Chapter 12

Big "O" notation

12.1 Introduction

When analyzing algorithms it is often sufficient to get a reasonable bound on their performance. For example, when analyzing the bubble sort, one could say "it takes roughly n^2 steps for an input of length n"; as we've seen, the true value is

$$\frac{1}{2}n(n-1) \;=\; \frac{1}{2}n^2 - \frac{1}{2}n,$$

and so we would be ignoring the factor of $\frac{1}{2}$ in $\frac{1}{2}n^2$ and the term $\frac{1}{2}n$. We think of the function $\frac{1}{2}n(n-1)$ as growing "about as fast as" the function n^2. We want to make this idea more precise.

In general, let $S = \{x \in \mathbb{R} : x \geqslant 0\}$ and then let f and g be functions from S to S. We say that $f(x)$ is $O(g(x))$, or write $f(x) \in O(g(x))$, if there are constants $C > 0$ and $k \geqslant 0$ such that $f(x) \leqslant Cg(x)$ whenever $x \geqslant k$; see Figure 12.1. We think of $O(g(x))$ as being the set of functions that grow "no faster" than $g(x)$.

12.2 General properties

In this section we establish some general results about how functions beeave with respect to this notation.

First note that, if $f(x) \in O(g(x))$ and $g(x) \in O(h(x))$, then we have constants C_1, k_1, C_2 and k_2 such that $f(x) \leqslant C_1g(x)$ for $x \geqslant k_1$ and $g(x) \leqslant C_2h(x)$ for $x \geqslant k_2$. So, if we let k be the maximum of k_1 and k_2, then we have

$$f(x) \leqslant C_1g(x) \leqslant C_1C_2h(x) \;\text{ for } x \geqslant k,$$

and so $f(x) \in O(h(x))$. We see that, if $g(x) \in O(h(x))$, then

$$f(x) \in O(g(x)) \Rightarrow f(x) \in O(h(x)),$$

and so $O(g(x)) \subseteq O(h(x))$. In particular, if $g(x) \in O(h(x))$ and $h(x) \in O(g(x))$ then $O(g(x)) = O(h(x))$.

If functions $g(x)$ and $h(x)$ "differ by a constant factor" (i.e. if there exists a constant $D > 0$ such that $g(x) = Dh(x)$ for all x), then we have:

- $g(x) \in O(h(x))$, as $g(x) \leqslant Dh(x)$ for all x, and
- $h(x) \in O(g(x))$, as $h(x) \leqslant \frac{1}{D}g(x)$ for all x.

So $O(g(x)) = O(h(x))$ in this case.

Note that, if $f_1(x) \in O(g_1(x))$ and $f_2(x) \in O(g_2(x))$, then there are constants C_1, k_1, C_2 and k_2 such that $f_1(x) \leqslant C_1g_1(x)$ whenever $x \geqslant k_1$ and $f_2(x) \leqslant C_2g_2(x)$ whenever $x \geqslant k_2$. If C is the maximum of C_1 and C_2 and k is the maximum of k_1 and k_2, then

$$f_1(x) \leqslant Cg_1(x) \text{ and } f_2(x) \leqslant Cg_2(x) \text{ whenever } x \geqslant k.$$

We see that $f_1(x) + f_2(x) \leqslant C(g_1(x) + g_2(x))$ whenever $x \geqslant k$, and so $f_1(x) + f_2(x) \in O(g_1(x) + g_2(x))$.

