Sums and averages of arithmetic functions
(Ch. 3).

GOAL: given an arithmetic function f, to study the growth of sums like

(a) $\sum_{n \le x} f(n)$, (b) $\frac{1}{x} \sum_{n \le x} f(n)$

as $x \to \infty$.

Notes:
1) Always, (a) denotes the sum over $n \in \mathbb{Z}^+$ with $n \le x$.
2) (b) is a kind of average of the first $[x]$ values of f.

Some terminology:

(1) "sufficiently large" means $\geq a$, for some $a \in \mathbb{R}_+$.

(2) Suppose $f : \mathbb{R} \to \mathbb{C}$ and $g : \mathbb{R} \to \mathbb{R}$, with $g(x) > 0$ for x suff. large. We write

$f(x) = O(g(x))$ ("$f(x)$ is big oh of $g(x)$")

if $f(x)/g(x)$ is bounded for x suff. large (equivalently, $\exists M > 0$ such that $|f(x)| \le M g(x)$ for x suff. large).

(3) For such f, g and $h : \mathbb{R} \to \mathbb{C}$, we write $f(x) = h(x) + O(g(x))$ if $f(x) - h(x) = O(g(x))$.

(4) We write $f(x) \sim g(x)$ ("$f(x)$ is asymptotic to $g(x)$") if

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

Examples

(i) $x^2 = O(x^2)$

(ii) $x^2 + x = O(x^2)$

(iii) $x^2 + x = x^2 + O(x)$

(iv) $x^2 + x + 25\sin(x) = x^2 + x + O(1)$

(v) $x^2 + x \sim x^2$

(vi) $\displaystyle\sum_{p \le x} 1 \sim \frac{x}{\log x}$: this is the Prime Number Theorem.

(we'll prove this later).

(vii) Harder result (we won't prove this):

$$\sum_{p\le x} 1 = \frac{x}{\log x} + O\Big(\overbrace{\frac{x}{(\log x)^2}}^{\text{"error term"}}\Big).$$

A primary tool for deducing "$O(\ )$" and "$\sim$" results for sums like (a) and (b) above is:

Theorem 3.1: Euler summation formula, or ESF.

If $f : [y,x] \to \mathbb{R}$ is continuously differentiable on $[y,x]$, and $0<y<x$, then

$$\sum_{y<n\le x} f(n) = \int_y^x f(t)\,dt + \int_y^x (t-[t])f'(t)\,dt + f(x)([x]-x) + f(y)([y]-y).$$

$<$, not $\le$

Proof.

Denote the sum by $S(f;y,x)$, and note that

$$S(f;y,x)=\sum_{n=[y]+1}^{[x]} f(n).$$

Then

$$\int_y^x f(t)\,dt+\int_y^x (t-[t])f'(t)\,dt$$

$$=\int_y^x f(t)\,dt+\underbrace{\int_y^x tf'(t)\,dt}-\int_y^x [t]f'(t)\,dt$$

integrate by parts with $u=t$, $dv=f'(t)\,dt$

$$=\int_y^x f(t)\,dt+tf(t)\Big|_y^x-\int_y^x f(t)\,dt-\int_y^x [t]f'(t)\,dt$$

$$=xf(x)-yf(y)-\int_y^x [t]f'(t)\,dt$$

break up $[y,x]$ into subintervals

$$=xf(x)-yf(y)-\int_y^{[y]+1}[t]f'(t)\,dt-\int_{[x]}^x [t]f'(t)\,dt-\sum_{n=[y]+1}^{[x]-1}\int_n^{n+1}[t]f'(t)\,dt$$

$[t]$ is constant on each given subinterval

$$=xf(x)-yf(y)-[y]\int_y^{[y]+1}f'(t)\,dt-[x]\int_{[x]}^x f'(t)\,dt-\sum_{n=[y]+1}^{[x]-1}n\int_n^{n+1}f'(t)\,dt$$

$$=xf(x)-yf(y)-[y]\big(f([y]+1)-f(y)\big)-[x]\big(f(x)-f([x])\big)-\sum_{n=[y]+1}^{[x]-1}\big(nf(n+1)-nf(n)\big)$$

make sum go to $[x]$; then compensate

$$=xf(x)-yf(y)-[y]\big(f([y]+1)-f(y)\big)-[x]\big(f(x)-\cancel{f([x])}\big)+[x]f([x]+1)-\cancel{[x]f([x])}\ \ldots$$

$$-\sum_{n=[y]+1}^{[x]} (nf(n+1) - n(f(n))$$

add and subtract $S(f;y,x)$

$$= xf(x) - yf(y) - [y](f([y]+1) - f(y))$$

$$- [x]f(x) + [x]f([x]+1) + S(f;y,x)$$

$$- \sum_{n=[y]+1}^{[x]} (nf(n+1) - (n-1)f(n))$$

the sum telescopes

$$= xf(x) - yf(y) - [y](\cancel{f([y]+1)} - f(y))$$

$$- [x]f(x) + \cancel{[x]f([x]+1)} + S(f;y,x)$$

$$- \cancel{[x]f([x]+1)} + \cancel{[y]f([y]+1)}$$

$$= (x-[x])f(x) - (y-[y])f(y) + S(f;y,x),$$

as required. $\square$

Note: sometimes, for integral y (e.g. $y=1$), we want to sum over $[y,x]$ instead of $(y,x]$. This adds one more term to our sum; moreover, for such y, $[y]-y=0$. So ESF yields:

Corollary: ESF$'$. For $y \in \mathbb{Z}_+$ and everything else as above,

$$\sum_{y \le n \le x} f(n) = \int_y^x f(t)\,dt + \int_y^x (t-[t])f'(t)\,dt$$

$$+ f(x)([x]-x) + f(y).$$