

Cambridge University Press

978-0-521-76135-2 - Finite Precision Number Systems and Arithmetic

Peter Kornerup and David W. Matula

Index

[More information](#)

## AUTHOR INDEX

- Agarwal, R. C., 392  
 Akushskii, I. Ja., 625  
 Al-Khwarizmi, 444  
 Antelo, Elisardo, 200  
 Atkins, Dan, 391  
 Atrubin, A. J., 241, 270  
 Avizienis, Algirdas, 55, 200
- Baugh, C. R., 222, 270  
 Benschop, N. F., 624  
 Booth, A. D., 54, 115  
 Brabec, Tomáš, 687  
 Brent, Richard P., 201  
 Briggs, Willard, 392, 444  
 Brightman, Thomas, 444  
 Bruguera, Javier D., 202, 522  
 Burgess, Neil, 116, 392, 625  
 Burla, N., 200
- Chen, I.-N., 270  
 Ciminiera, Luigi, 443
- Dadda, Luigi, 270, 271  
 DasSarma, Debjit, 393, 444  
 Daumas, Marc, 116
- Edalat, A., 688  
 Edwards, D. B. G., 201  
 Ercegovac, Miloš, 115, 200, 271, 391,  
 392, 443  
 Even, Guy, 271  
 Even, Shimon, 270
- Fandrianto, Jan, 443  
 Farmwald, Paul M., 393, 522  
 Ferrari, Domenico, 392  
 Fischer, Michael J., 115, 201  
 Fit-Florea, Alex, 624  
 Flynn, Michael J., 391, 392  
 Freiman, C. V., 391
- Garner, Harvey L., 624  
 Goldschmidt, R. E., 392  
 Gosling, J. B., 443  
 Gosper, R. W., 634, 666, 687  
 Gregory, Robert T., 626  
 Gustavson, F. G., 392
- Hall, Marshall, 634, 687  
 Hehner, Eric C. R., 626  
 Hensel, Kurt, 623, 626  
 Hiasat, A. A., 624  
 Horn, B. K. P., 687  
 Horspool, R. N. S., 626
- Iordache, Christina S., 392, 444  
 Ito, M., 393
- Jenkins, W. Kenneth, 624  
 Jensen, Thomas A., 391  
 Jullien, Graham, 624
- Kahan, William, 392  
 Khinchin, A. Y., 687  
 Knowles, Simon, 201  
 Knuth, Donald E., 1, 54, 523, 623, 686  
 Koç, Çetin K., 116  
 Kogge, Peter M., 115, 201  
 Kornerup, Peter, 115, 202, 271, 391, 626, 633,  
 686  
 Krishnamurthy, E. V., 392, 626  
 Kung, H. T., 201
- Ladner, Richard E., 115, 201  
 Lang, Tomás, 115, 202, 271, 391, 443, 522  
 Ledley, R. S., 391  
 Leeser, Miriam, 391  
 Lehman, M., 200  
 Lester, David, 688  
 Litman, Ami, 270  
 Lyon, R. F., 239

Cambridge University Press

978-0-521-76135-2 - Finite Precision Number Systems and Arithmetic

Peter Kornerup and David W. Matula

Index

[More information](#)

692

## Author index

MacSorley, O. L., 115, 201, 270, 391  
 Majerski, Stanislaw, 200  
 Matula, David W., 54, 115, 116, 200, 271, 392,  
 393, 444, 624, 633, 686  
 Metzke, Gernot, 200, 201, 391  
 Montgomery, Peter L., 550, 624  
 Montuschi, Paolo, 200, 443  
 Muller, Jean-Michel, 115  
 Munk Nielsen, Asger, 115  
 Møller, Ole, 523

Nadler, M., 391  
 Neumann, John von, 54, 200, 391, 522  
 Newton, I., 444  
 Niqui, Milad, 687

Oberman, Stuart F., 391, 392, 444  
 Oklobdzija, Vojin G., 202  
 Orup, Holger, 624

Panhaleux, Adrien, 115  
 Parhami, Behrooz, 55  
 Phillips, Braden, 116  
 Piestrak, Stanislaw J., 624  
 Potts, Peter J., 688

