Chapter 16

Appendix: Number Theory
Tools for Floating-Point Arithmetic

16.1 Continued Fractions

Continued fractions make it possible to build very good (indeed, the best possible, in a sense that will be made explicit by Theorems 49 and 50) rational approximations to real numbers. As such, they naturally appear in many problems of number theory, discrete mathematics, and computer science. Since floating-point numbers are rational approximations to real numbers, it is not surprising that continued fractions play a role in some areas of floating-point arithmetic.

Excellent surveys can be found in [166, 384, 331, 218]. Here, we will just present some general definitions, as well as the few results that are needed in this book, especially in Chapters 5 and 11.

Let \( \alpha \) be a real number. From \( \alpha \), consider the two sequences \((a_i)\) and \((r_i)\) defined by

\[
\begin{align*}
    r_0 &= \alpha, \\
    a_i &= \lfloor r_i \rfloor, \\
    r_{i+1} &= \frac{1}{r_i - a_i},
\end{align*}
\]

(16.1)

where \( \lfloor . \rfloor \) is the usual floor function. Notice that the \( a_i \)'s are integers and that the \( r_i \)'s are real numbers.

If \( \alpha \) is an irrational number, then these sequences are defined for any
$i \geq 0$ (i.e., $r_i$ is never equal to $a_i$), and the rational number

$$\frac{P_i}{Q_i} = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots + \frac{1}{a_i}}}}$$

is called the $i$-th convergent of $\alpha$. The $a_i$'s constitute the continued fraction expansion of $\alpha$. We write

$$\alpha = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}$$

or, to save space,

$$\alpha = [a_0; a_1, a_2, a_3, a_4, \ldots].$$

If $\alpha$ is rational, then these sequences terminate for some $k$, and $P_k/Q_k = \alpha$ exactly. The $P_i$'s and the $Q_i$'s can be deduced from the $a_i$'s using the following recurrences:

$$\begin{align*}
P_0 &= a_0, \\
P_1 &= a_1a_0 + 1, \\
Q_0 &= 1, \\
Q_1 &= a_1, \\
P_k &= P_{k-1}a_k + P_{k-2} \text{ for } k \geq 2, \\
Q_k &= Q_{k-1}a_k + Q_{k-2} \text{ for } k \geq 2.
\end{align*}$$

Note that these recurrences give irreducible fractions $P_i/Q_i$: the values $P_i$ and $Q_i$ that are deduced from them satisfy $\gcd(P_i, Q_i) = 1$.

The major interest in the continued fractions lies in the fact that $P_i/Q_i$ is the best rational approximation to $\alpha$ among all rational numbers of denominator less than or equal to $Q_i$. More precisely, we have the following two results [166].

**Theorem 48.** ([166, p.151]) Let $(P_j/Q_j)_{j \geq 0}$ be the convergents of $\alpha$. If a rational number $P/Q$ is a better approximation to $\alpha$ than $P_k/Q_k$ (namely, if $|P/Q - \alpha| < |P_k/Q_k - \alpha|$), then $Q > Q_k$.

**Theorem 49.** ([166, p.151]) Let $(P_j/Q_j)_{j \geq 0}$ be the convergents of $\alpha$. If $Q_{k+1}$ exists, then for any $(P, Q) \in \mathbb{Z} \times \mathbb{N}^*$, with $Q < Q_{k+1}$, we have

$$|P - \alpha Q| \geq |P_k - \alpha Q_k|.$$ 

If $Q_{k+1}$ does not exist (which implies that $\alpha$ is rational), then the previous inequality holds for any $(P, Q) \in \mathbb{Z} \times \mathbb{N}^*$. 