4. Asymptotic analysis

As mentioned before, it is convenient to find lower and upper bounds for some of the functions which appear in number theory. To this end, it is expedient to introduce some convenient notation.

Definition 4.1. *Let $S \subset \mathbb{R}$ be unbounded from above. If we are given two functions $f\colon S \longrightarrow \mathbb{R}$ and $g\colon S \longrightarrow \mathbb{R}$, where g is always positive, then we write $f(x) = O(g(x))$ if there is a constant M such that*

$$\frac{|f(x)|}{g(x)} < M,$$

for all x sufficiently large.

We write $f(x) = o(g(x))$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0.$$

We write $f(x) \sim g(x)$ if

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1.$$

Example 4.2. *Suppose that we compare* $\sin x$ *and x.*

Both functions are defined for all real numbers. Note that all three of the following statements are correct:

- $\sin x = O(x)$
- $\sin x = o(x)$
- $\sin x = O(1)$.

It is not hard to see that the second statement implies the first and that the third statement implies the second statement. For the third statement note that

$$\frac{|\sin x|}{1} = |\sin x| \leq 1.$$

In fact $f(x) = O(1)$ is equivalent to the statement that $f(x)$ is bounded from above and below for x large enough.

We note the following comparisons

- $\varphi(n) = O(n)$.
- $\sqrt{x} = o(x)$.
- $x^k = o(e^x)$.
- $\log^k(x) = o(x^\alpha)$ for any k and any $\alpha > 0$.
- $\llcorner x \lrcorner \sim x$.

Note that we cannot replace O by o in the first statement, since if $n = p$ is prime, then $\varphi(p) = p - 1$. We list some simple rules for manipulating this notation.

Lemma 4.3.

(1) $O(O(g(x))) = O(g(x))$.
(2) $o(o(g(x))) = o(g(x))$.
(3) $O(g(x)) + O(g(x)) = O(g(x))$.
(4) $o(g(x)) + o(g(x)) = o(g(x))$.
(5) $(O(g(x)))^2 = O(g^2(x)))$
(6) $(o(g(x)))^2 = o(g^2(x)))$

Proof. (1) is shorthand notation for the statement that if

$$f(x) = O(g(x)) \qquad \text{and} \qquad h(x) = O(f(x))$$

then

$$h(x) = O(g(x)).$$

The first statement implies that there are real numbers x_0 and M_0 such that if

$$x > x_0 \qquad \text{then} \qquad \frac{|f(x)|}{g(x)} < M_0.$$

The second statement implies that there are real numbers x_1 and M_1 such that if

$$x > x_1 \qquad \text{then} \qquad \frac{|h(x)|}{f(x)} < M_1.$$

Suppose that $x > \max(x_0, x_1)$. Then

$$\begin{aligned} \frac{|h(x)|}{g(x)} &= \frac{|h(x)|}{f(x)} \frac{f(x)}{g(x)} \\ &\leq \frac{|h(x)|}{f(x)} \frac{|f(x)|}{g(x)} \\ &\leq M_0 M_1. \end{aligned}$$

This gives (1).

(2) is similar to (1).

(3) is shorthand notation for the statement that if

$$f_0(x) = O(g(x)) \qquad \text{and} \qquad f_1(x) = O(g(x))$$

then

$$f_0(x) + f_1(x) = O(g(x)).$$

The first and second statement imply that there are real numbers x_i and M_i, $i = 0$ and 1, such that if

$$x > x_i \qquad \text{then} \qquad \frac{|f_i(x)|}{g(x)} < M_i.$$

Suppose that $x > \max(x_0, x_1)$. Then

$$\begin{aligned}\frac{|f_0(x) + f_1(x)|}{g(x)} &\le \frac{|f_0(x)|}{g(x)} + \frac{|f_1(x)|}{g(x)} \\ &\le M_0 + M_1.\end{aligned}$$

This gives (3).

(4) is similar to (3).

(5) is shorthand notation for the statement that if

$$f(x) = O(g(x))$$

then

$$f^2(x) = O(g^2(x)).$$

The first statement implies that there are real numbers x_0 and M such that if

$$x > x_0 \qquad \text{then} \qquad \frac{|f(x)|}{g(x)} < M.$$

It follows that if $x > x_0$ then

$$\begin{aligned}\frac{|f^2(x)|}{g^2(x)} &= \frac{|f(x)|}{g(x)} \cdot \frac{|f(x)|}{g(x)} \\ &\le M \cdot M \\ &= M^2.\end{aligned}$$

This is (5).

(6) is similar to (5). □

Definition-Theorem 4.4. *There is a constant γ, called* ***Euler's constant****, such that*

$$\sum_{k=1}^{n} \frac{1}{k} = \log n + \gamma + O\left(\frac{1}{n}\right).$$

Proof. Let

$$\alpha_k = \log k - \log(k-1) - \frac{1}{k} \qquad \text{for} \qquad k \ge 2$$

and let

$$\gamma_n = \sum_{k=1}^{n} \frac{1}{k} - \log n \qquad \text{for} \qquad n \ge 1.$$

Note that

$$\begin{aligned}\sum_{k=2}^{n}\alpha_k &= \sum_{k=2}^{n}\log k - \log(k-1) - \frac{1}{k}\\ &= \sum_{k=2}^{n}\log k - \sum_{k=1}^{n-1}\log k - \sum_{k=2}^{n}\frac{1}{k}\\ &= \log n - \log 1 - \sum_{k=2}^{n}\frac{1}{k}\\ &= 1 + \log n - \sum_{k=1}^{n}\frac{1}{k}\\ &= 1 - \gamma_n.\end{aligned}$$

Thus

$$1 - \gamma_n = \sum_{k=2}^{n}\alpha_k.$$

Note that

$$\begin{aligned}\int_{k-1}^{k}\frac{1}{x}\,\mathrm{d}x &= [\log x]_{k-1}^{k}\\ &= \log k - \log(k-1)\end{aligned}$$

is the area under the curve $y = 1/x$ over the interval $k-1 \leq x \leq k$. On the other hand $1/k$ is the area over the interval $k-1 \leq x \leq k$ inside the largest rectangle inscribed between the x-axis and the curve $y = 1/x$.

It follows that α_k is the difference between these two areas, so that α_k is positive. Note that if we drop these areas down to the region between $x = 0$ and $x = 1$ then all of these areas fit into the unit square bounded by $y = 0$ and $y = 1$.

Thus $0 < 1-\gamma_n < 1$ is bounded and monotonic increasing. It follows that $1 - \gamma_n$ tends to a limit. Define γ by the formula:

$$\lim_{n\to\infty}(1-\gamma_n) = 1 - \gamma.$$

As each area is less than half of the natural bounding rectangle, in fact

$$1 - \gamma_n < 1/2.$$

Finally note that the difference

$$\begin{aligned}\gamma_n - \gamma &= (1-\gamma) - (1-\gamma_n) \\ &= \sum_{k=n+1}^{\infty} \alpha_k\end{aligned}$$

is represented by an area which fits inside a box with one side 1 and the other side $1/n$, so that it is less than $1/n$. Thus

$$\gamma_n - \gamma = O\left(\frac{1}{n}\right). \qquad \square$$

Conjecture 4.5. *Euler's constant γ is irrational.*

We can compute as many decimal place of γ as we want; we have

$$\gamma \approx 0.5772156649\ldots.$$