In a similar vein, we have that

$$f_1(x) \times f_2(x) \leqslant C_1C_2(g_1(x) \times g_2(x))$$

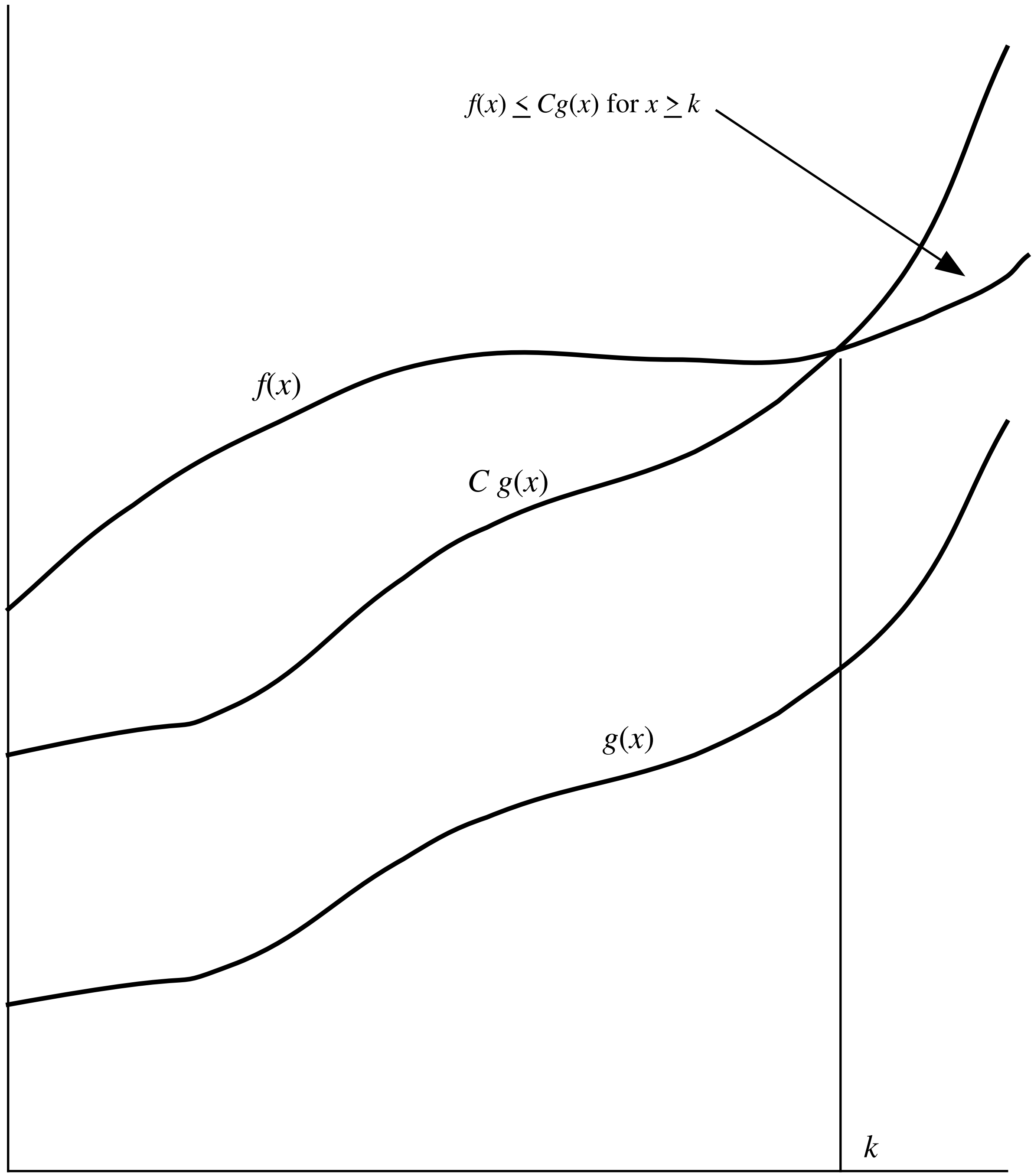

Figure 12.1: $f(x) \in O(g(x))$

if $x \geqslant k$ and so $f_1(x) \times f_2(x) \in O(g_1(x) \times g_2(x))$.

Let us sum these observations up as follows:

THEOREM. Let $S = \{x \in \mathbb{R} : x \geqslant 0\}$ and suppose that f, g and h are functions from S to S.

1. If $f(x) \in O(g(x))$ and $g(x) \in O(h(x))$ then $f(x) \in O(h(x))$.
2. If $g(x) \in O(h(x))$, then $O(g(x)) \subseteq O(h(x))$.
3. If $g(x) \in O(h(x))$ and $h(x) \in O(g(x))$ then $O(g(x)) = O(h(x))$.
4. If functions $g(x)$ and $h(x)$ differ by a constant factor then $O(g(x)) = O(h(x))$.
5. If $f_1(x) \in O(g_1(x))$ and $f_2(x) \in O(g_2(x))$, then $f_1(x) + f_2(x) \in O(g_1(x) + g_2(x))$.
6. If $f_1(x) \in O(g_1(x))$ and $f_2(x) \in O(g_2(x))$, then $f_1(x) \times f_2(x) \in O(g_1(x) \times g_2(x))$.

