Gareth Loy
foreword by John Chowning

the mathematical foundations of music

Musimathics

Volume 2

Musimathics

Musimathics

The Mathematical Foundations of Music

Volume 2

Gareth Loy

Foreword by John Chowning

The MIT Press
Cambridge, Massachusetts
London, England

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please e-mail <special_sales@mitpress.mit.edu> or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

This book was set in Times Roman by ICC Macmillan Inc.
Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Loy, D. Gareth.
Musimathics : the mathematical foundations of music / Gareth Loy.
p. cm.
Includes bibliographical references and indexes.
ISBN 978-0-262-12285-6 (v. 2: alk. paper)—ISBN 978-0-262-12282-5 (v. 1: alk. paper)
1. Music in mathematics education. 2. Mathematics—Study and teaching. 3. Music theory—Mathematics.
4. Music—Acoustics and physics. 5. Composition (Music). I. Title.

QA19.M87L69 2007
781.2—dc22

2005051090

10 9 8 7 6 5 4 3 2 1

This book is dedicated to the memory of
Maxine Flora Reitz Loy
Harold Amos Loy
Thomas Harold Loy

Contents

	Foreword by John Chowning	xi
	Preface	xiii
	Acknowledgments	xiv
1	**Digital Signals and Sampling**	**1**
1.1	Measuring the Ephemeral	1
1.2	Analog-to-Digital Conversion	9
1.3	Aliasing	11
1.4	Digital-to-Analog Conversion	20
1.5	Binary Numbers	22
1.6	Synchronization	28
1.7	Discretization	28
1.8	Precision and Accuracy	29
1.9	Quantization	29
1.10	Noise and Distortion	33
1.11	Information Density of Digital Audio	38
1.12	Codecs	40
1.13	Further Refinements	42
1.14	Cultural Impact of Digital Audio	46
	Summary	47
2	**Musical Signals**	**49**
2.1	Why Imaginary Numbers?	49
2.2	Operating with Imaginary Numbers	51
2.3	Complex Numbers	52
2.4	de Moivre’s Theorem	62
2.5	Euler’s Formula	64
2.6	Phasors	68

2.7	Graphing Complex Signals	86
2.8	Spectra of Complex Sampled Signals	87
2.9	Multiplying Phasors	89
2.10	Graphing Complex Spectra	92
2.11	Analytic Signals	95
	Summary	100
3	**Spectral Analysis and Synthesis**	**103**
3.1	Introduction to the Fourier Transform	103
3.2	Discrete Fourier Transform	114
3.3	Discrete Fourier Transfom in Action	125
3.4	Inverse Discrete Fourier Transform	134
3.5	Analyzing Real-World Signals	138
3.6	Windowing	141
3.7	Fast Fourier Transform	145
3.8	Properties of the Discrete Fourier Transform	147
3.9	A Practical Hilbert Transform	154
	Summary	156
4	**Convolution**	**159**
4.1	Rolling Shutter Camera	159
4.2	Defining Convolution	161
4.3	Numerical Examples of Convolution	163
4.4	Convolving Spectra	168
4.5	Convolving Signals	172
4.6	Convolution and the Fourier Transform	180
4.7	Domain Symmetry between Signals and Spectra	180
4.8	Convolution and Sampling Theory	185
4.9	Convolution and Windowing	187
4.10	Correlation Functions	191
	Summary	193
	Suggested Reading	194
5	**Filtering**	**195**
5.1	Tape Recorder as a Model of Filtering	195
5.2	Introduction to Filtering	199
5.3	A Simple Filter	201
5.4	Finding the Frequency Response	203
5.5	Linearity and Time Invariance of Filters	217
5.6	FIR Filters	218

