

# Figures, Tables, and Listings

Preface

About This Book xvii

---

Chapter 1

IEEE Standard Arithmetic 1-1

---

<b>Table 1-1</b>	Approximation of real numbers	1-4
<b>Listing 1-1</b>	Inverse operations	1-6
<b>Figure 1-1</b>	Parallel resistances	1-9
<b>Figure 1-2</b>	Graph of continued fraction functions $cf(x)$ and $rf(x)$	1-10
<b>Table 1-2</b>	Area using Heron's formula	1-11

Chapter 2

Floating-Point Data Formats 2-1

---

<b>Figure 2-1</b>	IEEE single format	2-3
<b>Table 2-1</b>	Names of data types	2-4
<b>Figure 2-2</b>	Normalized single-precision numbers on the number line	2-5
<b>Figure 2-3</b>	Denormalized single-precision numbers on the number line	2-6
<b>Table 2-2</b>	Example of gradual underflow	2-7
<b>Figure 2-4</b>	Infinities represented in single precision	2-8
<b>Table 2-3</b>	NaN codes	2-9
<b>Figure 2-5</b>	NaNs represented in single precision	2-10
<b>Figure 2-6</b>	Zeros represented in single precision	2-11
<b>Table 2-4</b>	Symbols used in format diagrams	2-11
<b>Figure 2-7</b>	Single format	2-12
<b>Figure 2-8</b>	Single-format floating-point numbers on the real number line	2-12
<b>Table 2-5</b>	Values of single-format numbers (32 bits)	2-12
<b>Figure 2-9</b>	Double format	2-13
<b>Table 2-6</b>	Values of double-format numbers (64 bits)	2-13
<b>Figure 2-10</b>	Double-format floating-point values on the real number line	2-14
<b>Figure 2-11</b>	Double-double format	2-14
<b>Figure 2-12</b>	Double-double format number example	2-15
<b>Table 2-7</b>	Summary of PowerPC Numerics data formats	2-16

Chapter 3

Expression Evaluation 3-1

---

<b>Figure 3-1</b>	Evaluating complex expressions without widest need	3-5
<b>Figure 3-2</b>	Evaluating complex expressions with widest need	3-7
<b>Figure 3-3</b>	Evaluating an expression with a function call	3-9
<b>Figure 3-4</b>	Evaluating an expression with arithmetic operations	3-10

Chapter 4	Environmental Controls	4-1
	<b>Table 4-1</b>	Examples of rounding to integer in different directions 4-4
Chapter 5	Conversions	5-1
	<b>Table 5-1</b>	Examples of floating-point to integer conversion 5-4
	<b>Table 5-2</b>	Double to single conversion: Possible exceptions 5-5
	<b>Figure 5-1</b>	Single to double-double conversion 5-6
	<b>Figure 5-2</b>	Double-double to single conversion 5-6
	<b>Table 5-3</b>	Double-double to single conversion: Possible exceptions 5-7
	<b>Figure 5-3</b>	Conversion cycle with first-time error 5-8
	<b>Figure 5-4</b>	Conversion cycle with correct result 5-9
Chapter 6	Numeric Operations and Functions	6-1
	<b>Table 6-1</b>	Comparison symbols 6-4
	<b>Table 6-2</b>	Arithmetic operations in C 6-5
	<b>Table 6-3</b>	Special cases for floating-point addition 6-6
	<b>Table 6-4</b>	Special cases for floating-point subtraction 6-7
	<b>Table 6-5</b>	Special cases for floating-point multiplication 6-8
	<b>Table 6-6</b>	Special cases for floating-point division 6-9
	<b>Table 6-7</b>	Special cases for floating-point square root 6-11
	<b>Figure 6-1</b>	Integer-division algorithm 6-12
	<b>Table 6-8</b>	Special cases for floating-point remainder 6-12
	<b>Table 6-9</b>	Special cases for floating-point round-to-integer 6-14
	<b>Table 6-10</b>	Examples of <code>rint</code> 6-14
Chapter 7	Numeric Data Types in C	7-1
	<b>Table 7-1</b>	Names of data types 7-3
	<b>Table 7-2</b>	<code>float_t</code> and <code>double_t</code> types 7-3
	<b>Table 7-3</b>	Class and sign inquiry macros 7-4
Chapter 8	Environmental Control Functions	8-1
	<b>Table 8-1</b>	Rounding direction modes in MathLib 8-3
	<b>Table 8-2</b>	Floating-point exception flags in MathLib 8-6
Chapter 9	Conversion Functions	9-1
	<b>Table 9-1</b>	Special cases for the <code>rinttol</code> function 9-4
	<b>Table 9-2</b>	Special cases for the <code>roundtol</code> function 9-5
	<b>Table 9-3</b>	Special cases for the <code>ceil</code> function 9-7
	<b>Table 9-4</b>	Special cases for the <code>floor</code> function 9-8
	<b>Table 9-5</b>	Special cases for the <code>nearbyint</code> function 9-9
	<b>Table 9-6</b>	Special cases for the <code>round</code> function 9-11
	<b>Table 9-7</b>	Special cases for the <code>trunc</code> function 9-12