Reitwiesner, G. W., 115  
 Robertson, James E., 55, 200, 305, 391

Schmookler, Martin, 392, 523  
 Schulte, Michael J., 271  
 Schwartz, Eric M., 523  
 Seidensticker, R. B., 687  
 Sklansky, J., 201  
 Smith, J. L., 201

Soderquist, Peter, 391  
 Soderstrand, Michael A., 624  
 Spira, Philip M., 120–122, 200  
 Stone, Harold S., 115, 201  
 Sun-Tsü, 571  
 Svoboda, Antonin, 392, 624  
 Swartzlander Jr., Earl E., 201, 523  
 Sweeney, D. W., 305, 391  
 Szabo, Nicolas S., 624

Takagi, Naofumo, 393, 624  
 Tanaka, Richard I., 624  
 Taylor, George, 391, 443  
 Thill, Marco, 687  
 Tocher, T. D., 305, 391  
 Trivedi, Kishor, 200, 271, 686  
 Tyagi, Akhilesh, 201

Vàzquez, Alvaro, 200  
 Valach, M., 624  
 Vuillemin, Jean, 688

Wallace, C. S., 271, 392  
 Walter, Colin D., 624  
 Weinburger, Arnold, 201, 271  
 Wheeler, D. J., 392  
 Willoner, R., 270  
 Wilson, J. B., 391  
 Winograd, S., 120, 123, 200  
 Wooley, B. A., 222, 270

Yajima, S., 393

Zurawski, J. H. P., 443  
 Zuse, Konrad, 522

Cambridge University Press

978-0-521-76135-2 - Finite Precision Number Systems and Arithmetic

Peter Kornerup and David W. Matula

Index

[More information](#)

## INDEX

- 1's complement addition, 171, 172
- 1's complement representation, 41
- 2's complement, 15
- 2's complement carry-save sign, 41
- 2's complement carry-save polynomial, 40
- 2's complement overflow, 171
- 2's complement representation, 39
- 3-to-2 adder, 140
- 3-to-2 addition, 159
- 3-to-2 counter, 135
- 4-to-2 adder, 160
- 7-to-3 counter, 135
- 9's complement representation, 41
- adder
  - borrow-save, 164
  - carry-in of 0/1, 155
  - carry-save, 159
  - digit-serial, 144
  - $G_i, P_i$  composition, 152, 155
  - on-line, 146
  - ppp, npn, pnp, npn, nnn, 142
  - ripple carry, 140
  - sub-linear time, 147
- adder/subtractor, 178
- addition
  - 1's complement, 172
  - complexity, 123, 155
  - mapping, 126, 127, 129–131, 135
  - multioperand, 136
  - non-regular, 135
  - on-line, 144
  - overflow, 167, 170
  - radix complement, 170
  - table, 126, 129, 131, 133–135, 161
  - with redundant digit sets, 129
- additive inverse, 40, 111, 531
  - 1's complement, 42
  - 2's complement, 40
  - radix complement, 113
  - radix polynomial, 112
- adjacent, 637, 639, 642
- admissible representation, 677
- admissible string, 677
- arithmetic shift, 222
- array multiplier, 210, 229
- Atrubin multiplier, 241
- B-EEA Algorithm, 540
- balanced RNS, 566
- base, 4
- base conversion, 60
- base vector
  - MR system, 50
  - RNS system, 564
- base and digit set conversion, 95
  - radix 2 to  $2^k$ , 98, 101, 106
- basic-32, 504
- basic-64, 504, 505
- basic-128, 504, 505
- basic digit set, 31
- Baugh and Wooley scheme, 222, 270
- BCD, 519
- BCD code, 138
- BCD4221 code, 138, 200
- best rational approximation, 642
- bi-homographic, 666
- bi-linear, 648, 687
- bias, 39, 42, 142
  - floating-point exponent, 504, 505, 517
- biconvergent, 660, 662, 665
- bijacent, 661, 662, 664
- binade, 451
- binary
  - convergent, 660
  - mediant, 663
  - number, 8
  - part, 657
  - polynomial, 5
  - single precision, 500
- bipartite modular multiplication, 556, 624
- bipartite reciprocal, 358

- bipartite table, 375
- bipartite table look-up, 283
- bit-serial subtractor, 183
- bits-of-accuracy, 428
- BNE Algorithm, 652
- Booth recoding, 88, 107, 212, 270
- borrow-save, 32, 74
  - adder, 164
  - encoding, 104, 143, 213
- Brent–Kung prefix tree, 154, 260
- broadcast elimination, 233
- BSSE Algorithm, 651
  
