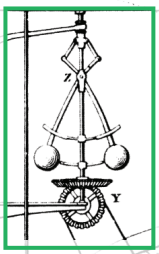
Physical Models of Living Systems

Probability, Simulation, Dynamics

Second edition



Philip Nelson

University of Pennsylvania
with the assistance of Sarina Bromberg,
Ann Hermundstad, Keith Kroma-Wiley,
and Jason Prentice

Brief contents

PART I First Steps

Chapter 1 | Virus Dynamics 6

Chapter 2 | Physics and Biology 25

PART II Randomness in Biology

Chapter 3 | Discrete Randomness 33

Chapter 4 | Some Useful Discrete Distributions 71

Chapter 5 | Continuous Distributions 105

Chapter 6 | Random Walks on an Energy Landscape 137

Chapter 7 | Model Selection and Parameter Estimation 165

Chapter 8 | *Excursion:* Single Particle Reconstruction in Cryo-electron Microscopy 199

Chapter 9 | Poisson Processes and Their Simulation 230

Chapter 10 | Randomness in Cellular Processes 259

PART III Feedback Control

Chapter 11 | Negative Feedback Control 293

Chapter 12 | Positive Feedback, Epidemics, and Genetic Switches 328

Chapter 13 | Cellular Oscillators 376

PART IV Nonlinear, Stochastic Dynamical Systems

Chapter 14 | Demographic Variation in Epidemic Spread 401

Chapter 15 | Bet-Hedging Via Stochastic, Excitable Dynamics 411

Epilogue 426

Appendix A | Global List of Symbols 429

Appendix B | Units and Dimensional Analysis 435

Appendix C | Numerical Values 441

Acknowledgments 442

Credits 445

Bibliography 447

Web resources

The book's Web site (`www.physics.upenn.edu/biophys/PMLS2e/`) contains links to the following resources in the Student section:

- *Datasets* contains datasets that are used in the problems. In the text, these are cited like this: Dataset 1, with numbers keyed to the list on the Web site.
- *Media* gives links to external media (graphics, audio, and video). In the text, these are cited like this: Media 1, with numbers keyed to the list on the Web site.
- *Errata* will appear as needed.

Detailed contents

Problems 1
To the student xvi
To the instructor xxi
Prologue: A breakthrough on HIV 1

PART I First Steps

Chapter 1 | Virus Dynamics 6

1.1 First Signpost: *Fitting* 6
1.2 Modeling the Course of HIV Infection 7
1.2.1 Biological background 7
1.2.2 A semilog graph can disclose an exponential relation in data 8
1.2.3 Physical modeling begins by identifying the key actors and their main interactions 10
1.2.4 Mathematical analysis yields a family of predicted behaviors 11
1.2.5 Most models must be fitted to data 13
1.2.6 Overconstraint versus overfitting 14
1.3 Just a Few Words About Modeling 15
Big Picture 16
Key Formulas 17
Track 2 19
1.2.4′ Exit from the latency period via proliferation of mutants 19
1.2.6′a Informal criterion for a falsifiable prediction 19
1.2.6′b More realistic viral dynamics models account for partially-effective drugs 19

1.2.6′c Latently-infected cells make eradication of HIV difficult 20
Problems 21

Chapter 2 | Physics and Biology 25
2.1 Signpost: *Inference* 25
2.2 The Intersection 26
2.3 Dimensional Analysis 27
Big Picture 27
Key Formulas 28
Problems 29