5.7	IIR Filters	218
5.8	Canonical Filter	219
5.9	Time Domain Behavior of Filters	219
5.10	Filtering as Convolution	222
5.11	Z Transform	224
5.12	Z Transform of the General Difference Equation	232
5.13	Filter Families	244
	Summary	261
6	**Resonance**	**263**
6.1	The Derivative	263
6.2	Differential Equations	276
6.3	Mathematics of Resonance	280
	Summary	297
7	**The Wave Equation**	**299**
7.1	One-Dimensional Wave Equation and String Motion	299
7.2	An Example	307
7.3	Modeling Vibration with Finite Difference Equations	310
7.4	Striking Points, Plucking Points, and Spectra	319
	Summary	324
8	**Acoustical Systems**	**325**
8.1	Dissipation and Radiation	325
8.2	Acoustical Current	326
8.3	Linearity of Frictional Force	329
8.4	Inertance, Inductive Reactance	332
8.5	Compliance, Capacitive Reactance	333
8.6	Reactance and Alternating Current	334
8.7	Capacitive Reactance and Frequency	335
8.8	Inductive Reactance and Frequency	336
8.9	Combining Resistance, Reactance, and Alternating Current	336
8.10	Resistance and Alternating Current	337
8.11	Capacitance and Alternating Current	337
8.12	Acoustical Impedance	340
8.13	Sound Propagation and Sound Transmission	344
8.14	Input Impedance: Fingerprinting a Resonant System	351
8.15	Scattering Junctions	357
	Summary	360
	Suggested Reading	362

9	**Sound Synthesis**	**363**
9.1	Forms of Synthesis	363
9.2	A Graphical Patch Language for Synthesis	365
9.3	Amplitude Modulation	384
9.4	Frequency Modulation	389
9.5	Vocal Synthesis	409
9.6	Synthesizing Concert Hall Acoustics	425
9.7	Physical Modeling	433
9.8	Source Models and Receiver Models	449
	Summary	450
10	**Dynamic Spectra**	**453**
10.1	Gabor's Elementary Signal	454
10.2	Short-Time Fourier Transform	459
10.3	Phase Vocoder	486
10.4	Improving on the Fourier Transform	496
10.5	Psychoacoustic Audio Encoding	502
	Summary	507
	Suggested Reading	509
	Epilogue	511
	Appendix	**513**
A.1	About Algebra	513
A.2	About Trigonometry	514
A.3	Series and Summations	517
A.4	Trigonometric Identities	518
A.5	Modulo Arithmetic and Congruence	522
A.6	Finite Difference Approximations	523
A.7	Walsh-Hadamard Transform	525
A.8	Sampling, Reconstruction, and the Sinc Function	526
A.9	Fourier Shift Theorem	528
A.10	Spectral Effects of Ring Modulation	529
A.11	Derivation of the Reflection Coefficient	530
	Notes	533
	Glossary	539
	References	543
	Index of Equations and Mathematical Formulas	547
	Subject Index	551

Foreword

During the 1960s and 1970s the Artificial Intelligence Laboratory at Stanford University was a multidisciplinary facility populated by enormously gifted and dedicated workers trained in the sciences, engineering sciences, social sciences, and music. The musicians were part of the Center for Computer Research in Music and Acoustics (CCRMA) and shared the use of the A. I. Lab's computer. While the scientists had no professional interest in music, their scientific and technical knowledge was of critical importance to the musicians in learning to make effective use of the ever-evolving hardware and software. We had been seduced by Max Mathews's now famous statement, "There are no theoretical limitations to the performance of the computer as a source of musical sounds, in contrast to the performance of ordinary instruments."[1] Music was to enter an entirely new domain.

An implicit understanding within this community of individual users went something like this: Any question asked would be cheerfully and completely answered, to any level of detail . . . *once!* It was assumed that the questioner would then make the effort to follow up and fully comprehend the answer and all its implications before returning with another question. For the musicians, disciplined and well educated but most having little mathematical or scientific training and none in the use of computers, this was an opportune and accommodating intellectual environment.