- canonical complement polynomial, 34
- canonical continued fraction, 641, 667
- canonical representation, 26, 53, 92, 115, 213
- canonical signed-digit form, 35
- carry
  - anticipation, 88, 90, 133, 135
  - $G_i$ ,  $P_i$  composition, 152, 155
  - generate, 147
  - generate signal  $G_i$ , 152
  - kill, 147
  - look-ahead generator, 157
  - propagate, 147
  - propagate signal  $P_i$ , 151
  - set, 67, 126
- carry-absorption table, 133
- carry-look-ahead adder, 151
- carry-relation, 78
- carry-save, 32, 103
  - 2's complement polynomial, 40
  - 2's complement sign, 41
  - adder, 159
  - addition, 134, 135, 159
  - digit set, 125
  - encoding, 103, 138
- carry-select adder, 149
- carry-skip adder, 148
- carry-transfer function, 70, 72, 82
- Cauchy sequence, 618
- channel, 528, 566
- Chinese Remainder Theorem (CRT), 571
- coding
  - complete, 137
  - non-redundant, 137
- combinational circuits, 120
- comparison, 185, 187
- compatible radices, 62
- complement digit-set, 33
- complement polynomials, 33
- complementary error factor, 344
- complementary root pair, 401
- complementary roots/tails/remainers, 402
- complementary  $(q, r)$ -pair, 291
- complete digit set, 16, 18–20, 24, 26, 27
- complete residue system, 16, 19, 21, 23, 24
- completeness, 14
- conditional-sum adder, 150
  
- constant-time adders, 159
- continuation, 46
- continuation function, 483
- continued fraction, 609, 641
  - convergent, 642
  - expansion, 609
- continued logarithms, 687
- convergence division, 281, 345, 350
- convergence square root, 434
- convergent, 609, 642, 648
- conversion diagram, 84
- conversion mapping, 67
- core
  - computation of, 592, 595
  - critical, 593
  - function, 586, 592
  - non-critical, 593
  - weights, 586
- correct rounded, 447
- corresponding tail, 401
- counter
  - 3-to-2, 135
  - 7-to-3, 135
- critical core, 593
- CRT, 571, 592
  - alternative form, 579
  - base extension, 579
  - for core functions, 592, 597
  - rank, 579, 593
- cube, 668, 672
  
- Dadda scheme, 253
- decade, 451
- decimal number, 8
- decimal polynomial, 5
- decleat, 138, 519
- decoding function, 137
- deferred carry-assimilation, 159
- denormalized, 469, 523
- Densely Packed Decimal, DPD, 138, 200, 519
- DGT Algorithm, 21, 23, 213
  - complete digit sets, 26
- digit addition, 125
- digit coding, 137
- digit-mapping function, 71, 72
- digit selection function, 308, 324, 408
- digit serial adder, 144
- digit set, 14
  - basic, 31
  - conversion, 60, 66, 75
  - conversion complexity, 123
  - extended, 31
  - maximally redundant, 31
  - minimally redundant, 31
  - standard, 31
  - symmetric, 31
- diminished radix complement, 41, 112
- Diophantine equation, 563
- direct coding, 137

- directed precise rounding, 460
- directed root, 400
- directed rounding, 460
- discrete logarithm, modular, 559
- discrete logarithm encoding, 562
- dividend, 275, 284
- division
  - $2n$ -by- $n$ -bit fixed-point, 304
  - binary SRT, 305
  - convergence, 281, 345, 350
  - digit serial paradigm, 279
  - Goldschmidt, 345
  - integer 2's complement, 305
  - invariant, 275, 277, 284
  - iterative refinement, 344
  - iterative refinement paradigm, 281
  - Newton–Raphson, 281, 346
  - $n$ -by- $n$ -bit fixed-point, 300, 301
  - $n$ -by- $n$ -bit integer, 305
  - non-restoring, 279, 298
  - non-restoring 2's complement, 300
  - postscaled, 282, 345, 354
  - prescaled, 280, 334; with remainder, 338
  - quarter-ulp, 344
  - recurrence, 290
  - restoring, 279, 293
  - short reciprocal, 280, 330
  - SRT, 280, 307
  - testing, 392
  - with remainder zero, 568
- divisor, 275, 284
- dlg factorization, 560
- dot-matrices, 162
- dot-products, 211
- double-base system, 52
- double-rounding, 90, 469, 523
- double precision, 504, 505
- DPD, Densely Packed Decimal, 138, 200, 519
  
