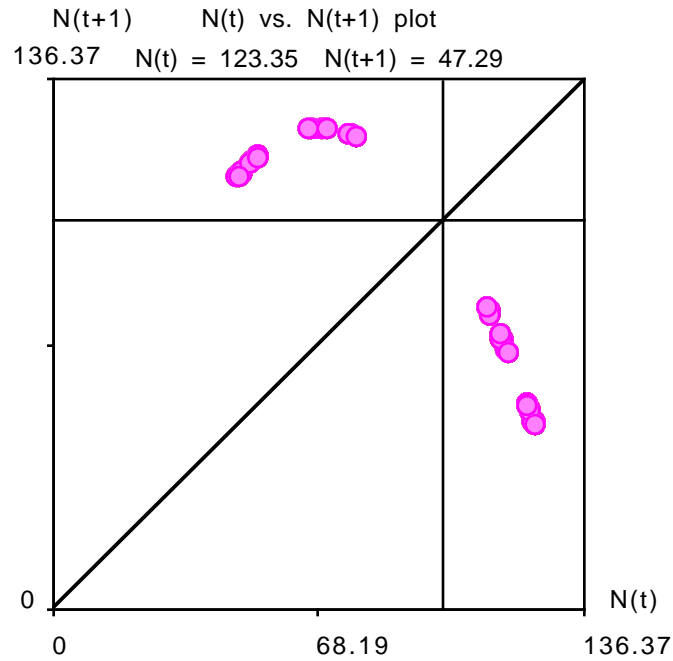


Figure 3-23. In successive numbers in this series, clusters wholly to one side of the $x = y$ line straddle it in the next higher phase portrait.

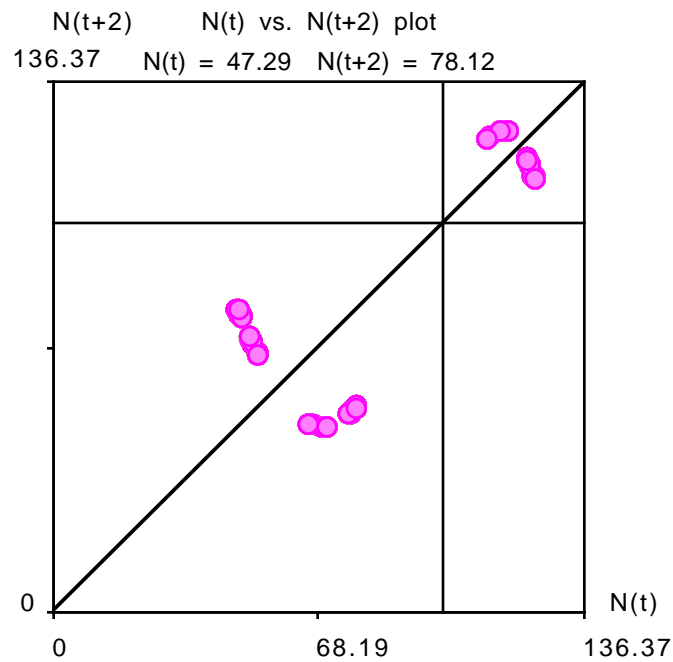


Figure 3-24. $r = 2.56995$, 128 generations, 3 trajectories for different N values at start, 200 \times 200 graph.

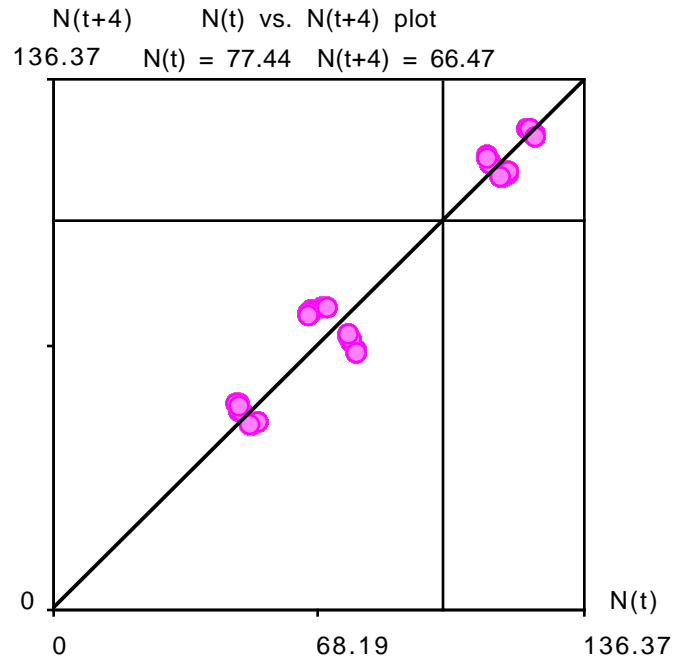


Figure 3-25. The same data in a lag-4 phase plot.

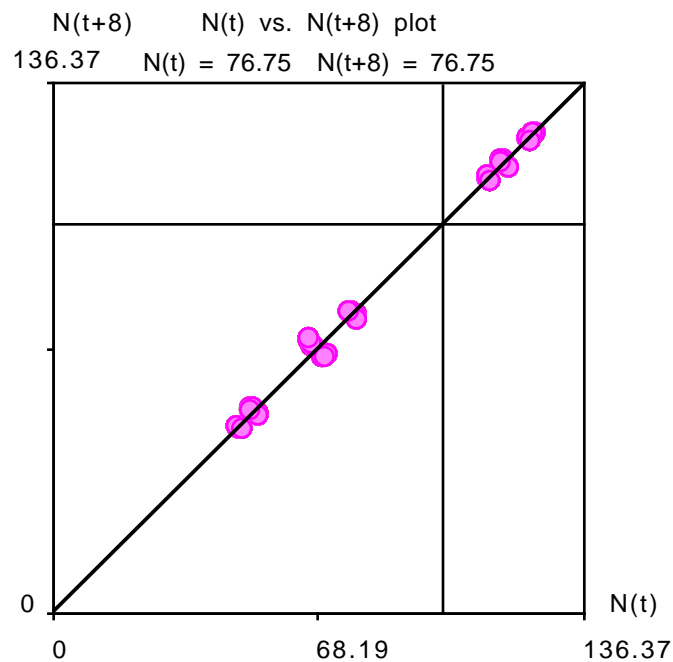


Figure 3-26. The same data, now in a lag-8 phase plot. Note how in each of these moves to a lag twice as great, the clusters normal to the stability line rotate parallel to it, producing a closer fit. This is characteristic of the hierarchical period 2^n bifurcations.

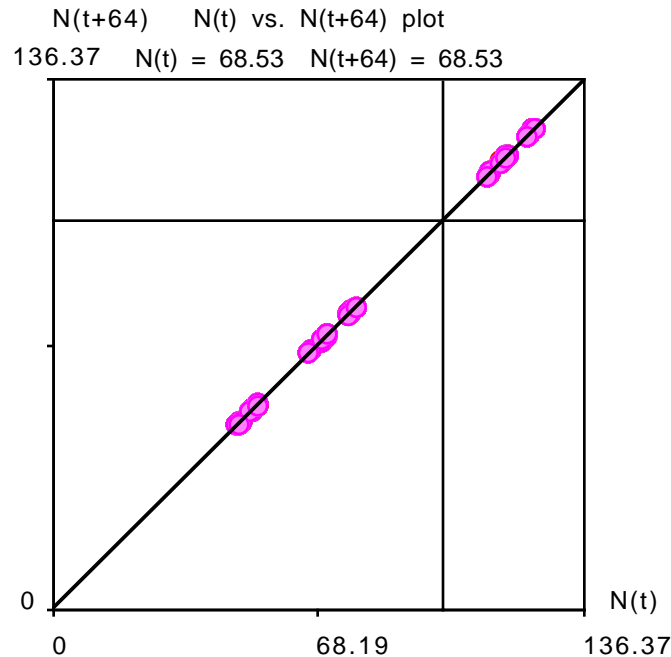


Figure 3-27. With lag 64, there is now an almost perfect fit with the stability line, though doubling the size of the graph, one can still see slight improvement between periods of 64 and 128 (if you plot the graphs long enough (i.e., for at least 256 points) to detect a mismatch).

Try periods 16 and 32 if you like. With the program, try periods one more and one less than the above periods (e.g., 65, 63; 33, 31; 17, 15; 9, 7; 5, 3) on these data. How about two more and two less? Can you explain the transformations of the patterns that you get when you do so? (Do you know why wheels appear to rotate backwards slowly when they are in fact going forward quite quickly in a motion picture or repeating strobe flash? Can you relate that to this behavior?)

Compare the behavior for r values in the chaotic domain (i.e., $2.57 > r \geq 3.00$)? The graph depicted in Figure 3-28 is intended to illustrate that the convergence we have seen in Figures 3-23 through 3-27 is not just an artifact of plotting $N(t)$ vs $N(t+m)$ for large m , i.e., for long lag times. This picture is taken in the chaotic regime, and indicates the kind of random scatter most such pictures show—unless they have a strong periodicity for the m in question. The pictures illustrated in Figures 3-28 and 3-29 have other uses too—they aren't always just random, but often show interesting patterns. Here's one (Figure 3-29) from just about where the two main spreading branches of the bifurcation diagram intersect—look at r -values around 2.67 or 2.68. What do you make of it? You might experiment a little (with different m 's and different r 's) to decide what to make of it. Check the same runs out in an N vs t plot. If you think you've got the explanation for why the points from one curve are found in all four quadrants, but the points of the other two curves are found (with one exception) in only two (they are all for the same r -value), try to generate curves that show this difference in an N vs t plot, and then convert it to this form to see if you are right.

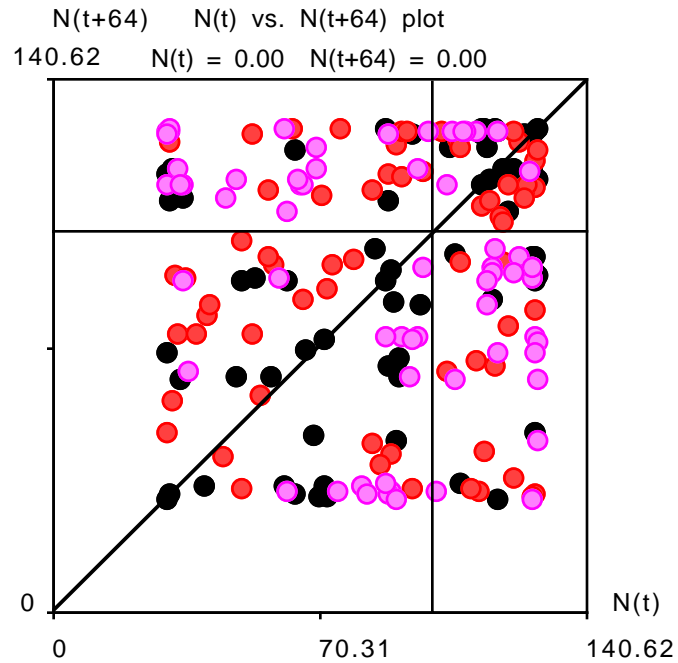


Figure 3-28. Phase plot for 3 different values of N_0 , where $r = 2.75$.

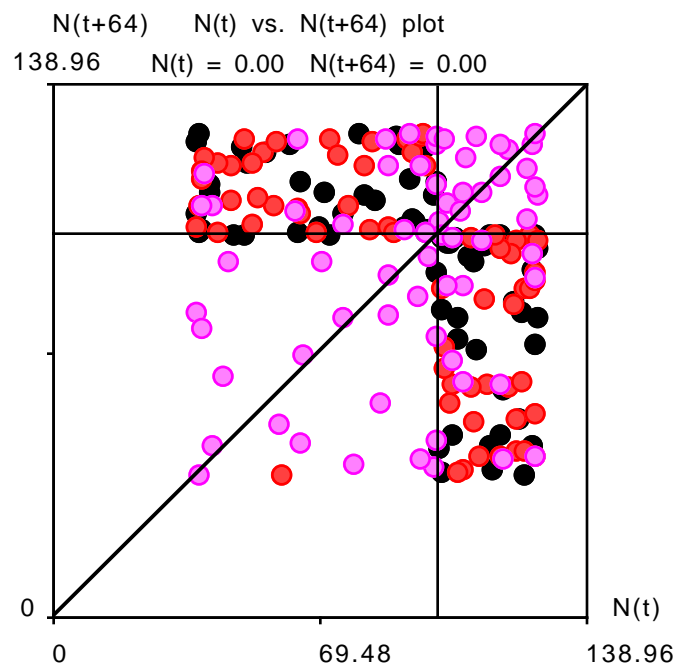


Figure 3-29. Can you explain why the darker points (black and red, in color) are found only in quadrants 2 and 4, whereas the lightest points (pink, in color) are found in all four quadrants?

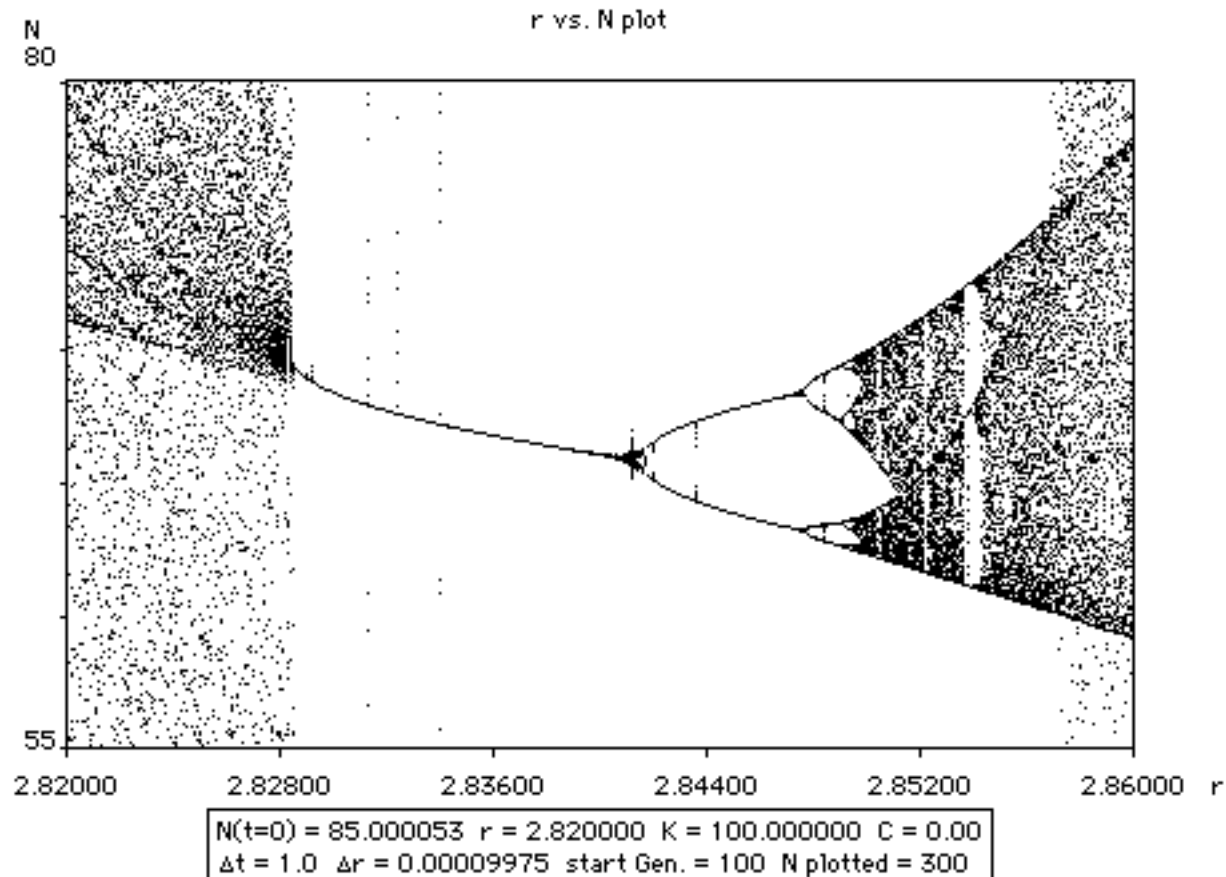


Figure 3-30. Blowup of a bifurcation plot or Feigenbaum tree, indicating loci of trajectories in between generations 100 and 400. (A large number of points were plotted so as to show a reasonable density of points in the blowup.) The blowup is of the middle branch in the period 3 window between r -values of 2.82 and 2.86, (compare, e.g., Figures 2-2, 3-2, or 3-6 above for full-scale bifurcation plots). This blowup shows the self-similarity of the middle branch with the whole tree, and also shows the locus of attention for the next two figures, 3-31 and 3-32, at $r = 2.856820$. Note also that some r -values are still showing transient behavior at generation 100, as indicated by the scattering of points in the space of the window and the transient convergences to the bifurcating twig indicated by the vertical projections to it from inside and outside at different r -values.

The next three plots (Figures 3-30 through 3-32) are sort of a graduation present for the interpretation of phase plots, and also indicate the usefulness—indeed, sometimes the absolute necessity—of having multiple ways of representing the same data to see what is going on. The first plot, Figure 3-30, is a blowup of a bifurcation plot to a horizontal or r -scale 25x greater than that of earlier bifurcation plots and a vertical scale about 6x as great, as indicated by the range of N and r values. It is centered on the middle branch of the period 3 window, but it also shows, in some more detail, what the "back" side of that window (where points start escaping from the more confined regions and again move chaotically throughout the whole range) looks like. It is that latter transition point, for a value of $r = 2.856820$, which is the subject of Figures 3-31 and 3-32; your charge is to figure out, from working back and forth between these three pictures (and particularly Figures 3-31 and 3-32), as much as possible about what is going on.

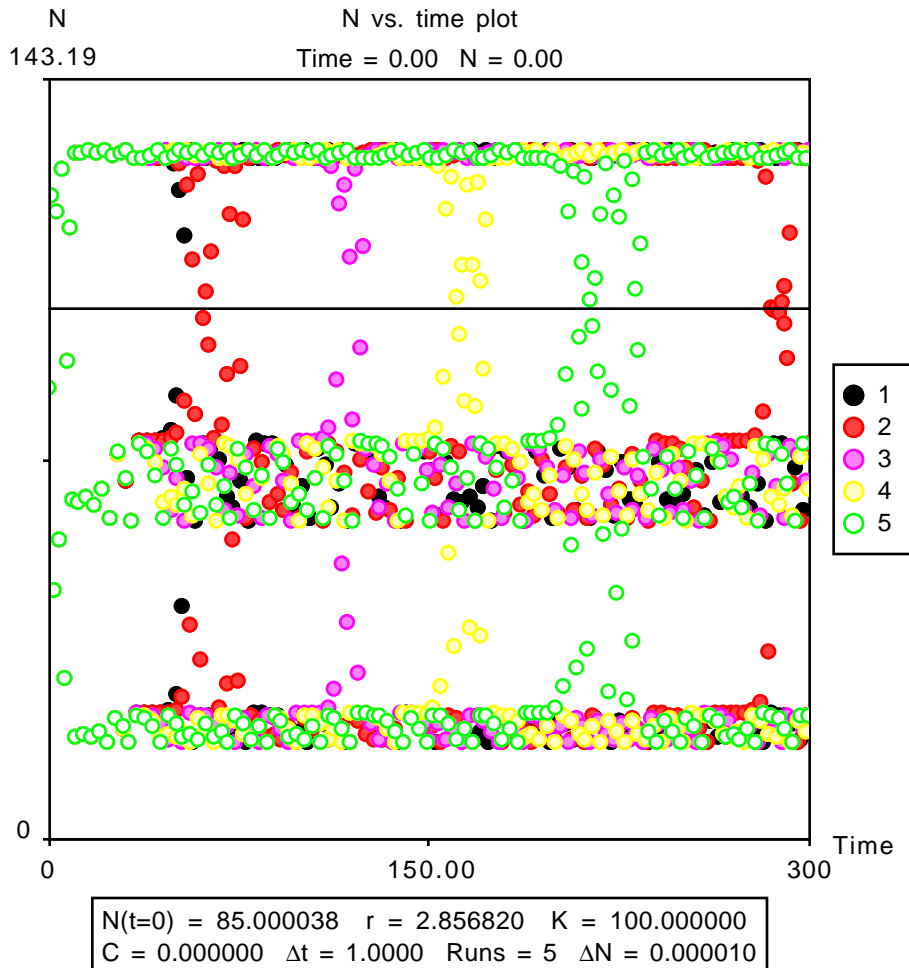


Figure 3-31. Population size vs time (N vs t) plot for 5 trajectories all with the same r -value which start very close together ($\Delta N = .00001$) and then diverge, showing the infinite sensitivity to initial conditions characteristic of chaotic dynamical systems. Note that in this case, the trajectories spend most of their time circulating regularly among three regions, with occasional excursions. Your job is to find out as much as possible about how they circulate among these regions.

Although it is easy enough—after the fact—to see what is happening from these pictures, there is no substitute in complex cases like this for running your own simulations, because then you can see how making your own changes in the parameters affects things (usually small changes, in a case like this) and you can run the graphs in slow motion, using the Redraw option, both of which are very useful to help you see what is going on. Although Figure 3-32 is perhaps most informative of all, there are ways in which each of Figures 3-31 and 3-32 are misleading without the other, so you will want to work back and forth between them to see what features in one correspond to observed features in the other. In a way, this interpretation job is an analog to the two parts of Figure 3-18 above; whereas that illustrated regular periodic behavior, these illustrate the interpretation of interesting statistical patterns in chaotic dynamics.

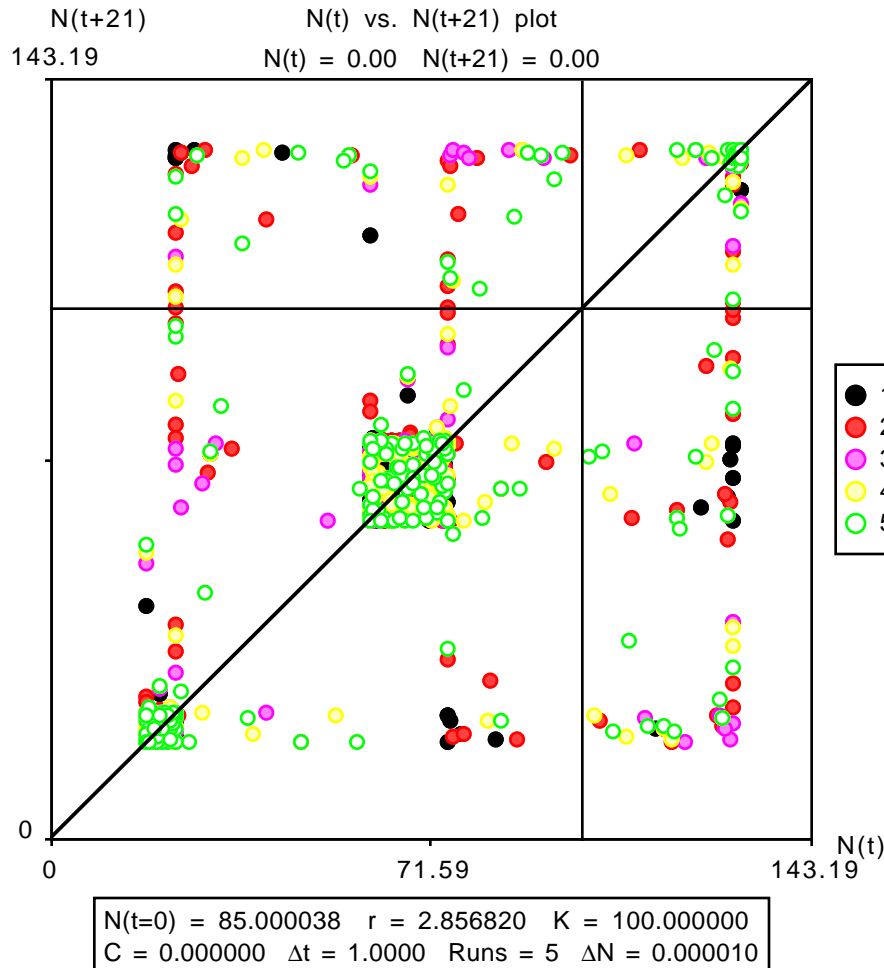


Figure 3-32. Phase plot of the same trajectories with a time lag of 21 generations. The regular patterns of this phase plot are very revealing of the statistical dynamics of the system. See the text.

In interpreting this phase plot, see if you can answer as many as possible of the following questions:

(1) Why are so many spots so tightly concentrated in the three rectangular areas symmetrically located on the stability line? Can you say anything qualitative about the relative numbers of points you would expect to find in each compared to the numbers in the others? Can you think how you would estimate what proportion of the total dots are in these areas? Why do the areas differ in size? (There's a quicker answer to this, and a deeper one.) What happens to the location of these areas when you increase (or decrease) the lag by 1? By 2? By 3? For what lags would you expect to find areas concentrated along the stability line? Go through these lags in order from smaller to larger: do you see any useful patterns in the changes in the distribution of the dots within the three regions?

(2) The overall pattern of Figure 3-32 looks much like a "scotch plaid" made up of dots. Given this characterization, can you explain vertical bars, horizontal bars, off-diagonal intersections, nearly white squares? How would you expect these to change for

different lags? Can you tell what would likely increase (or decrease) the whiteness of the squares?

(3) Why do most of the dots in the vertical bars appear to be along the right edge for the two leftmost bars, but along the left edge for the rightmost one? What can you learn from this? What can you learn from the fact that there don't appear to be any similar "edge effects" for the horizontal bars?

(4) Do you think that the phenomena of (3) constitutes a good statistical sample of the dynamics of the system? Give reasons for your answer which relate to how the patterns appear in both Figures 3-31 and 3-32.

Temporal Effective Population Size, $N_{\text{eff}}(t)$, the Logistic and Chaos

So far, there is one facility of the program that we have not yet talked about. This is the ability to calculate the temporal effective population size, $N_{\text{eff}}(t)$. We owe to Sewall Wright the development of the theory of evolution for finite population sizes, sizes small enough that things don't always average out (Wright, 1931, 1968-83, Provine, 1987). You can study some of the effects of population size on the evolution of populations using Frank Price's EVOLVE program. In this you can see that with a sufficiently small population, an allele may be lost simply due to random drift—or even, with a smaller chance, if it is selected for. This same process comes up when studying Mendelian inheritance, using John Jungck and John Calley's GCK—where you learn that an expected 3:1 ratio of phenotypes is consistent with substantial deviation from that, and that deviations tend to get more severe with smaller population sizes.

Unlike those programs, which have genetics built right in, this program makes no reference to or use of genetics whatsoever. Nonetheless, Wright derived a variety of formulas for how various processes involving fluctuations in population size affected the “average” population size for purposes of calculating the average rate of loss of alleles.⁵⁰ If a population fluctuated wildly in numbers, for example, the points when it had far lower numbers indicated a bottleneck. If there were N diploid organisms, then they together had between them $2N$ genes at each locus, so they could not pass on more than $2N$ distinct alleles, and would generally pass on far fewer than that.⁵¹

⁵⁰ Actually, the concept of effective population size (or number, as it is sometimes called) is better defined in terms of the flip-side of loss of alleles, namely the degree of inbreeding in the population. The recent molecular discoveries about “Eve”, the African hominid of 1.5 million years ago whose mitochondrial genes we all share shows that genes may become extinct even when populations do not. The human population shows 100% inbreeding—we are completely monophyletic—with respect to those genes that we all share from Eve. In this light, consider James Crow's (1986, pp. 43-4) definition of “effective population number” as “the size of an idealized population that has the same probability of identity [of alleles by descent from a common ancestor] as the population being studied.” In this idealized population, there is no selection, both sexes are equally frequent, it is of the same size in successive generations, any member of the population has an equal probability of mating with any other member, and the number of offspring of mating pairs is binomially distributed. Any deviations from these assumptions will cause the effective population number to be different from the actual population size.

⁵¹ Rather more generally if you have a large urn filled with multiple copies of a large number of

There are other complications. For sexual species, if one sex is far less frequent than the other, members of that sex would nonetheless contribute half of the genes to the next generation, and the genetic variability that could (or would) be passed on will be less than that possible (or probable) if both sexes were equally frequent. Ignoring the effect of sex ratio on effective population size in our calculations of temporal effective population size thus commits us to one of two assumptions: (a) that the species is asexual (monoecious) rather than sexual (diecious), or (b) that if it is sexual, the sexes are present in equal frequency, and all members of the population are mated.

Small effective population size is very important in evolution for several reasons, a few of which are:

- Small effective population size in a single population can result in the loss of genetic variability and thus the loss of an ability to adapt over evolutionary time. A number of species near extinction face exactly this danger. Recent publicity has focused on the cheetah, the panda, and other well known animals, but the danger is just as great with the widespread practice of monoculture using inbred and highly selected strains of domestic crops like corn, wheat, and rice.
- This may be due to natural factors or to biotic ones. New virulent plant diseases (fungi and other pests) can rip through uninterrupted monocultures like a forest fire, and without variability in the strain, it could become as extinct as the American chestnut has become, and more recently, the Dutch elm. Virulent strains may arise and spread because of the high frequency of a uniformly infectable host that would never do so in the face of a highly variable host population. (The study of these kinds of interactions is called epidemiology, and the breakout of epidemics is one area where chaotic dynamics are extremely likely—see Schaeffer, 1984). Therefore, it is important to notice that wild fluctuations in population numbers, by reducing genetic variability, may kill off a species as surely (though not as immediately) as outright immediate extinction.
- A species may subdivide itself into small populations with occasional migration between them as an adaptive strategy for a variety of reasons, or may do so for no selected reason at all. In such species, different populations go to fixation for different alleles, actually producing new genotypes which would never arise in a large population but do so in the smaller populations because of the concentration of the relevant genes there.

Studying effective population size is particularly interesting in the present context for at least three reasons:⁵²

different types of marbles, a smaller sample from this urn will tend to contain fewer types of marbles than a larger sample—and you will have lost variety, even if you are later allowed to enlarge your collection considerably by taking additional copies of any marbles that you have already.

⁵² There is actually a fourth. This work represents new and heretofore unpublished research resulting from the thoughtful questioning and superlative programming skills of Robert N. Farber, leading to joint research we did together in 1981 (when he was a sophomore in my biology class at the University of

- $N_{eff(t)}$ is a function of N , and because of this we can study chaos not directly as we have done in the first three modes (N vs t , r vs N , and $N(t)$ vs $N(t+m)$), but through its effect on another variable, which can sometimes be measured by other means. (See, e.g., Kimura, 1983.) This represents a common situation in science, where we often use variation in one variable as a way of assessing variation in another which is less convenient (or sometimes impossible) to measure. Indeed, discovering new ways of indirectly measuring things in this way often heralds major new scientific discoveries, and is a significant form of scientific invention.
- Comparison of r vs N , r vs $N_{e(t)}$, and both of these with t vs N in the light of the formula(s) for temporal effective population size provide a striking demonstration of the advantages of visualization in discovery.
- We give (and calculate) by two different means, one that can be found relatively widely in textbooks (e.g., Crow and Kimura, 1970; Roughgarden; 1979, Hartl, 1988) and described as an approximate formula, and another (given by Ewens, 1969, pp. 35-36) which is described as an exact formula.⁵³ Since much of the work in theoretical areas is constructing and then evaluating useful ways of approximating things that are difficult to calculate in other ways, it is useful to compare these two formulas to see under what conditions they produce divergent results.

The inexact formula for effective population size for a temporally (and cyclically) fluctuating population from generation 1 to k is given by equation (4):

$$\frac{1}{N_{eff(t)}} = \frac{1}{k} \left[\frac{1}{N_1} + \frac{1}{N_2} + \frac{1}{N_3} + \dots + \frac{1}{N_k} \right]. \quad (4)$$

The so-called "exact" formula for the same conditions is given by equation (5):

$$\left\{ 1 - \frac{1}{2 N_{eff(t)}} \right\}^k = \left\{ 1 - \frac{1}{2 N_1} \right\} \times \left\{ 1 - \frac{1}{2 N_2} \right\} \times \left\{ 1 - \frac{1}{2 N_3} \right\} \times \dots \times \left\{ 1 - \frac{1}{2 N_k} \right\}. \quad (5)$$

These formulae, rearranged so as to give explicit solutions for $N_{eff(t)}$ are used in the " r vs $N_{e(t)}$ " plot mode as follows: A starting value of N , a range of r values, and a range of N values to be plotted are specified. By analogy with the " r vs N " plot mode, the first set of values are calculated, but ignored, up until "start Gen." (These values are not irrelevant, of course, since they determine what the value of N will be at start Gen.)

Chicago). It was Farber's idea to do effective population size calculations for the logistic, and our joint idea that it would be interesting to study its behavior in the chaotic realm, which we did on his North Star "home brew" Z-80 machine. With this we discovered for ourselves many of the basic phenomena you can find out for yourselves with the $N_{e(t)}$ vs. r plot mode, though we hadn't yet seen a bifurcation plot through its whole range, and thus didn't realize that the period 3 window was the cause of the trench or "crash" in effective population size. Since then, undergraduate research in my classes has never been quite the same!

⁵³Interestingly, only in Ewens, is this formula described as applying in a cyclical environment—presumably to damp out transients from prior stages.

Then the next “Gens plotted” generations are used in the calculation for effective population size, as N_1 through N_k . Both formulas are plotted, the exact equation with black solid lines, and the inexact equation with dotted lines (which are red with a color monitor). The exact equation is plotted last, so that it overlays the inexact equation. They are thus distinguishable only when their results are different by at least a pixel width.

Mathematical Artifacts Revisited—Case 2: When Is an Exact Formula Less Exact Than an Inexact One? When You Least Expect It!

Exercise 12: Mathematical Artifacts

As you can see from these formulas, with either one, small values of N have a disproportionately larger effect on $N_{eff}(t)$ than large values. The exact formula is supposed to be more accurate than the inexact one for smaller values of N , but how small? Try $r = 0$ to 3, start Gen = 0, and Gens plotted = 10 for various starting values of population size, N . (Try starting N at 50, 25, 10, 5, 2, and 1.) Try a value of N between 0 and 1/2. What happens? Which formula is acting strange? Why is this happening? (Look at the formula and you can see why!) Is this biologically realistic? (This has two parts: (a) Is it realistic that population sizes should be this small? (b) Is it realistic that $N_{eff}(t)$ should behave like this?) Now look at the behavior of the two formulas close to $r = 3$ (for $C = 0$). (You will have to get very close.) Can you explain what is going on? Which formula would you trust, and why? Do you think that the exact formula *is* exact? (In other places in Wright’s theory where terms like $\{1-(1/2N)\}$ appear, there is often a qualifier that N should be ≥ 5 .) When you are doing model-building, you will sometimes have to check for yourself some things that others tell you on faith. On the other hand, with the power of a computer, that isn’t always such a hard thing to do.

Relationships Between r vs N Plots and r vs $N_e(t)$ Plots

These two kinds of plots show interesting parallels in their behavior. Some connections are to be expected, as they are both graphs relating the behavior of N to r . Formally, only the second (r vs $N_e(t)$) is a (single-valued) function, since the r vs N plots can generate many values for a given r value. But in fact, for a given value of r , and for the same settings of start Gen and Gens plotted, $N_e(t)$ is a function of exactly those points that would be plotted in the “ N vs r plot.” Thus comparison of these plots can be particularly informative.

Here are some of the comparisons you can make:

- In the chaotic regime, the exact trajectory depends sensitively on the initial value of N , the exact value of r , and the ratio $(1-C/K)$. You will be plotting curves for a range of r -values, but the starting values of N , and the ratio $(1-C/K)$, which acts to modulate r can generate different results. Run comparable N vs r and $N_e(t)$ vs r plots for different values.

- One systematic way of changing N is to set the starting value of N , and, for a given value of Gens plotted, vary start Gen to pick out different temporal windows of behavior for comparison using the different graph types. In fact, as you will see in the series below, a series of these temporal windows shows, in either plot type, systematic changes that exemplify the temporal unfolding of chaotic divergence.
- Applying these two points together, these patterns of unfolding can be expected to change for different starting values of N .

We will start by exhibiting parallels between the two plot types over the whole range of r :

Note that in the first plot (Figure 3-33), the exact (lower) curve and the inexact (higher) curve *are* distinguishable. This is because of the very low starting value of N chosen. (Verify this for yourself.) This picture was taken using color (and solid) patterns, so the different patterns for the two curves do not show up. The second plot (Figure 3-34) is the now familiar bifurcation plot for the corresponding broad range from $r = 0$ to $r = 3$.

Note in these two plots which broad features correspond to which. As you might expect, $N_{eff(t)}$ gets the jaggies as r gets up into the chaotic regime. But why do you think that N climbs so slowly with increasing r while N in the N vs r plot goes rapidly up to 100? You might experiment with different viewing windows for this—say start Gen of 0, 20, 40, 100 for Gens plotted of 10. An understanding why this happens is very important for understanding how the $Ne(t)$ vs r plots relate to the N vs r plots.

There follows a gallery of corresponding N vs r and $N_{eff(t)}$ vs r plots for your inspection. (See Figures 3-33 through 3-48 below.) Compare these or generate your own, and see what you can find.

