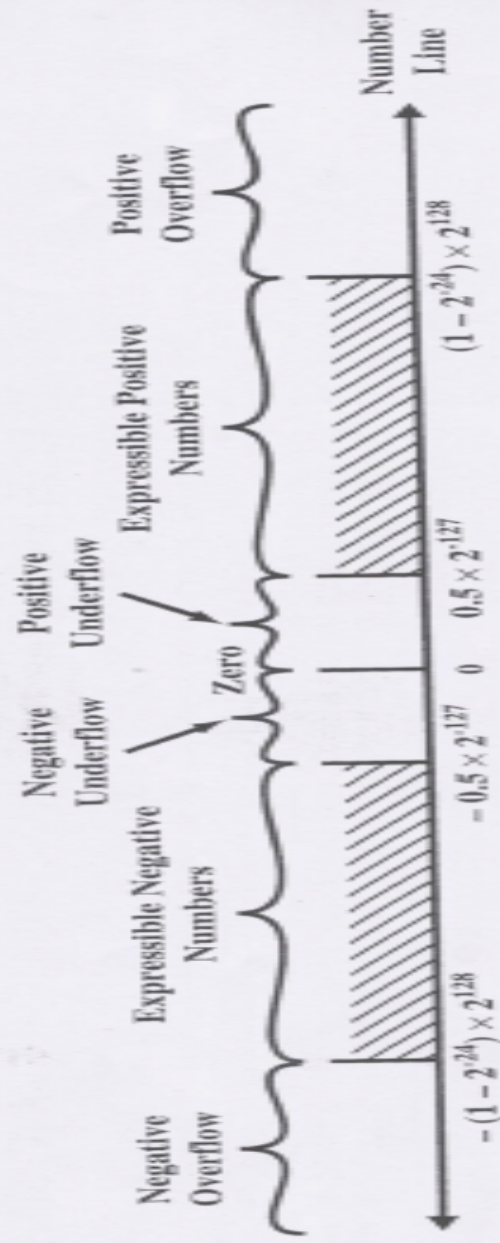


(a) Twos Complement Integers



(b) Floating-Point Numbers

Figure 8.19 Expressible Numbers in Typical 32-Bit Formats.

- **Exponent overflow:** A positive exponent exceeds the maximum possible exponent value. In some systems, this may be designated as  $+\infty$  or  $-\infty$ .
- **Exponent underflow:** A negative exponent is less than the minimum possible exponent value (e.g.,  $-200$  is less than  $-127$ ). This means that the number is too small to be represented, and it may be reported as  $0$ .
- **Significant underflow:** In the process of aligning significands, digits may flow off the right end of the significand. As we shall discuss, some form of rounding is required.
- **Significant overflow:** The addition of two significands of the same sign may result in a carry out of the most significant bit. This can be fixed by realignment, as we shall explain.

1. Check for zeros.
2. Align the significands.
3. Add or subtract the significands.
4. Normalize the result.

Table 8.4 Interpretation of IEEE 754 Floating-Point Numbers

Single Precision (32 bits)				Double Precision (64 bits)			
Sign	Biased exponent	Fraction	Value	Sign	Biased exponent	Fraction	Value
0	0	0	0	0	0	0	0
1	0	0	-0	1	0	0	-0
0	255 (all 1s)	0	$\infty$	0	2047 (all 1s)	0	$\infty$
1	255 (all 1s)	0	$-\infty$	1	2047 (all 1s)	0	$-\infty$
0 or 1	255 (all 1s)	$\neq 0$	NaN	0 or 1	2047 (all 1s)	$\neq 0$	NaN
0 or 1	255 (all 1s)	$\neq 0$	NaN	0 or 1	2047 (all 1s)	$\neq 0$	NaN
0	$0 < e < 255$	f	$2^{e-127}(1.f)$	0	$0 < e < 2047$	f	$2^{e-1023}(1.f)$
1	$0 < e < 255$	f	$-2^{e-127}(1.f)$	1	$0 < e < 2047$	f	$-2^{e-1023}(1.f)$
0	0	f $\neq 0$	$2^{e-126}(0.f)$	0	0	f $\neq 0$	$2e^{-1022}(0.f)$
1	0	f $\neq 0$	$-2^{e-126}(0.f)$	1	0	f $\neq 0$	$-2e^{-1022}(0.f)$

