N	c_N/N^2
10	0.63
100	0.6087
1000	0.608383
10000	0.60794971
100000	0.6079301507

TABLE 3.1: The probabilities that two random positive integers below N are coprime.

We conclude this section with the following question: what is the probability that two random integers are coprime? More precisely, when N gets large and $c_N = \#\{1 \le x,y \le N : \gcd(x,y) = 1\}$, we are interested in the numerical value of c_N/N^2. Table 3.1 gives c_N/N^2 for some values of N; it seems to approach a limit which is a little larger than $3/5$. In fact, the value is

$$\frac{c_N}{N^2} \in \frac{6}{\pi^2} + O\Big(\frac{\log N}{N}\Big) \approx 0.6079271016 + O\Big(\frac{\log N}{N}\Big).$$

Interestingly, a similar approximation holds for the probability that a random integer is **squarefree**, so that it has no square divisor p^2:

$$\frac{\#\{1 \le x \le N : x \text{ is squarefree}\}}{N} \in \frac{6}{\pi^2} + O\Big(\frac{1}{\sqrt{N}}\Big).$$

Exercises 4.18 and 14.32 answer the corresponding questions for polynomials over a finite field.

In Figure 3.2, we see a two-dimensional coordinate system where the point $(x,y) \in \mathbb{N}^2$ for $x,y \le 200$ is colored white if $\gcd(x,y) = 1$ and gray otherwise. The intensity of a pixel is proportional to the number of prime factors in the gcd. The probability that two random integers below 200 are coprime is precisely the percentage of the area of the 200×200 pixels that is colored white. Thus about $3/5$ of all pixels are white, and about $2/5$ are gray.

If you hold the page horizontally in front of your eyes, you can see (almost) white horizontal and vertical lines corresponding to prime values of x and y, and dark lines through the origin corresponding to lines $ax = by$ with small integers a,b, the most clearly visible being the line $x = y$.

3.4. (Non-)Uniqueness of the gcd

The nonuniqueness of the gcd is a harmless nuisance from a mathematical point of view. But in software, we have to implement a *function* gcd with a unique output. In this section, we discuss one way of achieving this.

Since $\mathbb{Q}$ is a field, every nonzero rational number is a unit in $\mathbb{Q}$, and so $ua \sim a$ in $R = \mathbb{Q}[x]$ for all nonzero $u \in \mathbb{Q}$ and all $a \in R$. If we want to define a single

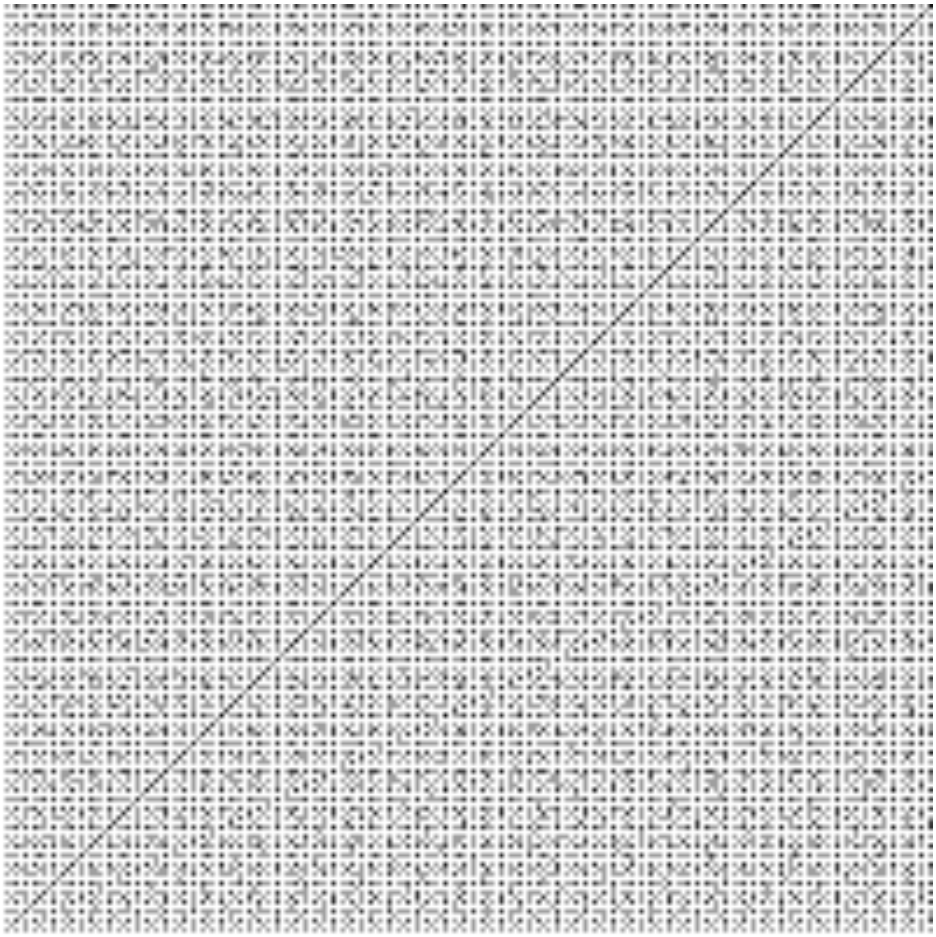

FIGURE 3.2: The greatest common divisors of x and y for $1 \leq x, y \leq 200$.

element $\gcd(f,g) \in \mathbb{Q}[x]$, which one should we choose? In other words, how do we choose *one* representative from among all the multiples of a? A reasonable choice is the **monic** polynomial, that is, the one with leading coefficient 1. Thus if $\mathrm{lc}(a) \in \mathbb{Q} \setminus \{0\}$ is the leading coefficient of $a \in \mathbb{Q}[x]$, then we take $\mathrm{normal}(a) = a/\mathrm{lc}(a)$ as the **normal form** of a. (This has nothing to do with the "normal EEA" on page 51.)