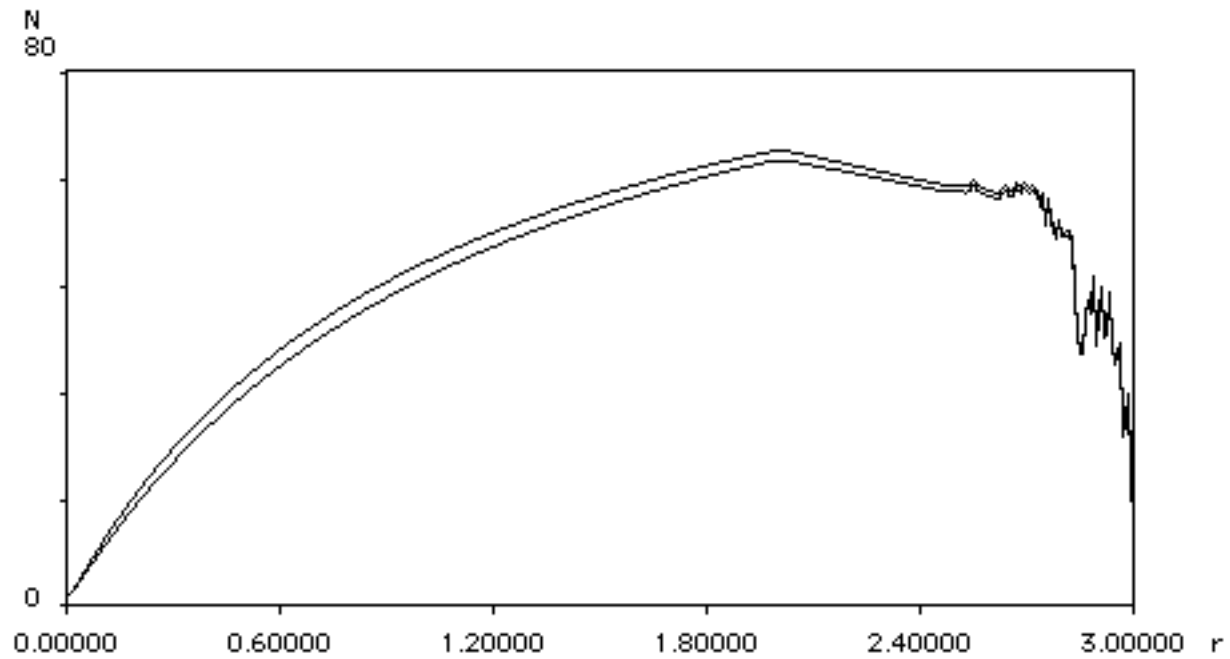


Figure 3-33. $Ne(t)$ vs r plot, for $0 \leq r \leq 3$, with a starting N of 2, start Gen = 200, Gens plotted = 100.

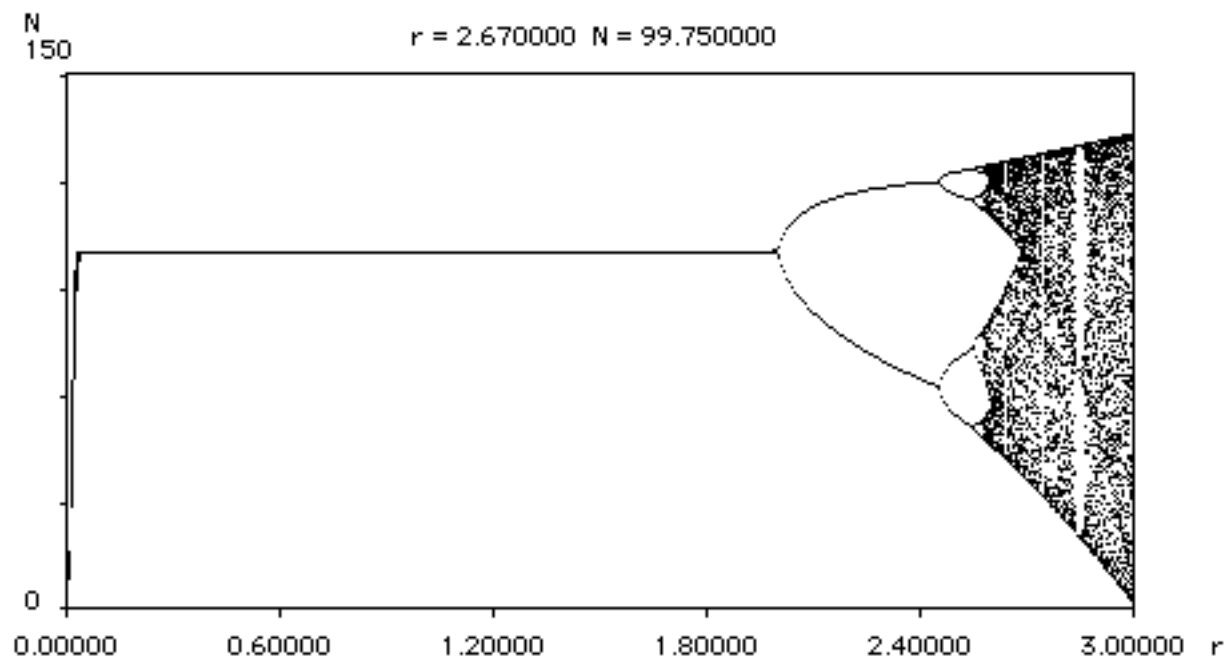


Figure 3-34. The corresponding N vs r plot. Try these for a larger starting value of N .

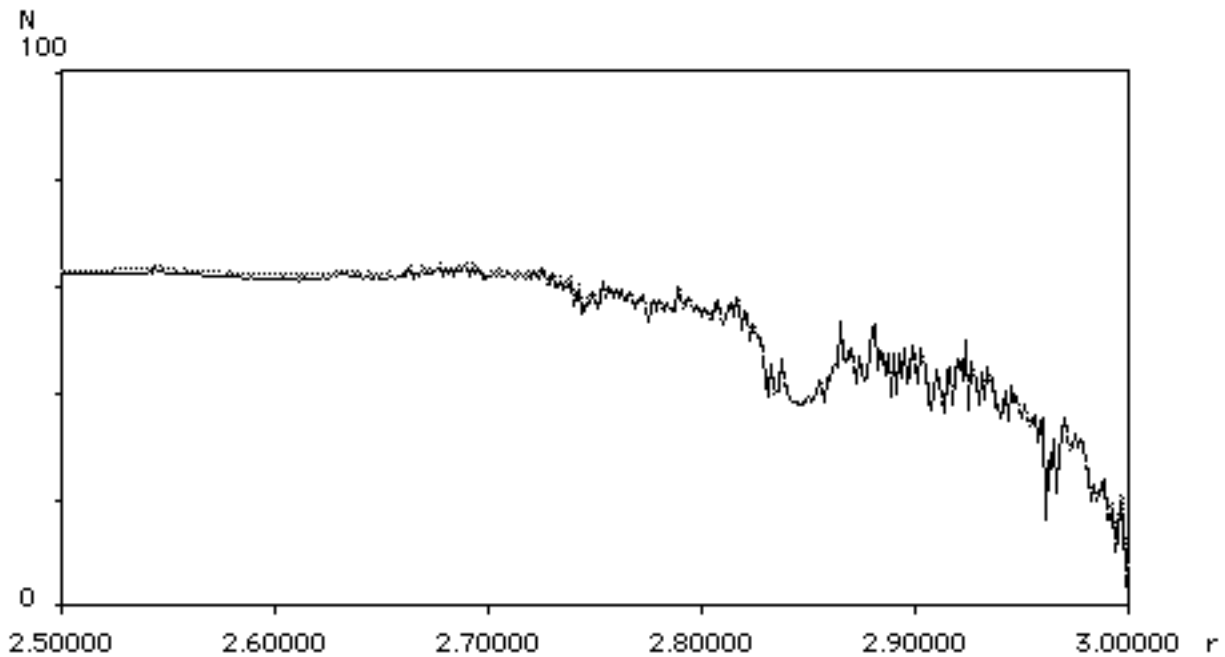


Figure 3-35. $N_e(t)$ vs r plot for $2.5 \leq r \leq 3.0$, a starting N of 2, start Gen = 0, and Gens plotted = 100. Note that the approximate and exact curves are distinguishable, particularly to the left of the graph.

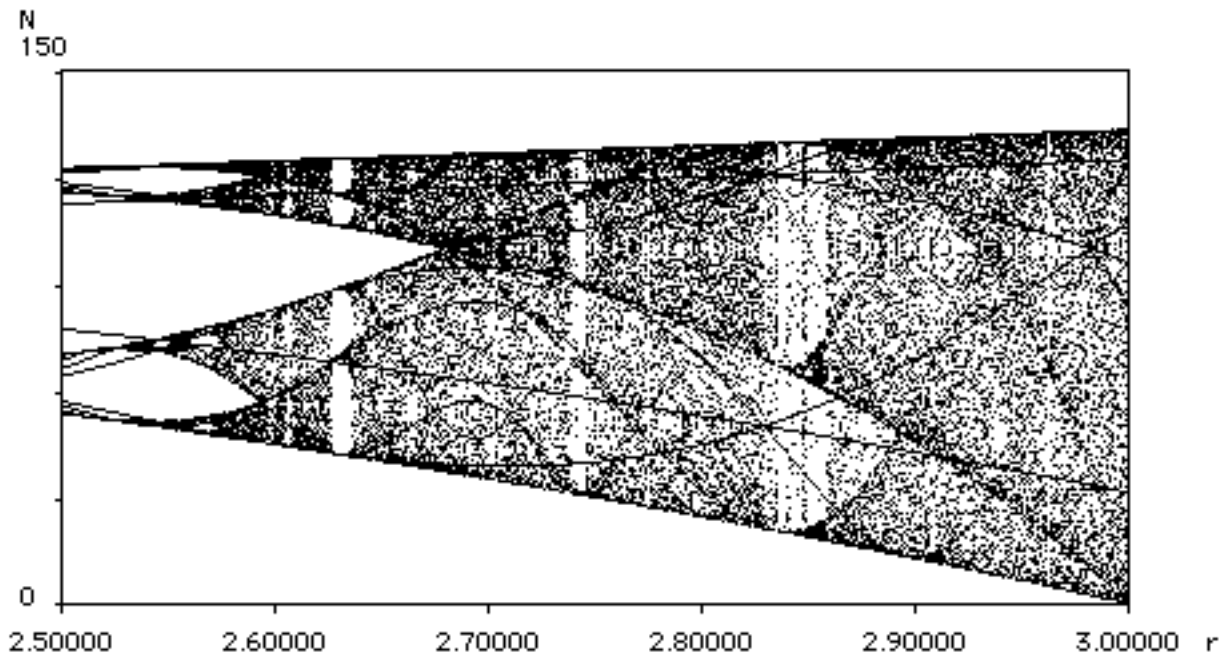


Figure 3-36. N vs r plot for the same values. Compare bands in the N vs r plot with minima in the $N_e(t)$ plot. Do you see any other features you can explain?

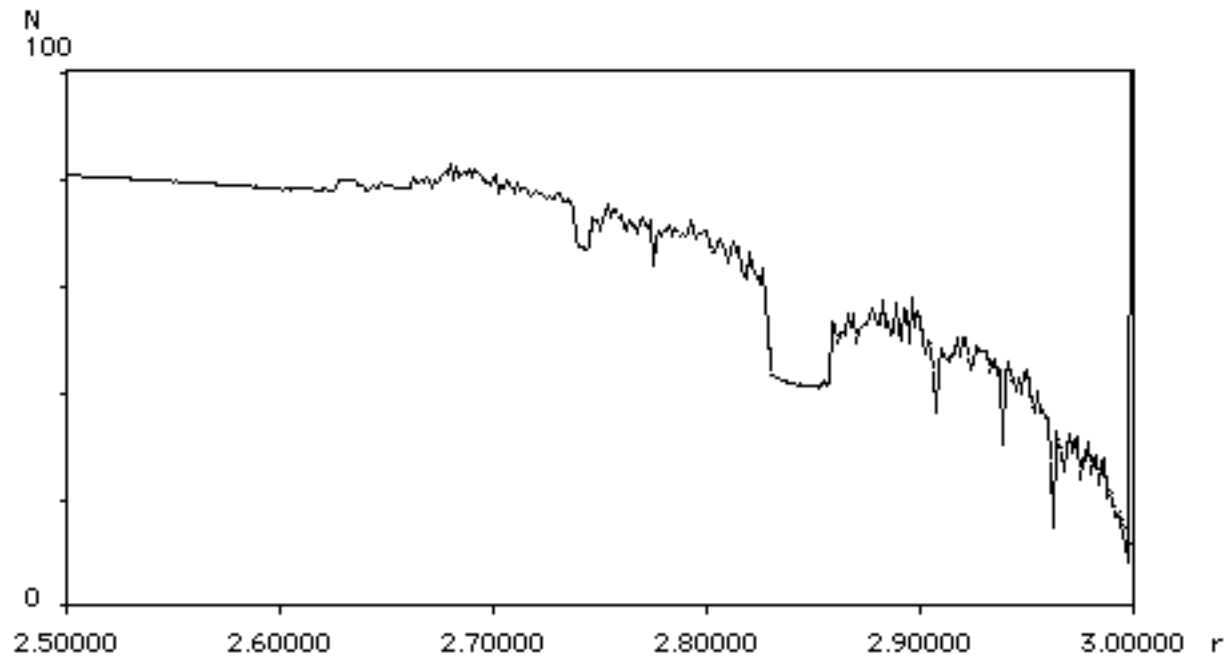


Figure 3-37. $N_e(t)$ vs r plot for $2.5 \leq r \leq 3.0$, a starting N of 2, start Gen = 200, and Gens plotted = 100. Notice the differences with Figure 3-35. Can you explain them in the light of the differences between Figures 3-36 and 3-38?

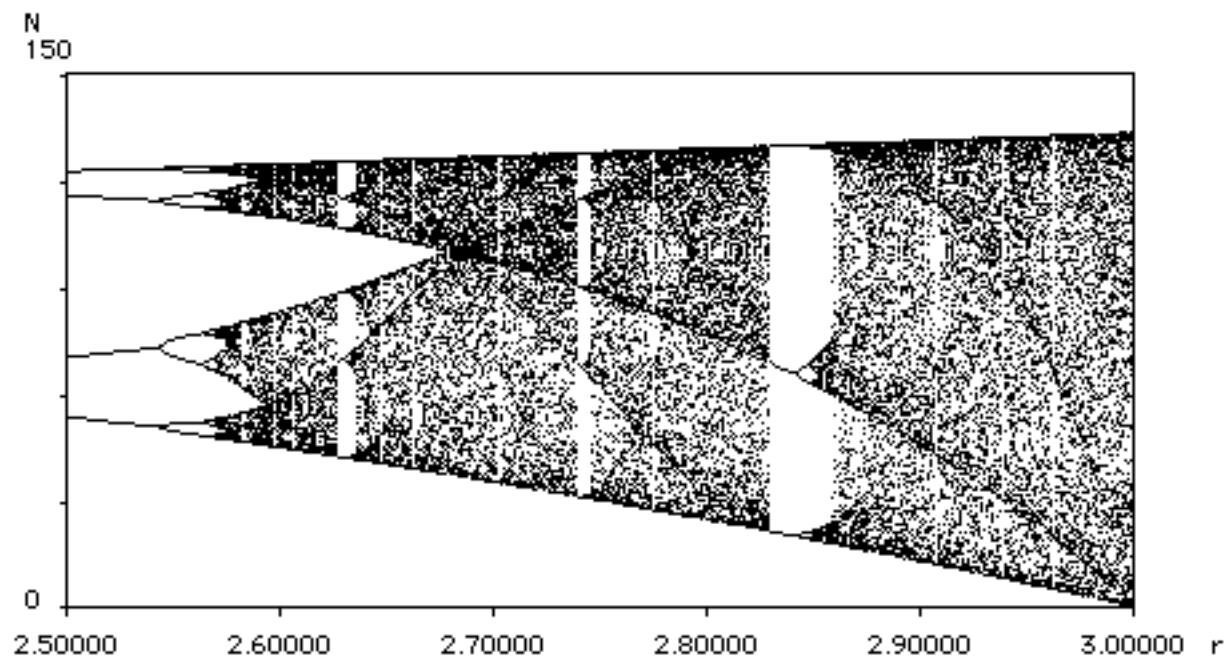


Figure 3-38. N vs r plot for the same values. Compare bands in the N vs r plot with minima in the $N_e(t)$ vs r plot again. Do you see anything else interesting to compare?

The next series of graphs is a series of N vs r plots for very narrow windows at different starting points.

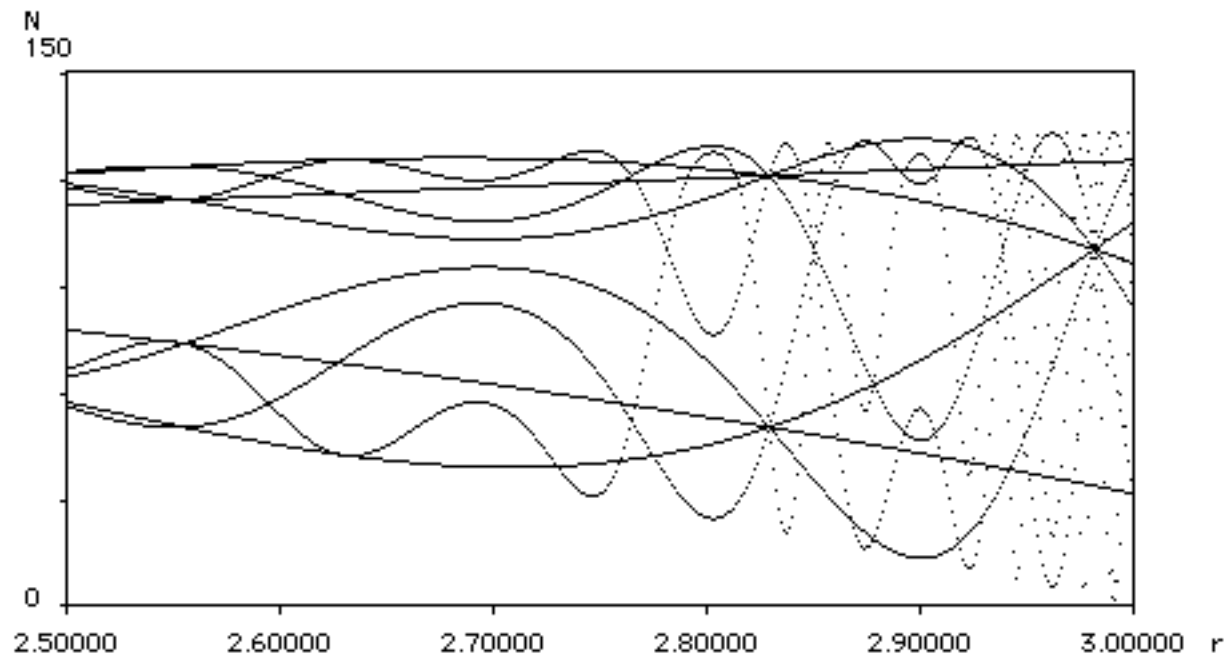


Figure 3-39. N vs r plot; $2.5 \leq r \leq 3.0$; start Gen = 0; Gens plotted = 10. I haven't said what the starting value of N is because I lost it. This program now automatically saves all relevant parameters with the graph. Figure 2-3 above is a plot of the first 5 generations. A comparison of this with Figure 2-3 will show you which trajectories are the first 5.

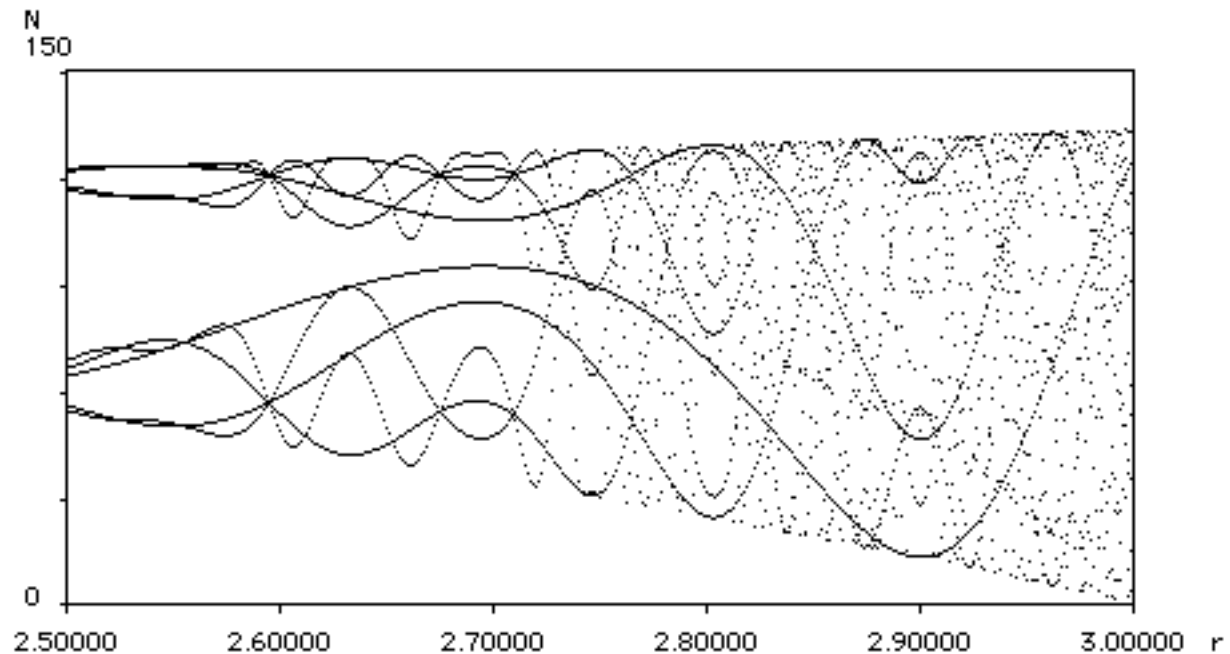


Figure 3-40. N vs r plot; $2.5 \leq r \leq 3.0$; start Gen = 10; Gens plotted = 10. Note that later generation trajectories get curlier for lower r values. Can you explain this? You may also want to look at what happens in the chaotic regime for $n(t)$ vs $N(t+m)$ plots for increasing values of m .

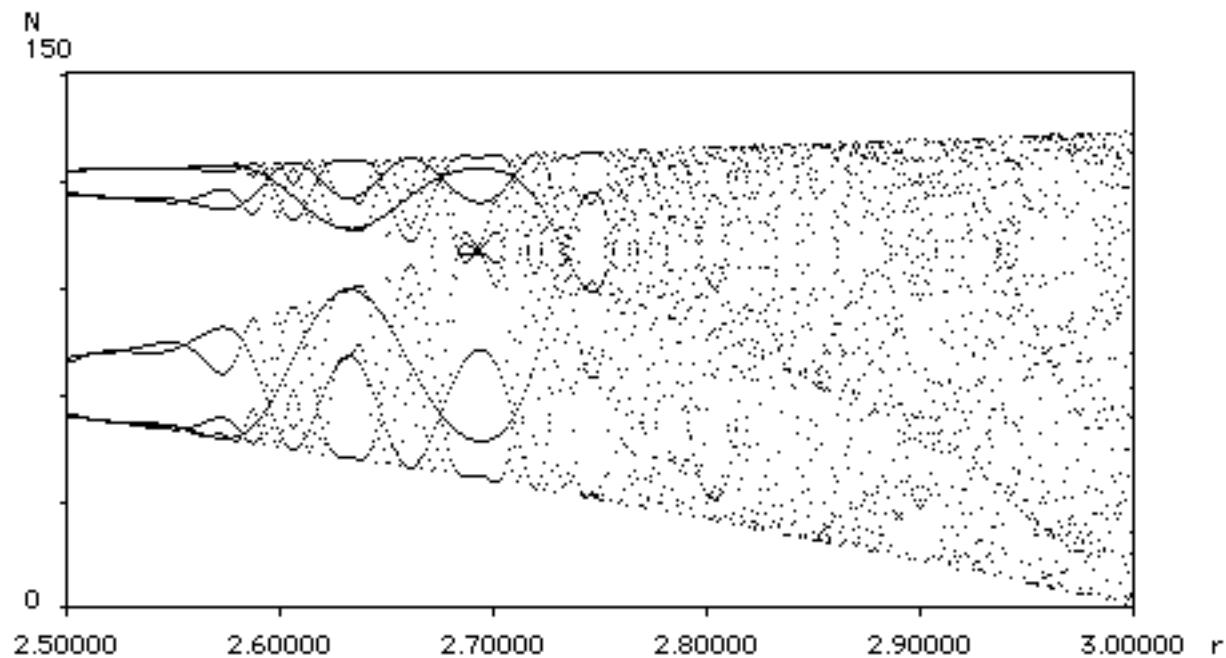


Figure 3-41. N vs r plot; $2.5 \leq r \leq 3.0$; start Gen = 20; Gens plotted = 10. The expanding and crossing chaotic envelopes are now clearly visible.

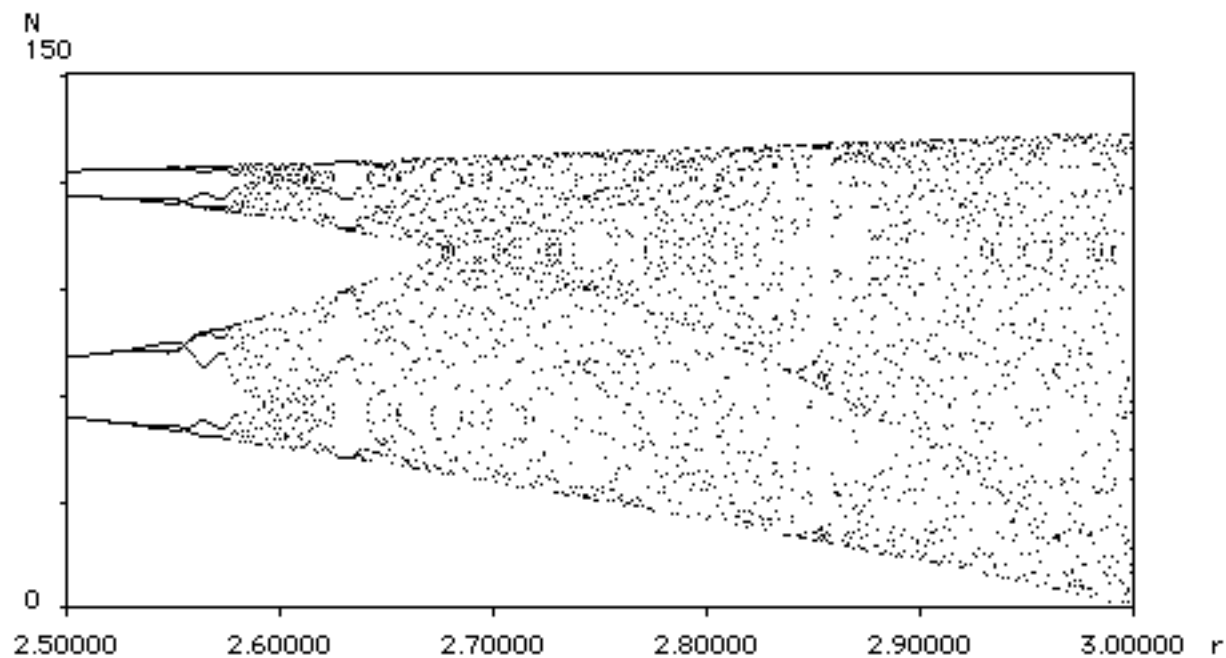


Figure 3-42. N vs r plot; $2.5 \leq r \leq 3.0$; start Gen = 30; Gens plotted = 10. Bands are beginning to emerge. Look for self similar features on different scales for different values of r in these successive time slices.

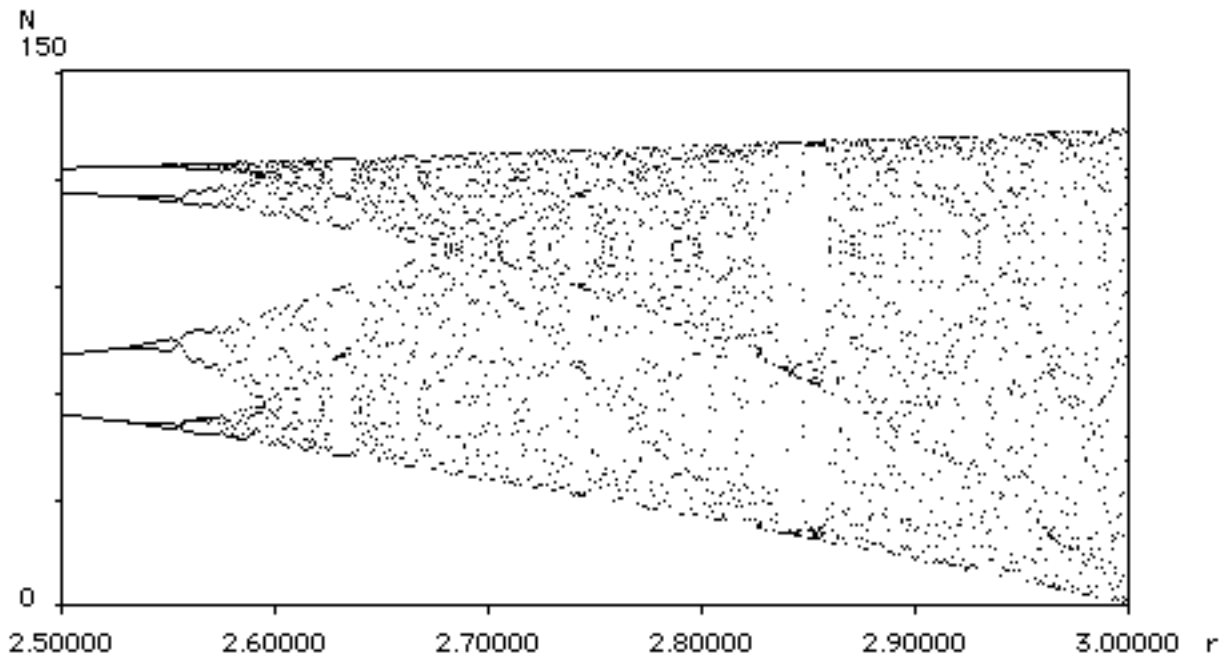


Figure 3-43. N vs r plot; $2.5 \leq r \leq 3.0$; start Gen = 40; Gens plotted = 10. Bands and envelopes are becoming more well defined.

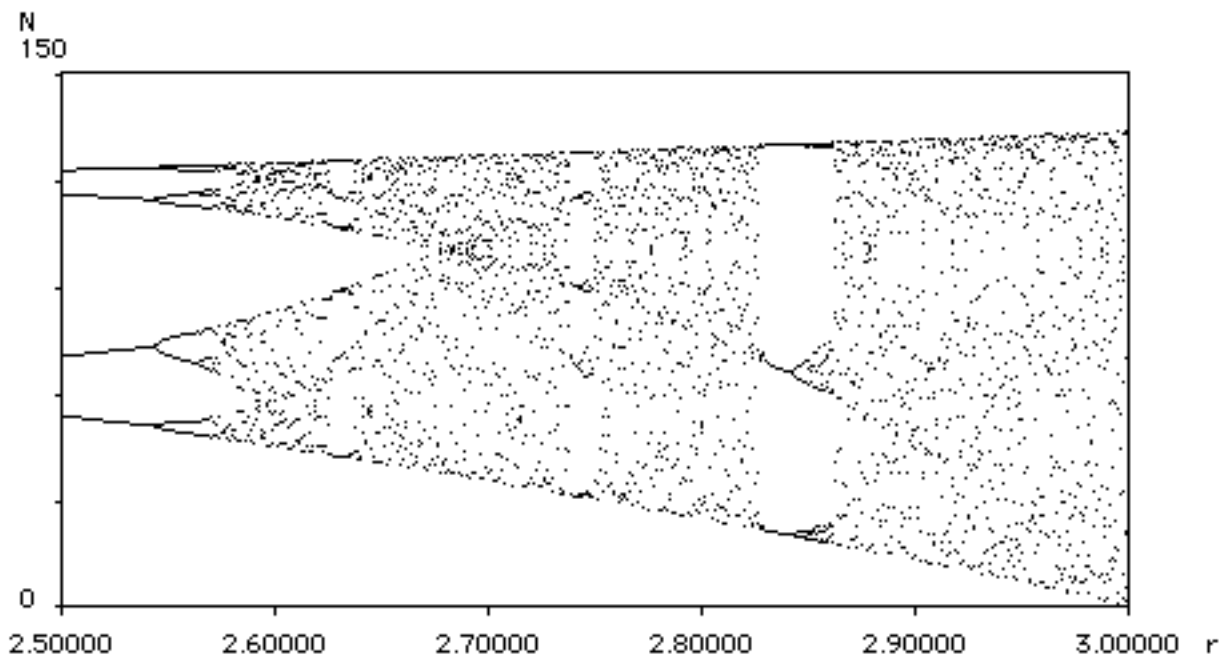


Figure 3-44. N vs r plot; $2.5 \leq r \leq 3.0$; start Gen = 200; Gens plotted = 10. This is the last in the series, and skips 150 generations until most of the transients have disappeared. Bands are as well defined as they can be with only 10 points plotted per r -value. Try experimenting with plotting fewer points than there are twigs in a band. What happens? Add generations to start Gen, 1 at a time. What happens? Can you explain it? What does the fact that a twig is either present or absent (and not, e.g., dotted) tell you?

The following series of graphs are $N_{\text{eff}}(t)$ vs r “strip charts” (note the versatility of the graph resizing capability) for most of the same values as the preceding N vs r plots. Compare them with one another and with the corresponding N vs r plots.

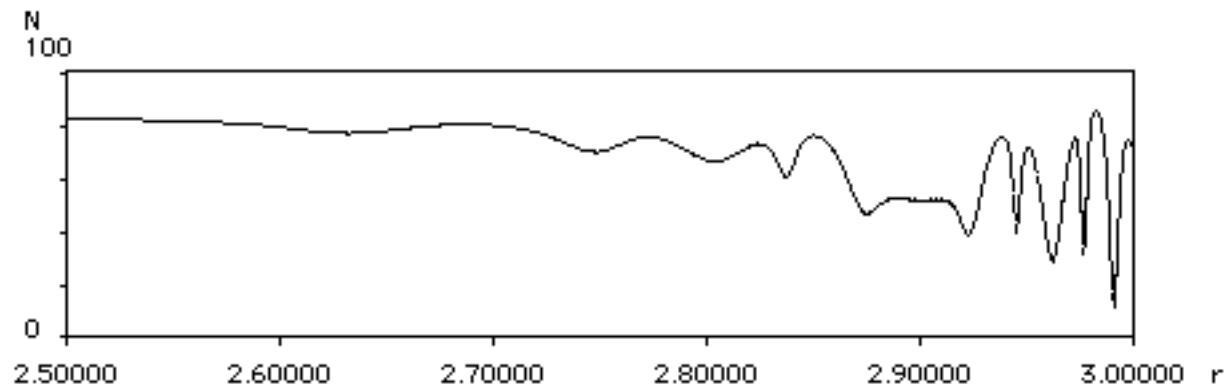


Figure 3-45. $N_{\text{eff}}(t)$ vs r : $2.5 \leq r \leq 3.0$, $N_{\text{start}} = 2$, start Gen = 0, and Gens plot = 10

Note the “false valley” from about 2.87 to 2.93, which will disappear in later generations.

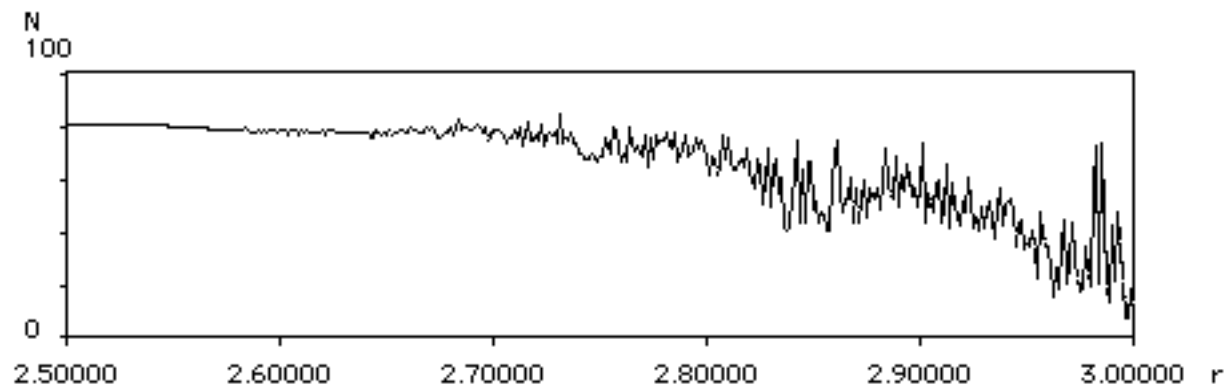


Figure 3-46. $N_{\text{eff}}(t)$ vs r : $2.5 \leq r \leq 3.0$, $N_{\text{start}} = 2$, start Gen = 10, and Gens plot = 10

Note how rapidly the curve gets jagged with increasing generations. Can you explain why? There are some big minima (but also big maxima) covering a broader region including the period 3 valley.

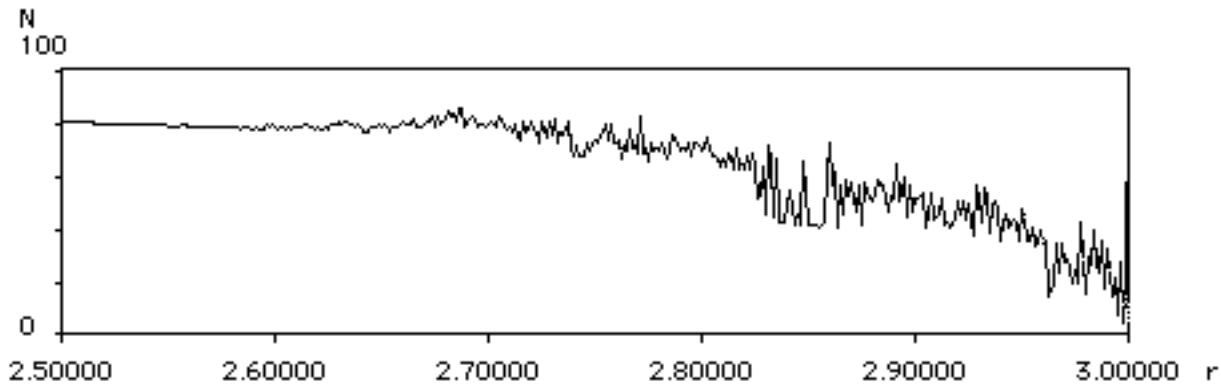


Figure 3-47. $N_e(t)$ vs r : $2.5 \leq r \leq 3.0$, $N_{\text{start}} = 2$, start Gen = 20, and Gens plot = 100.