Positive zero

Negative zero

Plus infinity

Minus infinity

Quiet NaN

Signaling NaN

Positive normalized nonzero

Negative normalized nonzero

Positive denormalized

Negative denormalized

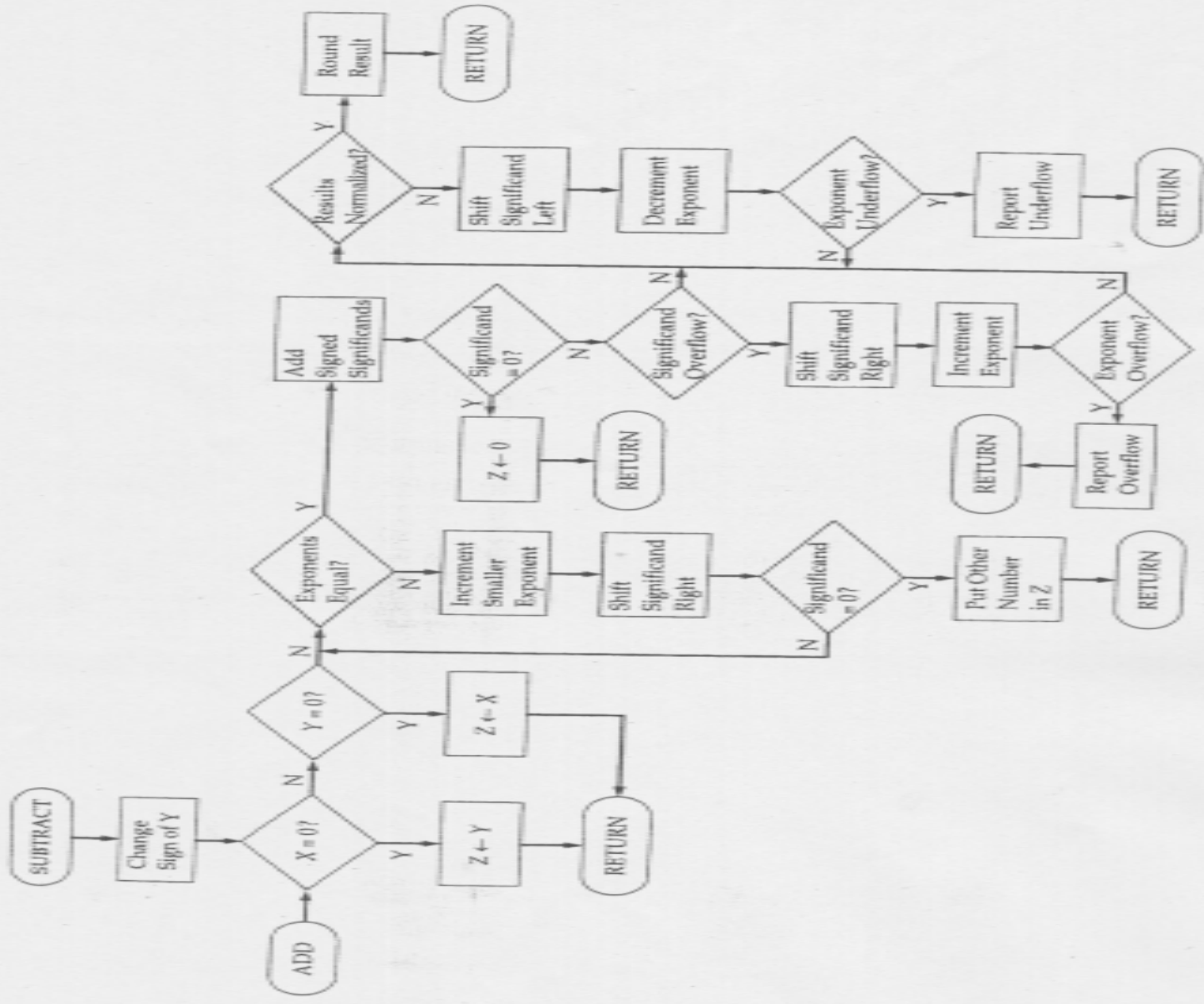


Figure 8.22 Floating-Point Addition and Subtraction ( $Z \leftarrow X \pm Y$ ).