12.3 Polynomials

Let us consider polynomials. For example, we have:

$$\text{if } m < n, \text{ then } x^m \in O(x^n) \text{ as } x^m \leqslant x^n \text{ for } x \geqslant 1. \tag{12.1}$$

EXAMPLE. $x \in O(x^3)$ and $x^2 \in O(x^3)$; so $x + x^2 \in O(x^3 + x^3) = O(2x^3) = O(x^3)$. □

In general, if $n_1, n_2, \ldots, n_r \leqslant m$ and $a_i > 0$ for each i with $1 \leqslant i \leqslant n$, then $a_i x^{n_i} \in O(a_i x^m) = O(x^m)$ for each i, and so

$$a_1 x^{n_1} + a_2 x^{n_2} + \ldots + a_r x^{n_r} \in O(x^m + x^m + \ldots + x^m) = O(rx^m) = O(x^m).$$

So we have:

THEOREM. If $p(x)$ is a polynomial of degree at most m, then $p(x) \in O(x^m)$.

For example, we have that $x^2 \in O(x^2)$, $x^2 + x \in O(x^2)$ and $x^2 \in O(x^3)$. Note that $1 = x^0$ so that $O(1) \subset O(x)$.

Note that, if $n > m$, then we do not have $x^n \in O(x^m)$; we cannot find constants C and k such that $x^n \leqslant Cx^m$ for $x \geqslant k$ because we would have to have $x^{n-m} \leqslant C$ for $x \geqslant k$ which is clearly impossible. So we have:

THEOREM. If $n > m$ then $x^m \in O(x^n)$ but $x^n \notin O(x^m)$. In particular, we have:

$$O(1) \subset O(x) \subset O(x^2) \subset O(x^3) \subset \ldots\ldots$$

12.4 Logarithms and exponentials

There is a slight issue when considering logarithms. We have taken our functions as being from S to S where $S = \{x \in \mathbb{R} : x \geqslant 0\}$; however, for any base c, $\log_c(x)$ is negative for $0 < x < 1$. This is not really a problem as we are only interested in the behaviour of our functions from some point k onwards, and we can regard $\log_c$ as being a function from $T = \{x \in \mathbb{R} : x \geqslant 1\}$ to S.

Note that the choice of base c in $\log_c(x)$ is not important when considering $O(\log_c(x))$. We saw in Chapter 8, that, for any two fixed bases c and d, we have that $\log_d(x) = \log_c(x)\log_d(c)$. So $\log_c(x)$ and $\log_d(x)$ differ by a constant, and hence $O(\log_c(x)) = O(\log_d(x))$. We will take our base to be 2 in what follows (i.e. we consider $\log_2(x)$).

One can show that $x < 2^x$ for $x \geqslant 1$ (for natural numbers x we can do this by induction). Taking logarithms of both sides of this inequality gives that $\log_2(x) < x$ for $x \geqslant 1$; we are using the fact that, if $a < b$, then $\log_2(a) < \log_2(b)$. So we have that

$$O(\log_2(x)) \subseteq O(x) \subseteq O(2^x);$$

see Figure 12.2.

In fact we have:

$$O(1) \subset O(\log_2(x)) \subset O(\log_2(x)^2) \subset O(\log_2(x)^3) \subset \ldots\ldots \subset O(x) \subset O(x^2) \subset O(x^3) \subset \ldots\ldots \subset O(2^x), \tag{12.2}$$

and, in particular, that $O(\log_2(x)^d) \subset O(x)$ for any $d \geqslant 1$ and that $O(x^d) \subset O(2^x)$ for any $d \geqslant 1$. We can think of (12.2) as showing the relative growths of the functions listed there.

We can use the facts that, if $f_1(x) \in O(g_1(x))$ and $f_2(x) \in O(g_2(x))$, then $f_1(x) + f_2(x) \in O(g_1(x) + g_2(x))$ and $f_1(x) \times f_2(x) \in O(g_1(x) \times g_2(x))$ to establish some more relationships.