The period 3 valley is beginning to acquire some character, and the curve is losing some of its jaggies.

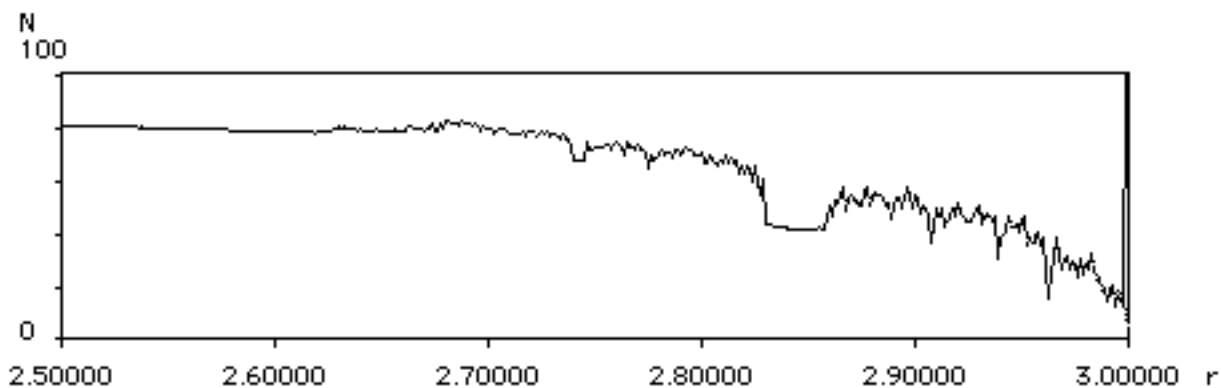


Figure 3-48. $N_e(t)$ vs r : $2.5 \leq r \leq 3.0$, $N_{\text{start}} = 2$, start Gen = 200, and Gens plot = 10.

This is the last in the series. Various valleys corresponding to various bands are quite well defined, and this curve has lost still more jaggies, suggesting some form of filtering of extremes, which I do not yet understand. Can any of you explain it?

Here is something else to try—compare corresponding strip charts for other starting values of N . They aren't all the same.

With this, we leave you to explore the logistic family of equations. There is a lot more to do, not all of which we have talked about. (Indeed, every time I think I've exhausted the topic, something else comes up, as often as not from my students, to prove that I'm wrong.)

Thought Questions on Visual Representation

1. In what ways do these aspects of visual analysis of the logistic model generalize to the analysis of other models?
2. Can the visual analysis of the logistic model suggest ways in which we can make truer models of population growth—or model other things not considered here?
3. Are there other visual modes of presentation which we have not yet explored? (Yes there are, and more of them may well show up in the release version of this software and manual. But if you can think of any you'd particularly like to see, let us know.)
4. In what ways can the visual analysis of the logistic model help us to interpret empirical data?
5. In what ways can we distinguish computer effects from mathematical consequences of the model?
6. In what ways is visual analysis useful in detecting artifacts?

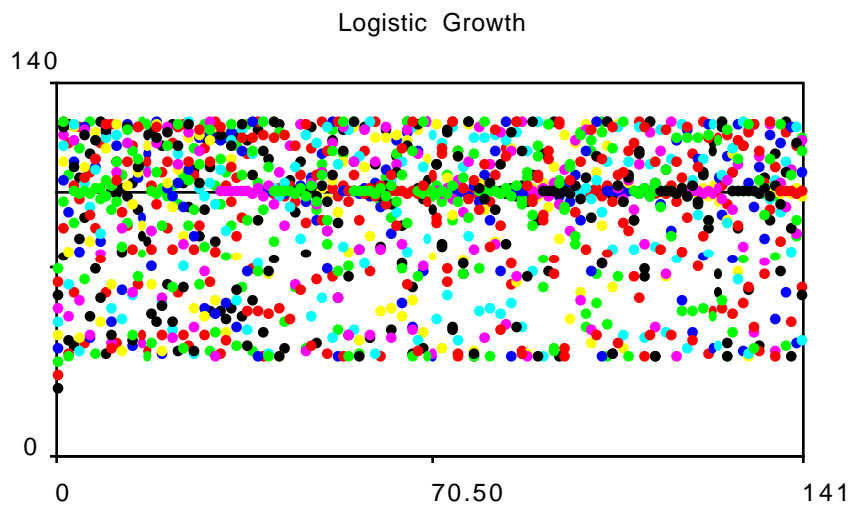


Figure 3-49. Mystery figure (relatively easy).

Where (i.e., for what conditions) do you think this is happening? What kinds of clues do you have (in general) for identifying graphs? A good starting point might be the bifurcation chart. You also have some other clues. What can you learn from the y -scale?

Other Food for Thought:

I. Simplifications of the Logistic and MDLS Equations

Appendix D contains as exhaustive a list of the simplifications of the logistic and MDLS equations as I am aware of. Think about these assumptions and try for each of them to determine:

1. Why someone would make such an assumption. Is it intrinsically plausible in general? Does it simplify computation? Are there ideal systems or limiting conditions for which it applies? Does it tie in with standard assumptions in other models? Are there real systems for which it is plausible?
2. If the assumptions are unrealistic, what biasing effects would you expect them to have? What changes would increase the realism of the assumptions? What questions do these changes raise? Do they in turn suggest further directions of elaboration?
3. Can you learn anything more about the use of these models or variants on them which we have not discussed? What motivated their use, and what shortcomings, if any, did they have?

II. Behavior of Dynamical Systems

These are behaviors of dynamical systems in general which are exemplified in the equations of the MDLS model. How many can you find? Specify how you would tell or test for them.

1. Continuous growth/decay
2. Damped oscillations
3. Stable 2^n point oscillations
4. Chaos (according to which criterion?)
5. Finding 3, 5, 6, 9, etc. cycles
6. Finding 2^n point cycles.
7. Finding cycles that are longer than 400 generations.
8. Infinite sensitivity to ΔN
9. Infinite sensitivity to Δr
10. Dependence on Δt ; $r_{\text{eff}}\Delta t_{\text{eff}}$ as a sufficient parameter for chaos.
11. δ and estimation of bifurcation points; λ and estimates of spread of branches.
12. Universality, effective r , and comparison of logistic and MDLS bifurcations.
13. Some other values to look at:
 - a. 2.8537-9
 - b. 2.56995
 - c. 2.83-2.85
 - d. 2.72
- e. How would you test the conjecture: all stable equilibria in the 2^n range remain as unstable equilibria as they bifurcate with increasing r . Aside from testing it, how would you argue for it?
- f. How many different periodicities can you find?

III. Other Thought Questions

1. Suppose someone claimed that the system you are investigating is not totally deterministic, but had a small amount of normally distributed random variation added at unknown (periodic or random) intervals, which was what was responsible for the apparent sensitivity to initial conditions. (a) Does this claim make sense? (b) How could you tell whether there was a small random noise input to your deterministic system? What test or tests would you make? (c) Can you totally refute this claim? (d) Can you show that (or under what conditions) it is implausible? (e) Can you constrain it (e.g., by showing that the variation added has to be less than a certain amount, and/or with less than a certain frequency)? Justify or explain your answers.
2. How can you tell where roundoff error is acting in the simulation? (Think of yourself as an auto mechanic, listening and looking for the characteristic signs and sounds of different kinds of breakdowns. Can you tell anything about the source of the roundoff errors by the conditions under which they appear, and the form of mistakes they involve or produce? What do you know about the computer or the program (from what we have told you about it) that indicates possible sources of breakdown)?
3. From your study of chaotic dynamical systems, do you think that chaos is important in nature, or do you think it's just a mathematician's pipe dream? If you think it is important, say where (under what real-world conditions) you would expect to find it. If you think it isn't, say why not. How would you tell chaos from randomness in nature? (See Wimsatt, 1980, for discussion of this question.)

Understanding the Role of Mendelian Inheritance in Evolution by Simulating Blending Inheritance

In this simulation, unlike Logistic Growth and most other simulations you will find in BioQUEST Modules, you will be dealing with organisms which are imaginary in two senses:

(1) Many simulations or models in science assume “ideal organisms” which have some features—those feature selected for modeling—which are like those of organisms found in nature, but may in detail, and in most other respects, be quite unlike the organisms being modeled. The Logistic and Minimum Density Limited Logistic Growth models discussed in the last two chapters are of this sort. In the other BioQUEST modules, every attempt has been made to make and base simulations on ideal organisms which are governed by laws and mechanisms which are, in relevant ways, like the laws and mechanisms governing the behavior of organisms in nature—at least as far as the theoretical perspective being modeled is concerned. (Given that nature is more complex than we are able to handle easily, we'll usually simplify extensively in all other details, and usually also make any simplifications that we think we can get away with even for the phenomena under study.) Thus, the theoretical focus of GCK is on Mendelian genetics, so its ideal organisms exemplify these laws and mechanisms with substantial accuracy and detail but don't, for example, say anything about the biochemistry, the physiology of digestion, or mechanisms for thermoregulation of the ideal organisms it simulates. Less happily, GCK at present also does not model pleiotropy or epistasis, though both are clearly relevant to classical genetics. μ GCK uses biochemical information in the specification of plate media, and Sequence it! and

Purify it! use much more biochemical information because this knowledge and these mechanisms are more relevant to what is being studied in those simulations. (Of course, they may leave out or simplify other details, for computational or other reasons.) All of the other BioQUEST simulations are of this form: *such simulations model "ideal organisms" as realistically as possible relative to the theoretical perspective of concern.*

(2) A second kind of simulation involves supposing as correct certain laws, mechanisms, or features of the world that we believe to be false in order to draw out the different consequences of making or not making the assumptions that we typically do about the world. The blending inheritance simulation discussed here is of this often-ignored but very important type.⁵⁴ It explores the consequences of a mode of inheritance which was championed by Darwin and virtually all of his contemporaries, and later by others (the so called "Biometricians") who opposed the Mendelian point of view when it first emerged.⁵⁵ The kinds of patterns of inheritance that made the "blending" models plausible were, in a triumph of the new Mendelism, reconceptualized as Mendelian multilocus additive traits, and are today the subject matter of quantitative genetics. This kind of simulation (i.e., a simulation based on what we now accept as a false model of inheritance) is not directly relevant for deriving and understanding the consequences of theories of inheritance that we hold to be true, but it may be indirectly relevant, in that it can be crucial for understanding why we accept certain theories as opposed to their competitors. It may also be crucial for understanding the similarities and differences between these theories—usually a necessary requirement for designing an adequate experimental test of the theories—a matter that is often more difficult than our *ex post facto* histories suggest.⁵⁶ For a collection of essays which explicitly recognizes this kind of use of models, see Nitecki and Hoffmann, 1987.

But this is not just a matter of historical interest. Such fundamentally false models shed light on current and future investigations in science as well. What happens if we don't know the laws of nature or the mechanisms involved? What if we are trying to develop, test, or elaborate, a theory of what these laws and mechanisms are? What if one of the theories we have been taking for granted turns out to be false, as has happened many times in the history of science, and certainly will continue to happen in

⁵⁴Other examples of this counterfactual type of "false model" in physics include the various types of perpetual motion machines, each one of which violated one of the basic laws of physics, and which lead to enthusiastic puzzle solving to figure out why they were impossible in the 18th and 19th centuries, and Newton's "demonstration" that the solar system was stable for the inverse square law of gravitational attraction, but not if force varied as the inverse 1st or 3rd power. (This "demonstration" was later shown to break down for the many-body problem even for the inverse square law by Poincaré—whose results presaged the discovery of Chaos. See Stewart, (1989).)

⁵⁵ See, e.g., Provine, 1971, on the causes and consequences of the controversy between Mendelians and Biometricians, which had a significant influence on the form and some of the key assumptions of the modern theory of evolution.

⁵⁶The latter (i.e. false and rejected theories) are rarely if ever discussed in science textbooks, so many of the reasons for accepting the theories that we have, as well as the logic of their confrontation, are often lost to students. The need to take this broader perspective—discussing past theories that we now regard as false—is why it is often argued that science students should study history of science as well as the current state of the art. For further discussions of this broader perspective of science and the relevance of teaching history of science to scientific practice, and for an explanation of the usually inaccurate "reconstructed" history of science taught by and for scientists, see Kuhn, 1970.

the future? Science is more than doing experiments and gathering data—it also involves the structured exploration of new phenomena, the choice and designing of experiments, the construction, testing, rejection, and modification of theories, hypotheses, and models, and the re-evaluation and sometimes rejection of data as irrelevant or as collected under conditions which render it useless.

The truth of almost anything in science has been up for grabs at one time or another, and false models are often effective means for getting to truer theories—particularly if we know something about the ways in which they are false. Indeed, we must every day work with models that we know are false in some (often in many) details. Model builders know and use this fact effectively, but it is seldom presented in science textbooks or popular presentations of science, for very interesting reasons.³ (See the appendix on “Functions of False Models as Means to Truer Theories” or Wimsatt, 1987). False theories may be (but usually aren't) the result of bad science, and in any case we have to learn to live with them—in many cases, they are all that we have. We need to learn more about how to reason productively with models which are true in some ways, false in some others, and unknown (we do not know, and would like to find out!) in still others.

In a small way, of course, this hidden aspect of science is required for effect to cause reasoning, as exemplified in the BioQUEST program GCK, where we take the laws of Mendelian genetics as given, but we are uncertain what the mechanisms are that produce the particular phenotypic patterns we are trying to analyze. (This pattern of reasoning is also present in the LOGISTIC simulation when we are trying to explain what is producing a given pattern, or change of pattern.) Mechanisms produce systematic patterns of behavior, so in some ways, they are kind of like local special laws. Or maybe it is more correct to say that once we understand the mechanisms behind a set of phenomena, we understand the conditions under which regular behavior is to be expected, and the factors that can cause deviations from expected regularities. In trying to analyze a mating pattern, we guess hypothetically what the underlying mechanisms are which produce it. Do the phenotypic patterns we see involve 2 alleles or more? Variant alleles at 1 locus, or 2, or more? Are the patterns a product of sex-linkage? Linkage? Dominance or codominance, etc.? Data can be thrown away as well as hypotheses. Some of the data we produce in trying to discover the mechanisms are crucial, but a fair amount of it we later decide are irrelevant, artifactual, biased, or worthless. But there can be deeper uncertainties, as when the very mechanisms of inheritance are unknown (as before the rediscovery of Mendel's work in 1900), or only partially known (as with the period between then and 1920, when the basic mechanisms of classical genetics were still being worked out and the subject of heated debates). In dealing with such questions we are dealing with the problem situation found at the cutting edge of science; the establishment of today's givens were the cutting edges of past investigations and debates.

The basic rationale behind constructing and evaluating hypotheses is, in most respects, the same in both cases—that small modifications engender and add up to larger ones, and that scientific revolutions differ in size, but not in kind, from the kind of piecemeal daily manipulations, discoveries, and reconceptualizations of lab, field, and problem set. But it is important to look at these larger changes, lest we think that science is always a matter of proceeding from basic, unchanging, given assumptions which are not themselves ever challenged. In looking at large, or fundamental,

scientific changes—such as the change from a theory of blending inheritance like that Darwin held to the particulate non-blending account that Mendel held, we will see several things:

- Fundamental changes tend to affect a broader range of problems and perspectives than less fundamental ones. Thus one needs to consider the effects of a fundamental change on a large variety of other issues and the explanation of many different kinds of phenomena. (This is why good science fiction does not freely suppose deep modifications in the basic laws of nature—such deep modifications are likely to have multiple far-reaching consequences which are not easily anticipated by either the author or the reader.)
- It is usually the case that older theories that we now hold to be incorrect had a lot going for them—we should expect that they will provide plausible (if incorrect) explanations for many phenomena, and basically correct explanations for some others. Thus, in his excellent discussion, John Moore⁵⁷ lays out 10 classes of phenomena that Darwin thought he had to explain, and how his theory does so. Indeed, Darwin's theory does quite well, in its own terms, at explaining these phenomena—many of which are legitimate and still interesting regularities involving complex multilocus interactions in inheritance. We will see in this lab that some of the qualitative effects of blending inheritance look surprisingly like what we observe, and can be distinguished from the effects of Mendelian inheritance only with careful and well-specified controls that don't become obvious unless and until one has already accepted Mendelism. It is often not a trivial matter to ferret out just how these older theories are wrong—and in doing so to figure out just what questions to ask (the important questions usually change with the theories) and how to answer them as unambiguously as possible.
- One of the main features of fundamental changes of this sort is that fairly direct observational facts which are widely taken for granted come to be rejected. Unfortunately for Darwin, some of the "facts" he thought he had to explain (like the inheritance of acquired characters), though widely believed, were not facts at all, so didn't need any explanation. Some other "non-facts" (like his belief that the heredity of large variations or "sports" was fundamentally different from that of small (or "individual" variations) prevented him and his advocates (as well as the initial rediscoverers and exponents of Mendelism) from seeing the applicability of Mendelian ideas to his theory. As a result, it took some 20 to 30 years after the rediscovery of Mendelism to give a unified (and basically modern) account of Mendelian genetics and evolution.
- Questions concerning how factors not included in the model can interact with features of the model, and with other unincluded factors, are

⁵⁷ *Heredity and Development*, Oxford, 1972, pp. 7-17. The example of blending inheritance and the reasons for its plausibility are also discussed in Jim Stewart and John Jungck's introductory module for BioQUEST.

often posed and answered much more simply (and basically correctly) by using an incorrect or oversimplified model. Thus, for example, near the end of the exercises (below), we will consider how selection, finite population size, assortative mating, migration, and overlapping generations would act in a model of evolution with blending inheritance. The effects they have are all in the same direction as the effects they would have in a Mendelian population, but it is easier to see these effects in a blending inheritance model without the combinatorial complexities introduced by Mendelian inheritance and segregation.

- Modern scientists often do not have a correct understanding of the rejected older theory, or of why it was rejected--largely because they never had to deal with it as a serious competitor, and the scanty history they got of the past controversy was a biased story told by the winners. This has at least two bad consequences: (1) In not understanding the problems that led to the rejection of the prior theory and the acceptance of the newer one as they appeared at the time, they fail to appreciate some of the strongest possible evidence for the current theory. (2) If they never had to deal with the older theory as a serious competitor, and the reasons why it seemed plausible, they may as a result fall into very similar traps in their current research—committing errors similar to those that fueled the last controversy. Thus, it is generally recognized (as part of the folk wisdom of evolutionary genetics) that blending inheritance makes evolution virtually impossible.⁵⁸ It is, however, also generally (and incorrectly) believed that particulate Mendelian inheritance rules out blending inheritance. This is not so. It is not just the particulate or continuous nature of the hereditary material which makes a theory a blending theory. Rather, blending is a joint product of (1) mechanisms through which an uncontrolled and statistically irreversible mixing of the hereditary material occurs, and (2) the size of the hereditary contribution of each unit.⁵⁹ Discussions of blending inheritance tend to focus on the second, although the first is in many contexts even more important, and not at all affected by the transition between blending and Mendelian theories. An appreciation of this fact yields a new understanding of why the total variance of previously isolated subpopulations decreases under random mating (Wahlund's principle), of why the Hardy-Weinberg equilibrium is an equilibrium, of the evolutionary significance of population structure, and of other things as well.⁶⁰

⁵⁸That is, since R. A. Fisher, *The Genetical Theory of Natural Selection*, Cambridge, 1930, chapter 1.

⁵⁹The smaller the per unit contribution, and the larger the number of hereditary units affecting the given trait of the individual, the smaller is the equilibrium variance with uncontrolled mixing, and the more the conditions of the blending inheritance model are approximately met. As the last exercise shows, even the paradigm case of blending, mixing beakers of colored water, is technically a case of particulate inheritance.

⁶⁰It is a maximally mixed state, with a minimum of variance for hereditary units of that "size"—in Darwin's model, there was no minimum size, and thus the equilibrium variance is 0. This same insight also explains why Mendelian polygenic traits respond more slowly to selection than single factor traits when both are "additive". A failure to appreciate the deep analogies between random mating or

The Genotype-Phenotype Distinction

Before developing a model of blending inheritance, in order to better understand why we accept Mendelian inheritance today, it is necessary to make a detour and comment on the relation between genotype and phenotype, both for standard Mendelian inheritance and for our imaginary “blending inheritance” model. For Mendelian inheritance, there is a natural distinction between genotype (the partial or complete genetic specification of the organism) and phenotype (the partial or complete specification of the characters or traits of the organism). But, as we will see below, there is no such natural distinction between genotype and phenotype for blending inheritance.

Genotype and Phenotype

Although technically, genotype and phenotype refer to complete specifications of the genetic (combinatorial) and phenotypic (potentially selective) properties of the organism, they are commonly used to refer also to partial specifications. Thus, e.g., geneticists speak of the *AA*, *Aa*, and *aa* genotypes, referring to the genetic complement, the 2 alleles, at the *A* locus, even though it is taken for granted that the genotype includes many other loci. Similarly, they talk about the white eye vs red eye phenotypes in the fruit fly, *Drosophila*, although the flies' phenotypes have many other biochemical, morphological, physiological, and behavioral properties. This has the advantage that while two complete genotypes or phenotypes are rarely identical, with less demanding partial specifications of one to a few traits or loci, a number of individual organisms may fall under each of the relevant types. This is an essential feature for statistical studies and for the reproducibility of results, though problems can sometimes arise when the phenomena under study depend upon variations in the unstudied traits or loci.

In what follows, where symbols are used to refer to genotypic or phenotypic states, genotypes or their partial specifications will be given in italics (e.g., *AA*), whereas the phenotypic characters they produce will be written in bold when referred to as a character (e.g., **A**, for the dominant character and **a** for the recessive character produced by this genotype). With Mendelian dominance, both the *AA* and *Aa* genotypes will correspond to the same phenotype, **A**. This forces a recognition of the distinction between the two concepts, since phenotypically identical crosses (e.g., *AA* X *AA* vs *Aa* X *Aa* or *Aa* X *Aa*), all appearing as **A** X **A**, will produce different results in the offspring phenotypes because of the different composition of the genotypes. If the same phenotypes yield different outcomes, there must be differences elsewhere, by the

"panmixia" at the group level and blending inheritance at the individual level lead a whole generation of our best population geneticists to build modeling assumptions equivalent to assuming an unusually strong (even stronger than Darwin's!) form of blending inheritance at the group level in their models of group selection. They erroneously concluded that group selection was virtually impossible--not too surprising given that their simplifying assumptions essentially simplified the groups out of existence! (For extensive discussion of this "migrant pool" assumption, as well as others and their significance, see Michael J. Wade, 1978, and W. C. Wimsatt, 1980b).

principle "same cause, same effect"—or, more exactly, its *modus tollens* form, different outcomes require different causes.⁶¹

To illustrate the point that there is no natural distinction between genotype and phenotype in this blending inheritance experiment, consider the following model as an interpretation of the process: Start with an imaginary indefinitely large population of beakers, each filled with the same volume of water. In some of the beakers (a proportion, p , of the total beaker population) a red dye is added producing a "red" phenotype and the others (in proportion $q = 1-p$) are left clear, or "white." We can imagine that "mating" two beakers consists of mixing the liquid from the two parent beakers and then redistributing the mixed liquid equally among the two beakers, which are now called "offspring." The parents in effect "die" when the offspring are born—a useful simplification to avoid the complexities introduced by overlapping generations, and which also guarantees both that there is no selection (no differential reproduction or viability) and that the population stays the same size. In this experiment, there is no dominance, since the genotype-phenotype distinction of a mating between red and white beakers (or, more generally of any mating) is a shade intermediate between the parental shades—in this case an intermediate pink. To put it in another way, this would (in at least one respect) be a kind of special case of what quantitative geneticists call an "additive trait," since the phenotypic trait (beaker color) is a linear or additive function of the number of genetic units (dye particles) contained in the beaker. Thus in this imaginary blending inheritance experiment, the mixtures of dye particles constitute both the genotype and the phenotypes of the beakers in this population.

In Mendelian genetics, dominance and intermediacy (or additivity) represent only two special kinds of cases of the relation between genotype and phenotype. The same genotype can also make different phenotypes in different environments. This effect is known by a variety of different names in different kinds of cases, including:

(1) (when the response is adaptive), (a) "environmental plasticity" or (b) "phenotypic switching" (the latter when quite different phenotypic responses are keyed by different environmental inputs early in development), and

(2) (when the response is not adaptive) (a) "teratogenesis" (when an environmental input, commonly a chemical stimulus, leads a "normal"—this term is misleading!—genotype to produce a phenotype commonly produced by a specific mutant genotype) or (b) "deprivation effect" (when the absence of a factor normally present in the environment leads to the loss of a normally present phenotypic capacity).

⁶¹In Darwin's time, no distinction was made between genotype and phenotype, so the way of formulating genetic problems discussed in this section could not arise. Indeed, the development of a genotype-phenotype distinction was an important part of the conceptual development of classical Mendelian genetics. (Historically, this distinction required first the denial of the inheritance of acquired characters, which Darwin and most of his contemporaries accepted, and second, the distinction between germ-line and somatic cells and its consequences which came to be known, after its promulgator, as Weismannism.) Indeed, the name for the modern theory of evolution ("neo-Darwinism") which has come to mean the fusion of Darwin's theory with the populational consequences of Mendelian genetics, was first coined by G. J. Romanes as a term to describe Weismann's theory of evolution in the 1890's—before the rediscovery of Mendelism.

There is no general exhaustive classification for kinds of environmental effects, since they are as varied and as complex as the variety of environments and the variety of phenotypes.

Furthermore, the same phenotype can be produced by different genotypes in the same environment ("genetic canalization"), or by the same genotype across different environments ("environmental canalization"). These effects are much more common than were supposed by geneticists. (They must be, since particularly in sexually reproducing species, recombination is constantly producing new genotypes (offspring) which are raised in new environments but which nonetheless resemble their parents in varied but readily recognizable ways.) They are often ignored in introductory genetics courses, being put off for more advanced developmental genetics courses on grounds that they involve a variety of often unknown interactions among a multiplicity of genes. This is unfortunate, because they are very important. These kinds of stability of phenotypic traits are in fact required by evolution, which demands significant heritability of fitness for systematic evolutionary change. Thus, in Mendelian inheritance, there are many different possible relationships between genotype and phenotype, whereas in blending inheritance there is only the relationship of intermediacy.

Suppose we (temporarily) ignore the effects of the environment (which plays a crucial role in producing the phenotype), or control for its effects, by raising each of our test organisms in the same environment or developmental sequence of environments, then each genotype should correspond to a phenotype, but as illustrated in Figure 4-5 there can be mappings of arbitrary complexity between genotypic and phenotypic states.⁶² In Figure 4-1, we have depicted three mappings between genotype and phenotype for the simplest case of just one-locus with two alternative alleles:

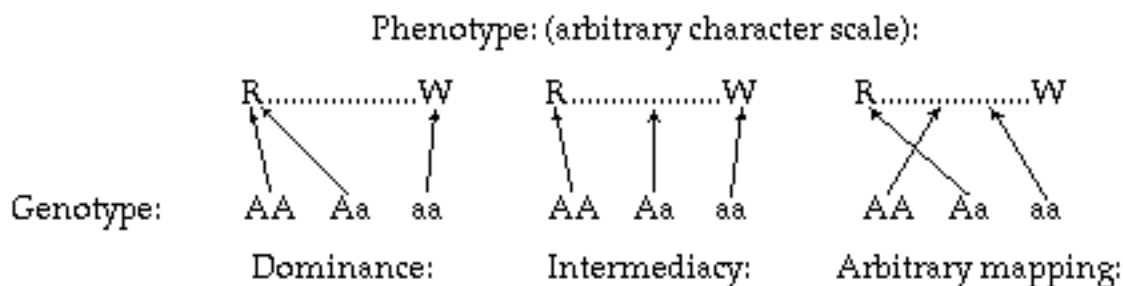


Figure 4-1. Genotype-phenotype mapping for 2 alleles at 1 locus.

The first two mappings we have already described: dominance and intermediacy. The third mapping labeled "arbitrary mapping" is but one of many possible alternative mappings. In this case it is called "overdominance," because the heterozygote has a more extreme value of the phenotypic property than the dominant heterozygote. There are, however, no general names for the types of arbitrary

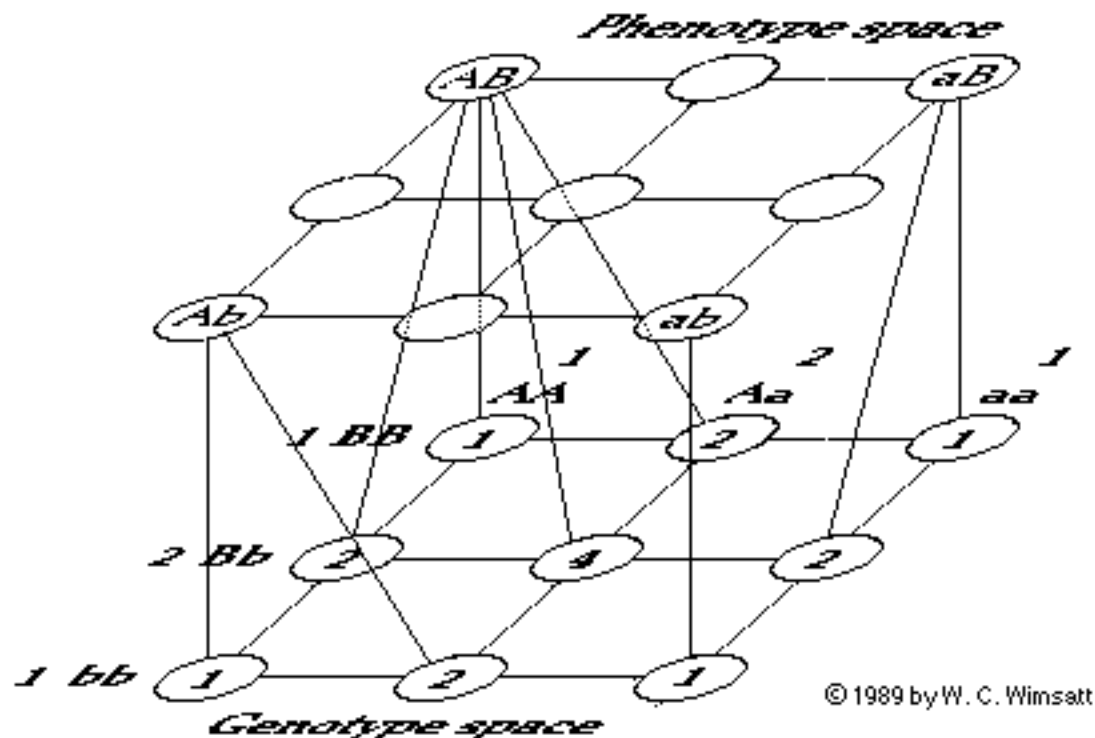
⁶²The terms "map" and "mapping" as used here have nothing directly to do with the common genetic use of that term to denote the location of an allele or locus in a linkage map of a chromosome. It is here used to denote the transformation through which a genotype (in the appropriate environment) produces a phenotype. This use is similar to that of Lewontin, (1974), chapter 1.

mapping possible in multilocus cases, except that if the genotypic value at one locus has an effect on the phenotypic mappings for another locus (because of interactions between gene products) this is called *epistasis*.

We can also graphically represent the possible mapping relationships between genotype and phenotype for the case where there are two alternative alleles at each of two loci. For the 2-locus case, we will consider 4 out of a very large number of the possible mappings: dominance at each of 2 loci, 2 kinds of intermediacy (one of which is analogous to what we find in the blending case), and an arbitrary mapping.

Consider first the classical Mendelian case of 2 factors, A or a and B or b , with the capital letter denoting the dominant gene. The results of selfing a double heterozygote (or "dihybrid") genotype ($AaBb$) with independent assortment for the two factors and Mendelian dominance for A and B are diagrammed in Figure 4-2 in a format we will use for the next 4 cases.⁶³ In Figure 4-2, we have represented a *phenotypic space* (the top plane) and a *genotypic space* (the bottom plane). The lines going from circles in phenotype space to circles in genotype space represent a particular mapping from genotype space to phenotype space and vice versa. These correlations are the product of the interactions of genetic and environmental factors in development, but we will not talk about these causal interactions here. For example, the circle labeled "4" in genotype space is the genotype $AaBb$ which is mapped into the circle in phenotypes space labeled **AB**. The mapping in Figure 4-2 indicates that phenotypically, A is dominant to a and B is dominant to b .

⁶³This form of diagram is suggested by figure 1, chapter 1 of Lewontin (1974), where the life cycle as modeled in evolutionary genetics is represented as a series of mappings back and forth between genotype space and phenotype space, represented as a pair of planes. The diagram here differs from Lewontin by including qualitative and quantitative information implied by Mendelian genetics.



Cross with two Mendelian dominant factors showing mappings between genotypes and phenotypes to produce 9 : 3 : 3 : 1 ratios of observed phenotypes.

Figure 4-2. Mappings in genotype-phenotype space for 2 Mendelian traits.

We can also extract quantitative information from this kind of mapping about both the ratio of genotypes and phenotypes of a large randomly mating population of organisms. First, there are four possible results of allowing a population of AA and aa individuals to mate at random: AA , Aa , aA , and aa . But since Aa and aA are genetically the same type, we have a 1:2:1 ratio of the genotypes AA , Aa , and aa in a monohybrid cross (similarly for B). These ratios are shown above the A genotypes and to the left of the B genotypes in the lower or "genotype" space of Figure 4-2. Phenotypically, however, since A is dominant over a , AA and Aa are both phenotypically observed as A . This is represented graphically in Figure 4-2 by lines connecting genotypes AA and Aa with phenotype A . This indicates that the phenotypic traits A and a occur in a ratio of 3 : 1 for the dominant and recessive phenotypes. This holds for the B locus as well.

We can use Figure 4-2 and the 1-locus ratios to generate the genotypic and phenotypic ratios for 2-alleles at 2-loci. The genotype corresponding to a circle in the genotype plane is obtained by combining the relevant A -locus genotype (in columns) and B -locus genotype (in rows) for that circle. Thus, the last circle to the right in the bottom middle row is in the aa column and the Bb row, giving a genotype of $aaBb$. The ratio of each different genotype is simply the product of the row and column ratios for that genotype's circle. Doing this calculation for each genotype yields, in normal reading order (left to right and top to bottom in the lower plane), the ratios: 1: 2:1;

2:4:2; 1:2:1 as illustrated by the number in each circle of genotype space. The phenotypic ratios of **AB**, **AB**, **aB**, and **ab** are easily obtained by summing the number of lines leading from genotypic space to a type in phenotypic space in Figure 4-2. For example, there are four lines leading from different genotypes (i.e., *AABB*, *AABb*, *AaBB*, and *AaBb*) space into the phenotype **AB**. The sum of these lines is $1 + 2 + 2 + 4 = 9$. If we do this for each phenotype in phenotype space we get 9: 3:3:1 as the ratio of the phenotypes **AB**, **AB**, **aB**, and **ab**. In short, an account of the production of a ratio of phenotypes produced by a mating requires two things: *first*, an account of the ratios of the genotypes produced by a given cross, and *second*, a mapping from genotype space to phenotype space, describing for each genotype which phenotype it produces. From these, we can derive the expected phenotypic ratios. This figure is designed to visually separate these two parts of the problem.

Keeping Track of Genotypes and Phenotypes

We need to keep track separately of both genotypes and phenotypes in a problem such as this, so the diagram represents the 9 possible genotypes in the lower plane. The maximum number of phenotypes these can produce is 9, and is represented by the circles in the upper plane. This is ignoring interactions with the external environment, the somatic environment, and the environment composed of the rest of the genotype (Williams, 1966) or keeping these constant. Variations in any of these can increase the number of distinct phenotypes produced by 9 genotypes. The number 9 also assumes that each 2-locus genotype has 2 allelic variants, and each different 2-locus genotype produces a distinct phenotype.

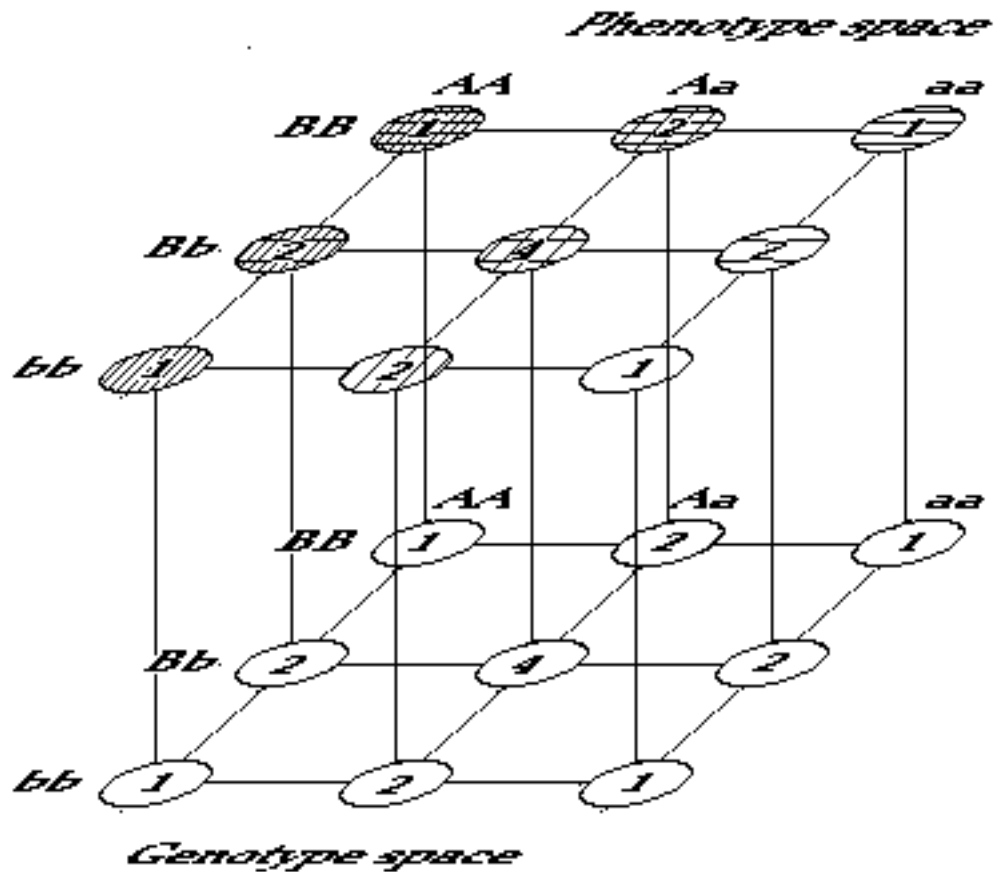
The phenotypic ratio 9:3:3:1 is often described as the ratio for 2 Mendelian factors with independent assortment, but there are in fact two separate independence assumptions required to produce this ratio of phenotypes. One is statistical independence of the factors in the production of gametes, such as would be the case if the factors are found on different chromosomes; this is why this is described as independent assortment. Not usually noticed is another assumption—that there is no epistasis, or interaction between the genotypes at different loci in determining which phenotypes are produced. This is a condition not on assortment, but on the nature of developmental interactions which affect the structure of genotype-phenotype mappings. Neither independent assortment (the first independence assumption) nor the absence of genetic interaction in the production of phenotypic characters (the second independence assumption) has any necessary connection with the other. Either can be met or not independently of the other. The second kind of independence is not entailed by Mendelian dominance, though it is by the stronger condition that the dominance relation is independent of the genetic (and environmental) contexts in which the locus finds itself. We will see a case below where this assumption is violated. (Some of these interactions with the genetic, somatic, and environmental contexts, particularly the somatic and environmental ones, are termed *penetrance*, rather than epistasis, though the distinction is somewhat arbitrary.)