- EAA Algorithm, 648, 650, 651
- ECA Algorithm, 643, 647, 650
- EEA Algorithm, 538, 607, 610, 616, 622
- encoding function, 137
- end-around carry, 172, 182, 202, 541
- equality testing, 185, 187
- equivalent digit formula, 65
- eqv-interval, 449, 465–467, 470
- error factor
  - complementary, 281
  - relative, 281
- Euclidean Arithmetic Algorithm (EAA), 648, 651
- Euclidean Convergent Algorithm (ECA), 643, 647, 650
- evaluation mapping, 3
- exceptional value, 514
- excess-3 code, 138, 141
- extended digit set, 31
- extended monomial, 2
- extended polynomial, 2
- extended precision, 504, 505
  - extended-40, 504
  - extended-80, 504, 505
- Extended Euclidean Algorithm (EEA), 538, 607, 610, 616, 622
  - binary (B-EEA), 540
- fan-out restriction, 120
- far path, 473, 479, 481
- Farey fraction, 605, 606, 636
- Farey set, 637
- fault tolerance, 625
- Fermat's Little Theorem, 534
- Fibonacci number, 665
- fixed-point system, 36
- fixed-slash, 636
- floating point
  - directed rounding, 460
  - exception, 508
  - exceptional value, 451
  - exponent, 450, 500
  - exponent bias, 517
  - factorization, 450, 499
  - Hensel code, 622
  - IEEE rounding modes, 510
  - infinity, 454
  - normalized, 451, 500
  - not-a-number, 454
  - number, 450
  - polynomials, 453
  - range-width, 503
  - rounding, 460
  - scale factor, 450
  - significand, 450, 500
  - sign bit, 450
  - sign factor, 450
  - single precision, 500
  - subnormal, 500
  - unit (FPU), 450
- floating-slash, 636
- fma: fused multiply-add, 211, 471, 475, 482, 522
- forward mapping, 607
- FPU, 450
- fraction, 634
- full-adder, 135, 140, 160
- fused multiply-add, 211, 471, 475, 522
  
- gap, 455, 637, 639, 664
  - relative, 502
- gap function, 502
- gates, 120
- GCD algorithm
  - binary, 623
  - classical, 537
  - extended, 538
- generalized matrix multiplication, 672
- generalized residue class, 604
- generator, 548
- Goldschmidt division, 345, 350

Cambridge University Press

978-0-521-76135-2 - Finite Precision Number Systems and Arithmetic

Peter Kornerup and David W. Matula

Index

[More information](#)

696

Index

Goldschmidt square root, 434  
 gradual underflow, 469  
 guard bits in reciprocal tables, 368, 372  
 guard digits, 166, 176, 323, 329, 332, 429, 447, 486, 487  
 leading, 166–168, 174, 176

$h_j$ -separable set, 122, 125  
 half-adder, 138  
 half-period of modulus, 533  
 half-ulp approximation, 35  
 half-ulp radix approximation, 43  
 half-ulp root, 400  
 half-ulp rounding, 463  
 Hensel code, 613, 618, 621  
 hereditary function, 558  
 hexadecimal number, 8  
 hexadecimal polynomial, 5  
 high radix multiplication, 217  
 hyperbolic chain, 637

IEEE rounding, 89, 286  
 IEEE rounding modes, 510  
 IEEE Standard, 462, 498  
 index table, 548  
 inheritance property, 558  
 interleaved modular multiplication, 549, 624  
 interpolation in table support, 383  
 interval arithmetic, 447  
 inverse  
   additive, 531  
   multiplicative, 531, 534, 535, 538  
 inverse mapping, 610  
 irreducible factorization, 8  
 irreducible fraction, 635  
 iterative multiplication, 210  
 iterative multiplier, 228  
 iterative refinement division, 344