EXAMPLE. As $\log_2(x) \in O(x^2)$ and $x \in O(x^2)$, we have that

$$x + \log_2(x) \in O(x^2 + x^2) = O(2x^2) = O(x^2).$$

We also have that, as $\log_2(x) \in O(x)$ and $x \in O(x)$. then

$$x \times \log_2(x) \in O(x \times x) = O(x^2).$$

Given that $\log_2(x)^2 = \log_2(x) \times \log_2(x) \in O(x)$ by (12.2), we have that

$$\log_2(x) \;=\; \frac{\log_2(x)^2}{\log_2(x)} \;\in\; O\left(\frac{x}{\log_2(x)}\right).$$

It is clearly possible to derive a wide variety of similar facts. □

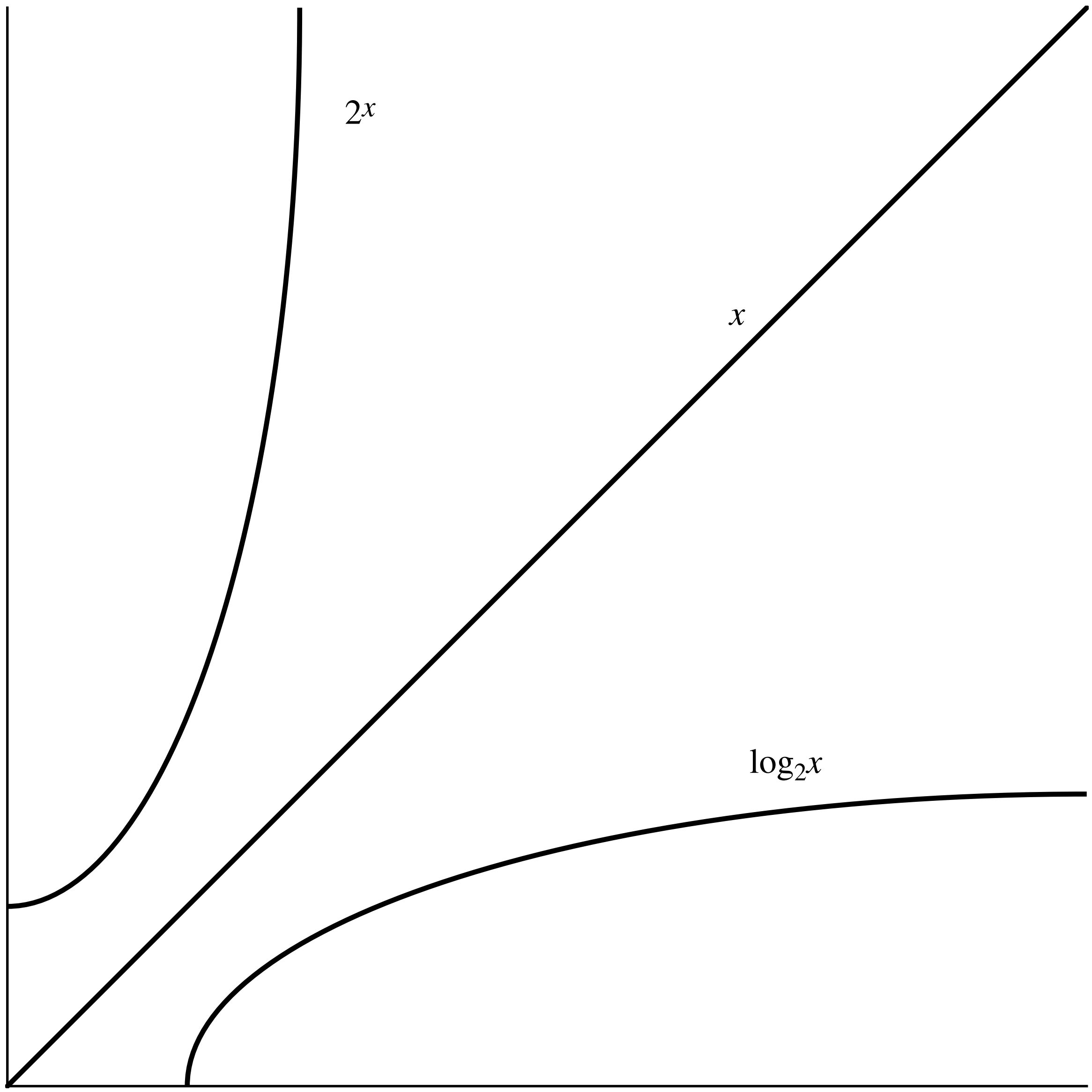

Figure 12.2: $\log_2(x)$, x and 2^x