The next case we will consider violates Mendelian dominance, though Mendel himself noted a case like this in the inheritance of flower color in peas.⁶⁴

Historical Note on Dominance and Mendel's Experimental Design

Mendel himself never spoke of dominance as a law or law-like relation. The perception that he did was promulgated by William Bateson. (Perhaps he should be called "Mendel's bulldog"—he surely played a similar role in the propagation of Mendel's ideas, as Huxley--"Darwin's bulldog"--did for Darwin). Bateson was an ardent advocate of Mendelism after Mendel's results were rediscovered and made public in 1900. He spoke of dominance as Mendel's "third law" (with segregation in the F₂ and independent assortment being the first and second). The evidence is very strong (both from a close reading of his original paper, and from other sources—see Robert Olby's excellent book, *The Origins of Mendelism*) that Mendel picked character pairs that showed a dominance-recessive relation in their inheritance rather than intermediacy (or any other pattern) so that the phenotypes he observed would be more readily and unambiguously classified without error. (With two readily distinguishable phenotypes, rather than three, with the third intermediate between the first two, there would have been much less chance for misclassification due to environmentally or genotypically induced overlapping phenotypic distributions.) Mendel is obviously worried about this possibility, as he explicitly talks about the importance of selecting characters whose expression is relatively insensitive to variations in environmental conditions. The role of dominance for Mendel thus seems to have been as an essential part of his experimental design, rather than as a claim that inheritance patterns were generally like that in nature, and he never claimed that all or even most characters showed the dominance-recessive pattern of the characters he studied. Mendel says that he chose 7 out of 22 traits of his peas as suitable for his experiments. We don't know what patterns of inheritance these other traits would have shown, but it seems likely that many of them would have been polygenic, epistatic in their effects, shown substantial genotype-environment effects, closely linked, or had other characteristics which would have made them much harder to analyze. Many of the more complex interactions (e.g., epistatic ones) were likely eliminated in Mendel's search for characters which were relatively insensitive to environmental effects—a remarkable unanticipated benefit of his experimental design. Darwin argued (only partly incorrectly) that the characters important for evolution did not show a dominant/recessive pattern, but showed blending or intermediacy, and would have been happier with one of the following two kinds of intermediacy described below. Ironically, Bateson's formulation of the "law of dominance" was taken as essential to Mendelism, and its easy and early refutation in the first decade of the 20th century was given as a reason to reject or to modify the theory—something which would not have been necessary for Mendel's original formulation.

⁶⁴Mendel, (1866) in the Harvard reprint of Mendel's paper.



Intermediacy in each of two Mendelian factors. Notice that in this case, each genotype produces a unique phenotype.

Figure 4-3. Intermediacy mapped in genotype-phenotype space.

In the first kind of intermediacy to be described here, each heterozygote (Aa and Bb) has a phenotypic character state exactly halfway in between the character states found for the two homozygotes at that locus (AA and aa , and BB and bb , respectively), but the two loci code for different phenotypic characters—e.g. (no longer following Mendel), plant height, and flower color. (Suppose that in this case, $AA \rightarrow 6$ ft., $Aa \rightarrow 4$ ft., and $aa \rightarrow 2$ ft.; and $BB \rightarrow$ red, $Bb \rightarrow$ pink, and $bb \rightarrow$ white.) In this kind of case, the 3 genotypes at each locus are distinguishable, and there are 9 different phenotypic states produced by the 9 different possible genotypes. This is a special case, in that each genotype codes for or maps to a unique phenotype, as is suggested by the different cross-hatchings for the 9 circles in the phenotypic plane of Figure 4-3.

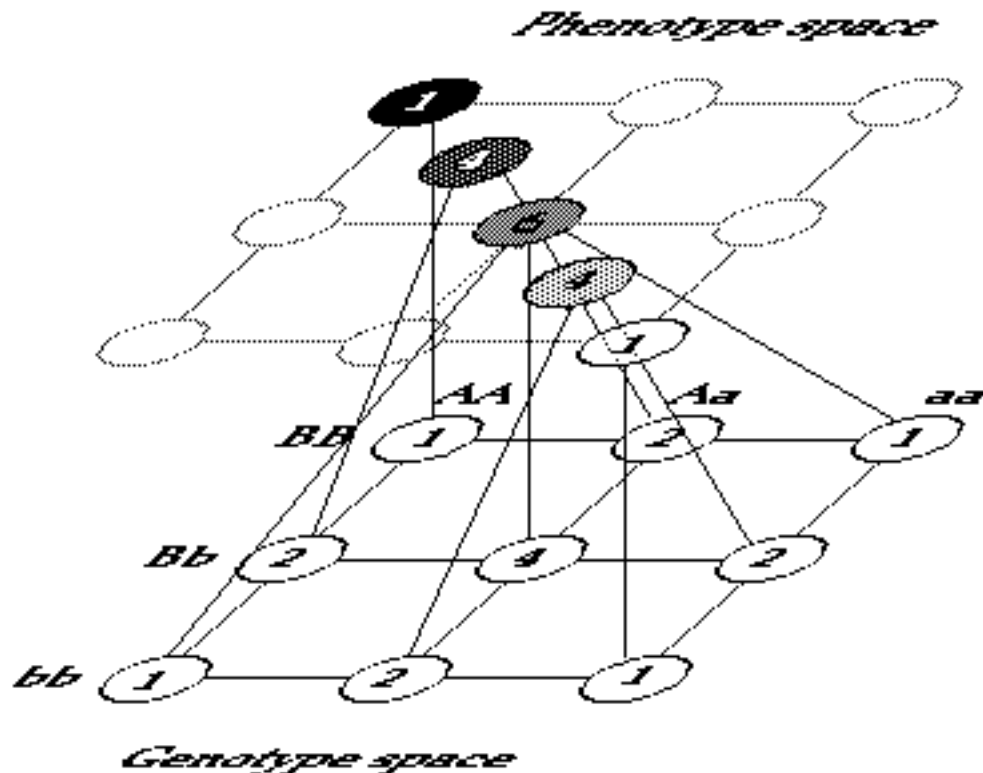
At times in the past history of genetics, this kind of 1-1 mapping between genotype and phenotype has been supposed to be a common situation. (See Lewontin, 1974, chapter 2.) Following the development of gel electrophoresis as a technique for discovering (enzyme polymorphism) variability in apparently (macroscopically) phenotypically homogeneous natural populations (see Hubby and

Lewontin, 1966, and literally hundreds of studies since), this view can no longer be accepted—at least for macroscopically characterized phenotypes. Whether such enzyme polymorphisms have subtle fitness effects in the same or in different environments and thus need to be counted as different phenotypes, or whether they are actually or approximately adaptively neutral is currently unknown in all but a very few cases.⁶⁵

In the second kind of intermediacy we will consider, the alleles at both loci contribute to the formation of the same character, producing a mapping like that illustrated in Figure 4-4. This is a particularly important case for us, because with many factors or genes contributing equally to the same character it is possible to explain the Mendelian inheritance of a quantitative trait.⁶⁶ This is a common situation where the rate of production of an enzyme, affected by the number of genes coding for that enzyme, affects the magnitude or intensity of the phenotypic character. In the case diagrammed below, there are 5 character states, where, in effect, the number of capital letter genes in the genotype determines the phenotype produced. (With 2 alleles at 2 loci there can be from 0 (in the case of *aabb*) to 4 (for *AABB*) capital letters.) This case is also interesting for its consequences for the ratios or distribution of phenotypes produced. In this case the ratio is 1:4:6:4:1; these are also the binomial coefficients for the binomial expansion $(p + q)^4$, which means, as explained below, that this is the expected proportion of types found in the second generation under blending inheritance where the two parental types (red and white) start out in equal proportions in the population. In Figure 4-4, the resulting phenotypic state is indicated by the shade in the phenotypic circles.

⁶⁵See Hartl (1988) chapter 1 for a summary review of findings.

⁶⁶See Hartl, 1988, chapter 4.

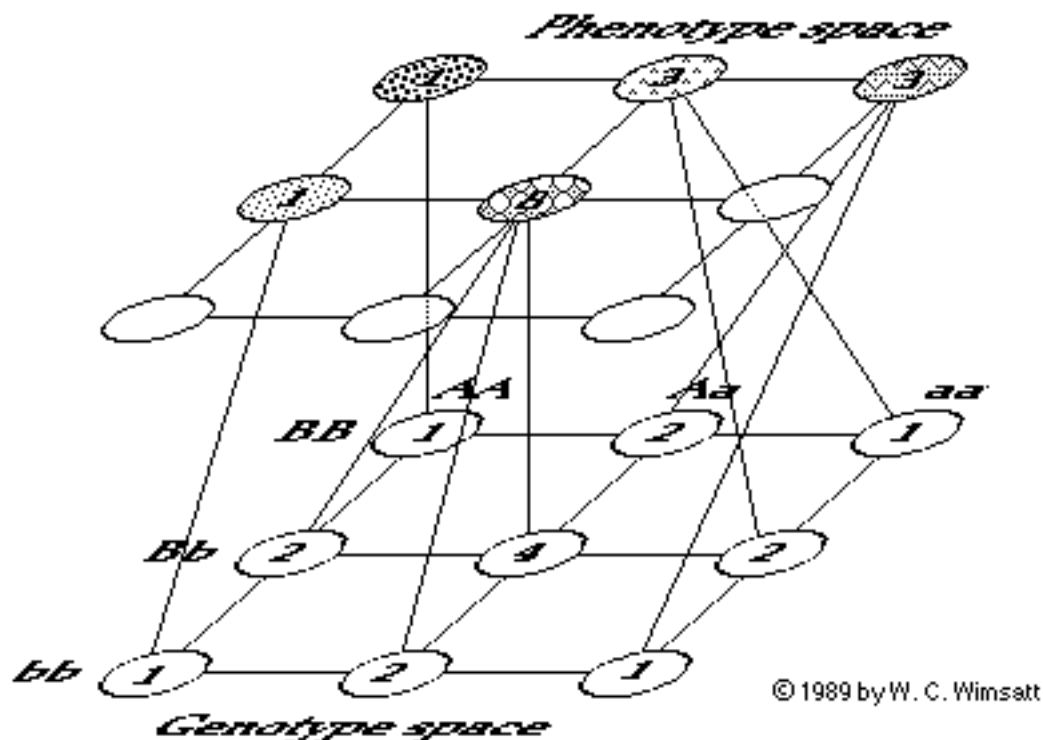


Genotype-Phenotype mapping where 2 loci contribute equally to a single trait, such as the corn kernel color studied by East. The mappings and the distribution of traits also fit the results of the second generation with blending inheritance, although in the Mendelian 2-factor case, the mating represents the proportions expected for a cross between two double heterozygotes (**AaBb**).

Figure 4-4. Mapping in genotype-phenotype space where 2 loci contribute equally to a single trait. Shade indicates magnitude or intensity of the trait.

Corn and Melanin Production in Humans

The situation depicted in Figure 4-4 is like that found for the inheritance of coat color in grains of wheat (where 2 loci are involved); in corn (a historically important case first studied by E. M. East, where it is now known that 3 loci are involved); and for melanin production in humans, where it affects both skin color and (the biologically relevant factor) resistance to skin damage from ultraviolet radiation in sunlight (where it is estimated that 5 loci are involved—see Jacquard, 1976). East's work was crucial to the ultimate reconciliation of Mendelian and Darwinian views by demonstrating that Mendelian inheritance was capable of explaining the behavior of so-called quantitative traits which blending inheritance had been invoked to explain. See Provine, 1971.



Arbitrary mapping between genotype and phenotype for 2 alleles at 2 loci. This mapping will produce an 8 : 3 : 3 : 1 : 1 ratio of phenotypes from a selfing of double heterozygotes. Notice that the ratios of any given phenotype is simply the sum of ratios for any genotypes which map to it. All 5 phenotypic patterns are intended to be different, as is indicated by the different fill patterns. No simple patterns of dominance and recessiveness are present, and there is significant epistasis.

Figure 4-5. Arbitrary mapping in genotype-phenotype space.

The last case we will discuss is perhaps (in theory, or on *a priori* grounds) the most general kind of case, where the mappings from genotype space to phenotype space are quite arbitrary. In Figure 4-5 (which does not correspond to any known case), 5 qualitatively distinguishable phenotypes (indicated by the different fill patterns in the phenotypic circles) are produced in the ratio 8 : 3 : 3 : 1 : 1. This kind of mapping is the most general in the mathematical sense in that there are no assumed constraints on the kinds of mappings between genotype and phenotype. This case shows a great deal of epistasis, or gene interaction, where the phenotype produced by the genetic specification at one locus depends upon or varies according to the genetic specification at the other locus. (It is qualitatively the most common kind of mapping one would get by randomly drawing arrows from genotypic to phenotypic circles.)

This does not mean that this is the most biologically general or common kind of mapping, since there may be biological constraints, of a biochemical, developmental, or evolutionary character (and these are not exhaustive) which makes certain kinds of mappings more likely or plausible in certain circumstances. Thus, for example, standard biochemical accounts of the action of enzymes which suppose that a

phenotypic trait results if more than a threshold amount of an enzyme is produced will generate a dominant-recessive pattern of inheritance (with **A** dominant over **a**) if the ratio of enzyme produced for the three genotypes *AA*, *Aa*, and *aa* is 2:1:0.⁶⁷

Similarly, Fisher argued that dominance will evolve and be selected for as a means of protecting against the expression of deleterious mutations (so that most new mutations will be recessive, and will not be expressed when rare). Wright and Fisher argued over the reasons for the evolution of dominance. The point here is not to decide who was right, but to point out that there may be evolutionary constraints or tendencies favoring different modes of gene expression. Lewontin's (1979) requirement of quasi-independence, together with the arguments of Charlesworth (1989) on the evolution of mimicry, could similarly be seen as an argument that long-range selection on the architecture of gene-expression favors some local but not too much global epistatic interaction, i.e., epistatic interactions among relatively small clusters of genes (particularly those contributing to the construction and fine tuning of particular adaptations) but not general epistasis involving interactions among a large proportion of genes in the genome.⁶⁸

Finally, various developmental constraints on evolution suggest that genes expressed earlier in development will tend to have larger effects, and mutations in them will be more likely to be deleterious, resulting in greater evolutionary conservatism for such traits. This is supported by a variety of experimental, observational, and theoretical considerations, and has a variety of consequences for the characteristics of evolutionary change and selective constraints on the architecture of genetic developmental programs (Arthur, 1984, 1987, 1988; Wimsatt 1986; Schank and Wimsatt, 1988; Wimsatt and Schank, 1988). It is too early, however, to say how or whether these considerations will translate in detail into the character of genotype-phenotype mappings of the sort discussed here.

⁶⁷For example, if *A* produces a functional enzyme and *a* does not) and the critical threshold amount of the enzyme is between 0 and 1. If the threshold is between 1 and 2, there will also be dominance, with **a** dominant over **A**, and if the phenotype is proportional to the amount of enzyme, there will be intermediacy. Thus various supposed modes of biochemical interaction can produce various inheritance patterns, including also various types of epistasis, which will not be discussed here.

⁶⁸See Kauffman (1991) for a general theoretical argument, and simulations to support this conclusion.

Fleeming Jenkin's Review of Darwin's *Origin of Species*

In 1867, a Scottish engineer, Fleeming Jenkin, published in the *North British Review* an extended discussion of the 5th edition of Darwin's *Origin*. Jenkin's review was, by Darwin's own admission, perhaps the most searching critical review his work received—and was most penetrating on two points: *First*, arguments from the best physics of the day applied to the rates of cooling of the sun and earth suggested that there were not billions, but only a few (estimates ran from 3 to 100) million years for the evolution of life on earth, before which the earth would have been too hot. These arguments were due to Jenkin's friend, Lord Kelvin (Sir William Thompson, with Maxwell, one of the two greatest English physicists of the 19th century), and seemed very compelling to Darwin's contemporaries. Only the subsequent discovery of radioactivity (in the 1890s) and nuclear fission and fusion gave time estimates consistent with what we know today. *Secondly*, Jenkin's pointed out that if one assumed, as Darwin did, that the offspring traits were the average of the parental traits, there would be no way for a superior mutant to pass on its superior trait undiluted. Furthermore, he argued that if a mutant were rare, even if it was much fitter, intense (viability) selection would still leave it more likely to perish without descendants—anticipating modern conceptions (due to Sewall Wright) on the role of random genetic drift in the loss of rare favored alleles in evolution. *Third*, he argued that with the superior mutant's advantages rapidly diluted out in its descendants, the outcome would be that the whole population would move but an infinitesimal amount in the favored direction. As he put it, the major fitness advantages of even a highly superior mutant, if rare, would be overwhelmed by the far greater numbers of average organisms. Each bit of supposed added superiority could be readily outweighed by adding just a few more average organisms, so the magnitude of a mutation had little moment for its ultimate effect on evolution.

This was (and still is) a telling criticism; 63 years later in 1930, Sir Ronald Fisher, one of the 3 founders (with Wright and Haldane) of the modern mathematically based population genetic theory of evolution, began his major work, *The Genetical Theory of Natural Selection*, with a consideration of blending inheritance. He argued that with it, a randomly mating bisexual population would lose half of its variation (literally: half of its variance) in each generation. But Fisher had just proved his "Fundamental Theorem of Natural Selection"—that the rate of evolution was proportional to the (additive) variance in fitness. This meant, in the absence of any source of new variation, that evolution would go as far in one generation as it would in all successive generations, in accordance with the series $1, 1/2, 1/4, 1/8, \dots$. Similarly, even if variation were being produced anew in each generation at a rate V , blending inheritance alone would prevent accumulation of more than twice that amount—limiting evolution to a relative snail's pace. (See Wimsatt, 1980b for further details.) Jenkin's critiques—basically that Darwin had far less time than he thought, and his favorite horse was lame to boot, lead Darwin to depend more and more on the inheritance of acquired characters to speed up the acquisition of new variation, so that evolution could proceed more rapidly—an erroneous move. It was not appreciated until after 1908 (and the formulation of the Hardy-Weinberg equilibrium principle) that Mendelism provided a way out—that at least for a 1 locus trait, under certain assumptions (see, e.g., Hartl, 1988, Ch. 1, pp. 23-28) the population went in one generation to an equilibrium level of variance, and did not lose any more thereafter due to a balance between Mendelian segregation and the formation of heterozygotes. Jenkin's mathematical sophistication in reasoning about heredity and evolution were unparalleled, except perhaps by Galton, in the 19th century. But perhaps this is not too much to expect from an engineer who played a significant role in laying the first transatlantic cable, helped to promulgate the acceptance of reforms which for the first time placed the teaching of basic science in the engineering curriculum, and first developed graphical methods for the solution of supply and demand curves in economics. Jenkin's review is reprinted in D. L. Hull, ed., *Darwin and His Critics*, Chicago: University of Chicago Press, 1987. For more on the age of the earth controversy, see J. Burchfield, *Kelvin and the Age of the Earth*, Chicago: University of Chicago Press, 1988.

Investigating the Consequences of Blending Inheritance

In this section, we will use the binomial distribution to construct a model of blending inheritance. The mode of presentation in this module is much more structured than with your investigations of the logistic and other growth equations, or the standard investigations characteristic of other BioQUEST modules, or supported by the "3P's" philosophy. Investigating a fundamentally false model like this one involves perhaps a 4th P—*exploring Possibilities*—which is ultimately usually in service either of the 1st P (problem-posing and formulation, because it helps to suggest useful questions to ask), or the 3rd P—persuasion. (Persuasion in science almost always involves a comparative judgment of the merits of alternative competing accounts, all but one of which have to be at least partially wrong--and not too infrequently, *all* of the posited alternatives are wrong.) We must explore possibilities part because we are investigating a false theory—and as a result there are no given standard models, and no data in the literature to use as a check on our speculations. Since we cannot depend on the structure of the real world to guide our investigations, and we want to convince you of the value of investigating the behavior of a fundamentally incorrect model, you will have to indulge our conceptions of what we can get out of it.

There are no texts on model building which have exercises like this one, so follow our lead, and then perhaps this investigation can act as a paradigm for parallel investigations of your own. This investigation has the added benefit that it both aids in the understanding of a historically important—indeed centrally important dispute—and has important implications for current models of inheritance used in models of group selection, cultural evolution, and problem-solving in artificial intelligence.

This simulation experiment of blending inheritance and its writeup consists of three parts. In the first, you are asked to do some things which familiarize you with the ideas behind the experiment and with the binomial distribution. It is useful to do these before you do the simulations on the computer. In the second, you will do the simulations on the computer and record the results. This involves copying plots generated for different simulation experiments. In the third part you will be asked to analyze the data, comment on their significance for Darwin's theory, and discuss ways in which Darwin could have avoided or minimized the effects of the conclusions pointed to in Fleeming Jenkin's review.

Basic Description of the Model

To see what blending inheritance does to the amount and nature of variation in a population, imagine a physical analogue for a model population with "blending" inheritance: this is a population of indefinitely large size (so there will be no "sampling error" effects) which consists equally of beakers of clear (W) and red-colored (R) water. A mating is a thorough mixing of the contents of two beakers, which is then redistributed equally between them. Thus there is no mortality, differential reproduction, or change in population size--hence no selection. Since two beakers are

involved, mating is biparental, but since any two beakers can mate, we have to assume either that the sex of the mating individuals doesn't matter, or that there always happen to be enough beakers of both sexes that no beaker has trouble finding a beaker of the opposite sex. This implies that there are no excess males or females, so that everyone can mate. (It also implies that there is an even number of individuals in the population.) Matings are assumed to occur through random choice of beakers (thus, no assortative mating, i.e., "panmixia" is assumed) so that the frequency of any type of mating is simply the product of the frequencies of the two mating types involved.⁶⁹

Building a Binomial Distribution Model of Blending Inheritance

Consider the binomial expansion for $(p + q)^m$. When this formula is expanded and multiplied out we get:

$$\begin{aligned}
 (p + q)^m &= p^m + \\
 &mp^{m-1}q^1 + \\
 &[m(m-1)/(1 \cdot 2)]p^{m-2}q^2 + \\
 &[m(m-1)(m-2)/(1 \cdot 2 \cdot 3)] p^{m-3}q^3 + \\
 &\dots + \\
 &[m(m-1) \dots (m - (m-2))/(1 \cdot 2 \cdot 3 \cdot \dots \cdot (m-1))] p^{(m-(m-1))}q^{(m-1)} + \\
 &q^m.
 \end{aligned}$$

Note that, because of cancellations in the exponents (e.g., $m-(m-2) = 2$), and in the numerators and denominators, the next to last term is just $= mp^{m-1}q$, and is symmetric with the first term. Similarly, the next to last term is symmetric with the second term, and so on.

This expansion has three important characteristics.

- First, it has $m + 1$ terms.
- Second, the coefficients of p and q in these terms (called binomial coefficients or combination numbers) are denoted by $C(m, 0)$, $C(m, 1)$, ..., $C(m, k)$, ..., $C(m, m-1)$, $C(m, m)$. $C(m, k)$ is the number of different collections of k objects one can get from a collection of m objects (if one draws them k at a time and replaces them before the next draw).⁷⁰

⁶⁹If the population size were finite instead of indefinitely large, the product of the frequencies would give the expected frequency of that mating combination, rather than the actual frequency.

⁷⁰Note that for each draw of k objects, one is also leaving behind $m-k$ objects, so the only difference between draws of k objects and $m-k$ objects is in which are picked up and which are left behind. Thus there should be the same number of combinations for draws of k objects as for draws of $m-k$ objects, or in

- Third, and this is especially important for modeling blending inheritance, if we suppose that $p + q = 1$, then $(p + q)^m = 1^m = 1$.

In this special case of the binomial expression let us make the following assumptions: (a) Let p be the proportion of red beakers in the original population. (b) Let q be the proportion of white beakers in the population. (c) Since any beaker is either red or white—no beaker can be both (exclusivity) or neither (exhaustiveness)—then $p + q = 1$. (d) *Claim*: If we start with pure red and pure white beakers in any given proportions, the behavior of this population in a blending inheritance model is given by the binomial expression for $(p + q)^m$ where $m = 2^G$, where G is the generation number ($G = 0$ at the start). Then the following correspondences hold between the binomial distribution and the imaginary case of blending inheritance:

- There are $2^G + 1$ types in the G th generation.
- The powers of p and q in the terms give the relative proportions of R and W in that term. Thus the 4th term in the third generation contains the expression $p^5 q^3$, and corresponds to the type $R^5 W^3$.
- Evaluation of a term in the binomial expression gives the frequency of the corresponding type. Thus, for $p = q = 0.5$, the 4th term in the expansion for $G = 3$ has the form $C(8, 3)p^5 q^3$, and gives the frequency of type $R^5 W^3$ (having proportions .625 R and .375 W), which is 0.2187.
- The types in the G -th generation are spaced at intervals of $1/2^G$ of the total or are $100/2^G\%$ apart in composition.
- The sum of all of the proportions of all of the types for any generation is 1.

The binomial distribution for any given generation thus contains all the necessary information on the number, composition, and relative frequencies of types produced in that generation from a blending inheritance process with initial proportions p and q of R and W beakers. In successive generations, you are in effect squaring the distribution of types produced in the last generation. So starting with a distribution $(p+q)$ of types in the 0th generation, we get $(p+q)^2$ types in the 1st generation, $((p+q)^2)^2 = (p+q)^4$ types in the 2nd generation, $((p+q)^4)^2 = (p+q)^8$ types in the 3rd generation, and so on. This is why m in the G -th generation $= 2^G$.

As noted above, with this blending inheritance model, there is no natural distinction between genotype and phenotype, since both genotype and phenotype correspond to the amount of red-making material (or proportion of dye particles) in the beakers. In effect, we are choosing as our phenotypic property the visible

general, $C(m, k) = C(m, m-k)$. You can see this for the first and last terms, where the coefficients are 1 (since there is only one way of drawing no objects or all of the objects). You can also see this for the second and next to last terms if you cancel out like terms in the numerator and denominator of the coefficient for the next to last term.

genotype—the shade of the water in the beaker, which is a direct product of the density of genetic units (dye particles) in the mixture. In this experiment, there is no dominance, since the genotype-phenotype of a mating between red and white beakers (or, more generally of any mating) is a shade intermediate between the parental shades—an intermediate pink. To put it in another way, this would be a kind of special case of what quantitative geneticists call an "additive trait," since the phenotypic trait (beaker color) is a linear or additive function of the number of genetic units (dye particles) contained in the beaker.

Simulating Blending Inheritance with the Binomial Model

Doing the Blending Inheritance Simulation

In your writeups of exercises 1–4 below you should include all the graphs you generated using Blending Inheritance. Be sure to do these four exercises before proceeding on to the final section, since the exercises in the next section presuppose that you have done these exercises and read the material in the boxed texts.

The program you will use calculates the relevant values for the binomial distribution $(p + q)^m$. Open the program call Blending Inheritance as described in Appendix A. Double click the blending inheritance icon, go to the **Simulation** menu and select "Initialize & Plot." You will then see a dialog box that look like Figure 4-6. Character x is a variable in this simulation. In all simulation experiments the distribution of phenotypes resulting from the crosses of beakers is determined by setting character x to the red beaker type. Thus the frequency p of character x in the population is the frequency of red beaker types. And so the frequency of white beaker types is $1 - p = q$.

Figure 4-6. Parameter settings upon opening Blending Inheritance.

Change the last generation to 2 (i.e., in Figure 4-6 start should be set to 1 and stop to 2 when you are finished) as shown in Figure 4-7 and click on the OK button or hit the <Return> key.

The image shows a 'Parameters' dialog box with a title bar. Inside, there are several input fields and radio buttons. The first section is 'Frequency P of character x, in the Population ($0 \leq P \leq 1$):' with a text box containing '0.5'. Below this are two radio buttons: 'Binomial Distribution, trials:' with a text box containing '8', and 'Blending Inheritance, generations:' which is selected. Under the selected option, there are two text boxes: 'From (start ≥ 1)' containing '1' and 'to (stop ≤ 10)' containing '2'. At the bottom are three buttons: 'Plot Range...', 'OK' (which is highlighted with a thick border and a mouse cursor), and 'Cancel'.

Parameters	
Frequency P of character x, in the Population ($0 \leq P \leq 1$):	
	0.5
<input type="radio"/> Binomial Distribution, trials:	8
<input checked="" type="radio"/> Blending Inheritance, generations:	
From (start ≥ 1)	1 to (stop ≤ 10) 2
Plot Range...	OK Cancel

Figure 4-7. Parameter settings for two generations of blending inheritance with the starting frequency of character X (e.g., Red) set at 0.5.

Investigating Blending Inheritance with the Binomial Model: Exercises

Exercise 1: Type Distributions under blending inheritance for Different Initial Frequencies in Successive Generations

Do runs for $p = 0.5, 0.75, 0.9$, and 0.95 for generations 2 through 7. We assume that by now you know what will happen in the first generation; you may go further if you wish.⁷¹ You can resize the resulting graph, apply grid lines, include or remove lines connecting points, and change point size as described in Appendix A--the "How To" manual. Use Copy from the **Edit** menu to copy the graph so that it can be pasted into another program, or save it as a PICT directly if that better meets your needs, or you can print it out directly. See the "How To" manual for more details.⁷² Next, go again to the **Simulation** menu and select the menu option Variance. This will plot a distribution of the contribution of variance of each character type in the population relative to its frequency. Copy this graph as you did with the binomial graph. There are two other plotting options under the **Simulation** menu: SumOfVar and SumOfBinomial. Choose both of these options and save or copy and paste the resulting graphs to the scrapbook.

Variance and the Binomial Distribution

Recall that the *variance* of a population is a statistical measure of variation and is defined as the sum of the squared deviations from the mean of a character times its frequency. We have assumed that White from Pink to Red can be measure on a scale from 0 to 1, where 0 is White, Red is 1, and x , the degree of Pink, is some number between 0 and 1, depending on how much red is in it. $f(x)$ is the frequency of type x . Thus the variance plot shows the contribution to variance $f(x) \times (\mu - x)^2$, where μ is the mean, for each of the $2^G + 1$ values of x , from $x = 0$ to $x = 1$ in the binomial distribution for generation G . The SumOfVar option under the **Simulation** menu provides a cumulative total, starting at the leftmost class (which is all "red") of the variance contribution of different phenotypes in the population. The last plot point in the distribution is therefore the total variance in the population for generation G .

⁷¹Note that by looking at the grid after you do this run, the setting of the y-scale will be normalized such that the highest point plotted will be at the 10th division. Thus the spacing between y-scale intervals is 1/10th of the high point. This allows you to get out a more accurate reading of the value of that point (important for your work with the sum of variances). You may also want to set the x-increment to some decimal value for convenience.

⁷²For example, it can be copied and pasted into the scrapbook, into word processing programs such as MacWrite or Word, or graphics programs such as Super Paint. If you have a scrapbook in the apple (🍏) menu, then it will be convenient to paste your figures into the scrapbook for later use when you are writing up the results of these simulation experiments.

Exercise 2: Changes in Variance under Blending Inheritance

From these runs, can you give (or guess) the equation for the sum of variance as a function of frequencies p and q (which = $1-p$) and generation number, G ? Try out your guess for at least 2 other values in that interval. (If you do the values .4, .6, .7, .8, you should have with the preceding values more than enough for a guess.) Is your function symmetric? Should it be? Under what condition(s) is it equal to zero? Tabulate and graph your values (a) for $0 \leq p \leq 1$ in a given generation, and (b) for generations 1-10 for $p = .5$. (You can use the autoscale feature and the vertical grid readout to get accurate readings for sum of variance if you manipulate plotting options so that value is the maximum value plotted.) Show by substitution in your equation how you can get two of the values you got in the simulation from your equation. If you have a book on population genetics (e.g., Hartl, 1988, or Crow, 1986) look up the formula for the change in gene frequency (Δp or Δq) per generation with two alleles at one locus. What does that formula and the formula you derived for variance above have in common? What do you make of the similarity you find? Look up the formulas for the variance of the binomial and the normal distributions in a statistics textbook. How do you explain the similarities (and the differences) between the formula here and those formulas?

Exercise 3: Comparing Blending and Mendelian Inheritance in the First Generation

Demonstrate that the first mating (at random, with a pool of 50% R and 50%W beakers) will produce the offspring proportions: 25%RR:50%RW:25% WW. Hint: do a "Punnett Square" (see the box below) with the different possible parental mating types and their frequencies arranged along the top (for columns) and side (for rows). The offspring proportions can then be obtained by calculating the frequencies in each cell as the product of the appropriate row and column frequencies, and summing the frequencies of like types. Assume, in this and in subsequent generations, that a "type" is a given proportion of R and W water, so that RW, RRWW, RWRW, RRRRWWWW, . . . are all of the same type (R^1W^1), RRRW, RWRR, RRRWRRRW, RRRRRRW, are all of the same type (R^3W^1), and so on.

Do these offspring proportions differ in the first generation from what Mendel would have obtained with his "particulate inheritance" theory for a 1 factor cross if R were a dominant (and W a recessive) allele at the same locus?

What if he were looking at a case which showed perfect intermediacy of the heterozygote rather than dominance, i.e., if RR = red, RW = pink, and WW = clear or white?

What would have resulted on Mendel's account in the 1st generation from a two-factor model, starting with mating types RRRR and WWWW, and with perfect intermediacy so that RRWW = RW = pink? Assume in this case and below that both factors code for flower color, and that his law of independent assortment holds.

Optional: What would he have obtained in the 2nd generation: (a) with a random mating population composed of all of the 1st generation offspring, and (b) with 2 parents who were both double heterozygotes (RWRW). Do you know enough to answer this question? See the following 2 pages.

Note on Using Punnett Squares:

1 allele at 1 locus

Gamete Types: male

pA

qa

pA	p ² AA	pqAa
qa	pqAa	q ² aa

2 allele at 1 locus

Gamete Types: male

pAB

qAb

raB

sab

pAB	p ² AABB	pqAABb	prAaBB	psAaBb
qAb	qpAABb	q ² Aabb	qrAaBb	qsAabb
raB	rpAaBB	rqAaBb	r ² aaBB	rsaaBb
sab	spAaBb	sqAabb	sraaBb	s ² aabb

Figure 4-8. Two examples of Punnett Squares in Mendelian genetics.

In Figure 4-8, two examples of Punnett Squares are provided for 2 alleles at 1 locus and 2 alleles at 2 loci. In the 2 alleles at 1 locus case, we have two alleles, A and a , occurring with frequencies p and q , respectively. Columns represent the gametic contribution by males and rows represent gametic contribution by females. The genotypes of offspring in the next generation are determined by combining the possible gametic contributions of males and females which result in three genotypes: AA , Aa , and aa . The frequency of these genotypes is determined by multiplying the row frequency by the column frequency for each of the four cells and summing across identical genotypes. This results in genotype frequencies of p^2 , $2pq$, and q^2 , the Hardy-Weinberg equilibrium proportions for the AA , Aa , and aa genotypes. The ratio of these three genotypes when all gametes are equally frequent reduces to the familiar Mendelian 1:2:1 pattern as can be seen from the Punnett Square. Similar points can be made for the 2 alleles at 2 loci case in Figure 4-8. (One must be careful, however, for the 2 allele at 2 locus case: One can use Punnett Squares to determine genotypes from gametic types under random mating, but one cannot determine gametic frequencies from genotype frequencies without knowing the linkage relations of the alleles. Even the case of unlinked loci contains a provocative surprise, in that one does not go immediately to Hardy-Weinberg equilibrium for 2 loci in 1 generation, because of the constraints implied by haploid gametic and diploid genotypic packaging. See the discussions in Hartl (1988) or Crow (1988), and the "recombination worksheet" on the next page.) Also notice that the "Punnett Square" as a strategy for calculating the genotypes of offspring in the next generation and their frequency applies equally well to the case of blending inheritance, since a Punnett Square uses only gamete types and their frequencies. In its most general formulation, it is basically a computational tool for determining and calculating the frequencies of a large number of distinct possible compound events, each of which is composed of two simpler events. Taking the probability of the compound event to be the product of the probabilities of the simpler events amounts to assuming that their probabilities of occurrence are independent, and using the multiplicative rule for compound probabilities.

