

# Contents

Figures, Tables, and Listings xi

Preface **About This Book** xvii

---

What's in This Book xviii  
Conventions Used in This Book xviii  
    Special Fonts xix  
    Types of Notes xix  
For More Information xix

Part 1 **The PowerPC Numerics Environment**

---

Chapter 1 **IEEE Standard Arithmetic** 1-1

---

About the IEEE Standard 1-3  
    Starting to Use IEEE Arithmetic 1-5  
    Careful Rounding 1-5  
    Exception Handling 1-6  
        Example: Finding Zero Return Values 1-7  
        Example: Searching Without Stopping 1-8  
        Example: Parallel Resistances 1-8  
Using IEEE Arithmetic 1-9  
    Evaluating Continued Fractions 1-9  
    Computing the Area of a Triangle 1-11  
About the FPCE Technical Report 1-12  
PowerPC Numerics Versus SANE 1-13

Chapter 2 **Floating-Point Data Formats** 2-1

---

About Floating-Point Data Formats 2-3  
Interpreting Floating-Point Values 2-4  
    Normalized Numbers 2-5  
    Denormalized Numbers 2-6  
    Infinities 2-7  
    NaNs 2-8  
    Zeros 2-10

Formats	2-11
Single Format	2-11
Double Format	2-13
Double-Double Format	2-14
Range and Precision of Data Formats	2-16

---

**Chapter 3**                      **Expression Evaluation**      3-1

---

About Expression Evaluation	3-3
Evaluating Expressions Without Widest Need	3-3
Evaluating Expressions With Widest Need	3-5
Comparisons of Expression Evaluation Methods	3-8

---

**Chapter 4**                      **Environmental Controls**      4-1

---

Rounding Direction Modes	4-3
Rounding Precision	4-4
Exception Flags	4-4
Invalid Operation	4-5
Underflow	4-5
Overflow	4-5
Divide-by-Zero	4-6
Inexact	4-6

---

**Chapter 5**                      **Conversions**      5-1

---

About Conversions	5-3
Converting Floating-Point to Integer Formats	5-3
Rounding Floating-Point Numbers to Integers	5-4
Converting Integers to Floating-Point Formats	5-5
Converting Between Floating-Point Formats	5-5
Converting Between Single and Double Formats	5-5
Converting Between Single and Double-Double Formats	5-5
Converting Between Double and Double-Double Formats	5-7
Converting Between Binary and Decimal Numbers	5-7
Accuracy of Decimal-to-Binary Conversions	5-7
Automatic Conversions	5-8
Manual Conversions	5-10
Converting Between Floating-Point and Decimal Structures	5-10
Converting Between Floating-Point and Decimal Strings	5-12

---

Chapter 6                      **Numeric Operations and Functions**                      6-1

---

Comparisons	6-3
Comparisons With NaNs and Infinities	6-3
Comparison Operators	6-3
Arithmetic Operations	6-5
Auxiliary Functions	6-14
Transcendental Functions	6-15

---

Part 2                              **The PowerPC Numerics C Implementation**

---

---

Chapter 7                      **Numeric Data Types in C**                      7-1

---

C Data Types	7-3
Efficient Type Declarations	7-3
Inquiries: Class and Sign	7-4
Creating Infinities and NaNs	7-5
Numeric Data Types Summary	7-6
C Summary	7-6
Constants	7-6
Data Types	7-7
Special Value Routines	7-7

---

Chapter 8                      **Environmental Control Functions**                      8-1

---

Controlling the Rounding Direction	8-3
Controlling the Exception Flags	8-5
Accessing the Floating-Point Environment	8-9
Environmental Controls Summary	8-14
C Summary	8-14
Constants	8-14
Data Types	8-14
Environment Access Routines	8-15

---

Chapter 9                      **Conversion Functions**                      9-1

---

Converting Floating-Point to Integer Formats	9-3
Rounding Floating-Point Numbers to Integers	9-6
Converting Integers to Floating-Point Formats	9-12
Converting Between Floating-Point Formats	9-13
Converting Between Binary and Decimal Numbers	9-13

Converting Between Decimal Formats	9-19
Conversions Summary	9-24
C Summary	9-24
Constants	9-24
Data Types	9-24
Conversion Routines	9-25

---

Chapter 10	<b>Transcendental Functions</b>	10-1
------------	---------------------------------	------

---

Comparison Functions	10-3
Sign Manipulation Functions	10-9
Exponential Functions	10-12
Logarithmic Functions	10-21
Logarithmic Functions	10-25
Trigonometric Functions	10-31
Hyperbolic Functions	10-42
Financial Functions	10-50
Error and Gamma Functions	10-55
Miscellaneous Functions	10-60
Transcendental Functions Summary	10-65
C Summary	10-65
Constants	10-65
Data Types	10-65
Transcendental Functions	10-65