<b>Table 9-8</b>	Format of decimal output string in floating style	9-15
<b>Table 9-9</b>	Format of decimal output string in fixed style	9-15
<b>Listing 9-1</b>	Accounting program	9-21
<b>Listing 9-2</b>	Scanning algorithm	9-22
<b>Table 9-10</b>	Examples of conversions to decimal structures	9-23

## Chapter 10

## Transcendental Functions 10-1

---

<b>Table 10-1</b>	Special cases for the <code>fdim</code> function	10-4
<b>Table 10-2</b>	Special cases for the <code>fmax</code> function	10-6
<b>Table 10-3</b>	Special cases for the <code>fmin</code> function	10-7
<b>Table 10-4</b>	Special cases for the <code>relation</code> function	10-8
<b>Table 10-5</b>	Special cases for the <code>copysign</code> function	10-10
<b>Table 10-6</b>	Special cases for the <code>fabs</code> function	10-11
<b>Table 10-7</b>	Special cases for the <code>exp</code> function	10-13
<b>Table 10-8</b>	Special cases for the <code>exp2</code> function	10-14
<b>Table 10-9</b>	Special cases for the <code>expm1</code> function	10-15
<b>Table 10-10</b>	Special cases for the <code>ldexp</code> function	10-16
<b>Table 10-11</b>	Special cases for the <code>pow</code> function	10-18
<b>Table 10-12</b>	Special cases for the <code>scalb</code> function	10-21
<b>Table 10-13</b>	Special cases for the <code>frexp</code> function	10-22
<b>Table 10-14</b>	Special cases for the <code>log</code> function	10-23
<b>Table 10-15</b>	Special cases for the <code>log10</code> function	10-26
<b>Table 10-16</b>	Special cases for the <code>log1p</code> function	10-27
<b>Table 10-17</b>	Special cases for the <code>log2</code> function	10-28
<b>Table 10-18</b>	Special cases for the <code>logb</code> function	10-30
<b>Table 10-19</b>	Special cases for the <code>modf</code> function	10-31
<b>Table 10-20</b>	Special cases for the <code>cos</code> function	10-33
<b>Table 10-21</b>	Special cases for the <code>sin</code> function	10-35
<b>Table 10-22</b>	Special cases for the <code>tan</code> function	10-36
<b>Table 10-23</b>	Special cases for the <code>acos</code> function	10-37
<b>Table 10-24</b>	Special cases for the <code>asin</code> function	10-38
<b>Table 10-25</b>	Special cases for the <code>atan</code> function	10-39
<b>Table 10-26</b>	Special cases for the <code>atan2</code> function	10-41
<b>Table 10-27</b>	Special cases for the <code>cosh</code> function	10-43
<b>Table 10-28</b>	Special cases for the <code>sinh</code> function	10-44
<b>Table 10-29</b>	Special cases for the <code>tanh</code> function	10-45
<b>Table 10-30</b>	Special cases for the <code>acosh</code> function	10-46
<b>Table 10-31</b>	Special cases for the <code>asinh</code> function	10-47
<b>Table 10-32</b>	Special cases for the <code>atanh</code> function	10-50
<b>Table 10-33</b>	Special cases for the <code>compound</code> function	10-51
<b>Table 10-34</b>	Special cases for the <code>annuity</code> function	10-54
<b>Table 10-35</b>	Special cases for the <code>erf</code> function	10-56
<b>Table 10-36</b>	Special cases for the <code>erfc</code> function	10-57
<b>Table 10-37</b>	Special cases for the <code>gamma</code> function	10-58
<b>Table 10-38</b>	Special cases for the <code>lgamma</code> function	10-59
<b>Table 10-39</b>	Special cases for the <code>nextafter</code> functions	10-61
<b>Table 10-40</b>	Special cases for the <code>hypot</code> function	10-63