Gareth Loy was one of us who in the mid-1970s, as a graduate student at CCRMA, made full use of the opportunities presented to him in this extraordinary environment. He assumed responsibility for writing the software for the Samson Box, by far the most powerful and complex digital synthesizer/processor of the day, named after its primary designer, Peter Samson.[2] Because Loy was trained as a composer, composition was his ultimate purpose, and on completing the software he composed *Nekyia,* a beautiful and powerful composition in four channels that fully exploited the capabilities of the Samson Box. As an integral part of this community, Loy paid back many times over all that he had learned, by conceiving the system with maximal generality such that it could be used for research projects in psychoacoustics as well as for hundreds of compositions by a host of composers having diverse compositional strategies. These accomplishments, both musical and technical, further revealed to Loy the profundity of his capabilities and led him to pursue his interest in this shared territory of music and mathematics, ultimately to the benefit of all of us.

The two volumes of *Musimathics* are a kind of instantiation of the process of learning that had such a powerful facilitating effect on the work at CCRMA in those years: explanations presented with wit and in great detail but here logically ordered and ever available, not just *once!* This second volume of *Musimathics* is comprehensive. Loy focuses on the digital domain, from elemental binary numbers through digital signal processing and synthesis to such heady topics as Gabor and acoustical quanta, all in terms of their mathematical underpinning and all clearly explained with elegant and illuminating graphics. Reflecting his intellectual journey—the questions, the answers, the study, and most important, the motivation: music!—but now with the wisdom from years of teaching and study, Loy is an extraordinarily gifted guide. Excellent texts inspire, and this one certainly does.

John Chowning

Preface

This second volume of *Musimathics* continues the story of music engineering begun in volume 1. It takes a deeper cut into the mathematics of music and sound, including

- Digital audio, sampling, binary numbers
- Complex numbers and how they simplify representation of musical signals
- Fourier transform, convolution, and filtering
- Resonance, the wave equation, and the behavior of acoustical systems
- Sound synthesis
- The short-time Fourier transform and the wavelet transform

The material in volume 1 was all preparatory to the subjects introduced in this volume, although this volume can certainly be read independently. Cross-references to volume 1 occur wherever there is an antecedent concept required in this volume. Additional mathematical orientation is provided as necessary.

Musimathics takes an uncommon approach to presenting mathematics. It cultivates the reader's common sense. I believe that enlightened common sense and inference are the whole of mathematics and that inference itself flows from enlightened common sense. The cure for any lack of mathematical preparation on the reader's part is simply to focus on what makes the most sense, and the rest will follow. This is my personal experience and a major premise of this book.

Inference without common sense leads nowhere. But to the naive reader, this is exactly where treatises on mathematics seem to lead. The problem is how mathematics is presented in print. If authors had to state explicitly all the assumptions that underlie an argument, even trivial mathematical assertions would be too long-winded to print. Instead, the commonsense foundations of mathematical arguments are assumed so that the focus can be placed on the interesting and possibly surprising inferences that are being reported. This means that *most of the common sense has been removed on purpose,* rendering this splendid and remarkable subject off-limits to the mathematically unprepared reader. The cure for this is to "rehydrate" the common sense back into it. By seeing the rationale alongside the mathematics, the reader can gain a deeper insight into both music and mathematics.

Musimathics provides two aids to the reader. The first is common sense about music and sound; the second is patience about the process of inference. *Musimathics* provides complete derivations of important concepts together with explanations of the steps. Breathtaking vistas can be opened up by starting from humble assumptions and climbing the ladder of inference. But it's easy to fall off the ladder if the reader misses a step. I know what that fall feels like, so I've labored to make the climb as secure and straightforward as possible.

The Web site http://www.Musimathics.com contains additional source material, animations, figures, and sources for other program examples in this book. Also, try saying "Musimathics" to your favorite Web browser and see what happens.

Acknowledgments

I am grateful for the loving support I have received from my wife, Lisa, and my family and friends over the decade it has taken to write *Musimathics*. Thanks to all, including Bernard Mont-Raynaud, Mark Dolson, Dana Massie, and Charles Seagrave for reviewing chapters of this volume. Thanks to Linda Graham and Barbara Cook Loy for inspiration and support.

I am continually grateful to all whose scholarship and insight have fed into the rich stream of knowledge that this book can at best sample and summarize. The enormous list of these individuals begins with the bibliography of this book and extends recursively through all the influences they cite. If there is anything to praise in this work, it is because it reflects these antecedents; if there is fault, it is mine alone.