PART II Randomness in Biology

Chapter 3 | Discrete Randomness 33
3.1 Signpost: *Distributions* 33
3.2 Avatars of Randomness 33
3.2.1 Five iconic examples illustrate the concept of randomness 33
3.2.2 Computer simulation of a random system 38
3.2.3 Biological and biochemical examples 38
3.2.4 False patterns: Clusters in epidemiology 39
3.3 Probability Distribution of a Discrete Random System 39
3.3.1 A probability distribution describes to what extent a random system is, and is not, predictable 39
3.3.2 A random variable associates numerical values to points in sample space 41
3.3.3 The addition rule 42
3.3.4 The negation rule 43
3.4 Conditional Probability 43
3.4.1 Conditional probability is the ratio of two probabilities 43
3.4.2 Independent events and the product rule 44
3.4.3 Crib death and the prosecutor's fallacy 45
3.4.4 The Geometric distribution describes the waiting times for success in a series of independent trials 46
3.4.5 Joint distributions 47
3.4.6 The proper interpretation of medical tests requires an understanding of conditional probability 49
3.4.7 The Bayes formula streamlines calculations involving conditional probability 51
3.5 Expectation and Other Moments 52
3.5.1 The expectation expresses the average of a random variable over many trials 53
3.5.2 The variance of a random variable is one measure of its fluctuation 54
3.5.3 The standard error of the mean improves with increasing sample size 56
3.5.4 Correlation and covariance 57
Big Picture 59
Key Formulas 59
Track 2 61
3.4.2′a Negation rule extended to conditional probabilities 61
3.4.2′b Product rule subject to a condition 61
3.4.2′c Independence property subject to to a condition 61
3.4.7′ Bayes formula subject to a condition 61
3.5.2′a Skewness and kurtosis 62
3.5.2′b Limitations of the correlation coefficient 62
Problems 63

Chapter 4 | Some Useful Discrete Distributions 71
4.1 Signpost: *Simulation* 71
4.2 Binomial Distribution 71
4.2.1 Drawing a sample from a reservoir can be modeled via Bernoulli trials 71
4.2.2 The sum of several Bernoulli trials follows a Binomial distribution 72
4.2.3 Expectation and variance partially characterize a random variable 73
4.2.4 How to count the number of fluorescent molecules in a cell 74
4.2.5 Computer simulation 75
4.3 Poisson Distribution 76
4.3.1 The Binomial distribution becomes simpler in the limit of sampling from an infinite reservoir 76
4.3.2 The sum of many Bernoulli trials, each with low probability, follows a Poisson distribution 77
4.3.3 Computer simulation 80
4.3.4 Determination of single ion-channel conductance 80
4.3.5 The Poisson distribution behaves simply under convolution 82
4.4 The Jackpot Distribution and Bacterial Genetics 83
4.4.1 It matters 83
4.4.2 Unreproducible experimental data may nevertheless contain an important message 84
4.4.3 Two competing models for the emergence of resistance 86
4.4.4 The Luria-Delbrück hypothesis makes testable predictions for the distribution of survivor counts 87
4.4.5 Perspective 90
Big Picture 90
Key Formulas 91
Track 2 93
4.2.4′ Asymmetric cell division 93
4.4.2′ No lysogenic state 93
4.4.3′ More about the Luria-Delbrück experiment 93
4.4.5′a Analytical approaches to the Luria-Delbrück calculation 93
4.4.5′b Other genetic mechanisms 93
4.4.5′c Epigenetic mechanisms 94
4.4.5′d Direct confirmation of the Luria-Delbrück hypothesis 95
Problems 96

Chapter 5 | Continuous Distributions 105
5.1 Signpost: *Long tails* 105
5.2 Probability Density Function 105
5.2.1 The definition of a probability distribution must be modified for a continuous random variable 105
5.2.2 Three key examples: Uniform, Gaussian, and Cauchy distributions 107
5.2.3 Joint distributions of continuous random variables 109
5.2.4 Expectation and variance of the three example distributions 110
5.2.5 Convolution and mixture distributions 112
5.2.6 Transformation of a probability density function 114
5.2.7 Computer simulation 116
5.3 More about Gaussian Distributions 116
5.3.1 Gaussian distributions also arise as limits of Binomial 116
5.3.2 The central limit theorem explains the ubiquity of Gaussian distributions 118
5.3.3 When to use/not use a Gaussian 119
5.3.4 The diffusion law 120
5.4 More About Long-Tail Distributions 122
5.4.1 Many complex systems generate long-tail distributions 122
5.4.2 A log-log graph can disclose a power-law relation in data 123