Kogge–Stone prefix tree, 154

$L_M$ : multiply latency, 345  
 $L_T$ : table look-up latency, 345  
 last bit problem, 47  
 last place, 5  
 latency, 231  
 LCF, 657  
 LCF tree, 660, 664  
 leading zeroes determination (LZD), 193, 474  
 least-significant digit first (LSD), 209  
 lexicbinary encoding, 657  
 lexicographic continued fraction (LCF), 657  
 lexicographically adjacent, 661  
 linear-time adders, 136  
 LR optimal recoding, 94  
 Lyon multiplier, 239, 270  
 LZD, leading zeroes determination, 193

M-residue, 550  
 Manchester carry chain, 149

Maniac-III, 523  
 mapping between residue and radix, 569  
 maximally redundant digit set, 31  
 mediant, 638, 663  
 mediant rounding, 644, 654  
 Mersenne prime, 532, 612  
 minimally redundant digit set, 31  
 mixed-radix representation (MRR), 574  
 Möbius transformation, 648, 687  
 modified Booth recoding, 34, 90, 107, 213, 270  
 modular addition, 544, 547  
 modular addition select adder, 542  
 modular division, 531, 539  
 modular exponentiation, 557  
 modular multiplication, 532, 541, 548, 551, 561  
   bipartite, 556, 624  
   interleaved, 549, 624  
   Montgomery, 552, 555  
 modular multiplier, 543  
 modulus operator, generalized, 39  
 Montgomery modular multiplication, 552, 555  
 most-significant digit first (MSD), 209  
 MR base extension, 579  
 multioperand modular addition, 544, 547  
 multipartite table, 381  
 multipartite table look-up, 283  
 multiple-modulus rational systems, 613  
 multiple-modulus residue representation, 125  
 multiplicand, 208  
 multiplication complexity, 124  
 multiplicative inverse, 531, 538  
 multiplier, 208  
   2's complement, 222  
   array, 210, 229  
   Atrubin, 241  
   Baugh and Wooley, 222  
   broadcast elimination, 233, 236  
   integer with overflow detection, 258  
   iterative, 210, 228  
   logarithmic time, 252  
   Lyon serial/parallel, 239  
   Lyon serial/serial, 246  
   on-line, 211, 249  
   pipeline, 237  
   radix-4, 214, 225  
   recoded, 224  
   recoding, 212, 216, 218  
   rectangular, 219  
   redundant, 215  
   retiming, 231  
   serial, 211  
   serial/parallel, 232  
   tree, 210, 256  
 multiply-add operation, 211, 475  
 N-mapping, 96, 105  
 NAF, non-adjacent form, 35, 92, 115, 213  
 NaN, not-a-number, 508, 635

Cambridge University Press

978-0-521-76135-2 - Finite Precision Number Systems and Arithmetic

Peter Kornerup and David W. Matula

Index

[More information](#)

- nearest precise rounding, 461
- near path, 472, 479
- nega-binary polynomial, 5
- negated digit set, 111
- Newton–Raphson Division, 281, 345, 346
- Newton–Raphson Root-Reciprocal, 432
- Newton–Raphson square root, 428
- non-adjacent form (NAF), 35, 92, 115, 213
- non-critical core, 593
- non-redundancy, 14
- non-redundant digit encoding, 104
- non-redundant digit set, 15, 17–19, 26, 31
  - conversion into, 66, 72
  - on-the-fly conversion into, 69
- non-redundant representation, 125
- non-restoring division, 279, 298
- normalization, 171, 192, 199
  - quasi-, 193, 471, 654
- normalized, 500
- normalized remainder, 290
- normalized residue vector, 616
- normalized rounding eqv-interval, 468
  
- octal number, 8
- octal polynomial, 5
- on-line, 45
- on-line adder, 146
- on-line addition, 144
- on-line algorithms, 108
- on-line converter, 109
- on-line delay, 109
- on-line multiplier, 211, 249
- on-the-fly conversion, 69, 413
- on-the-fly rounding, 495, 523
- one-ulp root, 400
- one-ulp rounding, 463
- one ulp radix approximation, 43
- order, 6, 452
  - LCF expansion, 658
- ordering relation
  - complexity, 188, 190
  - three-way, 189, 190
  - two-way, 189
- overflow
  - 2's complement integer multiplication, 262
  - 2's complement, 171
  - redundant addition, 166
  - unsigned integer multiplication, 259
  