In closing, let me express my heartfelt thanks to the *Musimathics* team at the MIT Press: Doug Sery and Valerie Geary (acquisitions editor and assistant), Deborah Cantor-Adams (production editor) and Alice Cheyer (freelance copyeditor), Sharon Deacon Warne (designer), Janet Rossi (production coordinator), Mary Reilly (graphics coordinator), and Patrick Ciano (cover designer). I am especially grateful to Doug Sery, whose clear vision and steady hand helped guide me from initial contact through completed project. His belief in the value of this effort has sustained me and helped make publication possible.

Gareth Loy
Corte Madera, California, October, 2006

Musimathics

1 Digital Signals and Sampling

The gods confound the man who first found out how to distinguish the hours! Confound him, too, who in this place set up a sundial, to cut and hack my days so wretchedly into small portions!
—Plautus

Digital audio has fundamentally changed the way music is made, distributed, and shared. It is now so pervasive that most of the music we hear is digitally stored and processed. A good deal of it is also created digitally.

The public has benefited enormously from the technological advances of digital audio, but at a price. Legal efforts to limit and regulate music copying have come about largely because of digital audio's ability to make perfect copies of recordings. Digital audio has become the proverbial lion in the pathway that our society must address in order to restore balance between artistic, commercial, and social aims. To fully understand its promise and pitfalls, we all—musicians, audio engineers, and listeners alike—must come to grips with this technology.

This chapter introduces the digital representation of signals, which (in the words of Plautus) consists of cutting and hacking time into very small portions. The focus is on intuitive understanding. The mathematical underpinnings of sampling theory are delayed until section 4.8.

1.1 Measuring the Ephemeral

The dimensions of a wooden board remain relatively static over time, but the dimensions of sound waves are transitory, changing from moment to moment. How are we to measure something so ephemeral? To begin with, let's consider the ocean's tides, which have the explanatory advantage over sound waves of changing slowly and being visible.

In order to record wave motion, we could construct a float attached to a pier in the ocean (figure 1.1). An angled bar attached to the float is pushed up and down by the waves. The rollers that connect the bar to the pier constrain it to travel only vertically. A pen mounted on the end of the bar leaves a mark on a piece of paper wrapped around a rotating drum. The apparatus monitors wave motion continuously—at every instant. The track mark created by the wave fluctuations is a *continuous function* of time.

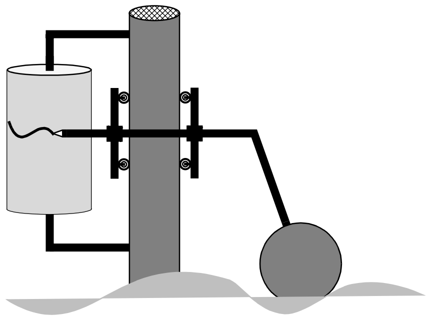

Figure 1.1
Float system for measuring waves.

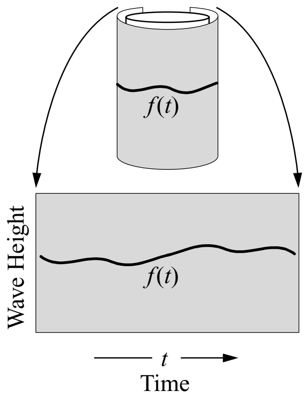

Figure 1.2
Unwrapping the paper from the drum.

Suppose we stop the drum after it has rotated once around, and unwrap the paper to examine the mark left on it by the pen (figure 1.2). Because of the paper's position on the drum, the x-axis represents the passage of time and the y-axis represents the fluctuating height of the waves.

If we let $f(t)$ represent the track mark, we can determine the height of the waves at any time t by evaluating $f(t)$ for the particular value of t that we wish to examine. For example, suppose it took 1 minute for the drum to revolve once, and the width of the paper is 1 meter. Then the height of the wave that occurred 30 seconds after we began is analogous to the height of the mark at 0.5 meter.