To make this work in an arbitrary Euclidean domain R, we assume that we have selected some normal form $\mathrm{normal}(a) \in R$ for every $a \in R$ so that $a \sim \mathrm{normal}(a)$. We call the unit $u \in R$ with $a = u \cdot \mathrm{normal}(a)$ the **leading unit** $\mathrm{lu}(a)$ of a. Moreover, we set $\mathrm{lu}(0) = 1$ and $\mathrm{normal}(0) = 0$. The following two properties are required:

- two elements of R have the same normal form if and only if they are associate,
- the normal form of a product is equal to the product of the normal forms.

finite fields, Ma & von zur Gathen (1990)) give worst case and average case analyses of several variants of the Euclidean Algorithm.

The fact that two random integers are coprime with probability $6/\pi^2$ is a theorem of Dirichlet (1849). Dirichlet also proves the fact, surprising at first sight, that for fixed a in a division the remainder $r = a \text{ rem } b$, with $0 \le r < b$, is more likely to be smaller than $b/2$ than larger: If p_a denotes the probability for the former, where $1 \le b \le a$ is chosen uniformly at random, then p_a is asymptotically $2 - \ln 4 \approx 61.37\%$. For Dirichlet's theorem, and also the corresponding statement about the probability of being squarefree (due to Gegenbauer 1884), see Hardy & Wright (1985), §§18.5 and 18.6. A heuristic argument goes as follows. A prime p divides a random integer x with probability $1/p$, and neither x nor y with probability $1 - 1/p^2$. Hence $\gcd(x,y) = 1$ happens with probability $\zeta(2)^{-1} = \prod_{p \text{ prime}}(1 - 1/p^2) = 6/\pi^2$; see Notes 18.4 for a discussion of Riemann's zeta function. The value of $\zeta(2)$ was determined by Euler (1734/35b, 1743); see Apostol (1983) for a simple way of calculating this quantity.

3.4. The Euclidean Algorithm 3.14 with monic remainders (for univariate polynomials) appears in the 1969 edition of Knuth (1998), and in Brown (1971).

The calculation of the Bézout coefficients via the EEA in general is in Euler (1748a), §70. See also Notes 6.3. Gauß (1863b), articles 334 and 335, does this for polynomials in $\mathbb{F}_p[x]$, where p is prime.

Exercises.

3.1 Prove that two odd integers whose difference is 32 are coprime.

3.2 Let R be an integral domain. Show that

$$a \sim b \iff (a \mid b \text{ and } b \mid a) \iff \langle a \rangle = \langle b \rangle,$$

where $\langle a \rangle = Ra = \{ra : r \in R\}$ is the ideal generated by a.

3.3 Prove Lemma 3.4. Hint: For (v) and (vi), show that any divisor of the left hand side also divides the right hand side, and vice versa. What are the corresponding statements for the lcm? Are they also true?

3.4* Show that $\gcd(a,b) = 1$ and $\gcd(a,c) = 1$ imply that $\gcd(a,bc) = 1$.

3.5** We consider the following property of a Euclidean function on an integral domain R:

$$d(ab) \ge d(b) \text{ for all } a,b \in R \setminus \{0\}. \tag{9}$$

Our two familiar examples, the degree on $F[x]$ for a field F and the absolute value on $\mathbb{Z}$, both fulfill this property. This exercise shows that every Euclidean domain has such a Euclidean function.

(i) Show that $\delta : \mathbb{Z} \longrightarrow \mathbb{N}$ with $\delta(3) = 2$ and $\delta(a) = |a|$ if $a \ne 3$ is a Euclidean function on $\mathbb{Z}$ violating (9).

(ii) Suppose that R is a Euclidean domain and $D = \{\delta : \delta \text{ is a Euclidean function on } R\}$. Then D is nonempty, and we may define a function $d : R \longrightarrow \mathbb{N} \cup \{-\infty\}$ by $d(a) = \min\{\delta(a) : \delta \in D\}$. Show that d is a Euclidean function on R (called the **minimal Euclidean function**).

(iii) Let δ be a Euclidean function on R such that $\delta(ab) < \delta(b)$ for some $a,b \in R \setminus \{0\}$. Find another Euclidean function δ^* that is smaller than δ. Conclude that the minimal Euclidean function d satisfies (9).

(iv) Show that for all $a,b \in R \setminus \{0\}$ and a Euclidean function d satisfying (9), we have $d(0) < d(a)$, and $d(ab) = d(b)$ if and only if a is a unit.