Recombination Worksheet: 2 alleles at 2 loci

gametic types:

	<u>A</u> <u>B</u>	<u>A</u> <u>b</u>	<u>a</u> <u>B</u>	<u>a</u> <u>b</u>
<u>A</u> <u>B</u>	<u>A</u> <u>B</u> <u>A</u> <u>B</u>	<u>A</u> <u>b</u> <u>A</u> <u>B</u>	<u>a</u> <u>B</u> <u>A</u> <u>B</u>	<u>a</u> <u>b</u> ⁺ <u>A</u> <u>B</u>
<u>A</u> <u>b</u>	<u>A</u> <u>B</u> <u>A</u> <u>b</u>	<u>A</u> <u>b</u> <u>A</u> <u>b</u>	<u>a</u> <u>B</u> ⁻ <u>A</u> <u>b</u>	<u>a</u> <u>b</u> <u>A</u> <u>b</u>
<u>a</u> <u>B</u>	<u>A</u> <u>B</u> <u>a</u> <u>B</u>	<u>A</u> <u>b</u> ⁻ <u>a</u> <u>B</u>	<u>a</u> <u>B</u> <u>a</u> <u>B</u>	<u>a</u> <u>b</u> <u>a</u> <u>B</u>
<u>a</u> <u>b</u>	<u>A</u> <u>B</u> ⁺ <u>a</u> <u>b</u>	<u>A</u> <u>b</u> <u>a</u> <u>b</u>	<u>a</u> <u>B</u> <u>a</u> <u>b</u>	<u>a</u> <u>b</u> <u>a</u> <u>b</u>

The 4 boxes along the reverse diagonal are the only cases in which new gametic combinations can be produced through recombination. For all other boxes, the recombination products are the same as the inputs to recombination. The signs in the boxes indicate the direction of change towards Hardy-Weinberg equilibrium produced by recombination events in these boxes, assuming that the population starts out composed completely of AABB and aabb genotypes. Where x and y are the frequency of the two gametes in the genotype, the frequency of recombinants from that genotype is xyr , where r is the recombination frequency. If $p = f(AB)$, $q = f(ab)$, $r = f(Ab)$, and $s = f(aB)$, then H-W equilibrium occurs when $pq = rs$.

Exercise 4: Comparing Blending and Mendelian Theories in the Second Generation

2. What will the 2nd generation produce with the blending inheritance model? Do a Punnett Square again, remembering that you now have to consider all possible matings among 3 types.

- a. List the types produced. How many are there?
- b. What are the frequencies of these types?
- c. What are the ratios of the types if $p = q$? (Let the rarest type be 1, with other types scaled proportionately.)
- d. Do these offspring types and proportions differ from what Mendel would have obtained with his "particulate inheritance" theory in the 2nd generation starting with 1 segregating factor? What are the differences?
- e. *Optional:* Could Mendel have captured the 2nd generation results if he had been allowed to construct a 2-factor model? What proportions of the mating types would he have to assume in the 1st generation to do so? Could this 2-factor model simultaneously explain both the 1st and the 2nd generation results for the same mating conditions assumed in the blending inheritance model? Why? (It will help you to think about this question if you do Punnett squares for the 2nd generation for both blending inheritance and for the 2 factor Mendelian cross and compare the results.)

Further Exercises and Analysis

After you have written up your answers to exercises 1–4 you should be well-prepared to discuss more general issues about the implications of blending inheritance for evolution and the role of idealized assumptions in biology. Therefore, you should be ready to discuss and analyze the following questions:

5. Describe what happens in successive generations. How is it the same and how does it differ for the two frequencies that you chose?
6. How do these results raise problems for Darwin's theory? Do they confirm Jenkins' criticism of blending inheritance? (For further work on this you might want to read Jenkins' review.)
7. This experiment involves a number of idealized assumptions which would not generally be true for populations in nature (even if they *did* have a blending mode of inheritance.) Here are 4 which would have been relevant in Darwin's day, and a 5th which would occur to modern ecologists or demographers. Consider whether relaxing or modifying these assumptions has any effect on the outcome, and particularly, whether Darwin could have used this to make at least a partial counter argument to Fleeming Jenkin.

- a. This model assumes an infinitely large population, so that the frequencies of the types are actual frequencies, rather than expected frequencies. What effects would sampling error due to a finite population size have on these results? You can describe these qualitatively, though this is a difficult question unless you know some probability or statistics.
- b. All of the effects in this model are due to blending inheritance. No provision is made for selection. If selection is incorporated into this model could it be made to oppose the effects of blending inheritance? For example, what would happen to the population in successive generations if RW were lethal or sterile? How does the outcome of this differ if in the two cases first, $p = q = .5$, and $p \neq q$? (Think about this carefully--there's more than immediately meets the eye.) This mimics the effects of an important mechanism of species isolation. Could selection be made to worsen the effects of blending inheritance? What if there is "stabilizing" selection with fitness highest at the mean?
- c. This model assumes that organisms (or beakers!) mate randomly--the so-called "panmixia" assumption. What if they did not do so? Suppose that there were a strong tendency for organisms to select organisms like themselves as mates (assortative mating)? Suppose that there were a strong tendency for organisms to select organisms unlike themselves as mates? Would either of these have helped Darwin? There are other processes in nature which don't depend upon mating preference but which have the same effect. Suppose that the species is subdivided into populations which are relatively homogeneous, and that the probability is much greater that organisms will mate with others in the same population than with "outsiders." Or suppose that the species has a large range (say from north to south), that one extreme variant is found at the north and the other is found at the south, and that individuals migrate freely, but that the 1-generation migration distance is much less than the species range (mating is assortative by distance).⁷³
- d. Suppose that new variants are produced through spontaneous mutation. What kinds of assumptions would you have to make about the mutations if they are to oppose the effects of blending inheritance? Are these assumptions plausible in the light of Jenkins' criticisms?
- e. What if organisms don't die when they produce offspring, but continue to survive to produce one more generation (a so-called "overlapping generations" model)? Consider two different cases: (You may limit your consideration to qualitative effects--detailed predictions are not necessary, though you can construct such a model if you like.) (i) suppose they stay mated for life, and thus produce the same genotypes or phenotypes in successive generations. (ii) suppose they mate at random anew each generation, with no bias toward mating with same-generation organisms. Assume that two offspring are produced by each mating in whatever generation. Once you've got this down, suppose they survive and mate for n generations? Suppose they are immortal?

⁷³This range difference in phenotype or genotype is called a "cline". Such differences are usually maintained by selection for locally optimal phenotypes, and opposed by random migration of individuals from regions where other phenotypes are favored.

(8) How good is the beaker model as a model of blending inheritance? Assume that there is about an Avogadro's number of molecules of water in each beaker (that's 6×10^{23} .) Suppose that there are a finite number of dye particles, say 10^{18} , in each red beaker. How many generations would be required to show that inheritance in the beakers is not truly blending but that the number of dye particles is truly quantized? Would this be a plausible experiment to do, i.e., do you think that practically one could get data in this way which would produce an unambiguous result? If not, why not? Consider (a) distinguishability of types, (b) No. of beakers required. Using the formula for variance you derived above, can you estimate the variance in this generation, (a) if $p = .5$; (b) if $p = .01$? (Express your results in terms of the formula rather than calculating them out.) Could you modify this experiment to use fewer beakers and tell if the number of states was much larger than the number of beakers?

(9) Suppose that one had a Mendelian genetic system for height determination involving 10 loci, each of which could have a "tall" or a "short" gene for each of its 2 alleles, and such that the height of the organism was a linear or additive function of the number of "tall" genes it had. Could you tell from the behavior of the system under various conditions that you could set up that this was a Mendelian or a blending system? How? (There is more than one way.) What would you have to assume about the effects of the environment on height in order to get a clear answer using this method?

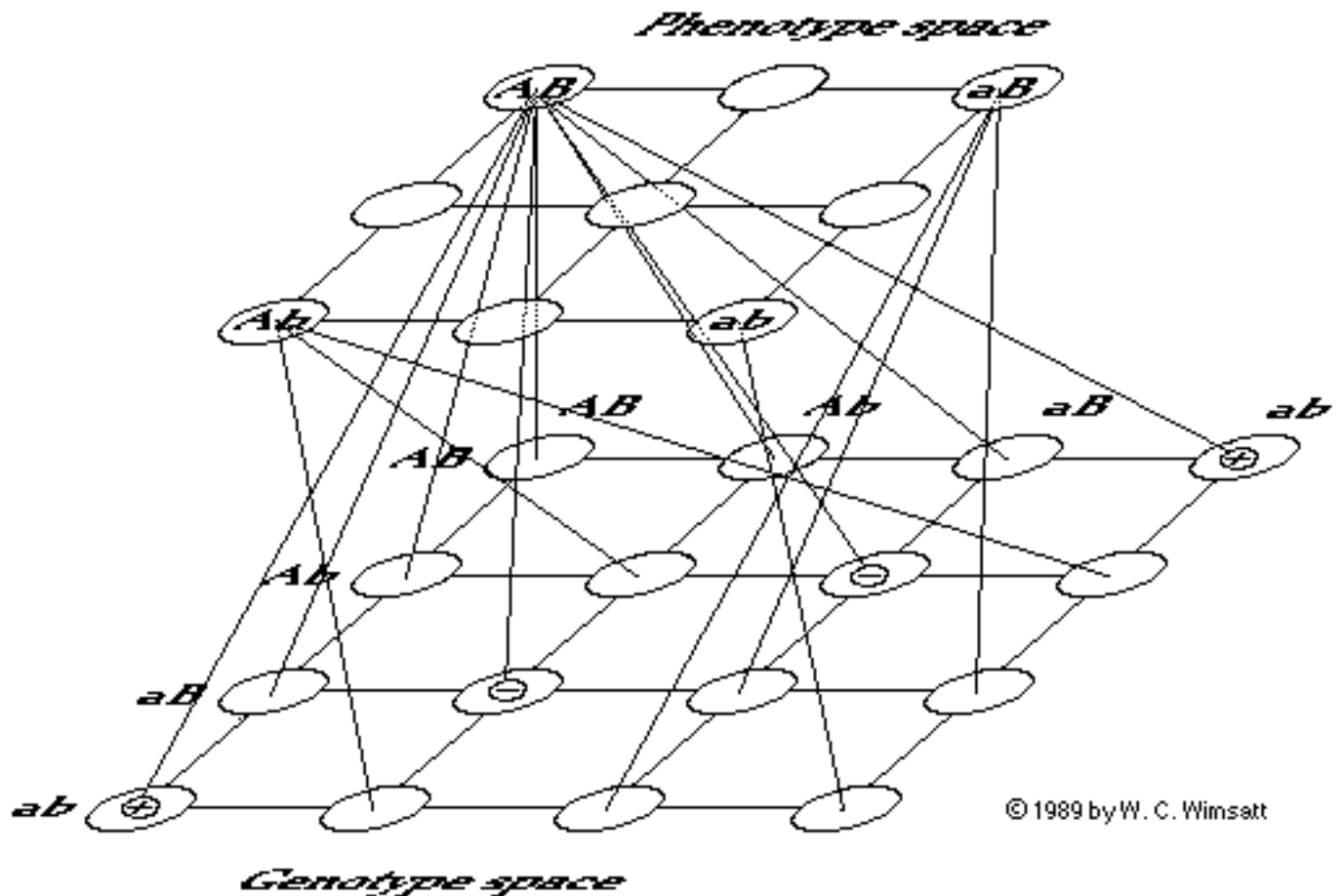
(10) Optional: If you know something about probability distributions, explain how you can get the normal distribution and Poisson distribution as limits of the binomial distribution. Demonstrate distributions which get successively closer to these limits using the program.

Projects Involving Current Disputes

(11) Discuss the "migrant pool" assumption and its effects on the efficacy of group selection in models which assume it. (See Wade (1978), and/or Wimsatt (1980b, 1981b) for relevant background.)

(12) A "new paradigm" for problem-solving in artificial intelligence is the so-called "genetic algorithms" or GA approach pioneered in the late 1960's by John Holland at the University of Michigan. In this approach, proposed problem solutions are supposed to be made up of partial solutions which are strung together in chromosome-like entities. Better solutions (chromosomes) leave more copies around than poorer ones, and chromosomes are picked from the population at random and allowed to cross over and recombine at random to generate new combinations of partial solutions. The mechanisms here are based upon those of classical Mendelian genetics and population genetics. (For an introduction to this field, see Goldberg, 1989.) A competing paradigm, the so-called "evolutionary strategies" (or ES) approach has arisen in Germany. (See the first paper (by Bäck, Hoffmeister, and Schwefel) in R. Belew and L. Booker, Eds., 1991, for an example of this approach.) The ES approach involves a form of blending inheritance. Its proponents claim that it converges to solutions much more rapidly than the GA approach, and is therefore more efficient. Can you explain why and in what circumstances this should be so? What kinds of solutions should it be good at finding? What kinds should it be bad at finding? The ES approach characteristically

involves much higher mutation rates than the GA approach. ES advocates have also shown a particular interest in assortative mating schemes. Can you explain why these might be so? If you can get your hands on the Belew and Booker volume, try to characterize as precisely as possible how the model of blending inheritance assumed in the ES approach is similar and how it is different from the model discussed here, and what consequences the differences could be expected to have.



Appendix A: How To Use Logistic Growth and Blending Inheritance

This appendix is a copy of the User's Manual for the Modeling Tools module. The Modeling Tools module contains the simulation programs Logistic Growth and Blending Inheritance and is intended to accompany the Modeling module. These applications can be found in the Modeling Tools folder inside the Collection Candidates folder on The BioQUEST Library CD.

Appendix A has been removed from this copy of Modeling to reduce the size of the document. It can be found in the Modeling Tools folder inside the Collection Candidates folder titled "Modeling Tools User Manual".

Appendix B: Functions of False Models as Means to Truer Theories

(with examples from biology)
W. C. Wimsatt 12-10-85 (revised: 6-26-91)

(1) An oversimplified model may act as a starting point in a series of models of increasing complexity and realism. (Note that these often form branching trees, rather than linear series in the structure of their assumptions in complex cases.) An important question here: how do we evaluate goodness of a starting point?

1. Population growth: exponential, logistic, MDLS, age structure, density-dep. age str., N-species Lotka-Volterra, etc.
2. Population genetics: deterministic models for 2 alleles at 1 locus, 3 alleles at 1 locus, inbreeding, frequency dependent selection, density-dependent selection, migration, 2A2L (linkage), stochastic models of same.
3. Classical genetics: single factor cross, with dominance, over- and under- dominance, co-dominance, partial penetrance, multiple factor crosses, sex-linkage, linkage, pleiotropy, epistasis, linkage mapping with multiple crossing over and interference.
4. Sequence of more realistic models in 19th century kinetic theory and statistical mechanics of gases.
5. Genetic nets (Kauffmann, Wimsatt, and Schank): a. directed graph models: constant fitness decrement models, both linear (K, W/S) and power functions (K), variable decrement models (to reflect generative entrenchment) (W/S), truncation selection models in which fitness decrements may sum to > 1 (W/S); b. Boolean automata models: (K, W/S).

(2a,b) A known (or believed?) incorrect but otherwise suggestive model may undercut the too ready acceptance of a preferred hypothesis by

- (a) suggesting new alternative lines for the exploration or explanation of the phenomena, or
- (b) providing an existence proof to show that models of a given class are capable of explaining a phenomenon which is held to be inexplicable by any models which fall into that or a broader class.

1. Use of "neutral models" in testing selectionist explanations.
2. Use of automata theory to show that any specifiable behavior (of finite length) can (in principle) be explained by a neural network, contra those who claimed that certain behaviors could never be explained mechanically (cf. McCulloch and Pitts (1943), and Kleene (1956)).
3. Kauffman's random networks and generic properties (but cf. also "robustness analysis" under (10) below).

(3) An incorrect model may suggest new predictive tests or new refinements of an established model, or highlight specific features of it as particularly important.

1. blending inheritance and preservation of variance under H-W equilibrium, and higher-level segregation analogues.
2. Goldschmidt's "pea-pod" model, Castle's "distributed chromosome" model.

(4) A model which is incomplete may be used as a template which captures larger or otherwise more obvious effects which can then be "factored out" to detect phenomena which would otherwise be masked or too small to be seen.

1.
 - a. Linear expectations vs HMF due to multiple crossovers,
 - b. HMF vs Muller data due to interference,
 - c. "mapping function" approach of early Bridges, Muller, Haldane vs "no general function" approach from third chromosome data of Bridges and Morgan (1923).
 2. H-W equilibrium expectations as reference state to detect and look for explanations of deviations from it.
- (5) A model which is incomplete may be used by comparison with the data as a template for estimating the magnitude of parameters which are not included in the model.
1. HMF vs Muller data to estimate interference distance.
 2. H-W equilibrium expectations as reference state to detect and look for explanations of deviations from it.
- (6) An oversimplified model may provide a simpler arena for answering questions about properties of more complex models which also appear in this simpler case, and answers derived here can sometimes be extended to cover the more complex models.
1. 2D arguments against Castle's 3D model.
 2. Analysis of chaos in single-species growth equations to predict (qualitatively) chaotic behavior in multispecies systems, and in more complex single-species systems (compare Logistic with MDLS).
 3. Roughgarden's use of 2A1L model with high mutation rates to explain qualitative features of 2A2L system.
- (7) An incorrect simpler model can be used as a counterfactual reference standard to evaluate causal claims about the effects of variables which are left out of them but included in more complete models, or in different competing models to determine how these models fare if these variables are left out.
1. Use of HMF-transformed data to calculate what Castle's map would look like in the absence of interference.
- (8) Two (or more) false models may be used to define the extremes of a continuum of cases in which the real case is presumed to lie, but for which the more realistic intermediate models are too complex to analyze or the information available is too incomplete to guide their construction or to determine a choice between them. In defining these extremes, the "limiting" models specify a property of which the real case is supposed to have an intermediate value.
1. Bracketing of linkage cases by Mendelian single-factor inheritance with pleiotropy or 2 linked factors without recombination on one side and Mendelian two-factor case with independent assortment on the other.
 2. Haldane's (1919) use of "infinitely flexible" (HMF) and "rigid" models of the chromosome to bracket the real case (interpretation 1).
 3. Note that Haldane's actual procedure (see Wimsatt, 1992) is actually more sophisticated, and involves use of three limiting cases to derive a parameterized scale-dependent causal meta-model -- a kind of inductive (or better adductive) generalization from the limiting cases (interpretation 2).
- (9) A false model may suggest the form of a phenomenological relationship between the variables (a specific mathematical functional relationship which gives a "best fit" to the data,

but which is not derived from an underlying mechanical model). This "phenomenological law" gives a way of describing the data, and (through interpolation or extrapolation) making new predictions, but also, if its form is conditioned by an underlying model, may suggest a related mechanical model which is capable of explaining the phenomena.

1. Virial equation of state
2. Logistic growth equation?
3. Interpretation of significance of b in Hassell's growth equation.
4. Haldane's (1919) predictive equation as decidedly non-causal (because it gives map distance as a function of recombination frequency), derived as a linear combination of non-causal forms of two limiting case causal models--so see also (8) above.

(10) A family of models of the same phenomenon, each one of which makes a variety of false assumptions, may be used in a variety of ways: (a) to look for results which are true in all of the models, and therefore presumably independent of the various specific assumptions which vary from model to model (Levins' (1966) "robust theorems") and thus are more likely trustworthy or "true"; (b) in the same manner to determine assumptions which are irrelevant to a given conclusion; and (c) where a result is true in some models and false in others, to determine which assumptions or conditions a given result depends upon. (See Levins (1966, 1968) and Wimsatt (1980a, 1981) for further discussion.)

1. Species tend to broader niches in uncertain environments (Levins, 1966).
2. Population growth is roughly exponential at low to intermediate densities.
3. Chaotic behavior results in non-linear systems with time lags when $r\Delta t$ is too large, and is produced more readily in larger, more complex systems (because of longer time lags).

(11) A model which is incorrect by being incomplete may serve as a limiting case to test the adequacy of new, more complex models. (If the model is correct under special conditions, even if these are seldom or never found in nature, it may nonetheless be an adequacy condition or desideratum of newer models that they reduce to it when appropriate limits are taken.) Note also that in complex models, unrealistic (often extreme) values of parameters may be used to test limiting behavior of models, or to understand how it works -- e.g., assumption of equal fitnesses or of no variability in a deterministic selection model should produce a Δq of 0.

1. Felsenstein's use of the HMF as a limiting case for more complex mapping functions.
2. H-W equilibrium limits for more complex selection models.
3. Ideal gas law as limit of Van der Waals equation of state.
4. Use of deterministic limit to test stochastic models of selection (this is a more general strategy for related stochastic and deterministic models).
5. Setting l 's and m 's in age structure model to values necessary to get synchronous non-overlapping generation model.

(12) Limiting case models may also play a role when a conservative underestimate or overestimate is desired. Thus Muller's (1916) procedure for estimating interference from distance between crossovers used as a conservative underestimate the shortest distance that could not be ruled out, rather than the most likely distance. This led him to underestimate interference (a good conservative procedure for demonstrating its existence to a skeptic) but probably also led to underestimates of recombination peak at modal distance.

(13) Use of an incorrect model under conditions where it does not deviate too strongly from the correct state of affairs (e.g., for the "right" parameter ranges, or aspects of the model which are not too sensitive to the incorrect aspects of the model) to calibrate apparatus, or provide correction factors for calculations involving better models.

1. Use of the HMF to estimate a minimum discriminable distance for deviations from linearity which must be subtracted from the real data in using Muller's data to estimate interference distance.

2. Use of incorrect linkage map + correction curves + new data to iterate to half- new correction curves, half-new map, 3/4-new correction curves and map, etc., as recommended in Bridges and Morgan (1923), in which an incorrect starting value (but one that is "in the ballpark") is used with an iterative procedure to converge on the correct or better answer. (This iteration is purely computational, and thus does not count as a case of type 1, above, and isn't used to calibrate or (by itself) to provide correction factors, so perhaps it deserves to be distinguished as a separate kind of case -- i.e., . . .

(14) Use of an incorrect model to "prime the pump" by providing a reasonable starting value for an iterative numerical computational procedure (one that may not converge, or may converge too slowly if starting with a value that is too far from the right one).

(15) Where optimization or adaptive design arguments are involved, an evaluation of systems or behaviors which are not found in nature, but which are conceivable alternatives to existing systems, can provide explanations for the features of those systems which are found, or for why alternative systems are not found. This may lead to the discovery of new selection forces. If the found alternative is not optimal, this may lead to the discovery of new constraints on the design problem.

1. Use of 3 or n-sex models to explain why there are (not more than) 2 sexes (Wimsatt, 1987).
2. Use of models of chaotic behavior to explain why it should be selected against, and is not found more widely in nature (Wimsatt, 1980a).
3. Use of blending inheritance models to explain adaptive advantages of Mendelian inheritance and higher-level segregation analogues (Wimsatt, 1980b, 1981b).
4. Use of Simon's "lock models" (1962) to explain why near-decomposeability and quasi-independence (the simple lock, rather than the complex one) should be found in phenotypic design (see also Lewontin, 1978).
5. Use of "symmetry breaking" argument to show that state of equal selection coefficients for all genes in genome is an unstable state, and that variable generative entrenchments in gene control networks is a doubly generic property (Wimsatt and Shank, 1988).

Appendix C: Reductionistic Problem-Solving Heuristics and their Biases

W. C. Wimsatt (11-15-84) rev. 7-9-91

I. Properties of Heuristics

(1) By comparison with truth-preserving algorithms or with other procedures for which they might be substituted, heuristics make no guarantees (or if substituted for a heuristic, weaker guarantees) that they will produce a solution or the correct solution to a problem. A truth-preserving algorithm correctly applied to true premises must produce a correct conclusion. But one may correctly apply a heuristic to correct input information without getting a correct output.

(2) By comparison with the procedures for which they may be substituted, heuristics are very "cost-effective" in terms of demands on memory, computation, or other limited resources. (This of course is why they are used.)

(3) The errors produced by using a heuristic are not random, but systematically biased. By this, I mean two things: (a) The heuristic will tend to break down in certain classes of cases and not in others, but not at random. Indeed, with an understanding of how the heuristic works, it should be possible to predict the conditions under which it will fail. (b) Where it is meaningful to speak of a direction of error, heuristics will tend to cause errors in a certain direction, which is again a function of the heuristic and of the kinds of problems to which it is applied.

(4) The application of a heuristic to a problem yields a transformation of the problem into a non-equivalent but intuitively related problem. This means that answers to the transformed problem may not be answers to the original problem, even though various cognitive biases operative in learning and science may lead us to ignore this.

(5) Heuristics are useful for something--they are purpose relative. Tools which are very useful for one purpose may be very bad for another (Levins, 1968). This often gives a useful way of identifying or predicting their biases: one would expect a tool to be relatively unbiased for the applications it was designed for, and perhaps quite biased for others. One might also expect that increases in performance in one area will be accompanied by decreases elsewhere.

II. The Origins of Bias in Reductionistic Heuristics

Reductionistic problem-solving strategies have systematic biases. How these biases arise can be easily seen from a general characterization of the problem-solving context of a reductionistic problem-solver:

(1) One starts by choosing, designating, or constructing a system for analysis. This immediately partitions the world of study into that system and its environment.

(2) Second, we must make Simon's (1957) "assumption of bounded rationality": That any real world system is too complex to study in all of its complexity, so we must make simplifications--through selection of properties or objects for study, simplified assumptions about relationships between these properties or objects, assumptions about what variables must be controlled or randomized, and the like.

(3) Third, I will assume a very general characterization of what it is to be a reductionist, merely that a reductionist is interested in understanding the character, properties, and behavior of the studied system in terms of the properties of its parts and their interrelations and interactions. (This is a sufficiently inclusive description that it probably captures any analytic methods in general, even that of many who would not call themselves reductionists. It should in any case be acceptable to any reductionist.)

From (3), the reductionist is primarily interested in the entities and relations internal to the system of study. But this fact, together with the assumption of bounded rationality, has an interesting consequence. While simplifications will in general have to be made everywhere, the focus of the reductionist will lead him to order his list of "economic" priorities so as to simplify first and more severely in his description, observation, control, modeling, and analysis of the environment than in the system he is studying. This systematic bias is characteristic of all reductionistic problem-solving strategies (see list from Wimsatt, 1980b, p. 235). Each one is used in some circumstances because its adoption transforms the initial problem into one which is easier to analyze and to solve. Each of them can be seen as an application of the general schema for making simplifications to a specific scientific activity.

A particularly crucial factor in explaining the inability to see the biasing effects of these strategies (whose effects are not only cumulative, but also "cover" for each other, since they all produce biases in the same direction) is what I have called "perceptual focus. (Wimsatt, 1980b, pp. 248-249.) This is the tendency to see all phenomena in terms of properties of entities at the focal level of interest. Thus, "We are nothing more than bags of genes" or "Groups are nothing more than collections of individuals" indicates lower level foci, while the "invisible hand" of free market economics indicates a higher one. The only way to eliminate level-relative biases is to change the boundary between system and environment by going to a higher or lower level. The same heuristics will then have different effects, and different-level models of the same system can be used to check each other.

III. Specific Reductionistic Biases

In the processes of reductionistic analysis, the standard procedure is to analyze a complex system into its parts, analyze the behavior of these parts in isolation, and then resynthesize these parts and the explanations of their behavior into a composite explanation of the behavior of the whole system. In doing so, a number of heuristic strategies are employed which carry with them systematic biases which lead to ignoring or downplaying context-sensitivity of results and the importance of the environment. These are:

A. Biases of Conceptualization

(1) Descriptive localization: describe a relational property as if it were monadic, or a lower order relational property. Thus, e.g., describe fitness as if it were a property of phenotypes or genes, ignoring the fact that it is a relation between organism and environment.

(2) Meaning reductionism: assume that lower-level redescrptions change the meanings of terms, but higher-level descriptions do not. Result: philosophers (who view themselves as concerned with meaning relations) are inclined to a reductionistic bias. (cf. also multidimensional context-sensitivity as an intermediate between positivist "objectivity" (context-independence of observations) and total meaning holism.

(3) Interface determinism: Assume that all that counts in analyzing the nature and behavior of a system is what comes or goes across the system-environment interface. This has two versions: (a) black-box behaviorism -- all that matters about a system is how it responds to given inputs; and (b) black-world perspectivalism -- all that matters about the environment is what comes in across system boundaries and how the environment responds to system outputs. Either can introduce reductionistic biases when conjoined with the assumption of "white box" analysis -- that the order of study is from a system with its input-output relations to its subsystems with theirs, and so on. The analysis of functional properties, in particular, is rendered incoherent and impossible by these assumptions.

(4) Entificational anchoring: Assume that all descriptions and processes are to be referred to a entities at a given level, which are particularly robust, salient, or whatever. This is the ontological equivalent of assuming that there is a single cause for a phenomenon. Thus the tendency to regard individual organisms as primary, and more important than entities at either higher or lower levels (or similarly for genes for some reductionist neo-Darwinians). cf. perceptual focus and multilevel reductionistic modeling.

B. Biases of Model-Building and Theory Construction

(4) Modeling localization: look for an intrasystematic mechanism to explain a systemic property rather than an intersystemic one. Corollary 4a: Structural properties are regarded as more important than functional ones.

(5) Contextual simplification: in reductionistic model-building, simplify environment before simplifying system. Thus the environment may be treated as constant (in space or in time), regular in some other way, or random. This strategy often legislates higher-level systems out of existence, or leaves no way of describing systemic phenomena appropriately.

(6) Generalization: When starting out to improve a simple model of the system in relation to its environment, focus on generalizing or elaborating the internal structure, at the cost of ignoring generalizations or elaborations of the existing structure. Corollary 6a: If a model doesn't work, it must be because of simplifications in description of internal structure, not because of simplified descriptions of external structure.

C. Biases of Observation and Experimental Design

(7) Observation: Reductionists will tend not to monitor environmental variables, and thus will often tend not to record data necessary to detect interactional or larger-scale patterns.

(8) Control: Reductionists will tend to keep environmental variables constant, and will thus tend to miss dependencies of system variables on them. (*Ceteris paribus* is regarded as a qualifier on environmental variables.) Mill's methods applied with this heuristic will yield as a systematic bias apparent independence of system variables from environmental variables, though the right experiments won't have been done to establish this.

(9) Testing: Test a theory only for local perturbations, or only under laboratory conditions, rather than testing it in natural environments, or doing appropriate robustness or sensitivity analyses to suggest what are important environmental variables or parameter ranges.

(10) Abstractive reification: Observe only those things that are common to all cases; don't record individuating circumstances. Losses: (1) sense of natural (or populational) variation; (2) lose detail necessary to explain variability in behavior, or exploitable in experimental design.

(12) Articulation-of-Parts (AP) coherence (Kauffman/Taylor/Schank): Assuming that studies done with parts studied under different (and often inconsistent) conditions are context-independent, and thus still valid when put together to give an explanation of the behavior of the whole. (Schank: Checking this gives a non-trivial use for computer simulation.)

(13) Behavioral regularity (Schank/Wimsatt): The search for systems whose behavior is relatively regular and controllable will result in selection of systems which may be uncharacteristically stable because they are relatively insensitive to environmental variations (Schank: regular 4-day cyclers among Sprague-Dawley rats are insensitive to con-specific pheromones; Mendel's selection of 7 out of 22 characters which are relatively constant and insensitive to the environment probably resulted in unconscious selection against epistatic traits.)

D. Functional Localization Fallacies

(14) Assuming that the function of a part is to produce whatever the system fails to do when that part is absent (e.g., Spark plugs as "sputter suppressors"), or generally, reifying added or subtracted behaviors of the system as functional properties of the manipulated unit.

(15) Assuming simple 1-1 mappings between recognizable parts and functions.

(16) Ignoring interventive effects and damage due to experimental manipulation.

(17) Mistaking lower-level functions for higher-level ends, or misidentifying system which is benefited.

(18) Imposition of incorrect set of functional categories (common in philosophy of psychology which ignores ecology and evolutionary biology).

E. Other Important Biases

(No.'s 10, 11, and these can generate either reductionistic or holistic biases in different contexts.)

(19) Perspectival blindness: Assuming that a system can be exhaustively described and explained from a given perspective because it has been very successfully and powerfully described. (Not all problems of biology are problems of genetics, molecular biology, physiology, or anatomy.)

(20) Tool-binding: Becoming sufficiently bound to a specific (usually very powerful) tool that one chooses problems for it, rather than conversely ("The right job for the organism," rather than "The right organism for the job!") This may be an efficient division of labor if mastery of the tool is very demanding -- it is problematic only if it facilitates errors No. 11 or 16.

Appendix D: Analyzing the Assumptions of the Logistic and MDLS Models

Analysis of Assumptions

An explicit list of the assumptions of a model or class of models can serve as a reminder of the dimensions and nature of the simplifications in the models, (a) to indicate the limitations of the models, (b) to provide a resource for directions to increase the realism of existing models or for generating new models, (c) where a given model is advanced as providing an explanation for a particular case, to provide a list of simplifications which can suggest other possible mechanisms to explain the result when the simplifications are relaxed. (d) Making changes in one respect often naturally suggests changes in others, so that in seeing a list of simplifications together, the whole may be more than the sum of the parts.

Thus, going from a non-overlapping-generations to an overlapping-generations (age structure) model conceptually opens up a much wider variety of mechanisms and functional forms for density dependence than are suggested by the logistic or MDLS models. Thus, how does density dependence affect birth rates (m 's) and death rates (l 's)? The question does not arise when they are not distinguished but lumped into a single r . Is the density dependence itself a function of age structure? (Different aged organisms may eat more or less, bring in more or less food, be more or less likely to share with selected others, or more or less likely to give it up in competitive interaction).

Seeing a list of assumptions may help at all stages with the "3P's." The elaboration of density-dependent interactions in the preceding paragraph could serve as the beginning of a classification of possible density and age-structure-dependent causal interactions among individuals in a population prior to constructing a model. Here it is helping to structure the space of possible interesting interactions which we seek to try to model. It is aiding in posing the problem or *Problem-posing*. In the explanation of the behavior of a particular interaction in nature—say, cooperativity in which more individuals actually produce young at a higher effective growth rate—it can give us the resources to construct alternative possible explanations, an exercise in *Problem-solving*. And in the last stage, that of *Persuasion*, when we are trying to validate and convince others of the explanation we have fixed upon, it provides a list of parameters we need to measure to test the sufficiency of our hypothesis, and a list of possible alternative causal mechanisms which need to be ruled out to demonstrate its necessity.

It is important to remember in all of these cases that these methods are heuristic, and that the judgments generated from them are conditional. In the real world, we never can be sure that we have got all of the assumptions, that we have figured out all of the alternatives that arise by relaxing them, or that we have assessed correctly the results of the assumptions we have made. But that is part of the seductiveness of science. It is the richest source of problems, and the source of the richest problems that we know.

Assumptions of the Logistic and MDLS Models

The assumptions here are somewhat arbitrarily divided into mathematical and biological assumptions. There is however usually a complex interaction among mathematical and biological assumptions and measurement problems. Testing a model involves both an adequate solution to these measurement problems, and a comparative analysis of other relevant, possibly competing, models, both of which are discussed after listing the assumptions.

Mathematical Assumptions

1. r is positive (note strange consequences if it is negative, even though this is OK for exponential growth). This is not required for the exponential growth model, but is for the logistic and MDLS models. Note that adding assumptions to a model may require changes in the assumptions that are already there. This problem disappears if we separate out birth and death rates and assume that the intrinsic birth rate is ≥ 0 .

2. All parameters can meaningfully take on real values--allowing, e.g., fractional population sizes.

Biological Assumptions

3. Non-overlapping generations with organisms which either die while giving birth or are immortal and do not age.

4. All organisms have the same fitnesses and other parameters (thus no variability, differential selection, or genetics; note, e.g., that small population size can produce inbreeding depression, and that selection may change the phenotype with effects on r , K , and C). Common implicit assumption used to justify this: gene frequency change is slow on the scale of population size changes.

5. All organisms have the same needs and consume the same resources, independent of age. (So K is compared just with N , rather with some weighting for different ages.) In the worst case, an organism's needs as well as the available resources vary with the numbers and age distribution, as might be true if there are cooperative effects; e.g., cooperative nesting may (1) reduce individual effort required for nest construction; (2) increase abilities for predator avoidance or defense; (3) reduce metabolic demand through reduced heat loss, etc.

6. r , K , C are constants over space, time, and the same for all individuals. (K could fluctuate seasonally, could be different for different activities (e.g., food, nest sites, territory sizes), could decrease over time as the environment is poisoned or depleted, or as other species change in their numbers--cf. e.g., succession.) C is not a function of sex ratio, or of other environmental variables which could change over time. Note also that there could be different C 's for different purposes--mating, predation, defense against predation, and the like. Similarly for r . Spatial heterogeneity of environments could also have a variety of effects.

7. Synchronous reproduction.

8. Linear dependence of density effects for both K and C.
9. Independence of all parameters on context (no other species affect them).
10. Reproduction is asexual.
11. Time lag for feedback of density-dependent effects is 1 generation (and is constant across different effects).
12. Competition for resources is of the "scramble" rather than of the "contest" variety.
13. No other species affect growth of the population in question in ways which depend on their densities, or their densities are constant.

Measurement Assumptions and Problems

Meanings of and values to put in for r , K , C are not always clear, both in testing a model and also in evaluating the significance of empirical studies:

1. r is measured in the laboratory under low-density conditions--not clearly a good estimate for r in nature. r has both viability and fecundity components, and measurements of either in nature have problems. (Talk about marker genes, and mark-recapture methods.)
2. K is Carrying Capacity, but a variety of different things can be limiting in nature (e.g., resources (and note that these fluctuate over time in a seasonal environment), nest sites, and territory sizes, all of which may vary both spatially and temporally.)
3. Similarly for C --e.g., probability of finding a mate for sexual species (and notice that this depends upon sex ratio), critical size needed for efficient feeding (e.g., wolf packs need critical minimum size to hunt large prey), or neutralization of toxins in environment (for certain bacteria). Also hidden genetic variables here--e.g., if population size is too small, get inbreeding depression, which acts to reduce r (over several generations).
4. The problem here is not just what factors to include in measuring r , K , and C , but also whether the assumed functional dependence of growth rates on these parameters should have the same functional form for all of these factors.