The function $f(t)$ is analogous in two senses: it analogizes time to place (on the x-axis), and wave height to place (on the y-axis). So $f(t)$ is an *analog function* of time. If $w(t)$ is the actual wave height, we can relate it to $f(t)$ by writing

$$f(t) \propto w(t),$$

where $\propto$ means "is proportional to." Analogies are very useful but can sometimes be misleading. For example, whereas real time flows inexorably forward, the paper analogy of time (represented by the variable t on the x-axis) is not similarly restricted. We can select any position along the x-axis, but we cannot return to the corresponding moment in real time. So the analogy is not perfect. We must remain alert to the limitations of analogies lest they confuse our thinking.

1.1.1 Sampling

An entirely different approach to measuring ocean waves is to sample wave height at periodic time intervals. Suppose once a day precisely at noon we go to a lagoon by the sea and look at a tide-measuring pole in the water showing the median height of the waves (figure 1.3). Numbered marks on the pole indicate the waves' height. We identify the mark that seems to be nearest to wave height, and record the measurements sequentially in a log book. After a month, we may end up with a list of measurements like those in figure 1.4.

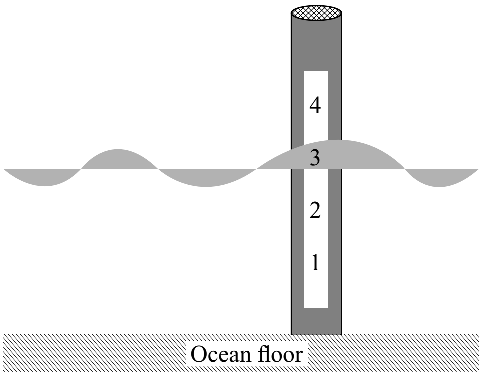

Figure 1.3
Sampling wave height.

January Tide Log

Date	Height
Jan. 1	3
Jan. 2	4
Jan. 3	3
⋮	⋮
Jan. 29	2
Jan. 30	3
Jan. 31	1

Figure 1.4
Tide log.

Whereas the float recorder measures wave height continuously, sampling measures it discontinuously, "every so often." Taken together, the samples represent a *discrete function* of time. If the discrete function g consists of the samples in the tide log in figure 1.4, then $g(n)$ indexes the nth tide sample record, where n is an integer. For example, if $n = 1$, then by figure 1.4, $g(n) = 3$. Similarly, $g(2) = 4$, $g(3) = 3$, and so on.

Though they have similar mathematical notation, continuous functions like $f(t)$ and discrete functions like $g(n)$ are very different. Whereas a continuous function can be evaluated at any real-valued index, there is nothing in between the integer-valued indexes of a discrete function. Nothing at all. Whatever the waves were doing while we weren't measuring is lost forever. The only way to tell for sure if a function is discrete or continuous is to check the type of its argument: if it is an integer, the function is discrete; otherwise it is continuous. Throughout this book, if I don't state otherwise, assume that the index of a function is a real value, that is, taken from continuous measurements.

1.1.2 Sampling Rate

If a constant time interval T elapses between observations, the process is called *periodic sampling* with sampling period T. The expression nT, where n is an integer, corresponds to the moment when sample n was taken. The rate at which samples are taken, called the *sampling rate* or *sampling frequency* f_s, is the reciprocal of the sampling period:

$$f_s = \frac{1}{T}. \qquad \textit{Sampling Frequency} \ (1.1)$$

1.1.3 Capturing Frequency Information

Periodic sampling not only captures amplitude information about the changing height of the waves, it also captures frequency information. Increasing the sampling rate increases the highest frequency that can be recorded.

For example, suppose we wish to study the ebb and flow of the ocean's tides, the vertical rise and fall of the sea level surface that is caused primarily by the change in gravitational attraction of the moon and, to a lesser extent, the sun. Sampling the tide once a month is not frequent enough to get meaningful data because the tides have a period of about 24 hours, 50 minutes. If we want to capture useful tidal information, we are *undersampling*.