- $p$ -adic expansion, 619
- P–D diagram, 310
- $P$ -mapping, 96, 105
- parallel prefix
  - adder, 153
  - computation, 72, 151, 153, 201
  - structure, 73
- parent, 639
- partial product, 208
- partial product array, 252
- partial product generator (PPG), 212, 228, 257
  
- partial quotient, 609, 641
- partial quotient distribution, 657
- partial remainder, 290, 310
- partial squarand, 267
- partial squarand generator, 267
- period of modulus, 533, 546
- pipeline, 237
- PLA, 196, 308
- place digit, 76
- positional number systems, 1
- postscaled division, 282, 345, 354
- PPG, 212, 228, 257
- PQRP-algorithm, 291, 301, 305
- precise rounding, 460
- precision
  - binary floating-point, 504
  - decimal floating-point, 505
  - wobble, 457
- precision conversion, 467
- precision wobble, 502
- preconvergent, 662
- prescaled division, 280
- prescaled short radicand-reciprocal, 426
- prescaled square root, 418, 422
- prescaling SRT, 319
- prime decomposition, 9
- principal binade, 451
- projection of residue vector, 572
- proper quotient, remainder pair, 289
- pseudo-overflow, 607
  
- $Q$ -mapping, 101, 106
- quadruple precision, 504, 505
- quantization error, 583
- quarter-square method, 547
- quarter-ulp division, 344
- quasi-normalization, 193, 471, 654
- quotient, 275, 284
  - rounding, 483
- quotient–remainder representation, 573
  
- radical, 400
- radicand, 400
- radicand-reciprocal, 422
- radix, 4
  - $\beta$  monomials of order  $j$ , 6
  - $\beta$  polynomials, 5
  - complement, 33
  - complement addition, 170
  - complement representation, 38, 169, 188
  - digit-string, 11
  - factorization, 8
  - fraction part, 12
  - integer part, 12
  - point, 11, 12
  - vector (MR-system), 50
- radix- $\beta$  digit set, 16
- radix- $\beta$  factorization, 450
- radix- $\beta$  number, 7
- radix- $\beta$  rational, 7

- radix polynomial
  - addition mapping, 127
  - canonical symmetric, 34
  - subtraction mapping, 178
- range-width, 503
- rank, 579
- reciprocal
  - best approximation, 366
  - best real approximation, 364
  - bipartite approximation, 375, 381
  - by interpolation, 383
  - direct table lookup, 363
  - monotonic approximation, 371, 373, 374
  - multipartite approximation, 381
  - table support, 361
- reciprocal bit, 658
- recoded multiplier, 224
- recoding, 212
  - in table support, 382
- rectangular multiplier, 209
- redundancy
  - class, 9
  - factor, 308, 408
  - index, 86
- redundant digit set, 17
  - conversion into, 76, 77, 86
  - on-line conversion into, 108
- redundant representations, 3
- regular addition table, 126
- relative error factor, 344
- relative gap, 456
- relgap, 640
- remainder, 275, 284
- residue
  - class, 3
  - determination, 532
  - mapping of rationals, 604
  - standard, 530
  - symmetric, 530
  - vector, 564
- residue number system (RNS), 564
- restoring division, 279, 293
- retiming, 231, 233
- ripple carry adder, 140
- RL canonical recoding, 93
- RN, RU, RD, RZ, RA roundings, 48, 460
- RN codings, 115
- RNS
  - addition, 567
  - additive inverse, 567
  - balanced, 566
  - base, 564
  - base  $2n - 1$ ,  $2n$ , and  $2n + 1$ , 625
  - base extension, 578
  - channel, 566
  - core: base, 592; computation, 592; function, 586, 592
  - divide and conquer conversion, 577
  - division, 568, 600, 602
  - error detection, 568
  - mapping into, 569
  - multiplication, 567
  - multiplicative inverse, 568
  - orthogonal base, 572
  - overflow, 567, 584
  - overflow check, 590
  - parity check, 590
  - projection, 572, 576
  - representation, 564
  - residue recovery, 591
  - scaling, 582, 583
  - sign determination, 584, 597, 598
- Robertson diagram, 297, 308, 391
- round bit, 465
- round digit, signed, 467
- round-to-nearest, 89
- round-to-nearest representation, 89
- round-away-from-zero, RA, 461
- round-down, RD, 461, 510
- round-to-nearest-away,  $RN_a$ , 462, 510
- round-to-nearest-even,  $RN_e$ , 461, 510
- round-towards-zero, RZ, 447, 461, 510
- round-up, RU, 461, 510
- rounded fp-arithmetic, 463
- rounding, 48, 89, 460
  - directed, 48
  - equivalent, 465
  - precise, 460
  - rule, 635
- rounding interval
  - normalized, 468
  - quasi-normalized, 471
- rounding-equivalent interval, 465
- RPQ Algorithm, 678
- RPQ cell, 683
- RPQ radix representation, 682
- RPQ representation, 677
- scaled remainder, 290, 300
- scientific notation, 448, 451
- seed matrix, 537, 608, 647, 648, 656, 657
- selection function, 310
- self timed, 147
- self delimiting, 657, 677
- semi complete, 33
- sequential circuit, 120
- serial/parallel multiplier, 232
- serial multiplier, 211
- sgn (signum function), 46
- short radicand-reciprocal, 424
- short reciprocal, 331, 332
- short reciprocal division, 280, 330
- short reciprocal square root, 418, 419
- short root reciprocal, 419
- sign determination, three-way, 190, 199
- sign extension, 177
- sign magnitude
  - addition and subtraction, 180, 183
  - conversion into, 157
  - encoding, 137, 213; of digits, 104