Big Picture 124
Key Formulas 124
Track 2 126
5.2.1′ Notation used in mathematical literature 126
5.2.4′ Interquartile range 126
5.4′a Terminology 127
5.4′b The movements of stock prices 127
5.4′c Lévy flights 129
Problems 130

Chapter 6 | Random Walks on an Energy Landscape 137
6.1 Signpost: *First passage* 137
6.2 One particle 137
6.2.1 Review: The free random walk as a model for molecular diffusion 137
6.2.2 The random walk with drift as a model for driven motion 138
6.3 Random Walk in a Trap 140
6.3.1 A force field can be modeled via a position-dependent stepping probability 140
6.3.2 The Boltzmann distribution emerges after equilibration 140
6.4 Escape over a Barrier 142
6.4.1 First-passage time gives a quantitative, single-molecule replacement for the notion of rate 142
6.4.2 In simple situations, pulling speeds up unbinding 145
6.5 Bonds. Catch Bonds. 146
6.5.1 More complex molecular pairs can have multiple unbinding pathways 146
6.5.2 Single-molecule experiments yield the entire distribution of unbinding times 148
6.5.3 Implementation in biomolecules 149
6.6 Biophysical Implications 150
6.6.1 Immune cell activation involves catch bonding 150
6.6.2 Leukocyte rolling also relies on catch bonds 152
6.6.3 Cell adhesion complex formation 154
Big Picture 155
Key Formulas 155
Track 2 157
6.2.2′ Low Reynolds number motion 157
6.4′a Mean first-passage time from simulation data 157
6.4′b Kramers approach 158
6.4.1′ A physical origin for a Power-law distribution 158
6.6.1′ More about T cell activation 159
Problems 160

Chapter 7 | Model Selection and Parameter Estimation 165
7.1 Signpost: *Likelihood* 165
7.2 Maximum Likelihood 166
7.2.1 How good is your model? 166
7.2.2 Decisions in an uncertain world 167
7.2.3 The Bayes formula gives a consistent approach to updating our degree of belief in the light of new data 168
7.2.4 A pragmatic approach to likelihood 169
7.3 Parameter Estimation 170
7.3.1 Intuition 171
7.3.2 The maximally likely value for a model parameter can be computed on the basis of a finite dataset 171

7.3.3 The credible interval expresses a range of parameter values consistent with the available data 172
7.3.4 Summary 174
7.4 Likelihood Analysis of the Luria-Delbrück Experiment 175
7.5 Localization microscopy 175
7.5.1 On seeing 175
7.5.2 Fluorescence imaging at one nanometer accuracy 175
7.5.3 Complete images: PALM/FPALM/STORM 178
7.6 An Extension of Maximum Likelihood Lets Us Infer Functional Relationships from Data 180
Big Picture 183
Key Formulas 183
Track 2 185
7.2.1′ Cross-validation 185
7.2.3′a Objective priors 185
7.2.4′a Binning data reduces its information content 185
7.2.4′b Odds 186
7.3.2′a The role of idealized distribution functions 186
7.3.2′b Improved estimator 187
7.3.3′a Credible interval for the expectation of Gaussian-distributed data 187
7.3.3′b Confidence intervals in classical statistics 188
7.3.3′c Asymmetric and multivariate credible intervals 189
7.5.2′ More about FIONA 190
7.5.3′ More about superresolution 190
7.6′ What to do when data points are correlated 190
Problems 193