Chapter 11	Introduction to Assembly-Language Numerics	11-1
	<b>Table 11-1</b>	Load and store floating-point instructions 11-6
	<b>Listing 11-1</b>	Polynomial evaluation 11-8
Chapter 12	Assembly-Language Environmental Controls	12-1
	<b>Figure 12-1</b>	Floating-Point Status and Control Register (FPSCR) 12-3
	<b>Table 12-1</b>	Bit assignments for FPSCR fields 12-4
	<b>Figure 12-2</b>	Condition Register 12-5
	<b>Table 12-2</b>	Branch instructions using the Condition Register 12-6
	<b>Table 12-3</b>	Values for FPSCR bits 15 through 19 12-7
	<b>Listing 12-1</b>	Determining the class of an assembler instruction result 12-8
	<b>Table 12-4</b>	Rounding direction bits in the FPSCR 12-9
	<b>Table 12-5</b>	Floating-point exception bits in the FPSCR 12-10
	<b>Listing 12-2</b>	Testing for occurrence of floating-point exceptions 12-13
	<b>Figure 12-3</b>	SRC and DST fields for <code>mtfsf</code> instruction 12-15
	<b>Listing 12-3</b>	Saving and restoring the floating-point environment 12-15
Chapter 13	Assembly-Language Numeric Conversions	13-1
	<b>Listing 13-1</b>	Converting a number from integer format to floating-point format 13-4
Chapter 14	Assembly-Language Numeric Operations	14-1
Appendix A	SANE Versus PowerPC Numerics	A-1
	<b>Table A-1</b>	Class and sign inquiries in SANE versus MathLib A-6
	<b>Table A-2</b>	Environmental access functions in SANE versus MathLib A-7
	<b>Listing A-1</b>	Using environmental controls in SANE and PowerPC Numerics A-8
	<b>Table A-3</b>	<code>float_t</code> and <code>double_t</code> definitions A-9

Appendix B                      Porting Programs to PowerPC Numerics      B-1

---

Appendix C                      MathLib Header Files      C-1

---

Appendix D                      FPCE Recommendations for Compilers      D-1

---

Appendix E                      MathLib Reference      E-1

---

<b>Figure E-1</b>	Floating-point data formats	E-1
<b>Table E-1</b>	Interpreting floating-point values	E-2
<b>Table E-2</b>	Class and sign inquiry macros	E-2
<b>Table E-3</b>	Environmental access	E-3
<b>Table E-4</b>	Floating-point exceptions	E-3
<b>Table E-5</b>	Rounding direction modes	E-3
<b>Table E-6</b>	Arithmetic operations	E-4
<b>Table E-7</b>	Conversions to integer type	E-4
<b>Table E-8</b>	Conversions to integer in floating-point type	E-4
<b>Table E-9</b>	Conversions between binary and decimal formats	E-5
<b>Table E-10</b>	Conversions between decimal formats	E-6
<b>Table E-11</b>	Comparison operations	E-6
<b>Table E-12</b>	Sign manipulation functions	E-6
<b>Table E-13</b>	Exponential functions	E-6
<b>Table E-14</b>	Logarithmic functions	E-7
<b>Table E-15</b>	Trigonometric functions	E-7
<b>Table E-16</b>	Hyperbolic functions	E-8
<b>Table E-17</b>	Financial functions	E-8
<b>Table E-18</b>	Error and gamma functions	E-8
<b>Table E-19</b>	Miscellaneous functions	E-8

Appendix F                      PowerPC Assembly-Language Numerics Reference      F-1

---

<b>Figure F-1</b>	Floating-point data formats	F-1
<b>Table F-1</b>	Interpreting floating-point values	F-1
<b>Table F-2</b>	Bit assignments for FPSCR fields	F-2
<b>Table F-3</b>	Rounding direction bits in the FPSCR	F-3
<b>Table F-4</b>	Class and sign inquiry bits in the FPSCR	F-3
<b>Table F-5</b>	FPSCR instructions	F-4
<b>Table F-6</b>	Load instructions	F-4
<b>Table F-7</b>	Store instructions	F-5
<b>Table F-8</b>	Conversions to integer format	F-5
<b>Table F-9</b>	Conversions from double to single format	F-5
<b>Table F-10</b>	Comparison instructions	F-5
<b>Table F-11</b>	Arithmetic instructions	F-6
<b>Table F-12</b>	Multiply-add instructions	F-6
<b>Table F-13</b>	Move instructions	F-6