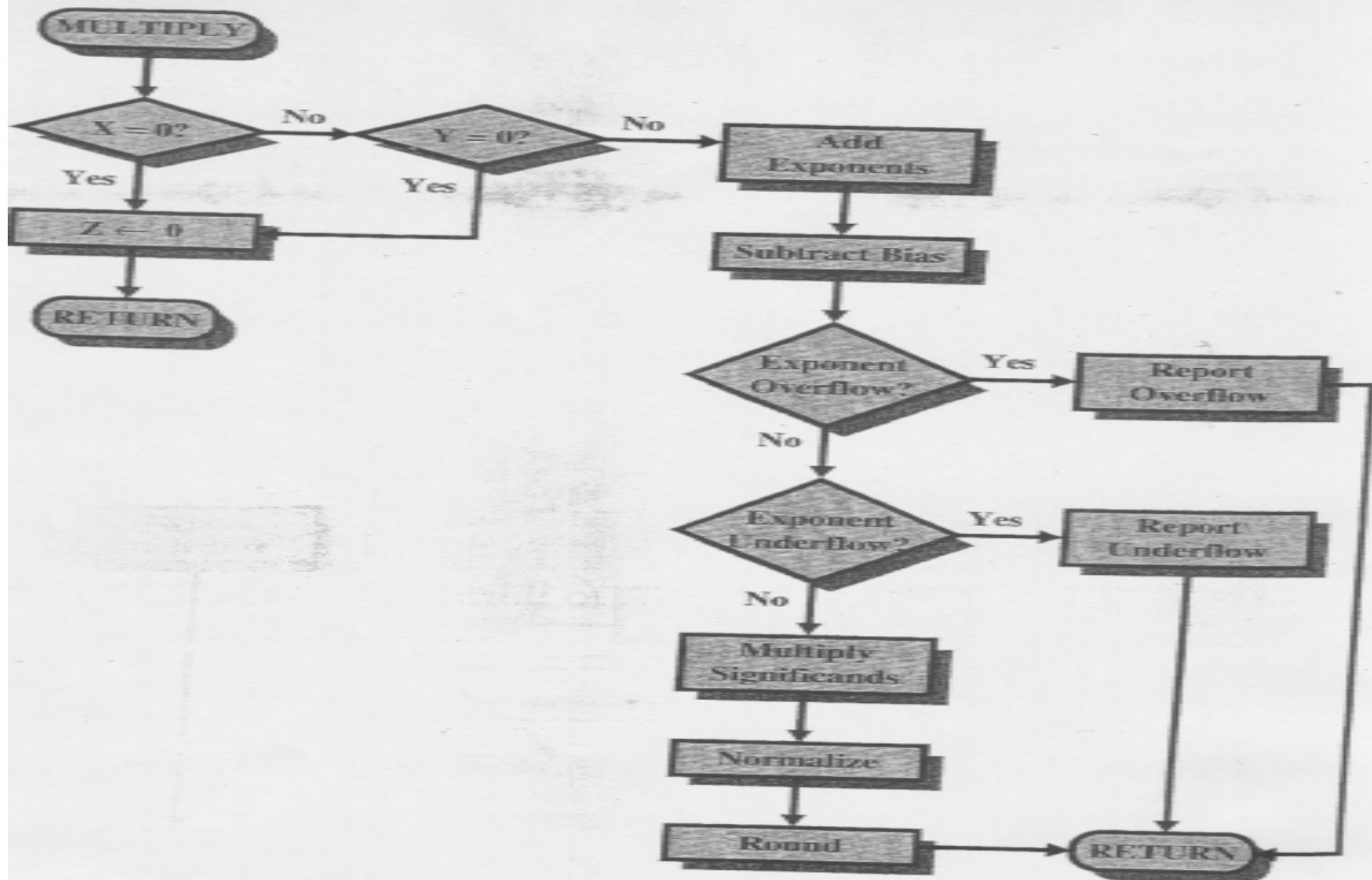


Figure 8.23 Floating-Point Multiplication ( $Z \leftarrow X \times Y$ ).