Chapter 8 | *Excursion:* Single Particle Reconstruction in Cryo-electron Microscopy 199
8.1 Signpost: *Alignment* 199
8.2 A Powerful New Tool 199
8.2.1 The coronavirus spike protein is a key therapeutic target 199
8.2.2 Many macromolecules of interest cannot be crystallized 201
8.3 Extracting a Signal from Very Noisy Data 203
8.4 Cross-Correlation 205
8.4.1 A peak in cross-correlation identifies the best match of two signals 205
8.4.2 Numerical implementation 205
8.5 Approach to 1D Alignment via Cross-correlation 207
8.6 Improved Approach via Maximum Posterior 207
8.6.1 To extract an image, marginalize latent shift variables 208
8.6.2 The cross-correlation again enters, this time in a weighting function 212
8.6.3 Numerical implementation of convolution 213
8.6.4 Iterative construction of the inferred image 213
8.6.5 Summary and results in 1D 214
8.7 Approach to the 2D Problem Via Cross-Correlation 214
8.8 Improved Approach Via Maximum Posterior, 2D 217
8.8.1 To extract a 2D image, marginalize latent shift and rotation variables 217
8.8.2 Summary and results in 2D 219
Big Picture 221
Key Formulas 221
Track 2 223
8.2.2′ Other aspects of the cryo-EM revolution 223

8.6′a Estimating other parameters 223
8.6′b Expectation–maximization algorithm 224
8.8′a Contrast transfer function 224
8.8′b Details of posterior maximization in 2D 224
8.8′c Fourier methods 226
8.8′d Heterogeneous samples 227
Problems 228

Chapter 9 | Poisson Processes and Their Simulation 230
9.1 Signpost: *Mean rate* 230
9.2 The Kinetics of a Single-Molecule Machine 230
9.3 Geometric Distribution Revisited 232
9.4 A Poisson Process can be Defined as a Continuous-time Limit of Repeated Bernoulli Trials 234
9.4.1 Waiting times are Exponentially distributed 235
9.4.2 Counts are Poisson distributed 237
9.5 Useful Properties of Poisson Processes 239
9.5.1 Thinning a Poisson process results in another Poisson process 239
9.5.2 Merging two Poisson processes also results in another Poisson process 239
9.5.3 Significance of thinning and merging properties 241
9.6 More Examples 242
9.6.1 Enzyme turnovers at low concentration follow a Poisson process 242
9.6.2 Neurotransmitter release 242
9.7 Multistep Processes and Convolution 245
9.7.1 Myosin-V's stepping times display a dual character 245
9.7.2 Relative standard deviation can be used to reveal substeps in a kinetic scheme 247
9.8 Computer Simulation 247
9.8.1 Simple Poisson process 247
9.8.2 Poisson processes with multiple event types 248
Big Picture 248
Key Formulas 249
Track 2 250
9.2′ More about motor stepping 250
9.6.1′a More detailed models of enzyme turnovers 250
9.6.1′b More detailed models of photon arrivals 250
9.7.2′ The randomness parameter 251
Problems 252

Chapter 10 | Randomness in Cellular Processes 259
10.1 Signpost: *Inventories* 259
10.2 Random Walks and Beyond 259
10.2.1 Situations studied so far 259
10.2.2 A more realistic model of Brownian motion includes both random step times and random step directions 260
10.3 Molecular Population Dynamics As a Markov Process 261
10.3.1 The birth-death process describes population fluctuations of a chemical species in a cell 262
10.3.2 In the continuous, deterministic approximation, a birth-death process approaches a steady population level 264
10.3.3 Stochastic simulation via Gillespie's direct algorithm 265
10.3.4 The birth-death process undergoes fluctuations in its "steady" state 267
10.4 Gene Expression 268
10.4.1 Exact mRNA populations can be monitored in living cells 268