We are still undersampling if we sample once a day because the time between a low tide and subsequent high tide is somewhat more than 6 hours. If we sample at least four times a day, we start getting reasonable data about the flow of tides. Sampling every hour provides a better view but requires making more measurements (and storing more information). If we sample every minute, we are probably *oversampling* because the tide changes less than about 1 centimeter per minute at its fastest rate. If we sample every second, we are now recording the individual waves as they splash past the measuring pole.

Thus, increasing the sampling rate increases the highest frequency of fluctuation that can be tabulated.

1.1.4 Improving the Process

There are a couple of problems with the tide sampling method just outlined. When we look at the tide-measuring pole, a jumble of waves obscures our observation of the slow-changing average tide. Let's remove these extraneous waves so that they don't interfere with the measurements.

Lowpass Filtering We can create a sampling system to record only slow-changing tide fluctuations. Mount a hollow tube vertically in the ocean floor. Its bottom rests on the sea floor and is closed; its top rises above the highest tide and is open. Sea water can flow in and out of the tube through a small-diameter pipe that is attached to its side well below sea level (figure 1.5). The narrow pipe restricts the rate at which water flows into and out of the tube, preventing rapid changes in water level. Since the small pipe prevents rapid fluctuation of water level—and since frequency is proportional to rate of fluctuation—the small pipe blocks high-frequency wave energy from entering the tube. The small-diameter pipe acts as a *lowpass filter,* meaning it only allows low frequencies to enter the tube.

Sample-and-Hold Although the lowpass filter slows down the water's rate of fluctuation inside the tube, the water level is still constantly changing, albeit at a slower rate. If we could make the level inside the tube as stationary as possible during measurements, accuracy would improve. One last refinement takes care of this: install a hinged lid that can seal the small pipe shut on command. The system in figure 1.6 now functions as a *sample-and-hold*.

Ordinarily, the lid is left open so that water can flow in and out of the tube as before. When we wish to measure the tide, we close the lid and wait a while for any turbulence to die down so that the water level becomes constant. Now we can measure the water's height with confidence.

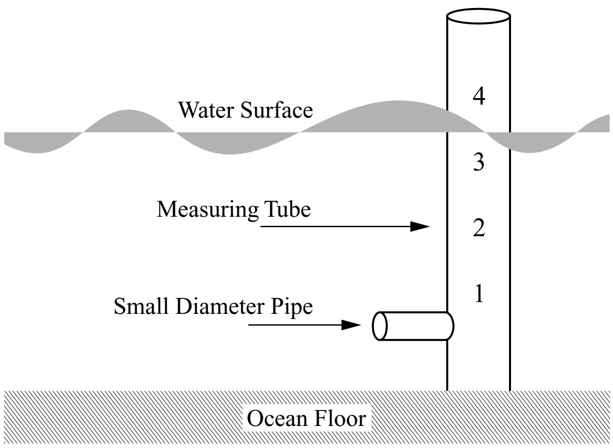

Figure 1.5
Lowpass filtering.

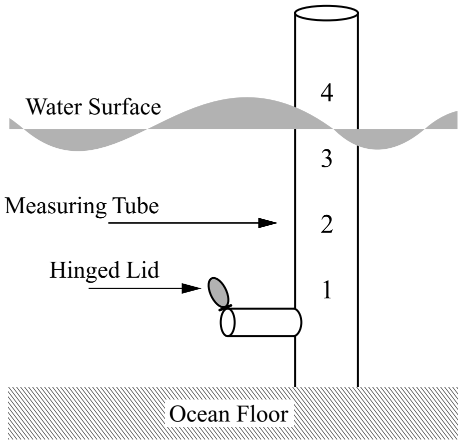

Figure 1.6
Sample-and-hold.

The sample-and-hold gets its name from its two states: sampling and holding. The hold state occurs when the valve is closed, while the tube holds the water level steady to be measured. During the time the valve is open, it is continuously sampling the height of the water level. The characteristic response of the sample-and-hold is shown in figure 1.7. The time required for the valve to close—to go from sampling to holding—is called the *aperture time* T_{ap}. The time required to go from holding back to sampling—from the moment the valve is opened again until the water level inside the tube is the same as the tide level outside—is called the *acquisition time* T_{acq}. The switching transients shown in the figure are the shudders sent through the system when the valve is suddenly opened or closed. If there is any leakage in the system when the valve is closed, the hold value will droop, as shown in the figure. A well-designed sample-and-hold will have little droop and mild transients.