Comparative Evaluation of Assumptions--A Look at Other Models

1. Overlapping age-structure model.
2. Overlapping age-structure model with density-dependent effects on l 's and m 's.
3. Overlapping age structure model with above and contribution to K as an age-dependent phenomenon.

4. How about required balance among age-classes? e.g., without parents, young might all starve to death. (Suppose we also have a helping coefficient for different age classes? Also a dependency coefficient--e.g., which could be higher for the very young and the very old? Also, if there is more dependency need than helping capacity, how do we want to distribute helping effects--i.e., how does this affect l's and m's of various classes? Also, consider scramble vs contest competition for aid as well as food.)

5. Density-dependent selection model.

6. Model in which r , K , and C might be seasonally variable (e.g., sinusoidal or more complex?).

7. Stochastic fluctuations in r , K , and C ? (See Biota and PopDyn for models which allow this kind of variation.)

8. Hassell model for contest vs scramble competition. (Note that in other respects, the Hassell model is a curve-fitting model, with no clear interpretation for the variables in it.)

Topics for Discussion

1. Consider classification of above simplifying assumptions with respect to how easy they are to model--some are easily corrected for, others (e.g., spatial heterogeneity assumptions) are very complex or computationally demanding to simulate.

2. Look at explosion of difficulty in understanding interactions as the number of variables in the models grows, and also in finding good sets of values of the parameters to study (combinatorial explosion in sets of variables to investigate). Consider also the increase in number of parameters to be measured in experimental tests of the model in real situations. This suggests that it is better to deal with a family of simpler models including the effects of a few factors at a time rather than an all-encompassing model in which everything is variable.

3. Look at variables for (1) how easy they are to measure, (2) ambiguities in the definitions, (3) whether they are biologically meaningful, (4) whether they enter realistically in the equations.

4. Look at results of the models to see how sensitive they are to (1) specific parameter values, and (2) assumptions in the models. (If the natural phenomena are stable, but the model requires very special values to produce the phenomena, this is grounds for suspicion of the model.)

5. Evaluate models (both in general and in specific situations to which they are to be applied) to see how well they meet (Taylor's) conditions of quasi-independence and coherent aggregation.

6. Look at reductionistic research strategies and their biases for these models (Wimsatt, 1980b).

7. Look at strategies/functions of using false models (Wimsatt, 1987).
8. Identify sufficient parameters, and the variables that they aggregate. Identify conditions under which they may fail to be sufficient (Levins, 1966, 1968).
9. Look for robust theorems, models in which they fail, and conditions under which they may break down (Levins, 1966, 1968).

Appendix E: The Use of Graphics Options To Explore and Enhance Pattern-Detection and Analysis, and Examples of the Diagnosis of Causes from the Character of Patterns

These simulations have deliberately been designed with a lot of graphical options so that you can play with them to emphasize those aspects of the phenomena you wish to focus upon. In addition to the flexibility inherent in the 4 different modes of data analysis, you have additional degrees of flexibility available within each of them appropriate to that means of analysis. You can change the rate of presentation of points, speeding them up (within the limits of your computer) to see more holistic patterns, or slowing them down to follow particular changes. (In some cases it is possible to discern as many as three different levels of order in the patterns by changing their presentation rate relative to our "natural" flicker fusion and perceptual frequency filters.) The size and patterns of the spots can be varied, and connected with lines if you choose. Direct manipulation of graph coordinates allows for analysis of moire effects (enhanced by colors). Control of the grid scale facilitates not only measurement but also evaluation of periodic and nearly periodic patterns.

We will discuss some of our particular visual talents for the analysis of data, and then proceed to discuss examples which illustrate and use these talents, or involve particularly interesting diagnostic patterns of behavior.

One of the most important capabilities of our visual system is the processing of visual information about motion. There are at least three important components here: (1) motion detection, (2) the integration of similar, temporally ordered, and spatially contiguous stimuli into ordered motion of identifiable though possibly changing objects, and building on these reifications, (3) "hard wired" tendencies to hypothesize causal relationships--though not always correctly. All three of these condition our perception of patterns. (The fact that they may sometimes lead to incorrect conclusions--generating "illusions"--is not problematic if we can either be warned when errors are likely, or restrict their application, through careful analysis and design, to situations where they work. See the discussion of *heuristics* in Wimsatt, 1981 or 1984.) In the modeling simulation used to generate the pictures of chaotic behavior, motion detection is useful in detecting equilibrium, approach to equilibrium, divergence from unstable equilibrium, bifurcations, replication of arbitrary periodicities, roundoff error effects, and differential sensitivity of variables to parameter changes in the same and different parts of its range. (Motion can't be shown here, unless you are reading this manual on a computer, in which case the graphics will be redrawn in the order originally produced--effectively re-running the simulation before your eyes. Even so, you have no control over the presentation rates of the data unless you use the redraw option in the **Simulation** menu to change the plot delay.) Seen as static objects, these figures often show a confusion of points. But presented in temporal order, the trajectories of these same points are integrated as ordered transitions--demonstrating the second point above. Finally, all of the simulations with this program are designed to be run as families of curves in which the same system is run many times for systematically changing values of a key parameter. The visual impression of a causal relation between changes in the parameter value and in the behavior of the system is both overwhelming and usually justified.

Visual presentation is particularly appropriate when one needs to see or to track (where they are systematically changing) similarities and differences among complex entities which are changing in space, time, or--generalizing from these--in some subset of properties comprising the base components of an n -dimensional property space. These tasks make visual or spatial representation particularly appropriate:

- (1) Analysis of highly multidimensional data. (In effect, this means two or more dimensions.)
- (2) The need to see, or to find, in such data (a) similarities in diverse cases in spite of their differences, or (b) differences in spite of their similarities, or (c) to factor perceived global dissimilarities into localizable similarities and differences. These are all instances of template matching.
- (3) The need to look for relationships among patterns on several different size scales or levels of organization (Wimsatt, 1976; Tufte, 1983).
- (4) The need to combine or use jointly information from different perspectives on (i.e., using different ways of accessing or measuring) the same object (Wimsatt, 1974), or
- (5) To determine where there are boundaries between objects, or to identify whether there is a common object, property, or cause, behind apparently correlated patterns, i.e., to localize objects (Wimsatt, 1981).

The visual system has been selected for the efficient, reliable, and rapid solution of tasks like those of (2)-(5) above in a complex variegated natural environment. The discrimination, identification, and reidentification of objects, the identification of camouflaged predators, and the tracking of prey--all in the rapid service of feedback-controlled locomotion--require an on-board computer of enormous power that is older than and probably much more powerful than our specialized language capacity—which undoubtedly draws on it for processing in ways that we are just beginning to discover. Even the term "perspective" here, in (4), betrays the visual origin of this metaphor for the multidimensional integration of diverse kinds of information about the same object. The connections among realism, objectification, and multiple independent means of access to a common robust object--beginning with our three spatial senses, vision, hearing, and touch--go very deep (Wimsatt, 1974, 1976, 1981), and almost certainly represent an interleaved family of heuristics with a long evolutionary history. We need to learn how to draw on the enormous power of our visual system by representing data in ways that allow us to see the patterns in it effectively.

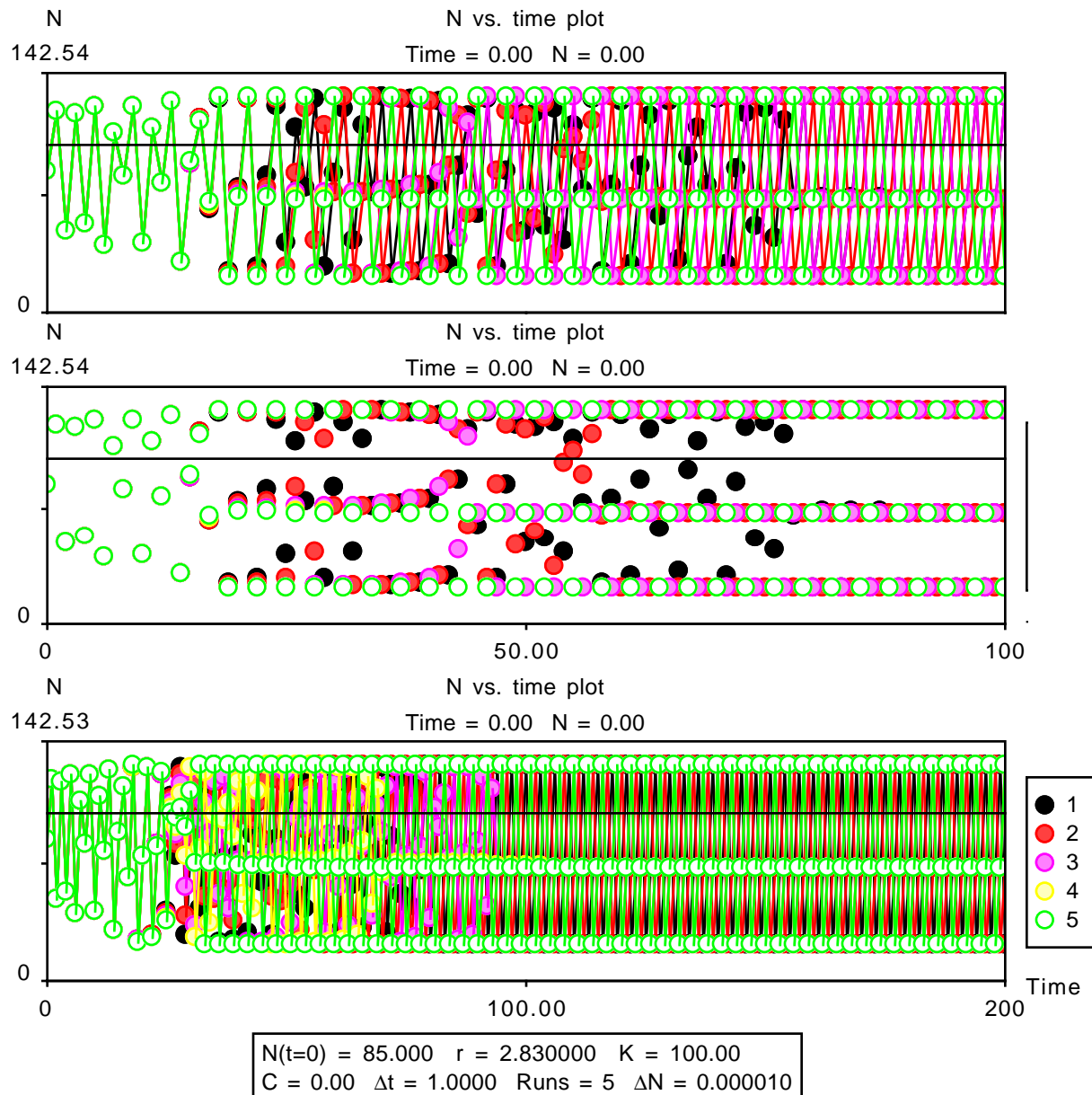
Many of the most complex cases of pattern detection require the comparative analysis of images, (1) looking for similarities and differences, (2) for transformations between images, or (3) for a common reference frame to use when integrating the information from different images. As noted by Tufte (1983), these cannot be done effectively unless the images are juxtaposed so that they can be seen together. This is plausibly due to two factors: (A) we have the highest sensitivity for comparative differences and similarities in local analysis. (B) More global comparisons require some way of transporting local information, but we aren't very good at doing this simultaneously for multiple dimensions of that information. So multi-dimensional comparisons have to be done locally, or one dimension at a time, using or supplemented by symbolic and analytical techniques. Thus, for example, a graph grid

allows local comparison with grid elements (providing a common reference frame) to give more global comparisons of the location of features, involving comparison of only one spatial component at a time. We can change from local to global comparisons via a scale change, preserving multidimensional information by making what was global local, but do so under pain of losing fine detail.

Let us see in more detail how some of these abilities work. In most cases, this will involve comparisons of information from a variety of closely related graphs. In some cases, the discussion will not illustrate, but presuppose the claims made in this section. I doubt, for example, whether the bifurcation plots or the effective population size plots below could be meaningfully compared unless they were presented together. When you do your investigations, try to keep some of these facts (and the examples we have discussed) in mind, and use them to your advantage in your own presentations and analyses.

Spurious Patterns Which Are Artifacts of Mode of Presentation

The next eight graphs all involve the same data set, but are presented differently to illustrate some of the graphics options and some perceptual and mechanical effects which can have a significant effect on how they appear. In the first two, the first 100 generations of 5 runs are plotted. In all the others the first 200 generations are shown. These are actually the same data used in the first Monte Carlo simulation exercise (see Figure 3-1)—depicting infinite sensitivity and divergence, followed by capture of the trajectories by the period-3 attractor. Here they are plotted as strip charts, and redundant chart labels are deleted (by resizing the images to cover them up) to get several of them on a page to facilitate direct visual comparison.



Figures E1 through E3. The same data at different densities and with/without lines.

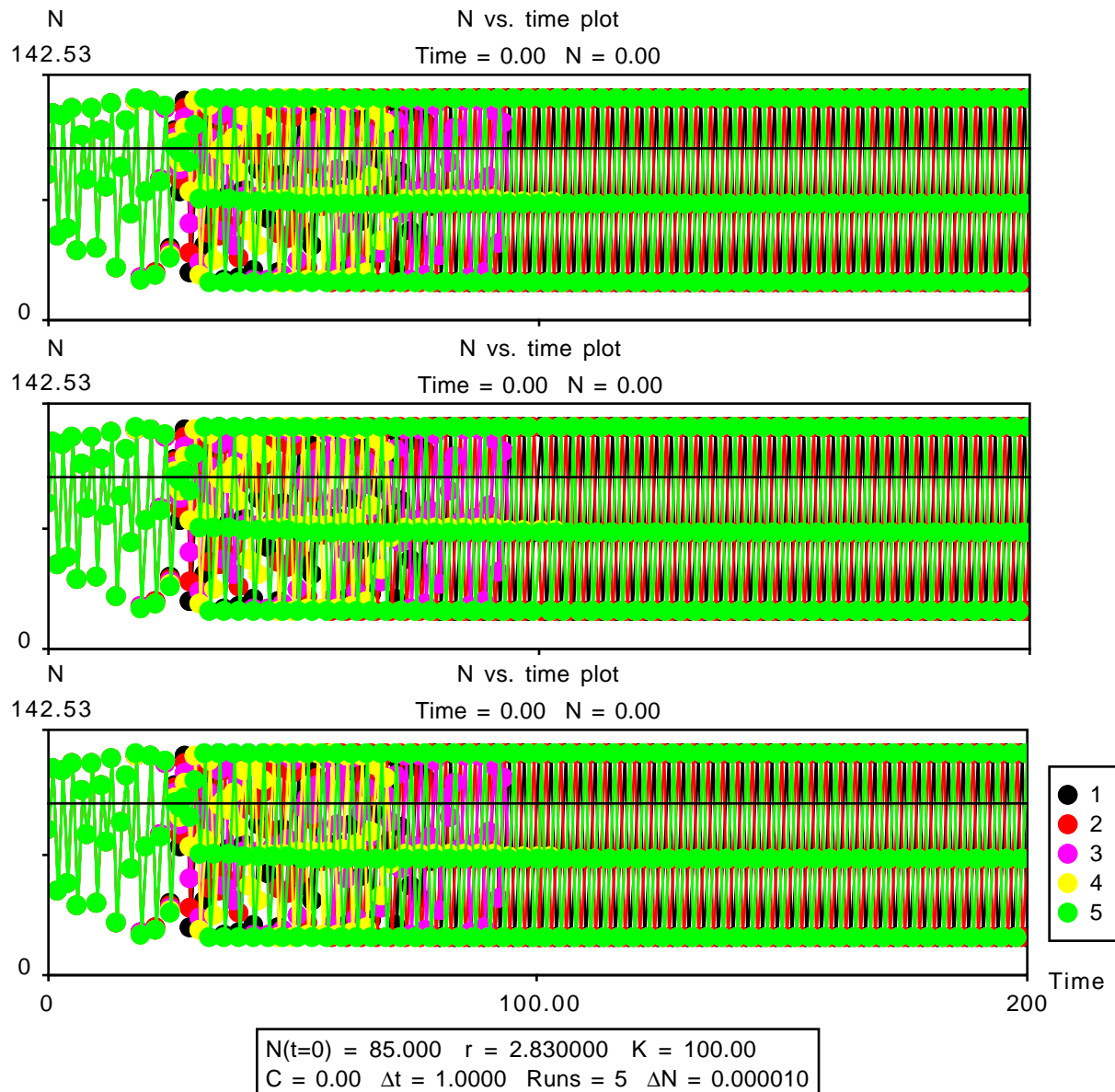
Comparison of the first two charts illustrates that lines are an ambiguous advantage. For the depiction of the gradual divergence of the trajectories in the first 20 generations, the lines act to cluster what points should go together and be compared. This is before the divergence of trajectories produces a thicket of lines. When the lines are deleted in the second chart, it is much easier to tell where various points are and to see how the divergence of the trajectories produces phase shifts (e.g., in trajectory No. 1, between generations 27 and 29). This can also be seen clearly in the partial overlap of the spots from the different trajectories in the 3-bands of the 3-point equilibrium. (Since trajectories cycle through the three points, returning to each one only every third generation, partial overlap indicates that the different trajectories are returning there in different generations, and are thus out of phase.)

In the 200 generation chart, viewed in black and white, you can no longer see that the trajectories are out of phase, since period-3 returns are plotted so close together that the last trajectory plotted covers over all of the preceding ones once they have all fallen into the stable 3-point cycle. (This could be partially compensated for by choosing smaller spot sizes.) If you are viewing this on a color monitor, you will still be able to see a sign of the different phase relations, however, because the different colored lines of the different runs show a period 3 repeat, like a woven fabric. The third chart also acts as a reference standard for the charts that follow, in that it indicates how these patterns look when 200 generations are plotted.

In all of the subsequent 200 generation plots--designed to illustrate moire effects (or other effects that look more like crystal dislocations) in the patterns produced by the lines, the point patterns have been changed to solid. The trajectories are still distinguishable by color if you are reading this manual on a computer with a color monitor, but black and white are just fine to illustrate the effects we will demonstrate now. These effects show how pattern can sometimes be used as a very sensitive indicator of dimensional magnitudes—in this case of the width of the graph.

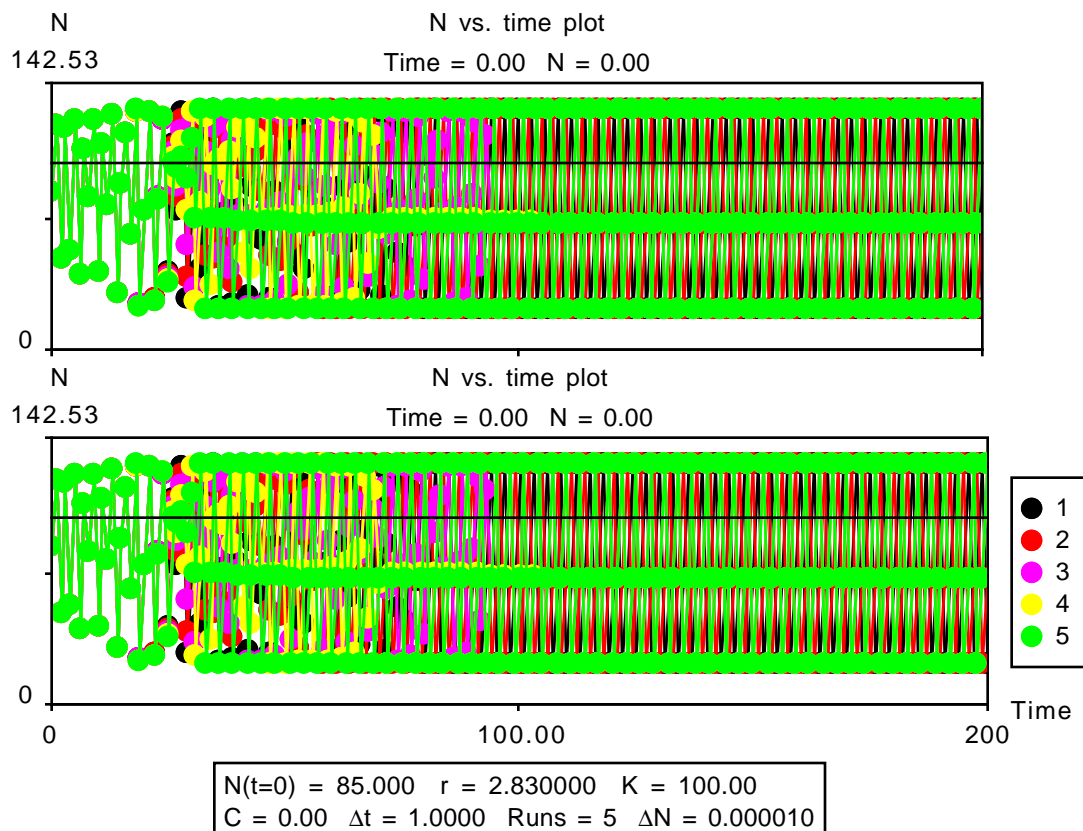
In the first of the graphs below, the length of the graph, including both end bars, is 401 pixels (left and right graph coordinates of 25 and 426, respectively). Depicting generations 0 through 200 involves plotting 201 points. With 401 pixels, this means that the point for each generation is plotted 2 pixels to the right of its predecessor. As a result, after all of the trajectories have settled into their places in the 3-point cycle, a very regular pattern ensues. The second graph is one pixel longer (402 pixels) and the third is one pixel shorter (400 pixels.) In these cases, the graphics software tries to place the points at equally spaced intervals, but it can't because it has to plot at an integral number of pixels. Roundoff error means that the extra pixel (or the missing one, for the third graph) has to be stuck in (or removed) somewhere, and given that the difference is just 1, roundoff at .5 will make it occur in the middle, at generation 100. And sure enough--there is a white band at that point in the second graph, and a corresponding black band there in the third. It was the desire to be able to eliminate these extraneous bands (especially problematic given the complexity of the patterns produced with chaotic behavior) that led us to the decision to allow direct control over the graph coordinates.

Where do you think the bands will occur if you put in or take out 2 pixels? or 3? or n ? Try it! Here mathematical relations and the fact that the computer is a digital machine are used to make differences in length visible which could not be otherwise perceived or measured directly. Similar processes of tuning the stimulus to the visual system by changing the ratio of two quantities are used in a variety of other measurement processes throughout science. Thus, with Vernier calipers, matching intermediate marks on 2 scales which divide equal lengths into 9 vs into 10 equal divisions gives an additional decimal place in accuracy of measurement.



Figures E4 through E6. Moire or "crystal dislocation" effects arising when a graph (E4) is lengthened (E5) or shortened (E6) by 1 pixel.

(See also Figures E9 through E14 below.) "Newton's rings" provide accurate measurements of the distance between two glass lenses (one flat, the other curved) in units of the wavelength of light through the production of interference bands of varying intensity--used by lens grinders to produce symmetrically ground lenses. Pilots can fine-tune the relative speed of aircraft engines in multi-engine aircraft by listening to the beats produced by the differences in their speeds. And variable speed strobe flashes can accurately measure the speed of a rotating wheel by tuning the strobe to "freeze" the image--which happens when the strobe frequency matches the time it takes the next radially symmetric pattern in the wheel to rotate into the position of the last. (If the strobe is a little too fast, the wheel will appear to be rotating slowly backwards--a scene that appears often also in movies or videos.)



Figures E7 and E8. Confusion of moire effects when there is no small common divisor.

Too large a mismatch, producing too many bands, makes problems, because the extra variability overloads the visual system, and this process of counting bands loses diagnostic effectiveness. (This is true for all of the examples mentioned above—the ratio of the two quantities must be tuned to a small integral multiple, or the difference between the two quantities being compared is lost in a welter of confusing pattern.) This is illustrated in the last two graphs in this series, Figures E7 and E8. These are 200 generation graphs which have pixel widths of 350 and 351, respectively, producing pixel ratios of 1.75 or 1.755 per generation. This generates a repeat pattern of 4 generations every 7 pixels in the first which produces a confusion of visual complexity. Can you see a difference in their patterns? You usually can find differences by looking in one for a variety of very small features found in the other. Is it a difference you can use to tell you how to change the length to get more regular patterns? I've never been able to tell which is which by looking. (The first should produce the more regular appearance, but to my eyes, it is the second which appears to do so, and in any case, it is not clear which way to go.)

For those of you working with a laser printer, which prints at 300 dots per inch (or DPI), the process described above presents new complexities. The standard screen pixel ratio for Macintoshes (and the ImageWriter printer) is 72 DPI; 300 DPI is not an exact multiple of this, so trying to print images designed to be shown at 72 DPI at 300 DPI can produce additional interference effects, because the printer driver will aim to

preserve the size of the image rather than its exact proportions. To best see this, print the graphs on one of these pages (probably the three of 401, 402, and 400 pixels) out at the different resolution settings on the printer (300, 150, and 75 DPI), both with and without the "best bitmap" option checked. The best bitmap option shrinks the image to 96% of its original size, producing resolution settings of 288, 144, and 72 DPI, which are integral multiples of 72, and removes these extraneous interference or moire effects. (In printing these out, particularly for the r vs N plots below, make sure that you do NOT have the "smooth bitmap" option on. This can introduce all kinds of spurious forms, particularly in the clouds of random points produced in the chaotic regions of bifurcation plots. You might try it both ways to see what will happen.)

So the net effect of all of this is two warnings: (1) use the graph coordinate options to set the graph width to eliminate extraneous patterns in the screen pictures and your scrapbook and PICT files; and (2) use the "best bitmap" option where it is available (on all laser printers, and most other higher resolution (e.g., inkjet) printers) to keep from reintroducing interference patterns in your printed output.

All of this might seem out of place here. Isn't this more appropriate to a printer manual or a handbook on desktop publishing? It might well be appropriate there too, but it is included here because it directly illustrates again one of the most important lessons of this module—that the computer is just another kind of instrument, and works no better than the assumptions put into it. This includes not just the models simulated, but the program written to simulate them, the language the program is written in, the screen and printer drivers written to display the results of the program, and the computer, screen, and printer hardware which turn these instructions into visible output. Until you understand the kinds of errors, biases, and noise that can be introduced by each of these, you cannot reliably tell artifactual errors from "real" phenomena (meaning phenomena which are products of the system under study and not of the instruments used to study it) under conditions where artifacts can arise. You need to be able to tell these various kinds of errors apart if you are going to be able to eliminate or correct for them, and you need to know how they can arise if you are going to design your experiments and design or calibrate your instruments to show what you hope they can show, and to know when you can trust their results. Patterns are an important diagnostic tool for us, and interference effects change patterns, so we need to study them here.

But why, if we know how to eliminate or correct for these errors, don't we just eliminate all of them? Unfortunately, just as there is no universal set of variables to be controlled for in an experiment in order to eliminate error, there is no general procedure for eliminating all of these biases. Thus, in an experiment, you can't control for the variable you are studying, because controlling it means keeping it constant, and if you are studying it, you need to vary it to see the effects of changing it. Thus, you can't control for everything because one person's experimental control is bound to be someone else's variable of interest. Computers, programming languages, software applications, and printer drivers are all cost-benefit compromises designed to produce tools which are usable by a wide variety of individuals for a large range of tasks. If you try to automatically correct for everything, you will make the devices larger, more complex, and more expensive, and—undoubtedly—you will introduce constraints on how you do things which a variety of different people will need to violate in different

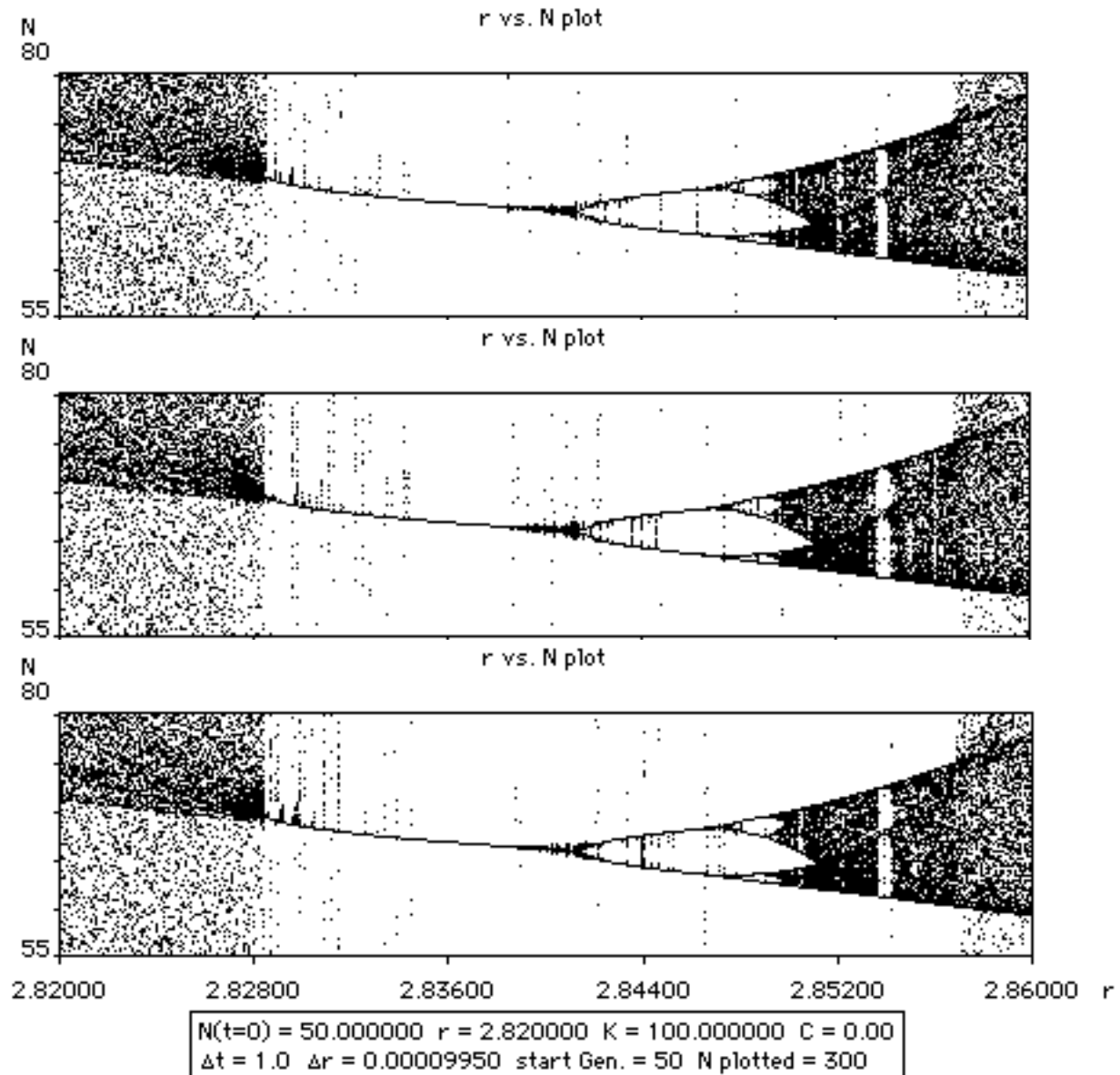
ways. Special purpose instruments and special market software are just devices for which a certain group has an interest in taking certain constraints as given because they don't have to fiddle with them to do what they want to do, and not having to fiddle with them makes their task easier. Given this, the only defense we have is to learn to recognize when we are coming up against these constraints, to know either what we can and cannot trust if the constraints are ineliminable, or how to get around them to eliminate or minimize their biasing effects if they can be finagled.

Real Differences Which Are Products of How the Data Are Sampled

To make good on this promise, we will illustrate a kind of case which it is important to distinguish from the preceding interference or moiré effects, although it seems similar. This case also involves changing the width of the graph, but because of how the r vs N viewing mode works, this actually changes the data produced by the simulation. It is easy to understand why the data are changed in this case, but we will follow this with a "mystery effect" whose cause we have not proven (though there are good reasons to believe what it is) which could well be a case worth substantial further study.⁷⁴

The three pictures below (Figures E9 through E11) show the middle strand in the period-3 window in the r vs N or bifurcation plot for the logistic. Graph height has been reduced to 100, and the bottom labels for the top two have been deleted in order to aid comparison. These pictures differ in the width of the graph (set using the **graph coordinates** option at Right - Left = 400, 401, and 402 pixels, respectively). These small differences in graph width aren't directly noticeable as length differences by eye, but they make a profound difference in the r -values used to construct the graph. All three graphs use the same lower and upper limits for r -values, but they interpolate 400, 401, and 402 r -values in between them, yielding corresponding Δr values of .00010000, .00009975, and .00009950, respectively. (The number of pixels for a strange graph can be calculated by dividing the range in r ($2.86 - 2.82 = .04$ in this case) by the value of Δr found in the data box at the bottom of the graph. Rounding this to the nearest integer should yield the number of pixels.)

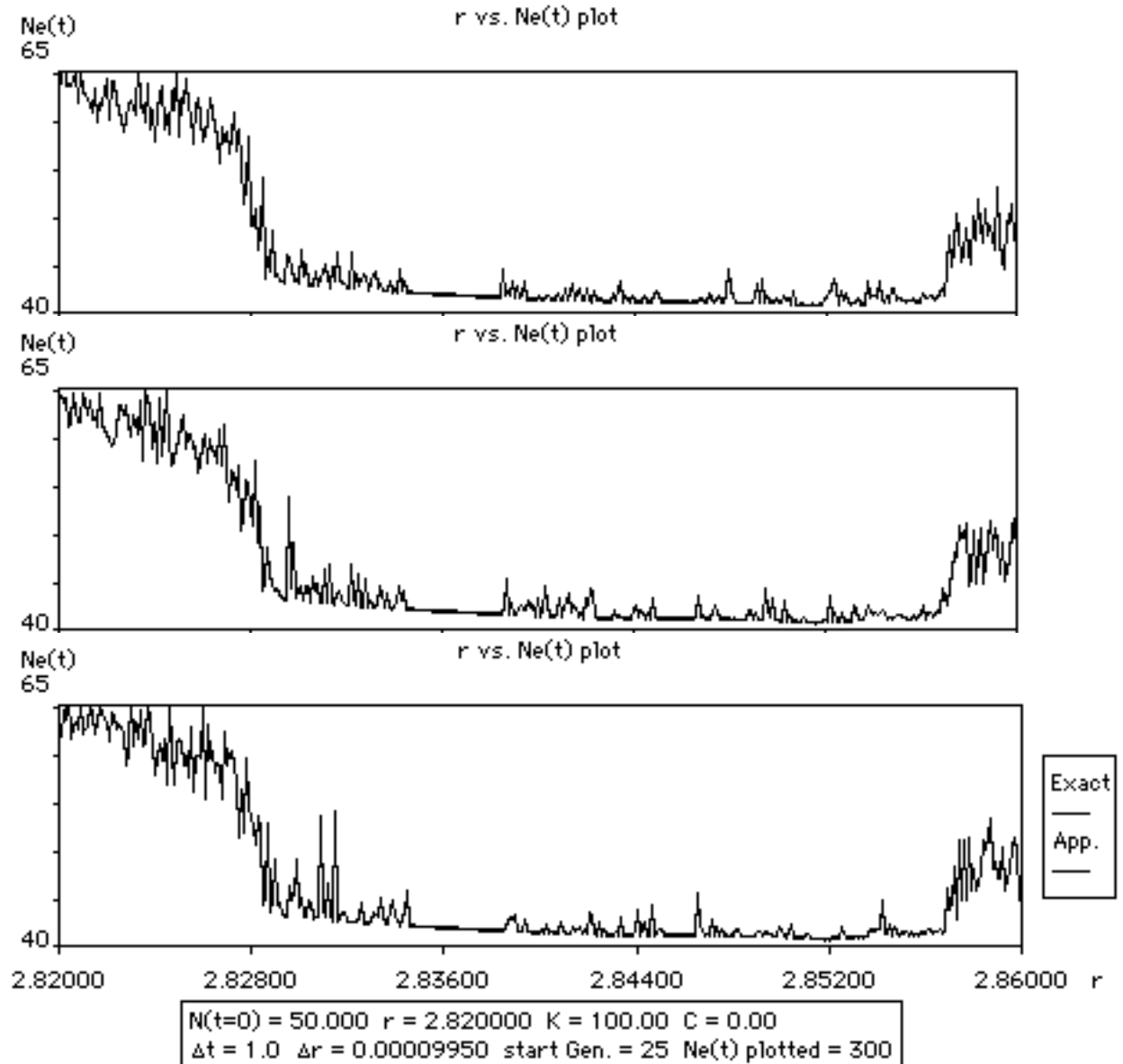
⁷⁴Although you can't test this explanation directly, you can test it indirectly using the "round N" or the "Numerical precision" options to see if you can replicate phenomena like those found here. See the discussion of the use of these options in chapter 3.



Figures E9 through E11. Vernier interpolated r -values with three bifurcation plots of widths 400, 401, and 402 pixels.

The first and second graph have no r -values in common except at the endpoints, and the sampling points are maximally out of phase in the middle. It is a measure of the infinite sensitivity of the fine structure of chaos that the graphs differ noticeably across their entire range. The first and third bifurcation graph differ in width by 2 pixels, rather than 1, and are in phase in the middle as well as at the ends, and maximally out of phase at the $1/4$ and $3/4$ points. For plotting smooth curves or patterns these differences would have no effect, but the infinite sensitivity of chaotic behavior guarantees that these minuscule differences in where the r -values are sampled produce noticeable changes. (These differences can never be larger than half of the width of the broadest band sampling of r —half of .0001, or .00005, and will range down from there to 0 in 200 increments) Major features remain roughly the same in all three,

but smaller features (particularly the presence and width of transient dark bands of points, and correspondingly, the structure, width, and even the presence of the equilibrated lighter bands of points) have changed.



Figures E12 through E14. Vernier interpolated r -values with three effective population size plots of widths 400, 401, and 402 pixels.