---

Part 3	<b>Numerics in PowerPC Assembly Language</b>
--------	--

---

Chapter 11	<b>Introduction to Assembly-Language Numerics</b>	11-1
------------	---	------

---

PowerPC Floating-Point Architecture	11-3
Floating-Point Data Formats	11-3
Floating-Point Registers	11-3
Floating-Point Special-Purpose Registers	11-4
The Machine State Register	11-4
Floating-Point Instructions	11-4
Load and Store Instructions	11-5
Numerics Example Using PowerPC Assembly Language	11-7

---

Chapter 12      **Assembly-Language Environmental Controls**      12-1

---

The Floating-Point Environment	12-3
The Floating-Point Status and Control Register	12-3
The Condition Register	12-5
Inquiries: Class and Sign	12-7
Floating-Point Result Flags and Condition Codes	12-7
Example: Determining Class	12-8
Setting the Rounding Direction	12-9
Floating-Point Exceptions	12-10
Exception Bits in the FPSCR	12-10
Signaling and Clearing Floating-Point Exceptions	12-11
Enabling and Disabling Floating-Point Exceptions	12-12
Testing for Floating-Point Exceptions	12-12
Saving and Restoring the Floating-Point Environment	12-14

---

Chapter 13      **Assembly-Language Numeric Conversions**      13-1

---

Conversions From Integer to Floating-Point Formats	13-3
Conversions From Floating-Point to Integer Formats	13-4
Conversions From Single to Double Format	13-5
Conversions From Double to Single Format	13-5

---

Chapter 14      **Assembly-Language Numeric Operations**      14-1

---

Comparison Operations	14-3
Arithmetic Operations	14-4
Arithmetic Instructions	14-4
Multiply-Add Instructions	14-6
Move Instructions	14-7
Transcendental and Auxiliary Functions	14-8

---

Appendix A      **SANE Versus PowerPC Numerics**      A-1

---

Comparison of SANE and PowerPC Numerics	A-1
Floating-Point Data Formats	A-1
Conversions	A-1
Expression Evaluation	A-2
Infinities, NaNs, and Denormalized Numbers	A-2
Arithmetic and Comparison Operations	A-2
Environmental Controls	A-3
Transcendental (Elementary) Functions	A-3
Porting SANE to PowerPC Numerics	A-3
Replacing Variables of Type comp	A-4

Using MathLib Instead of the SANE Library	A-4
Replacing Extended Format Variables	A-5
Using MathLib Functions	A-5
Differences in Transcendental Functions	A-5
Differences in Class and Sign Inquiries	A-6
Differences in Environmental Controls	A-7
Compatibility Tools in MathLib	A-9
Portable Declarations	A-9
Macros	A-10

---

**Appendix B**                      **Porting Programs to PowerPC Numerics**      B-1

---

Semantics of Arithmetic Evaluation	B-1
Mixed Formats	B-2
Floating-Point Precision	B-2
The Rules of Evaluation	B-2
The Invalid Exception	B-3

---

**Appendix C**                      **MathLib Header Files**      C-1

---

Floating-Point Header File (fp.h)	C-1
Constants	C-1
Inquiry Macros	C-2
Data Types	C-3
Functions	C-4
Trigonometric Functions	C-4
Hyperbolic Functions	C-5
Exponential Functions	C-5
Power and Absolute Value Functions	C-7
Gamma and Error Functions	C-7
Nearest Integer Functions	C-8
Remainder Functions	C-9
Auxiliary Functions	C-9
Maximum, Minimum, and Positive Difference Functions	C-9
Internal Prototypes	C-10
Non-NCEG Extensions	C-10
Floating-Point Environment Header File (fenv.h)	C-12
Constants	C-12
Floating-Point Exception Flags	C-12
Rounding Direction Modes	C-12
Data Types	C-13
Functions	C-13
Controlling the Floating-Point Exceptions	C-13
Controlling the Rounding Direction	C-13
Controlling the Floating-Point Environment	C-13

Environmental Access Switch	D-1
Contraction Operator Switch	D-2
Hexadecimal Floating-Point Constants	D-3
Implementing an Expression Evaluation Method	D-3
Expression Evaluation Without Widest Need	D-4
Expression Evaluation With Widest Need	D-5
Floating-Point Constant Evaluation	D-5
Initializing Floating-Point Objects	D-7
Compiler Extensions for Expression Evaluation	D-8
Determining the Expression Evaluation Method	D-8
Widening for Efficiency	D-8

Floating-Point Data Formats	E-1
Environmental Controls	E-3
Operations and Functions	E-3

Floating-Point Data Formats	F-1
Floating-Point Status and Control Register	F-2
Instructions	F-4