A control function labeled Hold in figure 1.7 indicates the beginning of the aperture time and the beginning of the acquisition time. When the Hold function is low, the sample-and-hold is continuously sampling, and when it is high, the sample-and-hold is holding. Measuring the level and converting its height to a discrete value can commence anytime after the aperture time ends and before the acquisition time begins. Conversion is signaled by the Convert control function, as shown in the figure. It goes high in the middle of the hold period to trigger sample capture. The *conversion time* T_{conv} is the minimum sample time.

Figure 1.8 shows the operation of the sample-and-hold through several samples. The Hold and Convert signals are shown below the input and output functions. When the Hold function goes high, the sample-and-hold starts to hold its current output value. We see the output function of the sample-and-hold flatten and start to droop a little during the time Hold is high (marked with vertical lines). In the middle of the hold time, after the aperture time has elapsed, the Convert signal goes high. During that time (marked with boxes) the magnitude of the output function is converted. When

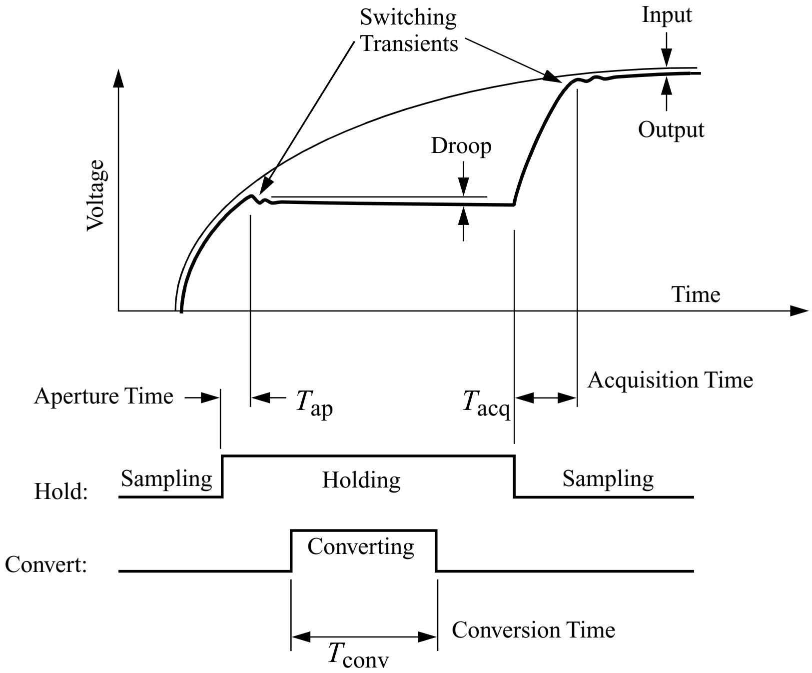

Figure 1.7
Sample-and-hold response characteristics. Adapted from Ramsay (1996).

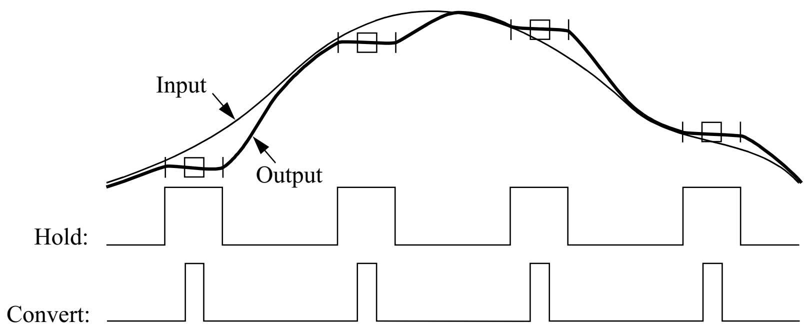

Figure 1.8
Sample-and-hold operation.