Figures E12 through E14 are the r vs $N_{eff}(t)$ plots for the same parameter values and pixel widths as the bifurcation plots, and show corresponding differences from one another. Detailed comparisons within this group show the same patterns of macro-similarities and micro-differences. Cross comparisons between the two sets are also useful, since they give different reflections of the differences produced by changing the sampling interval. These differences tend to show up in the same places, i.e., for the same r -values, indicating that effective population size is a good indicator variable for at

least some aspects of the structure of the bifurcation plots. It should thus be clear in any case that graph width is an extremely important variable in the r vs N and the r vs $N_{eff}(t)$ plots, and that reproducing a result in these plot modes requires use of the same graph width.

These last 6 plots in effect represent a generalized version of Vernier interpolation: If 2 scales cover the same interval, each dividing it into v and $v+1$ equal intervals, respectively, one can use the proportional distance along the whole scale to estimate how far apart (what fraction of a unit) the corresponding small divisions are out of synchronization. Vernier interpolation can be used to give a more accurate reading of a variable which changes monotonically (ideally, linearly) between the two endpoints. Chaotic behavior fails to meet that condition, and so you might think that Vernier interpolation would be useless. In this case, it serves the opposite conclusion: the fact that the three plots are no more alike half-way or thirds of the way along the plot gives a much higher-resolution confirmation of the infinite sensitivity to initial conditions characteristic of chaotic behavior. There is behavioral divergence not only between neighboring pixels ($\Delta r \approx .0001$) but also (contrasting the divergent behaviors of corresponding points on homologous graphs) for differences as small as $1/400$ of that, or .00000025.

Roundoff Errors Unmasked—Quantized Escape Orbits

We now turn to a more difficult case—one where the argument for the causes of the phenomenon has to be indirect, but in which it is fairly certain none-theless, as can be shown by subjecting the phenomenon to various tests. (See also note 1, 3 pages above.) This is a phenomenon which you will probably discover when you do the problem suggested above of localizing the boundary of an attractor by doing successive sweeps at successively higher resolution. In these runs, just as in those, r is fixed, and the differences are differences in the starting values for population size, N_0 . These trajectories do not appear to straddle the boundary, as the trajectories you would choose for that investigation (they don't because they all escape in the same way at about the same time--there are no "holdbacks") but they must be very close to it as it takes 39 generations before the first deviation becomes observable in the graph. This deviation is a motion of a point in trajectory 10 (white or hollow circles) up 1 pixel in the middle line, showing an edge of trajectory 6 (black or solid light blue circles) below it. The puzzling case is shown below:

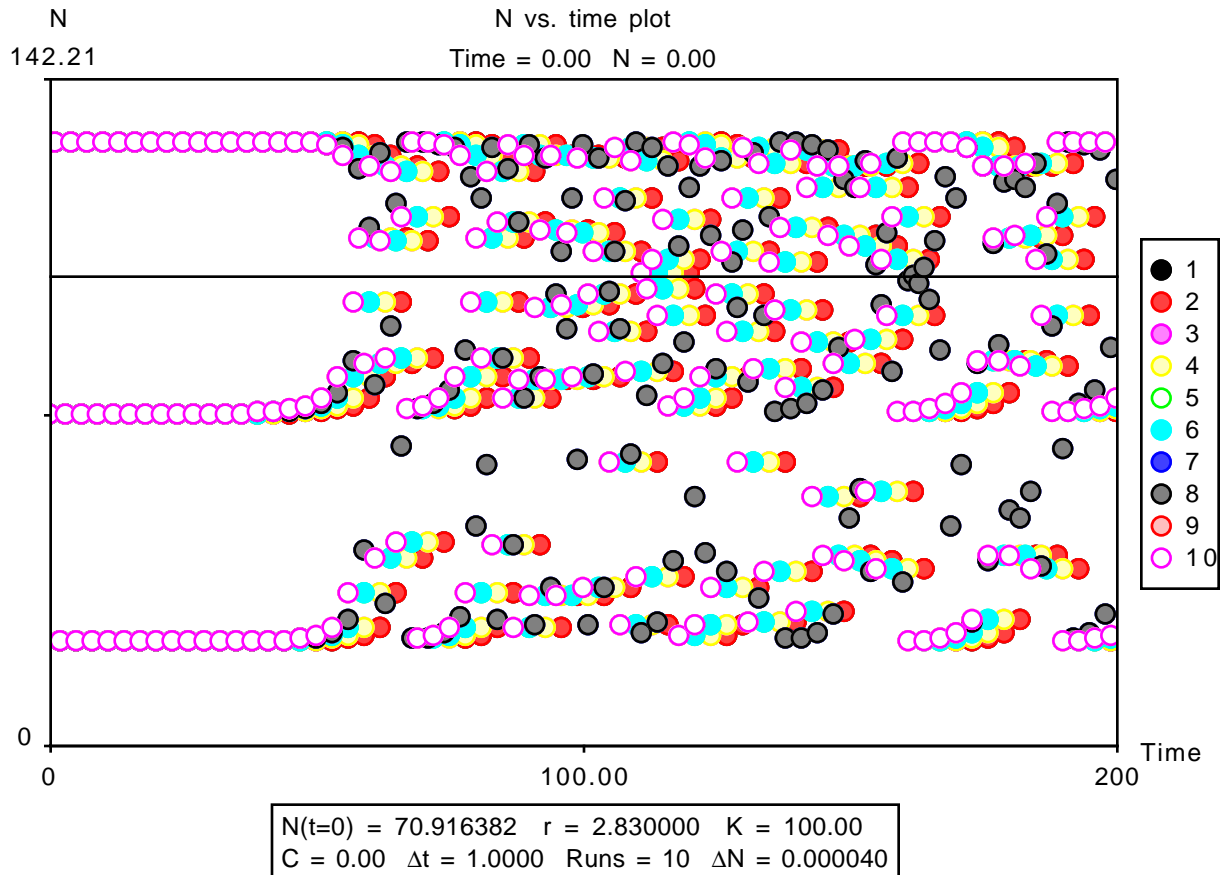


Figure E15. Nearly synchronous phase-shifted divergence in trajectories near an attractor boundary.

What is puzzling about this case is that there is entirely too much regularity in it for what we expect with chaotic behavior. In fact, if you ignore the medium gray spots, you will see that the entire rest of the picture is made up of repeated quartets of spots in which dark gray is covered by light gray, which is covered by black, which is covered by white. (This is much more obvious when seeing the PICT redrawn on the computer screen, as when scrolling the image into view. There the repetition and phase shifts of the trajectories become objects for our visual processing of motion detection and integration, which immediately pick out the similarities in the shifted trajectories—sensed most immediately as global color or gray-shifts as the successive trajectories are drawn. If you can't see this, try plotting 400 generations in the same width graph—401 pixels. This will pull the quartets closer together and make it easier to see them as a single (fused) object.)

In order to figure out what is going on, we need first to be able to identify the trajectories, so that we can relate changes in them to changes in the starting values for population size, N_0 . Then we need to be able to characterize the patterns of changes in these trajectories as N_0 increases. This will involve in part doing additional runs for families of trajectories with different values of ΔN . Finally, we need to combine what we know to analyze the causal interactions which produce the patterns.

We start by noting that the identity of trajectories seems ambiguous in black and white—from the trajectory labels it seems that the hollow circle could be 5, and the solid one 1, instead of the postulated 6 and 10. In fact, from the trajectory labels at the right of the graph, it seems that each type of spot from black to white has two possible trajectories that it could be, one in the interval 1-5, and another in the interval 6-10. This and other similar ambiguities can be removed in different ways:

First, the apparent ambiguity can be removed by deduction, if we note that higher numbered trajectories are plotted after lower numbered ones and can partly or completely cover but cannot be at all covered by them. The solid circles cover others and thus cannot be from the first trajectory. So they are 6's. Thus the open circles, which cover them, can't be 5's and must be 10's. Similar reasoning reveals that a darker can cover a lighter spot only if the lighter is from the first 5 trajectories, and the darker is from the second 5. (Look at the key for earlier and later trajectories.) Now look at the repeated multiples of 4 trajectories ganged together with separations of 3. The solid 6 (overlapped by the 10) overlaps a lighter colored spot, which therefore has to be a 4, rather than a 9, which overlaps a darkest gray, which thus has to be a 2, rather than a 7. Further looks for overlaps reveal that the medium gray has to be an 8, rather than a 3. Thus the visible trajectories are 2, 4, 6, 8, and 10.

There are two other means of validating these inferences, and you should use both, though either will do; the third can be done on any machine.

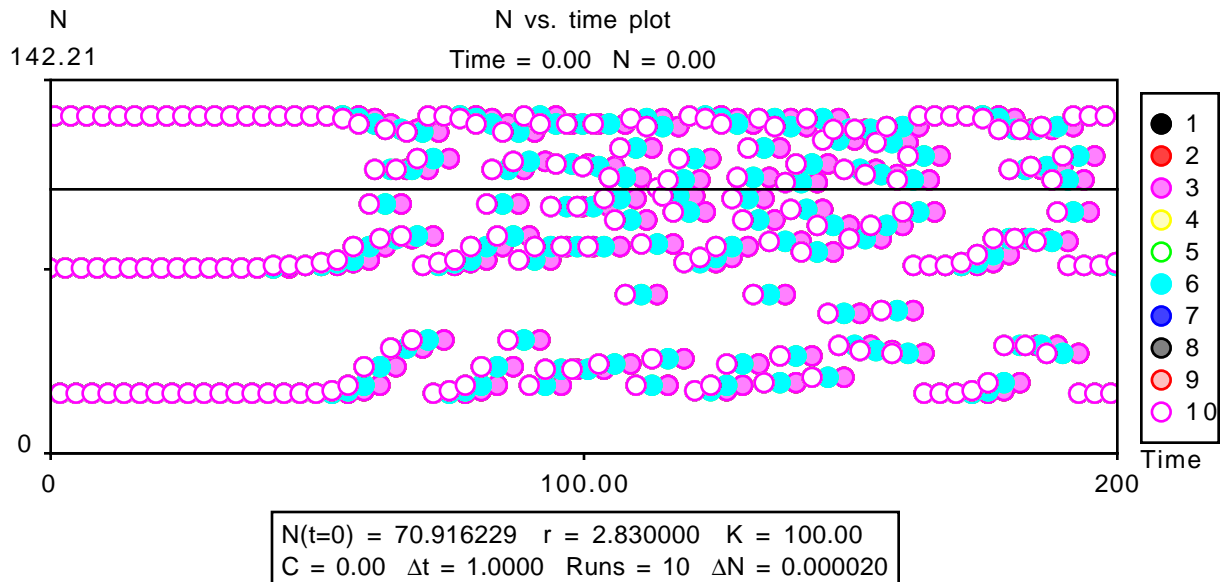
Second, the ambiguities are completely removed on a color monitor, where the visible spots can all be seen to be from trajectories 2 (dark gray), 4 (light gray), 6 (black), 8 (medium gray), and 10 (white), with colors red, yellow, light blue, black, and pink, respectively.

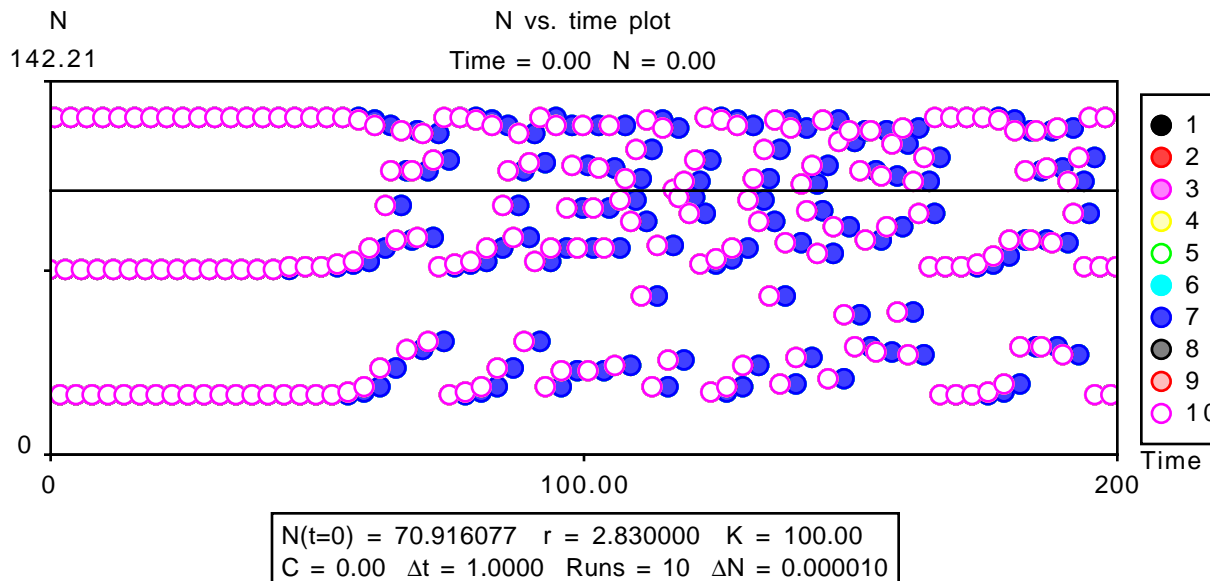
Third, if you remember the plotting order and play (*redraw*) the runs at a slower speed, (3 works fine, but try your own speed) you can see what covers up what in what order. This will show that odd-numbered runs are covered by their successors, and that each odd number establishes a new trajectory. So in the series, {1,2}, {3,4}, {5,6}, {7,8}, and {9,10}, the second number in each bracket covers the first, and there are 5 distinct trajectories, initiated by 1, 3, 5, 7, and 9.

A sense of scale can be obtained by looking at the spot separations at the beginning before the trajectories escape the period-3 cycle and remembering that successive spots in each line have to be 3 generations apart. But this is also just the separation between the quadrupled spots in the above figure. What does it mean for the spots to be quadrupled together in this way? It means that most of the trajectories (8 of them to be exact) are the *same* trajectory--differing only in *when* they escape from the period 3 lines. Trajectories 9 and 10 are followed 3 generations later by trajectories 5 and 6, they in turn 3 generations later by 3 and 4, and they 3 generations later by 1 and 2. Trajectories 7 and 8, though identical to each other, appear to be different from all of the others. So in all, we see only two basic trajectory types, and the most common change between trajectories appears to be just a phase shift. That this phase shift is always by 3 and is significant. We will return to it below.

Now it is time to notice just how anomalous this is. Our experience with chaos so far has been to demonstrate that (A) any differences in starting values, no matter how small, cause exponentially increasing divergences between trajectories with increasing time, and (B) there seems to be an arbitrarily large number of different trajectories—all apparently random in appearance. But here, we saw (1) that neighboring increments in ΔN yielded identical trajectories, (2) that accumulating increments caused phase-shifted trajectories, but not different trajectories, and (3) that 10 runs yielded only 2 basic trajectory types. Taken together, these seem to contradict both properties (A) and (B) which we have learned to expect from chaotic systems.

I want to suggest an explanation which seems to explain almost all of the phenomena observed. Note that the differences in ΔN between successive trajectories are very small (.00004), and that right at the boundary of the attractor, trajectories are delicately poised between falling in and flying out—so delicately that they move very little from one generation to the next (or actually every third generation) so they diverge relatively slowly. I want to suggest (1) that given the small size of ΔN , roundoff error is removing the significant differences between successive trajectories (2) that roundoff error is similarly reducing the number of available "escape" trajectories available in that neighborhood (i.e., in the boundary region). If this explanation is correct, then (a) making ΔN still smaller for this value of N_0 should reduce the variety of trajectories still further, and (b) ΔN 's of this size should not have a similar constraining effect on trajectory number for values of N_0 not near the boundary.





Figures E16 and E17. Reduced variety of phase shifts with smaller Δt 's.

We can see exactly this in the next 2 graphs, in which N_0 is the same as before (set at 70.916) but ΔN is set at .00002 and .00001, respectively. We see that in both of them, there is now only one trajectory, which is phase shifted—twice for .00002, (after trajectories 3 and 6) and once for .00001 (after trajectory 8). The variety decreases (as predicted) with smaller values of ΔN , and in fact disappears entirely for slightly smaller values. You can see for yourself that clause b is met—that we do not get the same restricted array of trajectories for other values of N_0 .

There is one other revealing fact suggesting the same thing more directly. The values for N_0 showing in the parameter listings below the graph will be those of the last trajectory, including possible roundoff errors. All of these plots started with N_0 set for 70.916. What they end up with should be 9 times the increment value of ΔN they actually used greater than this value. (To run 10 trajectories, the starting value together with the value from 9 successive increments are used.) The differences are .000382 (instead of .000360) for $\Delta N = .00004$; .000229 (instead of .000180) for $\Delta N = .00002$, and .000077 (instead of .000090) for $\Delta N = .00001$. There are two interesting things to notice: first, that the percentage errors in these increment sums from what they should be are enormous—low for 1 and high for 2 and 4×10^{-5} . Second, the summed increments are in a ratio very close to 5:3:1 instead of 4:2:1. This suggests that the smallest effective unit (the roundoff quantum) that the computer can use in incrementing N at $N_0 = 70.916$ is approximately $1/9$ of .000077, or 8.5×10^{-6} , and that it used 1, 3, and 5 of these units as "best fits" to the desired 1, 2, and 4×10^{-5} .

The second question is where this rounding off error occurs which reduces the variety of trajectories possible, and here the patterns give us another intriguing clue. Phase shifts are the changes produced by the smallest differences in ΔN . They are the last changes to disappear before all trajectories look alike. One of the remarkable things about the phase-shifted trajectories is that the phase-shift is always by a multiple of 3 generations. This suggests that the roundoff error is always occurring in the same

place, i.e., when the trajectories are in the same phase of their 3-point cycle. Once it has occurred, producing identical values in the different trajectories, they will be constrained to follow each other from there on. And where might this be? In fact, it has to occur before they diverge from the 3-point cycle (both because of the period-3 phase shifts, and also because the trajectories are clearly (i.e., observationally) already entrained when they leave the 3-point cycle. But which point of the cycle is to blame? We can answer this one too: it has to be the top point because there the envelope of the attractor is mapped into the narrowest region and roundoff errors will have the largest effect in gathering diverse trajectories as one. This also explains why the "roundoff quantum" should be so large measured at $N_o = 70.916$. It is actually the much smaller spread around the top point at 129.575 that gives the real size of the roundoff error. And in which cycle does it occur? Barring means whereby the divergences of trajectories can get smaller then larger (in the topmost point of the trajectory, where they are already maximally close), we have to suppose that the quantization occurs the first time they pass through the top region.

An unresolved question which I leave for you to puzzle over is where the aberrant path followed by trajectories 7 and 8 came from—flanked as it was by the common phase-shifted trajectories on both sides. And investigating these phenomena more, how phase shifting and different trajectories can happen in this realm of deterministic quanta, and how the patterns of appearance of diverse trajectories and phase shifts, as ΔN is increased, are to be explained. You might want to explore this phenomenon in general using the "round N" option and with different K values to see if you can generate roundoff errors that are large enough to investigate further. Good luck!

(* Appendix F *)

(* This entire appendix can be used as a Mathematica program. All text enclosed between the "parenthesis/asterisk" symbols, such as the symbols around the title of this appendix, are considered comments, and Mathematica will ignore them. Do not remove those symbols if you wish to run this program in Mathematica. *)

(* NOTE ON USING THIS APPENDIX IN MATHEMATICA: First, you should be familiar with the basic Mathematica user interface and the basic syntax and commands for performing symbolic manipulations and plotting functions. Second, in order to reproduce--as closely a possible--the printed version of this appendix, you should do four things:

(i) Save this document as a Text Only file.

(ii) Choose "Select All" from the Edit menu, change the Font to Palatino, and then simultaneously press the Shift-Apple-d keys in order to make the text Plain and not Bold.

(iii) Next, after each Mathematica instruction (e.g., $\text{LeftSide} = \text{Integrate}[1/N, N]$) choose from the Cell menu "Divide Cell" to divide the cells after each instruction. In this way, Mathematica will place the output of an instruction below each instruction.

(iv) Again choose "Select All" from the Edit menu and either press the "Enter" key or evaluate the cells from the Cells menu.

In this appendix, we will use the computer program Mathematica to symbolically solve and plot the continuous versions of the exponential, logistic, and minimum-density limited logistic growth equations:

$$(1) \quad dN/dt = rN \quad (\text{exponential growth})$$

$$(2) \quad dN/dt = rN(1 - N/K) \quad (\text{logistic growth})$$

$$(3) \quad dN/dt = rN(1 - N/K)(1 - C/N) \quad (\text{minimum-density limited logistic growth})$$

where N = population size or density, r = growth rate, K = Carrying Capacity, C = minimum population size or density.

For each equation we can rewrite it in separable form:

$$(1) \quad 1/N \, dN = r \, dt \quad (\text{exponential growth})$$

$$(2) \frac{1}{N(1 - N/K)} dN = r dt \quad (\text{logistic growth})$$

$$(3) \frac{1}{(N(1 - N/K)(1 - C/N))} dN = r dt \quad (\text{minimum-density limited logistic growth})$$

For each case, we will integrate the left and right sides of these equations and solve for N under the initial value condition that $N(0) = N_0$.

First, for equation (1) we integrate the left and right sides for the variables N and t, respectively: *)

$$\text{LeftSide} = \text{Integrate}[1/N, N]$$

$$\text{RightSide} = \text{Integrate}[r, t] + A$$

(* where A is an arbitrary constant of integration. *)

(*We can now solve for the arbitrary constant A under the initial condition $N(0) = N_0$, which is simply: $A = \text{Log}[N_0]$. We then substitute $\text{Log}[N_0]$ for A: *)

$$\text{RightSide} = \text{RightSide} /. A \rightarrow \text{Log}[N_0]$$

(* Finally we exponentiate both sides to solve for N *)

$$\text{LeftSide} = \text{Exp}[\text{LeftSide}]$$

$$\text{RightSide} = \text{Exp}[\text{RightSide}]$$

(* Note that Mathematica does not always write an equation in its simplest form even when the function Simplify is called explicitly. In simpler form, the right side of the above equation becomes: *)

$$N_{\text{exp}}[t_] = N_0 * \text{Exp}[r * t]$$

(* We can now use Mathematica to plot a family of curves for exponential population growth by varying the value of r for each plot: *)

```
Show[Plot[Nexp[t] /. {N0 -> 25.0, r -> 0.1}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nexp[t] /. {N0 -> 25.0, r -> 0.2}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nexp[t] /. {N0 -> 25.0, r -> 0.3}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nexp[t] /. {N0 -> 25.0, r -> 0.4}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nexp[t] /. {N0 -> 25.0, r -> 0.5}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nexp[t] /. {N0 -> 25.0, r -> 0.6}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nexp[t] /. {N0 -> 25.0, r -> 0.7}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nexp[t] /. {N0 -> 25.0, r -> 0.8}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nexp[t] /. {N0 -> 25.0, r -> 0.9}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nexp[t] /. {N0 -> 25.0, r -> 1.0}, {t, 0, 10}, DisplayFunction -> Identity],
DisplayFunction -> $DisplayFunction];
```

(* Next, consider the logistic growth equation (2). We will use the same procedure as before. *)

```
LeftSide = Integrate[1/(N*(1 - N/K)), N]
```

```
RightSide = Integrate[r, t] + A
```

```
Solve[LeftSide == RightSide /. {t -> 0, N -> N0}, A]
```

(* We then substitute for the arbitrary constant A back into the right side. *)

```
RightSide = RightSide /. A -> (-Log[K - N0] + Log[N0])
```

(* We can now exponentiate each side and solve for N *)

```
Solve[Exp[LeftSide] == Exp[RightSide], N]
```

```
Nlogistic[t_] = (Exp[r * t] * K * N0)/(K - N0 + Exp[r * t] * N0)
```

(* We can again plot a family of curves for the logistic growth equation, and again we will vary the value of r for each plot. *)

```
Show[Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 0.1, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 0.2, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 0.3, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 0.4, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 0.5, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 0.6, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 0.7, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 0.8, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 0.9, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[Nlogistic[t] /. {N0 -> 25.0, r -> 1.0, K -> 100}, {t, 0, 10}, DisplayFunction -> Identity],
PlotRange -> {20, 100}, DisplayFunction -> $DisplayFunction];
```

(* Note again that Mathematica does not always find a solution in its simplest form. A simpler form for the solution to the logistic equation is given below: *)

$$Nlogistic[t] = K / (1 + (K/N0 - 1) * \text{Exp}[-r * t])$$

(* Finally, we solve the minimum-density limited logistic equation. *)

$$\text{LeftSide} = \text{Integrate}[1/(N * (1 - N/K) * (1 - C/N)), N]$$

$$\text{RightSide} = \text{Integrate}[r, t] + A$$

(* Solving for A we get: *)

$$\text{Solve}[\text{LeftSide} == \text{RightSide} /. \{N \rightarrow N0, t \rightarrow 0\}, A]$$

(* Now, we can simplify the left side by simple multiplication and division: *)

$$\text{LeftSide} = \text{Simplify}[\text{LeftSide} * (-C + K) * 1/K]$$

$$\text{RightSide} = \text{Simplify}[\text{RightSide} * (-C + K) * 1/K]$$

(* Finally we exponentiate both sides, substitute for A, and solve for N: *)

```
Simplify[Solve[Exp[LeftSide] == Exp[RightSide] /. A -> (K * (-Log[C - N0] + Log[K - N0]))/(C - K), N]]
```

```
NMDL[t_] = ((-C + (Exp[((-C + K) * r * t)/K] * K * (C - N0))/(K - N0)))/((-1 + (Exp[((-C + K) * r * t)/K] * (C - N0))/(K - N0)))
```

(* We can again plot a family of curves for the logistic growth equation, this time varying the value of N0 close to C: *)

```
Show[Plot[NMDL[t] /. {N0 -> 24.0, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 24.2, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 24.4, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 24.6, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 24.8, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 25.0, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 25.2, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 25.4, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 25.6, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 25.8, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
Plot[NMDL[t] /. {N0 -> 26.0, r -> 0.5, K -> 100, C -> 25}, {t, 0, 10}, DisplayFunction -> Identity],
PlotRange -> {-60, 60}, DisplayFunction -> $DisplayFunction];
```


References:

General Mathematical Background:

Batschelet, Edward, (1979); ***Introduction to Mathematics for Life Scientists***, 3rd. ed., New York: Springer-Verlag. (paper: Springer Study Edition) (E/M) *This is an excellent source for relevant mathematical background for both of the model units in this module (and for most other biological topics), beginning with topics from high school algebra and running up through probability, matrices, calculus, and differential equations, omitting only programming and statistical techniques. It illustrates all topics with biological examples, and contains worked and unworked problems from the current research literature. This is a clear and well motivated book, and is also particularly well suited for self-study. For this module, see particularly his discussion of exponential functions, growth equations, and of the binomial, normal, and Poisson distributions.*

Boyce, William E. and Richard C. DiPrima, (1991); ***Elementary Differential Equations and Boundary Value Problems***, 5th ed., New York: Wiley. (M/H) *This is a modern text, emphasizing the use of the computer in studying differential equations. Chapters 2 (especially Sections 2.6 on population growth models and 2.12 on first-order difference equations) and 9 (on nonlinear differential equations, stability analysis, and chaos--see especially sections 9.4 through 9.8) contain systematic discussions of the exponential, logistic, MDLS, and Lotka-Volterra equations and their analysis. Section 2.12 contains an unusually clear exposition of stability analysis for the logistic using the lag-1 phase plot. Many of the problems complement the exercises suggested here.*

References on Visualization:

Tufte, Edward M., (1983); ***The Visual Display of Quantitative Information***, Cheshire, CT: Graphics Press. (E/M) *There is no better source to convince you of the importance of visualization in science (and how it can be manipulated for good or for bad) than these two books. Contains extensive discussion and illustration of desiderata for good graphics. The first book is also a history of the evolution of statistical graphics. Both of these books are full of paradigms of good and bad graphic design, and are printed with a quality and care seldom (if ever) found in the commercial publishing world.*

Tufte, Edward M., (1990); ***Envisioning Information***, Cheshire, CT: Graphics Press. (E/M) *Broadens the scope of investigation beyond quantitative graphics to consider other uses and dimensions of visualization. Like the first book, an illuminating (and illuminated) visual feast.*

Wimsatt, W. C., (1991); "Taming the Dimensions—Visualizations in Science", in M. Forbes, A. Fine and L. Wessels, eds., ***PSA-1990: volume 2***, East Lansing, MI: The Philosophy of Science Association, pp. 111-135. *Documents the role of visualization in scientific problem-solving, with paradigmatic examples of different tasks for which visualization is particularly well-suited.*

References on Population Growth Models and Population Genetics and Evolution:
 (These texts are included not only for their relevance to the topics discussed here, but because they are also an excellent source of other models for further study.)

Charlesworth, Brian and Deborah Charlesworth, (1990); "Mimicry", in M. Nitecki, ed., ***Evolutionary Innovation***, Chicago: The University of Chicago Press. *A readable and up-to-date account of the evolution, population genetics, genetic basis and complications of mimicry--a finely tuned multi-locus adaptive trait.*

Crow, James F., (1986); ***Basic Concepts in Population, Quantitative, and Evolutionary Genetics***, New York: Freeman. *A masterful presentation of the relevant mathematical theory; a somewhat simpler (and more up to date) version of Crow and Kimura, and a nice complement to Hartl, which is more empirically oriented.* (M).

Crow, James F. and Motoo Kimura, (1970); ***An Introduction to Theoretical Population Genetics***, Harper and Row: Evanston. (Chapter 1) (M/H/V) *The classic and standard modern graduate level text in population genetics by two of the world's leading population geneticists. The first chapter, using growth equations as an example, is explicitly about model-building.* (M for this chapter).

Ewens, W. J., (1969); ***Population Genetics***, London: Methuen. (H/V) *Elegant but technically demanding.*

Goldberg, David, (1989); ***Genetic Algorithms in Search, Optimization and Machine Learning***, Reading, MA: Addison Wesley. (E/M for most things, some H) *Genetic algorithms are problem-solving methods in artificial intelligence modeled on the population genetics of bacteria--haploid sexual organisms with recombination. An interesting variant of standard population genetic theory, which sometimes benefits from the different perspective. This is the clearest introduction to work in genetic algorithms available. Contains Pascal programs for all of the major systems.*

Hartl, Daniel C., (1988); ***A Primer of Population Genetics***, Sinauer: Stamford, Conn., 2nd. ed. (M for most things, some E, H) *A very clear presentation of basic relevant genetics, population genetics, quantitative genetics, and molecular evolution. Good worked problems; covers more advanced material (simply) than other comparable texts. The natural next step up from Wilson and Bossert on population genetics.*

Hartl, Daniel C. and Andrew G. Clark, (1989); ***Principles of Population Genetics***, Sinauer: Stamford, Conn., 2nd. ed. (M for most things, some E, H) *A very clear and full presentation of relevant genetics, population genetics, quantitative genetics, and molecular evolution, with more attention to advanced topics than Hartl above, though the level of difficulty of presentation is surprisingly not that much greater. Excellent graphics. Probably now the standard graduate text.*

Kingsland, Sharon B., (1986); ***Modelling Nature: The Origins of Theoretical Ecology***, Chicago: The University of Chicago Press. (E) *An excellent and readable history of the rise of mathematical modeling in ecology, from the late 19th century up through about 1970. Contains extended discussion of the development and interpretations of the logistic and related*

multi-species Lotka-Volterra population growth models. Complements Provine (below) for ecology.

Hubby, J. L. and R. C. Lewontin, (1966); "A molecular approach to the study of genetic heterozygosity in natural populations. I. The number of alleles at different loci in *Drosophila pseudoobscura*," **Genetics**, 54: 577-594.

Jacquard, A., (1976); **The Genetic Structure of Populations**, New York: Springer-Verlag, series in Biomathematics, vol. 5. *Probably the best French text on population genetics--particularly distinguished for its clarity, the use of demographic concepts and problems, and the focus on problems in human genetics--all unusual for texts in population genetics.* (M/H)

Lewontin, Richard C. **The Genetic Basis of Evolutionary Change**, New York: Columbia University Press. (mostly M, some E, H) *An elegant and methodologically sophisticated presentation of the application of population genetics to problems of evolutionary change and speciation, focusing on methodological problems of constructing and testing the theory, and the "neutralist-selectionist" controversy. Also has provocative views on model-building. Mostly readable with population genetics at the level of Hartl, 1988, though occasionally much more advanced, and many parts are readable without even this much.*

Lewontin, Richard C., (1979); "Adaptation", **Scientific American**, September.

Lotka, Alfred J., (1956); **Elements of Mathematical Biology**, New York: Dover. (M/H) *Published in 1924 as "Elements of Physical Biology" this book by one of the founders of mathematical ecology is still fascinating reading. Particularly useful is his extended discussion of types of equilibrium and stability, which is still one of the best available anywhere.*

MacArthur, Robert H. and Edward O. Wilson, (1967); **The Theory of Island Biogeography**, Princeton: Princeton University Press. (M) *One of the most influential works of modern theoretical ecology, deals with rates of colonization and extinction of species on islands. Relates more generally to the migration of animals through "patchy" environments, and the effects of the area of suitable patches to the number of species found there. Has led to the development of much of modern theory relating to conservation biology.*

Provine, William B., (1971); **The Origins of Theoretical Population Genetics**, Chicago: University of Chicago Press. (E/M) *Now a little dated in some respects (a great deal of historical work has been done in this area since), this is still a very readable introduction to the origins of population genetics from the time of Darwin to the publication of the major works of Fisher, Wright, and Haldane ushering in the synthetic theory of evolution around 1930. Good on the Mendelian/Biometrician debate.*

Provine, William B., (1986); **Sewall Wright and Evolutionary Biology**, Chicago: University of Chicago Press. (M) *An exhaustively researched and scientifically revealing intellectual biography of Sewall Wright, one of the greatest scientific minds of this century, and of the emergence of population genetics and its central role in the evolutionary synthesis. A masterful history of how we got where we are today which, unlike many scientific biographies, goes to great pains to explain the major ideas as well as the controversies surrounding their acceptance. Unfortunately no comparable biographies exist for Fisher and Haldane.*

Raup, David A., (1991); ***Extinction: Bad Genes or Bad Luck?***, New York: W. W. Norton. (E) *An engaging and fascinating discussion of different theories of and factors relating to extinction by a leading theoretician of paleontology. Focuses primarily on mass extinctions, (and extra-terrestrial causes--collisions with asteroids or comets), but contains material relevant to the discussion of extinction in general. Raup was one of the very first (1957) to use computer models in biology, and much of his work since has had to do with building and evaluating models.*

Roughgarden, Jonathan, (1979); ***Theory of Population Genetics and Evolutionary Ecology: An Introduction***, MacMillan: New York. (M/H/V) *A major reference work for its large variety of models in population genetics and ecology, this is also liberally illustrated with graphs and contains conceptually sophisticated, strikingly original, and revealing insights derived from the comparative analysis of sometimes surprisingly juxtaposed models. Contains an excellent theoretical discussion of chaotic behavior with the logistic growth equation in Chapter 14--remarkably modern in tone for its occurrence in 1979. The chapter on quantitative genetics is also useful background for the blending inheritance simulation.*

Williams, George C., (1966); ***Adaptation and Natural Selection***, Princeton: Princeton University Press. (E/M) *A deceptively straightforward argument and reductionistic interpretation of population genetics that genes are the only significant units of selection, and one need not consider adaptive organization at any higher levels. This book started a debate which continues to this day. For arguments against this view, see Wade, 1978 and Wimsatt, 1980b.*

Wilson, Edward O. and William Bossert, (1971); ***A Primer of Population Biology***, Sinauer: Stamford, Conn. (E) *Although it appears a bit dated (largely for its assumption that all calculations will be done by hand, and consequent choice of easy problems), this is a conceptually clear and very nice basic introduction to the exponential, logistic, MDLS, and age-structure models. It was revolutionary for its time in being self-consciously aware of the difference between models (idealizations which make known unrealistic assumptions) and theories (the currently most accurate general accounts that we have), and focusing on the former as the tools of theory construction in science.*

Wright, Sewall A. (1968, 1969, 1977, 1978); ***The Genetics of Populations***, Chicago: University of Chicago Press. (M/H/V) *This four volume set summarizes Wright's work and visions of evolutionary theory, and its relation to others. It covers in detail virtually all parts of evolutionary and population genetics, comprising in effect a 1-man authoritative and uncompromising encyclopedia of the field, and is still the only easily accessible source (beyond Wright's original papers) for many of these topics. Occasionally suffers from Wright's self-taught and idiosyncratic modes of mathematical representation.*

References on Chaotic Behavior:

Baker, Gregory L. and Jerry P. Gollub, (1990); ***Chaotic Dynamics: an introduction***, New York: Cambridge University Press, (paperback). (H) *An introductory text designed for second or third year math, physics, or engineering students, presupposing multivariate calculus, linear differential equations, and a year of physics. Clearly for majors, it gives a working knowledge of a number of the mathematical techniques used for analyzing chaotic behavior, including Fourier analyses, power spectra, and Lyapunov exponents. Contains*

problem sets and code in True BASIC™ for most of the analytical procedures. Discusses the logistic map, but examples and orientation are otherwise completely from physics. For those who want to go beyond the level of the analysis presented here to a deeper analytical understanding.

