Smooth Integers

Integers with all their prime factors small are called *smooth* integers. If all prime factors of an integer are $\leq$ some limit z, this integer is called *z-smooth*. The occurrence of smooth integers is of utmost importance for several modern factorization methods, such as the quadratic sieve and the number field sieve. The optimization strategy for these methods demands some estimate of the number $\psi(x, z)$ of z-smooth integers up to x, since the running time of an implementation of the method depends on the supply of smooth integers.

How many z-smooth integers below x are there? We are looking for integers which are products of primes p_i, all less than z, with $\prod p_i \leq x$, i.e., with $\sum \ln p_i \leq \ln x$. Since $\sum_{p \leq z} \ln p \approx z$, see (2.6A), and the number of primes $\leq z$, $\pi(z) \approx z/\ln z$, the average value $\overline{\ln p}$ of $\ln p$ for $p \leq z$ is

$$\overline{\ln p} \approx \frac{\ln z}{z} \sum_{p \leq z} \ln p \approx \ln z.$$

Thus we expect the average number of terms in $\sum \ln p_i \leq \ln x$ to be about $u = \ln x / \ln z$. Thus a z-smooth number $\leq x$ contains about u prime factors, and these can be picked out from $\pi(z)$ primes in about

$$\binom{\pi(z)}{u} \approx \frac{(\pi(z))^u}{u!}$$

different ways, since $\pi(z)$ is large. Using the value of u above and Stirling's formula for $\ln u!$, we find

$$\ln \psi(x, z) \approx u \ln \pi(z) - u \ln u + \text{ smaller terms} \approx \ln x - u \ln u. \tag{5.28}$$

If this is rewritten as $\psi \approx x u^{-u}$, we see that (5.28) implies that the *proportion* as z-smooth numbers below x is u^{-u}, with $u = \ln x / \ln z$. This approximation is often sufficiently accurate to estimate the running times of computer programs, which depend on the supply of smooth integers.—Because of the crude estimates made, this formula is, however, not quite as accurate as it could be. A more careful analysis [8] shows that, for $z < \sqrt[3]{x}$,

$$\ln \psi(x, z) \geq \ln x - u \left(\ln u + \left(1 + \frac{1}{\ln u}\right)(\ln \ln u - 1) + C \left(\frac{\ln \ln u}{\ln u} \right)^2 \right), \tag{5.29}$$

for some absolute constant C. Our crude estimate coincides with the two leading terms of this result.—Because of the inequality sign we have a supply of smooth numbers, which is at least this large.—For future reference, we shall give an important particular case, namely

$$z = e^{a\sqrt{\ln x \ln \ln x}}, \quad \text{leading to} \quad \psi(x, z) = x z^{-1/(2a^2)+o(1)}. \tag{5.30}$$

If we introduce the function

$$L(x, u, v) = e^{v \ln^u x (\ln \ln x)^{1-u}} = x^{v(\ln \ln x / \ln x)^{1-u}}, \tag{5.31}$$

then z in (5.30) is $= L(x, \frac{1}{2}, a)$, and (5.30) can be transformed to

$$\frac{xz}{\psi(x, z)} \to L(x, \frac{1}{2}, a + \frac{1}{2a}), \quad \text{as} \quad x \to \infty, \tag{5.32}$$

a result, which is quite frequently used in running time analyses of those factorization and prime proving methods, which depend on the use of smooth numbers.—See e.g. [8′].

Searching for Factors of Certain Forms

It is sometimes known that the factors of a composite number are of a certain mathematical form. The number may, e.g. have the form $N = 6a^2 - b^2$ with $(a, b) = 1$, and hence we know that 6 is a quadratic residue of N, implying that all prime factors of N (if N can be factorized) take either of the forms $p = 24k \pm 1$ or $p = 24k \pm 5$. (See Table 22.) Several of the methods for factor search can quite easily be modified to take advantage of such a situation. In the method of trial division for instance, described on p. 142, it is quite simple to generate trial divisors of a given linear form and no others. Just set the initial values `p1:=-5; p2:=1; d1:=14;` and `d2:=2;` and in the division loop write

```
d1:=24-d1; d2:=24-d2; p1:=p1+d1; p2:=p2+d2;
```

subsequently checking whether N is divisible, first by `p1` and then by `p2`.

Exercise 5.3. An alternative to the above construction is to write *four* division statements in the program loop. Try this (compare with the computer program `TrialDivision` on p. 143).

Later we shall briefly hint at how several of the more important factorization methods can be modified to use shortcuts when the factor has the form $p = 2kn+1$ and where n is a given integer, and $k = 1, 2, \ldots$ To indicate the importance of this case, we give

Legendre's Theorem for the Factors of $N = a^n \pm b^n$

It has long been known that, under certain conditions, the prime factors of $N = a^p \pm b^p$ (p being an odd prime), all have the form $2kp + 1$. This result has been generalized by Legendre, who proved

Theorem 5.7. Legendre's Theorem. All prime factors p of the number $N = a^n \pm b^n$, with $\text{GCD}(a, b) = 1$, have the form $p = kn + 1$, apart from those which