10.4.2 mRNA is produced in bursts of transcription 271
10.4.3 Perspective 274
10.4.4 Vista: Randomness in protein production 275
10.5 Intracellular Mechanics 275
10.5.1 Kinetochores form catchbonds to microtubules 275
10.5.2 Kinesin stepping depends on applied load and supplied ATP 276
Big Picture 278
Key Formulas 278
Track 2 280
10.3.4′a The master equation 280
10.3.4′b Time development of the mean 282
10.3.4′c Steady-state distribution 282
10.3.4′d A caveat 283
10.4′ More about gene expression 283
10.4.2′a The role of cell division 283
10.4.2′b Stochastic simulation of a transcriptional bursting experiment 284
10.4.2′c Analytical results on the bursting process 285
10.4.3′a Mechanism of bursting 286
10.4.3′b From single cells to single genes 286
Problems 287

PART III Feedback Control

Chapter 11 | Negative Feedback Control 293
11.1 Signpost: *Fixed points* 293
11.2 Mechanical Feedback and Phase Portraits 294
11.2.1 The problem of cellular homeostasis 294
11.2.2 Negative feedback can bring a system to a stable setpoint and hold it there 294
11.3 Wetware Available in Cells 296
11.3.1 Many cellular state variables can be regarded as inventories 296
11.3.2 The birth-death process includes a simple form of feedback 297
11.3.3 Cells can control enzyme activities via allosteric modulation 297
11.3.4 Transcription factors can control a gene's activity 298
11.3.5 Artificial control modules can be installed in more complex organisms 301
11.4 Dynamics of Molecular Inventories 302
11.4.1 Transcription factors stick to DNA by the collective effect of many weak interactions 302
11.4.2 The probability of binding is controlled by two rate constants 303
11.4.3 The repressor binding curve can be summarized by its equilibrium constant and cooperativity parameter 304
11.4.4 The gene regulation function quantifies the response of a gene to a transcription factor 307
11.4.5 Dilution and clearance oppose gene transcription 308
11.5 Synthetic Biology 309
11.5.1 Network diagrams 309
11.5.2 Negative feedback can stabilize a molecule inventory, mitigating cellular randomness 310
11.5.3 A quantitative comparison of regulated- and unregulated-gene homeostasis 312
11.6 The *trp* Operon is a Natural Example of Negative Feedback Control 315
11.7 Some Systems Overshoot on the Way to Their Stable Fixed Point 316
11.7.1 Two-dimensional phase portraits 317
Big Picture 319
Key Formulas 319

Track 2 321
11.3.1′a Contrast to electronic circuits 321
11.3.1′b Permeability 321
11.3.3′ Other control mechanisms 321
11.3.4′a More about transcription in bacteria 322
11.3.4′b More about activators 322
11.3.5′ Gene regulation in eukaryotes 323
11.4.4′a More general gene regulation functions 323
11.4.4′b Cell cycle effects 323
11.5.1′a Simplifying approximations 324
11.5.1′b The Systems Biology Graphical Notation 324
11.5.3′ Exact solution 324
11.7.1′ Taxonomy of fixed points 325
Problems 326

Chapter 12 | Positive Feedback, Epidemics, and Genetic Switches 328
12.1 Signpost: *Watersheds* 328
12.2 Positive Feedback 328
12.2.1 Runaway growth 328
12.2.2 Limited resources: the chemostat 329
12.2.3 Nondimensionalization can reduce clutter in complicated equations 331
12.2.4 Perspective 334
12.3 The Spread of Epidemics 334
12.3.1 No immunity: The SI model 334
12.3.2 Permanent immunity: The SIR model 336
12.3.3 Temporary immunity: The SIRS model 338
12.4 Bacteria Have Behavior 339
12.4.1 Cells can sense their internal state and generate switch-like responses 339
12.4.2 Cells can sense their external environment and integrate it with internal state information 341
12.4.3 Novick and Weiner characterized induction at the single-cell level 341
12.4.4 Quantitative prediction for Novick-Weiner experiment 344
12.5 Positive Feedback Can Lead to Bistability 348
12.5.1 Mechanical toggle 348
12.5.2 Electrical toggles and neural excitability 350
12.5.3 A 2D phase portrait can be partitioned by a separatrix 351
12.6 A Synthetic Toggle Switch Network in *E. coli* 352
12.6.1 Two mutually repressing genes can create a toggle 352
12.6.2 The toggle can be reset by pushing it through a bifurcation 354
12.6.3 Perspective 356
12.7 Natural Examples of Switches 357
12.7.1 The *lac* switch 357
12.7.2 The *lambda* switch 361
Big Picture 362
Key Formulas 363
Track 2 365
12.3.3′ Other epidemic models 365
12.4.3′ More details about the Novick-Weiner experiments 365
12.4.4′ Mosaicism 365
12.6.1′a A compound operator can implement more complex logic 366
12.6.1′b A single-gene toggle 367