Karl-Heinz Becker and Michael Dörfler, ***Dynamical Systems and Fractals: Computer Graphics Experiments in Pascal***, New York: Cambridge University Press, 1989. (H/(M for programmers)) *A fascinating discussion of the whole range of chaotic and fractal phenomena, with lots of pictures, suggested exercises, and, most usefully, bits of code and programs for investigating them in Pascal on a variety of computers, including Apple's Macintosh. A very useful book.*

Feigenbaum, Mitchell, (1980); "Universal Behavior in Non-linear Systems", ***Los Alamos Science***, #1, Summer, pp. 1-27. (H) *A somewhat less technical but still difficult article by the discoverer of universality (and of course, Feigenbaum's numbers!) Discusses only applications in physics.*

Gliesh, James D., (1988); ***Chaos: the Making of A New Science***, Viking. (E) *Eminently readable, intriguing narrative of the developing awareness of chaotic phenomena and the emergence of tools for their analysis, with emphasis on biography and discovery, and almost no mathematics. It has become the common public introduction to the field. Flawed only by its strong bias towards happenings in the last 30 years and dearth of mathematics. (contrast Stewart's book).*

Hofstadter, Douglas. R., (1981); "Strange Attractors," ***Scientific American***, (mathematical games section), November. (E) *A compact and relatively readable introduction to chaos and strange attractors.*

May, Robert. M., (1974); "Biological populations with non-overlapping generations: stable points, stable cycles, and chaos," ***Science***, **186**: 645-647. (M) *The first article on chaos in ecology, by a physicist turned ecologist--and one of today's leading ecological theorists. Very much worth reading as an introduction to the phenomena. This paper was followed rapidly by many more (partial early bibliography on these through 1979 in Wimsatt, 1980).*

Schroeder, Manfred, (1991); ***Fractals, Chaos, Power Laws: Minutes from an Infinite Paradise***, Freeman. (M/H) *A recent discussion focusing on self-similarity as a unifying theme. Uses a lot of number theory. The style is stimulating, eminently readable, and informative (and rates an E) and the discussion of the logistic map contains very useful advanced theoretical material not available in any of the other sources given here (M), but it sometimes presupposes familiarity with mathematical material or ideas not discussed in the book (H).*

Stewart, Ian, (1989); ***Does God Play Dice?***, New York and Oxford: Basil Blackwell. (E/M) *A marvelously clear exposition of chaos and fractals by a British mathematician, starting over 300 years ago, with the rise of Newtonian deterministic and probabilistic paradigms, through Poincaré, up to the present. Sometimes conceptually a little difficult, but this is because it attempts more than other introductory books. This is the clearest and deepest simple presentation I know, and my personal favorite for a presentation at what I conceive as the level of these simulations.*

Schaffer, W. M., (1984); Stretching and folding in lynx fur returns: evidence for a strange attractor in nature? ***American Naturalist***. 124: 798-820. (H) *A more recent analysis of chaos in ecology, using some of the latest analytical tools for detecting chaos.*

Wimsatt, W. C. , (1980); "Randomness and Perceived Randomness in Evolutionary Biology", ***Synthese***, 43: 287-329, 1980. (M) *A general discussion of the problem of predicting patterns of environmental change, of our biases towards simplicity in model-building, and of the significance and robustness of chaos in ecology. (Also a case study of the strategies of robustness analysis advocated below in Levins, 1966, and Wimsatt, 1981.)*

References on Model-Building:

Ashby, W. R. (1956); ***An Introduction to Cybernetics***, New York: Wiley-Interscience, (E/M) *Now long forgotten, this is one of the best introductions to the study of dynamical systems (and theories about their evolution and self-organization) that I know. Well worth a closer look.*

Cairns-Smith, Christopher A. G. (1971); ***The Life-Puzzle***, Toronto: University of Toronto Press (E/M) *Written by a long-time advocate of the now increasingly popular theory that life developed with clay as an essential chemical-mechanical substrate. Included here as one of my favorite examples of the use of simple models to illustrate more complex ideas.*

Cartwright, Nancy. (1983); ***How the Laws of Physics Lie***. London: Oxford University Press. (M/H) *One of the few books by a professional philosopher of science to discuss model-building in any detail--in physics, though many of her points apply to mathematical model-building everywhere.*

Eigen, Manfred and Ruthilde Winkler-Oswatitsch, (1981); ***The Laws of the Game***, New York: Alfred A. Knopf. (E/M) *Co-written by a Nobel prize-winner in physical chemistry and one of the leading theoreticians in the origin of life, this eminently readable and visual book (often in color) illustrates 10 different principles or causal factors important in the origin, evolution, and interaction of life--each with a simple board game readily adaptable to the computer. (A workbook for varieties of cellular automata.) A paradigm of the pedagogical virtues of simple modeling, this book cries out for someone to program the games on the computer. A good project book for computer programmers.*

Jones, B., Sterner, W., & Schank, J. (1994) "Biota: an object-oriented tool for modeling complex ecological systems." ***Mathematical and Computer Modeling***, 20: 31-40.

Kauffman, Stuart, (1972); "Articulation of parts explanations in biology and the rational search for them," in R. C. Buck and R. S. Cohen, eds., ***PSA-1970***, Boston Studies in the Philosophy of Science, Vol. 8., pp. 257-272. (E/M) *A classic exposition of the strategy of constructing and testing models of articulated parts for the explanation of systems composed of these parts by a leading developmental biologist and successful modeler. The strategy is applicable from molecular biology to ecology, and all points in between.*

Kauffman, Stuart A., (1993); ***Self-organization, complexity and the emergence of order***. Oxford: Oxford University Press. (probably M). *An elegant and enormously fruitful case*

study of the use of simple models to study the evolution of complex organization, by one of the best theoreticians of developmental biology in the business. A paradigm of the fruits of building simple models of complex systems, and probably destined to be a major text in a new emerging science of the study of complex systems. See also his article in August, 1991 Scientific American.

Langton, Christopher G., (1989); **Artificial Life**, Santa Fe Institute Studies in the Sciences of Complexity, vol. 6, Redwood City, CA: Addison-Wesley. *Proceedings of the first biennial conference on artificial life. A smorgasbord of approaches to modeling life, with a liberal dose of divergent approaches to model-building. Fascinating reading for anyone interested in model-building. 655 pp. + many illustrations, some in color. Available in paperback for around \$20. A second volume (with associated videotape) has also been published (in 1991), and future volumes are expected. (E/M)*

Levins, Richard, (1966); The strategy of model-building in population biology. **American Scientist**. 54: 421-431. (E/M) *One of the best articles ever written on model-building--a revolutionary piece when it first appeared, the views it espouses have now become more widely accepted.*

Levins, Richard, (1968); **Evolution in Changing Environments**. Princeton: Princeton University Press. *Chapter 1 of this book is a slight extension of the preceding article, and the models in the book exemplify Levins' model-building strategies. (Chapter 1: E; rest of book: M/H).*

Nitecki, M. and A. Hoffmann, Eds., (1987); **Neutral Models in Biology**, New York: Oxford University Press. (E/M) *An excellent collection of papers illustrating the use of "neutral" models--models which are believed false--as "null hypotheses" to test the efficacy of forces left out of the models.*

Petroski, Henry, (1992); **To Engineer Is Human: The Role of Failure in Successful Design**. New York: Random House (Vintage Books, paper), 1992. (E). *I would argue that any evolutionary biologist should take a course in engineering design (and conversely, that any engineer should take a course on biological adaptation), since both are working on complementary parts of the same problem. The engineer has or constructs the design and needs to know how it will perform. The biologist sees some the performance, and knows perhaps aspects of the design, and is trying to figure out what other parts of it are for. Both of them must consider how to evaluate it, and both must evaluate it in part by considering how it performs in comparison with other variations on the design, and revisions on it through time. This is a charming, delightful, and revealing book about the processes of engineering design with equal applicability to the design and analysis of adaptations, and to the process of building and evaluating models. Compare also Simon, 1981, Wimsatt, 1987, (and Jenkin, 1867!) below.*

Schank, J. C. & Alberts, J. R. (1997a) "Self-Organized Huddles of Rat Pups Modeled by Simple rule of Individual Behavior." **Journal of Theoretical Biology** (in press)

Schank, Jeffrey C. , (1991); **Computer Simulation, Model Building, and Experimental Design in Biology**, Ph.D. dissertation, Committee on the Conceptual Foundations of Science, University of Chicago. *A more extended analysis of the preceding, with detailed case*

studies of model-building and experimental design in several areas of biology and an "object oriented" approach to mechanistic explanation.

Schank, Jeffrey C. and Martha K. McClintock, (1990); "Computer simulation of ovarian cycle synchrony in female rats: an object-oriented integration of computer simulation and laboratory experiment." In: B. Schmidt (ed.), **Modeling and Simulation**, The Society for Computer Simulation, San Diego, California: 590-596. *Argues that there are parallels between experimental design and design of a computer program, and for an object-oriented approach to model-building and testing as the way to implement Kauffman's strategy of model-building.*

Schank, Jeffrey C. and M. K. McClintock, (1992), "A Coupled-Oscillator Model of Ovarian-Cycle Synchrony Among Female Rats." **Journal of Theoretical Biology**, 157: 317-362.

Schank, J. C. & C. Twardy (1997) "Mathematical Models." In: Volume 6 of the Cambridge History of Science Series: The Life and Earth Sciences, (Peter J. Bowler and John Pickstone, eds.) Cambridge University Press

Simon, Herbert A., (1981); **The Sciences of the Artificial**. 2nd. ed. Cambridge: MIT Press. (E) *A classic exposition by the father of heuristic programming in artificial intelligence (and Nobelist in economics) of the role of a heuristic vision of man as problem-solver--one of the key assumptions of the approach to model-building advocated here. Chapter 7 is independently one of the most important founding papers on hierarchical organization and complexity.*

Singh, Jagjit, (1968); **Great Ideas of Operations Research**, New York: Dover Publications. (E) *Singh is one of the great (and unfortunately nearly unknown) expositors of the mathematical sciences. In this book, see his Chapters 2 (on statistical description of data and probability distributions--including the binomial) and 7 (on Monte Carlo simulations). He has other excellent books, all published by Dover, Great Ideas of Cybernetics, Information Theory and Artificial Intelligence, and Great Ideas of Modern Mathematics. He manages to give simple and elegant descriptions of complex ideas with clear examples and without distortion as well or better than anyone I have read.*

Sober, Elliot, (1984); **Conceptual Issues in Evolutionary Biology**, Cambridge: MIT Press. (E/M) *This is the best single collection of methodological, conceptual, and philosophical articles on evolutionary biology. Contains Levins, 1966 (#3), Wimsatt, 1976 (#26), and Wimsatt, 1980b (#11). Other clearly relevant articles on model-building or critical discussions of the influence of broad theoretical perspectives which have lead to many models include those of Lewontin (#1) Gould and Lewontin (#15), Oster and Wilson (#16), Maynard Smith (#17), Sober and Lewontin (#13), and Felsenstein (#34).*

Taylor, Peter J., (1985); **Construction and Turnover of Multi-species Communities: A Critique of Approaches to Ecological Complexity**. Ph.D. dissertation. Department of Organismal and Evolutionary Biology. Harvard University. 328 pp. (M/H) *The long first chapter of this is an extended elaboration and critical discussion of various model-building strategies in ecology, with a close examination of Levins' strategies.*

Wimsatt, W. C., (1974); Complexity and Organization, in K. F. Schaffner and R. S. Cohen, eds., **PSA-1972** (Boston Studies in the Philosophy of Science, volume 20), Dordrecht: Reidel, pp. 67-86. (M) *Builds on Kauffman's analysis to consider problems and strategies of problem formulation and solution when we have multiple partial perspectives on the object of study (a biological organism!), no one of which is adequate by itself to solve the problem--a common, if not the universal, situation in biology, psychology, and the social sciences.*

Wimsatt, W. C., (1976); Reductive Explanation--A functional account, in R. S. Cohen, et. al., **PSA-1974**, Dordrecht: Reidel, pp. 671-710. (M) *Argues for a distinction between two different kinds of reduction in science in terms of the roles that they serve in construction, elaboration, and testing of theory. Both types of reduction have been important sources of model-building. The heuristic account of explanation offered there is also closely connected with the views of model-building argued for here. Reprinted in Sober, 1984.*

Wimsatt, W. C., 1980a. (see above)

Wimsatt, W. C., (1980b); Reductionistic research strategies and their biases in the units of selection controversy. in T. Nickles, Ed. **Scientific Discovery-vol. II: Case Studies** Dordrecht: Reidel, pp. 213-259. (M) *A systematic examination of 9 reductionistic model-building strategies, and their strengths (and failures) in a now classic controversy in population genetics. The last two sections expose the role of incorrect blending inheritance-like assumptions in the construction of mathematical models of group selection which led to biased underestimates of its importance. See Wade, 1978. (See also appendix C-III). Reprinted in Sober, 1984.*

Wimsatt, W. C., (1981); Robustness, reliability and overdetermination. in M. Brewer and B. Collins. eds. **Scientific Inquiry and the Social Sciences**. San Francisco: Jossey-Bass, pp. 124-163. (E/M) *A discussion of one of the most important methodological strategies in science--the use of multiple independent means to measure properties, detect entities, evaluate models, or demonstrate conclusions--crucial to the development of trustworthy theories, models, explanations, and measurements. (See also Cartwright, for related views).*

Wimsatt, W. C., (1985a); "Heuristics and the Study of Human Behavior," in D. W. Fiske and R. Shweder, Eds., **Metatheory in Social Science: Pluralisms and Subjectivities**, Chicago: University of Chicago Press, pp. 293-314. (E/M) *A compact discussion of reductionistic problem-solving heuristics and their biases building on Wimsatt, 1980b.*

Wimsatt, W. C., (1986); "Forms of aggregativity." in A. Donagan, N. Perovich, and M. Wedin, Eds. **Human Nature and Natural Knowledge**. Dordrecht: Reidel, pp. 259-293. (M) *An analysis of the conditions under which a system property is or is not analyzable as an aggregate of parts' properties. Relevant to the analysis of reductionistic model-building heuristics and their biases.*

Wimsatt, W. C., (1987); "False Models as Means to Truer Theories," in M. Nitecki and A. Hoffmann, (above), pp. 23-55. *The falsity of a model needn't be a handicap--it can in fact be a powerful tool in the construction of better ones. Documents 12 different strategies for the use of false models, with examples from the history of genetics. (See also appendix B for a more up to date list.)*

References on Blending Inheritance and Binomial Distribution:

The readings in this section assume that the reader is familiar with the basics of classical genetics, such as can be gotten from most genetics textbooks. An excellent book with a historical development of the relevant material, beginning with Darwin (but surprisingly, failing to mention Jenkin) is Moore, 1972. For the binomial distribution, see the Batschelet reference given at the beginning. It is also useful to consult one of the Hartl books, one of the Crow books, or Roughgarden for the discussion of quantitative genetics to provide a perspective on how a Mendelian account can be given for the phenomena which seemed most suggestive to some of the blending inheritance model.

Belew, R. and L. Booker, Eds., (1991); ***Proceedings of the Fourth International Conference on Genetic Algorithms***, San Mateo, CA: Morgan Kaufmann. *The latest current research and applications papers on genetic algorithms and evolutionary strategies—Mendelian and blending modes of inheritance for the population genetics of different possible worlds.*

Darwin, Charles, (1868); ***Variation of Plants and Animals Under Domestication***. See especially Chapter 27, "The provisional hypothesis of pangenesis." Reprinted widely. (E/M)

Eiseley, Loren, (1967); ***Darwin's Century***. Doubleday (Anchor). (E) *A very readable overview of the conceptual changes leading up to Darwin's theory of evolution by natural selection, it also discusses problems with Darwin's theory of heredity. It has a chapter on Mendel, and discusses Jenkins' criticism of blending inheritance.*

Fisher, Ronald A., (1930); ***The Genetical Theory of Natural Selection***, Cambridge: Cambridge University Press. (M/H). *One of the three classic founding texts of the mathematical theory of population genetics by a man who was also probably the leading theoretical statistician of the century. Often deceptively easy to read, though the derivations are usually sketchy. Fisher opens the first chapter with a discussion of the deleterious consequences of blending inheritance and how they are avoided by Mendelian segregation.*

Jenkin, Fleeming, 1867, review of Darwin's *The Origin of Species*, *The North British Review*, reprinted in D. L. Hull, Ed., ***Darwin and His Critics***, Chicago: The University of Chicago Press, 1987. (E/M) *Jenkin's review is the classic critique of blending inheritance and is worth closer attention than it is usually given. Also, in a case where the physicists came out wrong, it provides a clear presentation of Kelvin's arguments that the earth couldn't have been habitable for long enough for life to evolve. These were very influential in leading Darwin to place more weight on the inheritance of acquired characters to "speed up" the rate of evolution. (Kelvin's arguments and time scale had to be completely rewritten with the discovery of nuclear fission and fusion processes, and the inheritance of acquired characters died with the rise of Weismannism.)*

Moore, John A., (1972); ***Heredity and Development***, 2nd ed., Oxford: Oxford University Press. (E/M) *The best short historical presentation of the development of genetics designed for use as a textbook that I know. Starts with Darwin's theory of pangenesis, and goes in detail through the development of cytology and the interacting role of theories of heredity and*

development in localizing the genes on chromosomes. Excellent up through the development of classical genetics at the hands of the Morgan school.

Olby, Robert, (1987); ***The Origins of Mendelism***, 2nd rev. ed., Chicago: The University of Chicago Press. (E/M) *The best history of the origins and rediscovery of Mendelian genetics. Reprints many of the original sources as appendices.*

Wade, M. J., (1978); "A critical review of the models of group selection," ***Quarterly Review of Biology***, pp. 301-314. (M) *A classic review of five simplifying assumptions built into standard models of group selection which strongly bias the case against the efficacy of group selection as a process. Includes the infamous "migrant pool" assumption which reintroduces a particularly strong form of "blending inheritance" at the group level. One of the most centrally important articles in the group selection controversy. Discussed in Wimsatt, 1980b.*

Wimsatt, W.C. , (1981b); "Units of selection and the structure of the multi-level genome." in P. D. Asquith and R. N. Giere. Eds. ***PSA-1980, volume 2***. East Lansing, Michigan: The Philosophy of Science Association, pp. 122-183. (M) *In the last parts extends the discussion of blending inheritance of Wimsatt 1980b to look at higher level genetic models and analogues to Mendelian segregation in population structure.*

References on Developmental Constraints in Evolution and Models of Finite Automata:

Arthur, W. , (1984); ***Mechanisms of Morphological Evolution: A Combined Genetic, Developmental and Ecological Approach***, New York and London: John Wiley and Sons.

Gelfand, A. E. and C. C. Walker, (1984); ***Ensemble Modeling: Inference from Small-Scale Properties to Large-Scale Systems***. Marcel Dekker Inc.: New York and Basel.

Kauffman, Stuart A., (1969); Metabolic stability and epigenesis in randomly constructed genetic nets. ***Journal of Theoretical Biology***, 22: 437-467.

Kauffman, Stuart A. , (1985); Self-Organization, Selective Adaptation and Its Limits: A New Pattern of Inference in Evolution and Development in ***Evolution at a Crossroads: The New Biology and the New Philosophy of Science***, edited by David J. Depew and Bruce H. Weber, MIT press, Cambridge, Massachusetts.

Kauffman, Stuart A., (1992); ***The Origins of Order...*** See under refs. on "model-building."

Schank, J. C. and W. C. Wimsatt, (1988a); Generative entrenchment and evolution, in A. Fine and P. K. Machamer, Eds., ***PSA-1986, volume 2***. East Lansing: The Philosophy of Science Association, pp. 33-60.

Wimsatt, W. C. , (1986a); Developmental constraints, generative entrenchment, and the innate-acquired distinction, in P. W. Bechtel, Ed. ***Integrating Scientific Disciplines***. Dordrecht: Martinus-Nijhoff, pp. 185-208.

Wimsatt, W. C. and J. C. Schank, (1988b); Two constraints on the evolution of complex adaptations and the means for their avoidance, in M. Nitecki, ed., ***Evolutionary Progress***, Chicago: The University of Chicago Press, pp. 231-273.

Index

A

Acquired characters, inheritance of 100, 115
Additive trait 119
Age structure 23
Analysis of periodicities 72
Analytic classification 22, 23
-Solution 24
Analytically intractable 6
Apparently random 15
Apple (🍏) menu 158
Approximations 4, 6, 10, 11, 25
Arthur, W.; 113, 219
Ashby, W. 45, 213
Assortative mating 101, 126
-Assortative by distance 126
Assumptions of a model 178
Attractor 45, 47
-Domain of 59
-Width of 47
-Integer as quantized attractor 52
Autoscaling 40, 42, 139

B

Bad Data 133, 134
Bad parameters 132
Baker, G. 211
Bamboo 26
Bateson, W. 108
Batschelet, E. 207
Beach ball cursor 134
Beats 188
Bechtel, W. 61
Becker, H. 14, 212
Begin r 149
Belew, R. 128, 218
"Best bitmap" option 190
Biases 6, 21, 99, 190
-Sources of 8, 9, Appendix C
-Systematic 17
Bifurcation diagram 17, 32, 34
-Map 45
-Plots 51, 130, 148
Bifurcations 33, 37, 38, 64
Binomial 164
-Coefficients 110, 118
-Distribution 116, 118, 119, 121, 162, 166, 217

- Expansion 110
- Biological assumptions 179
 - Constraints 112
- Biometricians 98
- BioQUEST 7, 22, 97, 99
- Biota™ 7, 21, 22, 23, 38
- Birth and death rates 26
- Bitmap pict 161
- Blending inheritance 1, 2, 3, 12, 97, 101, 102, 103, 104, 110, 115, 116, 118, 119, 120, 125, 126, 127, 129, 134, 144, 161, 164, 166, 217
 - Icon 119
 - Model 101
 - Simulation 119
- Blow up 53
 - Of a bifurcation plot 77
 - Computational time in r vs. N plots 150
- Booker, L. 218
- Boolean automata, networks of 61
- Bossert, W. 211
- Bottleneck 81
- Bottom 140
- Boyce, W. 68, 207
- Boyd, R. 12
- Bracketing 5
- Bugs 147
- Burchfield, J. 115
- Butterfly effect 15

C

- Cairns-Smith, W. 213
- Calibration 24
 - Of pixel width 43
- Calley, J. 80
- Carrying Capacity 11, 135, 138, 139
- Cartwright, N. 213
- Century plant 26
- Change N by [increment] 131
 - R by [increment] 131
 - X-axis by [increment] 139
 - Y-axis by [increment] 139
 - Memory available to application 150
- Chaos, chaotic: 3, 5, 14, 15, 16, 25, 28, 30, 38, 64
 - As a random number generator 48
 - In discrete state systems 50, 59
 - Behavior 5, 25, 29, 31, 210, 211
 - Regime 47, 84, 85
- Charlesworth, D. and B., 113, 208
- Check mark 136
- Chromosomes 107
- Circadian rhythms 8
- Clark, A. 209

Classical genetics 99
 Classifications of deviations 21
 Click the mouse 129
 Close 157
 Colonization 210
 Color 196
 Color codes for generation number 35
 Color pict 34
 Combinatorial complexity 20, 21
 -Explosions 22
 Comparative evaluation of assumptions 181
 Competition 23
 Complexity 17
 Computational roundoff error 43, 49
 Computer simulations 4, 5
 Connectionist networks 61
 Copy 121, 157
 -Plot parameters 161
 Copying and pasting graphs 157
 -Text 161
 Count strips 55
 Crick, F. 4
 Critical density, C 62
 Crow, J. F., 82, 122, 123, 208
 Cursor 129
 Cybernetics 5

D

Darwin, C. 3, 12, 26, 39, 98, 99, 100, 108, 115, 116, 125, 126, 218
 Darwinian evolution 12
 Data, multiple ways of representing 77
 Δt (or "delta-tee") 26
 Densities 27
 Density dependence 27, 178
 Dependence on starting value of systems With Small numbers of states 52
 Deprivation effect 104
 Deterministic models 5
 -And finite cycle length 58
 -Principle of determinism 28
 -Principle of smooth determinism 28
 -Systems 15
 Developmental constraints 113
 -In evolution 219
 Dialogue box 43, 119
Dictyostelium discoideum 8
 Diecious 81
 Difference equations 24, 25
 Differential equations 24, 25, 207
 Dihybrid 105
 DiPrima, R. 68, 207
 Directed graphs 59
 Discrete state automata (modeled using "round N" option) 55, 61

Discussion or thought questions 181
 Do color pict 150
 Dominance 102, 103, 105, 108, 113, 119, 122
 Double click 119, 129
 Dörfler, M. 212
Drosophila 26, 102
 Dynamical systems 212

E

E. coli 26
 East, E. M., 111
 Ecological web 38
 Ecology 1
 Eddington, A. 22
 Edit menu 121, 158
 Effect to cause reasoning 99
 Effective growth rate 27
 Effective population size 80, 82, 130, 154, 194
 -"Exact" and inexact formulae for 83
 Eigen 213
 Eiseley, L. 218
 End r 149
 Environmental canalization 104
 Environmental Decision Making 7
 Environmental plasticity 103
 Enzyme polymorphism 110
 Epidemiology 81
 Epistasis 105, 107, 108, 112, 113
 Equilibrium 30, 33, 44
 -Level of variance 115
 -And stability 209
 Errors 21
 Evolution 208
 Evolutionary constraints 113
 -Strategies 128
 EVOLVE™ 80
 Ewens 82, 208
 EXP 14
 Experimental control 191
 -Design 10, 39, 108
 Exponential growth 11, 14, 17, 23, 24, 25, 26, 27, 36, 38, 39
 -Decay 38
 Extend™ 7
 Extinction 62, 210
 -Half-life 63

F

False theories 99
 -Models 10, 11, 18, 98
 Feigenbaum, M. 15, 212

File menu 157
 Finite difference equations 24, 25
 -Population size 101
 -Automata 219
 Finite-state automata 50
 Fisher, R. A., 113, 115, 218
 Fitness as a property of genes 9
 Flicker fusion 183
 Flow charts 59
 Fluctuations in population size 80
 Focal level of organization 21
 Footprints 2
 Fractals 212
 Fundamental theorem of natural selection 115

G

Galton, F. 115
 Gametic contribution 123
 GCK™ 22, 80, 97, 99
 Gel electrophoresis 110
 Gelfand, N. 219
 Gene control networks 61
 Generating new models 178
 Genetic algorithms 127, 208
 -Canalization 104
 -Developmental programs 113
 -Drift 115
 -Interaction 107
 Genome 113
 Genotype and phenotype 102, 103, 104, 105, 109, 119
 -Environment effects 108
 -Phenotype distinction 12, 103
 -Phenotype mappings 105, 107, 113
 -Phenotype space 106, 109, 111, 112
 Genotypic space 105, 107
 Gens plotted 83
 Get Info 150
 Get N 41, 143, 150, 166
 Gliesh, J. 14, 212
 Goaway boxes 157
 Goldberg, B., 128, 208
 Gollub, J. 211
 Graph coordinates option 35, 40, 43, 54, 140, 191
 -Dimensions 43
 -Grid 40, 42, 46
 -Allows local comparison 185
 -Lines in only one direction 139
 -Parameters 41
 Grids menu 42
 Group Selection 12
 -Efficacy of 127
 Growth equations 14

-Rate 11, 24, 27, 38, 44

H

Haldane 115
 Half-life, extinction 63
 Hardy-Weinberg equilibrium 101, 115, 123
 Hartl, D. 82, 115, 122, 123, 208, 209
 Hassell, R. 181
 Help facility 144
 -File 145
 Heterozygote 109
 Heuristic 48, 178
 Homozygotes 109
 Hereditary material 101
 Heterozygotes 115
 Heuristics 5, 183
 Hierarchy, 'kind of' 19, 20
 -'Part of' 19, 20
 History of science 98
 Hoffmann, A. 98, 214
 Hofstadter, D. 212
 Holland, J. 127
 Hubby, J. 110, 209
 Hull, D. 115
 Hypothesize causal relationships 183

I

Icon 129
 Idealizations 4, 6, 12, 24
 Idealized assumptions 125
 Ideal organisms 97
 Illusions 183
 Imagewriter™ 161
 Independent assortment 107, 108
 -Events, probability of 22
 Infinite accuracy 16
 Infinite sensitivity of chaotic behavior 16
 Infinitely large population 126
 Initial N 150
 Initialize & Plot 119, 130, 139, 159
 Intermediacy 103, 104, 105, 108, 109, 110, 122
 Interference effects 190
 interrupt a simulation 134
 Intrinsic growth rate 26

J

Jacquard, A 111, 209
 Jenkin, F. 115, 116, 125, 126, 218
 Jungck, J. 12, 80

K

K & C lines 138
 Kauffman 45, 50, 61, 214, 219
 Kelvin 115
 Kimura 82, 208
 'Kind of hierarchy 19, 20
 Kingsland, S., 14, 25, 209

L

Laboratory experiment 5, 10
 Laboratory biology 10
 Lambda, λ 72
 Langton, C. 214
 Law of dominance 108
 Laws and mechanisms 97
 -Of Mendelian genetics 99
 -Of nature 98
 Levels of abstraction 18, 19, 21
 -Of organization 8, 9, 18, 19, 184
 Levins, R. 37, 182, 213, 214
 Lewontin, R., 110, 113, 209
 Limitations of the models 178
 Limiting behavior 27
 -Cases 4
 -Value 28
 Linear extrapolation 26, 27
 Lines 139
 Literally false models 12
 Local analysis 185
 Local to global comparisons 185
 Localize objects 184
 Logistic automata 60, 61
 Logistic growth 1, 2, 3, 11, 14, 15, 17, 23-27, 29, 30, 36, 39, 72, 97, 129, 134, 144, 149 155, 161, 162, 166, 167
 -Equations 130
Logos 25
 LOGS 14
 Lorenz, E. 15
 Loss of genetic variability 81
 Lotka, A. 209
 Lotka-Volterra equations 23

M

MacArthur 210
 Macro-similarities 193
 Major causal factors 6, 8, 9, 11
 Malthus 39
 Mandelbrot set 17
 Mappings between genotype and phenotype -104, 112
 -Arbitrary mapping 105

Mass-extinctions 210
 Mathematica 24
 Mathematical assumptions 179
 -Models 4, 11, 24
 -Artifacts 38, 83
 -Background 207
 Maximum value plotted 48, 139
 Maximum yield 69
 Maxwell, J. 115
 May, R. 14, 15, 30, 62, 212
 McClintock, M. 215
 MDLS 14, 15, 17, 18, 23, 24, 26, 27, 36, 37, 38, 39 (See "minimum density limited logistic")
 Measurement assumptions and problems 180
 Mechanisms 6, 8, 9, 10, 97, 98, 99, 101
 Melanin production in humans 111
 Mendel 99, 108, 109, 122, 125
 Mendelian 110
 Mendelian dominance 102, 105
 -Factors 107
 -Genetics 3, 97, 100, 103, 123
 -Inheritance, 12, 101, 102, 104
 -Laws of 12
 -Population 101
 -Segregation 115
 Mendelism 100, 108, 115
 μ GCKTM (or "micro"-GCK) 97
 Migration 101
 "Migrant pool" assumption 127
 Mimicry 113, 208
 Minimum density 138, 139
 -Threshold 36
 Minimum density limited logistic 11, 14, 17, 36, 39, 97, 129, 131
 Monoecious 81
 Multidimensional data 184
 Model building 39, 208, 213
 -Strategies 24
 -Models and theories 3
 -How different formulations of can have different consequences 50
 -Increase realism of 178
 Model organisms 10
 Models in population genetics 210
 Moire or "crystal dislocation" effects 43, 187, 188
 Monoculture 81
 Monohybrid 106
 Monte Carlo simulation 14, 23, 46, 47, 62
 Moore, J. 100, 218
 Motion detection 183
 -Pictures 159
 Multi-species interactions 23
 Multi-level character of biological systems 9
 Multiple ways of representing data 77
 Multiplicative rule for probability, 22
 Mutation 113, 126

- 'Individual' variations 100
- Sports 100

N

- N vs. r plot 32, 34, 47, 73, 161 (see also r vs. N plot)
- N vs. t plot 75, 130
- N(t) vs. N(t+m) or "phase" plots 14, 34, 36, 64, 72, 75, 82, 130, 136, 159
- Near decomposability 22, 39
- Networks of Boolean automata 61
- Neutralist-selectionist" controversy 209
- Newton's rings 188
- Nitecki 98, 214
- N vs. t plot 34, 82, 136
- Non-linear 15
- Non-linear differential equations 15
- Normal distribution 127
- Number of allowable states in the system 58
- Numerical precision 49
- Numerical solution 24

O

- Odum, E. 7
- Olby, R. 108, 219
- Opening PICT files 156
- Opening Text files as Notebooks 157
- Oscillations 44
- Overdominance 105
- Overlapping generations 101, 103, 126, 178
- Overshoots 42, 44
- Oversimplifying 12

P

- Panmixia 117, 126
- Part of hierarchy 19, 20
- Particulate inheritance 122, 125
- Paste 121, 158
- Patchy environments 210
- Pattern-detection 183 (see also perception)
- Patterns 164
- Patterns and colors 142
- Penetrance 108
- Perception of objects 183 (see also visualization)
- Period doubling bifurcations 30, 33
- Personal help file 145
- Perspectives 184
- Persuasion 12, 178
- Petroski, H. 214
- Phase shift 197
- Phase lags 64
 - Plots 14, 34, 36, 64, 72, 75, 82, 136, 159

- And abstract art 138
- relations 187
- shifts 195
- Phenotypic mappings 105
 - Space 105, 107
 - Switching 103
- Pict 159, 160, 161
- Pixel roundoff errors 40, 41, 42, 43, 49, 63
- Pixels 140
- Plot delay 183
 - K & C 166
 - Window 144
- Plotting order 196
 - Speed 142
- Point grabber 41, 42, 48, 49, 63
 - Size 40, 141, 164
- Poisson distribution 127
- Polygenic 108
- Population crashes 62
 - Density 137, 149
 - Genetics 1, 208
 - Growth 14
 - Growth Models 208
 - Size 24, 26, 28, 42
 - Structure 101
- Postscript™ 160
- Predation 23
- Price, F., 80
- Printing Windows 157
 - "Smooth bitmap" option 190
- Probability of independent events,
 - Multiplicative rule for 22
- 3-P's 12, 116, 178
- Problem-posing 12, 178
- Problem-solving 12, 39, 178
 - Strategy 39, Appendix C
- 4th P—exploring Possibilities 116
- Provine, W. 80, 111, 210
- Punnett Square 122, 123, 125

Q

- Qualitative analysis 16
- Quantized escape orbits 194
- Quasi-independence 22, 113
- QuickDraw™ 160

R

- r vs Ne(t) plots: 14, 82, 83, 84, 130, 152, 154, 159 (see also "effective population size")
- r vs. N or "bifurcation" plots 14, 33, 36, 82, 83, 84, 130, 148, 149, 152, 153, 159
- r-end 149
- Radioactivity 115

Random number generator 46
 Randomness 213
 Raup, D. 210
 Realism 36
 Recessive 102, 108, 113, 122
 Recombination 104, 123
 Record coordinates 34, 41, 43, 143, 146, 150
 ReDraw option and plot speed 78, 142, 183, 196
 Reductionistic research strategies 39
 Representative sampling 22, 23
 Richerson, P. 12
 Robust 184
 Roughgarden, J. 63, 82, 210
 Round N option (see also discrete state models) 49, 50, 52, 53, 150
 Roundoff error 49, 131, 187, 197, 199
 -And finite automata 131
 Roundoff quantum 199
 Run Record utility 34, 41, 146, 150

S

Save PICT 34
 Saving Text files 156
 Scaling Graphs 134
 Schaffer, W. 31, 81, 213
 Schank, J. 113, 215, 220
 Schroeder, W. 212
 Scrapbook 53, 121, 158, 159
 Segregation 108
 Self-organization 214
 Self-similarity 77
 Sensitivity to initial conditions and chaos 15, 47, 78
 Sex ratio 81
 Show Grid 139, 166
 Signatures 2
 Simon, H. A. 39, 215
 Simplifications 6, 8, 10, 11, 12, 178
 Simulation 7, 11, 97, 116
 Simulation menu 34, 119, 121, 136, (Figure A8) 138, 158, 162, 164, 165, 166
 -Initialize and Plot 134
 Singh, J. 215
 Size scales 184
 Size vs. time 136
 Sober, E. 216
 Spanning trajectories 48
 Spatial senses, vision, hearing, and touch 184
 Species, subdivision of 82
 Stability analysis 42, 207
 -Line 64, 66, 137
 -Of an equilibrium 69
 -Of phenotypic traits 104
 Stabilizing selection 126
 Stable and unstable cycles 32

Stable equilibrium 29, 30
 Standard deviations 166
 Start Gen 32, 83
 Starting generation 149
 State transitions 59
 State-transition diagrams 60
 Statistical patterns in chaotic dynamics 79
 Stella™ 7
 Stewart, J., 12, 14, 15, 68, 212
 Strategies:
 -Divide and conquer 39
 -Model-building, Appendix C
 Strip charts 91
 Strobe frequency 188
 Successional reduction 28
 Sum of binomial 166
 Sum of Variance 165
 SumOfBinomial 121
 SumOfVar 121
 Symbiosis 23

T

t vs. N plots (see N vs. t plots)
 Taylor , P. 182, 216
 Template matching 72, 184
 -Capabilities of our visual system 53
 Teratogenesis 104
 Trajectories, merging of and determinism 58
 -Maximum length of 58
 Theoretical perspectives 9
 Theory of the instrument 9
 Thinner appearance in the "bush" of the tree 52
 -"twig" 33
 Thompson, W. 115
 Time lags 37, 38
 Transient behavior 33
 Transition to chaos 61
 Tufte 184, 185, 207

U

Ultraviolet radiation 111
 Using lines to give transition directions 137

V

Variance 121, 164
 Vernier calipers 187
 Vernier interpolation 192, 193, 194
 Visual information about motion 183
 Visual processing of motion detection and integration 195
 Visualization 1, 2, 5, 11, 12, 16, 207

-Use of "lines option" is used to aid 56

W

Wade, M. 127, 211, 219

Wahlund's principle 101

Walker, C. C., 219

Watson, J. 4

Wiener, N. 5

Williams, G. C. 107

Wilson, E. O., 210, 211

Wimsatt, W. 12, 18, 28, 31, 39, 48, 96, 113, 115, 182, 184, 208, 211, 213, 216, 217, 219, 220

Windows 144

Winkler-Oswatitsch, R. 213

Wright, S. 80, 84, 113, 115, 210, 211