Cambridge University Press

978-0-521-76135-2 - Finite Precision Number Systems and Arithmetic

Peter Kornerup and David W. Matula

Index

[More information](#)

- fixed point system, 38
- radix polynomials, 32
- representation, 32, 112
- signed continued fraction, 653
- signed digit, binary, 32
- signed digit encoding, 104, 143
- signed digit set, 14
- signed infinity, 514
- signed zero, 514
- significance arithmetic, 523
- significand, 192, 500
- signum function, 46, 190
- sign bit, 499
- sign factor, 450
- simpler-than relation, 635
- simple chain, 635
- single-modulus integer arithmetic, 530
- single-modulus integer system, 529
- single precision, 504, 505
- SLCF, 659
- squarand, 266
- square root
  - binary non-restoring, 406
  - combined with SRT division, 409
  - convergence, 427, 434
  - digit serial, 400
  - Goldschmidt, 427, 434
  - invariant, 400
  - Newton–Raphson, 427, 428
  - remainder form, 430
  - remainder update, 403
  - restoring, 404
  - root reciprocal, 427, 432
  - scaled remainder, 404
  - school method, 404, 429
  - shifted tail, 404
  - short reciprocal, 416
  - SRT digit selection, 411
- squaring, 262
- SRT division, 280, 307
- SRT square root, 407
- standard
  - binary, 32
  - complement digit set, 33
  - digit set, 31
  - mixed-radix system, 51
  - residue system, 530
- stereographic projection, 676
- Stern–Brocot tree, 641, 687
- sticky bit, 465
- sticky digit, 475
  - signed, 467
- storage-16, 504
- storage-32, 505
- string signum function, 190
- subnormal, 469, 500
- subtractor, 143, 178
- suffix, 191
- switch bit, 657, 658
- symmetric digit set, 31
- symmetric residue system, 530
- system modulus, 529
- systolic, 233
- systolic array, 235
- table look-up
  - bipartite, 283
  - direct LUT, 283
  - multipartite, 283
- table maker’s dilemma, 47, 55
- Taylor diagram, 310
- tensor, 671
- tensor product, 671, 672, 687
- ternary number, 8
- ternary polynomial, 5
- transducer, 68, 144, 145
- transfer digit, 67
- tree-multiplier, 210
- truncated  $p$ -adic system, 618
- truncation rule, 49
- two-level radix system, 52
- ulp (unit in the last place), 6, 43, 48, 454
- ulp function, 454
- Ulp Approximation Lemma, 47
- unary part, 657
- unit in the last place (ulp), 43
- unit monomial, 6
- Wallace scheme, 253
- weight, 53
  - minimal, 94, 116
  - of a digit set, 217
  - of a digit string, 92
- weighted number system, 50
- zero polynomial, 28
- Zrange*, 669, 673, 679