12.6.2′ Adiabatic approximation 371
12.7.1′ DNA looping 372
Problems 374

Chapter 13 | Cellular Oscillators 376
13.1 Signpost: *Limit cycles* 376
13.2 Many Individual Cells Have Circadian Clocks 376
13.3 Synthetic Oscillators in Cells 377
13.3.1 Negative feedback with delay can give oscillatory behavior 377
13.3.2 Three repressors in a ring arrangement can also oscillate 378
13.4 Mechanical Clocks Can be Represented by Their Phase Portraits 379
13.4.1 Adding a toggle to a negative feedback loop can improve its performance 379
13.4.2 Synthetic-biology realization of the relaxation oscillator 384
13.5 Natural Oscillators 384
13.5.1 Protein circuits 385
13.5.2 The mitotic clock in *Xenopus laevis* 386
Big Picture 389
Key Formulas 390
Track 2 392
13.4.1′a Attractors in phase space 392
13.4.1′b Deterministic chaos 392
13.4.1′c Linear stability analysis 392
13.4.1′d Noise-induced oscillation 394
13.5.2′ Analysis of *Xenopus* mitotic oscillator 394
Problems 397

PART IV Nonlinear, Stochastic Dynamical Systems

Chapter 14 | Demographic Variation in Epidemic Spread 401
14.1 Signpost: *Superspreaders* 401
14.2 Stochastic SIR Model 402
14.2.1 Some outbreaks fizzle out 402
14.2.2 Infections attributable to a single individual follow a Geometric distribution in the SIR model 403
14.3 Superspreading via Overdispersion 405
14.3.1 For some illnesses, infectivity is highly variable 405
14.3.2 Even a tiny minority of superspreaders has a big impact 406
Big Picture 406
Problems 409

Chapter 15 | Bet-Hedging Via Stochastic, Excitable Dynamics 411
15.1 Signpost: *Excitability* 411
15.2 The Strategy of Bet Hedging 412
15.2.1 Variety is more than the spice of life 412
15.2.2 Bacterial persisters arise randomly and spontaneously 412
15.2.3 *Bacillus subtilis* has several stress-response strategies 413
15.3 The Competence Switch 414
15.3.1 Competence is an individual, all-or-none decision made at a randomly chosen time 414

15.3.2 The competence network has a single, stable fixed point in the continuous, deterministic approximation 415
15.3.3 Random molecular fluctuations initiate a stereotyped program 419
15.3.4 The ComK/ComS model makes testable predictions 421
15.4 Vistas 422
Big Picture 422
Key Formulas 422
Problems 424

Epilogue 426

Appendix A | Global List of Symbols 429
A.1 Mathematical Notation 429
A.2 Graphical Notation 430
A.2.1 Phase portraits 430
A.2.2 Network diagrams 430
A.3 Named Quantities 431

Appendix B | Units and Dimensional Analysis 435
B.1 Base Units 436
B.2 Dimensions Versus Units 436
B.3 Dimensionless Quantities 438
B.4 About Graphs 438
B.4.1 Arbitrary units 439
B.5 About Angles 439
B.6 Payoff 439

Appendix C | Numerical Values 441
C.1 Fundamental Constants 441

Acknowledgments 442

Credits 445

Bibliography 447
Index 460